

## 研究論文

Wason 選択課題におけるデータ選択傾向：  
推論におけるファラシーと合理性をめぐって服部 雅史<sup>1) 2) 3)</sup>The Tendency of Data Selection in the Wason Selection Task:  
Human Fallacy and Rationality in Reasoning

HATTORI Masasi

**Abstract**

In contrast to other theories in the psychology of reasoning, according to the optimal data selection (ODS) model (Hattori, 2002; Oaksford & Chater, 1994), data selection performance in the Wason selection task is rational. In this model, the tendency to select data is defined by a *selection tendency function* (STF). The research reported here considered the hypothesis that there may be minor (cultural) variation in the tendency to select data. This hypothesis was tested by estimating the parameters of the STF from new experimental data using a group of Japanese participants and from previous data in the literature using Western and Japanese participants. The results revealed that the ODS model can rationally capture the performance of participants across different experimental groups by a simple adjustment of the STF.

**Key words** : optimal data selection model , information gain model , selection tendency function , Bayesian belief updating , cultural and individual variations in reasoning

**キーワード** : 最適データ選択モデル , 情報獲得モデル , 選択傾向関数 , ベイズの信念更新 , 推論の文化差と個人差

1) 立命館大学文学部

2) 本研究の一部は、英国カーディフ大学心理学部における在外研究になされた。カーディフ大学滞在中は、マイク・オクスフォード教授、および他のスタッフの方々のおかげで、極めて良好な研究環境を享受できた。また、当在外研究は、立命館大学心理学科の同僚の理解を得て実現した。ここに記して心からの謝意を表したい。

3) 本稿の準備にあたり、大和日英基金、日本学術振興会科学研究費第14310045号(基盤研究B)、立命館大学人間科学研究所(プロジェクト研究B)の資金援助を受けた。

Wason 選択課題 (Wason, 1966) は、非常にシンプルな課題にも関わらず多くの人が間違えるのはなぜかという点をめぐって、これまで、膨大な数の実験的研究と多くの議論がなされてきた。そもそも、エラー (間違い) やファラシー (錯誤) とは、その判断基準に依存した相対的概念であり (松田, 2003), 基準は、ア・プリオリにも一意的にも定められるものではない。Wason 選択課題は、推論研究においてこれまで

最も精力的に研究されてきた課題であるが、そこでの人間のファラシーは、主に論理的観点から論じられてきた。この課題に関して論理学は、客観的かつ自明な基準としての役割を果たしてきたと言える。しかし論理は、コンピュータ科学の発展以降に提唱された様々な非古典的体系まで含めたとしても、計算論的水準 (Marr, 1982) において、人間の推論の理論として適切な枠組みではない (Oaksford & Chater, 1991)。われわれの日常生活においては、論理的意味で確実なものはごくわずかであり、多くのものごとは、せいぜい尤もらしいと言えるだけであることを考えれば、むしろ人間の推論過程は、確率論的基準に照らし合わせてみた場合に、その合理性が見えてくるかもしれない。

本研究は、Wason 選択課題における人間の行動が、文化などの背景の異なる被験者のパフォーマンスの違いまで含めても、合理的な振る舞いとして説明されるかどうかを検討した。ここでは特に、定量的最適データ選択 (quantitative optimal data selection; QODS) モデル (Hattori, 2002) の枠組みに則り、被験者の選択傾向の違いをパラメータとしてモデルに組み込むことにより、単一の理論的枠組みによる説明を試みた。

### 最適データ選択モデル

Wason 選択課題において被験者には、「もしカードの片面に母音書かれている ( $p$ ) ならば、もう一方の面には偶数が書かれている ( $q$ )」という論理的規則と、 $E(p)$ ,  $K(\neg p)$  (これは not  $p$  と読む),  $4(q)$ ,  $7(\neg q)$  という4枚のカード (あるいは、その図柄) が与えられる。各カードの片面には文字、その裏面には数字が書かれている。被験者は、規則が真か偽かを確かめるためにどのカードを裏返すべきかを問われる。論理

的に正しい解は「Eと7」( $p$ と $\neg q$ ) であるが、通常10%以下の被験者しか正解できないとされてきた (Evans, Newstead, & Byrne, 1993; Wason & Johnson-Laird, 1972)。

Oaksford & Chater (1994) は、この課題を、純粋な論理的課題としてではなく、仮説検証のためのデータ選択という意味決定課題として定式化することに成功した。これまで、この課題における人間の行動は、論理のエラーや単なるバイアスとして認識されてきたが、彼らは、そのようなファラシーが、むしろ環境に適應した合理的な行動であることを明らかにした。彼らのモデルによれば、被験者のカード選択は、与えられた規則「もし  $p$  ならば  $q$  である」( $r_1$ ) の前件の確率  $P(p)$  と後件の確率  $P(q)$  によって定まる。被験者は、検証すべき仮説  $r_1$  に対立する仮説「 $p$  と  $q$  は独立」( $r_2$ ) を持つと仮定し、 $P=\{p, \neg p\}$ ,  $Q=\{q, \neg q\}$ ,  $R=\{r_1, r_2\}$  とする。すると、例えば  $p$  カードを裏返すことにより期待される不確実性の減少量は、 $I(p)=H(R|p) - H(R|Q, p)$  と書ける。ここで  $H(X)$  は、任意の確率変数  $X=\{x_i\}$  のエントロピーをあらわす。すなわち、 $X$  は、確率分布  $P(x_i)$  によって定められ、 $\sum_i P(x_i)=1$  として、 $H(X)=-\sum_i P(x_i) \log_2 P(x_i)$  と定義される。課題における各カードの選択は競合的となし得るので、カード  $x$  の期待獲得情報量  $I(x)$ ,  $x \in \{p, \neg p, q, \neg q\}$  は、全カードの期待獲得情報量の総和により尺度化され、次のようにあらわされる。 $I_s(x)=I(x)/\sum_i I(x_i)$ 。この値は尺度化情報量と呼ばれ、カード選択率に対応すると仮定される。

Hattori (1999, 2002) は ODS モデルを修正・拡張し、与えられた規則の前件と後件の被験者の主観的確率を、データから推定することを可能にした。選択傾向関数 (selection tendency function; STF) と呼ばれる関数により、各カードの尺度化情報量は、それが裏返される確率に写像されると仮定された。STF は2つのパラメ

ータ  $a, b$  を含むロジスティック関数によって定義され、次のようにあらわされる。

$$T(x) = f[Is(x)] = \frac{1}{1 + e^{-(a+bs(x))}} \quad (1)$$

以上の仮定より、各カードが裏返される確率は、前件の確率  $P(p)$  と後件の確率  $P(q)$  によって定まることになる<sup>1)</sup>。

式1のSTFパラメータ  $a, b$  が  $P(p)$  および  $P(q)$  と独立と仮定すると、最小二乗法によって、カード選択確率の実験データから、 $a$  と  $b$  を推定することが可能となる。 $N$  を全被験者数とすると、観察されるカード選択率は、二項分布  $B(N, T)$  によってあらわされるので、観察されるデータの尤度関数は次のようにあらわされる。

$$like(\mathbf{d}) = \prod_x \binom{N}{d_x} T(x)^{d_x} (1 - T(x))^{N-d_x} \quad (2)$$

ここで、 $d_x$  はカード  $x$  の選択された頻度、 $d$  は  $d_x$  のベクトルをあらわす。Hattori (2002) は、過去の文献の実験データを用いて、式2の右辺を最大化するような  $a, b$  を STF パラメータとして推定した。Figure 1 に示されている通り、推定値は  $a = -2.43, b = 9.27$  で、そのとき、 $P(p) = .23, P(q) = .25$  であった。

## 実験

STFパラメータは、カード反転によって獲得される情報の相対量、およびカード間の情報量の違いに対する敏感さをあらわす。それは、人々の背景知識、文化、あるいは、個人のパーソナリティによってさえも変動し得ると考えられる。Hattori (2002) はこの点を考慮して、考えられる差異を注意深く除去して STF を推定した。その結果、過去の文献のデータに対するモデルのフィットは極めてよかった。Read & Cressie (1988) に従って  $P \geq .01$  を基準としたとき（モデル棄却基準としては .05 は不合理に高いため）、47 個の実験のうちたった 1 個

（Table 3 の No. 8）だけがモデル棄却域のボーダー上にあり、残りの実験はモデルを支持した。

一方、日本人を被験者として得られた Hattori (2002) 自身の実験データは、モデルのフィットがあまり芳しくなかった。例えば、モデルのフィットのよさ（goodness-of-fit）の指標としての  $G^2(2)$  は、オリジナル抽象課題のデータに対しては、15.2（実験1、統制群）、および 14.1（実験2）であり、いずれも  $P < .01$  であった（Table 2 の No. 52, 53、また、Hattori, 2002, Table 5 も参照せよ）。これは、STFパラメータを推定するためのデータがすべて、西洋の文献から得られたものであったことと関係があるかもしれない。そこで、本実験は、日本人被験者を用いて STF パラメータを決定し、その結果が、Hattori (2002) を含めた文献のデータに対するモデルのフィットを高めるかどうかを検討するために実施された。

## 方法

**被験者** 立命館大学の学部生 419 人が被験者として実験に参加した。実験は、一般授業の 7 つのクラス（各クラス 45 人から 67 人）において、授業に先立って集団で実施された。以前に選択課題を見たことがあると答えた 22 人、および、年齢が 30 歳以上であると答えた 21 人（うち 2 人は年齢無回答）は分析対象から除外した。残る 376 人のデータが分析された。

**材料および手続き** 課題は、オリジナル標準課題（Wason, 1966）と等価なもので、教示は以下の通りであった。

一方の面にアルファベットが、別の面に数字が書かれているカードセットがあります。このカードセットについて、佐藤さんは、次の規則が成り立っていると主張しています。

「もしカードの片面が母音であるならば、別の面は偶数である。」  
ただし、母音とはA, E, I, O, Uの5文字を指します。

さて、カードセットから4枚のカードを抜き出して机の上に並べました。佐藤さんの言う規則が真か偽かを確かめるためには、下のどのカードを裏返す必要があるでしょうか。どうしても裏返す必要があると思われるカードに印をつけて下さい(複数可)。

教示の下に4枚のカード(E, K, 4, 7)の並びの図が示された。カードの並び方は、 $(p, \neg p, q, \neg q)$ ,  $(q, p, \neg q, \neg p)$ ,  $(\neg p, \neg q, p, q)$ ,  $(\neg q, q, \neg p, p)$ の4種類が用意され、被験者はいずれか1つに無作為に割り当てられた。

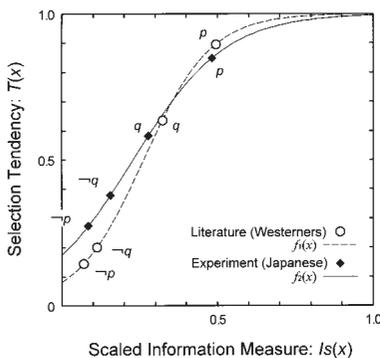


Figure 1. The selection tendency function (STF).

### 結果および考察

各カード選択率は、 $p, \neg p, q, \neg q$ の順に、.85 (319), .27 (103), .58 (219), .38 (142)であった(括弧の中の数値は376人中の頻度をあらわす)。この実験結果から推定されたSTFパラメータは、 $a = -1.56$ ,  $b = 6.81$ であった(Figure 1参照)。また、そのときの前件と後件の確率推定値は、 $P(p) = .27$ ,  $P(q) = .32$ であった。

Figure 1は、この実験において日本人被験者を用いて同定されたSTFと、西洋の文献データから推定されたSTFを示す。図より、両者がやや異なった形をしていることがわかる。明らかな違いは、前者がやや緩い傾きを持っていることである。一般に、傾きが緩くなるほど、尺度化情報量を異にする任意の2点(2枚のカード)間の選択傾向の差が小さくなる。したがって、西洋の文献データにおいては、本実験のデータに比べて、尺度化情報量が最も多い $p$ カードはより選択率が高く、最も少ない $\neg p$ カードはより選択率が低くなっている。この結果は、もしかするとNisbett (2003)において論じられているような何らかの文化差の側面と関係があるかもしれない。ただし、現時点では、このような一面的な結果から文化差一般に関する安易な結論を導くようなことは慎むべきであろう。

4枚全てのカードを選択した被験者の割合は、.10 (38)であった。これは、以前の西洋の研究のデータ(例えば、Wason, 1968; Hoch&Tschirgi, 1983, 1985; Chrostowski & Griggs, 1985)と比べてかなり高い。これらの文献の平均選択率は.05 (28/534)である。STFの形の違いは、この「全カード」反応の比率の違いも説明できる。Figure 1に示されるように、尺度化情報量の値がおよそ.35よりも小さいときは、日本人被験者のカード選択傾向は西洋人の場合より高い。尺度化情報量の平均値は.25であることから(全カードの総和が1ゆえ)、このことは、一般に、日本人の被験者のカード選択傾向を高めることになり、延いては、全カード反応率も高めたと考えられる。

同様に、日本人被験者の論理的正解カードの選択率が、西洋人の結果に比べて若干高い傾向があることも、STFの形の違いによって説明できるだろう。本実験において、論理的に正しい選択をした被験者の比率は.10であった。山(1999)のa/a条件(Table 2のNote参照)で

は.14, 堀内(2000)の“ $p \sim q$ ”条件では.11であった(いずれも通常の条件文を用いた条件)。Figure 1を見るとわかるように, $p$ カードも $\sim q$ カードも(特に後者が顕著),いずれも本実験データ(日本人)の方が文献データ(西洋人)に比べて高い選択率を示している。したがって,このことが,日本人被験者の比較的高い“ $p \& \sim q$ ”選択率(正答率)の主要な原因になっていると考えることができる。

### メタ分析

ここで,われわれは2つのSTFを得たことになる。一つは,本実験(日本人被験者)によって同定されたもの,もう一つは,西洋の文献データからHattori(2002)により推定されたものである。そこで,被験者群に応じて別々のSTFを想定するのが妥当か,あるいは,どちらか一方の推定結果が妥当性が高いかを吟味するため,メタ分析を行った。これらの各STFを用いて,過去の文献の様々な実験データから,被験者の潜在的な主観的確率 $P(p)$ と $P(q)$ を推定し,両結果を比較した。

Oaksford & Chater(1994)およびHattori(2002)は,オリジナル抽象選択課題に関する過去の実験データを再分析したが,それらはすべて西洋の文献のものであった。そこで,被験者群の同一性と差異を調べるため,またモデルの一般性を吟味するため,NACSIS-IR(国立情報学研究所)とPsycINFO(アメリカ心理学会)の文献データベースを用いて,日本人被験者を使って行われた直接法的選択課題(義務論的課題は除外)の実験を検索した。選択課題に関する14篇の論文が見つかり,うち9篇が直接法的課題を扱っていた。それらの中で,カード別の選択頻度が報告されていた4つの研究が分析された<sup>2)</sup>。

Table 1は,日本と西洋のデータ(Table 2と

3)に対する推定値のフィットのよさ(goodness-of-fit) $G^2$ を示す。フィットのよさは,「非常によい」(very good:  $P \geq .20$ )から「悪い」(poor:  $P < .01$ )までにカテゴリー化されている。この表では,モデルが予測する推定確率値の高低に応じて,推定確率値がHIGHまたはLOWに分類されている。課題の中で,否定辞を伴ったり明示的に高確率と述べられたりした前提(前件と後件の両方を含む)の推定値はHIGHに分類され,それ以外はLOWに含められた。ただし,次の2つのタイプのデータ,すなわち,Table 2のNo. 55, 57, 59, 61, および,Table 3のn/a条件(No. 38 - 45)は,この表から除外された。その理由は,前者の実験においては,実験の目的上,相矛盾する要因が意図的に含まれたが,そのことがこれらの実験の分類上の位置づけを難しくするためである。後者の実験は,モデルの前提条件, $P(p) \leq P(q)$ に反するからである。

LOWとHIGHの推定値の間には,すべての組合せにおいて,統計的に高度に有意な差が見られた。西洋のデータについては,文献データ(西洋の被験者を使用)(Table 3のNo. 1-20)から推定したSTFを用いた場合,LOWとHIGHの各平均値は.23 vs .45で, $t(76)=7.64$ であった。実験データ(日本の被験者)から推定したSTFを用いた場合,.16 vs .46, $t(76)=8.33$ であった。日本のデータについては,同STFを用

Table 1  
Frequencies of Categorized Goodness-of-Fit and Mean Values of Estimators for Data Shown in Table 2 and 3

Goodness-of-fit	Western		Japanese	
	Lit	Exp	Lit	Exp
$P \geq .20$ (very good)	38	30	11	16
$P \geq .05$ (good)	6	10	5	10
$P \geq .01$ (fair)	3	4	5	2
$P < .01$ (poor)	0	3	9	2
LOW	.23	.16	.33	.30
HIGH	.45	.46	.55	.58

Note: “Western/Japanese” indicates the region of data collected; “Lit/Exp” indicates whether estimation were conducted using STF fixed by data in literature (No. 1-20 in Table 3) or by Experiment reported in the text.

いた場合、それぞれ .33 vs .55,  $t(58) = 6.13$ , および .30 vs .58,  $t(58) = 6.50$ であった(すべて  $P < .01$ )。ただ、Hattori (2002) および Oaksford & Chater (2003) が彼らの推定結果において観察したように、高い確率は、本来期待されるほど高くはなかった。例えば、Hattori (2002) の推定結果に従って、もし  $P(q)$  を .25 とするならば、通常、 $P(\neg q)$  は  $1 - .25 = .75$  であることが期待されるが、HIGH の確率は、Table 1 に見られる通り、.45 - .58 の範囲に分布しており、.75 よりずっと小さい。推定におけるこのような「穏健性」の原因となる要因については、次節で詳しく論ずることにする。

### 考察

概して、西洋の文献のデータは、西洋の文献データから推定された STF をもつモデルによってよく説明され、同様に、日本の文献のデータは、日本人被験者を用いた実験データから推定された STF によってよく説明された。この結果は、これらの2群の被験者(すなわち、西洋人と日本人)は、平均的に見て、情報獲得に対して異なる敏感さを持つことを示唆する。

しかし、言語、背景知識、教育レベル、民族性など、結果に影響を及ぼし得る他の多くの要因が存在するので、この結果は、単純に文化の違いのみに帰属されるべきものではないだろう。この問題に関しては、多様なアプローチが不可欠であり、もっとずっと広い範囲の被験者からのデータを集める必要がある。現段階においては、いかなる決定的結論を導くのも時期尚早であろう。重要な点は、文化差を含む群間の差異の一側面が、STF の形態の違いとして客観的に捉えられたことにある。被験者群間の差異の客観的測度を提供するこのモデルは、人々の行動の差異を同定し、また、その要因を探るために、きわめて有効な手段の一つとなるだろう。

う。

メタ分析の結果は、ODS モデルの高い説明力を示した。しかし、モデルのデータに対するフィットのよさが、人々が情報獲得に敏感であるというモデルの仮定自体を直接的に証明するわけではない。よって、メタ分析の節において見られた「穏健な」推定結果は、ODS モデルの反対派の攻撃の材料となる可能性がある。しかし、この点については、以下で論ずるように、関連する様々な要因に留意する必要がある。

パラメータ推定における穏健性、すなわち、被験者の確率情報に対する感度の鈍さは、確率的教示の実験的効果を曖昧なものにする大きな要因と考えることができる。これに関連する要因は、いくつか考えられる。まず第1に、主題的材料は、暗黙のうちに人々の一般的知識を喚起し、明示的に与えられる確率情報をも上書きするため、明らかに、人々のパフォーマンスに強い影響を与える(Oaksford, 2002; Oaksford, Chater, & Grainger, 1999)。第2に、解釈が決定的な影響力を持ち得る。実験者がいかなる確率的情報を与えようと、被験者は常にそれを無視する可能性がある。被験者が与えられた確率情報を無視すれば、稀少性を満たすデフォルト値が採用されると考えられるが、そのような被験者が混在すれば、全体の平均値はデフォルト値の方へ引き寄せられ、結果として穏健な値に落ち着くことになる。したがって、いかなる情報も、被験者に理解しやすい形で与えることが重要となる。時には、Gigerenzer & Hoffrage (1995) が主張するように、頻度(非確率)形式で与えるべきかもしれない。また、何らかの現実的経験(例えば、Pollard & Evans, 1983)を伴う形で与えるのが有効かもしれない。

第3に、穏健な信念更新は、Bayes 的規範の枠組みの中でも正当化される。いま例えば、前件の確率  $P(p)$  が(未知の)真の前件の確率  $\theta$  の周りに標準偏差  $\sigma$  で分布すると仮定する。ここ

で、 $\theta$ 自体が、平均 $\mu$ 、標準偏差 $\tau$ の確率変数であると仮定すると、前件確率の事後分布は次のようになる<sup>3)</sup>。

$$\mu + \left(1 - \frac{\sigma^2}{\sigma^2 + \tau^2}\right) [P(p) - \mu].$$

これは、 $P(p)$ が $\mu$ に向かって $\frac{\sigma^2}{\sigma^2 + \tau^2}$ だけ縮小されることを意味する。本来的には、 $\sigma^2$ と $\tau^2$ はデータから推定されるべきであるが、大まかな傾向を見るために、ここでは仮に $\sigma^2$ と $\tau^2$ が等しいとし、Hattori (2002) の推定に従って $\mu$ を.23とする。このとき、仮に「前件の確率は.80」という情報が与えられたとすると、Bayes 計算に従えば、前件の確率の事後信念は.52になる。つまり、デフォルト値 $\mu$ を基点とするような穏健な信念更新が実現されることがわかる。

同様に、われわれは条件文の前件と後件の等確率性に関する強い信念を持っていると考えられる (Hattori, 2002, 2003; Hattori & Oaksford, 2005)。よってわれわれは、この点に関しても同じタイプの穏健性を有していることになる。すなわち、与えられた後件の確率が高いとき、われわれは高い確率を低めるか、あるいは与えられた前件の確率が低ければそれを高める傾向を持つだろう。このことは、課題中の確率的情報を操作することによって論理的正解を高めるのがそれほど容易ではない、という事実を説明する。ODS モデルによれば、論理的に正しい反応 ( $p$  と  $\neg q$  の選択) の確率を高めるためには、 $P(p)$ と $P(q)$ の値を大きく引き離す必要がある (Hattori, 2002)。しかし、これらの値の差は、Bayes 推論によれば 0 の方向に縮小されるべきものとなる。したがって、實際上、2つの値を引き離すにはしばしば困難が伴うことが予想される。

以上をまとめると、次のようになる。ODS モデルは、適切な STF パラメータを使用すれば

データに非常によくフィットする。適切なパラメータとは、ターゲットのデータと類似した選択傾向を持つ被験者群のデータから推定されたものである。これまで、ODS モデルを支持しないように見えたいくつかの経験的証拠は、信念更新および解釈の点においてモデルを精緻化した上で再吟味したところ、いずれも、モデルに矛盾なく説明されることが明らかとなり、ODS モデルの妥当性が再確認された。

## 引用文献

- Beattie, J., & Baron, J. (1988). Confirmation and matching biases in hypothesis testing. *The Quarterly Journal of Experimental Psychology*, 40A, 269 - 297.
- Carlin, B. P., & Louis, T. A. (1996). *Bayes and empirical Bayes methods for data analysis*. London: Chapman & Hall.
- Chrostowski, J. J., & Griggs, R. A. (1985). The effects of problem content, instructions, and verbalization procedure on Wason's selection task. *Current Psychological Research and Reviews*, 4, 99 - 107.
- 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., Newstead, S. E., & Byrne, R. M. J. (1993). *Human reasoning: The psychology of deduction*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Gigerenzer, G., & Hoffrage, U. (1995). How to improve Bayesian reasoning without instruction: Frequency formats. *Psychological Review*, 102, 684 - 704.
- Griggs, R. A. (1984). Memory cueing and instructional effects on Wason's selection task. *Current Psychological Research and Reviews*, 3, 3 - 10.
- Griggs, R. A., & Cox, J. R. (1982). The elusive thematic-materials effect in Wason's selection task. *British Journal of Psychology*, 73, 407 - 420.
- Hattori, M. (1999). The effects of probabilistic information in Wason's selection task: An analysis of strategy based on the ODS model. In

- Proceedings of the Second International Conference on Cognitive Science and the 16th Annual Meeting of the Japanese Cognitive Science Society Joint Conference (ICCS/JCSS99)*(pp. 623 · 626). Waseda University, Tokyo.
- Hattori, M.(2002). A quantitative model of optimal data selection in Wason's selection task. *The Quarterly Journal of Experimental Psychology*, 55A, 1241 · 1272.
- Hattori, M.(2003). Adaptive heuristics of covariation detection: A model of causal induction. In *Proceedings of the Fourth International Conference on Cognitive Science and the Seventh Australasian Society for Cognitive Science Joint Conference (ICCS/ASCS 2003)*(Vol. 1, pp. 163 · 168). The University of New SouthWales, Sydney, Australia.
- Hattori, M., & Oaksford, M.(2005). Adaptive heuristics for covariation detection: Model comparison and rational analysis. (Manuscript under review)
- Hoch, S. J., & Tschirgi, J. E.(1983). Cue redundancy and extra logical inferences in a deductive reasoning task. *Memory & Cognition*, 11, 200 · 209.
- Hoch, S. J., & Tschirgi, J. E.(1985). Logical knowledge and cue redundancy in deductive reasoning. *Memory & Cognition*, 13, 453 · 462.
- 堀内正彦.(2000). Wason の選択課題の発展 試論的実験。『基礎心理学研究』, 18, 156-164.
- Kirby, K. N.(1994). Probabilities and utilities of fictional outcomes in Wason's four-card selection task. *Cognition*, 51, 1 · 28.
- Manktelow, K. I., & Evans, J. St. B. T.(1979). Facilitation of reasoning by realism: Effect or non · effect? *British Journal of Psychology*, 70, 477 · 488.
- Marr, D.(1982). *Vision*. New York: W. H. Freeman.
- 松田隆夫.(2003). 知覚判断における「基準」の多様性とヒューマン・ファラシーの諸相。『立命館人間科学研究』, 6, 67 · 76. 立命館大学人間科学研究所。
- Nisbett, R. E.(2003). *The geography of thought: How Asians and Westerners think differently ... and why*. London: Nicholas Brealey.
- Oaksford, M.(2002). Contrast classes and matching bias as explanations of the effects of negation on conditional reasoning. *Thinking and Reasoning*, 8, 135 · 151.
- Oaksford, M., & Chater, N.(1991). Against logicist cognitive science. *Mind & Language*, 6, 1 · 38.
- Oaksford, M., & Chater, N.(1994). A rational analysis of the selection task as optimal data selection. *Psychological Review*, 101, 608 · 631.
- Oaksford, M., & Chater, N.(2003). Optimal data selection: Revision, review, and reevaluation. *Psychonomic Bulletin and Review*, 10, 289 · 318.
- Oaksford, M., Chater, N., & Grainger, B.(1999). Probabilistic effects in data selection. *Thinking and Reasoning*, 5, 193 · 243.
- Oaksford, M., & Stenning, K.(1992). Reasoning with conditionals containing negated constituents. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 18, 835 · 854.
- Pollard, P., & Evans, J. St. B. T.(1983). The effect of experimentally contrived experience on reasoning performance. *Psychological Research*, 45, 287 · 301.
- Read, T. R. C., & Cressie, N. A. C.(1988). *Goodness - of - fit statistics for discrete multivariate data*. New York: Springer-Verlag.
- Valentine, E. R.(1985). The effect of instructions on performance in the Wason selection task. *Current Psychological Research and Reviews*, 4, 214 · 223.
- Wason, P. C.(1966). Reasoning. In B. M. Fos (Ed.), *New horizons in psychology* (pp. 135-151). Harmondsworth, Middlesex, UK: Penguin Books.
- Wason, P. C.(1968). Reasoning about a rule. *The Quarterly Journal of Experimental Psychology*, 20, 273 · 281.
- Wason, P. C., & Johnson-Laird, P. N.(1972). *Psychology of reasoning: Structure and content*. London: Batsford.
- Yachanin, S. A.(1986). Facilitation in Wason's selection task: Content and instructions. *Current Psychological Research and Reviews*, 5, 20 · 29.
- 山 祐嗣.(1999). Wason 選択課題における選択の主観的理由。『教育心理学研究』, 47, 11 · 18.
- Yama, H.(2001). Matching versus optimal data selection in the wason selection task. *Thinking and Reasoning*, 7, 295 · 311.

## 注

1) 厳密には、仮説の確信度  $P(r_1)$  を変化させることも可能であり、実際、それは選択傾向に影響するが、モデルの振る舞いの変化の点においてその影響力は非常に小さい (Hattori, 2002, p. 1247 参照)。そこで、本論文では  $P(r_1) = .5$  として固定した。

- 2) 堀内正彦氏には、堀内 (2000) に報告されていた実験のローデータを提供して頂いた。ここに記して感謝する。0
- 3) この式の導出に関するより詳しい記述については、例えば、Carlin & Louis (1996) を参照されたい。

(2005.3.14 受理)

Table 2

Maximum Likelihood Estimates (MLEs) of the Model and their Goodness-of-Fit to Data in the Japanese Literature

No	Study	Experiment, Condition	N	Literature			Experiment		
				$P(p)$	$P(q)$	$G^2(2)$	$P(p)$	$P(q)$	$G^2(2)$
48	Yama (1999)	a/a	36	.28	.32	3.26**	.19	.24	1.76***
49	Horiuchi (2000)	$p \rightarrow q$	36	.18	.21	0.51***	.04	.05	1.32***
50	Hattori (2002)	expt 1, orig, LH	45	.32	.34	8.14*	.26	.28	2.03***
51	Hattori (2002)	expt 1, orig, HH	46	.34	.36	12.45	.27	.31	4.13**
52	Hattori (2002)	expt 1, orig, C	41	.36	.36	15.20	.28	.29	5.97**
53	Hattori (2002)	expt 2, orig	82	.25	.26	14.14	.13	.14	3.80**
54	Yama (2001)	expt 1, a/a, *L	34	.22	.25	1.96***	.26	.29	2.38***
55	Yama (2001)	expt 1, a/n, *L	35	.24	.35	4.96**	.28	.37	4.89**
56	Yama (2001)	expt 2, a/a, *L	76	.30	.35	2.18***	.33	.37	0.48***
57	Yama (2001)	expt 2, a/n, *L	73	.41	.45	8.22*	.42	.46	5.05**
58	Yama (2001)	expt 3, a/a, *L	50	.24	.27	0.31***	.28	.31	0.37***
59	Yama (2001)	expt 3, a/n, *L	51	.40	.42	0.28***	.43	.44	1.21***
60	Yama (2001)	expt 4, a/a, *L	48	.24	.30	23.55	.29	.34	17.58
61	Yama (2001)	expt 4, a/n, *L	47	.32	.48	15.31	.34	.47	10.06
62	Hattori (2002)	expt 1, prob, LH, non-PC	24	.43	.44	1.53***	.39	.41	0.48***
63	Hattori (2002)	expt 1, prob, HH, non-PC	31	.42	.47	4.96**	.39	.47	1.03***
64	Hattori (2002)	expt 2, prob, non-CC	58	.39	.40	4.20**	.34	.36	0.44***
65	Yama (1999)	a/n	36	.32	.50	1.94***	.22	.59	1.41***
66	Horiuchi (2000)	a/n	36	.28	.53	1.85***	.05	.89	1.83***
67	Yama (2001)	expt 1, a/a, *H	36	.45	.55	0.77***	.46	.54	0.30***
68	Yama (2001)	expt 1, a/n, *H	36	.03	.97	9.10*	.03	.97	4.04**
69	Yama (2001)	expt 2, a/a, *H	75	.36	.37	11.86	.39	.40	5.61**
70	Yama (2001)	expt 2, a/n, *H	73	.41	.49	8.92*	.43	.49	3.98**
71	Yama (2001)	expt 3, a/a, *H	50	.46	.53	9.78	.47	.52	5.72**
72	Yama (2001)	expt 3, a/n, *H	51	.51	.56	1.53***	.51	.55	0.19***
73	Yama (2001)	expt 4, a/a, *H	46	.20	.35	9.12*	.25	.37	5.25**
74	Yama (2001)	expt 4, a/n, *H	47	.27	.65	12.62	.33	.59	8.04*
75	Hattori (2002)	expt 1, prob, LH, PC	21	.44	.44	3.06***	.40	.40	1.11***
76	Hattori (2002)	expt 2, prob, CC	24	.43	.51	12.71	.40	.54	7.47*
77	Hattori (2002)	expt 1, prob, HH, PC	15	.59	.60	5.50**	.60	.62	1.87***

Note. In "Experiment, Condition" column, a/a means affirmative antecedent and affirmative consequent condition, a/n means affirmative antecedent and negative consequent condition, n/a means negative antecedent and affirmative consequent condition, and n/n means negative antecedent and negative consequent condition; expt means Experiment, orig or prob means original or probabilistic version of the selection task; HH (LH) means high (low)  $P(p)$  and high  $P(q)$  condition, \*H (\*L) means high (low)  $P(q)$  condition, C means control condition; PC means probability-conscious group, Non-PC means the complement of the PC group, CC means consistent conditional interpretation group, Non-CC means the complement of the CC group.  
In  $G^2(2)$  column, \*\*\*  $P \geq .20$ ; \*\*  $P \geq .05$ ; \*  $P \geq .01$ .

Table 3

MLEs of the Model and their Goodness-of-Fit to Data in the Western Literature

No	Study	Experiment, Condition	N	Literature			Experiment		
				$P(p)$	$P(q)$	$G^2(2)$	$P(p)$	$P(q)$	$G^2(2)$
1	Wason (1968)	expt 1, expt group	18	.18	.21	5.85**	.05	.06	3.77**
2	Wason (1968)	expt 1, control group	16	.04	.05	1.28***	.03	.05	4.07**
3	Wason (1968)	expt 2	26	.06	.10	1.15***	.02	.05	3.00***
4	Evans & Lynch (1973)	a/a	24	.28	.34	1.50***	.20	.28	4.04**
5	Manktelow & Evans (1979)	expt 1, abstract, a/a	24	.24	.29	2.72***	.14	.19	2.17***
6	Manktelow & Evans (1979)	expt 2, abstract, a/a	24	.27	.29	0.08***	.19	.21	1.12***
7	Manktelow & Evans (1979)	expt 3, abstract, a/a	16	.35	.38	2.81***	.29	.33	0.84***
8	Manktelow & Evans (1979)	expt 4, abstract, a/a	16	.25	.25	9.14*	.16	.17	14.02
9	Manktelow & Evans (1979)	expt 5, abstract, a/a	16	.12	.13	0.62***	.05	.05	2.57***
10	Griggs & Cox (1982)	expt 1, abstract, a/a	32	.04	.06	2.10***	.03	.05	10.55
11	Griggs (1984)	expt 1, t/f, non-mem	25	.23	.26	4.87**	.09	.12	8.07*
12	Chrostowski & Griggs (1985)	expt 2, t/f, non-mem	60	.04	.05	0.78***	.04	.05	14.59
13	Hoch & Tschirgi (1985)	bachelor, standard	25	.32	.36	2.00***	.26	.31	0.29***
14	Valentine (1985)	t/f, a/a	24	.28	.31	0.54***	.19	.22	2.98***
15	Yachanin (1986)	expt 2, t/f, abstract	20	.24	.27	2.65***	.14	.17	0.83***
16	Beattie & Baron (1988)	expt 1, 4-card, a/a	18	.04	.05	1.36***	.03	.05	5.25**
17	Oaksford & Stenning (1992)	expt 2, abstract, a/a	24	.31	.32	0.64***	.23	.25	1.96***
18	Oaksford & Stenning (1992)	expt 3, colored shape, a/a	24	.36	.36	0.32***	.32	.32	2.40***
19	Oaksford & Stenning (1992)	expt 3, vowel-even, a/a	24	.23	.25	3.08***	.10	.13	1.36***
20	Oaksford & Stenning (1992)	expt 3, control, a/a	24	.25	.29	1.27***	.13	.18	3.11***
21	Evans & Lynch (1973)	a/n	24	.22	.61	2.62***	.02	.96	6.71*
22	Manktelow & Evans (1979)	expt1, a/n	24	.23	.64	4.58**	.02	.97	4.83**
23	Manktelow & Evans (1979)	expt2, abst, a/n	24	.31	.55	7.60*	.18	.71	7.61*
24	Valentine (1985)	t/f, a/n	24	.30	.35	1.33***	.22	.29	0.22***
25	Beattie & Baron (1988)	expt 1, 4-card, a/n	18	.02	.95	1.01***	.02	.95	5.85**
26	Oaksford & Stenning (1992)	expt 2, abstract, a/n	24	.39	.54	1.15***	.31	.59	1.77***
27	Oaksford & Stenning (1992)	expt 3, colored shape, a/n	24	.36	.36	1.60***	.32	.32	5.36**
28	Oaksford & Stenning (1992)	expt 3, vowel-even, a/n	24	.28	.32	0.90***	.20	.25	2.00***
29	Oaksford & Stenning (1992)	expt 3, control, a/n	24	.29	.42	1.09***	.22	.42	3.20***
30	Evans & Lynch (1973)	n/n	24	.54	.59	0.68***	.55	.62	0.28***
31	Manktelow & Evans (1979)	expt 1, abst, n/n	24	.39	.46	6.08*	.36	.46	3.60**
32	Manktelow & Evans (1979)	expt 2, abst, n/n	24	.44	.49	0.35***	.42	.50	0.45***
33	Valentine (1985)	t/f, n/n	24	.33	.37	1.02***	.26	.32	0.01***
34	Oaksford & Stenning (1992)	expt 2, abst, n/n	24	.34	.38	1.43***	.27	.34	0.07***
35	Oaksford & Stenning (1992)	expt 3, colored shape, n/n	24	.36	.36	0.93***	.28	.28	3.70**
36	Oaksford & Stenning (1992)	expt 3, vowel-even, n/n	24	.31	.34	1.91***	.24	.27	2.88***
37	Oaksford & Stenning (1992)	expt 3, control, n/n	24	.41	.47	1.18***	.37	.47	4.08**
38	Evans & Lynch (1973)	n/a	24	.40	.41	0.80***	.40	.40	2.37***
39	Manktelow & Evans (1979)	expt 1, abst, n/a	24	.36	.38	5.30**	.32	.34	1.96***
40	Manktelow & Evans (1979)	expt 2, abst, n/a	24	.28	.29	2.92***	.21	.21	0.38***
41	Valentine (1985)	t/f, n/a	24	.32	.32	1.50***	.24	.25	3.08***
42	Oaksford & Stenning (1992)	expt 2, abst, n/a	24	.40	.40	1.76***	.36	.36	1.42***
43	Oaksford & Stenning (1992)	expt 3, colored shape, n/a	24	.36	.36	2.96***	.32	.32	6.72*
44	Oaksford & Stenning (1992)	expt 3, vowel-even, n/a	24	.21	.22	1.84***	.08	.09	2.75***
45	Oaksford & Stenning (1992)	expt 3, control, n/a	24	.39	.40	0.03***	.35	.37	1.14***
46	Kirby (1994)	L*	45	.38	.41	4.12**	.34	.38	0.90***
47	Kirby (1994)	H*	44	.53	.56	3.73**	.54	.58	0.23***

Note. In "Experiment, Condition" column, t/f means the task in which participants should determine whether the rule is true or false.

All the other notations in "Experiment, Condition" column are same as those in Table 2.

In  $G^2(2)$  column, \*\*\*  $P \geq .20$ ; \*\*  $P \geq .05$ ; \*  $P \geq .01$ .