

哲學研究

第四百五十一號

第三十九卷
第五冊

心理テストについて

矢田部達郎

私は十二年前にこの會で話をしたことがあり、十二年振りであるが、それは京都に來たてのときであつたから、私の京都生活も十二年になるわけだ。その間、哲學會の講演はいつも難しい話ばかりだつたが、今日は心理テストという卑俗な話することになつた。しかしその話のうちの或る部分は科學方法論の一こまになるもので、哲學研究者にとつても、全く興味の無いものではないと思われる。また一方心理學者の方からいつても、哲學研究者から大きな助力を乞わなければならない領域だと思つてゐる。

テストは一時非常な流行になつたが、その時分から既にもう激しい反對論が方々に起つてゐた。今日ではますます人氣がない。それは大別して三つの方面から擧げられたといえる。第一に、一般の人々から、第二には、哲學者或いは科學方法論的な學問に興味を持つ人々から、第三には、心理學とは限らないが、テストに近い問題を取扱つてゐる、特に臨床家たちからの反對である。

第一に、一般知識人からの反對は、心理テストだけではなく、大體テストというもの一般についての攻撃である。その代表的なものとしては例えば、今年（一九五六年）八月の「アメリカナ」に翻譯が出たハーバード大学のソロ

ーキンものなどを擧げることができる。

かれによると、現在はテストマニアによるテストクラシーの時代であるといわれる。しかもそのテストたるやことごとくくだらない。自然界には自然淘汰的なテストがあつて、社會的な地位・階層・身分・職業などが自然に選擇される。このようなテストは立派である。或いは社會成員間にリーダーになつたり、使われる身分になつたりするような生存競争が繼續的に行われているが、これも非常に重要なテストであるといふことができる。これらをかはは自然的な永續的テストと名づける。さらに自然社會にはもつと現實的なテストがある。例えば或る緊急事態に直面すると人はそれによつてその能力をテストされる。或いは或る職業だの宗教團體だのに入る場合に行われる見習いの試練のようなものにあつては、人はやはりその實力をテストされる。これらの自然發生的なテストは多くは實質的で有效なものであるけれども、今日のテストマニアが用いるテストはすべて甚だ人爲的なテストなのである。ただその内一つ、實演的テストと名づくべきものはまあよろしい。例えば、オーケストラのメンバーを募集する時にヴァイオリンを弾かしてみる。自動車の免許を與える場合に自動車を運轉させてみて、うまく運轉できれば與える。これはよろしい。しかし流行のテストはいわゆるペーパー・ペンシルテストであり、筆答試験か、口述試験である。これらはすべて大きな缺點をもつていと主張された。

そういう缺點を數えてみると、まず第一にそういうテストは一回限りで或る人間の能力をテストしようとする。これはやり損う公算が非常に大きい。一回でなくても、能力の判定は難しいもので、例えばトルストイや、プーシキンは、學校の國語の成績が非常に悪かつた。ヘーゲルはあらゆる學科が非常に良かつたけれども、ただ哲學だけは全然駄目であつたといわれる。ニュートンは大學を出る時に、碌な學科は一つも無かつたし、ベートーヴェンは作曲家としての才能を、先生のハイドンから認められなかつた。このように長い間その人間について知つてゐる先生とか、或いは友達という場合であつてさえ能力の判定はむずかしいものである。それを一回或いはただか數回で判るはずが

ない。さらに試験となれば、時間の制限がある。一時間で書けといつてもろい人は書けない。しかしのろま必ずしも馬鹿とは限らないのである。それにさらにもう一つ悪いことは、試験というものは大體假定狀況についての試験であつて、こんな場合にはどうするかと聞くわけであるが、しかしいざとなればほんとうはどうするかわからない。また試験問題には不適當なものが非常に多い。まず第一に試験委員が間違つた答案を正しいと思つて出している場合がしばしば起る。ハープと太鼓とヴァイオリンとピアノとの繪を描いておいて、このうち一つ違ふものはどれかと問うタイプの問題がある。試験官はタイコ（打楽器）といわせるつもりだろうが、ピアノだつて打楽器だともいえるわけである。或いは質問のカテゴリーがはつきりしない場合も多い。「あなたは宗教書を讀むのと、經濟書を讀むのとどちらが好きか」という問が與えられる。これは思想調査とか性格検査などによく出てくる問題の形で、どちらか一方にきめて答えなければならぬように書いてあるのであるが、場合によつては、經濟書を讀みたくない時もあるし、宗教書が讀みたくなる時もある。そのほかテストの問題は多くは知識の切賣りだの、單純な記憶のテストに關するものであつて、そんなもので人間の本當の價値を計ることはできないという非難も行われる。また特に人文領域では、學説が異なれば正答と誤答とが全然違つてくる場合も多い。ハイデガーの“Nichts”と鈴木大拙の“Nichts”とは同じニヒツでもニヒツがちがう。先生というものは大概頭が狭いから、大抵はどちらかを要求している。私の知っている先生で偉い先生があつて、自分の説でないのを書くとき非常にいい點をくれた先生があつたが、そんな先生はめつたにいない。

以上のような缺點は極めて簡單であつて別段の説明を必要としないが、この他やや説明を要する問題が二つばかりある。その一つはテスト結果に對する解釋の曖昧性ということであり、その二はテスト成績の量化に關する問題である。それについて素人の意見の代表という意味でソローキンの説くところから従つて述べてみよう。

ソローキンによると、人文關係のテスト結果にはどうしても解釋が入る。自然科學のものだと、例えば今日は三十

二度だということは寒暖計を見さえすれば誰にでもわかる。ところが、心理検査などでIQが百三十だといつてもその意味は極めて漠然としている。ひどいものになると、フロイドの夢判断や、色んなコンプレックスの解釋のようなもので、全く勝手な解釋が跳梁している。これがソローキンなどのような素人の反對であるが、これは或る意味では非常な間違えといわなければならない。テストの結果は解釋されなければならぬということは正しい。しかしそれは自然科學のテストにおいても異なるところがない。熱を測つて四十度あつたとする。なるほど四十度にはちがいないが、それが何の熱であるかは解釋を必要とする。チフスの熱であるか、猩猴熱であるか、異種蛋白を血管中に注入したためであるか、それは別に解釋されなければならぬのである。これと同じように知能検査をすればIQは百三十なら百三十とちゃんと出てくる。ただ百三十度といわないだけである。それから先の解釋はもとより必要であるが、それは自然科學の場合と異なるものではない。解釋が怪しいかどうかはその科學の進歩の段階によるので、心理學でも相當信賴するに足る結果もあれば、自然科學でもまだ極めて曖昧な領域もある。そこに何等本質的な相異はないのであつて、自然科學であるのと人文科學であるのとの別はなく、すべての科學がみなこの曖昧さを征服すべく日々の努力を傾けているわけなのである。

第二の量化的問題については精神的なものを量化しようとするのは全く無意味であるというのがソローキンなどの意見である。もちろんどうにもならない場合も多々あるのであるが、しかし極めて有效な場合もまた少くない。われわれの問題になつている試験成績の量化なども、通例の取扱いは甚だ不合理であるので、心理學者はそれをいろいろ工夫して改良しようとしている。

ソローキンなどはそれも結局量化できないものを無理に量化しようとするのであるから意味がないというように説くのである。しかし次のような改良を無意味であるという人はおそらくあるまいと思われる。例えば入學試験のとき數學の得點と國語の得點とをそのまま加算して合計點を出すとする。この場合もし例えば數學の問題が非常にむずか

しくてほとんどが五〇點以下であり、國語の問題が非常にやさしくてほとんどが五〇點以上であつたとすればどうか。原理的には數學と國語とは同じウェイトと考えているにしても、實際には國語の出来る者が得をし、數學の出来る者は相對的に損をすることになるのは明らかといわなければならない。こういう不平等は學科の性質によつても起り、例えば數學では〇は稀れでないのに、作文ならまず〇をつけることは少ないであらう。同様に出題者の難易、採點者の辛甘によつても著しい差ができる。そこで心理學者は標準得點というものを算出する方法を工夫したのである。その原理は各學科成績を受験者の成績順位で表わし、その總計を總合點とすることに盡きる。しかし實際には細點において種々なる工夫が行われる。その最も簡単なやり方は百分段階（パーセンタイル）といわれ、受験者の成績順位を自らの位置より下に來る受験者の百分比で表わすものである。つまり五十人中の二十五番は五〇パーセンタイルといわれる。これをかりに五〇點として他の學科の同様の處理による得點と合計すれば、よし學科の性質、問題の難易、採點者の辛甘があるにせよ、また受験者數が學科によつて幾分の差がある場合にも、極めて合理的な總合得點を算出することができるわけである。しかし心理學者はさらに得點の分布の不規則性などを考慮にいれて、いわゆる標準得點なるものを使用するようになった。これは原理的にいつて百分段階と異なるものではないが、それよりもより公平な採點を可能にするものと考えられている。標準得點すなわちスタンダード・スコアといわれるものは得點の分布の標準偏差を基礎として作られる。それにはいろいろな表示法があるが、そのうち最も基礎的なのは z スコアと呼ばれ、次のように表わされる。 $z = \frac{(x - \bar{x})}{s}$ この場合 x はある人の得點、 $\bar{x}$ は得點の平均、 s は分布の標準偏差である。しかしこの表示法では中位者が零點となり、その上下にプラスとマイナスの得點を與えられる者が並ぶことになるので、中位者を五〇點とする操作を行つて、次のような式が使用されることが多い。即ち $T = 50 + 10z$ これを T スコア或は偏差値と名づける。 T スコアは標準偏差が 15 であれば一〇〇點から〇點まで分布することになるが、通例は標準偏差はもつと小さい。それで $H = 50 + 14z$ といふ式を用いこれを H スコアというが、これならば標準偏

差が H_{95} であつても得點は九九點から一點まで分布することになる。或は知能指數と一致させるために知能等價量 $100+17.5x$ という式を用いることもある。これらのうち何れを用いるかはその時時の都合によるわけである。これらの採點法は心理學者にとつては極めて初歩的な知識に屬するものであるが、この方面に縁遠い人々のために簡単に説明した。

これはほんの一例に過ぎないが、量化的有効である場合はいたるところに見出される。もつともこの例は評價そのものの妥當性を證明するものではない。不完全な評價を如何にして改良し、役に立つ程度にまで引き上げるかという問題であつた。評價の不完全性は例えば人格の全體的評價の場合などには殊に著しい。しかし或る特定の目的に對して或る人がどの程度適當な能力をもっているかということの評定するような場合には、かなり適切な評價が可能であることも少なくない。この多少とも存在する不完全性を有用な程度にまで除去することは量化によつてはじめて可能であるということが出来る。これを精神的なものの量化は本質的に不可能だからといつて全面的に排斥するような態度は餘りにも見解が狭い。自然界には自然淘汰があると同時に人爲淘汰が行われるということも極めて自然的であるといえるのであるから、人爲淘汰としてのテストを無視することは適當でない。不完全であるからといつても試験を全廢することはできないのであるから、これを改良する努力はどこまでも眞剣にしなければならぬ。量化ということがこの改良の最も有效な手段の一つであることには疑いをさしはさむ餘地はありえないのである。

第二に、哲學者或いは科學方法論的な學問に興味を持つ人々からの反對論について考えよう。このうちにもまたいろいろな立場からの反對論が存在するが、その一例として一昨年出版されたパウ・メール P. E. Mehl という人の説を取り上げて見よう。

この人はミネソタの心理學の教授で、精神病學も教えているが、この人が "Clinical versus Statistical Prediction, 1954" という本を出した。この本の中でメールはまず二つの對立する意見のあることを述べている。一つは多くの

臨床家達の説で、臨床家は直觀的に診断を下せばよいのであるから、テストなどは役に立たないという意見である。それに對して他は統計家の意見で、結局直觀的といつても何にかやはり統計的な基礎の上に、別の言葉でいえば過去經驗の上に、判断を下しているのだから、全然單一の判断ということはありえないという立場である。これはサーモン T. R. Sarbin とカランドバーク G. A. Lundberg とかいう人々の意見である。メール自身は、結局両方とも必要だといふのであるけれども、形式的な統計説の説くもの以上に何か直觀的なものの存在を力説するところに特徴が認められる。オールポート G. W. Allport などは統計輕視説であつて、統計的な豫言はうまく當らないが、臨床的な豫言はうまく當ると主張した。そこでメールは從來の心理學的論文を集めて、オールポートのいうことがはたして正しいかどうかを調べてみた。その結果はやはり統計的な材料からの豫言の方が、事例研究のようなインフォーマルなデータによる豫言よりもよく當つていることが明らかにされた。しかしそれにも拘らず心理學的な名醫とも稱せられるような臨床家はしばしば立派な診断を下す場合のあることも否定できない。そこでこの *clinical method* と *statistical method* (*actuarial method*) の差はどこにあるかということが問題になる。これは結局昔からの、*uniqueness* の *lawfulness* の問題、*idiographisch* なのと *nomothetisch* なの、或は *verstehen* と *erklären* の問題といつたことに通じる問題である。簡単にいうと、臨床家は統計家が使用することのできない多くの特殊因子を考へる可能性があるから、それがうまく見つかつた場合には名診断がつくことになる。統計家の方は、一般因子だけしか考へられない。あらゆる因子を全部式の中に入れて計算するといふことはできないから、個別的な豫言では當らない場合が多々出てくる。しかし全體からいえば、統計的處理の方が當る率が多くなる。従つて結論的にいえばこれは両方とも使わなければならないといふことはつきりしているように思われるのであるが、實際にはどうも二つの立場が對立する傾向が著しい。これはしかし、醫者の場合を考へると極めて簡単に解決がつく問題ではあるまいか。醫者は多くのテストを用意している。熱を測り、血沈を測り、その上さらに打診もする。醫者にテストを止

めるとか、聴診器をすてろとかいう人がないように、心理學者に對してもテストをやめろ、統計をやめろというのは随分おかしい話ではないかと思われる。何故このような馬鹿氣た議論が行われるか不思議であるが、現實には世間ではこのためにどれだけの時間をつぶしているかわからないのである。

そこでもう少しこの問題を分析して考えてみよう。ディルタイ一派の人々は法則的理解即ち説明は合理的に行われるのに對して、了解は直觀的に行われると考えているようである。しかし、合理的とか直觀的とかいう語義の問題はしばらく論外におくことにしたいが、法則的理解に對しても直觀は不可缺の因子であるし、逆に了解にとつても或る種の合理性は到底缺くことのできない前提となるものである。まず法則を作り出すためには通例實驗的操作を必要とするわけであるが、その際如何なる條件を實驗的變數に選ぶかは、全く研究者の直觀に委ねられるのである。これを直觀といつても、類推といつても、或はまた研究者の勘といつても同じことであるが、とにかくその人の科學的研究經歷によるその領域の知識がバックになつて、或る特定の條件がその現象の決定因子であろうということを洞察し、當該條件を操作することによつて、もしその洞察が適當なものであつた場合には、そこに法則の定立が成功するのである。のみならず一旦法則が見出された後においても、これを個々のケースに適用するに當つては多くは再びそこに直觀の助けを求めなければならない。例えば銅の融點は約千度だという法則的知識をもっているにしても、與えられた金屬が果して銅であるか否かを鑑別することができなければ、この知識は何ら豫言に寄與しないであろう。銅の鑛石の條痕が綠であることを知つていても、さらにその色のニュアンスによつてもつと精密な判定を下すためには、多少ともエクスパートでなければならぬであろう。このように法則的知識も全く合理性のみの基礎の上に立つてゐるものではない。

これと同様に了解法の根據にも常に或る一定の合理性を、即ち法則性を見出すことができる。例えば日本の封建制度を理解するためには歐洲や中國における同種の制度が歴史の流れのうちにどういう位置を占めるかを明らかにし、

これとわが國の事情とを比較綜合することによつて、より完全な理解がもたらされるであらう。これは一般的な歴史の流れのうちに問題の研究對象を定位してみることであり、取りもなおさず或る種の法則的理解であるといふことができる。或いは或る醫者が全く特異の症狀を呈する患者を與えられたとき、よしそれがかれにとつて一回限りの経験であるにしても、例えば多くの抗生物質の無効であることによつて細菌性疾患ではあるまいとか、或る種の症候がないからこの種の疾患ではあるまいとか、種々な既成知識によつて診斷の方向を導いていけるのは、やはり合理的思考といわなければならない。

このように法則的理解にも、またいわゆる了解にも、共に直觀性と合理性とが含まれるとすれば、これらはかの人々の主張に反して、科學研究法の二つの異なるアプローチの仕方ではなく、同じ一つの研究法の異なるフェーズに過ぎないと考へべきではなからうか。一般に理解は或る事物を既成の知識體系に關係づけることによつて行われる。それは組織的な法則的知識體系であれ、或いはもつとルースな單位に分類され、整頓されたというに止まる常識的経験の集合(典型)にせよ、とにかくそういう既成知識の特定の場處に定位されることによつて、新経験はわれわれによつて理解されるのであり、その経験はそれによつて説明されたといわれるのである。だから典型はいわばインフォーマルな法則に過ぎない。法則なり典型なりは前にも述べたような創造的活動によつて経験のうちから組織され、かかる組織された知識に新事實を包攝させる場合に、説明なり了解なりが起る。法則定立學というのはこの初めの過程を強調したものであり、了解的な個性記述學というのはこの後の過程を強調したものに過ぎない。自然科學ではこの初めの過程が目立つのは事實であり、精神科學や歴史學ではこの後の過程が目立つのは事實であるが、法則を新事實に適用しようとする自然科學がないのと同様に、基礎となる経験體系を缺如する精神科學もまた考へられない。この経験體系は法則體系というにはあまりに精密でないかも知れないが、比較、概括等いわゆる歸納的處理によつてえられるものであることは法則的知識の場合と異なるところはないのである。

このようにこの二種類の科學または方法といはれたものは、實は何ら本質的差異をもたないのであるが、しかし見かけの上でかなりの相異を示すこともまた事實である。特に自然科學と歴史學とを對立させるような場合にはその差は益々甚だしい。これがウィンデルバントをして法則定立學と個性記述學とを對立させた事情であらう。しかし歴史學がその資料の聚集のために歴史的事實の記述をすることは、原理的には自然科學がその資料聚集のための記述をするのと少しも異なるところはない。また歴史學がただ一回きりの時間的現象を取扱うということも、たとえば天文學に對しても同様にあてはまる事柄に過ぎない。文化的價值が問題となるということも、天文學等においては天文學的な價值が問題になつてゐるのだということを見逃した立言である。各科學はそれぞれの價值を追及するものであつて、自然科學においてはただそれが人文的價值とはちがうというに止まるのである。一般的にいつて如何なる科學も廣義における歴史的敘述を必要としないものはいえるだろう。

そこでウィンデルバント等が自然科學と歴史學との差と考えたことは、實はそうではなくて、法則定立をその主要任務とする理論學 *Theoria* と、個別的事實を理解することを重要任務とする技術學 *Techné* もしくは臨床學との間の差であつたのではあるまいか。この理論學と技術學との區別は從來は基礎學と應用學若くは臨床學との區別といわれた。しかし技術學はいわゆる基礎學の單なる應用ではなく、それ自身の領域において構成される法則體系をもち、各自獨特の基礎的部門をもつものである。例えば建築學は力學の單なる應用ではなく、内科學は生理學や解剖學の單なる應用ではない。その特徴は一般的法則的知識體系の構成にあるのではなく、むしろ個別的ケースに對する實用的効果を目的とするところにある。建築學ならば或る特定の建造物を作ることが目的であり、内科學ならば或る特定の病人をなおすことが目的である。しかしいづれも立派な獨立の學問であつて、他のいかなる基礎學に從屬するものではない。歴史學はこれらとはやや異なるもののようにも見える。それには歴史哲學の如き歴史の流れに對する一般的方法則を問題とする理論學をも含むが、通例は個々の出來事を歴史的知識の體系中に定位して、それを理解すること

を任務とする個別的理解の學を意味することが多い。この種の學問は他の技術學や臨床學とちがつて未來における實施、治療に向けられるものではなく、むしろ過去への連關を問題するものであるとも考えられるが、それにも拘らず或る一般的知識の個別的ケースへの適用である點において三者とも重要な共通點を有するものといふことができる。このことは歴史學中でも考古學などを考えると、或る出土品の時代的判定などにおいて、臨床醫學などと極めて類似した操作が見られることによつても知られるであらう。

そこでこれら技術的學問においては個別的現象に對して、既知のいかなる知識を適用すべきであるかということが中心問題となるわけであるが、その場合には勘とか、經驗とか、熟練とか、洞察とか、直觀とかいふ、いわゆる名人藝が働く餘地が甚だ大くなる。問題をテストに關連した領域に限定して考えていこう。いま或る一人の問題児が與えられたとしよう。臨床家はおそらくその子のケースヒストリーを熟讀するであらう。またその子の一般的舉動を注意深く觀察するであらう。それからその子とのラッポールを親密にすることを努力するであらう。そしてその後になん種な分析の技術に訴え、プロジェクトヴィテクニク等を使用して、診斷を下だし、治療に取りかかるに違ひない。治療の或る方法が失敗すれば、さらに他の思考方向に添うて再び別の努力を繰りかえす。その際いわゆる統計的テストというようなものは全く問題の外に置かれる。極言すればそんなものは全く不用であるともいえるだろう。

なるほど今日の心理學において統計的テストといえば、幾種類かの知能検査や適性検査と、インヴェントリー式性格検査、内田クレペリン検査、向性検査等極めて限られた武器を有するに過ぎない。しかもそのうちの或るものはテスト構成上に缺陷があつて、使用に堪えないものも含まれている。従つて實際上それらに頼る必要がある場合は極めて稀れであるのが實情なのである。しかし或る兒童のアイキューと成績との間に大きな開きがあれば、その兒童に何か問題があることは明らかであり、或るクラスの向性點が著しく片よつていれば、そのクラスには何か考慮すべき問題が存在することに間違ひはないであらう。前にも述べたように醫學の領域においては多くのテストが作られており、

それと臨床上の種々なる所見との綜合評價として、正しい診斷が行われている。また或る治療法がそれ自身診斷の資料たる役割を演じていることもある。これらの事實を参考にすれば、心理學のテストが無能なのはテストそれ自身の罪ではなく、この領域が未だ充分に開拓されていないことに主たる理由があると考えられるのである。もとより醫學においてもテストが充分な効果を示さぬ領域においては、種種なる直觀的手段が用いられる。

これに反して或る種のテストはそれだけでもほとんど確實に診斷の基礎を提供しうるものもある。例えばワッサーマン反應がはつきりプラスであるならば、多少の例外はあるにしても、或る病氣の存在を確認して差支えないであらう。このようにテストの効果性はそれに關する技術の發達度に應ずるのであつて、一〇〇パーセントの確率をもたない限り、そのみにたよることはできないにしても、綜合評價の資料として極めて重要なものであることは論をまたない。

現在の心理學の發達段階においては、テストの信頼性があまり大きくないということは事實であるといわなければならないまい。これに對してインフォーマルな資料の上に營まれる心理的診斷が案外有效である場合が少なくないこともまた事實のようである。しかしこの後の場合は結局長期間に亘る患者との接觸中において、種々の假説を試してみる機會があり、その間に動物的本能的直觀によつて、病因をかき出すことに成功するのであつて、出来るならば診斷の基盤を科學的なものとし、テストの形にまとめ上げることが望ましい。もちろんテストが充分に發達しても、その適用に直觀を必要とすることは繰り返して述べた。またその點については次項においてさらに詳しく考察してみたいと考えているのであるが、ここにはただ直觀は常に組織化された一般的知識の基盤の上においてのみ有効に働さうることを強調しておきたい。

さて第三は比較的専門的な臨床家が特に統計的なテストに對して表明する不信の問題である。テストには知能検査や性格診斷目錄のような統計を基礎とするものと、プロジェクト・テクニクのように主として直觀にたよるもの

との二種を分けることができるが、臨床家達のうちには後者に對しては信をおくが、前者に對しては全くその重要性を認めない人達が多い。それは結局統計的テストが實際の役に立たないからだといえるだろう。しかしそれが役に立たないということには色々な理由があり、また誤解もあるようである。まず第一にテスト構成が未熟のために起る缺點で、或るものはテストそれ自身のうちに矛盾のあるものもあり、また或るものは現場から妥當性に關するテスト資料が充分戻つて來ていないために期待されるほどの効果を發揮しえないものである。特に殘念なことは、わが國で最も盛んに用いられた向性検査が古い時代の製作であるために、内的整合性において缺けるところがあり、各アイテムが綜合點に寄與する仕方に矛盾が認められる。こういう不完全なテストが現場で効果を發揮しえないことはいうまでもない。今日ではしかし京都で作られたYK及びYA兩向性検査の如きはこの點では缺陷が除かれるようになった。ただしそれらは成立後日が浅いために現場からフィードバックされる資料がまだ充分でなく、診斷的應用の面はまだ限られているといわなければならない。このように統計的テストは無用だというのはなく、現在でも役に立つている面もあるのであるが、さらにもつと役に立つように、學者も現場も協力して改良に努力をいたすべきであるというのが、この問題に對する正しい理解であると思われる。

次に問題となるのは自己診斷テストには根本的な缺陷が内在するという考え方である。つまり自分に對して判斷を下だす場合人には多少とも身びいきがあり、或は故意にいつわりをいうことさえできるというのである。しかしこの考え方ほど批評者の無知を暴露するものはないともいえる。あらゆる發達した科學と同様に自己診斷法もまた見かけの回答をそのまま學問的知識として取り上げるものではない。あなたは正直ですかと問うてイエスと答えた人はすべて正直であるなどとは誰れも考えないであらう。だから極端に言えば正直性テストで高い得點をえたものも必ずしも正直であるとはいえないのである。そういう人は見かけの上での（即ち顯型としての）正直性の高い人とはいえる。しかし顯型といわゆる原型とは必ずしも一致しないのである。原型を決めるには別の科學的處理を必要とする。それ

はテストの場合なら、別の見地から或る一定の特徴を有することがわかつている集團に對してそのテストを施行してみても、その集團の得點が他の集團の得點に比して特に高いか低いによつて、そのテストの得點の高低がその集團の特徴を代表するものと決めるわけである。だからかりにあなたは正直ですかと問われてイエスと答える人數が非行少年の集團において著しく多いとすれば、この顯型としての正直性は原型としては非行少年性を意味することになる場合もありうるわけである。この科學方法論の初歩的知識さえもたずに科學的處理に對する批評をする人が少なく、しかも教育界においてはいわゆる人格者（これも顯型に過ぎないが）といわれる人々において特に著しいのは全く心外にたえぬところである。もつと虚心坦懷に教育界に科學的方法を導入するようにしたいものである。これと同様の誤解は、自己診斷テストは年少者には適用できないという意見のうちに認められる。自己判定は高級な精神機能であるから青年期に至つて始めて發現するものと考えられる。しかし自己診斷テストで用いられる自己診斷は「見かけ」の自己判定であつて、自分は正直であるか否かを眞に反省しうるか否かはここでは問題ではなく、自分は正直であるか「思う」かどうかで回答はできるわけであり、その回答が正直という概念に正しく合致するか否かはこの場合には問題にならない。つまりこの場合には前にもべたようにそういう回答を多く、或は少く與える集團が或る一定の特徴（正直性と無關係でもかまわない）をもっているかどうかということだけが問題にされるわけだからである。現に上述のY A 向性検査の如きは小學校二年生に對しても施行可能であり、一年生に對してさえ個人テストならば或る程度の結果を導出することができるものと考えられる。このようなことは自己診斷を顯型的に考えていたひと昔前までには夢にも考え及ばなかつた事項といえるだろう。

以上においては統計的テストが診斷法として本質的には何ら不都合なものを含まない點を強調したのであるが、その點を承認した上でもなお臨床家からの非難は終らないのである。それは統計的に豫言できるのは高々一般的傾向であつて、個別的なケースの診斷にはそれは全く無力であるというのである。或るテストで或る個人に或る特定の指標

(徴表) a があることが認められたとして、その徴標 a を示す個人は例えば三〇%或る特性 A を有することがわかつていたとしても、今臨床家の診断を求めている一定の個人が果して特性 A を有するか否かは七〇%の誤りを犯す覚悟でなければ豫言できない。これでは豫言でもなんでもないといわなければならぬ。もちろん一〇〇%の確率をもつて豫言を可能にするテストがないわけではないが、これはむしろ例外であつて、多くは他のテストやその他色々のインフォメーションを集めて総合的診断を下すのが通例である。わかりやすくするために醫者の場合に例を取ろう。小兒結核のレントゲン像は影のあるもののうちの三〇%に過ぎないといわれる。他はむしろウィールス性肺炎と診断した方が間違いが少ない。しかし熱型だの、血沈だの、或は家族に結核患者があるかどうかというようなインフォーマルな問診などを参考とした綜合評價の結果として診断が下される。その際レントゲン像は三〇%の確率しかないから、むしろ撮らない方がよいという主張はなり立たないであらう。のみならず一個のテストの確率は三〇%でも、通例は幾つかのテストを同時に適用するのであつて、あたかも三つ或は四つの交叉する圓のいずれにも含まれる領域が極めて狭い領域に限られるのと同様に、テストだけでもその組合わせの如何によつては極めて有效な判定資料を提供することができるのである。醫學において症候群といわれるものはこの理を診断に適用したものに過ぎない。テスト構成の仕事がまだ充分發達していない科學においては、症候群は漠然たる印象によつて、或いはいわゆる名人藝によつて捉えられる。これは *physiognomic cue* 或いは *Gestalt cue* と名づけることができるであらう。學問が進むにつれて症候群を規定する因子が明らかにになり、それらを發見するためのテストが構成される。テストができれば直観はいらぬというのではない。如何なる法則でもその個別的ケースへの適用には直観を必要とするものであることは前述の如くである。しかし出來うる限りは直觀的印象を科學的方法で客觀的に規定することが學問の理想であり、それによつて名人藝を萬人のものにすることができるのである。

この間の事情は醫學の進歩を振りかえることによつて明らかになる。百年ほど前には醫學の領域にも疎なテストは

出来ていなかった。否、最近までかの體溫表の如きが診断の主たる材料になつていたのである。チフスの熱型はこうである、肺炎の熱型はこうであるというように、熱型のプロフィールから診断が行われていたことは今日の心理診断の水準に極めて近いものであつたといふことができる。ところが病型がだんだん複雑になつて熱型が亂れてくると、このような簡単なテストでは間に合わなくなり、レントゲン検査だの、血液検査だの、その他色々なテストが作られるようになった。これは餘談にわたるが、筆者が喉頭癌を患つたとき、京大耳鼻科の後藤光治博士の診断を乞うたが、博士はまず喉頭鏡で筆者の口の中をのぞき込み、すぐこれはという顔付をされた。しかしその後約十日にわたつて色々なテストをされる勞をも惜まれなかつたのである。レントゲン像を取つて結核性の影のないことから喉頭結核ではないことを確かめ、大動脈の肥大を認めないことから舌咽神経の壓迫でもないことを確かめ、その後には醫者の義務としてワッサーマン反應をたしかめることを要求された。その結果がわかる日筆者は教授室に博士を訪ねたが、留守を守る女事務員からおめでとうという言葉とともに結果が陰性であることを告げられた。筆者は苦笑しながら自分の場合はおめでとうではなく、死刑の宣告にも等しいのだといつて笑つたことであつた。このように宣告が殆んど確定した後始めて博士は患部の組織をとつて検鏡に付せられた。筆者の場合は天野重安博士と太田邦夫博士とが検鏡されて、共に立派な癌であることを認定されたのであるが、この検鏡などにはまだ名人藝を必要とする面が残つてゐることであり、これも近い將來にはインターンの學生にでも委かせられるようなテスト化が望まれる領域といえるだろう。このように最初の診察で癌であることは直觀されたのであろうにも拘わらず、後藤博士がその後慎重なる科學的處理に訴えられたことは、當然のこととはいひながら、筆者に深い感銘を與えたことであつた。このことは内科ばかりではない。外科手術の如きも昔は脈を取りながら、顔色を見ながら、出来るだけ早くすることを努力していた。しかし今日の大手術では脈搏も血壓もポリグラフに記録しながら、はつきりその程度が眼に見えるようにし、血糖濃度やペーハーなどを絶えず測定しながら、もし足りなければ輸血によつて適當に補なうことができるようにしてい

る。だからあわててガーゼを縫いこむようなことも起らないようになった。これはやはりテスト技術の進歩とも密接に關連した事柄であるといえる。

テストには通例の診斷的テストの他に治療がそのままテストになる治療的テストというものも考えられる。數年前新しい抗生物質がどんどん發見されたとき、醫者はまずペニシリンを用ひ、それがきかなければオーレオマイシンを用ひ、それがきかなければまた何にというように投藥の順序を決めていると皮肉られたことさえあつた。これは治療がそのまま診斷になる好例ということが出来る。治療的テストの働く機構は理論學における假說演繹法の機構に似ている。すなわちいずれも假說を個別的現象に適用して、一方はその假說の正當性を證明しようとし、他方は治療法の出發點を確立しようとするのである。これは結局試行錯誤の原理によるものといふことができる。この種のテストは若しそれが成功する場合には同時に治療が完成するのであつて、臨床學にとつて極めて重要な手段であるといわなければならない。しかも或る治療法によつて治療する疾患はその治療法によつて定義することも出来るわけであつて、その治療學の組織化にとつても極めて有效である。しかし現在の心理學においてはこの種のテストはまだ殆んど構成されていない。かのババンスキーが説得可能症と名づけたヒステリーの如きはかかる考え方に對する例外であるが、その後この方向に向けられた努力はあまり著しくないのである。プロジェクトヴィテクニクの如きはもちろんこの種のテストではなくて、統計法を重視せざる通例の診斷テストに過ぎない。

このようにテストの種類は數多くあるのであるが、實際上ただ一つのテストで診斷を下しうる場合は上述の治療的テストを除いてはあまり多くはない。ジフテリアのワクチンで治るものはジフテリアであるとか、キニーネで治るものはマラリアであるとかいうことはできる。しかし診斷的テストの場合には通例幾つかのテストの結果から綜合的評價を行わなければならない。この場合には色々の治療上の所見や、インフォーマルなインフォメーションやを參考として最後の診斷が下される。それにはもちろん臨床家としての經驗や洞察が大いに物をいうわけである。醫學上の

テスト結果は數年間の基礎醫學の學習をし、さらに何年間かの臨床經驗をつんだ醫者という専門家の解釋に委ねられるのに、心理學上のテスト結果は現在の日本では時には一冊の心理學概論をさえ讀んだことのない教員に委ねられる。そういう人達がテストから有効な結論を引き出しえないとしても何の不思議があるのであらうか。またそういう人達がテストに對する不信を表明したとしてもそこに何の權威があるといえるのであらうか。特に古い教育者たちにはこういう教養を受ける機會がなかつた人が多いのであるから、かれらが醫學の學說に對して敬虔であるように、心理學の問題に對してもつと謙虛であることが望まれる。そしてかれらの理解によつて、日本の若い教員たちが今日の日本の醫者と同じ程度の心理學的教養を身につける日が來たならば、おそらく心理テストもまた世間から疑惑の眼を以つて見られるようなことはなく、日本人の精神的健康を増進し、精神的壽命を延長するために大きな役割を演ずることができるようになるであらう。

(了)

附記

この論文は昭和三十一年十一月二十四日午後、京都哲學會公開講演會に於ける同博士の講演を基礎とし、更に書改められたものである。なお講演の録音には心理學教室の木下助手が、録音の整理には同教室の高田登、名倉啓太郎兩君が當たられた。

(編輯者)

筆者 前京都大學文學部「心理學」教授、現
早稻田大學大學院文學研究科「心理學」教授

THE OUTLINES OF THE MAIN ARTICLES IN THIS ISSUE

The outline of such an article as appears in more than one number of this magazine is to given together with the last instalment of the article.

On Mental Tests

By Tatsuo Yatabe

Distrust in mental tests which has been expressed among various kinds of people may be classified into that of 1) non professionals, 2) of spiritual scientists and 3) of clinicians. R. A. Solokin's argument on Test Mania (Harvard Educ. Rev. 1955), if cited as an example of mistrust among non-professionals, is of course sufficiently convincing in some respects, but his criticism is justified rather in blaming the mental test because of its immaturity, than in distrusting it as such. His most layman-like reproach is expressed in his claim that the results of the psychological test has many ambiguity and an interpretation should be always necessary, while there are no rooms for interpretation concerning with the results of test in the natural science. I cannot endorse his opinion that the results of test in natural sciences do not need interpretation. His another amateurish criticism is the claim that the quantification of the mental in general should be impossible. This claim is only due to ignorance of many instances where the quantification is effectively performed.

The spiritual scientist's criticism says that the statistical method ought not to be used in the idiographic sciences. But every science has two aspects in its structure, the nomothetic process i. e. the constructing processes of the general knowledge, on the one hand, and the applying process of the knowledge to a particular case, on the other. Statistical operation of thought prevails in the former and intuitive thinking in the latter. It cannot hold good to consider these two as-

pects as two different sciences by cutting off them from each other. Therefore, as the nomothetic sciences need intuition in their process of application, so the idiographic sciences should be regarded as they need more or less statistical thinking in the process of construction of the knowledges which form the very basis of the description.

Against the clinician's attack on the statistical tests, which has been focused principally on the self-diagnostic inventories, we can stand on three points. a) There are, indeed, now in use some imperfect tests, but such a test must be improved, needless to say. b) The criticism says that the self-diagnosis is inevitably biased. But it matters little with the apparent bias, because the responses on each item of the diagnostic test should not be used in their phenotypical meanings, but in the genotypical meanings which are induced from the special validity testings. c) The criticism says that the statistical prediction can not be applied to individual cases. To this criticism we can answer that the prediction should be enough convincing, because the area, which several tests point out in common, is to a degree limited. Besides, the diagnosis must be always a synthetic evaluation and diagnostic tests are to be used together with other data, i. e. with therapeutic tests as well as other in formal informations. In such an evaluation test score plays rather the part of an indispensable component.

Eindrücke aus der Reise in Europa und Amerika

von Masaaki Kôsaka

Masaaki Kôsaka, Professor der Philosophie der Pädagogik, der auch Dekan der pädagogischen Fakultät der Kyoto Universität ist, reiste vom Frühling bis Sommer in 1956 durch USA, England, und Deutschland, besuchte viele Universitäten, und unterredete sich vertraulich mit vielen Denkern. Seine Eindrücke über die europäischen und amerikanischen gegenwärtigen Tendenzen der Philosophie sind folgende:

- 1) In USA. a) Die Philosophen in USA haben sich nun wieder neu