

|               |   |
|---------------|---|
| Title         | キャッカン テスト ニ オケル タイモン  |
| Author(s)     | ヒグチ, シンゴ  |
| Citation      | 大阪大学人間科学部紀要. 7 p83-p.101  |
| Issue Date    | 1981-03   |
| oaire:version | VoR   |
| URL           | <a href="https://doi.org/10.18910/11690">https://doi.org/10.18910/11690</a> |
| DOI           |   |
| rights        |   |
| Note          |   |

*Osaka University Knowledge Archive : OUKA*

<https://ir.library.osaka-u.ac.jp/>

Osaka University

# 客観テストにおける対問

樋 口 伸 吾

## 客観テストにおける対問

### I 問題

“思考力のみられない”という言葉は、ほとんど客観テストの枕言葉として定着してしまっている。たしかに、解答としていくつかの選択肢が用意されており、そのうちからひとつを正解として選ぶ行動のうちには、純粋な思考力の作用も含まれ得るであろうが、あてずっぽうな、もしくはでたらめな反応もある程度は存在すると考えなければならないであろう。しかし、それだけの理由で客観テストの存在理由を否定するのは無理である。というのは、国立大学入試共通第一次テストに代表される多量のデータ処理の場合をみればわかるように、客観テストはその可否を観念的に論ずる段階を既に過ぎ、今日の社会に存在する現実の事態となっているからである。問題は、むしろ客観テストの測り得るものの正確な判断と処理ならびにその問題点の検討・改善にあると思われる。本論文は、まず、そのひとつの問題であるでたらめ解答によるテスト結果の汚染防止の方略を検討する。あわせて、“思考能力のみられない”という通説に対し、客観テストにおいてもそれが可能であると主張するテストの内容を吟味する。それは、能力開発研究所で実施した進学適性能力テストである。このテストの内容の検討は、そのでたらめ解答防止の方略すなわち対問の問題と深くからみあっており、両者をわけては論じ得ないので、議論は常にその間を往復することになるであろう。

### II 得点分布と偶然得点分布との重なり

いわゆる統計的分布の検討においては、まず理論的にみて純粋に極端な場合を想定し、その場合の分布と現実の分布との対比によって、解決に向うのが科学的思考の常道のひとつである。ここでは、でたらめ解答による得点の純粋な極限的分布として、偶然得点分布を想定することにする。

偶然得点分布とは、選択肢の数・その組合せ等を含む客観テストの解答形式に対して、受験者総員がまったくでたらめに解答を選んだ場合に生じると期待される得点分布をいう。一番簡単な場合、すなわち各問について選択肢の数が一定数  $n$  であるテストにおいては、偶然得点分布は  $\frac{1}{n}, 1 - \frac{1}{n}$  を各項とする 2 項分布にしたがう。その他のより複雑な場合には各問

の形式に応じ偶然得点の確率を算出し、それにその問題の配点を乗じ、各問を独立に組み合せて、偶然得点分布を求めることができる。ただし、この場合は、配点の組み合わせによってある得点は表われないことがある。

次に示す図1～4は、昭和42年度に高校3年生（人員数はNで示される）に対して行われた能力開発研究所の学力検査のうちの一部について、偶然得点分布を点線グラフ、実際の得点分布を棒グラフで示したものである<sup>1)</sup>。

斜線を付した部分は、偶然得点分布の上部を5%で切断する線を求め、その内部に入ってくる実際の得点分布の出現状況を示したものである。この5%という数値は、特に理論的に設定したものではなく、危険率の目安として使用されている1%、5%のうち5%のほうが大まかな判断の基準となると考え、これを準用したものである。

斜線の部分を全体に対する比率で示すと、国語Ⅰの5.54%、同Ⅱの2.71%のように、この部分が極微な場合と、倫理・社会の17.39%、政治・経済の22.75%のように、でたらめ反応の存在が必ずしも無視できない場合とが存在する。その他、ここにあげない図では、同じ斜線の部分が、日本史16.32%、世界史29.71%、地理8.44%、数Ⅰ39.35%、数Ⅱ31.48%、数学甲17.70%、物理22.95%、化学16.67%、生物25.71%、地学35.63%、英語29.74%、ドイ

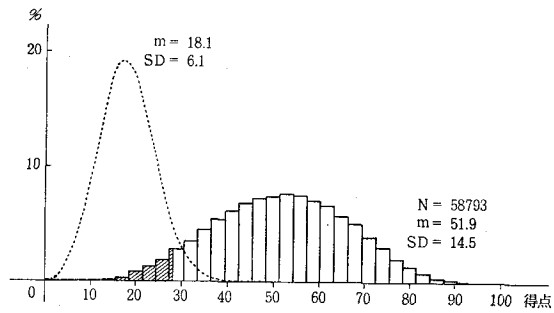


図1 国語Ⅰ

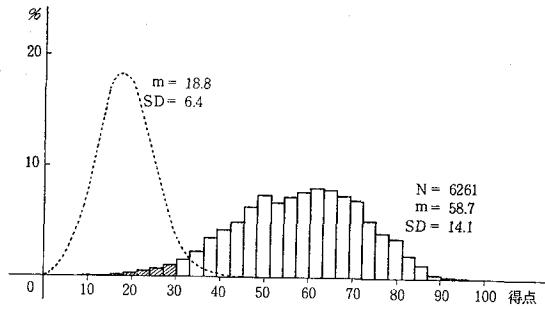


図2 国語Ⅱ

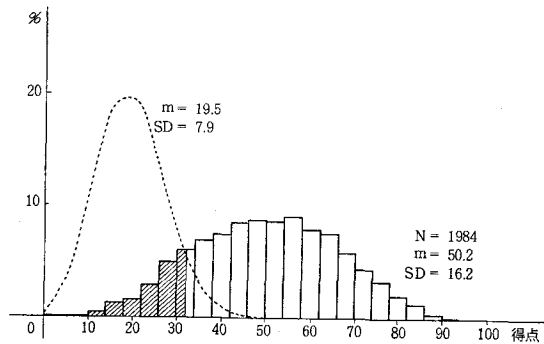


図3 倫理・社会

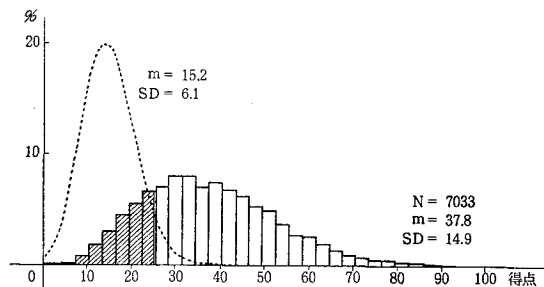


図4 政治・経済

ツ語35.37%, フランス語33.33% という比率がみられる。

なかには、極端な場合がある。第5図はわたくしの作成した職業適応能力テスト第Ⅱ部のB(20点満点)の実際の得点分布および偶然得点分布を示したものであって、実測値の83%が偶然得点のうちに入っている。すなわち、受験者約3,000名がいわゆる“チンパンジーでもできる”といわれるでたらめ反応を行った場合に発生する分布とほとんど変わらない得点分布がみられるのである。しかし、この現象は、必ずしもこれらの得点の無意味性を証明するものではない<sup>2)</sup>が、低い得点者群における各問の弁別性・斉合性などは一般に非常に悪いことが実証されている。このことだけでも、偶然得点の問題は重要である。

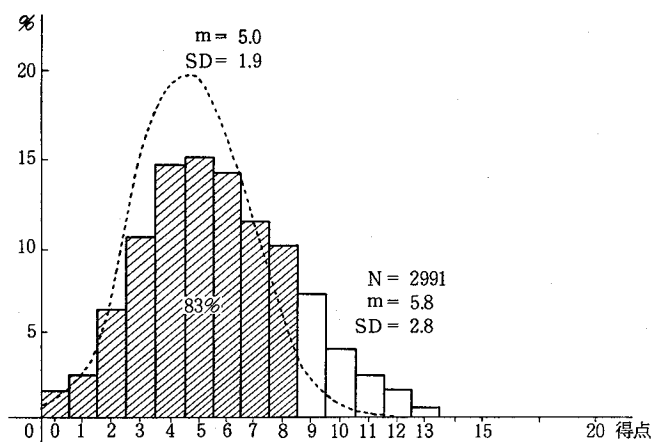


図5 職業適応能力テストⅡのB

以上の統計的実態は、受験者の学力の程度、問題の難易等の複雑なからみあいから発生するものであるが、客観テスト得点における偶然得点の混入が全面的に否定されないことを物語るのである。

### Ⅲ 対問による偶然得点の防止

偶然得点は、このように客観テストにつきまとう厄介な問題である。ただし、機械的に実際の得点分布を偶然得点分布からひきはなすのは簡単である。それは、誰でも正しい解を与えることのできるやさしい問題を多数出題すればすむのであって、その場合、テスト得点の実際の分布は図ではいちぢるしく右の方向に移動し、偶然得点分布とははっきりした差を示し得るからである。しかし、こういう姑息な手段をさげ、正面から堂々と、かつ、組織的に偶然得点を排除しようと試みた例がある。それは、能力開発研究所で行われた進学適性能力テストである。

このテストにおいて偶然得点を排除しようとする努力は、“思考能力のみられない”という非難に対応する必要からも発生している。次に、同僚の一人、鮫島の説明<sup>3)</sup>を引用し、このテストの特徴と偶然得点の防止策とについての説明をきくことにする。

まず、この検査の性格について、鮫島は次の4点をあげる。

(1) 大学に進学して高い水準の教育を受けるのに必要な知的能力のうち、特に一般的であり、また、中核的な役割をになっているとみなされる言語的推理能力および非言語的推理能力が測定にかけられており、このテストの第Ⅰ部は言語的推理能力、第Ⅱ部は非言語的推理能力の測定を、それぞれ担当している。

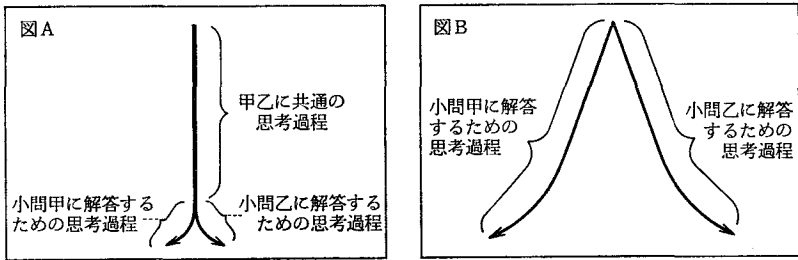
(2) 受験生が個々の問題の内容を理解するために必要な知識としては、義務教育の期間、すなわち、中学3年までに学んだもので、十分である。

(3) 全体として、じっくりと課題にとり組んで内容をよく理解した上で解かなければならない問題が、中心となっている。

(4) 各問題は、すべて甲、乙二つの小問から成り立っており、さらに、これらの小問には、それぞれAからEまでの5個の選択肢が与えられている。解答者は、各小問について、これら5個の中から、正しい答を1個だけ選ばなければならない。採点時には、各問題について、甲、乙ともに正答を選んでいる場合に限り、その問題に正解したものとみなされることになる。

このうち(4)が偶然得点排除のための技術的工夫である。その内容は、昭和41年版進学適性能力テスト解説（以下解説Ⅰという）にかみくだいて次のように解説される<sup>4)</sup>。

“小問は甲と乙の二つに分かれています。問題にとってもっとも本質的な思考は共通しており、その思考過程を経た者であれば、甲、乙ともにらくに正解を導くことができるように、問題形式に、できる限りのくふうがほどこされています。小問甲を解くための思考過程とはまったく別個に小問乙を解くための思考過程が必要とされるのではなく、その本質的な部分は、甲、乙に共通しているわけです。このことを図で表わしたものが図Aであり、この関係は図Bのようにはなりません。



各問題の採点方法は、甲、乙の両方に対して正しい答が選ばれている場合のみを正答と認め、いずれか一方だけが正しく答えられている場合は正答として認めません。したがって、

甲、乙ともにまったくでたらめに解答した結果が偶然に正答と一致する確率は、 $\frac{1}{5} \times \frac{1}{5} = \frac{1}{25} = \frac{4}{100} = 0.04$  というひじょうに小さい値をとります。問題数は第Ⅰ部、第Ⅱ部ともにせいぜい30題ていどですから、その全部に対してでたらめの解答をしたとすれば、その結果、たまたま正答として扱われることになる問題の数はそれぞれのテストについてたかだか1題ていどであり、でたらめ解答はあまり効果的な手段であるとは言えません。そんなことをするよりも、これらの問題のうちから解けそうな問題を選んで、それらに全部の時間をかけてじっくりとり組んだほうが成績が上がる公算が大であると言えます。また、テスト時間が残り少なくなったために、たとえば残りの5題にでたらめな解答をしたとしても、そのうちの1題ですら正答と一致することはあまり期待できませんから、それよりも、残り時間に全力をあげて1題ずつ順にとり組んだほうが、成績が上がる公算が大であるわけです”。

以上の説明は、いわゆる能研進学適応能力テストが偶然得点の出現に対して、ほとんど完璧に近い理論的武装をほどこしたという宣言とみられよう。それが客観テストの信用性に与えた功績は、非常に大きい。しかし、この理論的かつ実践的業績は同時に新しい問題をなげかけたのである。

ひとつは、このテストにおいては思考過程の一義性が前提となっているが、それと、この対問による得点結果の正確さとは必ずしも結びつかないこと。他のひとつは、対問を加えることによって問題の性格に変容が与えられることである。

#### Ⅳ 問題解決の多様性

鮫島は、進学適性能力テストにおける思考過程について次のように述べる<sup>5)</sup>。

“テスト理論にしたがって、ここで次の仮設を置く。

仮設1：アイテム  $j$  を解くために必要とされる心的機能  $X_j$  とアイテム  $k$  を解くために必要とされる心的機能  $X_k$  に関して、高3受験生および高2受験生は、それぞれ二変量正規分布に従って分布する。同様のことが、アイテム  $u$  とアイテム  $v$  に関しても成立する”。

この仮定に基づき、各問題とアイテムを解く心的機能  $X_i$  の間の相関係数の推定が行われ、次の仮設がおかれる<sup>6)</sup>。

仮設2：言語的推理能力  $C_I$  と、アイテム  $j$  を解くために必要とされる心的機能  $X_j$  に関して、高3受験生および高2受験生は、それぞれ二変量正規分布に従って分布する。同様のことが、非言語的推理能力  $C_{II}$  と、アイテム  $u$  を解くために必要とされる心的機能  $X_u$  に関しても成立する”。

ここから各問の相関マトリックスについて主因子解による因子分析が行われ、その結果に基づいて次の結論が得られる<sup>7)</sup>。

“以上の結果は、各層ともに、第Ⅰ部、第Ⅱ部のアイテム群が測定している心的機能に、

それぞれ単一共通因子が存在するとみなしてよいことを示すものである。これは、テスト理論のモデルに対する適合の重要な必要条件であるとともに、進学適性能力テストを第Ⅰ部・第Ⅱ部の２部門に分けて作成・実施した意図がよく生かされていることを示す一つの根拠ともなり、このことから、テスト実施上のさまざまな不利な条件にもかかわらず、所期の目的は一応達成されていると考えることができるであろう。高３，高２のデータからそれぞれ抽出された第Ⅰ因子は、その構造がたがいによく似ているので、両者を同じものとみなし、このようにして抽出された各部門ごとの単一共通因子を、改めて、それぞれ、言語的推理能力  $C_I$ 、非言語的推理能力  $C_{II}$  と名づけ、以下、これらを各部門の測定対象たる能力として扱うことにする”。

この結論は、各問題はそれぞれ唯一最善の解決法があり、それを発見し得る共通能力  $C_I$ 、 $C_{II}$  が正規分布するという立場を明らかにしたものである。この立場に基づき、テストの得点は、普通の粗点・標準点等による表示は行われず、 $C_I$ 、 $C_{II}$  に関する真の得点がある信頼限界の中で幅をもって示されることになっている。

しかし、この理論に基づいて、 $C_I$ 、 $C_{II}$  を正確に測定すべく作成された知能検査といえども、ひとつひとつの問題が彼女によって仮定された一義的な精神作用のみによって解決されるとは限らないのである。このことは適性問題の解決過程に関する実証的研究によって確認される。次にⅠ例を示す。まず問題である。

a, b, c, d という仲よし４人組みがいる。ある日、４人はそれぞれ仲間のだれかの家へ、遊びに出かけた。めいめいが、同じとき、別々の友だちの家に着いたが、どの家でも友だちは留守で、他の友だちとそこで出会うこともなかった。

・ a 君が、d 君の家をたずねたのではないということは、出かける前に、行く先をおかあさんに告げたことからわかった。

・ c 君の家をたずねた人をかりに x 君とすると、b 君がたずねたのは x 君をたずねた人の家である。

甲 b 君は、だれの家をたずねたのか。

- A a の家
- B c の家
- C d の家
- D a の家か d の家か決めることはできない。
- E この条件からだけでは、決まらない。



乙 d 君は、だれの家をたずねたのか。

- A a の家
- B b の家
- C c の家
- D a の家か c の家か決めることはできない。
- E この条件からだけでは、決まらない。

この問題の作成のヒントとなったのは、ブール代数における回路に関する考えかたであった。この問題に対する推理の方法について、昭和43年1月発行の「進学適性能力テストの解説と例題」(以下解説Ⅰという。)から示そう<sup>8)</sup>。その最初の部分は後で説明することにして、次に同書に書かれた“正解”をうつすことにする。

“ところでこの問題を正しく推理するのは、かなりわずらわしいことです。仮定を設けて推論し、矛盾が生ずればその仮定をすて他の仮定を検討しなければなりません。しかも何度かこれをくりかえしておこなう必要があります。それだけに、強<sup>じん</sup>靱な推理能力が必要とされます。そこで段階に分けてこまかく考えてみましょう。また手がかりとなる二つの条件を①、②とあらわすことにします。

(1) ①により a は d の家以外つまり b の家か c の家をたずねている。ここで a が b の家をたずねたと仮定しよう。②から c の家をたずねた x とは b か c であることになる。もし前者であるとすれば②の x に b を置きかえると、b は c の家と a の家とを同時にたずねたことになり矛盾がおこる。

(2) そこで c の家をたずねた x とは d であると仮定しよう。最初に a が b の家をたずねたという仮定はそのままだから、d の家をたずねたのは b か c である。前者は②に当てはめると自分の家をたずねることになり矛盾がおこる。一方、c が d の家をたずねたとすると②に当てはめて b は c の家をたずねることになり、c の家には先に d がたずねることになっているので、これも矛盾する。したがって、さかのぼって(1)で a が b をたずねたという最初の仮定から導かれる推論はすべて矛盾し否定される。

(3) a が c をたずねたことは確かである。すなわち、x とは a である。さて a の家をたずねるのは c か d である。前者とすれば②に当てはめて c の家を a と b とがたずねることになり矛盾がおこる。

(4) 故に a の家をたずねたのは d であることが確定する。b がたずねたのは  $x(=a)$  をたずねた人、すなわち d であることも確かである。こうして3人の行先が決定し、残る c は b の家をたずねたことも確定する。

以上ですべての可能性が検討しつくされ、矛盾なしに成立する1組の答が見いだされたこ

とになります。選択肢DやEのように不確定となることはないわけです”。

この問題の正答者が1割に達しなかったことは、その解決が“強靱な推理能力”を必要とする証明のようにも思われる。そして、ここで述べられている推理過程は、ほとんど唯一絶対の解のようにみえ、対問は、正解者がその過程を正確にふんできたことを保証するように思われる。

ところが、先に省略した文（それは、わたくし書き加えたのであるが）のように“この問題は、与えられた条件を整理して、直観的な解法によって解くこともでき”るのである。まず与えられた条件について、 $\rightarrow$ で方向、 $\nrightarrow$ でその否定を示すと、

$$1. b \rightarrow y \rightarrow x \rightarrow c$$

$$2. a \nrightarrow d$$

$$\text{したがって } x=a, y=d$$

$$\text{もしくは } x=d, y=a$$

$$a \nrightarrow d \text{ であるから}$$

$$x=a, y=d$$

この解法は、わたくし自身による。テスト終了後、この問題を解くために“強靱な推理能力”のみが一貫して流れる必要があるかという疑問が生じたので、わたくしは、20人の被験者に、この問題を解くプロセスについて言語報告をさせながら、その筋道を明らかにしようとした。その際、みずからも被験者となり解答しようとしたのであるが、わたくしは知らず知らずのうちにテストの内容をできるだけ明確に集約しようと努力していた。そして、それが完了した瞬間に解が自然に出現したのである。その過程を上記のように数式で書いてみると、少々時間を要したようにみえるかもしれないが、実は瞬時に解が得られたのである。あまりにもやさしいので、これが一体問題となり得るかと疑ったことを今なお鮮明に記憶する。しかるに、わたくし以外の20名の被験者は解決に早くて10分、おそくて20分を要し、なかには30分を経てしかも解を得られなかったものも多かったのである。

彼らの解法を辿ってみると、おおむね先に引用した解説Ⅱのバリエーションを示すといいたい得よう。しかしこのモデル通りの表現を報告したものは皆無であり、いずれも、どこかに飛躍がみられた。しかし、総員が“強靱な推理能力”を発揮しようと努力しており、わたくしのような直観的な解を与えたものは1人もいなかった。わたくしは、自分がまちがっているのかとしばらく疑い、また、解説者はわたくしの推理不足な点を指摘した。それは、わたくしの証明では、cがbに行くことが明らかでないこと、ならびにx,yがbもしくはcであり得ることをはじめから否定していることの2点であった。

しかし、問題の範囲内では、どう考えてもこの批評は詭弁であり、2点とも自明で証明を

要しないと思う。そして、今に至るまで、わたくしは、この解が正当な解法のひとつであると確信している。では、わたくしの推理はどのようにして発展したのであろうか。

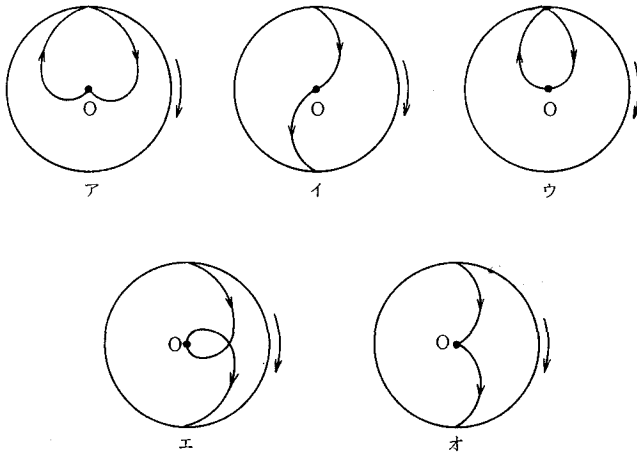
うすれた記憶を辿ってみると、それには、いくつかの前提条件があったようである。わたくしは、まず、研究部長として約20種類におよぶ学力テスト、進学適性能力テストⅠ、Ⅱ、職業適応性能力テスト、追跡調査等の総責任者となっており、さらに基礎研究部長をかねて多忙な日々を送っていた。このテスト問題を解いたのは、その多忙の間に辛じて得た機会においてである。しかし、この問題がブール代数からヒントを得たという事実がわたくしの興味を絶えず刺激していた。そして、大分以前に、問題の原案をみせられた時、解いてみようとは思わなかったがゴタゴタした表現に困惑し、できるだけ簡潔に問題の主旨を徹底する方策を論じてほしいと指示を与えた。その結果、問題は、前に示された通り、非常に明快な形をとり得たのである。そして、被験者としてその解決にあたった時、わたくしは、まず問題の内容をさらにはっきりとイメージ化するよう努力し、それが完了した時に問題が解けていた。つまり、より簡潔に、より本質をというわたくしの身体にしみこんだ思考法のひとつが、この解を生み出したのである。したがって、この問題を瞬時に解いたと前に述べたのは間違いであって、実は長い間無意識な試行・探索が行われていたにちがいない。それはともかく、この問題を解くには“強靱な推理能力”を必要とすると断定するのはあやまりであり、それと異なる方向からのアプローチも可能であることになろう。この二つの考え方を、かつてのわたくしの用語で説明すれば、アナログ型思考とデジタル型思考とに区分できよう<sup>9)</sup>。今は、この区分をイメージによる思考と言語による思考と書きなおしたほうがよりわかりやすいかとも考えている。とまれ、この二種の根本的に異なる思考方法がこの問題でみられるのであって、対問は、それをひとつの方向に規制する力をもっていないのである。

## V 対問の影響

第二の問題点をひとくちでいえば、対問も問いの一部になってしまう現象であろう。

その事例を端的に示す問題を次にあげる。

半径 10 cm の円板がある。この円板は中心  $O$  を軸として、1 回転 24 秒でゆっくり回転している。円板の、ある直径の一端から他端まで、虫が等速度で直進しているのを真上から観察した。



甲 虫が直径の一端から他端まで動くのに24秒かかったとすれば、虫の動きはどのように見えるか。

- |   |   |   |   |
|---|---|---|---|
| A | ア | D | エ |
| B | イ | E | オ |
| C | ウ |   |   |

乙 虫が直径の一端から他端まで動くのに12秒かかったとすれば、虫の動きはどのように見えるか。

- |   |   |   |   |
|---|---|---|---|
| A | ア | D | エ |
| B | イ | E | オ |
| C | ウ |   |   |

この問題を作るヒントはソビエトの高校参考書から得られた。それは軌跡の概念を教えるために書かれたので、飛行機がモスクワから北極を通過してワシントンまで地球の自転速度と同じ速度で飛ぶことを想定し、それを北極直上の遠い場所から見た場合に飛行機の軌跡がどのように表われるかという設問であった。

この発想がさまざまな議論のすえ、先に示した問題に結実したのであるが、原題の痕跡は、1回転24秒（地球の24時間1回転から変換した）などという表現のうちに残っている。わたくしは、それにしてもこの問題はもとの主旨をたくみに捉えた良い問題であると考え、出題することに決定したが、やがてこの問題が軌跡を求める作業とはまったく異なった見地からも解決し得ることに気付かされたのである。その間の事情を説明するために、とりあえず解説Ⅱに示されている解を示そう<sup>10)</sup>。

“最後に図形を見ながら考えるものを例題として示しました。図形にもとづく推理にはもっとやさしいものもありますが、上に示したものは図形もその上の点も移動するので、かなりむずかしいといえるでしょう。しかし、五つの選択肢のうちのどれか一つに正しい経路が

与えられているので考える手がかりは十分に与えられています。しかも、甲と乙とでは点の移動する速度が異なるので、基本的な考え方は同じでもそれぞれ別に考える必要があります。

まず甲から考えましょう。円板の周期と虫が直径を通るのに要する時間とがひとしいことに注目すると、円板が一回転したとき虫がちょうど一端から他端に達していることがわかります。このことから選択肢イ、エ、オのうちどれかが正しい経路であることとなります。ところで半分まで円板がまわったとき虫がどこにいてどちらを向いているかを調べると、直径の中心つまり円板の中心にいて出発点の方を向いているはずですが、これに適合するのはエだけです。さらに途中の経過もこまかく考えれば確かにエであることがわかります。

乙の条件にかなうのは、円板が半回転したとき他端に達しているわけですから実際の経路は再び出発点にもどっているようにみえるはずですが、したがって、アカウのどちらかが正しいはずですが、そこで円板が4分の1回転したとき虫がどちらを向いているかを考えると、円板の中心にいて上から見たとき出発した方向に対して直角の向きにあることがわかります。これに適合するのはウだけです。

この問題は、以上のように分析的にある時点の状態を考えるやり方とは別にすこしずつ頭の中で連続的に動きをとらえていって、これと合う正しい経路をかいてみてもよいわけです。この問題もテストの配列の終りに近いところにあるので、時間がなくて手をつけられなかったと思われる人が3割ほどあるようです。全体の正答者は1割に達しません”。

この解説を読みながらわたくしは暗然とせざるを得ない。はじめの主旨からいえば被験者は円板を少しずつ移動させながら虫の位置を確認してゆき、それをつなぎあわせて正解を得るべきであろう。その手順をふむことを強制してはじめて対問を設定し、同一思考過程を保証しようとするテストのねらいが生きてくる。しかるに解説Ⅱは、“それぞれ別に”、“五つの選択肢を手掛りとして解く”のが正しいやりかたであると強調している。しかも、解説Ⅱにおいては、同書7ページに前にあげた解説Ⅰの説明をそのまま踏襲し、思考過程の本質的な部分は甲乙に共通していると繰り返して説かれている。その主張にもかかわらず、ここでは選択肢を利用する解法の説明が行われているのである。すなわち、前文とこの解法との完全な矛盾がはばかりなく押し出されている。その上に、この解説Ⅱは、研究部長としてのわたくしの責任のもとで監修されている。さらに悪いことには、その矛盾がはっきり意識されたのは発行後10年以上を経た今日のことであった。穴があったら入りたい。

今考えてみると、もっと洗練された選択肢を選ぶことも可能であった。しかし、そう努力しても、結果はおそらくブール代数をモデルとした前問の分析と同じような結論へと指向したであろう。例えば、問題を甲だけにしぼり、虫が6秒後どこにいるか、18秒後どこにいるかという対問をつくり、選択肢としてできるだけまぎらわしい5個の図形を示すならば、今述べた問題は回避され得たかもしれない。ところがこの場合にも、軌跡的な考え方と、選択

肢に示される点の位置を利用し解決する方法と、出発点からの距離による解法との、少なくとも3種の思考過程の元型ならびにその混合型が表われる可能性がある。そして、選択肢をもつ解答を呈示すること自体が問題の本質を変えてしまうという事態は依然として続いたであろう。もっとも、思考過程の同一性を保証するためには、上記のような手段あるいは有効性を発揮し得たかもしれない。したがって、この点については、出題前に徹底した検討が行われるべきであった。しかし、解説Ⅱは、この問題にまったく触れず、前述の解法を推奨しているのである。

以上の錯誤について弁明にならない弁明をあえて行うとすれば、次の理由があげられるであろう。

第1の理由として、テスト実施後改めてこの問題の解法を検討した際の議論と結論とが解説者に強い影響を与えたことをあげられよう。わたくしは、先の問題とひとしく、この問題についても同様の実験を行ったのであるが、その際得られた最も敏速な思考過程は次の通りである。

“24秒は12秒の倍である。したがって線の長さを比較すれば、ほぼ倍の長さをもつア、エが甲の解に含まれ、乙の解としてはイ、ウ、オの3者が想定される。甲の場合、到着点は出発点の反対側だから、エが正解、乙の場合、出発点と到着点と同じはずであるから、ウが正解となる”。

この解は、与えられた選択肢としての解答を逆用し、むしろそこから発想した、全体としてみれば明快きわまりない解決である。この解決は“すこしずつ頭の中で連続的に動きをとらえていって、これと思う正しい経路をかいてみる”方法とはまったく異なる思考過程によっている。

以上の二つの解法を両極としてさまざまな解法が10種類ほどみられた。この事実が、解説Ⅱの著者の脳裏に強烈に反映したのであろう。彼は、これらの解法が存在することを自明のこととし、解説を書いてしまった。そこで解説書が全体としては前後撞着していることに気付かなかったのである。

第2の理由として、解説Ⅰと解説Ⅱの得点表示方法の微妙な差異がここに反映していると考えられる。得点は、前者では Lord—鯨島理論により、後者では Lord 理論に基づいて表示されているが、前者の仮定はⅢ節に引用されている通り、Lord のそれよりさらにきびしい。したがって理論自体が正解の一義性・正確性を強く要請しているので、5つの選択肢からなる対問の必要性はきわめて高い。これに対し、後者の立場は、Lord の使用したベータ関数のもつ可塑性が示唆するように、条件の規制がよりゆるやかであり、そのためやや弛緩した対問の扱い方が生じたとみられる。両者の立場のちがいが、あるいはさきほどの矛盾の一因を形成したのかもしれない<sup>11)</sup>。

その上、今のわたくしからみると、Lord に心酔した Lordian たちは、ひとつの問題を解決するためにさまざまな解法があり得ることについて十分に注意を払っていなかったように思われる。彼らは Lord とひとしく、過去の経験から既に与えられたもの、形成されたものの影をまったく逆に a priori な実在として天上に昇らせてしまった。そしてプラトンのなイディアの世界に遊び、真の姿を夢みている。この楽園に没入すれば、影は影であって、分析してもほとんど意味がない—こういう考え方が、影、すなわち現実のテストの解法の研究を怠らしめたのであろう。

“山並遠に春は来て こぶしの花は天上に”と詩人は歌う。三好達治は春の美しさという厩大な母集団のうちから、この風景を選びだしたのではない。こぶしの花にめざめるような驚きを感じ、春の本質をみたのであり、春は、逆に彼の詩によって内容をさらに深化させるのである。テストにおいても現実の問題の鋭い分析と新しい問題の作成への努力とが思考能力のさまざまな姿を浮びあがらせる。Lordian たちは、春から発想してこぶしの花を歌おうとする。詩はできあがるかもしれないが、おそらく論理にささえられた論文のような詩しか生まれまい。

第3の理由としてあげられるのは、わたくしの立場である。わたくしは、研究部長として先に述べたように、約20種類の問題の作成・実施・吟味等の仕事に追われていた。そのためにはずかしいことには、この解説書はほとんど責任者にまかせきりであった。しかし、進学適性ないしは思考能力を追求する問題は念頭をはなれることなく、そのためにひとつひとつの問題を作成・検討することよりも、真の思考能力をみることのできるテストの作成は果して可能か、そもそも思考能力の本質は何かという研究の必要を痛感し、そのために研究所に基礎研究部を確立すべく全力を注いでいたのである。当時の書類をみると毎年のように1億内至2億円の金額を基礎研究に使用する建議を行っている。今は反故と化したこの書類と2つの解説書を見比べ、ただ嘆息するほかはない。とはいえ、当時のわたくしが解説Ⅱを注意して読んだとしても、以上の過誤に気付いたか否か、さらに気付いたとしても対問の影響を逆用し、真の効果のあるテスト技法を開発し得たか否かは疑問である。それは、若干経験を積んだ今の目から見ると、これらの適性能力検査は、“詩なき科学”の研究能力検査に終始しており、わたくしはそのことに盲目であったからである。

## VI 問題と対問との調和

対問は、parallel items と訳されよう。ところで、この parallel という形容詞は、古典的テスト理論にあっては、真の得点 $T$ を共有し、独立でひとしい誤差分散をもつテストに与えられる術語である。進学適性検査の対問においても、この主旨にそって、甲・乙のそれぞれ

につき並列性・互換性が証明されるならば、これを *parallel items* と称して差支えあるまい。

ところでこの形容詞からただちに連想される言葉は *parallel prose* である。それは、中国の六朝時代に行われた駢儷体の文章をいう。各語それぞれ対をなした美しい表現であるが単なる言葉の遊戯にも墮ししやすい。その対句を音声の平仄のもつ厳格な規則とともに重要な構成要件としたのが律詩である。すなわち律詩においては、少なくとも第三句と第四句、第五句と第六句とは対句でなければならないのである。律詩は唐初に胚胎し、杜甫に至って完成する。したがって、杜詩の研究は *parallel* という言葉の本質について示唆するところが多い。

吉川は、杜甫の五言律詩「喜達行在所三首」を論じて次のようにいう<sup>12)</sup>。

“…この短促にして装飾的な詩形は、この発生の当初に於いては、筵宴、遊覧、贈答など、人生に於ける装飾的な事件を即興的に詠ずるものではあっても、まとまった思想、複雑な事件を、詠ずるものとしては、用いられていないようである。…ところが今や詩人は、その一生に於ける大事件として、自ずからも意識したに相違ないところの、九死一生の旅行の経験を、敢えてこの詩形によって、歌っている。詩人は、この作品に至って、いかなる題材をも、この短促にして装飾的な詩形によって歌い得るという自信、すなわち消極的には、その短かさと装飾とによって束縛を受けることなく、また積極的には、その短かさと装飾とを有効に作用させて、充実した新しい芸術を作り得るという自信を得たようであって、且つ美事それに成功している。……恐らく杜甫は今や、自ずからが律詩の完成者であることを、自ずからも覚ったことと思う”。

まさにその通りであろう。

ここに吉川がこの新しい律詩の「先駆」と見做した「月夜」、特にその対句を彼の解説により味ってみよう<sup>13)</sup>。

|       |                                  |
|-------|----------------------------------|
| 今夜鄜州月 | こんや ふしゅう つき<br>今夜 鄜州の月           |
| 閨中只獨看 | けいちゆう た ひと み<br>閨中 只だ独り看る        |
| 遥憐小兒女 | はる あわ しょうじよ<br>遥かに憐れむ小兒女の        |
| 未解憶長安 | い ちようあん おも かい<br>未まだ長 安を憶うを解せざるを |
| 香霧雲鬢湿 | かうむ うんかんうるお<br>香霧 雲鬢 湿い          |
| 清輝玉臂寒 | せいき ぎよくひさむ<br>清輝 玉臂寒し            |
| 何時倚虚幌 | な とき きよこう よ<br>何んの時か虚幌に倚り        |
| 雙照淚痕乾 | そうしよう るいこんかわ<br>雙照 淚痕 乾かん”       |

吉川は特に対句について次のように解説する。



“…第二聯〔遥憐小兒女，未解憶長安〕，遥一未，憐一解，小一憶，兒女一長安と，対句の形を保ちつつ，意味は上下一連，小兒女の未まだ長安を憶うを解せざるを遥かに憐れむ，と，いわゆる「流水対」である。英詩でいえば，enjambment であり，上下句の間に，関係代名詞 who を補って解釈すべき場合である。近体の律詩におけるこの「流水対」の詩法，卷二の詩をかえりみれば，…いずれもなおこの詩ほど大胆自由でない。…ただし自由であるだけに，兒女と長安とは，厳密には対しない。また上の句は，遥<sub>平</sub>憐<sub>平</sub>小<sub>仄</sub>兒<sub>平</sub>女<sub>仄</sub>と変格である。しかし二句は，もとより対句としての意識から作られているのであって，…（以下森槐南の説の引用）”。

さらに第三聯についての説明がある。“…またこの聯，上下句とも，光線に対する杜詩の敏感と技術，凡手には難しいものなのを，ここでも具現するが，上の句の〔香霧雲鬢湿〕は，雰囲気的，拡散的なのに対し，下の句の〔清輝玉臂寒〕が，凝視的，集中的なのは，一そうの妙である。そうして上下二句が，意義的にも文法的にも相互に厳密な対句であり，さきの第二聯の「流水対」と対比的なものも，月下の妻の容姿と悲哀を一そう鮮烈にする”。

以上の美事な対句を重要な構成要素として成立した，“この詩，〔今夜鄜州の月〕と，おもいを彼の地の月にはせるをはじまって，屈折によって悲哀を深化すること”を吉川は明代の批評を引用して讃えるのである。彼の説明の通り，この詩における対句は，単なる言葉の遊戯でなく律詩全体のきびしい法則性のうちに錬磨され，その制約が逆に利用され，緻密きわまる構成をもちつつ，人生の悲哀を結晶させているのである。

ここにいささか乱暴なアナロジーを行ってみると，進学適性能力テストの例題に示された対問のうち，前者は，対偶の典型的なもの，後者は，吉川の述べた流水対にあたるであろう。流水対は，彼の記したように上下両句を通じて，一意貫串する形式であって，杜甫の作品のひとつの特徴をなす。例えば「客至」の「花径不曾緑客掃 蓬門今始為君開」のもつ美しさをみればよい。また対偶といっても彼の詩の場合は機械的な対応を示すのではなく，「閣夜」の「五更鼓角声悲壯 三峽墨河影動搖」のように，平板な対句とはまったく異なった絶妙の深みをもっている。したがって仮にアナロジーが成立するとしても，当然ながらわれわれの対問は，彼の詩のもつ深さには到底及び得ないのである。

おそらく，杜詩の凜然とした美しい調和の感覚——それが進学適性能力テストを作成する際に欠けていたのではなかろうか。少くともわたくしには，責任者であった時，完全にこの感覚が欠如していた。今，彼の歿年とひとしい年齢を迎え，はじめて，その詩が単にかよわい詠嘆に留るのではなく，厳密な法則にしたがって鍛えに鍛えられ，“金石に発する力強いひびき”に満ちていることを悟るに至った。対句といい，対問といい，全体のはりつめた緊張の部分であるとともに逆に微妙な調和のうちに全体を構成すべき性格をもっている。この事実を自覚したのは，まったく最近のことである。この意識がたまたまわたくしの過去の

仕事の欠陥に反省を促し、この文を書かしているのであろうか。

理想的な客観テストが成立するとすれば、その問題と解答とは、律詩における各聯のごとく終始一貫してひとつの全体を構成しているはずである。したがって、被験者が問題だけをみて解答を与えることは、まずあり得ない。むしろ逆に、解答が問題を規制しつつ全体に微妙な調和をもたらしている。この理想の姿が、とりもなおさず、科学的思考の本質そのものを示しているのではなからうか。科学者がひとつの独創的な仕事をなしとげる場合、彼は、常に結論を意識しつつ、しかもそれを固定化せず流動的な状況において捉え、かつ絶えずそのフィードバックのもとに現実の実験・調査の結果を解釈している。彼の行動を支配しているのはデジタルな思考ではない。そこには、詩の精神にまがうリズムカルな心的活動がある。それは「真」に迫ろうとする美的衝動といってもよい。この当然の現象に思い到らなかったのは、わたくしの人間としての未熟さに起因する。

こういう研究部長によって設計された進学適性能力テストを回顧し、その特性をひとことで表わせば、それは“詩なき科学”への適性検査であった。すなわち、ひややかな論理のみの構成する世界への適合者を発見する技術であって、科学的創造の一面は捉え得るかもしれないが、その本質を衝き得なかったのである。もっとも追跡調査、理論的研究等で、わたくしは、当時、文部省・大学・高校等の求めている早急な断定をできるだけ避けさせていたことも事実であるが。

それから10余年、このごろ対問という技術自体が変化しつつある。すなわち item の内容を変化させながら思考過程やその中に共通な考え方をみようとする多重対問形式テスト(multi-cross item testing)が出現しはじめた。これは過去の経験にかんがみると好ましい事象であるが、この形式においても、今述べたわたくしの考えをいれなければ、結局、“詩なき科学”に汲々としている人々を選抜するだけのテストに墮してしまうことであろう。

## 文 献

- 1) 能研テストの妥当性に関する研究—追跡調査資料 I—能力開発研究所 1968, 6-7. なお、これらの図は、当時の所員、現筑波大学中山和彦教授、現岐阜大学藤田恵重教授の発案と計算の結果による。
- 2) 職業適応能力テストの妥当性については次の論文参照。樋口伸吾 適性, 肥田野直編 現代教育研究13巻 教育評価所収 1968, 123-145
- 3) 鮫島央子, アイテム特性関数にもとづくテスト結果の分析, 能力開発研究所研究紀要 I. 1967, 6
- 4) 進学適性能力テスト解説 能力開発研究所 1966, 7-8.
- 5) 鮫島央子, *ibid* 8.
- 6) 鮫島央子, *ibid* 18.
- 7) 鮫島央子, *ibid* 29.
- 8) 進学適性能力テストの解説と例題 能力開発研究所 1968, 28-29.
- 9) 樋口伸吾「創造性」と二つの思考型 同著 人間科学への道所収 平河出版社 1979, 45-100.
- 10) 進学適性能力テストの解説と例題 能力開発研究所 1968, 31.
- 11) Lord の理論についての著者の考えは次の著者の論文参照 テスト—その理論的展開 同著 人間科学への道所収 平河出版社 1979, 7-40.
- 12) 吉川幸次郎 杜甫詩注 第三冊 筑摩書房 1979 328-329.
- 13) 吉川幸次郎 *ibid* 132-135.

## PARALLEL ITEMS IN THE OBJECTIVE TESTS

Shingo HIGUCHI

In objective tests high scores might be got easily, because there are always chances where a blindly selected alternative happens to be the right answer.

To clarify this contamination by chance, chance score distributions of the Noken achievement tests were calculated and compared with the observed score distributions. The results showed there was a danger that some chance scores might affect the total scores of the applicant's answers.

The Scholastic Aptitude Test of Noken was planned especially in order to avoid this deficiency, i. e., the questions were constructed with paired parallel items of A and B. An applicant must select one from five alternatives in each group A and B. If an applicant selects at random one alternative respectively for parallel items, the probability he gets a true answer for this question is  $1/25(1/5 \times 1/5)$ . So there is almost no possibility that he may get high score in this test.

It is very sophisticated method to avoid chance score. This formula seems to be a creative trial. But the planners of it dared to say that this method could warrant even the consistency of thinking at the problem-solving processes. The author suspects this idea and wishes to examine closely the validity and effectivity of the parallel items technique.

Two selected test questions of SAT were given to 20 subjects. They were asked to report orally about the contents of their thinkings, while trying to get an answer. Their reports were analyzed and compared with the interpretation of the test planners. Two facts were demonstrated :—

1. The parallel items technique could not always be taken as an useful way of ascertaining that there was only one and best problem-solving process.
2. The parallel items themselves might become some cues for problem-solving. Many applicants tried to get an answer by making full use of the alternatives.

Getting these results, the author concluded that there might be misleading ideas in the constructions of these questions, one of which was the Load's matrix sampling theory assuming that the problem-solving processes are independent in any test items, and that the only one and best thinking way exists always in every problem-solving activities. Further consideration of interaction between question and answer is necessary for making test items.