

教育研究所報

佐賀県立教育研究所

4号

特集 教育と評価

……◇もくじ◇……

教育評価とテスト	(1)
東京教育大学教授 橋本 重治	
教育評価の機能	(2)
所長 須古 将宏	
能力診断のための知能検査	(2)
所員 久保山幾男	
評価と指導のための学力検査	(4)
所員 久保山幾男	
知能と学力の対比による学力の診断と評価	(5)
所員 木下 巧	
診断と評価のための基礎的な教育統計法	(6)
所員 向井正之・田中 照	
所員に托して	(3)
第八代所長 碓 敏雄	

教育評価とテスト

東京教育大学教授 橋本 重治



近年、わが国の教育に対する歯切れのよい批判として「テストあって教育なし」とか「テストあって授業なし」「授業あって教育なし」ということばがある。このような批判は、一刀のもとに今日の教育のあり方を切ったようで、胸がすく思いがしないでもないが、あまり切れ味がよすぎはしないだろうか。と言うのは、こういう批判をする人の意識の中には「教育」はりっぱでたいせつなものであるが、「テスト」はつまらないものであるとしてこれを軽べつしたり、あるいは悪者視したりするような考えがかくれてはいないだろうか。かりにこういうことばをはく人自身は、テストの意義と機能を正当に理解しているとしても、こういう批判的言説をきく人々に対しては、「テストは非教育的なものである」とか「テストは悪者である」といった考え方を、自然に植えつける危険があることは否めないであろう。

今日、わが国の小中学校教育が、またテストや試験が、このようなむごい批判をうけるにいたったその背景は、理解することができなくもない。それはおそらく、これまで毎年行なわれてきた文部省学力調査の影響と、激化する上級学校入学試験の影響に起因するであろう。

しかし、これはたぶん主なる理由ではないであろう。最も大きな理由は、長いわが国の教育の歴史において、テストや試験というものの本義が理解されなかったり、発見されなかったりしていて、一般にテストの機能に対する偶像や誤解が存在しているということにあるであろう。

一部のよく考えた人を除いて、わが国の大部分の教師や父兄が試験やテストの目的や役割や意義をどう考えているであろうか。筆者の観察するところでは、それは通信簿記載や成績報告のために必要であり、そのためのものであるというようにきわめて事務的に理解している人々が案外に多い。

またテスト観が日本人の意識の底に少なくとも明治以来ひそむところのテストに関する一種の邪道的方法観が、今日、きびきた風な教育評論家達の「テストあって教育なし」という名文句(?)の生まれる最大の根源ではなからうか。

一体、テストの最深の意味は何であろうか。教師の指導も生

徒の学習も、その他人間の営む一切の仕事は過去・現在・未来の時間の流れにそって営まれる。こうした事業の時間的経過において、現在の時点にたつて、過去の成敗等の実情を確認し、それで得た知識を将来にフィードバックしてより効果的な行動方針をたてることを可能にすることである。テストの本義は、現在の時点で過去と将来をつなぐことであり、過去の経験を将来に生かすことであり、フィードバックによる今後の行動の効果的な規制を可能にすることである。

こういっても、テストの意義と機能は、このようなフィードバックの機能ばかりではなく、テストの最大の機能は、生徒側からみた学習機能と教師側からみた指導機能との二つにあると考えてよいが、そのいずれにしても、根本は、前にも述べたように、英知的な営みであるところの——人間の仕事を連続的・発展的な時間的経過であるとして、その中で過去に関する知識を求めてこれを将来に生かすということである。こう考えると、決してテストは邪道でもなければコンソクとした裏参道的方法でもなく、また教育方法としてでも方法でもない。実に堂々たる表参道的方法であり、教育の本道であるといわなければならない。「テストあって授業なし」、「テストあって教育なし」ではなくて、むしろ正しくは、「テストなくして授業なし」とか、あるいは「テストなくして教育なし」というべきであろう。

しかしながら、こういったからとて、わたくしはなにも無条件で、テストの教育的価値を謳歌しているのではない。

テストが体質的に伴いやすい弊害の一つは、通常ペーパーテストであるために、これのみによる評価が観念的な測定となり勝ちである点である。第2は、テストが生徒に及ぼす心理的影響のあること、第3は、テスト目的に妥当な評価法の組合せがないと所期の目的を達することはできない。

こういうわけで、テストは原則として効果的教育方法の重要な一部であるといっても、その具体的方法については、いろいろとじゅうぶん研究してかかる必要があることはいうまでもないことである。

1 教育評価の機能

教育評価は「教育の目標がどのように達せられているか」の確認と、その改善を目的として、教育上の事象をテストや観察等によって数値的にとらえ、これらの測定数値を、教育目標に照らして、価値的に解釈するものである。

教育評価の本質的な機能としては、次のようなことがあげられる。

- ① 集団の構成にどのような特徴があるか。(平均・分散)
- ② 他とくらべてどのくらいか。(相対的位置づけ)
- ③ 何を成就し、何を残しているか。(質的)
- ④ どこまでわかり、どこがわからないか。(系統的)
- ⑤ 何を選択し、何を排除しているか。(個性的性向)
- ⑥ どこが長所で、どこが短所か。(内部差異)

さらに、以上の第1次の評価結果を組み合わせ

- ⑦ 対応する事象の相互間に関連があるか。(相関度)
- ⑧ 現在の能力は素質相応のものだろうか。(成就度)

これらは、選別的機能と診断的機能とに大別される。

選別的機能は、測定結果を包括的・全体的に概観することによって達せられるが、診断的機能には、測定結果を部分的・分析的に検討したり、相互の関係を多角的に考察することが必要である。

2 教育評価の実施目的

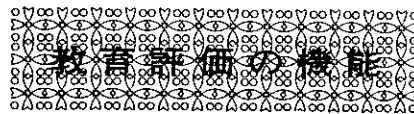
実施の具体的目的からは、次の4類型に分類される。

(1) 管理目的 教育の能率的管理を目的として、学校の教育条件整備状況の評定をすとか、優劣順位の判定を主とする入学者選抜試験などがこれに属する。

(2) 指導目的 これは、①教師の立場から、生徒の実態や変化(進歩・つまづき)を診断的にみて、直接指導に活用する場合と、②用いた指導法や教育課程が適切であったかどうかを自己評価する場合とがある。

①はあらかじめ計画された学習指導の一過程として行なうものであるが、②は指導計画の修正・改善をねらいとして行なわれる。

以上の評価は、通常単元学習の間に、自作テストによって行なうのが普通であるから、出題は指導の観点に絞り、簡単なテストを頻繁に行なって、指導効果の確認と、生徒の学習意欲高揚高揚をねらいとす



所長 須古将宏

べきである。

(3) 学習目的 指導目的とは表裏の関係にある。

指導目的のテスト結果を、児童、生徒の側から、自己評価や相互評価をさせ、学力の確認、学習方法の規制、学習意欲の喚起をねらわせる。

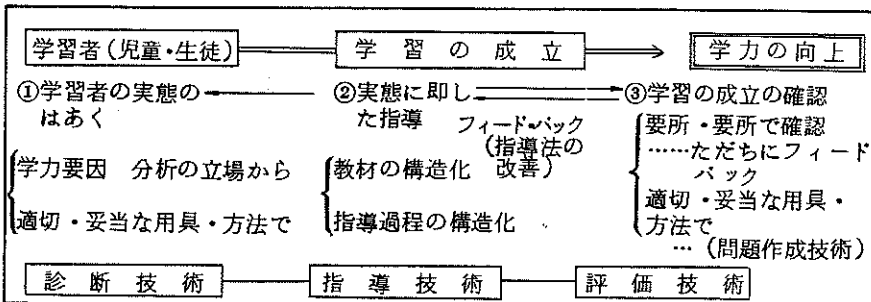
(4) 研究目的 研究機関・学者・研究者等により、実験・調査を主として実施される場合である。この場合は、結果の数値の信頼度を高めるため、高度な統計的処理を必要とすることが多い。

能力診断のための

知能検査

所員 久保山 幾男

学習指導の基本問題は、「学習の成立」を効率化することであろう。そのためには学習者の診断すなわち実態のはあく、その実態に即した指導、そして学習の成立の状態を確認し、指導法の改善へフィードバックするための評価等の有機的連係が大事なことであろう。……下表参照



1 診断としての知能検査

学力要因は、学習者の知能・性格・身体および健康度等の主体的要因と、社会・家庭・学校などの環境的要因とに大別できるであろう。そして知能検査は学習者個人について、学力の主体的要因としての知能を測定し、それによって学習の可能性を判断する資料を得るものといえよう。知能検査のえらび方・実施法・結果の処理・解釈等に関する知識・技術は、診断技術の重要な部門であろう。

2 知能検査のえらび方

知能検査を内容・適用範囲等から次のように分けることができる。

団体検査か 個別検査か	時間制限法か 作業制限法か	年齢別式か 通年式か
単式か (A式かB式か) 複式か (A・B式)	診断的検査か 概観的検査か	

普通学級で知能検査を行なうとすれば、当然団体検査(時間制限法)がえらばれるであろう。特殊学級等の場合も年齢

相当の団体検査を実施し、その結果と個別検査(主として作業制限法)の結果とを照らし合わせて解釈すべきであろう。

知能検査によって、知的能力を概観的にとらえるだけでよければ、概観的検査でよからうし、知能を分析的にとらえ学習指導に有効に生かそうとすれば、診断

的検査を使うのが妥当であろう。また知能をB式(非言語式)A式(言語式)いずれかの一面からとらえるか、A・B式両面からとらえるかによって、単式・複式に分かれるが、学習者の特性の診断・理解のためには、A式・B式の両面からとらえることが妥当であろう。……(知能A・Bと学力の相関については、知能検査の手引を参照されたい。)

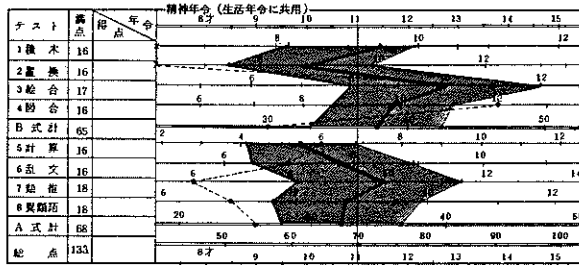
年齢別式か、通年式かについては、特別の研究目的以外は年齢(学年)別式が妥当であろう。

○小学校 5 年 A 組

S-43 6.12~14

児童 番号	知能(教研小5)										国語(教研小4G)					算数(教研小4G)					国語 SS	算数 SS	平均 SS	知能 SS		
	1. 積	2. 置換	3. 総合	4. 図	5. B	6. 式	7. 算	8. 文	9. 乱	10. 類	11. 果	12. A	13. 総	1. 読	2. 話	3. 読	4. 作	5. 書	6. 総	1. 関					2. 考	3. 理
1	13	12	12	14	51	8	9	14	11	42	93	12	10	23	16	14	75	9	11	13	11	44	58	50	54	62
2	10	9	5	6	30	6	10	6	8	30	60	7	3	12	13	14	49	9	5	15	6	35	40	45	43	45
3	11	8	12	13	44	8	9	10	8	35	79	13	9	21	17	8	68	14	9	24	13	60	53	59	56	55
4	3	4	11	12	30	6	7	6	7	26	56	7	5	11	6	12	41	7	2	8	8	25	35	39	37	43
5	10	11	14	11	46	8	9	12	10	39	85	13	9	21	16	11	70	12	10	20	12	54	55	56	56	59
6	9	8	2	7	26	2	3	6	2	13	39	8	7	8	15	8	46	8	6	13	8	35	38	45	42	35
34	10	10	12	14	46	5	8	6	8	27	73	10	12	19	18	14	73	11	9	15	10	45	57	50	54	55
35	9	10	9	7	35	7	9	11	9	36	71	12	10	17	15	12	66	10	1	11	7	28	52	41	47	54
M	9.4	8.7	10.0	9.8	37.9	5.5	7.7	9.9	8.7	31.9	68.4	10.3	8.5	16.0	13.3	11.8	68.4	9.4	7.9	13.9	9.4	47.5	48.0	47.7	51.7	
SD	2.3	2.7	4.0	2.6	8.5	2.6	2.5	3.3	2.2	8.4	10.3	2.1	3.1	5.2	4.0	3.9	10.3	2.8	4.8	6.3	3.8	9.6	9.3	8.4	7.9	

●知能検査診断プロフィール●



3 知能検査実施上の留意点

知能検査は条件観察法である。したがって標準化されたときの条件どおりに実施することが大事である。たとえば時間制限法の検査では、実施時間の厳守は不可決の条件である。また被験者が解答要領について、じゅうぶん理解していることも前提条件の一つである。実施時間の厳守にとられるあまり、講堂等での学年一斉実施とか、拡声装置を通しての一斉指示などは適切な方法とはいえない。原則として担任が学級別に正確な計時のもとで実施すべきである。

4 結果の処理およびその利用

(1)採点が終わったら、各テストごとの得点を一覧表にあげる。もし各テストごとの得点に0点があれば、解答要領につ

いての理解が不じゅうぶんか、いちじるしく知能が劣る場合であろう。

- (2)知能偏差値平均および標準偏差の計算
知能概観を全国標準と比較し、学級の特性を知る。
- (3)各テストごとに、平均・標準偏差を求めプロフィール作成。
学級の特性を知能構造の上から分析・考察し、これを基準として個人の知能特性を解釈・診断する。

5 処理および利用の例

下表はA小学校で、教研式知能検査(年齢別・複式(A・B)・診断的)5年用と、同時実施した教研式G型式4年用(一学期まで一学年くり下げ)国語・算数学力検査の結果である。

この学級の知能概観は、 $M=51.7$ $SD=7.9$ で平均は全国標準よりやや高く、ちらばりはやや小さくまとまっている。これをさらに診断的・分析的に考察するため、 T_1 積木から T_8 異類語まで、各テストのM・SDを求めプロフィールをかいてみる。ぬりつぶした部分は $M \pm 0.5 SD$ すなわち5段階評定の3の段階(約38%)の範囲を示している。これからみて T_3 総合(空間関係はあく力)にはすぐれているが、 T_2 置換・ T_6 計算・ T_6 乱文・ T_8 異類語等に含まれる思考・推理の能力に劣ると思われる。図中○...○は、4番の児童の知能診断プロフィールで、 T_1 ・ T_2 ・ T_7 ・ T_8 からみて、感覚の速さ・推理する能力を訓練しなければならないと思われる。



所員に托して

第八代所長 碓敏雄

昭和34年9月から、昭和39年3月までの在任は、すこぶる永いものであった。今までもそうであるが、今後ととも、所長として、こんなに長く、その任につく人はいないであろう。

それだけに、全国教育研究所連盟や九州教育研究所連盟の役員や、全国教育研究所協会の理事などをも勤めさせられて広範囲な動きの中に立たせられた。

ところで、研究所充実の第一は、所員の頭脳にかかっている。その点、私は、すばらしい研究員に恵まれて、所の充実が期せられたのだった。橋本、西山、宗、百武、山中、須古の各所員は、実にすぐれた頭脳をもって、研究成果をあげてくれた。(金ヶ江、畑瀬の事務的援助も大きかった。)そうしてまた所員らは、全国や九州の大会において、その稔りを流布して来たが、私は、このことを誇りとして過したのだった。

研究テーマとしては、学力についての要因の研究など、学力向上に関するもの、勤労青年教育に関するもの、不適応児の指導に関するものなど、今も記憶に残っている。在任末期

に、国際数学教育調査協議会が出来て、国際的な研究にも及んで来たが、どんな結果になったことだろうと、時に、思いだしたりする。

思い出すことと言えば、35年10月、九州地区教育研究所連盟の総会並に研究発表会を、嬉野で開催したことである。少数の所員で、よくもやりぬいてくれたと、感激であった。また、国立教育研究所の十周年記念式、全国教育研究所連盟の十五周年記念式なども、思い出深い。

教育研究所の充実は、教育を進める鍵であるが、それほど大事にされなかったように思う。定員の不足も、予算の不足も、所員の全くの努力によっておぎなわれて来た。それでも33年の予算409千円が、39年に1,854千円となったのは、所員の努力への手向であった。とにかく、優秀な所員の一人一人が今も心の中に生きている。

(現在筆者は鳥栖に住まい、
日本歌人クラブ委員であられる。)

評価と指導のための

学 力 検 査

所員 久保山 幾男

1 学力の分析……能力概念によって次のように分析できる。

(1)主として知的な学力

- 理解…事象間の関係はあくの成立ということが理解の最も重要な面である。
- 知識…理解されているとともに、的確に記憶され再生され得るといふ特質をもっている。
- 思考…問題場面において、その場面を分析し、解決のための既得知識・技能を動員し、解決のアイデアを生み出す。

(2)主として技能的学力

- 技能…理解・知識の面と、身についた実技としての行動化され定型化された能力
- 作品・表現…理解・知識・思考・技能・態度・鑑賞等の結集物または具体物。

(3)主として態度的学力

- 態度…ものの見方・考え方・行ない方の一貫的傾向性。
- 興味(関心)…態度より感情的色彩が強い。
- 習慣…具体的行動に関して、実行の意志が自動的に発動するよう傾向化されたもの。
- 鑑賞…芸術的価値等の評価や選択に関する傾向性。

2 診断・評価用具としての学力検査

(1)学力検査の種類……教師作成テスト・ワークブック・標準学力検査に大別できる。

○標準学力検査…問題内容・実施方法・採点方法・解釈方法について、あらかじめ定められた集団に実施し、その結果に基づいて基準が設けられた検査である。それは個人あるいは学級の特性を測定し、その特性(個性)に応じた指導をするためのもので、指導目的としての利用が望ましい。

(2)作問法の種類とその用い方……望ましい

問題作成のためには、テスト目標についての理解・テスト諸技術の知識理解・テスト目標とテスト技術の妥当な組み合わせと、実施時期の適正とが大事である。

○論文体テスト法…「××について説明せよ」

事態の説明・要約や概括・比較などの能力や、態度・価値観等の高度な学力をみるのに最適。採点の客観性が劣る。

○客観テスト法

- 再生形式…単純再生法…1 mは何cmか
- 完成法… $C + O_2 \rightarrow ()$
- 訂正法…誤りのところをなおせ。〔小G・4算〕
- 序列法…年代順にならべかえよ。

断片的事実や簡単な学習内容の知識を明確にとらえ得るが、連用すれば断片的事実の学習にのみおちいる。

再認形式

- 真偽法…鯨は(卵生・胎生)である
- 多肢選択法…正しい答を次から選べ
- 組み合わせ法…関係あるものを線で結べ
- 選択完成法…次の□の答を下から選べ

再生形式とほぼ同じ学力内容を測り得る。採点の客観性は最もまさるが、連用すれば断片的学習におちいる。

○問題場面テスト法

新しい問題場面を提示し、既得知識・技能等を動員し

問題の分析や総合・推理・判断を加えて、新しい解決のアイデアを生み出させるような思考力をみるのに最適。

○その他

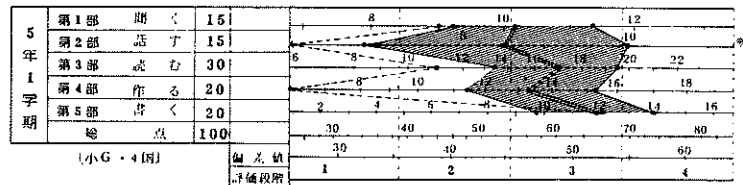
質問紙法・チェックリスト法・グスフーテスト法・レポート形式などの作問技術がある。

(3)テスト目標と技術の妥当な組み合わせ

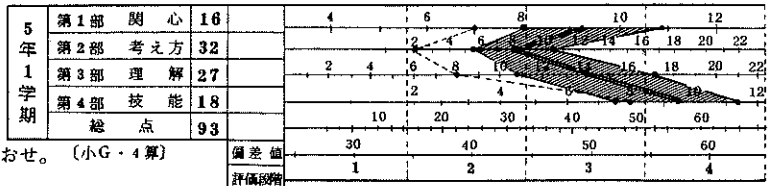
評価の目標		妥当する主要な技術
理 解		論文体テスト、客観的テスト(特に選択法、組合せ法、選択組み合わせ法)
知 識		単純再生法、選択法、組合せ法、選択組み合わせ法、真偽法、序列法、訂正法等の各種客観テストならびに論文体テスト
技 能	読み、書き 計算	各種客観テスト
	会話、討論などの 社会的技能	チェック・リスト、グスフー・テスト
能 力	図表、図解、機 械器具等の使 用技能	チェックリスト、各種客観テスト
	描画、描画、工 作、裁縫、実験 運動の技能	チェック・リスト、各種客観テスト
思 考 力		問題場面テスト(論文体と客観的との)
図画、工作などの作品 音楽、体操などの表現		チェック・リスト、評定尺度
態 度・ 興 味	価値観、意見	質問紙法、論文体テスト
	興味	質問紙法、グスフー・テスト
習 慣・ 興 味	学習習慣、健 康習慣	チェック・リスト、質問紙法、グスフー・テスト
	鑑 賞	チェック・リスト、評定法、質問紙法、真偽法

3 標準学力検査の利用の例

前記小学校の国語・算数の結果およびそのプロフィールから考察してみる。



国語… $M=47.5$ $SD=9.6$ で全国標準に比べてやや劣る。とくに「聞く・話す」の能力が劣り、かつちらばりも大きい。指導の重点をここにおくべきであろう。4番の児童(○……○)は、話す・作るの能力が劣るが、書く力はすぐれる。



算数… $M=48.0$ $SD=9.3$ で全国標準よりやや劣る。プロフィールからみて、考え方の能力は劣るが、技能の面はすぐれている。4番の児童は学級の傾向とほぼ一致しているが、技能の面はすぐれている。

学級の傾向として、知的能力・国語算数の学力とともに、推理・思考の面が劣っている。このことは今までの学習指導が技術の面に偏っていたのではないと思われる。科学的な学力分析にたつた学習指導法の改善が望まれる。学習指導にあたって、教材および指導過程の構造化とともに、指導効果の確認のしかたの研究・改善がきわめて大事なことであろう。

知能と学力の対比による

学力の診断と評価

所員 木下 巧

民主主義の信条は、個性の伸長にあり個性の尊重にある。個人内評価の立場は、学習は個人々の可能性、能力の発現を期待し、現実にごとまで発揮しえたかによって評価しようとするものである。従来、個人内評価の診断尺度として成就値 (AS) が広く用いられている。

$$\text{成就値 (AS)} = \text{学力偏差値} - \text{知能偏差値} \dots\dots\dots (1)$$

しかし成就値は一般的に知能偏差値の高いものほど低く、逆に知能偏差値が低いものほど高くなるという不合理な傾向が見られる。

新成就値の算出のしかた

そこで、現実的観点から同一知能偏差値をもつものの学力偏差値の平均値をもって、その知能偏差値に対応する学力偏差値の規準と規定するのが妥当である。これが新成就値 (RAS) の考え方、

$$\text{新成就値 (RAS)} = \text{学力偏差値} - \text{知能偏差値}$$

$$\text{対応学力偏差値} \dots\dots\dots (2)$$

の式で表わされる。

そこで規準となる知能対応の学力偏差値を算出すると、

$$Y = 0.7x + 15 \dots\dots\dots (3)$$

Y → 知能対応学力偏差値

0.7 → 知能、学力検査の相関係数

(相関係数は各学年、教科によって異なるが、大体、国社数理英では0.6~0.9の範囲にある。しかし測定誤差或いは、標準誤差などから0.7と規定してよい。)

x → 知能偏差値

(この場合、両検査とも同一生徒によって標準化される標準検査で全国母集団の偏差値平均はともに50、標準偏差も共に10である。)

(3)式の x に生徒個々の知能偏差値を代入するとその生徒の知能対応学力偏差値が算出され、更に(2)式に代入すると新成就値が導き出される。例えば表の新成就値の欄の如くである。

(3)式の回帰方程式から個々の生徒の規準となる学力推定値を算出することは煩わしいので、座標を利用した方が簡便でよい。

相関座標の利用のしかた

市販の方眼紙を購入し、右図の如く Y 軸に学力偏差値を、X 軸に知能偏差値を Y 軸と等間隔に目盛る。この場合、偏差値20から80までを記入し、偏差値1の目盛りは5mmぐらいが適当である。偏差値の目盛りが終ると、学力および知能偏差値のそれぞれを評価段階に区切り人数を記入できるようにする。次に知能対応学力偏差値の直線 $Y = 0.7x + 15$ を赤色で太く記入する。この直線は、学力、知能偏差値50の交点および、知能偏差値80と学力偏差値71の交点を結んだ延長線となる。これで一応作図が終了したので、個々の生徒の知能偏差値と学力偏差値の交点をプロットし、生徒番号を付していく。その結果、この赤線上にプロットされた生徒は能力相応の学力、赤線より上は能力以上の学力、赤線より下の生徒は学力が能力に及ばないものであることが一見できる。

しかしながら、新成就値の標準偏差は7であるので、規準学力水準(赤太線)より $\pm 1 \text{ SD} (+7 \text{ と } -7)$ の間にある生徒は普通に見られるもので特に注目する成績とは云えない。その

外に出る生徒が特異な存在として指摘されねばならない。そこで、赤太線と平行に $+7$ (知能偏差値50、学力偏差値57の交点と、知能偏差値80、学力偏差値78の交点を結んだ延長線) と -7 (知能偏差値50、学力偏差値43の交点と知能偏差値80、学力偏差値64の交点を結んだ延長線) に赤点線を引く。

そうすると上下の赤点線に囲まれる(例B君)生徒は普通学力児である。上の赤点線より上にある生徒(例C君)は特に学力が能力より優れている所謂学力過進児で注意を要する。又、下の赤点線より下の生徒(例A君)は特に学力が能力に及ばない生徒で、学力不振児と称し、特に注意が必要である。

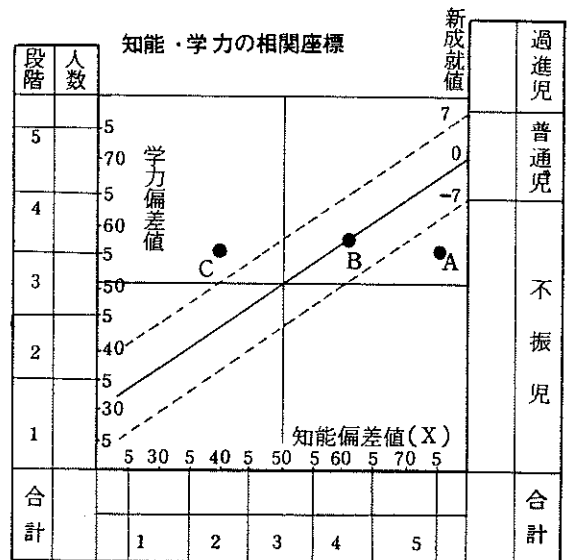
自作テストと知能を対比するとき

標準化されていない教師自作テストの結果と知能検査結果とを対比させても、そのままでは新成就値は算出されない。面倒でも自作テストの各得点と、知能偏差値をそれぞれ級を母集団とする偏差値に換算(本紙 6.7頁参照)してから前述の方法で診断しなければならない。

また、標準化されている学力検査でも、学力検査の標準化と、知能検査の標準化がそれぞれ異なる生徒を対象にしてある場合は、(手引きを見れば解る)この手続きをしたがよいであろう。

全国標準検査の測定値から得られる個々の生徒の新成就値は、全国の同一学年における個々の生徒の学力を診断し、自作テスト等の場合はそれを偏差値換算した集団(級)における学力を診断している。(本紙 2~4頁参照)

知能・学力の相関座標



(事例)	生徒名	知能SS	学力偏差値(SS)					知能対立SS	新成就値	
			国語	社会	数学	理科	英語			平均
	A	75	60	58	53	52	52	55	68	-13
	B	60	57	61	54	55	58	57	57	0
	C	40	58	60	49	53	55	55	43	+12

おわりに

このように個々の生徒の学力を診断すると、どうして良く出来たのか、悪いのか原因を認識せねばならない。つまり教師自身の教授法の反省とともに、生徒の心理・身体的要因、社会的要因、家庭・学校等の要因を多面的にとらえて適切な指導がなされねばならない。しかる後に上の方法で評価されるべきであろう。

診断と評価のための

基礎的な



所員 向井 正之
田中 照

1 度数分布表の作り方

テストを実施して、その成績(粗点)を名簿に記入しただけでは、個々の生徒の粗点は知ることができても、学級全体の特性をはあくすることは困難である。個々の生徒を理解するために学級全体の特性を知る必要がある。またAテストとBテストの共通な尺度を作る上からも度数分布表を作ることが必要である。第1表は50名の生徒のテストの結果を度数分布表にまとめたものである。

度数分布表の作り方は

(1) 級間(階級)の数をきめる

一般に級間の数は多過ぎては全体の分布の状態が見にくくなるし、逆に少な過ぎては精密さを欠くことになるので、10前後から20ぐらいまでがよいとされている。40~50名程度の学級で正規分布型を予想するような場合は、7~10ぐらいまでが適当と思われる。第1表は8階級にしている。

(2) 級間の幅をきめる

最低点および最高点を見つけ、全得点のひろがりを級間の数でわると、級間の幅が決まる。級間の幅としては2・3・5・10・20等が普通用いられる。人数が少ないときは級間の幅を大きく、人数が多いときは級間の幅を比較的小さくした方がよい。またその級間の中央値(中心点)が整数で示されることが望ましい。第1表の級間は10点である。

度数分布表は、測定値の中心傾向と得点の分布状況をみることが容易にしてくれるが、さらに、これが統計的操作の基本となるものである。

2 代表値の算出

度数分布表の中心的傾向を単一の数値であらわして、それでその分布全体を代表させる。こういう数値を代表値という。代表値として最も多く用いられるのは算術平均と中間数である。算術平均は、各測定値(得点)の総和を全体の

第2表

級間	度数f	偏差x	fx
100	1	4	4
90~99	3	3	9
80~89	7	2	14
70~79	10	1	10
60~69(A)	12	0	0
50~59	9	-1	-9
40~49	5	-2	-10
30~39	3	-3	-9
	N=50		fx=9

第1表 度数分布表

級間	度数(f)
100	1
90~99	3
80~89	7
70~79	10
60~69	12
50~59	9
40~49	5
30~39	3
	N=50

度数で除したものである。いま各人の得点が $X_1 \cdot X_2 \dots X_i$ であるとすると、その平均Mはつぎの式で示される。

$$M = \frac{1}{N} \sum X$$

(Nは総度数、 $\sum X$ は x_1, x_2, \dots, x_i の総和)

度数分布表から仮平均を用いて平均を求める方法

得点が度数分布表に整理されてい

る場合には、個々の得点が不明でもつぎのような計算式に従って平均を概算することができる。

$$M = A + \frac{\sum fx}{N} \times h$$

(A...仮平均、f...度数、
x...級間の仮平均からの順位、
h...級間の幅)

第2表で、仮平均Aを60~69の中心点64.5とする。偏差xは仮平均の級間を0として大きい方に1・2・3...小さい方に-1 -2 -3...を記入する。度数fと偏差xとの積fxを求め、その総和 $\sum fx=9$ を求める。級間の幅 $h=10$ だから $M=64.5 + \frac{9}{50} \times 10 = 64.5 + 1.8 = 66.3$ となる。

3 散布度

第3表は社会と数学の度数分布表である。平均点はいずれも70点であるが、社会は平均点近くに得点がそろっていて、数学のほうは得点が大きくちらばっていて個人差が大きいことがわかる。このような得点のちらばり度合い(散布度)をあらわすものに、いちばんよく使われるのは標準偏差(SD)である。

標準偏差とは、個人の得点と平均との差の自乗の総和の平均の平方根をいう。

標準偏差はつぎの式であらわされる。

$$SD = h \sqrt{\frac{\sum fx^2}{N} - \left(\frac{\sum fx}{N}\right)^2}$$

$$= \frac{h}{N} \sqrt{N \sum fx^2 - (\sum fx)^2}$$

f...度数、N...総度数、x...偏差、h...級間の幅、 \sum ...総和の意味

第4表は数学の得点分布表(第3表)から標準偏差の計算法を表わしたものである。N=50、 $\sum fx=0$ 、 $\sum fx^2=298$ となるから公式にあてはめて

$$SD = 5 \times \sqrt{\frac{298}{50} - \left(\frac{0}{50}\right)^2} = 5 \times \sqrt{5.96} = 5 \times 2.441 = 12.2$$

同じようにして社会の標準偏差を求めると、SD=5.3となり、ちらばり度合いが大きいほど、標準偏差は大きくなる。

したがって標準偏差が小さくなるにつれて、平均点近くに得点がそろっていることになる。また標準偏差を共通尺度としてAテストとBテストをくらべることができる。

第3表

級間	中心点	社会人数	数学人数
93~98	95		2
88~92	90		2
83~87	85		4
78~82	80	4	6
73~77	75	12	7
68~72	70	18	8
63~67	65	12	7
58~62	60	4	6
53~57	55		4
48~52	50		2
43~47	45		2
		平均70	平均70

第4表

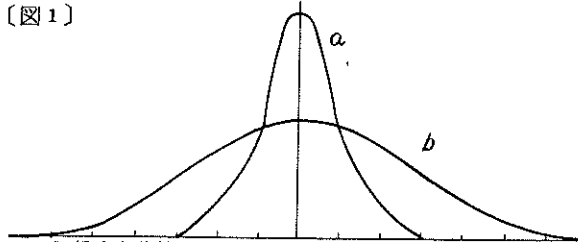
得点	度数f	偏差x	fx	fx ²
95	2	5	10	50
90	2	4	8	32
85	4	3	12	36
80	6	2	12	24
75	7	1	7	7
70	8	0	0	0
65	7	-1	-7	7
60	6	-2	-12	24
55	4	-3	-12	36
50	2	-4	-8	32
45	2	-5	-10	50
	N=50		fx=0	fx ² =298

4 正規分布について

(1) 正規分布曲線

多数の生徒について、身長や体重などの測定結果や、学力テスト等の得点の結果は多くの場合、中央付近に測定値があつまり、中央の平均から左右にはなれるにしたがって、度数が次第に減少して、分布の形が左右対称な山形をなすものであり、この曲線を正規分布曲線とよび、このような分布のしかたを正規分布という。

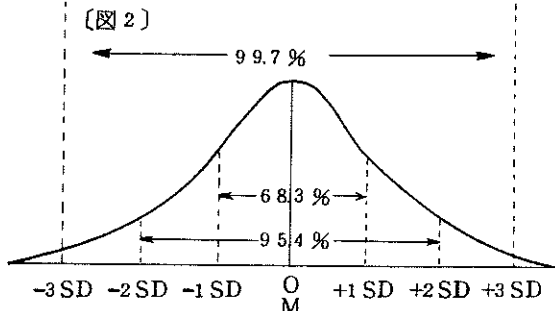
〔図 1〕



(2) 正規分布曲線の性質

イ. 標準偏差が大きいと分布の山は低くて横に広がり、標準偏差が小さいと、山は高く横の広がりは少ない〔図 1〕は a b の順に SD が大きい場合を示したものである。このように正規分布曲線はただ一つの固定形のものでなく、標準偏差の大きさにより形もかわるものである。

ロ. 標準偏差 (SD) をもとにして、正規分布の性質を考えると、〔図 2〕の正規分布曲線の基底上に平均 (M) を



基点として、右に +1 SD、+2 SD……、左に -1 SD、-2 SD……の点を取り、その上に垂線をたてると、その垂線と基底線と曲線とで囲まれる面積は、数学的計算の結果、M (平均) ± 1 SD までは、全測定値の 68.3% を占める。M (平均) ± 2 SD までは全測定値の 95.4% を占める。M ± 3 SD までは全測定値の 99.7% を占める。すなわち、測定値のほとんどすべてが、M ± 3 SD の範囲内にあると見なしてよい。

5 標準得点とその算出のしかた

ある生徒が国語のテストで 70 点を取り、数学のテストで 65 点であった場合、これらのテストをそのまま比較して優劣をいうことはできない。それは問題に難易があり、またそのクラスの得点のちらばりぐあいにより個々得点の価値もちがってくるからである。しかしその成績が、クラス全体でどんな位置を占めるかという、相対的な位置が示されるならばその比較も可能である。相対的位置の決め方にはいろいろあるが、ここでは標準得点 (Z 得点) について述べる。標準得点の算出の式は次のとおりである。

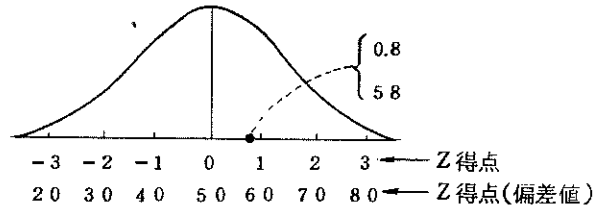
$$\text{標準得点}(Z) = \frac{X - M}{SD} \quad (\text{この式の場合、Z は小文字であらわす。})$$

(X……粗点(ある得点) M……平均

SD……標準偏差)

標準得点は、ある得点と平均との差(偏差)を、標準偏差

〔図 3〕



(SD) を単位として示したもの (SD で割った値) である。いいかえると、粗点が平均から SD の何個分 (何倍) はなれているかということである。これが偏差値の基本型で、Z 得点 (標準得点) といわれるものである。この Z 得点では平均が 0 で、標準偏差が 1 である。〔図 3〕

〔表 5〕

科目	得点	平均	S. D.	Z 得点	Z 得点 (偏差値)
国語	70	75	5	-1.0	40
数学	65	58	8.6	0.8	58

では、上の式により (表 5) の国語の成績を Z 得点になおすと次のようになる。

$$\text{国語 } Z = \frac{70 - 75}{5} = -1.0$$

数学も同じようにして 0.8 となる。〔図 3〕

しかしこのままではマイナスの符号や小数点がついて、不便であるので、1 SD を 10 等分し、さらに平均を 50 におきかえる。すなわち平均が 50 で SD が 10 の分布尺度である。〔図 3〕参照。これも標準得点とよび、Z 得点 (大文字の Z) または偏差値 (SS) ともいい、算出は次の式による。

$$Z (SS) = \frac{SD}{10} + 50$$

$$\text{変形して } Z (SS) = \frac{10(X - M)}{SD} + 50$$

〔表 1〕の数学の得点を上記□内の式により算出すると

$$\text{数学}(SS) = \frac{10(65 - 58)}{8.6} + 50 = 58$$

同じ方法で国語は 40 となり、すべて正の数で表現される。ここにおいて両科目の成績を比較することが可能であり、粗点では国語 70、数学 65 であるが、SS になおすと、数学が成績がよいことになる。このように平均点と標準偏差を用いて個々生徒の相対的位置を示すのが偏差値である。

すなわち表 5 の得点は間隔尺度による測定であるが、Z 得点 (偏差値) に換算することによって比率尺度となり、これによって国語と数学は共通な尺度で比較できるわけである。

与えられた得点を Z 得点に換算しても、粗点の分布が正規でないかぎり Z 得点も正規に分布しない。そこで、本来分布が正規であることが分っているが、測定法の都合で歪んで分布することが仮定されれば、粗点をパーセントイル法による T 得点に換算すればよい。

県立教育研究所主催で

初の教育評価研修会開催

＝ 4会場414名が集う ＝

8月16日から29日までの間の8日間、佐賀・唐津および武雄の4会場で初の研修会を催し、小・中・高校から414名の先生が参加された。

教育現場における教育研究をより科学的に進めるために、教育評価の知識と技術の習得をめざして研修会をもったのであるが、各会場とも教育熱にあふれる先生方に、担当所員も圧倒されるほどであった。

教育評価 研修会に参加して

佐賀市中学校 教員

研究所の皆さんが、情熱を傾けてご指導くださったことに対して感謝している。私の学校など教育評価の問題を特にとりあげねばならぬ段階に来ていると考える。

教育評価は教材、指導方法、指導過程、児童生徒、環境条件等さまざまな問題がからまるので、具体的な研修内容、方法になると誰れが立案しても困難な問題がふくそうすると考えられる。理想的にいえば、アンケート等適当な方法によって受講者の問題をとらえ、それを基底にしてそちらで計画、立案されることが望ましい。

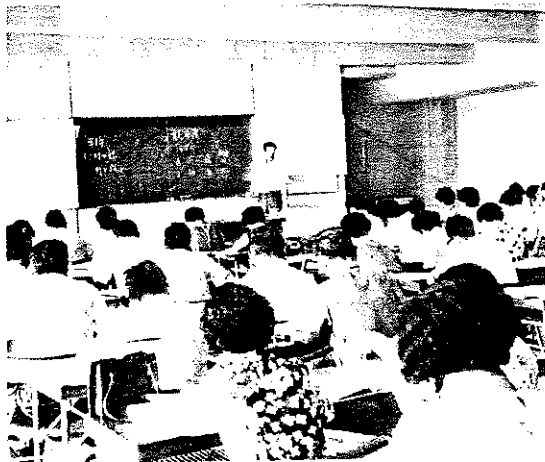
でも現場の実情等考えると、そこまで深刻に評価をみつめて指導に当たっている研究者はきわめて少ないように思われるので、今回のように研究所で計画立案され、継続的にやられることを望む。なお今回の講習で、実習を加味されたことは大賛成である。

教育研究には理論と実践が併行しなければ無意味である。受講者にとっては難解な点もなくはなかった。数的な処理、特にシグマなんて出てくると数に弱者等はめんくらう。したがって、あんなところは生徒に指導するような気持で、わかりましたか、わかりますかを連発してなっとくづくで指導

してほしい、今回の受講者はやる気があって参加している者ばかりであったと思う。

教師は、実質のあるこんな講習で鍛えられることを大いに希望するものである。なお、日程などにも若干問題があったように思う。

教育評価研修会の風景 (佐賀市民会館で)



教育評価主要参考図書案内

- 教育評価の技術 橋本 重治 図書文化社
 - 推計学による 新教育統計法 岩原新九郎 日本文化科学社
- (教師必読の基礎的な参考書として推せんします。)

所内人事

金子一司前所長の逝去にともない、去る6月1日付で、当所次長須古将宏の所長昇任が発令されました。



今回は、研究所創立以来初の研修会を開催したのを機会に、その備忘録の意

味も兼ねて、教育評価の特集号にしました。

巻頭には、教育評価の権威である東京教育大学教授橋本重治先生の玉稿を掲載できたことを感謝しています。

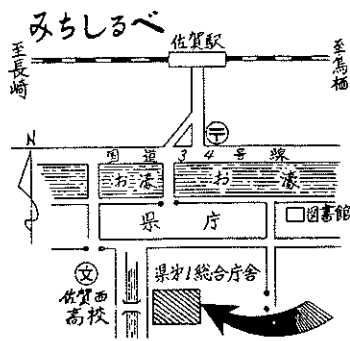
所員による執筆内容は、研修会当日の担当者が、それぞれの説明内容を要約したものです。字数の制限で、わかりにくい点もありますが、一般の方々にはどんな内容であった

かを知っていただくため、また出席された方々には記憶の整理のために、役立ててほしいと思っています。

「研修会に参加して」は、参加者全員にお願いしたアンケートの中から、掲載させていただきました。

毎回掲載してきた「ところどころ」、「内外通信」、「教育豆辞典」、「教育資料案内」は、特集記事の関係でいずれも割愛しましたが、次回からはまた復活させるつもりです。

研究指定校に限らず、自主研究の場合でも、発表会をされるときはご連絡ください。紙上に「研究発表会」だより、として概要を紹介させていただきます。



第 4 号
 発行年月日 昭和43年 9月 1日
 編集・発行 佐賀県立教育研究所
 佐賀市内1丁目6-5
 TEL ④2111内線437
 印刷 K K みづぼ印刷