『現代教育心理学大系／人間発達の基礎的理解』一九五八年九月（中山書店）

Ⅶ テストの価値と限界

3 教育学的立場

矢口　新

一　教育的テストの意義と役割

1　教育活動におけるテスト

教育活動の中においてテストは主として何に利用されるであろうか。いかなる目的で、どんなことが行われるのか。テストといってもメンタルテストもあれば、アチーブメントテストもある。今ここでは主としてアチーブメントテストを中心にして考えることにするが、もちろんメンタルテストについても必要な限りはふれるであろう。

ところでアチーブメントテストというと、日本ではいわゆるアチーブと言われて、上級学校の選抜考査みたいに考えられているが、それは誤解である。最近使われている学力調査とか、学力検査などというのも、結局はアチーブメントである。メンタルテストは、これに対し学習以前すなわちアチーブメント以前のものを調べると考えてよいであろう。

さて教育活動の中にアチーブメントテストやメンタルテストを置いてみると、共にテストであるが、その意味は、大いに異なるといわなくてはならぬ。今メンタルテストをいわゆる知能検査という狭義のものに考えてみる。これは結局は人間の相対的な位置づけである。人間といっても素質的なものである。そしてきわめて抽象的なものである。知能検査の中には、直観とか推理とか記憶とか再認とか、洞察とかいう分野が考えられているがそういってみても、それは能力の分野を言っただけで、具体的なものがどうかということをみているわけではない。つまりこれこれのことをたとえば記憶しているか、あるいは推理できるかどうかなどということをしらべるものではない。それを抽象的な素質としてみようとするのである。

アチーブメントテスト、ないし学力検査は、具体的にこれこれのことができるか、できないかをみることができるのである。一つの問題を出して、それができるかどうかを問題にすることができる。メンタルテストでも問題一つ一つについて、そうやればできるではないかということも考えられよう。しかしそれは意味がない。そういうことが意味があるような問題ではないのである。たとえば直観力の検査として一群の問題が出されている。机の脚が四本あるべきなのが一本かけている絵が出ている。人の指が六本ある絵が描いてある。そのまちがいを発見するといった問題が一〇も一五もある。そのまちがいを発見する早さを問題にして一〇個できたか、一五個かで、直観力という素質をみて相対的に人間に序列をつけるのである。この場合一つ一つができたとかできないとかにはあまり意味がない。それらの絵をみて判定の早さが全体として問題になるのである。つまり一定時間に全体としてどれだけできたかが問題になる性質のものである。一定時間内に判定した数が少ないこと、つまり判定の速度が遅いことが問題になる

のである。早いかおそいかを、数でみるのであって、一つ一つができたかできないかを問題にするのではない。そしてそれを直観力の素質というものにしぼってくるのである。だから知能検査では何個できたかという全体的な得点が重要な役目を果す。

アチーブメントテストでもそういう場合がないではない。しかし、本質的には、学力検査の場合には、その一つの問題ができたかできないかを問題にしていくことができる。たとえば漢字を書くという問題を出す。五〇のうち二五しかできないとか、三五できたとか、全体として得点がいくらということが問題にもなるけれども、それ以上に実はこの字が書けない、この字は書けたということが問題になりうる。否、実はアチーブメントをみるという点からみれば、そういうことがより大切なことである。どこまで到達したかをみるのである。これを得点にしてみるのは、比較をする必要上行われるのであろうが、その場合は相対的な位置づけを問題にしているのである。しかし、そういう相対的な位置づけをすることが教育活動としてどういう意味をもっているか。相対的な学力の位置づけは、もっとも普通には成績の評定として行われている。いちばん親しみのあるのは、通信簿への評定である。昔はこれは甲乙丙で表わされたが、今は五、四、三という五段階の評定になっている。これは余談であるが、父兄の間では五というのは昔の甲とか優に当ると誤解している人は非常に多い。甲と段階の五ではまるで意味がちがうが、要するに、それはいいことを表わす点だということで一致するのである。その意味では、父兄の誤解も大した誤解ではない。大きくみれば、大してちがわないとも言えるのである。共に得点という抽象的なものが土台になっているということである。

児童・生徒をそういうように位置づけることも全然意味がないことはない。あれはできる子どもとかいうことが全体としてわかっているということは、教育する場合役に立つという程度において価値はある。しかしそれ以上のものでない。

学習をさせるときは、具体的に、これこれのことがどれだけできるか、どれだけしかできないかがわかって、それではじめて具体的に学習の指導をすることができるのである。そこで児童・生徒の成就（アテーブメント）をみるのである。それは成就した結果、今もっている力というように考えれば学力という言葉を使ってもよい。ともかく教育活動の中に、アチーブメントをみるということは必然的になくてはならぬものである。そしてその場合の観点はできるかできないかということを評点で示すということが問題なのでなく、何ができ、何ができないかという具体的なことが問題なのである。ただ評点でできるとかできないとかいう相対的な位置づけがなされても、それは学習指導そのものには役だたない。しかし学習指導をする場合の一般的な心構えにはなるであろう。たとえば、この子どもは国語の成績が八〇点だとか、あるいは五段階にわけて五の所にいるとかいうことがわかっていれば、それだけの心構えができる。そこである具体的な教材があって、それを読ませるとき、教師はだいたいどのくらい読めるか見当をつけることができよう。しかしそれだけであって、具体的に指導をするということになれば、やはりその教材について、どこが読めて、どこが読めないかを明らかにしなくてはならない。それをテストして具体的に指導がきまってくるのである。この場合のテストというのはきわめて広い意味でいっているのである。調べるという意味である。

2　教育的テストの正しいあり方

以上アチーブメントのテストと知能検査とを比べながら、教育活動

の中に位置づけたのであるが、必ずしも現実に行われているアチーブメントテストが、そういうように教育活動の中に位置づけられているということではない。否むしろ反対の場合が多いのである。それはテストというものが教育活動として考えられていないということである。

テストは現在では一般に人間の評価と結びついて行われている。前に述べたように、成績がよいとかわるいとか、つまりできる子とかできない子とかいう人間のできに関して考えられている。何点とったか、さらにはその得点を基礎にして相対的比較をして人間を位置づけすることに関して使われる。きわめて抽象的な人間の相対評価の中に位置づけられている。テストというと頭の中にすぐピンとくるのはこれである。テストを受けるほうもそうであるし、するほうも結局はそういう意識がいちばん強い。これはアチーブメントをテストするという教育的な考え方からすると、全く誤った方向に考えられているといってさしつかえない。

この点でアチーブメントテストは、知能検査とは全く相反する考え方のものである。知能検査は、素質をみようとしているのであって、教育によって動かされないものを問題にしているのである。それは教育活動の条件として考えなければならぬかもしれないが、素質そのものをどうするということはできないという考え方のうえに立ってテストをしているのである。アチーブメントテストは、それと全く反対で、教育によって身につけさせようとしているものは、今どのくらい身についているか、どこが身についていないかといった、現在の到達点をみているのである。その到達点、アチーブメントはどしどし向上するという考え方でテストをしているのである。両者は同じくテストといっても、全く異なるといわなくてはならぬ。にもかかわらず、その本質が案外に理解されていない。混同されているのである。

それはアチーブメントテストが、全く便宜上から本質的には教育活動ではないような、たとえば入学試験などということに使われるからである。選抜して、できのよい子どもをとるなどというためのテストは、ほんとうはないほうがよいが、ともかく教育活動ではない。教育活動とは、できがよいならそれなりに教育し、悪ければ悪いなりに教育するので、それにはよいとは具体的にどうであり、悪いとは具体的にどうなのか、それを調べて、すなわちテストして、指導するのである。それが教育的なテストである。テストは本質的に考えれば、教育活動の中で立派な価値をもった一つの方法であるが、それを悪用しているものがある。つまり立派な武器も悪人が使えば、人間に害をなすようなもので、学校に入学せしめない者を選抜するというような一種の利己的な考え方から使われて、人々に誤った考え方をもたせてしまった。そういうことがなければ、もっとすなおに、自分は何ができて、何ができないかをしらべて勉強する問題点を発見するというように、皆の者が喜んでテストを受けるようになるはずであるが、事実はテストほどいやなものはない。馬鹿の程度を検査されているみたいで、劣等感を起させる以外に役だたない。これはテストの濫用の結果というべきであろう。

アチーブメントといってもさまざまな観点からとらえることができる。つまり学習の結果到達しているものをみる見方はいろいろあるのである。これこれのことを記憶しているかとみるのも一つの学習の結果に対するさぐりを入れる観点である。これこれこうと判断できるかというのも一つの視点である。これこれの技術的活動ができるかというのもまた一つの視点である。これらのことどれ一つとして、子どもが長い間に学習したことでないものはない。この場合の学習とは単に学校の教室の学習というのではない。遊びながら学習したこともあ

ろう。毎日の生活の中で学習などということを考えないでいろいろと行動をしている間に学習したこともあろう。ともかく子どもが現にあるところに到達しているのは学習の結果であると考えてよい。そこに子どもの現在の行動能力があるのである。あらゆる要素がよってもってその現在の能力をつくっている。それらを一つ一つ観点をたてて調べて行くことは、理論上はできるわけである。その観点は教育上の観点であることはいうまでもない。これこれの行為ができるか、というのはただ興味のうえから調べることもできるが、調べる必要があるのは、できなければできるようにする必要があると考えるからである。そこに教育的目標があるわけである。そういう教育のねらいがあって、それが視点となって、調べること、テストが行われ、そういうものが根底にないテストはこれは無意味なテストである。方法上ちゃんととのっていても意味がない、その意味でテストの価値をきめるのは、最後にテストの根底にある教育のねらいであるということができよう。

以上述べたことをこれまでの普通の概念でいえば、教育活動において意義を発揮するテストは、常に診断的な意味をもっているということである。一般に診断テスト、標準テストなどといわれて区別されているが、教育的意味においては、標準テストもまた診断の意味で行われ、診断の一つの方法として標準的なものを設けて、それによって診断するということである。標準を設けてそれの中への位置づけをするだけだというようなテストを行っても、それは教育的には何の意味もないのである。現在は標準テストとか診断テストの区別はテストの方法、手続の点からのみ考えられているが、目的的に考えれば、いずれも教育の問題としては診断的でなければならぬ。少なくとも教育活動の中で、教育者が考えれば、そういう考え方で使うべきものである。

その同じものを他の目的に使って、選抜試験に使ったり、人間の評価に使ったりしてもそれはかまわないし、現に行われているが、その場合は教育的意味で使用していないとさとるべきである。具合のわるいのは、その教育でない場合のテストも教育の場合のテストも形がおなじで異なるところがない。そうして教育的な意味のテストがだんだん影が薄くなるのである。悪貨が良貨を駆逐するようなものである。困ったことというべきである。

二　教育的テストの方法と問題点

1　教育的テストの現実形態

テストというのを今まで、きわめて漠然と調べるというほどの意味で考えてきたのである。アチーブメントを調べる、学力を調べるというように考えてきたのであるが、テストといえば一般には、紙の上に問題が書かれていて、それに答を出す種類のものがわれわれの頭の中にすぐ思い浮かぶのである。しかし本来テストというものはそういう紙を使ってなされなければならぬものでない。ただ一つの方式であるにすぎない。結果をはっきりと把握するためにそういう方式を使うのである。言葉だけを使って問答をすることによってテストを行うこともできるのである。しかし言葉は流れ去り、消え去ってしまうので、残るものとしては、テストをした人の主観的な印象だけになる。それでは都合の悪い場合もあるので、紙を使うわけであろう。そこには言葉が書かれているから、いつまでも〝資料〟として残るわけである。だれでもがそれを材料として解釈することができる、いわば多くの人がそれぞれの印象をもつことができる。それらの印象の共通なものが客観性をもつというように考えるのである。

しかしこれは紙でなくともできる。テープレコーダーのような便利なものができたから、そこへ問答を録音しておくこともできるのである。ある点では、つまり材料を残すという点では、録音のほうがより具体的な材料を残すこともできるのである。つまり問答をするほうが紙を使ってやるより、かゆい所に手のとどくやり方ができるということである。

しかし問答となると、紙でテストするようにいっせいに多くの人についてテストできない。テストする人が多くなればできないこともないが、一人一人によって問答のしかたが異なるであろう。それを全く同じにするということも考えられるが、そうなれば、問答という生きた個性をもった人間の話し合いという性格は多分に失われるから、紙のテストに近くなる。そこに限界があろう。現在は多数を教育し、多数をテストすることがどうしても要求されるので、そこに紙を使うテストが使われる。詳しくという要求を犠牲にしても、多数をしらべるということにならざるを得ない。こうしてテストといえば紙のテストがすぐ考えられるようになるまで紙のテストが普及したのである。

このように実際にテストが行われるときには、現実のさまざまな条件から制約されて、そこに現実の形が生れる。テストにはそれぞれの限界があるかというようなことが問題にされる。しかし一般にそういうことを言うのはあまり意味がないのである。とくに現実のテストの制約をいっているのか、そもそもテストということの限界を言っているのか、その辺が明らかでない。

いったい教育活動において、子どもの状態を調べるということは、絶対に必要なことである。必要というより、教育活動というものが教育する者とされる者との関係から具体的な形態をとってくるのである。調べるという意識がなくて、そんなことを全然問題にしないで教育活動が行われたとしても、結果として影響を与えることは、教育される当の人間が受けとりうる限界において受けとるにすぎない。これは教育の内容について考えても同様である。教師はある内容のことを理解させようとしても、その内容をどう受けとるかは常に児童・生徒の側のそのときの段階、あるいは状況によるわけである。この場合段階とか状況とかは一般的に言われる発達段階というようなことばかりでない。具体的にその内容を理解するに必要な現実の状態をいうのである。

だから、教育が能率的に行われ、また児童・生徒に即したものであるためには、実際に教育が行われるその具体のところで、児童・生徒がどのような状態であるかが調べられなければならない。

2　教育的テストの基本方式と問題点

テストをする場合には、何をテストするかがはっきりしていなくてはならない。つまりねらいである。ねらいはどうして立つかといえば、教えようとすることからきまってくる。これこれのことを教えようと思う。それがわかるかどうか、そこで調べてみるということになるのである。

こういうふうにしてテストするのは、テストを教育のレディネスを調べるのに使うという考え方でないかという人がいるであろう。確かにそうであるが、本質的に教育的テストはそういうものである。学習の結果をテストするという場合も、基本的にはこの考え方とかわらない。なぜなら結果をしらべて、それで終るのではない。結果はすぐつぎの学習につながるのである。結果が予想と反して、ある点に到達していなければ、学習をもう一度やり直すということにもなろう。結果が予想通りの点まで到達していれば、その上のつぎの段階の学習が考

えられるのである。結果だけ出して、それだけにとどまるというのは、教育活動の一環としてのテストとしては存在しえない。人間の位置づけをして成績の順位をつけるというのは前にも言った通り、教育的テストのなすべき仕事ではないからである。

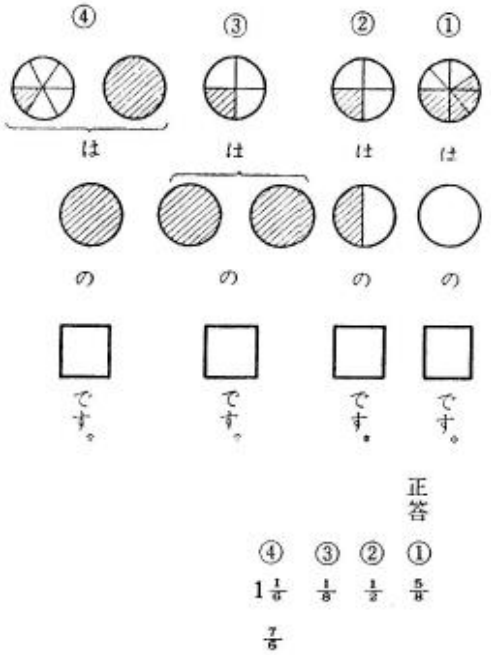
こう考えると、つまり教育的テストというのは、学習の目標と密接に関係がある。これこれのことができるようにしたいということが、これこれのことができるかどうか、そういう事態に児童・生徒がなっているかということをテストする必要を生み出させるのである。それによって、テストの内容がきまってくる。

つぎに具体的にそのことを考えてみる。

今、テストの問題を一つ例として、あげてみよう。

例題

つぎの図でななめの線をひいたところをくらべて、その割合を分数で□の中に書きいれなさい。



この問題には、一つのねらいがある。それは分数というものの意味を理解しているかどうかということである。すなわち、ある分数は基準となる量との関係においてあらわされているものだ、ということがわかっているかどうかをテストしようとしているのである。しかしこの一つの問題では、それがわからないかもしれない。そこで、つぎのようなものを出してみる。さらに第二の問題を出してみる。いずれも同じねらいをもった具体的な問題である。

さてここでテストしようとしていることは、分数の意味を理解しているかということである。それが具体的には、こういう図形の問題で出されているのであるが、この図形の問題で、はたして分数の意味がわかっているかどうか、を見ることができるだろうか。それは見ることができるとも言いうるし、できないとも言いうる。少なくとも最初の問題一つだけではそうであろう。〝これができないではほとんど基準ということがわかっていないとも考えられる〟しかしまたそれができなかったとしても斜線の部分を数えちがえたということもあるかもしれないのである。またできたとしても基準ということがはっきりわかっているとはいえない。それは常識的直観的にわかっている場合もある。こういう問題にぶつかることは多いのである。そこで、二つ、三つとたたみかけて聞いてみる。そうするとわかる。とくに、三つはいずれも基準となるものの量がちがっているのであるから、基準という考え方がどの問題を考える場合にもはっきり自覚されていることが必要である。だからかりに第一の問題ができても、第二、第二の問題ができない場合は、やはり基準の概念が明確でないのである。

こういうようにして、ねらっている分数の基礎的な概念である基準ということについて、理解しているかどうか、しらべるのである。これがわからなければ、そこで、それをわからせるような学習をさせな

ければならぬことになる。

このようにねらっているものは子どもの力である。その力を具体の問題に即してしらべるのである。問題ができたかどうかではない。できる力をもっていたかどうかをみているのである。

ところがこれは、一つのねらいをもった三つの問題をたたみかけて出して、テストしたわけである。こういうことは一人の対象をテストするときはかなりなところまで可能である。こういう種類の問題ならば、多数に対してもペーパーテストしうるが、一人との問答でなければならぬ問題もある。つぎの例題をみられたい。

例題

次の□の中にかいてある事がらと、とくに関係のふかいものを、下の1から8までの文の中から三つえらんで、その番号を○でかこみなさい。

江戸時代のおわりごろになると、手工業ではあるが、おおぜいの職人をつかって、分業によって製品を大量につくるような工場のしくみができてきた。

1　幕府は外国との取引や交通をやめさせてしまった。
2　城のまわりには、けらいの武士がすみ、そのまわりに町人や職人があつまってすむようになり、城下町が各地におこってきた。
3　大名の中には熱心に新しい産業をおこそうとしたものがあって、各地に工業がさかんになってきた。
4　各地でつくられた品物は、商人の手によって、日本国中に売りさばかれるようになり、商品の流通するはんいがひろがった。
5　農民は重い税をおさめねばならず、いろいろな天災などもあって、生活が非常に苦しくなってきた。
6　将軍吉宗は外国の学問をとりいれるようにした。
7　外国の解ぼうの本がほんやくされて、西洋の医学がさかんになってきた。
8　江戸や大阪には金持の商人があらわれ、武士や人々に金をかすようになった。

正答　3、4、8

これは大勢の対象に対してテストするためにつくられた問題である。答が出してあって、八つの中から三つを選択するようになっている。このねらいは、歴史事実を見る場合において、ある一つの社会的現象をとらえたときその現象の条件となっているものは何かということを考えて歴史的事実をとらえているかということをテストしようというのである。つまり歴史的事実を社会的因果関係の系列としてとらえているか、ということである。そしてその具体的な材料を江戸時代の末期の工業の発達の事実に関してとっている。

ところでこの問題について、三、四、八を選択して正しく答えられたとする。それは生徒がどの程度のことがわかっていることになるだろうか。非常におおざっぱに考えれば、江戸時代の後期に、マニファクチュアが発生したことについては、歴史的因果関係の系列としてとらえているといってもよいかもしれない。しかし、それはきわめておおざっぱな言い方であって、児童がどういう段階に到達しているかを十分につかんだことにはならない。つまりこの問題でテストできることはきわめておおざっぱなことだということである。もちろんこの問題ができたからといって、歴史事象の因果関係が理解できるなどという一般的なことは言えない。そういうことを明らかにするためには、もっとたくさん同じようなねらいをもった問題を出さなくてはならぬ。また、この江戸時代の後期のマニファクチュアについてのことだ

けに限ったとしても、この問題ができたというだけでは、どういう判断を下してよいかは厳密にいうとむずかしい。試みに四を選択した子どもに、どうして関係が深いかと反問してみるとよい。その答は随分まちまちで、決して正しいと思うものばかりは得られない。同様にして、三でも、八でも、そうである。また、一、二などをなぜ選択しなかったかを問うてみると、またその答がまちまちであって、このマニファクチュアの発達という歴史の事実について、ずいぶんさまざまな理解のしかたがなされていることがわかるのである。

このようなむずかしい問題になると、やはり一人一人の問答によって、児童がどういう理解をもっているかをテストしなければならないのである。大勢をいっせいにテストするようなペーパーテストでは、どうしても限界がある。

このように考えてみると、現実的に行いうる教育的テストによって明らかにすることはきわめて小さなことである。つまり、子どものアチーブメントといっても、大きく、国語の力がどうとか、算数の力がどうとかいうことは、とてもしらべられない。ごく小さなことである。たとえば、江戸時代のマニファクチュアの発生についても、実際に児童がどの程度の理解をもっているかを精密に調べようとしたら、相当な時間をかけて一人一人について問答をしてみなくてはなるまい。それをかりに一学級行うとしたら、たいへんなことになるのである。ましてや、生徒が歴史的因果関係について、理解をしているかというような一般的なことになったら、無数の問題が出されなくてはならない。それを一問で知るなどということは、全く非常識といわなくてはならぬが、そういう傾向が現在なきにしもあらずである。テストの限界を知らないのもはなはだしいというべきである。