

## ウェイソンの4枚カード課題に関する研究のレビュー

その1：1966年～1979年まで

柴 田 淑 枝<sup>1)</sup>

### はじめに

#### 4枚カード課題が登場してきた時代背景

「Wasonの4枚カード課題 (Wason's four-cards task : 以下4枚カード課題とする)」<sup>2)</sup>という名称で知られている課題が初めて紹介されたのは、 Wason (1966) である<sup>3)</sup>。Wason (1966) は、人間の推論の特徴を示すための実験的研究をいくつか紹介している(例: Brunerの概念学習・Wasonの2-4-6課題・三段論法など)。この中の1つとして4枚カード課題が紹介されており、それはWasonの行なった予備実験の結果を用いて説明されている。この4枚カード課題による研究は、演繹的推論の研究であり、主に「If ~ then ...」

という条件的推論に焦点をしぼっている。そのため人間の推論の特徴をより深く追究できるといえる。

さて、当時の研究で人間の推論能力について述べたものといえば、Piagetによる発達段階説を思い浮かべることができる。しかし、人間の発達を普遍的なものとして捉える方法は、今日まで多くの研究によって批判されてきている。Piagetも「平行的デカラージュ (decalage)」という概念を用いて、発達の非普遍的側面を説明しているのだが、Piaget自身の理論の枠組みの中で、それがうまく説明されているとは言い難い(川村, 1991)。4枚カード課題を用いた演繹的推論の研究も、このPiagetの発達段階説への批判に基づいている<sup>4)</sup>。

さらに、同じころHenleが、形式論理と人間の推論の関係についてレビューを行っている(Henle, 1962)。それによると、かつて論理は人間の推論の「お手本」のように考えられていたが、その後、人間の推論は実用的かつ具体的であり、形式論理とまったく異なるものとして捉えられるようになってきた。しかし、形式論理と人間の推論の間の「ズレ」は、人間の推論の「誤り」を示すものとは一概に言えない。状況に適応するために論理的な厳密性を切り捨てたり、あるいは、状況に含まれる暗黙の前提を取り入れた、いわば「良い誤り」と呼べるものもあるからである。つまり、推論のプロセス自体は人間の推論と形式論理の間にそれほど違いはない、ただ、

- 
- 1) 名古屋大学大学院教育学研究科博士課程(後期課程)
  - 2) この課題は、「4枚カード課題」以外にも、「Wason課題」と呼ばれることがある。しかし、広義には「Wason課題」は2種類あるので、「4枚カード課題」という名称に統一することとする。ちなみに、1つはこの「4枚カード課題」であり、もう1つは「2-4-6課題」と呼ばれている帰納的推論問題である。なお、「4枚カード課題」にもさらに2種類ある。(a)4枚カード選択課題と(b)4枚カード評価課題である。(a)は、本文中にFigure 1を用いて説明されているものである。これに対して(b)は、4枚のカードの中からあらかじめ2枚のカードが選択されており、その選択パターンがルールを真とするパターンであるか偽とするパターンであるかを同定する問題である。
  - 3) ちなみに、4枚カード課題に最初に名称がつけられたのは、筆者の調べたところでは Evans & Lynch (1973) であり、「Wasonの選択課題 (Wason's selection task)」と呼ばれた。しかし、2)にあるように、この課題は「選択課題」のほかに「評価課題」のバリエーションも持つ。「4枚カード課題」という名称が最初に用いられたのは、「4枚カード『問題』(four-cards problem)」という名称で Wason & Evans (1975) が最初に用いている。

4) 発達心理学の分野においても、ピアジェの発達段階と条件的推論を絡めた研究がいくつかある(Bucci, 1978; Roberge & Paulus, 1971; Shapiro & O'Brien, 1970; Staudenmayer & Bourne, 1977; Taplin, Staudenmayer, & Taddonio, 1974など)。問題はすべて三段論法の形式で提示されている。つまり、「カードをめくって確かめてみる」という設定は一切なされていない。

この理由については、4枚のカードという設定が問題構造を難しくしており、子ども向けの課題としては無理があるのではないかということを考えられる。

最初の前提条件として何を取り入れるのかが異なるというのである。このような人間の推論と形式論理との関係についての活発な議論を背景にして、4枚カード課題の研究が始まったとも言える。

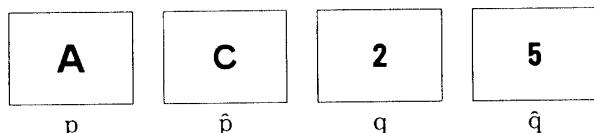
#### 4枚カード課題とは

4枚カード課題を簡単に説明すると、以下のようになる。

まず、4枚のカードを被験者に提示する。カードには、例えば、A・C・2・5という英数字がそれぞれ示されている (Figure 1)。被験者には、これらのカードの一方の面には文字が、もう一方の面には数字が書かれていると伝えられる。つまり、この両面に印刷されたカードのどちらか一方の面を表に向けて被験者に提示されていくわけである。そして、「もし、一方の面が母音ならば、もう一方の面は偶数である（一般的な表現として If p then q）」というルールがあるとしたときに、このルールが正しいかどうかを判断するためには、どのカードを裏返して確認しなければならないかを問うのである。ちなみに、正答は、「A(p)」と「5(̄q [=not q])」のカードである。上述のルールの場合、まず、母音が書かれている「A(p)」のカードについて、この裏が偶数になっているかどうかを確認しなければいけない。さらに、ルールは「もし、一方の面に偶数が書かれていなければ、もう一方の面には母音が書かれてい（If ̄q then ̄p）」と対偶の関係にあり、これらは互いに論理的に同値であるから、奇数が書かれている「5(̄q)」のカードを選択し、その裏が子音になっているかどうかを確認しなければいけない。

上に挙げた例は、基本的な問題であり、「抽象課題」と呼ばれるものである。研究が盛んに行なわれるようになるにつれて、それぞれの目的のためにさまざまのバリエーション（ルールの具体性を高める具体課題やルールに否定辞を加えたりしたものなど）が用いられるようになった。

なお、4枚カード課題については、1966年から現在まで約70人の研究者が約100本の論文を発表しているが、本論文では、「レビューその1」として、1966年から1970年代後半までにわたる約15年間の研究の流れを紹



もし一方の面が母音ならば、もう一方の面は偶数である

Figure 1 : Examples of four cards and the rule

介・検討することとする。

#### 4枚カード課題研究の黎明期： (1966年～1970年代初め)

初期の研究では Piaget の発達段階説に触れ、演繹的推論と最も関係の深い「形式的操作段階」について議論している。この段階には、通常11～12歳くらいに到達すると Piaget は主張した。しかし4枚カード課題の研究者は、成人に達した者でも形式的操作による思考をすることは非常に難しいという見解を示した (Johnson-Laird & Tagart, 1969; Wason, 1966; 1969)。彼らは、主に成人の被験者によるデータを分析し、正答率が非常に低いことを明らかにしている。もし Piaget の言うように成人に形式的操作が身についているとすれば、カードの裏側についての推論、つまり、対偶を考慮した推論も可能であり、容易に正答できるはずである。しかし、例えば Wason (1969) の実験では、ほとんどの被験者が推論に失敗した。「もし、一方の面に母音が書かれているならば、もう一方の面には偶数が書かれている（If p then q）」というルールを用いた結果、90.6%の被験者が、母音のカードのみ (p)，もしくは母音のカードと偶数のカード (p • q) を選択したのである。

このような失敗に関して、初期の研究では2つの説明が示された：

(1) 形式論理における真理値表 (Table 1) と被験者の考えるそれ (Table 2) との違いによるというもの。形式論理では、「If p then q」というルールの場合、カードの両面が「p • ̄q」という組み合わせのみ偽であり、それ以外はすべて真とみなす。ところが、実験結果から、被験者はルールの真値について「真」「偽」の他に「関係なし」を含めた3区分法をとっていると考えられる (Johnson-Laird & Tagart, 1969; Wason, 1966;

Table 1 Truth table on logic

	p	̄p
q	○	○
̄q	×	○

○=true ×=false

Table 2 Truth table of subjects  
(defective truth table)

	p	̄p
q	○	?
̄q	×	?

○=true ×=false ?=irrelevance

Wason & Shapiro, 1971)。つまり人間は、 $\hat{p}$ 、つまり前提に含まれない事象については、結果がどうであろうとルールとは「関係ない」とみなしているのである。Wason & Shapiro (1971) はこれを「不完全な真理値表 (defective truth table)」と呼んだ。

(2) Shapiro & O'Brien (1970) が行なった、子どもを被験者とした実験からヒントが得られる。被験児の多くが「If  $p$  then  $q$ 」を「If and only if  $p$  then  $q$ 」として、つまり、双条件的（逆も真なり）に解釈していた（Table 3 参照）。よって、成人も同じような解釈を行なっていると推察できる。

この 2 つのうち、どちらが正しい説明なのかを検討した研究はない。それは、被験者の反応が、どちらの説明にも当てはまるからである。被験者がルールを確証する

Table 3 Truth table of subject (biconditional)

	$p$	$\hat{p}$
$q$	○	×
$\hat{q}$	×	○

○=true ×=false

事例のみに焦点を当てている場合、(1)に従えば  $p$  カードのみを選択し、(2)に従えば  $p$  と  $q$  のカードを選択することになるはずである。実際に、被験者の最多反応は  $p$  もしくは  $p$  と  $q$  であり、どちらがより多いという結果は出されていない。ちなみに、もし被験者が反証事例にも焦点を当てているとすれば、(1)では  $p$  と  $\hat{q}$  のカードを選択し、(2)ではすべてのカードを選択することになる。しかし、上述の実験結果から、被験者は確証事例にのみ焦点を当てていることがわかる。Wason (1969) はこの様な推論の特徴を「確証バイアス」と呼んだ。

さて、以上のように成人における論理的推論問題の正答率が低いという結果が示されたわけだが、しかし、どのような問題でも正答率が低かったわけではない。当時の研究では、正答率を上昇させる要因として、問題に含まれる論理に気付かせることが最も重要であると考えられていた。そのために、2つの方法が考えられた。以下に詳しく述べていく。

### 論理に関するヒントの効果

Wason (1968) の実験 1 では、被験者に 4 枚カード課題を与えて選択させた。次に「カードの裏側に何が書いてあればルールを偽とするのか」を答えさせた上で、再度同じ 4 枚カード課題を与えた。その結果、1 度目の課題では正解者はいなかったが、2 度目の課題では、22.2% の被験者が正答できた。しかし、これは劇的な上昇

とは言えない。

そこで、Wason (1969) では、被験者にヒントを与えては問題を遂行させるということを繰り返して、誤答がどのように正答へと変化していくのかを観察した。被験者に与えたヒントは、反証することに気付かせるものであった。その結果、被験者の最多反応パターンは「 $p$ 」もしくは「 $p \cdot q$ 」（こういう反応パターンを考え出す推論を no insight と呼ぶ）→「 $p \cdot q \cdot \hat{q}$ 」（partial insight）→「 $p \cdot \hat{q}$ 」（complete insight）へと変化していく。最初は正解者がいなかったが、最終解答段階では約 46.9% の被験者が正答できた。なお、最初の反応パターンが「 $p \cdot q$ 」の被験者よりも「 $p$ 」の被験者の方が、最終的に正答できる割合が高かった。そのため、 $\hat{q}$  を選択しなければいけないことに気付くよりも  $q$  を選択しなくてもよいことに気付くほうが難しいことが示された。真理値表で言えば、真・偽・関係なしの 3 区分法をとる被験者の方が双条件の被験者よりもその後正答しやすいということになる。

Wason & Shapiro (1971) の実験 1 では、(a)カードの一面だけを被験者に提示し、その裏側に何が書いてあればルールを確証／反証するのか、を繰り返し被験者に考えさせる群（構成群）と、(b)カードの両面を一度に提示して（これを両面提示と呼ぶ）、それはルールを確証するカードなのか、反証するカードなのかを繰り返し被験者に考えさせる群（評価群）を用いた。そして、両群とも最後に抽象選択課題が与えられた。その結果、 $q$  の選択比率は構成群のほうが有意に低かった（フィッシャーイェーツ直接法検定 [片側] で  $p < .05$ ）。その結果は、実際には有意水準を満たさなかったが、正答率にも反映されているように見える（構成群 41.7%；評価群 16.7%）。<sup>5)</sup>

Wason & Johnson-Laird (1970) では、まず 4 枚カード課題を行わせ、次に、被験者が問題に含まれる論理に気がつくように「評価過程 (evaluation process)」と呼ばれる手順を踏ませた。評価過程とは「一方の面が  $p$  でもう一方の面が  $q$  という組み合わせのカードだったら、ルールを真とするのか偽とするのか」という評価をさせる手順をさす。そして、再度 4 枚カード課題を行わ

5) Wason & Shapiro (1971) によれば、 $q$  および  $\hat{q}$  カードの選択比率に関するフィッシャーイェーツの直接法検定のみ行なわれており、正答率に関する検定は行なわれていなかった（ $q$  カード： $p < .05$ ； $\hat{q}$  カード： $p > .05$ ）。正答率について著者が行なった検定では、 $p = 0.15$  であり、有意差はないとみなすことができる。

すると、有意に正答率が上昇した（1回目と2回目の選択課題の間の正答比率の差をマクニマーの検定（片側）によって分析した結果、 $p < .001$ ）。 $p$ と $q$ という組み合わせだけでなく、さまざまの組み合わせのカードのもとでこういう評価を行うことによって、被験者は、自分の選択について再考を促され、反証することに気づき、2度目の4枚カード課題では正答率が上昇したと考えられるのである。

Legrenzi (1971) では、ルールを与えて適切なカードを選択させるといった単なる選択課題による統制群と、ルールそのものも曖昧にしておいてまず被験者に発見させ、それから選択課題を与えるという発見+選択課題群との間で正答率を比較した。この発見+選択課題というのは、両面提示カードが被験者に示され、それがルールを確認するカードであれば+の符号が、ルールを反証するカードであれば-の符号が与えられた。被験者は、カードと符号の種類をあわせて考えながら、どういうルールのもとでそうなっているのかを発見し、その発見したルールを用いて選択課題を行なうのである。その結果、発見+選択課題群の方が有意に成績がよかった（正答率は、実験1では実験群43.3%；統制群23.3%，実験2では実験群83.3%；統制群16.7%。検定は、フィッシャーイェツ直接法検定[片側]で、実験1は $p < .01$ ；実験2は $p < .0025$ ）。のことから被験者は、ルールを発見するという過程を経ることで、ルールに含まれる論理の理解が促進されたと考えられる。

Goodwin & Wason (1972) では、前後2回の選択課題の間に、たとえば「 $\neg q \cdot p$ 」のような組み合わせの両面提示カードを被験者に示した。つまり、ルールの反例となるカードが存在しうることを示唆したのである。しかし、実験群と統制群の間に正答率の差は見られなかつた（統制群も実験群も18.8%）。これは、被験者が單にカードを示されただけだったので、問題に含まれる論理に気付くための能動的な手順を踏むことができなかつたからだと考えられる。

Lunzer, Harrison & Davey (1972) では、カードの裏側にはどういう情報がありうるかを考えさせる操作を行なった。しかし、1回しか行なわれなかつたからか、正答率にはほとんど影響しなかつた。

### 主題化効果

Wason & Shapiro (1971) の実験2では、抽象課題を統制群とし、実験群には「乗り物と行き先」といった具体的な名辞を用いた課題を準備した。これは、「マンチェスターへ行くときにはいつでも車で行く」といったようなルールについて「マンチェスター( $p$ )」「リーズ

( $\neg p$ )」「車( $q$ )」「列車( $\neg q$ )」といったカードで選択させるものである。その結果、両群の正答率に有意差が見られた（フィッシャーイェツ直接法検定[片側]で、 $p < .004$ ）。実験1で正答率には有意差が見られなかつたことと考えあわせると、問題に含まれる論理に気づかせようとするよりも、具体的な名辞によるルールを用いた方がより大きな効果が得られると結論づけた。そして、このような具体的な名辞による正答率上昇の効果を、Wason & Shapiro (1971) は「主題化効果」と呼んだ。

Johnson-Laird, Legrenzi & Legrenzi (1972)においても、主題化効果が同様に認められた。この実験では、「If  $p$  then  $q$ 」と「P only if  $q$ 」というルールの間の比較も行なわれ、同時に主題化効果との交互作用も検討された。と言うのも、if thenよりも only ifの方が対偶に対する推論が行ないやすいと考えられるからである。その結果、only ifの効果以上に主題化効果がはっきりと示された（only ifの抽象課題の正答率は20.8%であり、通常 if then の抽象課題で得られる正答率よりも高かつたが、有意ではなかった。しかし、抽象課題と具体課題の間は、サインテストで $p < .000002$ ）。

Lunzer, et al. (1972) では、主題化効果と論理的な考え方を引き出すためのヒントの効果との交互作用が検討されたが、主題化効果のみ実証された。

これらの研究以降主題化効果がより深く検討されていくようになる。

### その他

初期の研究ではその他に、カードの描き方など細かなところを修正しながら結果を検討していく研究がいくつ見られる。たとえば、Wason & Johnson-Laird (1970) では、「カードの一方の面」と「もう一方の面」と言った表現ではわかりにくいのではないか、というそれまでの反省をもとにカードを両面提示している。そして、「カードの一方の面」に当たる部分だけ見えるようにして、「もう一方の面」に当たる部分にはマスキングをかけて見えなくするという工夫をした。また、 $\neg p$ や $\neg q$ といったカードは、(a) $p$ や $q$ でない別の情報を書き込むことで表現できる一方、(b)白紙提示することもできる。そこで、(a)と(b)を比較した。しかし、こういったことは、結果的に正答率への影響がみられなかつた。Johnson-Laird et al. (1972) も、情報を書き込んだ封筒と白紙の封筒の両方を被験者に与えている。しかし、その選択率の違いなどは分析されていない。

### まとめ

1970年代の初めごろまでは、4枚カード課題の解決要

因は、問題に含まれる論理に気付かせることであると考えられてきた。そして、そのために大別して次の2つが考えられてきた。すなわち、(a)論理的な考え方を引き出すヒントを与えること、および(b)日常の具体的な名辞をルールに用いること、である。

(a)については、その後研究が衰退していく。なぜならば、Lunzer, et al. (1971) が示したように、主題化効果と比較して、その影響力が弱いと考えられるからである。

#### 4枚カード課題研究の成熟期1：(1970年代)

以後、主題化効果を中心とした研究の流れが続いた。同時に、別の流れとして、「マッチングバイアス(matching bias)」などの新たな考え方が浮上していく。というのも、当時4枚カード課題に失敗する原因是確証バイアスであるという考えが一般的であったが、これは、前件pと後件qの関係を正確につかめないことを意味している。とすれば、pとqの関係をルールの上でさまざまに変化させても有効な方法ではないかと考えられるからである。

#### 主題化効果

Bracewell & Hidi (1974) は、主題化効果の中でも、具体的にどういう側面が正答率を上昇させているかを検討している。pやqが英数字ではなく乗り物や行く先という具体的な名辞だから正答率が高いのか、pとqの関係がカードの表裏でしか表現されないのでなく、それに乗ってどこかへ行くというようにカードを離れても具体的な関係がイメージできるから正答率が高いのかを検討したのである。

そこで、以下のようなルールを用いて実験を行った：

- (1) オタワに行くときはいつでも自動車で行く（名辞=具体・関係=具体）
  - (2) Jへ行くときはいつでも2で行く（名辞=抽象・関係=具体）
  - (3) (カードの)一方の面がオタワのときはいつでも、もう一方の面は自動車である（名辞=具体・関係=抽象）
  - (4) (カードの)一方の面がJのときはいつでも、もう一方の面は2である（名辞=抽象・関係=抽象）
- その結果、関係の具体性に主効果がみられた ( $F(1,88)=21.83, p<0.01$ )。この結果について、Bracewell & Hidi (1974) は以下のように考察している。実験では教示の中で「このルールは双条件ではない」ということが強調された。つまり、「逆も真なり」と考

えてはいけないと忠告が与えられたのである。これは、問題に含まれる論理に気付かせるためのヒントであるといえる。しかし、実際にこのことを念頭において問題解決をするには、それなりの認知負荷がかかり、作業の負担となる。しかし、名辞間の具体的な関係が示されたことで、「逆は真に非ず」ということが認知負荷もなく当然のように考慮されることになる。例えば、「オタワに行くときはいつでも自動車で行く」というルールの場合、だからといって「自動車で行くときはいつでもオタワへ行く」とは限らないということはすぐに考えつくことができる。このようなことは、ルールの名辞を具体的にしただけでは被験者には伝えられない。

だから、名辞の具体性のみを操作している Roberge (1978) では、抽象課題と具体課題の間で正答率に差は見られなかった。

一方、Gilhooly & Falconer (1974) では、Bracewell & Hidi (1974) や Roberge (1978) とは逆に、ルールの名辞の具体性の方が有意に影響しているという結果が示された。この結果について Gilhooly & Falconer (1974) は、具体的な名辞によってその名辞に関係する過去の経験の想起、特に反例の記憶が想起されることによって確証バイアスに陥りにくくなるからであると説明している。しかし、もしその説明が正しければ、過去に反例に関する経験がない被験者には正答できないことになる。実際、この実験結果を支持するような結果はこれ以降ほとんど見い出されていない。しかし、主題化効果に直接経験が必要か否かという問題は重要であり、この後主題化効果に関する研究が展開していくきっかけとなっている。

#### マッチングバイアスと二重過程論

マッチングバイアスとは、「If p then q」というルールでの4枚カード課題の正答は「 $p \cdot q$ 」であるが、ルールの名辞に否定辞を加えていくと、正答も変わる (Table 4 参照)。ところで人間は、否定辞を含めない名辞のみに着目して、そのまま名辞と一致するカードを選択してしまう傾向があるという (Evans, 1972a; 1972b; Evans & Lynch, 1973)。つまり、ルールに否定辞がどのように含まれっていても、選択パターンは「 $p \cdot q$ 」ということになる。このような選択パターンの偏りを一般に「マッチングバイアス」と呼ぶ。この説明によれば、ルールが Table 4 の(c)の時に正答率が最も高いことになる。この仮説は、おおむね実験的検証によって支持されてきている。

マッチングバイアスに関する研究 Evans (1972a) は、

Table 4 Relation between negation of term and response pattern

rule/response	correct	matching	confirmation
(a) If p then q	p q̄	p q	p q
(b) If p̄ then q	p̄ q̄	p q	p̄ q
(c) If p then q̄	p q̄	p q	p q̄
(d) If p̄ then q̄	p̄ q̄	p q	p̄ q̄

マッチングバイアスに関する研究に先立ち、ルールに否定辞を加えた場合の影響について検討している。その結果、「If p̄ then ...」といったように、前件に否定辞が加わったものが推論を最も難しくしていることがわかった。彼の実験では、否定式推論 (*modus tollens*: たとえば「If p then q」というルールのとき, q̄ならば, p̄であるという推論)を行なわせている。だから、前提に否定辞が加わったルール(例えば If p̄ then q)だと否定式推論は「If q̄ then (p̄)」となり、帰結が2重否定となってしまう。これは、論理的には「If q̄ then p」と同値であるが、日常では、「p でないわけではない」は、必ずしも「p である」と同一ではない。こういったことが被験者を混乱させていると Evans (1972a) は考えた。

そこで、Evans (1972b) では、ルールを確証もしくは反証するようなパターンでカードを選択させたところ、確証事例はどのように否定辞を加えたルールでも正しく答えられるものの、反証事例となると、やはり、「If p̄ then q」の正答率が最も低かった。また、全体に、前件が p̄ の場合は反証事例の正答率は低く(サインテストで,  $p < .005$ )、後件が q̄ の場合は、反証事例の正答率は高かった(サインテストで,  $p < .012$ )。それは、ルールの否定辞に関わりなく「p」と「q」のカードを選択する比率が高いからである。Evans (1972b) は、これらの現象をまとめて「マッチングバイアス」と呼んだ。

Evans & Lynch (1973) は、この Evans (1972a; b) を受けて、もっとも基本的な4種類のルール(Table 4 参照)で実験を行っている。その結果、ルールの否定辞に関わらず、「p • q」というパターンを選択する割合が有意に高かった(サインテストでほぼ  $p < .002$ )。ちなみに、正答率は「If p then q̄」の時が最も高く、他が10%前後なのにたいして58.3%もあった。ただし「If p then ...」の様なルールでは p カード、「If p̄ then ...」の様なルールでは p̄ カードが選択される割合が比較的高かった。しかし、後件(qもしくはq̄)に関してははっきりとマッチングバイアス的な反応が確認されている。

ほかにも4枚カードの実験ではないが、条件文による三段論法に否定辞を加えて実験を行った Roberge

(1971) の研究でも、「If p then q̄」の正答率が最も高かった。

Evans (1977b) はさらに、この Evans & Lynch (1973) を始めとするいくつかの研究をもとに、被験者の反応パターンに対して(a)解釈的因子(*interpretational factor: I*)と(b)単純反応因子(*response factor: R*)を抽出した。(a)は被験者の解の論理的傾向(logical tendency)を示し、(b)はマッチング的傾向(matching tendency)を示すものである。被験者が具体課題を解決する場合は、ルールの理解が促進され、より論理的に考えることが可能になると Evans (1977b) は考えた。つまり、(a)因子の値が高いことになる。また、否定辞を操作した課題を解決する場合は、(b)因子の値が高いと考えた。このように、2つの因子は課題によって異なり、一方の値が高ければ一方の値は低くなる。そこで、この因子の値の高さを  $\alpha$  という重みづけで表現したのが式(1)である。

$$Pr(r) = \alpha \cdot I + (1 - \alpha) \cdot R \quad (1)$$

ここで、 $Pr(r)$  は、あるカードパターン反応が出現する確率を示す。たとえば、完全にマッチング的に反応した場合、 $\alpha = 0$  となり、 $Pr(r) = R$  となる。

この式による考え方をもとに、彼は、先ほどの Evans & Lynch (1973) のデータを再分析した。その結果、前件に関する判断では(a)因子が高く、後件に関する判断では(b)因子が高いということを見い出した。この結果は、Evans & Lynch (1973) の考察を支持した。

ところでマッチングバイアスは、2つの点から注目された。1つは確証バイアスを超えた点である。確証バイアスは、「If p then q」というルールの反応の予測しかできない。しかし、マッチングバイアスは、否定辞を含めたルールの反応の予測が可能である。もう1つは抽象課題の正答率の操作が可能になった点である。それまでの説明は、人間が論理的推論問題を解決するためには論理的推論を行なうことが必要であり、素材を具体的にすることで論理的推論が喚起されるでというものであった。しかし、マッチングバイアスが登場した結果、抽象課題の正答率の上昇も認められ、論理的推論問題を解決する別の過程の存在が示唆されるようになった(後述)。

しかし、マッチングバイアスが完全というわけではない。具体課題の正答率がなぜ高いのかについて、マッチングバイアスは何も説明していないからである。

Evans (1972c) は、マッチングバイアスは操作的要因 (operational factor) であり、具体課題ではそのバイアス効果が軽減されてしまうと示唆している。マッチングバイアスが抽象課題に適用できることは上述したが、抽象課題だからこそ影響力が強まったともいえる。さらに、Van Duyne (1973; 1974) や Roberge (1978) は、マッチングバイアスが、提示されるルールの表現によって起りやすかったり起りにくかったりすることを立証した。一般に、if then ルールにおいてもっともよく現われることが確認されている。この結果は、柴田 (1993) においても支持された。このようなマッチングバイアスの限界について、Evans は、大筋でそれを認めている (Evans, 1975)。

**二重過程説とマッチングバイアス** このように、マッチングバイアスは短期間のうちに注目を集め、そしてその限界が指摘されてきた。しかし、マッチングバイアスが、限定的な状況においてではあれ、存在することは事実である。そのためマッチングバイアスにより示唆されるようになった、論理的推論問題を解決する別の過程の存在についての研究が進められた。そのさきがけに相当するのが、Wason & Johnson-Laird (1970) の実験である。この実験では、2回の選択課題の間に「評価過程」と呼ばれる問題を被験者に与えて、その結果、正答率が有意に上昇している。この評価過程は、ルールとカードとの間の関係について評価させる問題であるが、実験者側の操作といえる評価過程の影響で正答率が上昇したということは、被験者自身の力では、評価過程のような認知行動をとれないことになる。では、マッチングバイアスによって正答率が上昇していると観察されたとき、被験者は問題に含まれる論理を理解するような認知行動をとっているといえるのだろうか。そこで、否定辞の操作を行った実験と同時にプロトコル分析を行って、被験者がどのように考えて問題解決をしたのかが詳細に検討された (Wason & Evans, 1975)。

Wason & Evans (1975) は、後件のみ否定辞の操作を行なった課題を用いて実験をし、さらに「なぜそういう選択をしたか」について被験者のプロトコルを記録した。その結果、被験者のカード選択は後件が肯定の場合も否定の場合もマッチングバイアス的であったが、選択理由については一見非常に論理的な理由を述べている被験者がいることがわかった。しかし、それらの理由に一貫性が見られなかった。例えば、「もし一方がBなら

ばもう一方は3である」というルールと「もし一方がBならばもう一方は3ではない」というルールに対して、ある被験者は、両ルールとも「B」と「3」のカードを選択している。これは明らかにマッチング的反応である。しかしその選択理由を尋ねると、前者に対しては“ルールはBと3に関するものだから、それ以外には言及せず、この2つのカードのみを選択すればよい”と述べている一方で、後者に対しては“もしルールが偽ならばもう一方の側は3であり、真ならば3ではないはずである。Bと3は一緒に考えなければならない”と述べている (p. 147)。別の被験者は、前者に対して “Bの裏側に3が、3の裏側にBがあるならば、ルールは真となるからである”と述べ、後者に対して “Bの裏側に3が、3の裏側にBがあるならば、ルールは偽となるからである”と述べている (p. 148)。このように、カード選択を行なった後、とてつけたようにして自分の選択について正当化するような理由をつけているのである。つまり、カードを選択する認知過程と「なぜそのカードを選択するのか」という理由づけの認知過程が別個に働いているのである。Wason & Evans (1975) は、こういった現象について、(a)行動主義者説 (behaviorist hypothesis) (b)合理説 (rational hypothesis) (c)二重過程説 (dual process hypothesis) を仮定した。

(a)は、プロトコルは言語的反応に過ぎなくて、カード選択という行動と結びついたものではないという仮説である。しかし、Wason & Evans (1975) は、確証的なプロトコルを述べた人のうち8.3%の人しか反証するようなカード選択をしなかったと認められるのに対し、反証的なプロトコルを述べた人の75%の人が反証カード選択をしたと認められる、という結果から(a)は疑わしいと主張した。

(b)は、プロトコルはカード選択理由そのものであるというのである。しかし、否定辞のある課題ではいかにも反証事例を考えたというプロトコルを述べたのに、否定辞のない課題では突然反証事例を考慮しないプロトコルを示してしまう被験者が観察された。また、Evans & Lynch (1973) の実験で見られるように、否定辞の位置によっては、まったく論理的正解とは異なる解を示してしまう被験者がいることから、これも疑わしいとした。そして、最終的に(c)と結論づけた (Wason & Evans, 1975)。

Evans & Wason (1976) では、あらかじめカード (正解のパターンのものもあれば、不正解のものもあり、群ごとに異なる) をルールとともに被験者に与え、「このルールの場合そういう解でよいのかどうか、またどうしてそう思うのか」を尋ねた。その結果、与えられたカ-

ドセットが正解であろうと不正解であろうと被験者はその選択を正しいと認めた。しかもその選択に関して, Wason & Evans (1975) に見られたように、一貫性のない理由づけを行っていたのである。さらに被験者は、自分の理由づけに対する自信度も非常に高かったことにより、曖昧な気持ちで答えているわけではないことも明らかになった。もし、被験者が、カード選択について表面的にしか考えておらず、そのことを自分で認識していたとしたら、自信度はもっと低くなるはずである。一方、自信度の高さに見合うほどしっかりと論理的に考えていたとしたら、与えられたカードセットについて、正解のものと不正解のものを見分けることができたはずであるし、与えられたカードセットごとに異なる理由づけをするような一貫性のない行動はとらないはずである。つまり、被験者のカード選択行動とその理由づけ行動は、別々の処理行動であり、被験者自身はそのことに関する認識はないことになる。

上述の2つの実験的研究から、EvansとWasonは、マッチングバイアス的なカード選択をしている被験者の実際の推論過程について直接は触れないものの、カードの選択理由として後から尋ねたものについては、実際の選択理由とは異なるものであると結論し、二重過程説を主張した。

### その他のルール表現効果

ルールの表現を変えるだけで正答率が変化するというマッチングバイアスの影響を受けて、さまざまにルールを変化させた研究がいくつも現われた。それは、上述したように、マッチングバイアスの限界を指摘することすらあった。

この種の研究は、次の2通りの研究の流れを確認できる。1つは、「If p then q」と論理的には同値であるさまざまな表現のルールにおける反応パターンの違いを考察するものである。もう1つは、pとqの順序関係に焦点を当てるものである。「If p then q」といった場合、必ずpが先行するような表現になってしまふ。それを「Q only if p」といったように表現し直して、被験者の反応を考察しようというのである。もちろんさまざまな表現において検討されている。

**論理的に同値なルール間の比較** まずは、4枚カード研究でも比較的初期の論文に注目することができる。Johnson-Laird & Tagart (1969) では、通常用いられる条件文的なルール ((a) If p then q) だけでなく、さまざまな表現によるルール ((b) There isn't p, if there isn't q • (c) Either there isn't p, or there is

q • (d) There is never p without there being q) を用いて実験を行っている。彼らは、両面提示カードを利用して、それらがルールを確証しているか、反証しているか、あるいはルールとは無関係であるかを決定させた(つまり評価課題)。その結果、ルールの種類によって確証カードと反証カードの決定に有意に偏りがあることがわかった(2項検定で $p<.01$ )。また、正答率は(c)が最も高く、(a)と(d)の反応が非常に似ていることも確認された。

Van Duyne (1974) は、(a) Every p is q, (b) If p then q, (c) Not p or q, (d) It isn't the case that p and not q の4つのルール表現を用いて、その正答率を比較している。その結果、抽象課題ではルール表現による正答率の違いはみられなかつたが、具体課題では、(a)と(b)が残りの2文よりもより正答した(ウィルコクスン(片側)で $p<.005$ )。Wason (1968)によれば、(a)と(b)のルールは、被験者が同値であると認識することが確認されており、(b)のかわりに(a)を用いて実験をしている論文もいくつか見られる(例: Bracewell & Hidi, 1974)ことから、この結果は妥当であるといえる。

しかしながら、Johnson-Laird & Tagart (1969)の結果と Van Duyne (1974) の結果は一致しない。理由を考えるとすれば、Johnson-Laird & Tagart (1969) では評価課題だったのにたいして、Van Duyne (1974) では選択課題だったので、そのあたりが結果の違いに影響したのかもしれない。どちらかといえば、Van Duyne (1974) の結果の方が支持されているようである。

Roberge (1978) では、(a) Either p or q (but not both), (b) Either p or q (or both), (c) If p then q, (d) P only if q の4文を用いて実験を行なった結果、(d)の成績が有意によいなど、文章表現ごとに正答率の差がみられた( $F(1, 60)=15.16, p<.001$ )。Johnson-Laird, et al. (1972) でも、この傾向は確認されている。

まとめると、被験者の反応はルールの表現によって異なるが、それによって正答率が上昇するようなルール表現というのは、課題間で一致しているとは言い難い。また、正答率に有意差をもたらすほどのルール表現は、なかなか現われにくい。むしろ、こうやってルール表現を変化させることで、誤答のパターンが多様化しているようである。

**pとqの順序関係** Wason & Golding (1974) では、ルールを(a) Above p is q と(b) Q is above p のように表現した。通常の「If p then q」という表現では、常に前提pが結果qに先行して表現される。誤反応の特徴

## 資料

は、 $q$  カードを選択して $\neg q$  カードを選択しないことに集中していることから、 $q$  が先行するように表現すれば、 $q$  へ注意がより向けられると彼らは考えた。そこで、(a) と(b)の 2 つのルールを比較したのである。結果は、(b)の方方が有意に正答率が高く、彼らの仮説が実証された（フィッシャーの直接法検定で  $p < .05$ ）。同時に、彼らは、「なぜそういう選択をしたのか」という被験者のプロトコルにも耳を傾けるべきだと主張している。被験者のプロトコルからは、一方向からしか思考できない人間の推論の特徴が浮き彫りにされている。つまり、「Above p is q」では  $p \rightarrow q$  のようにしか推論できないし、「Q is above p」では  $q \rightarrow p$  のようにしか推論できないのである。その結果、後者では、 $\neg q$  ならば  $\neg p$  であるという否定式推論を導き出せたと考えられるのである。

しかし、Bracewell & Hidi (1974) の実験では、逆の結果が出ている。彼らは、(a) Everytime  $p, q$  と(b) Q, everytime  $p$  の 2 つのルールを用いて実験した。その結果、具体課題では前者のほうが正答率が有意に高かった（分散分析の下位検定で  $p < .05$ ）。

2 つの結果の解釈として、前者は、 $p$  と  $q$  のどちらが先行しても基本的に考える過程に変わりはない。その場合、先行する名詞により多くの注意が払われると考えられるため、(b)の方が正答率が高かった。しかし後者では、 $p$  は明らかにルールの前提を意味し、 $q$  は結果を意味する。この場合、人間の考えるプロセスとしては、前提→結果という順序をたどるのが自然である。そのために、「Everytime  $p, q$ 」のほうが正答率が高いと考えられる。ただし、Evans (1977a) では、(a) If  $p$  then  $q$  (前提が先行するルール) と(b) P only if  $q$  (帰結が先行するルール) の 2 つの表現を用いて実験を行なったが、(a) ルールの方が有意に肯定式推論 (*modus ponens*: たとえば「if  $p$  then  $q$ 」というルールのとき、 $p$  ならば、 $\neg p$  であるという推論) が多く（ケンドールの S テスト（片側）で  $p < .002$ ）、(b) ルールの方が有意に否定式推論が多かった（同じく  $p < .05$ ）。

Bracewell & Hidi (1974) で用いられたルールを具体的に述べると、「オタワに行くときはいつでも車でいく」と「車でいく、オタワに行くときはいつでも」のように、前件と後件がひっくり返っているだけで、使われている表現はまったく同じである。だとしたら、前提から結果へという推論を自然にたどることができ、文章表現としても自然な「オタワに行くときはいつでも車でいく」の方が正答率が高いのも納得できる。しかし、Evans (1977a) の場合、「文字が L ならば数字は 5 である」と「数字が 5 であるときのみ文字は L である」は、表現そのものが異なっている。つまり、if ではなく

only if というところに否定式推論をおこさせやすい要因があるのではないかと考えられるのである。実際、Evans (1977a) の被験者は、過半数が(a)と(b)のルールを別の意味をもつものとして捉えていた。Roberge (1978) においても上述のように only if ルールの正答率は有意に高かった。

以上の結果から、正答率を上昇させるための  $p$  と  $q$  の順序は、(a) 人間の推論の順序に合致すること(b) ルールの意味が伝わりやすく、否定式推論を思いおこさせやすいこと、が条件であると考えられる。

## 本論文のまとめ

まず、4枚カード課題の説明および研究が始まった背景が述べられた。4枚カード課題の正答率はきわめて低いことで知られるが、その正答率が低い要因として確証バイアスが挙げられ、確証バイアスの克服のために、論理的に考えるためのヒントや主題化効果が示された。その後、さらに主題化効果の研究が盛んになるが、それとともに、否定辞の操作によるマッチングバイアスが登場してきた。しかし、マッチングバイアスの限界が示されると、今度はルール表現が推論に与える影響についての研究が盛んになった。

## 文 献

- Bracewell, R. J. & Hidi, S. E. 1974 The solution of an inferential problem as a function of stimulus materials. *Quarterly Journal of Experimental Psychology*, 26, 480-488.
- Bucci, W. 1978 The interpretation of universal affirmative propositions. *Cognition*, 6, 55-77.
- Evans, J. St. B. T. 1972a Reasoning with negatives. *British Journal of Psychology*, 63, 213-219.
- Evans, J. St. B. T. 1972b Interpretation and matching bias in a reasoning task. *Quarterly Journal of Experimental Psychology*, 24, 193-199.
- Evans, J. St. B. T. 1972c On the problems of interpreting reasoning data: Logical and psychological approaches. *Cognition*, 1, 373-384.
- Evans, J. St. B. T. 1975 On interpreting reasoning data: A reply to Van Duyne. *Cognition*, 3, 387-390.
- Evans, J. St. B. T. 1977a Linguistic factors in reasoning. *Quarterly Journal of Experimen-*

- tal Psychology*, 29, 297-306.
- Evans, J. St. B. T. 1977b Toward a statistical theory of reasoning. *Quarterly Journal of Experimental Psychology*, 29, 621-635.
- Evans, J. St. B. T. & Lynch, J. S. 1973 Matching bias in the selection task. *British Journal of Psychology*, 64, 391-397.
- Evans, J. St. B. T. & Wason, P. C. 1976 Rationalization in a reasoning task. *British Journal of Psychology*, 67, 479-486.
- Gilhooly, K. J. & Falconer, W. A. 1974 Concrete and abstract terms and relations in testing a rule. *Quarterly Journal of Experimental Psychology*, 26, 355-359.
- Goodwin, R. Q. & Wason, P. C. 1972 Degrees of insight. *British Journal of Psychology*, 63, 205-212.
- Henle, M. 1962 On the relation between logic and thinking. *Psychological Review*, 69, 366-378.
- Johnson-Laird, P. N., Legrenzi, P. & Legrenzi, M. S. 1972 Reasoning and a sense of reality. *British Journal of Psychology*, 63, 395-400.
- Johnson-Laird, P. N. & Tagart, J. 1969 How implication is understood. *American Journal of Psychology*, 82, 367-373.
- 川村久美子 1991 知識の獲得 丸野俊一編 新児童心理学講座5 概念と知識の発達 金子書房 3-58.
- Legrenzi, P. 1971 Discovery as a means to understanding. *Quarterly Journal of Experimental Psychology*, 23, 417-422.
- Lunzer, E. A., Harrison, C. & Davey, M. 1972 The four-card problem and the generality of formal reasoning. *Quarterly Journal of Experimental Psychology*, 24, 326-339.
- Roberge, J. J. 1971 Some effects of negation of adults' conditional reasoning abilities. *Psychological Reports*, 29, 839-844.
- Roberge, J. J. 1978 Linguistic and psychometric factors in propositional reasoning. *Quarterly Journal of Experimental Psychology*, 30, 705-716.
- Roberge, J. J. & Paulus, D. H. 1971 Developmental patterns for children's class and conditional reasoning abilities. *Developmental Psychology*, 4, 191-200.
- Shapiro, B. J. & O'Brien, T. C. 1970 Logical thinking in children aged six through thirteen. *Child Development*, 41, 823-829.
- 柴田淑枝 1993 論理的推論課題における論理的推論の利用：ウェイソン課題の解決過程の検討 名古屋大学大学院教育学研究科修士論文（未公刊）
- Staudenmayer, H. & Bourne, L. E. Jr. 1977 Learning to interpret conditional sentences: A developmental study. *Developmental Psychology*, 13, 616-623.
- Taplin, J. E., Staudenmayer, H. & Taddonio, J. L. 1974 Developmental changes in conditional reasoning: Linguistic or logical? *Journal of Experimental Child Psychology*, 17, 360-373.
- Van Duyne, P. C. 1973 A short note on Evans' criticism of reasoning experiments and his matching response hypothesis. *Cognition*, 2, 239-242.
- Van Duyne, P. C. 1974 Realism and linguistic complexity in reasoning. *British Journal of Psychology*, 65, 59-67.
- Wason, P. C. 1966 Reasoning. In B. N. Foss(Ed.), *New Horizons in Psychology* I. London: Penguin. 135-151.
- Wason, P. C. 1968 Reasoning about a rule. *Quarterly Journal of Experimental Psychology*, 20, 273-281.
- Wason, P. C. 1969 Regression in reasoning? *British Journal of Psychology*, 60, 471-480.
- Wason, P. C. & Evans, J. St. B. T. 1975 Dual processes in reasoning? *Cognition*, 3, 141-154.
- Wason, P. C. & Golding, E. 1974 The language of inconsistency. *British Journal of Psychology*, 65, 537-546.
- Wason, P. C. & Johnson-Laird, P. N. 1970 A conflict between selecting and evaluating information in an inferential task. *British Journal of Psychology*, 61, 509-515.
- Wason, P. C. & Shapiro, D. 1971 Natural and contrived experience in a reasoning problem. *Quarterly Journal of Experimental Psychology*, 23, 63-71.

(1996年9月13日 受稿)

## ABSTRACT

### Review of the research for Wason's four-cards task 1: From 1966 to 1979

Yoshie SHIBATA

This article is a review of the research of Wason's four-cards task, which was focused on the research from 1966 to 1979 as the first half of it. In the introduction, we described what was the background of the research and what is the four-cards task. In the chapter of "the dawn of the research", it was mentioned that the percentage of correct response by adults was very low, which was caused by the "confirmation bias". And, it was discussed that it was important to think logically for solving problems, and that it was effective (a)to give some hints to understand the logic of the problem and (b)to use the concrete theme in a rule(that is called "thematic effect") for logical reasoning. Especially, (b) was regarded as more important. In the chapter of "the mature period of the research", it is described that (b) was influenced by the relationship among terms in a rule, and that the response pattern of abstract task was varied with the expression of rules. "Matching bias", that arises by negation in a rule, was noteworthy because (a)response patterns of subjects could be predicted by it, and (b)it is applicable to abstract task. Then, it was tried to explain the process of thinking process as the "dual processes". It was pointed out, however, that (a)matching bias didn't contain the explanation about the process of solving concrete task, and (b)it is hard to appear in some rules. Then, especially for (b), it was discussed that some rule-expressions that agree with the order of human reasoning, is more effective.