

項目反応理論を用いた 4 枚カード問題の分析¹

中村 國則² 日本学術振興会 (東京工業大学)

An item response theory analysis of the four-card selection task

Kuninori Nakamura (*Japan Society for Promotion of Science, Tokyo Institute of Technology*)

The four-card selection task (Wason, 1966) is one of the best known tasks used in the literatures of human reasoning. This article analyzes this selection task by using item response theory (Lord & Novick, 1968). Japanese undergraduates ($N=323$) responded six types of the Wason's selection tasks including indicative task (Wason, 1966), beer task (Griggs & Cox, 1982), and cassava task (Cosmides, 1989). An exploratory categorical factor analysis revealed a one factor structure of the six tasks. The results of an analysis using a two-parameter logistic model indicated that the indicative tasks were similar to the beer task and the cassava task in terms of the discrimination parameter, and that the relative difficulty between the tasks would varied according to the value of the ability parameter estimated by the two-parameter logistic model.

Key words: Four-card selection task, item response theory, thematic-content effect, individual difference.

The Japanese Journal of Psychology
2009, Vol. 80, No. 5, pp. 436-441

4 枚カード問題 (Wason, 1966) は思考研究の中で最も多く用いられてきた課題の一つである。この課題のオリジナル版 (Wason, 1966) では、実験参加者には“E”, “K”, “4”, “7” の 4 枚のカードが与えられ、“片面が母音ならば、もう片面は偶数である”という規則が成立するかを判断するために最低限めくべきカードは何か、を答えることが求められる。この場合、検討すべき規則は“p ならば q”という論理的な条件文として考えられ、4 枚のカードはそれぞれ、“p である”, “p ではない”, “q である”, “q ではない”に対応する。したがって規則の真偽にかかわるカードは“p である”と“q ではない”を表すカードだけであり、ここでは正答は“E”と“7”になる。このように形式的には非常に単純な問題であるにもかかわらず、オリジナル版の正答率の低さ (Wason, 1966) や、問題に含まれている文脈の操作で正答率が変化する主題化効果 (thematic content effect; Johnson-Laird, Legrenzi, & Legrenzi, 1972; Wason & Shapiro, 1971) と

いった様々な知見が 4 枚カード問題には知られており、多くの研究者の注目を集めてきた。

4 枚カード問題に関する知見の多くは、課題間の正答率、あるいは各カードの選択比率の比較に基づいた分析から導かれてきた。たとえば、上に述べた主題化効果に対する説明として、実生活の中で帰納的に学習した知識構造に一致することが回答を促進するとする実用的推論スキーマ説 (pragmatic schema hypothesis; Cheng & Holyoak, 1985)、規則が社会交換を含んでいるときに裏切り者検知モジュールが発動するために生じるとする社会交換説 (social contract theory; Cosmides, 1989) などが知られている。これらの仮説を検討する場合、日常的に馴染みのある文脈を表す問題とそうでない問題、あるいは社会交換を含む問題とそうでない問題を作成して正答率を比較し、差の有無を検討する。このような比較による分析は他の側面の検討 (Kirby, 1994) などでもしばしば用いられ、4 枚カード問題の一般的なデータ解析手法といえる。

しかしながら比較という分析手段は、以下の二つの問題点をもつ。一つは、課題間の類似性を扱いきれない点である。課題形式に注目する限り、あくまで 4 枚カード問題は“p ならば q”という条件文を理解し、条件文の真偽にかかわるカードの選択を実験参加者に要求しており、その意味で、全ての 4 枚カード問題は何らかの形である種の論理的思考能力を反映する課題といえる。では、4 枚カード問題はどの程度、このよう

Correspondence concerning this article should be sent to: Kuninori Nakamura, Yamagishi Laboratory, Graduate School of Decision Science and Technology, Tokyo Institute of Technology, Ookayama, Meguro-ku, Tokyo 152-8552, Japan (e-mail: knaka@ky.hum.titech.ac.jp)

¹ 本研究は日本学術振興会特別研究員奨励費の補助を受けた。

² 統計解析について貴重なコメントを頂いた大久保 智哉 (大学入試センター)、原稿校正をしてくれた渡邊 匠子 (東京工業大学大学院) の両氏に感謝します。

な能力を反映するのだろうか。あるいは、能力の反映という点で様々な4枚カード問題は互いにどの程度似ているのだろうか。このような点を明らかにするためには、複数の課題に共通する性質を量的に表現する必要があるだろう。しかしながらこれまで、4枚カード問題の研究ではこのような検討は殆ど行われていない。

もう一つは、課題の性質を個人差と関連づけることが難しいことである。4枚カード問題について知られる様々な知見が全ての実験参加者に対して等しくあてはまるとは考えにくく、何らかの形で個人差があると考えるのが自然である。たとえば、ある実験参加者群にとっては容易な問題が、別な実験参加者群には困難な問題になることは十分あり得る。しかし、問題間の成績の比較ではこのような個人差の問題は扱いにくく、先行研究でも殆ど検討されていない。

数少ない例外として、Stanovichらの一連の研究(Stanovich & West, 1998a, b, 2000)がある。Stanovich & West (1998a)は、様々な4枚カード問題間の成績の関連を検討し、概して直説法課題と義務論的課題の間に有意な連関を見出している。またStanovich & West (1998b)では学業到達度テストと様々な4枚カード問題との間の関係を検討し、幾つかの4枚カード問題で正答した実験参加者と誤答した実験参加者との間で学業到達度テストの成績に有意な差がみられたことを報告している。これらの知見は、様々な4枚カード問題は実は互いに共通しており、しかもその共通した側面が学業到達度テストで測定可能な一般的思考能力と関係することを示唆している。しかしStanovichらの研究は特定の問題間の相関や正答率の差の検討に留まり、能力の反映といった側面から各問題の性質を検討してはいない。

このような背景の下、本研究の目的は項目反応理論(Item response theory; Lord & Novick, 1968)によって以上の問題を検討することである。項目反応理論は実験参加者のある項目に対する反応をモデル化し、実験参加者の性質を潜在特性によって記述する。この潜在特性は正規分布すると仮定され、項目の性質は潜在特性を正答確率と対応づける項目反応曲線によって表現される。本研究で用いるのは2母数ロジスティックモデルと呼ばれ、正解を1、誤答を0とするとき、能力値 θ を有する実験参加者が項目 v_i に正答する確率を表す以下の式で定義される;

$$P(v_i=1|\theta) = \frac{1}{1 + \exp(-1.7a_i(\theta - b_i))} \quad (1)$$

ここで a_i 、 b_i はそれぞれ項目の識別力と困難度と呼ばれるパラメータである。識別力とは $\theta=b_i$ において項目がどの程度実験参加者の能力を敏感に反映するかを表し、困難度はその項目が相対的にどの程度難しいか、具体的には正答確率0.5を与える時の実験参加者

の能力値を表す。

2母数ロジスティックモデルを用いることにより、これまで扱われていなかった4枚カード問題の様々な側面を検討できる。まず困難度のみならず識別力をみることによって、単純な正答率の比較だけでは明らかにならなかった問題間の類似性を議論することが可能になる。そしてもう一つ興味深いのが、問題の相対的な難度が個人差に応じて変動することを議論できる点である。ここでFigure 1中の困難度が同じで識別力が異なる二つの項目の項目反応曲線をみてみよう(困難度と識別力がそれぞれ、0.5、1.2の曲線と0.5、0.4の曲線、あるいは1.5、1.2の曲線と1.5、0.4)。これらの項目では、能力値が相対的に低い場合には識別力の低い項目の方が、逆に高い場合には識別力の高い項目の方が、高い正答確率を示すことがわかる。このような関係は実験参加者の能力によって“難しさ”が異なり、実験参加者の個人差に応じて問題の性質が異なることを表している。このような関係は困難度と識別力がともに異なる曲線同士でも成立し(たとえば識別力、困難度がそれぞれ1.2、0.5の曲線と0.4、1.5の曲線)、項目の相対的な難しさが個人差に応じて変動することがわかる。

以上の特徴は特に主題化効果を考える上で重要である。主題化効果とは先に述べたように、問題に含まれる文脈の操作によって正答率が変化することを指す。しかし、実験参加者の個人差によって、文脈の操作の影響が変動することは十分ありえる。即ち、ある能力値の実験参加者では主題化効果を確認できても、別の能力値の実験参加者群では確認できない可能性は十分考えられる。もし実際に問題間でこのような関係が成

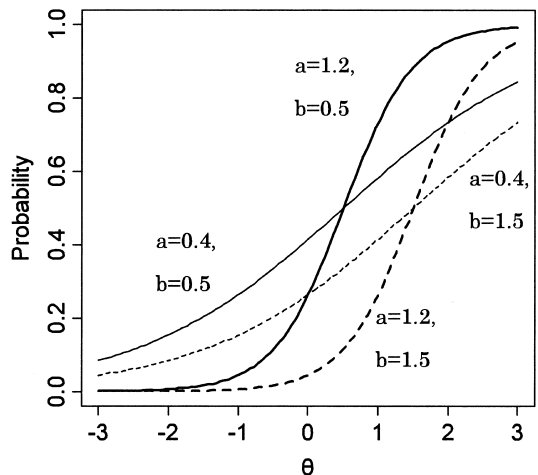


Figure 1. 様々な項目反応曲線：ここでは四つの項目反応曲線は、困難度・識別力をそれぞれ0.5と1.5、0.4と1.2に操作したものを示してある。図中の実線の曲線は困難度が0.5、破線の曲線が困難度が1.5のものに対応する。

Table 1
本研究で用いた4枚カード問題

	規 則	p	not-p	q	not-q
直説法 (偶数)	“カードの片面が母音である ならば、もう片面は偶数が書 かれている”	E	K	4	7
直説法 (A)	“もしカードの片面がAなら ば、もう片面は2が書かれて いる”	A	T	2	9
ビール	“ビールを飲んでいる人は、20 歳以上でなければならない”	ビール	コーラ	22歳	15歳
郵便	“もし郵便物に封がしてある ならば、表に20ドルの切手が 貼ってある”	封がしてある	封がしてない	20ドル切手	10ドル切手
キャッサバ	“ある男がキャッサバを食べ るならば、顔に刺青がなけれ ばならない”	キャッサバを 食べている	モロナッツを 食べている	顔に 刺青がある	顔に 刺青がない

立しているならば、課題間の成績の比較だけでは検討は不可能である。それに対し、2母数ロジスティックモデルは個人差に応じて問題の相対的な難度を表現するため、実験参加者によって主題化効果の出方が異なるかを検討できるのである。

本研究は2母数ロジスティックモデルを4枚カード問題に適用し、従来しばしば用いられてきた4枚カード問題の性質を検討する。無論、従来用いられたことのある全ての問題を網羅することは不可能だが、これまで頻繁に用いられてきた課題を取り上げ、それらの性質を検討する。具体的に本研究では、Table 1に示される6種類の代表的な4枚カード問題を題材として、項目反応理論を用いた解析例を示す。その中で、様々な4枚カード問題が互いに類似した課題であること、主題化効果の生起に実験参加者の個人差が影響することを示す。

Table 1中の2種類の直説法課題はともに Wason (1966) で用いられた、代表的な直説法課題である。二つの課題で異なるのは、片方の課題で言及されているのは“偶数”、“母音”といった集合であるのに対し、もう一方の課題では“A”、“2”といった特定の事例である点である。カードの属している集合の大きさが成績に影響することが先行研究から知られており (Kirby, 1994)、本研究ではこれらの二つの課題を直説法課題として採用した。

郵便問題とビール問題はともに、主題化効果を生じる課題として知られている。前者は主題化効果を示した初期の研究 (Wason & Shapiro, 1971) で用いられた問題であり、後者も主題化効果を生じる課題であると同時に、課題に含まれる文脈がある種の社会的な義務を表現したものであり、義務論課題として解釈される

課題である (Cheng & Holyoak, 1985; Manktletlow & Over, 1991)。

本研究では加えて、Cosmides (1989) で用いられた2種類のキャッサバ課題を用いる。Cosmides (1989) は、先に述べた社会交換説を検討するため、Cosmides は実験参加者にとって馴染みのない文脈で、課題内容が社会的交換を含んでいる場合とそうでない場合を比較し、前者の方が後者より高い正答率を示すことを明らかにしている。本研究では社会交換を含むキャッサバ課題と社会交換を含まないキャッサバ課題を用い、これらの問題の性質を他の課題との比較を通じて検討する。

方 法

実験参加者 都内私立大学経済学部学生 323 名が実験に参加した。実験参加者には4枚カード問題及び類似した論理的推論課題を解いた経験はなかった。

手続き 質問紙を用いて集団状況で、本研究とは別の目的で行われた他の課題とともに、Table 1に示してある規則、カードからなる6種類の4枚カード問題を刺激として提示した。実験参加者には口頭で、表裏双方に何か書かれている4枚のカードに関して、それらのカードに書かれている内容について、Table 1に記載されている規則が当てはまるかどうかを確かめる課題であると説明し、それを確かめるために最低限めくる必要のあるカードを選択させた。ただし2種類のキャッサバ課題については、規則、カード内容ともに同じものを用いるため、まず“自分がポリネシアのある部族の調査をしにきた調査員であると想像してください”という文脈を与えた上で、以下のような操作を与えて社会交換を含むものとそうでないものを区別し

Table 2
各問題の記述統計量, 因子分析・2母数ロジスティックモデルの結果

課 題	正答率 (%)	四分相関係数						項目母数	
		1.	2.	3.	4.	5.	6.	因子 負荷	識別力 (a) 困難度 (b)
1. 直説法課題(偶数)	9.6	—						0.84	1.56 1.57
2. 直説法課題(A)	6.5	0.72	—					0.91	2.22 1.70
3. ビール問題	52.3	0.25	0.23	—				0.87	1.79 -0.07
4. 郵便問題	14.6	0.46	0.53	0.34	—			0.83	1.50 1.27
5. キャッサバ問題(社会交換)	54.2	0.24	0.22	0.57	0.31	—		0.86	1.67 -0.13
6. キャッサバ問題(社会交換なし)	11.1	0.19	0.23	0.20	0.22	0.23	—	0.56	0.67 2.20

た。まず社会交換を含むものについては、調査対象の部族の決まりとして、キャッサバを食べるためには、顔に刺青を入れなければならないことになっており、この規則を破っているものがあるかどうかを確認することが課題であると教示した。社会交換を含まないものについては、調査対象の部族がキャッサバという植物かモロナッツという植物のどちらかを常食とし、かつ食べるものはどちらかであって、両方食べるという事はありえないということを背景情報として与え、その部族の常食が何であるかを確かめることが課題であると教示した。提示順序は事前に用意した2種類の無作為に配列したパターンのどちらかで提示した。実験参加者は他の課題を含めて30分以内で回答を終了した。

結 果

各課題の正答率と課題間の四分相関係数を Table 2 に示す。これをみると、2種類の直説法課題と比較してビール問題、社会交換ありのキャッサバ課題の正答率は高くなっていることがわかる。しかしながら郵便課題の正答率は15%、社会交換なしのキャッサバ課題の正答率は11%にとどまり、特に社会交換なしのキャッサバ課題の成績は Cosmides (1989) で得られた正答率 (20%) よりも低い。しかしこれらの課題でもわずかではあるが正答率は直説法課題よりも高く、本研究でも主題化効果は確認されたと考えられる。また、6種類の4枚カード問題の間の四分相関係数は全て有意であり (0.20—0.74)、Stanovich & West (1998a) と同様の知見を得ており、先行研究を再現したと考えられる。また、各課題の正誤とその課題を除く5つの課題の合計正答数との間の点双系列相関係数 (0.49—0.54) も有意であった。

続いて統計解析ソフト Mplus version5 (Muthen & Muthen, 2007) を用いて探索的カテゴリカル因子分析を行い、1因子性の検討を行った結果、固有値は第1因子から第6因子までそれぞれ、4.18, 0.76, 0.60, 0.25, 0.20, 0.02 であり、第1因子の固有値が非常に

高く、かつ第2因子以降の固有値が低かった。また、各問題に対する第1因子からの因子負荷は、どの問題に対しても強いことが分かる (Table 2)。以上の点から、本研究で用いた6種類の4枚カード問題に関する限り、問題間の1因子性は確認できたと考える。因子の内容は、4枚カード問題が“pならばqである”といった論理的規則を理解し、提示された4枚のカードの中から規則に関連するものを発見する課題であること、社会的文脈が付与された問題よりも2種類の抽象的課題の方が高い因子負荷を受けることから、ある種の論理的推論能力を表す因子であると解釈できる。

問題間の1因子性が確認されたため、Mplus version 5を用いて2母数ロジスティックモデルによって分析した³ (Table 2)。各問題の項目反応曲線を Figure 2 に示す。社会交換なしのキャッサバ課題を除く全ての代表的な4枚カード問題の識別力は高く⁴、困難度は概して直説法課題の方がビール問題、社会交換ありのキャッサバ課題よりも高く、正答率の傾向とほぼ一致する。

直説法課題とビール問題、社会交換ありのキャッサバ課題の関係に注目すると、 θ の値によらず直説法課題の方が低い正答確率を示している。このような結果は、社会交換を含む文脈の付与が正答率を向上させる主題化効果が、実験参加者の論理的能力の個人差に関わらず成立すると解釈できる⁵。一方、六つの問題の

³ 赤池情報量基準 (Akaike, 1974) を用いて2母数ロジスティックモデルを1母数ロジスティックモデルと比較し、識別力パラメータの設定の妥当性を検討すると、前者 (1 442.91) は後者よりも (1 463.10) よい当てはまりを示し、識別力パラメータの設定は妥当であると結論できる。

⁴ 社会交換なしの課題の識別力が低いのは、社会交換なしの文脈で“…なければならない”という文言を用いたことによって問題内容が理解しにくかった可能性が考えられる。しかしながら、実験参加者の能力値によって正答確率が異なる状況が4枚カード問題の中でありえることを示した点は重要と考えられる。

⁵ ここでいう1因子性はあくまで論理的能力の反映に関わるものであり、全ての問題で正答率が同じになることを意味しない。

中で一つ独特の振る舞いを示すのが社会交換なしのキャッサバ課題である。この課題は正答率が直説法課題と同程度であるものの、困難度は6問中最大であり、直説法課題より高い。また、識別力は6問中最低であり、能力の変化が比較的反映されにくい課題であることがわかる。

さらに興味深いのは、社会交換なしのキャッサバ課題の正答確率は、 θ がおよそ1.4より低い場合には直説法課題を上回るものの、それよりも高い場合には直説法課題を下回ることである。この傾向は、直説法課題と社会交換なしのキャッサバ課題の相対的な難易度が実験参加者の能力によって異なることを示している。すなわち、直説法課題との比較で考える場合、社会交換なしのキャッサバ課題では実験参加者の能力値が相対的に低い場合には主題化効果が生じるのに対し、能力値が相対的に高い場合は直説法課題の方が社会交換なしのキャッサバ課題よりも正答率が高くなるという、“逆”主題化効果を生じ得ることを示すものである。

考 察

本研究は直説法・義務論課題の双方を含む6種類の4枚カード問題をカテゴリカル因子分析、及び2母数ロジスティックモデルによって分析した。その結果、因子分析は1因子解を支持し、本研究で用いた4枚カード問題はいずれも論理的推論能力を強く反映する点で類似していることが明らかになった。また、2母数ロジスティックモデルによる分析の結果、識別力の点で多くの4枚カード問題は類似した性能を有すること、実験参加者の能力に応じて課題の相対的な難易度が変動する場合のあることが明らかになった。このような知見は、4枚カード問題が概して、条件文を理解するといった論理的思考能力を強く反映する課題であることを示しており、直説法課題と義務論課題という区別が一般的に受け入れられてきた背景を踏まえると興

味深いものであり、かつ以下の2点を示唆すると考えられる。

一つは、主題化効果の解釈に対する実験参加者の個人差の考慮の必要性である。主題化効果とは文脈の操作によって正答率が変化する現象を指し、課題間の正答率、あるいは各カードの選択比率の比較に基づいて議論されることが一般的であった。たとえば、主題が社会交換を含んだものであるときに正答率の向上がみられる(Cosmides, 1989)ことを検討する場合、社会交換を含むと仮定できる文脈と含まないと仮定できる文脈を構成した上で両者の正答率を比較し、差の有無を吟味して課題の性質を議論することが行われている。特に近年、社会交換と人間の思考の関係をめぐる研究の中で、4枚カード問題で検討される規則、及びカードの内容は多様化し、文化間比較(Sugiyama, Tooby, & Cosmides, 2002)なども行われるようになってきている。そのような背景の下、本研究は実験参加者の能力値によって成績の大小関係が変化する場合がありますを明らかにし、主題化効果を議論する場合、正答率のみならず実験参加者の個人差、特に論理的推論能力の個人差を考慮する必要性を示した点で興味深い示唆を与えるといえる。

そしてもう一つの示唆は、個人差の問題を踏まえた上でも、先行研究で確認された主題化効果は頑健に成立することを示した点である。ビール問題(Griggs & Cox, 1982)、キャッサバ課題(Cosmides, 1989)はオリジナルの直説法課題(Wason, 1966)よりも高い正答率を示すことが知られ、主題化効果を示す代表的な例としてしばしば言及されてきた。そして実験参加者の個人差によらず、これらの課題で主題化効果が成立すること自体は、実は経験的に知られていたことである。たとえば佐倉(1997)は、偏差値の高い学生を実験参加者としても主題化効果が生じることを報告し、論理的な推論が得意であると考えられる実験参加者にとっても、直説法課題よりも義務論課題の方が容易で

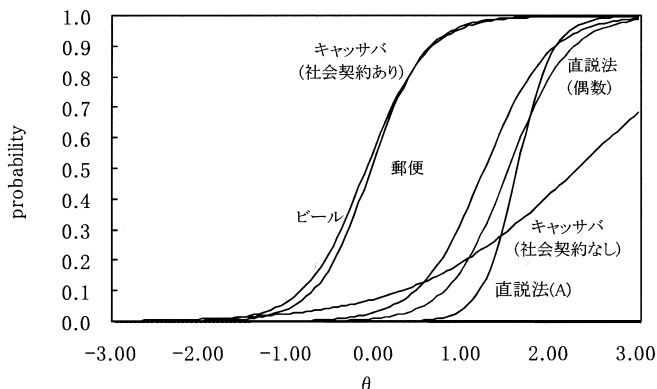


Figure 2. 各問題の項目反応曲線

あることを指摘している。このような知見は、実験参加者の論理能力によらない主題化効果の存在を示唆するものの、あくまで“経験知”に留まるものであり、体系的な検討は加えられていなかった。その点で本研究は2母数ロジスティックモデルの適用によって、従来確認されてきた主題化効果が実験参加者の能力値によらず頑健に生じることを示し、経験知をより厳密な形で確認したと評価できるだろう。

本研究には、以下の課題が残されている。一つは、課題のバリエーションの拡大である。本研究では扱う課題を6種類に限定していたが、本研究で扱ったもの以外にも4枚カード問題のバリエーションは多岐にわたり、かつ研究目的に応じて新たなものが作成され続けている (Gigerenzer & Hug, 1992)。そのような多様な目的・文脈を含んだ課題解析が重要な検討課題の一つとなろう。

もう一つは、正答率以外のデータ分析である。本研究では正誤の2値のみを検討対象としたが、誤答パターンの中にも興味深い統計的構造がある可能性は十分考えられる。項目反応理論の枠組みでも誤答パターンのモデルは可能であり、Bock (1972) のようなモデルや Thissen (1976) といった解析例も古くから存在している。正答誤答の2値データのみならず、誤答パターンを考慮した解析にも今後の興味深い展開を期待できるだろう。

引用文献

- Akaike, H. (1974). Information theory and an extension of the maximum likelihood principle. In B. N. Petrov & F. Csaki (Eds.), *Second international symposium on information theory*. Budapest: Akademiai Kiado. pp. 267-281.
- Bock, R. D. (1972). Estimating item parameters and latent ability when responses are scored in two or more nominal categories. *Psychometrika*, **37**, 29-51.
- Cheng, P. W., & Holyoak, K. J. (1985). Pragmatic reasoning schemas. *Cognitive Psychology*, **17**, 391-416.
- Cosmides, L. (1989). The logic of social exchange: Has natural selection shaped how humans reason? Studies with the Wason selection task. *Cognition*, **31**, 187-276.
- Gigerenzer, G., & Hug, K. (1992). Domain-specific

- reasoning: Social contracts, cheating, and perspective change. *Cognition*, **43**, 127-171.
- Griggs, R. A., & Cox, J. R. (1982). The elusive thematic-materials effect in Wason's selection task. *British Journal of Psychology*, **73**, 407-420.
- Johnson-Laird, P. N., Legrenzi, P., & Legrenzi, S. M. (1972). Reasoning and a sense of reality. *British Journal of Psychology*, **63**, 395-400.
- Kirby, K. N. (1994). Probabilities and utilities of fictional outcomes in Wason's four-card selection task. *Cognition*, **39**, 85-105.
- Lord, F. N., & Novick, M. R. (1968). *Statistical theories of mental test scores*. Reading, MA: Addison-Wesley.
- Manktelow, K. I., & Over, D. E. (1991). Social roles and utilities in reasoning with deontic conditionals. *Cognition*, **39**, 85-105.
- Muthen, B. O., & Muthen, L. K. (2007). *Mplus user's guide*. Los Angeles, CA: Authors.
- 佐倉 統 (1997). 進化論の挑戦 角川書店 (Sakura, O.)
- Stanovich, K. E., & West, R. F. (1998a). Cognitive ability and variation in selection task performance. *Thinking and Reasoning*, **4**, 193-230.
- Stanovich, K. E., & West, R. F. (1998b). Individual differences in rational thought. *Journal of Experimental Psychology: General*, **127**, 161-88.
- Stanovich, K. E., & West, R. F. (2000). Individual difference in reasoning: Implications for the rationality debate? *Behavioral and Brain Sciences*, **23**, 645-726.
- Sugiyama, L., Tooby, J., & Cosmides, L. (2002). Cross-cultural evidence of cognitive adaptations for social exchange among the Shiwiar of Ecuadorian Amazonia. *Proceedings of the National Academy of Sciences*, **99**, 11537-11542.
- Thissen, D. (1976). Information in wrong responses to the Raven Progressive Matrices. *Journal of Educational Measurement*, **13**, 201-214.
- Wason, P. C. (1966). Reasoning. In Foss, B. M. (Ed.), *New horizons in psychology*. Harmondsworth: Penguin.
- Wason, P. C., & Shapiro, D. (1971). Natural and contrived experience in a reasoning problem. *Quarterly Journal of Experimental Psychology*, **23**, 63-71.