

Gilbert De Landsheere

**Elementi di docimologia**

Valutazione continua ed esami

**La Nuova Italia Editrice**

Il caso in cui ci si crede infallibili nei giudizi non è infrequente. L'occhio clinico, come il buon senso, sembra una merce di cui nessuno lamenta la mancanza. Nel campo della scuola: sia in occasione di un esame risolutivo, istantaneo, condotto a tantum e con esito inappellabile, sia, con maggiori giustificazioni, nel caso di scolari seguiti per intere annate scolastiche, magari per un intero corso di studi. Poi saltano fuori i 'freddi' metodi statistici, e si scopre che per un'ovvia legge psicologica, della quale però tendiamo a non prendere coscienza, siamo portati — a parità di condizioni — ad abbassare il voto di uno scolaro ritenuto mediocre e a tenere alto quello del figlio del professionista di grido, fino ad attribuire l'insufficienza al compito eseguito di nascosto da un docente universitario. La *docimologia* non si presenta quindi solo come la scienza della giustizia distributiva ma anche come un *detector* delle nostre debolezze, dei nostri *idola theatri*. Quanto il dominio critico di queste manchevolezze o limiti oggettivi sia necessario lo si può capire solo pensando all'influenza che valutazioni su scala nazionale come esami o concorsi possono avere sul destino di migliaia di candidati.

## La nuova scuola media

Collana di pedagogia e didattica  
per gli insegnanti  
diretta da Aldo Visalberghi  
e Maria Corda Costa

15

Gilbert De Landsheere

**Elementi di docimologia**  
**Valutazione continua ed esami**



**La Nuova Italia Editrice**

## Indice

Introduzione	p. IX
I	
Definizioni	
I	
Introduzione	3
Docimologia, docimastica e dossologia, 3. Esami e concorsi: osservazione e valutazione continua, 4. Voti e punteggi, 6.	
II	
L'accusa e la difesa	
I	
Critica degli esami	11
Corpi estranei nell'educazione, al servizio di una pedagogia superata, 11. Ansia e stress, 12. Disuguaglianza-ingiustizia, 13. L'insuccesso, generatore di insuccessi, 17. Frattura fra insegnamento ed esame, 20. Disaccordo tra correttori, 21. Infedeltà di uno stesso correttore, 26. Stereotipi ed effetto di alone, 28. Mancanza di validità, 32. Uno strumento di molte esperienze docimologiche, 36. Altre critiche, 36.	
II	
Difesa del voto soggettivo e dell'esame	39
La misura rigorosa è forse impossibile, 40. Gli insegnanti giudicano bene i loro alunni, 42. Validità limitata ma reale	

Tutti i diritti riservati  
Printed in Italy

Titolo originale  
*Évaluation continue et examens. Précis de docimologie*

Traduzione di Anna Corda

© Copyright 1971 by Éditions Labor, Bruxelles  
and 1973 by « La Nuova Italia » Editrice, Firenze

1ª edizione: settembre 1973



degli esami tradizionali, 44. «Temprarsi per la vita», 44. Definirsi in rapporto agli altri, 45. Ampia sintesi e integrazione di conoscenze, 45. L'esame esterno controlla il professore, 45. L'esame esterno, «feedback» per il professore, 46.

## III

## Costruzione dell'esame

## I

L'oggetto e gli obiettivi . 53

L'oggetto, 53. (A) Il pronostico, 54. (B) L'inventario (prove di rendimento), 55. (C) La diagnosi, 56. Gli obiettivi, 58. (A) Gli obiettivi generali, 59. Gli obiettivi cognitivi, 60. Ambito affettivo, 66. (B) Gli obiettivi speciali, 68. Tassonomia di Bloom, 72.

## II

Formulazione di quesiti . 76

Osservazioni generali, 76. Risposte aperte o chiuse?, 83.

## III

La classificazione . 98

Un preambolo indispensabile: la curva di Gauss, 98. La classificazione soggettiva: la scala di valutazione, 110. Criteri per classificare le composizioni, 131. La classificazione obiettiva, 137. La taratura o misura della posizione relativa, 138.

## IV

Controllo della fedeltà dell'esame . 146

Evitare ogni ambiguità nei quesiti, 146. Quesiti in numero sufficiente, 147. Un controllo matematico, 147. Ripetizione della classificazione, 148.

## V

Controllo della validità . 149

La validità del contenuto, 149. La validità predittiva, 152. Fattori di validità di un esame, 153.

## IV

## Le procedure di moderazione

## I

Posizione del problema . 157

Definizione, 157. Moderazione volontaria o imposta?, 159. La moderazione comincia all'inizio dell'anno scolastico, 159. Non è possibile una confrontabilità senza una fedeltà elevata, 161. Si può fare affidamento sui test?, 161.

## II

Alcuni sistemi di moderazione degli esami . 163

La formula più liberale: il sistema svedese di moderazione per materia a partire da un test nozionale, 163. Sistema imposto di moderazione per materia a partire da un test di conoscenze, 166. Moderazione volontaria con ricorso ad una banca di item, 166. Un sistema semplice di moderazione globale al servizio della selezione, 167. In Inghilterra, un sistema di moderazione completo, 170. Il voto di fine anno: lavoro dell'anno + lavori pratici + test, 181.

## V

## Il mito della curva di Gauss

Il pericoloso mito della curva di Gauss, 189.

## I

Evoluzione della curva delle conoscenze . 191

La curva delle attitudini, 191. La curva delle conoscenze, 193.

## II

Una pedagogia di curva a J . 196

## III

La teoria della valutazione formativa . 198

Tracciare l'itinerario conoscitivo, 200. Guidare l'alunno, 201. Limite pratico del sistema, 202. Il sistema d'insegnamento messo in discussione, 203.

VIII	<i>Indice</i>	
IV		
Conclusioni e raccomandazioni		. 205
Appendici		
I		
Un quesito d'esame secondo il metodo tradizionale e secondo il metodo con quesiti a scelta multipla		. 211
Formulazione del quesito (metodo tradizionale), 211. Formulazione del quesito (metodo con quesiti a scelta multipla), 212.		
II		
Composizione in lingua materna		. 216
Prova I: due quesiti, un'ora e mezza, 216. Prova II: quattro quesiti, un'ora e 3/4, 217. Norme per la correzione; Prova I: massimo 50 punti, 218. Prova II: massimo 50 punti, 222.		
III		
Esempio di insegnamento semi-individualizzato		. 226
La New Trier Township High School di Winnetka, 226. L'individualizzazione dei programmi, 227. L'individualizzazione dell'insegnamento, 230. Lo studio di ciascun alunno prima del suo ingresso a New Trier, 237. I consiglieri pedagogici, 240.		
Bibliografia		. 243

## Introduzione

Nel nostro costume pedagogico si introducono nuovi metodi di valutazione. I risultati dell'*osservazione continua* prendono progressivamente il sopravvento sull'esame di fine d'anno.

L'innovazione è lodevole. Essa non significa affatto che si vuole sopprimere la *misura* a scuola, ma, al contrario, che si vuol renderla moralmente più giusta e scientificamente più esatta. Questo cambiamento va largamente al di là dell'ambito delle votazioni e coincide con profonde trasformazioni dell'educazione. L'avvenimento non è affatto fortuito. La civiltà contemporanea, l'economia della nostra società, richiedono un uomo provvisto di qualità e di conoscenze nuove: in queste condizioni le innovazioni pedagogiche si impongono irrevocabilmente.

Tuttavia la contestazione degli esami tradizionali ha creato un grave malinteso. Esso somiglia in maniera sorprendente a quello che sorse, soprattutto tra il 1920 e il 1940, quando l'adozione delle idee, mal digerite, della pedagogia funzionale dei Claparède, Dewey, ed altri « progressisti », condusse al culto dell'improvvisazione, al disprezzo della disciplina rigorosa, all'insistenza su interessi che non esistevano che nella mente dei teorici...

Con l'avvento di J. Dewey, la scuola avrebbe potuto prendere come motto: « Lo sforzo è morto, viva lo sforzo! ». Allo stesso modo saremmo tentati di proporre oggi: « Gli esami muoiono, gli esami sono morti, viva gli esami! ».

Senza dubbio i vecchi esami che avvelenavano l'atmosfera e i contenuti di cicli interi di studi e decidevano di una carriera scolastica o professionale in qualche ora — perfino in qualche minuto! — devono sparire. È anche indubbio che la maggior parte dei lavori docimologici pubblicati finora non sono stati che una messa a punto di sistemi perfezionati per continuare a far male le cose, perché non si basavano su una rimessa in discussione di tutto l'insegnamento.

Ma se invece si tiene conto, in ogni attività didattica di base, del posto che spetta alla valutazione, al *feedback* come dicono gli psicologi d'oggi, allora gli esami esistono in maniera *quasi* continua nella vita scolastica. Sia che consistano in brevi valutazioni, sia che prendano la forma di prove più lunghe, che riguardano talvolta numerose materie.

L'essenziale è che essi non si inseriscano come corpi estranei nel processo educativo, ma che ne facciano parte integrante. Per questa stessa ragione quindi essi non si riferiscono esclusivamente all'ambito della conoscenza, ma mirano a investire la personalità nel suo complesso.

Ogni riforma profonda si accompagna, quasi necessariamente, ad un periodo di adattamento, di tentativi, d'errori. Così si spiega la confusione di molti insegnanti che sperimentano attualmente il sistema della valutazione continua.

Così si spiega anche come una specie di tabù colpisca improvvisamente la parola esame e che, al momento in cui scriviamo queste righe, si rilascino qua e là certificati di studio sulla base di valutazioni puntuali, qualitativamente e quantitativamente insufficienti, per di più male integrate, basate su impressioni personali, accanto alle quali i vecchi esami napoleonici sembrano monumenti di obiettività! Ma non si tratta, ne siamo convinti, che di una crisi di adattamento.

Non è affatto paradossale quindi che dopo aver ripreso brevemente il processo agli esami ed esserci dichiarati ardenti partigiani dell'osservazione continua e del bilancio di fine ciclo, dedichiamo la presente opera... all'organizzazione degli esami. Ma il lettore ora sa che, se il termine è rimasto invariato, il suo senso è cambiato profondamente.

Molti insegnanti diffidano ancora dei test e degli altri strumenti che ambiscono a quantificare l'essere umano. Non senza motivo del resto, perché questi strumenti hanno fatto, negli ultimi decenni, le loro « malattie infantili ». E poiché coloro che li hanno utilizzati non sempre hanno saputo compensare l'imperfezione delle tecniche con una adeguata moderazione di giudizio ed una comprensione profonda delle situazioni d'insieme, molti errori sono stati commessi.

Ma se la prudenza resta necessaria, se lo spirito critico e l'occhio clinico non perderanno mai i loro diritti, tuttavia sono stati compiuti considerevoli progressi: ora possiamo misurare parecchi comportamenti umani in maniera soddisfacente.

Oggi alcuni potrebbero rifiutare la misura obiettiva, non tanto perché ogni validità le fa difetto, ma perché non ne comprendono l'economia. Negare ogni valore a ciò che ignoriamo, è una reazione di difesa ben conosciuta. L'aspetto matematico dei metodi da adottare e della letteratura sperimentale, spesso indigesto, complica ancora la situazione.

Passato è il tempo dell'opposizione tra psicometrismi e educatori « pratici ». I primi devono umanizzare le loro cifre; i secondi introdurre maggior rigore nei loro procedimenti; tutti e due devono unire ed armonizzare i loro sforzi per il maggior profitto degli studenti e della comunità.

Quelli che sfogliando questo libro hanno visto alcuni numeri si rassicurino: le quattro operazioni saranno un sufficiente bagaglio di cognizioni!

Come al solito, i processi che descriveremo sono più difficili da spiegare che da applicare.

Certo non saranno svelati tutti i segreti della misura e della statistica, ma le semplici nozioni che troveremo hanno un valore pratico provato. Esse hanno, inoltre, un innegabile potere di demitizzazione e faciliteranno ulteriormente la lettura di lavori docimologici più specializzati.

I  
Definizioni

# I

## Introduzione

### Docimologia, docimastica e dossologia

La *docimologia* è una scienza che ha per oggetto lo studio sistematico degli esami, in particolare dei sistemi di votazione, e del comportamento degli esaminatori e degli esaminati.

La *docimastica* è la tecnica degli esami.

All'inizio, la docimologia ha assunto un carattere negativo criticando i sistemi di votazione e dimostrando sperimentalmente la mancanza di fedeltà e di validità agli esami.

In seguito è entrata in una fase costruttiva cercando di proporre metodi e tecniche di valutazione più obiettivi o, quanto meno, più rigorosi, e mettendo a punto i mezzi per rendere le classificazioni confrontabili, in maniera da assicurare una maggiore giustizia scolastica.

La *dossologia* è lo studio sistematico della funzione che la valutazione ha nell'educazione scolastica.

Guillaumin<sup>1</sup> le assegna specialmente gli obiettivi seguenti:

- studio degli effetti inibitori o stimolanti delle differenti forme d'esami;
- studio delle reazioni emozionali degli alunni e, successivamente, delle reazioni intellettuali, ai giudizi del maestro;

<sup>1</sup> J. Guillaumin, *L'aspect interpersonnel de la notation: de la docimologie à la doxologie pédagogique*, in « Bulletin de la Société A. Binet et T. Simon » 86 1968, pp. 250-275.

- studio degli effetti che ha l'opinione del maestro sui propri alunni, sul suo insegnamento e sull'apprendimento scolastico;
- studio dei processi messi in moto e degli effetti ottenuti dall'automatizzazione, dalla votazione reciproca, dalla votazione di gruppo, dall'assenza di votazione.

### Esami e concorsi: osservazione e valutazione continue

È difficile dissociare l'esame e il concorso dall'idea di *prova*, divenuta del resto loro sinonimo, dopo aver indicato, più generalmente, la sofferenza, i guai, il pericolo, che mettono in luce coraggio e resistenza.

Nell'esame, l'ammissione è determinata da un voto che il candidato deve raggiungere o superare, mentre nel concorso il numero dei posti da coprire è fissato in anticipo. La presenza della *minaccia*, del *pericolo* — quello della bocciatura — è nei due casi innegabile. Per poco poi che la procedura sia viziata da gravi imperfezioni, non siamo molto lontani dall'« ordalie ».

Nella nozione di concorso e d'esame percepiamo una *carica aggressiva* totalmente assente dai concetti di osservazione e di valutazione continue. La serenità, la benevolenza, magari l'indulgenza, pervadono il maestro, che segue, con una simpatia da cui non è del tutto esclusa qualche severità, il lungo progredire dei suoi alunni verso l'equilibrio del momento e la piena accessione allo stato adulto, nel futuro.

Gli esami segnano il termine delle tappe; i concorsi aprono le porte agli eletti: sono dei punti fermi nel flusso, avvenimenti nel processo dell'educazione.

L'apprendimento che precede questi avvenimenti consiste fondamentalmente in una successione continua di comportamenti e di *feedbacks*, cioè di informazioni che danno lumi sulla loro validità e la loro pertinenza. Non è nostro compito intavolare qui una lunga discussione sulle modalità e sugli effetti

del « rinforzo » dei comportamenti. L'importante è che, senza di esso, l'apprendimento non sembra potersi produrre.

La *valutazione*, nel senso ristretto che le diamo nel presente lavoro, *merita dunque un posto importante nell'insegnamento, di cui essa fa parte integrante. Essa è sempre, direttamente o indirettamente, in rapporto col progresso, per estensione o qualità, dell'apprendimento.* La valutazione ha tre funzioni:

(1) Una funzione prognostica: l'alunno è provvisto o no delle qualità intellettuali e caratteriali e delle conoscenze necessarie per affrontare una nuova materia o un ciclo di studi superiori? È giunto al livello al quale dovrebbe trovarsi? Rispondere a queste domande equivale a predire il successo nella tappa che sta per cominciare.

(2) Una funzione di misurazione:

(a) controllo delle acquisizioni;

(b) valutazione del progresso, è il caso in cui l'alunno viene confrontato con se stesso;

(c) situazione dell'alunno in un dato momento: nella classe o nel gruppo di lavoro; nell'insieme delle classi parallele di una stessa scuola; in insiemi più vasti: città, cantone, provincia, paese.

Non si tratta necessariamente di procedere ad un esame o ad un concorso, ma di fare il punto, di determinare la posizione relativa.

(3) Una funzione diagnostica: perché non si è prodotto un apprendimento perfetto? di quali materie o tecniche lo studente non è sufficientemente padrone? quali sono i processi mentali in causa?

Non sono sempre gli strumenti necessari alla valutazione, all'esame o al concorso che differiscono, bensì la maniera di utilizzarli. Così, per evitare fastidiose precisazioni verbali useremo, nelle pagine che seguono, la parola esame in due diversi significati che il contesto chiarirà sempre: esami propriamente detti e, più generalmente, ogni procedura pedago-

gica avente per oggetto un metro di valutazione di apprendimenti o di conoscenze.

#### *Esami interni ed esami esterni*

In senso stretto, l'*esame interno* in una materia è elaborato e classificato dal docente che l'ha insegnata; è subito dagli alunni che hanno ricevuto questo insegnamento, nell'ambito della classe o della scuola.

In senso più ampio, sono qualificati interni gli esami organizzati indipendentemente in ogni scuola, che esista o no un coordinamento o una unificazione per materia e per livelli e sezioni.

Con il termine *esami esterni*, si indicano le prove organizzate e classificate da commissioni indipendenti dalle scuole, a livello locale, regionale o nazionale. Di queste prove le più note sono quelle del baccalaureato francese; citiamo ancora, in Belgio, gli esami cantonali, alla fine degli studi primari, e in Inghilterra, fino a questi ultimi tempi, l'Eleven+Examination all'ingresso nell'insegnamento secondario.

#### Voti e punteggi

Pensiamo che la distinzione tra voti e punti sarebbe molto utile nell'ambito dell'educazione.

In un dettato, l'alunno può commettere un certo numero di errori di cui non ci interessa molto valutare l'esistenza: essi ci sono o non ci sono. Tuttavia, il numero di errori rilevati non ha, in sé, alcun significato educativo: fare cinque errori nel dettato di Mérimée testimonia una straordinaria conoscenza dei misteri dell'ortografia; in altre condizioni, lo stesso risultato denuncia una grave insufficienza. È dunque necessaria anche una considerazione relativistica.

Con *punteggio*, indicheremo i risultati obiettivi ottenuti in un test o in qualunque altra forma di valutazione con calcolo e detrazione di punti, secondo regole fisse: numero di errori in dettato, risultati in un test standardizzato.

Per *voto*, intenderemo un apprezzamento sintetico che traduce la valutazione di una performance nel campo dell'educazione.

Il voto può essere obiettivo o soggettivo, ma esso è sempre relativo. Attribuire il voto A ad un alunno la cui performance si situa a un certo livello in una selezione nazionale, fa parte della prima categoria; scrivere un « buono » sotto la sua composizione fa parte della seconda.

II

L'accusa e la difesa



## I

### Critica degli esami

Una critica particolareggiata degli esami tradizionali è stata fatta da Piéron nel suo ottimo lavoro *Examens et docimologie*<sup>1</sup>. Tutti gli insegnanti dovrebbero averlo letto.

Nel presente capitolo vi sono due tipi di osservazioni. Alcune sono brevi richiami ad imperfezioni ben note, messe in luce instancabilmente dai docimologi; le altre si riferiscono a problemi studiati meno bene: stereotipia, effetto di alone, effetto edipico della predizione. Talora è in causa solo l'esame, talora tutta la pedagogia nella quale esso si inserisce.

#### Corpi estranei nell'educazione, al servizio di una pedagogia superata

Sia che si tratti di interrogazioni periodiche o d'esami trimestrali o annuali, la valutazione si riduce spesso ad un controllo di memorizzazione di conoscenze e lascia inesplorati non solo gli aspetti più importanti dell'intelligenza ma anche quasi tutti i tratti della personalità che una ben intesa educazione deve coltivare.

Ci troviamo, in realtà, davanti ai postumi di un sistema pedagogico in cui, come a Brameld piace dire, le lezioni servono da nastri trasportatori di conoscenze e di valori, selezio-

<sup>1</sup> Paris, PUF, 1963 (trad. it. *Esami e docimologia*, Roma, Armando, 1971<sup>2</sup>).

nati in funzione di un ruolo predeterminato in una società non democratica.

In questo quadro, l'esame costituisce una specie di marchio di fabbrica, di verifica di conformità allo stampo, alla sagoma, in breve alle indicazioni imposte dall'autorità.

Lo sviluppo della persona umana non trova posto in queste esigenze, perché l'educazione non è concepita per essa, ma al servizio di un regime.

Le infauste conseguenze di questa situazione sono state mille volte analizzate. L'esame, anziché servire all'alunno, informarlo in maniera funzionale del valore dei suoi comportamenti nel corso dell'apprendimento, far nascere in lui un miglior adattamento, dare dunque una valutazione accettata al punto che lo studente vi partecipi sinceramente e spontaneamente, è mal accetto come una dichiarazione di imposte, che non è certo nemmeno essa generatrice di civismo.

Inoltre il rigetto profondo del procedimento il cui effetto si combina con la povertà intellettuale delle domande poste, conduce ad un risultato facilmente immaginabile: nei quindici mesi che seguono un esame, dimenticare l'80% dei dati nozionistici non è eccezionale.

Così, se si prescinde dalla funzione sociale che ha l'esame, il fiasco è completo: non abbiamo né educato, né istruito.

### Ansia e stress

In una civiltà in cui il successo scolastico condiziona il successo materiale e sociale, al punto che Young ha creduto di dover denunciare i pericoli della « meritocrazia », l'esame che decide del passaggio di classe o del conseguimento del diploma è temuto dal bambino e dalla sua famiglia. Anche nel corso dell'anno, le prove di controllo di conoscenze e d'attitudini sono affrontate con tensione ed apprensione, ciò che non costituisce certo la condizione ideale, e snatura profondamente la funzione educativa della misurazione degli apprendimenti.

Circostanza aggravante, la tradizione vuole che da noi, i « grandi » esami, che si basano su tutte le materie, si svolgano consecutivamente in alcuni giorni. Durante una sessione universitaria, un anno, ebbi il triste privilegio di essere interrogato su quattordici corsi in due giorni consecutivi... Nell'insegnamento secondario, lo stesso fenomeno si produce spesso anche nel corso dell'anno, quando l'avvicinarsi della consegna ai genitori del « bollettino periodico »<sup>2</sup> suscita una fioritura di interrogazioni.

Nei due casi, l'accumulo delle prove e la loro preparazione, spesso con la difficoltà dell'apporto di nuova materia fino all'ultimo momento (alcuni professori non sempre rispettano i periodi di revisione o non li utilizzano nella maniera più felice), stancano il fisico e la mente in maniera considerevole.

Alcuni sostengono che anche queste difficili condizioni rivestono un valore educativo: la vita moderna non ci risparmia quasi mai lo stress, ed è bene esservi preparati. D'accordo, ma a condizione che almeno tutto questo sia chiaro!

Prima di organizzare un esame, bisogna definire chiaramente il suo scopo: si tratta di controllare l'acquisizione delle conoscenze? la resistenza allo stress? o la capacità di saper riferire conoscenze e di usarle in situazione di stress? Questi tre obiettivi sono differenti e richiedono prove differenti.

### Disuguaglianza - ingiustizia

Nel nostro sistema scolastico, ogni professore redige le domande d'esame destinate ai suoi allievi. Il principio è ottimo e in paesi come la Gran Bretagna dove, per tradizione, tutti gli esami chiave erano affidati a commissioni esterne alla scuola, si tende ad accordare ad ogni insegnante la responsabilità del controllo. Tuttavia, i sistemi anglosassoni sono più severi dei nostri, perché vengono presi precisi provvedimenti per as-

<sup>2</sup> Consegna delle pagelle a scadenze variabili, mensili, bimestrali, trimestrali, ecc. [NdT].

sicurare la confrontabilità degli esami nella regione e perfino in tutto il paese. Del resto una parte di questo libro è dedicata proprio ai processi utilizzati per raggiungere questo scopo.

La libertà, quasi totale, lasciata ai nostri insegnanti porta a situazioni ingiuste.

Le scuole del nostro paese costituiscono dei mondi isolati le cui popolazioni presentano caratteristiche a volte molto diverse. Un piccolo istituto d'istruzione secondaria non conta che pochi alunni, per la maggioranza handicappati dalle loro origini socioeconomiche. Al contrario, un altro istituto posto in una grande città vede le sue sezioni « forti » frequentate da una maggioranza favorita socialmente ed intellettualmente, essendo gli adolescenti in difficoltà orientati verso una vicina scuola, che ha fama di essere più comprensiva.

Col tempo, e col favore della concorrenza, si creano negli istituti tradizioni di severità o di « generosità ». Marion Coulon osservava recentemente che, quando scuole dello stesso tipo sono create successivamente in una stessa regione, la severità degli esami è inversamente proporzionale all'anzianità dell'istituto. Non possiamo certamente generalizzare, ma chi oserebbe dire che questa affermazione è priva di ogni fondamento? L'hanno capito molti genitori, che desiderano per i loro bambini un determinato diploma, senza preoccuparsi troppo del valore intrinseco degli studi: in caso di insuccesso è più facile cambiare scuola che organizzare un apposito trattamento o modificare l'orientamento.

Comunque, i docenti adattano il loro insegnamento e gli esami al livello della loro classe (cosa lodevole), con la conseguenza che un alunno, ben classificato in una scolaresca scadente, sarebbe fra i più scadenti o fallirebbe interamente in un gruppo ben preparato.

L'effetto cumulativo di tutti i singoli adattamenti porta a volte a situazioni scandalose. Qualche anno fa abbiamo potuto consultare delle statistiche le quali stabiliscono che, per uno stesso tipo e uno stesso livello d'insegnamento secondario, le possibilità di riuscita di un alunno variano da uno a otto secondo la scuola frequentata. Probabilmente è il caso

estremo; ancora ci mancano studi precisi che ci permettano di sapere ciò che accade realmente.

Bisogna insistere: il fatto che il maestro si adegui al livello dei suoi alunni non è, in sé, cosa criticabile, al contrario è caratteristico di un buon educatore.

L'ingiustizia appare quando voti puramente relativi sono utilizzati in situazioni di concorrenza esteriore, o sono l'unico criterio per il conseguimento di un diploma da cui può dipendere l'avvenire degli alunni.

In ogni modo constatiamo variazioni considerevoli nella *quantità delle materie su cui si è basato l'esame* e nella *qualità delle risposte richieste*. In un caso è sufficiente la ripetizione a memoria di una piccola parte di programma, mentre in un altro caso le domande esigono analisi, sintesi e giudizio personale in merito ad un ampio settore di materia.

Un insegnante pone solo una o due domande, scelte più per la facilità di correzione che offrono che per la loro importanza reale, un altro si sforza di scorrere tutto il programma col rischio di trasformare l'esame in una maratona. Nello sport, sappiamo che questa lunga corsa deve essere riservata ad individui eccezionali; perché, nel campo dell'istruzione, i trenta alunni di una classe dovrebbero essere tutti dei fondisti?

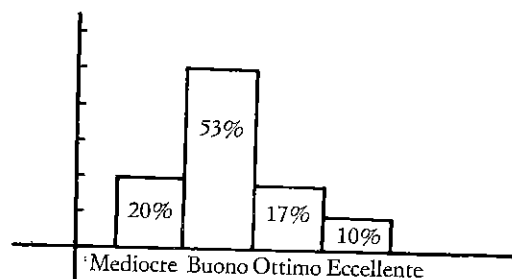
Prima di proseguire, ecco tre risultati di ricerche che ci permetteranno d'illustrare i fenomeni che stiamo per trattare.

Gjorgjevski<sup>3</sup> ha dimostrato, con una semplice esperienza, come, per relativismo alcuni professori valutino diversamente una stessa performance.

Cinque correttori, professori della stessa materia, hanno classificato indipendentemente l'uno dall'altro cento prove scritte provenienti da una scuola secondaria. In seguito sono stati estratti quindici compiti che avevano ottenuto tutti il voto « buono ». Sono stati affidati per una nuova correzione ad

<sup>3</sup> Vedere N. Rot e Z. Butas, *Les distributions des notes scolaires comparées aux distributions des résultats obtenus aux tests de connaissances*, in « Le travail humain » 1-2 1959.

altri quattro professori; questi, del tutto normalmente, hanno adottato nuovi criteri: solo nella metà circa dei casi il voto è rimasto « buono ». Ecco la nuova ripartizione <sup>4</sup>.



Proseguendo la sua esperienza, Gjorgjevski ha estratto dalla stessa serie di cento compiti, un gruppo di dodici giudicati ottimi, ed un gruppo di dodici giudicati mediocri.

In ogni gruppo, ha inserito tre compiti giudicati buoni. La valutazione è stata fatta da due gruppi di cinque professori. La media dei compiti « buoni » inseriti negli « ottimi » è scesa da 3 su 5 a 2,40; nel gruppo opposto la media è passata da 3 a 3,87.

Si deve, d'altra parte, a Reuchlin <sup>5</sup> uno studio che si basa su 4860 alunni distribuiti in 397 scuole primarie di Francia. Gli insegnanti sono stati invitati a dividere i loro alunni in quattro categorie: molto buoni, buoni, medi e mediocri. Poi è stato passato uno stesso testo di francese e di calcolo a tutti gli allievi. Lo spoglio dei risultati ha mostrato che lo stesso livello di conoscenze era conseguito da alunni che, secondo la scuola, erano giudicati ottimi, buoni, medi o mediocri.

Come Reuchlin sottolinea perspicuamente: « Si può dire che l'insegnante, certamente, conosce meglio di chiunque altro i punti del programma che sono stati assimilati o no da

<sup>4</sup> Il numero di compiti è troppo basso per condurre a risultati che abbiano un valore statistico.

<sup>5</sup> Vedere « Le travail humain » cit., p. 12 ss.

ciascuno dei suoi allievi. Ciò che ignora è la gravità da attribuire a ciascuna insufficienza, ad ogni lacuna, quando la si considera non più nel contesto di una classe che può essere « preparata » o « scadente », ma in rapporto all'insieme del paese. Da ciò, le divergenze di valutazione messe in luce dall'inchiesta ».

I grafici della pagina seguente sintetizzano i risultati.

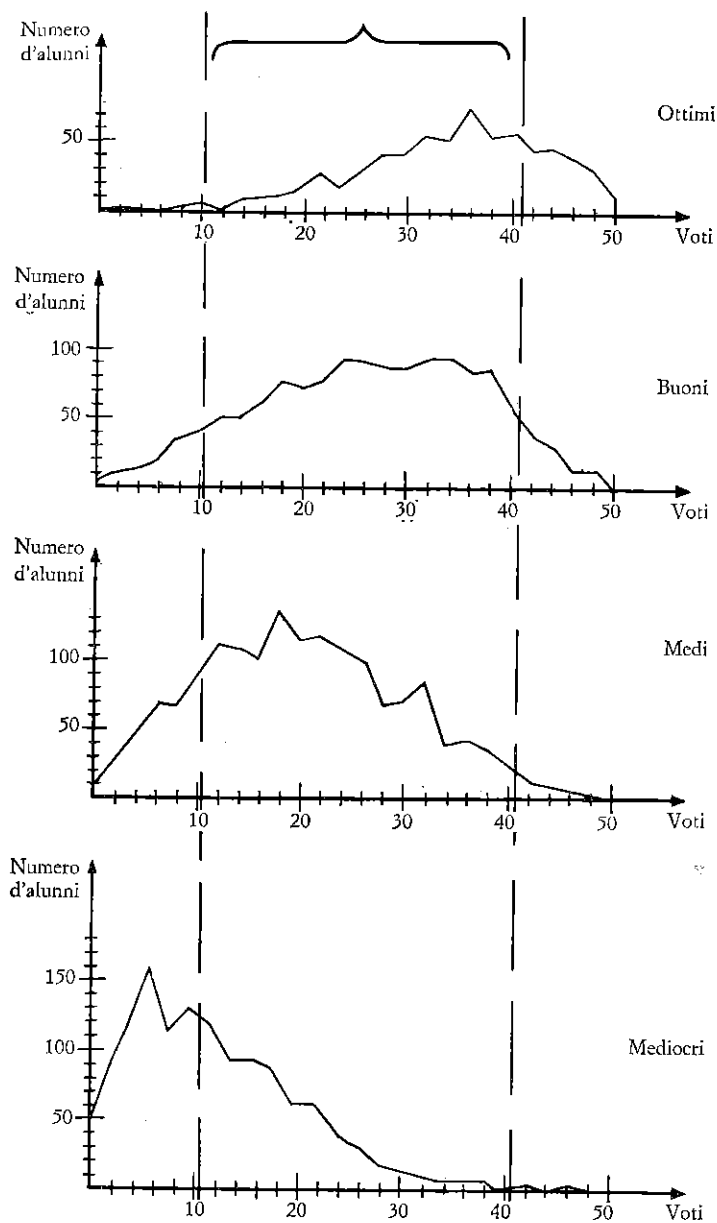
Ecco ancora due esempi dati da F. Bacher <sup>6</sup>: « In una inchiesta francese che si basa su un campione rappresentativo di alunni di terza (fine del primo ciclo secondario), si fa eseguire una prova di conoscenze letterarie ed una prova di conoscenze matematiche agli alunni di 406 classi. Si è constatato che le medie di classe si distribuivano, in lingua francese da 23 a 60 (per una prova il cui massimo teorico era 80), e da 7 a 38 in matematica (per una prova il cui massimo teorico era 44) (Reuchlin, Bacher 1968). Negli Stati Uniti, Flanagan (1964) segnala che alunni classificati nel quarto inferiore della loro scuola sarebbero classificati nel quarto superiore se essi frequentassero altre scuole della stessa regione ».

### L'insuccesso, generatore d'insuccessi

Si può obiettare che gli studi superiori o la vita si incaricheranno di ristabilire la giustizia che le classificazioni scolastiche non avrebbero rispettato. Disgraziatamente le cose non sono così semplici. Innanzi tutto perché gli alunni scartati o scoraggiati ingiustamente per esempio dall'insegnamento secondario in genere, non avranno forse più l'occasione di far valere la loro possibilità di accedere all'università. Poi perché i favoriti rischiano di vincere su tutti i fronti.

Le ricerche attuali confermano un vecchio adagio pedagogico: il successo genera il successo e l'insuccesso di oggi prepara un insuccesso domani. Sostenuti dal successo, alcuni

<sup>6</sup> F. Bacher, *La normalisation de la notion*, in *Docimologie et éducation*, numero speciale della rivista « Les Sciences de l'Éducation » 2-3 1969, pp. 51-52.



alunni valorizzano un capitale intellettuale limitato. Il passaggio di classe con una promozione « regalata » ha talvolta la doppia conseguenza salutare di non spingere i genitori a ritirare i loro bambini dalla scuola in cui si trovano e di permettere allo studente di ritrovare una sorta di nuova lena.

Quante volte gli insegnanti delle scuole secondarie vedono il rendimento di un alunno riprendersi in modo spettacolare quando la crisi fisica e psicologica dell'adolescenza non è più acuta?

Rosenthal e Jacobsen<sup>7</sup> hanno recentemente pubblicato uno studio sull'aspetto dell'*effetto edipico della predizione*, così definito perché se alla nascita di Edipo l'oracolo non avesse predetto che avrebbe ucciso suo padre, egli non sarebbe stato allontanato dalla sua famiglia. Conoscendo suo padre, non l'avrebbe ucciso. In altre parole, la tragedia si produsse *perché era stata predetta*.

Sembra che, in misura certamente difficile da calcolare, un alunno si comporti in funzione del giudizio che il professore formula su di lui. Anche gli animali da addestrare non sfuggono a questa regola. Rosenthal ricorda l'esperienza seguente: si formano due gruppi di topi bianchi geneticamente identici. Al momento in cui si consegnano agli studenti incaricati di addestrarli, un segno indica che il primo gruppo è composto di animali particolarmente ben dotati, mentre il secondo è di mediocre qualità. I risultati dell'addestramento confermeranno questo pronostico fantasioso.

Nello stesso contesto si inserisce l'esperienza seguente condotta da Rosenthal.

<sup>7</sup> *Pygmalion in the Classroom. Teacher Expectations and Pupil's Intellectual Development*, New York, Holt, Rinehart and Winston, 1968.

Figura 1. I quattro grafici permettono di confrontare la distribuzione dei voti ottenuti nella prova di calcolo da 654 alunni giudicati « ottimi » dai loro maestri, da 1303 alunni « buoni », 1551 « medi » e 1300 alunni giudicati « mediocri ». Si vede che le distribuzioni di voti si sovrappongono largamente: nella zona di voti che va da 10 a 40, lo stesso livello di conoscenze può, secondo i casi, essere classificato ottimo, medio o mediocre.

È stato annunciato a degli insegnanti delle classi primarie della Oak School che eminenti ricercatori avevano appena messo a punto un test di « crescita intellettuale » (è stato utilizzato in realtà un semplice test d'intelligenza<sup>8</sup>, poco conosciuto e senza caratteristiche diagnostiche particolari). Ad esso sono stati sottoposti tutti gli alunni della scuola, e sono stati segnalati quelli che « erano sul punto di " fiorire " intellettualmente » (pronostico naturalmente di fantasia).

Il quoziente intellettuale degli alunni così definiti si è elevato in maniera significativa nei tre anni inferiori; inoltre, progressi superiori alla media sono stati osservati in lettura ed in aritmetica<sup>9</sup>.

### Frattura tra insegnamento ed esame

*L'esame deve essere il riflesso dell'insegnamento dato.* Immaginiamo che invece di condurre i suoi alunni in piscina, un professore di educazione fisica abbia passato l'anno a disertare sul nuoto. Evidentemente sarebbe inammissibile che i voti d'esame venissero assegnati dopo le performance nel nuoto perché riuscirebbero a nuotare solo quelli che hanno imparato a nuotare indipendentemente dal corso.

Un simile modo d'agire sembra impensabile. Tuttavia, quante volte assistiamo ad una simile aberrazione! Il professore tiene un corso nel quale impone le informazioni, le opinioni e i giudizi, poi, all'esame, pone domande così dette d'intelligenza, domande cioè che mettano in gioco proprio capacità che non sono state promosse durante le lezioni. Perché gli studenti sarebbero all'improvviso capaci di scoprire da soli, all'esame, la soluzione di applicazioni geometriche nuove se non hanno potuto farlo durante l'anno?

Così si possono spiegare certe percentuali anormali di in-

<sup>8</sup> Misurato con l'aiuto del Toga di Flanagan (1960).

<sup>9</sup> Le cifre presentate da Rosenthal e Jacobsen sono contestate da R. L. Thorndike. Invece la tendenza dell'aumento dei QI e dei punteggi sembra riconosciuta da tutti gli specialisti. Evidentemente è questo che importa. Per ulteriori particolari vedere « American Educational Research Journal » 4 1968, p. 708.

successo. Una inchiesta provocata da numerose insufficienze date da un professore di chimica (solo 4 alunni su 22 avevano ottenuto più della metà dei punti all'esame di fine anno) mise ben presto in luce che il professore non faceva che due o tre interrogazioni scritte all'anno. Sulle tre domande d'esame, due implicavano interpretazione e transfert di apprendimento, caratteristiche che al livello di difficoltà in cui l'insegnante si poneva, avevano bisogno di una profonda comprensione dei fenomeni e di un esercizio, lungo e accuratamente controllato, per la soluzione di problemi nuovi.

Un'amichevole discussione col giovane professore in questione rivelò che egli credeva di essere obbligato a trattare in particolare tutti i punti del programma, che non si rendeva conto della necessità dell'esercizio sistematico o almeno semi-individualizzato per la soluzione dei problemi nuovi e che, in ogni modo, non aveva le nozioni psico-pedagogiche di base necessarie a guidare tale esercizio.

### Disaccordo tra correttori

Tutti i docimologi citano esempi di grandi divergenze tra insegnanti chiamati a valutare uno stesso lavoro di un alunno o uno stesso gruppo di lavori, ciò che d'altronde permette una relatività meno accentuata. Anche le fluttuazioni fra i voti di uno stesso professore che giudica lo stesso lavoro ad un certo intervallo di tempo (fedeltà) possono essere di una certa importanza.

Poiché tutte le opere di docimologia brulicano di esempi, ne faremo soltanto alcuni.

#### (a) Composizione

È l'accusata preferita.

Piéron riferisce che una stessa composizione, giudicata da 76 correttori, tutti professori di lingua materna, condusse ai seguenti risultati<sup>10</sup>:

<sup>10</sup> Piéron, op. cit., p. 123.

Voto in ventesimi			Numero di correttori che l'hanno attribuito
da	a	1	
0	1	1	1
2	3	6	6
4	5	20	20
6	7	34	34
8	9	10	10
10	11	3	3
12	13	2	2

Il Servizio di ricerca pedagogica del CRDP di Lione ha invitato nel 1967<sup>11</sup> 150 professori di francese dell'Académie de Lyon a correggere tre composizioni francesi che trattavano uno stesso soggetto.

Massimo 20	Media	Margine di variazione
I compito	10 $\frac{1}{4}$	4 $\frac{1}{2}$ - 13 $\frac{3}{4}$
II compito	6,5	2 $\frac{1}{2}$ - 12 $\frac{1}{2}$
III compito	11,6	5 $\frac{1}{2}$ - 17 $\frac{1}{2}$

L'intervento di uomini colti, estranei alla routine dell'insegnamento, migliorerebbe la situazione? Diederich ed i suoi collaboratori<sup>12</sup> hanno confrontato la maniera di correggere di 10 professori di lingua materna e di altri 43 classificatori: specialisti di scienze sociali, di scienze naturali, scrittori, editori, avvocati e uomini d'affari (quadri superiori).

Tutti sono stati invitati, senz'altra precisazione, a classificare 300 composizioni in 9 gruppi, in ordine di merito generale. Nessuno dei gruppi poteva contare meno di 12 compiti (4%). Il 94% dell'insieme si son visti attribuire 7,8 o 9 delle 9 possibili classificazioni. Nessuna composizione ha ricevuto meno di cinque classificazioni diverse.

<sup>11</sup> CRDP, *Docimologie et examens*, Lyon, IPN, 1969.

<sup>12</sup> P. Diederich e altri, *Factors in Judgments of Writing Ability*, Princeton, E.T.S., « Res. Bulletin » 1961.

## (b) Matematica

Il CRDP di Lione<sup>13</sup> ha anche incaricato 150 professori di matematica che insegnano in una terza classe, di classificare i compiti di 3 alunni che dovevano risolvere lo stesso problema.

Massimo 20	Media	Margine di variazione
I alunno	5,70	0,5 - 11,5
II alunno	16	11,5 - 20
III alunno	8	3,5 - 11,5

Se per l'alunno I, per esempio, si eliminano le due classificazioni estreme, restano ancora sia correttori che votano 2/20 sia correttori che votano 8/20, mentre ci si poteva attendere una classificazione obiettiva.

## (c) Medicina

I risultati non sono più sicuri a livello universitario.

Uno stesso candidato di secondo anno di medicina ha subito, in giugno ed in settembre del 1965, una prova scritta costituita da cinque quesiti semplici e precisi a cui rispondere in 1 ora e 30<sup>14</sup>.

Le risposte (anonime) sono state corrette indipendentemente da cinque correttori che avevano gli stessi titoli e le stesse caratteristiche.

Risultato totale su 100	Margine di variazione	Decisione *	
		Ammessi	Bocciati
Compito di giugno	47 - 78	3	2
Compito di settembre	32,5 - 73	3	2

\* Rileviamo che i tre ammessi furono gli stessi nei due casi.

<sup>13</sup> Op. cit.

<sup>14</sup> Vedere « Le Monde » del 6 settembre 1966.

Se si considerano i voti attribuiti a ciascuno dei 10 quesiti, si constata:

(1) scarto massimo	12	(da 3 a 15 su 20)
(2) scarto per 4 domande	11	
(3) più piccolo degli scarti massimi tra cinque correttori	7,5	

## (d) Varie

Agazzi<sup>15</sup> ricorda un esempio, ancor più sorprendente, che si riferisce ad un insieme di materie.

Sei correttori hanno, ogni volta, classificato gli esami del baccalaureato (voto d'insufficienza: meno di 10 su 20):

	Bocciati dai 6 correttori %	Ammessi dai 6 correttori %	Ammessi dagli uni e bocciati dagli altri %
Versione latina	40	10	50
Composizione francese	21	9	70
Inglese	37	16	47
Matematica	44	20	36
Filosofia	9	10	81
Fisica	37	13	50

## (e) Quanti correttori per stabilire il voto?

Calcolare la media tra due correttori porterebbe un miglioramento considerevole? No, se nessuna rigorosa misura è stata presa per mettere d'accordo i classificatori. Da quanti professori differenti bisognerà far valutare lo stesso lavoro per

<sup>15</sup> A. Agazzi, *Les aspects pédagogiques des examens*, Strasbourg, Conseil de l'Europe, 1967, p. 119. La ricerca è dovuta a H. Laugier e D. Weinberg e si basava sul baccalaureato francese.

ché la classificazione si stabilizzi? I numeri seguenti, proposti da Piéron, non hanno certamente valore assoluto. Danno tuttavia un'idea dell'ordine di grandezza dei numeri.

Composizione francese	78
Versione latina	19
Inglese	28
Matematica	13
Dissertazione filosofica	127
Fisica	16

La storia e la geografia non figurano nella lista; anche queste due materie danno luogo tuttavia a grandissime fluttuazioni.

Potremmo moltiplicare gli esempi di disaccordo tra classificatori. Ma, a che serve compiacersi nell'accusa facile, come hanno fatto troppi docimologi? L'importante è di essere coscienti del pericolo e, soprattutto, di cercare i rimedi.

Qual è la spiegazione delle divergenze osservate?

Essa è da ricercare principalmente nella molteplicità dei punti di vista, delle dimensioni secondo le quali una stessa prova può essere giudicata, e nell'assenza di direttive precise (a volte non ce ne sono per niente) date per la classificazione. Bacher<sup>16</sup> scrive:

« In una dissertazione, si può classificare l'organizzazione delle idee, la loro originalità, la correttezza dell'espressione, l'eleganza dello stile, ecc. Se si considera ogni aspetto possibile come una dimensione della classificazione, il compito di un allievo può essere situato in uno spazio a  $n$  dimensioni; la difficoltà deriva dal fatto che, in questo spazio, si stabilisce tra gli alunni solo un ordine parziale: un alunno che occupa una posizione elevata in tutte le dimensioni può essere dichiarato superiore ad un alunno che occupa una posizione meno elevata in tutte le dimensioni, ma non si può dire di un alunno che ha una posizione elevata in una dimensione e una posizione inferiore in un'altra dimensione se

<sup>16</sup> F. Bacher, *La normalisation des notes*, cit., pp. 53-54.



egli sia superiore o inferiore ad un alunno che occupa posizioni inverse ».

Inoltre, anche se considerano lo stesso aspetto, i classificatori variano in severità, in stabilità di giudizio, in resistenza all'effetto di alone, ecc.

### Infedeltà di uno stesso correttore

Uno stesso correttore è perlomeno uguale a se stesso? No. A fenomeni così evidenti come le variazioni di salute fisica e mentale e l'evoluzione del sapere si aggiungono un gran numero di fattori più o meno ben definiti: variazione nella qualità della relazione stabilita con l'alunno, nel contesto della valutazione (se abbiamo appena corretto un ottimo lavoro, quello che segue può essere sottovalutato,...), nella scala coscientemente o inconsciamente adottata, ecc.

Quattordici storici sono stati invitati a classificare una seconda volta quindici composizioni, da dodici a diciannove mesi dopo averle classificate una prima volta. Era stata cancellata ogni traccia di correzione. I professori dovevano non solo attribuire punti, ma indicare anche la riuscita complessiva o l'insuccesso.

In 92 casi su 210, il giudizio è stato differente la seconda volta<sup>17</sup>.

Bisogna tuttavia insistere sul fatto che risultati così deludenti sono dovuti alla mancanza di rigorose direttive che precisino gli aspetti che i classificatori devono considerare.

### Uno schema per continuare la ricerca

È utile che, di tanto in tanto, siano fatte ricerche sulla fedeltà di un classificatore in rapporto a se stesso o sull'accordo tra classificatori, non fosse altro che per richiamare alla co-

<sup>17</sup>Hartog e Rhodes, *An Examination of Examinations*, London, Macmillan, 1935, pp. 81 e 15.

scienza la minaccia di un pericolo. Per questo crediamo opportuno segnalare l'eccellente piano di ricerca dovuto a due dei migliori docimologi contemporanei, Yates e Pidgeon:

- sono stati costituiti sette gruppi di 50 bambini che finiscono la scuola primaria. Sono stati sottoposti a due prove di lingua materna: (a) componimento su un argomento scelto fra tre o quattro; (b) quesiti di comprensione di testi e quesiti di grammatica (uso della lingua);
- sette esperti esaminatori — cinque uomini e due donne — hanno classificato i lavori;
- ogni gruppo di bambini ha dapprima subito tre esami diversi, a una settimana d'intervallo uno dall'altro; poi, la prima delle tre prove è stata ricominciata ad una settimana d'intervallo.

Schema<sup>18</sup>.

Gruppo	Sessioni d'esame				
	I	II	III	IV	V
1	Aac	Bac	Dab	Aac	Aae
2	Bbf	Cbd	Ebc	Bbf	Bbf
3	Ccg	Dce	Fcd	Ccg	Ccg
4	Dda	Edf	Gde	Dda	Dda
5	Eeb	Feg	Aef	Eeb	Eeb
6	Ffc	Gfa	Bfg	Ffc	Ffc
7	Ggd	Agb	Cga	Ggd	Ggd

Sette serie di domande: A, B, C, D, E, F, G.

Sette esaminatori: a, b, c, d, e, f, g.

Non prendiamo in esame i risultati nei particolari; noi qui abbiamo voluto semplicemente indicare un esempio di controllo ben condotto, accompagnato, in particolare, da una misurazione dell'apprendimento.

<sup>18</sup>A cura di D. Finney, citato da Yates e Pidgeon, *Admission to Grammar Schools*, London, NFER, 1957, p. 99.

### Stereotipi ed effetti di alone

Nella situazione scolastica abituale, l'insegnante conosce ciascuno dei suoi alunni e può quindi dosare, sfumare le sue classificazioni in funzione di un effetto desiderato: da una parte si incoraggia sopravvalutando il lavoro, mentre altrove si dà prova di una eccezionale severità per dare uno choc che si spera salutare. In questi casi, l'insegnante agisce deliberatamente, in piena coscienza.

Nei fenomeni di stereotipia e di alone le cose vanno in tutt'altro modo.

#### (a) Stereotipia

Per *stereotipia*, intendiamo una inalterabilità più o meno dichiarata, che si instaura nel giudizio formulato nei confronti dell'allievo.

La stereotipia deriva da una contaminazione dei risultati. Un lavoro mediocre predispone a pensare che lo sarà anche il secondo; se questo si verifica, la tendenza a dare un voto mediocre al terzo lavoro si accresce ancora, e così di seguito. Nel professore sovraccarico di compiti da correggere, la deformazione si produce tanto più facilmente.

Si avrebbe torto a credere che la stereotipia influenzi solamente le valutazioni a base dichiaratamente soggettiva (dissertazione, composizione di storia e, in generale, risposte che implicano un giudizio di valore). Essa colpisce anche esercizi « oggettivi » come il dettato ortografico. L'esperienza seguente lo dimostra.

Un professore di lingua materna fa regolarmente dettati. Ben presto egli conosce gli alunni che abitualmente riescono meglio e quelli che riescono meno bene in questo esercizio. Se si determina la frequenza degli errori « dimenticati », non colti dal correttore, constatiamo che le dimenticanze in favore degli alunni bravi sono significativamente più elevate che quelle verso gli alunni meno preparati. Nel primo caso, l'inse-

gnante si aspetta di non trovare errori; nel secondo, li « aspetta al varco »<sup>19</sup>.

La situazione può diventare gravissima, se interviene una mancanza di coscienza professionale. Ci era stato segnalato (nell'ambito dell'insegnamento secondario) un caso di stereotipia dichiarata di cui un alunno era vittima per le versioni latine. A titolo di verifica, abbiamo fatto fare gli altri compiti successivamente da un altro alunno della classe, da un alunno dello stesso livello didattico, classificato primo in un'altra scuola, poi da un laureato in filologia classica: il voto non è cambiato di un mezzo punto su venti.

Abbiamo ripetuto questo genere di esperienze in altre occasioni. Quanti genitori, inoltre, a volte molto competenti nell'ambito della materia della quale facevano occasionalmente il compito dei loro bambini, hanno provato qualche delusione o meraviglia quando sono venuti a conoscenza del « loro » voto!

Tuttavia sarebbe ingiusto dedurne che tutti gli insegnanti accusati in tal senso manchino di coscienza e di responsabilità. Sono i metodi di valutazione più che le persone che bisogna mettere in discussione.

#### (b) Effetto di alone

L'*effetto di alone* presenta un dichiarato carattere passionale. Spesso, si sopravvalutano le risposte di un alunno di bel portamento, dallo sguardo franco, dalla dizione gradevole. Tuttavia non bisogna generalizzare. Conosciamo professori che, per anticonformismo, favoriscono una certa sciatteria o bizzarria che non sempre sono di buon gusto.

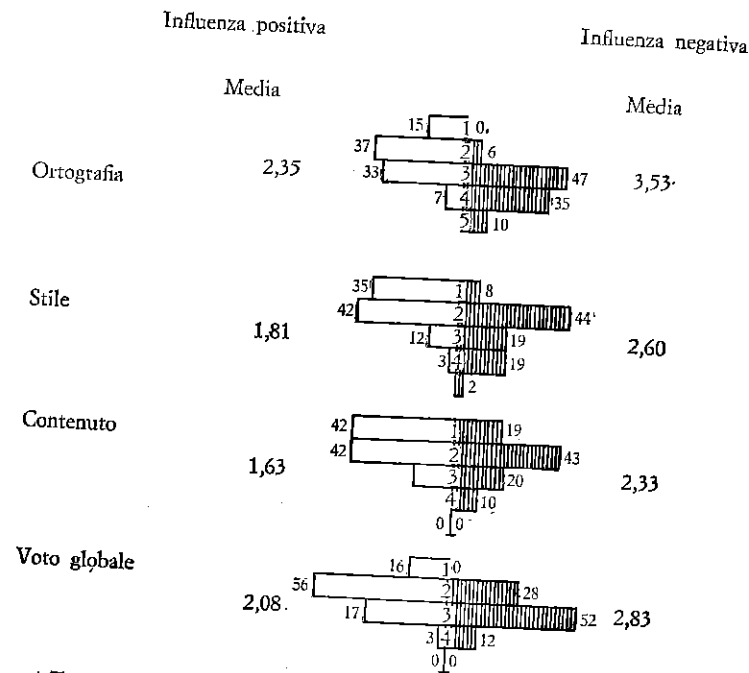
Anche la grafia può influenzare il correttore, sia per motivi di leggibilità, sia per motivi chiaramente affettivi. Gli specialisti della pubblicità sanno da molto tempo che il modo di presentare il messaggio esercita un'influenza considerevole sul suo effetto.

<sup>19</sup> Vedere a questo proposito M. Zillig, *Beliebte und unbeliebte Volksschülerinnen*, in « Arch. für die ges. Psychologie » 12 1934, p. 32, cit. da E. Hohn, *Der schlechte Schüler*, München, Piper, 1967, p. 105.

Le seguenti esperienze, dovute a Weiss<sup>20</sup>, dimostrano con quale facilità possa essere artificialmente provocata una combinazione di stereotipia e di effetto di alone.

Weiss ha selezionato due componimenti fatti da alunni di 4<sup>a</sup> elementare. I lavori sono stati dattilografati senza apportare alcuna modifica, poi sottoposti perché fossero corretti a due gruppi di 46 maestri che insegnano anche essi in 4<sup>a</sup> elementare.

Classificazione delle composizioni sotto l'influenza d'un pregiudizio favorevole o sfavorevole\*.



\* Tutte le differenze sono statisticamente significative a P.01 meno una (P.10).

<sup>20</sup> R. Weiss, *The Reliability of the Number Marking System. An Austrian Study*, in J. Lauwerys e D. Scalon, *Examinations*, London, Evans, 1969, pp. 101-107.

Il seguente commento accompagnava la distribuzione dei lavori al primo gruppo: « Ecco due composizioni scritte da alunni di 4<sup>a</sup> elementare. Il lavoro 1 è opera di un alunno medio cui piace leggere i fumetti; il padre e la madre sono impiegati. Il lavoro 2 è stato fatto da un bambino dotato; il padre è redattore di un noto quotidiano ». Per il secondo gruppo di insegnanti, i ruoli sono stati invertiti.

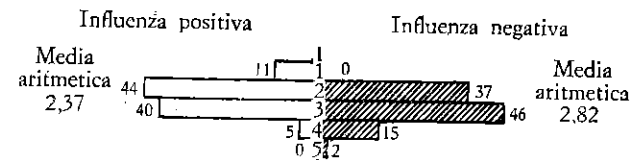
La correzione doveva esser fatta secondo una scala a cinque gradi: (1) ottimo; (2) buono; (3) medio; (4) sufficiente e (5) insufficiente. L'ortografia, lo stile, il contenuto di pensiero, poi l'insieme, dovevano essere classificati separatamente.

Per i quattro aspetti considerati, i voti attribuiti al lavoro per il quale si è creato un pregiudizio favorevole sono stati in modo significativo superiori agli altri. Per l'ortografia che sembrava dover sfuggire maggiormente all'effetto edipico della predizione, si osserva che al lavoro dell'alunno presentato come dotato, il 16% dei correttori danno la valutazione *ottimo* e nessuno la valutazione *insufficiente*; se lo stesso alunno è presentato come medio, i correttori non danno nessun *ottimo*, ma l'11% classificano *insufficiente*.

In una esperienza simile su problemi di aritmetica, sempre in 4<sup>a</sup> elementare, Weiss osserva:

Pregiudizio favorevole	Pregiudizio sfavorevole
11% ottimo	nessun ottimo
5% sufficiente	15% sufficiente
0% insufficiente	2% insufficiente

Problema d'aritmetica.



Si vede come un artificio relativamente grossolano esercita già un'influenza. Come sottolinea Weiss, la deformazione è molto probabilmente di gran lunga maggiore in una situazione reale, in cui l'insegnante conosce l'aspetto dell'alunno, la sua condotta, il suo modo di esprimersi, la professione e lo stato sociale dei genitori, ecc.

### Mancanza di validità

Gli esami tradizionali non sono solo poveri mezzi d'inventario, ma anche il loro valore prognostico sembra contestabile in molti casi.

Abbiamo osservato, assieme al « prefetto » di uno dei migliori atenei del Belgio, che due categorie di alunni riuscivano brillantemente negli studi universitari: quelli che durante tutti i loro studi sembrano aver superato la difficoltà con disinvoltura e sono rimasti i primi della classe, e una parte non trascurabile di quelli che non hanno quasi mai avuto insuccessi, ma si sono tuttavia mantenuti appena al di sopra del voto fatidico. Nei due casi, le performance negli studi secondari sono di buon auspicio, meno per la loro natura apparente che per la facilità di adattamento, per la plasticità intellettuale che ha condotto verso di esse.

Pronosticare la riuscita o l'insuccesso in facoltà universitaria sulla base della media ottenuta al termine degli studi secondari è dunque molto aleatorio.

### Uno strumento d'immobilismo sociale

L'esame tradizionale si presenta almeno sia come mezzo di selezione che come strumento di promozione dell'educazione individuale. Inoltre questa selezione si opera, in molti casi, meno in funzione della qualificazione tecnica che della qualificazione sociale.

Così l'esame diventa uno strumento privilegiato dell'immobilismo sociale, fenomeno analizzato in maniera penetrante da Passeron dal quale attingeremo l'essenziale delle seguenti considerazioni<sup>21</sup>.

#### (a) Effetti irreversibili della pagella scolastica

Il titolo scolastico, una volta ottenuto, segue l'individuo per tutta la vita, gli assicura grosso modo il suo livello nella gerarchia professionale, il suo livello di remunerazione, ed il potere che egli deterrà.

Al momento del conseguimento di un diploma, « la competenza si trova misurata non attraverso l'attività che sarà richiesta al lavoratore, ma attraverso un tipo di attività analogica, preparatoria, ludica e quasi fittizia. Ciò che misurano gli esami in rapporto alla domanda professionale, non è quello che si presume le persone sapranno fare, ma piuttosto il livello al quale si potrà esigere che la società le paghi » (p. 7).

Queste osservazioni non hanno certo valore assoluto; tuttavia corrispondono ad una realtà incontestabile. Per esempio non si attribuisce mai il titolo di ingegnere ad un operaio o ad un tecnico che ne ha acquisito la competenza funzionale.

L'ultimo esame che egli subisce con profitto, dunque, spesso fissa un individuo in una zona socio-professionale da cui, ancor oggi, gli è difficile uscire. La precocità di questa stabilizzazione è in profonda contraddizione con tutta la politica contemporanea d'educazione e dunque di promozione permanente.

Che l'esame dia più una certificazione sociale che una garanzia di competenza tecnica sarebbe d'altra parte confermato dal fatto che, in molte imprese, due persone occupano lo stesso posto: « l'una ufficialmente perché ha tutti i titoli scolastici richiesti per occuparlo (con la remunerazione, il prestigio ed il potere), l'altra di fatto, perché assume la parte tecnica del compito, cioè fa tutto il lavoro. Talvolta questa seconda per-

<sup>21</sup> J.-C. Passeron, *Sociologie des examens*, in « Éducation et gestion » 2 1970, pp. 6-16.

sona sarebbe tecnicamente più adatta del titolare a ricoprire l'incarico. Ma è proprio il certificato ottenuto all'esame che non gli consente di accedervi. Tale certificato, una volta per tutte e per tutta la vita, tende a stabilire ciò che vale la prestazione professionale di un individuo in base al suo valore sul mercato dei titoli scolastici ».

D'accapo, questa constatazione rende necessarie delle restrizioni. In una impresa, la sola competenza tecnica raramente è sufficiente ai quadri dirigenti: la capacità di comunicare, di esercitare funzioni di leader e di pubbliche relazioni, di armonizzare il lavoro di amministrazione e di produzione, ecc., sono altrettante qualità la cui manifestazione sfugge spesso a coloro che non vedono che l'aspetto tecnico delle cose. Ma queste qualità non sono sempre reali, tutt'altro. Spesso si ha a che fare con un formalismo di cui riparleremo.

(b) Gli esami non sono socialmente neutri

Un'inchiesta dell'INSEE, in Francia, mette in evidenza che se, nel corso degli ultimi anni, le possibilità di accesso agli studi superiori si sono accresciute per tutte le categorie sociali, « l'aumento globale del tasso di scolarizzazione si è distribuito tra le diverse categorie socio-professionali nella stessa proporzione che le ineguaglianze precedenti. In breve, si tratta di una semplice *traslazione verso l'alto* della struttura delle ineguaglianze » (p. 8). La tabella seguente è rivelatrice.

Categoria socio-professionale del padre	Possibilità di accesso agli studi superiori	
	1961-62	1965-66
Salariati agricoli	1,1	2,7
Agricoltori	3,4	8
Operai	1,3	3,4
Impiegati	9	16,2
Quadri dirigenti medi	24,9	35,4
Professioni liberali e quadri dirigenti superiori	38	58,7

Come si spiega questo fenomeno in apparenza contrastante con l'obbligo scolastico generalizzato per un numero di anni sempre maggiore, la gratuità degli studi e la ripartizione delle attitudini, secondo la legge del caso, in tutti gli strati della società?

Nel corso dei primi anni, la differenza d'educazione familiare creerebbe rapidamente differenze nei meccanismi intellettuali, nei mezzi linguistici e nelle attitudini di fronte al lavoro: « L'educazione data dalle famiglie predispone tanto meno alla riuscita negli esami quanto più queste famiglie appartengono ad una classe più lontana dalla cultura scolastica, dalla cultura erudita » (p. 11).

Gli esami « neutri » si limiterebbero a registrare gli effetti profondi di questa prima educazione. Questo è esatto solo in parte. La nostra opinione è meno drastica di quella di Passeron, che non riconosce abbastanza la realtà degli handicaps d'origine socio-culturali ai quali abbiamo alluso. Ma, come lui, pensiamo che gli esami non sono socialmente neutri.

« I procedimenti di classificazione e i tipi di prove utilizzati tengono conto delle attitudini tecniche almeno quanto di certi aspetti gratuiti della performance, che non hanno nessuna importanza tecnica, ma che invece sono fortemente legati alle abitudini culturali di una classe sociale piuttosto che di un'altra » (p. 12). Così, constata Passeron, i più prestigiosi esami francesi sono *meno prove di conoscenza che prove di stile o esercizi d'uso letterario del linguaggio*.

In questa prospettiva la dissertazione resta il mezzo privilegiato. Riacciandosi alla tesi sociologica di Durkheim, di Max Weber, di Clause, Passeron ricorda che il formalismo è sempre stato un mezzo di difesa delle classi privilegiate. In realtà, è sorprendente constatare che diamo la maggior importanza ad una prova, la dissertazione, che sfugge nella maniera più completa a qualunque classificazione obiettiva, tecnica...

Il formalismo consiste precisamente nel definire la cultura non per il suo contenuto obiettivo, ma per quegli imponderabili che sono lo stile, la sfumatura. Questa « indefinitezza »

permette di esercitare una *funzione di chiusura* a vantaggio di un gruppo privilegiato.

La neutralità degli esami scientifici non è del resto meglio garantita, perché anche essi spesso fanno intervenire il linguaggio, e pochi esaminatori restano indifferenti all'eleganza dell'espressione, tuttavia estranea al problema tecnico sul quale si basa fundamentalmente l'esame.

Quanto all'esame orale, esso non fa che rinforzare l'effetto della bella presentazione, del linguaggio corretto, della dizione elegante.

Per concludere, notiamo che la fioritura degli esami a carattere selettivo sembra caratteristica delle condizioni culturali del XIX secolo, in particolare dello sviluppo di una burocrazia molto gerarchizzata al servizio dell'economia capitalistica. È quindi naturale che la nuova forma di civiltà verso la quale ci evolviamo si accompagni ad una messa in discussione di un tipo d'esame concepito per altre condizioni.

### Debolezza di molte esperienze docimologiche

Dopo più di mezzo secolo, si accumulano le esperienze destinate a mostrare il disaccordo tra classificatori o la mancanza di fedeltà di uno stesso correttore.

Senza dubbio il problema esiste. Ma lo si complica inutilmente, in molti casi, quando si manca d'invitare i classificatori a mettersi d'accordo sugli aspetti da considerare e la rispettiva importanza da riservare loro. In altre parole si continua a registrare pazientemente i risultati di una politica anarchica degli esami.

### Altre critiche

Molte altre critiche vengono rivolte agli esami. Il rapporto della commissione consultiva sugli esami nell'insegnamento secondario di Gran Bretagna, depositato fin dal 1911, ce ne

fornisce una lista. Noi l'allungheremo un po', pur avendo l'impressione di non essere affatto esaurienti.

Per l'allunno:

- dedica troppa energia a ripetere le idee degli altri anziché sviluppare la sua creatività;
- è premiato per apprendimenti spesso effimeri (si è dimostrato che fino all'80% delle conoscenze, soprattutto nozionistiche, apprese per l'esame, sono scomparse quindici mesi dopo);
- impara l'obbedienza passiva alle ingiunzioni;
- la sua attitudine ad esprimersi prende talvolta il sopravvento sul contenuto (mostreremo più tardi quanto gli handicap socio-culturali si rivelano soprattutto nell'ambito dell'espressione verbale);
- spesso rinuncia ad esprimere un giudizio personale, per piegarsi alle idee del professore che valuterà l'esame;
- ciò che impara può essere viziato da uno spirito di competizione che rasenta a volte la piaggeria;
- impara a speculare sulla fortuna: nell'esame tradizionale il piccolo numero di domande riflette idiosincrasie del professore, apre la via al gioco dei suggerimenti;
- gli esami tradizionali, la loro lunga preparazione e il periodo di affaticamento che ne segue, accorciano considerevolmente l'effettivo anno scolastico;
- gli esami tradizionali impediscono il lavoro di gruppo, ed esaltano invece il valore della performance individuale, origine d'egoismo;
- la minaccia d'insuccesso che essi fanno pesare — spesso raddoppiata dalla paura di un provvedimento familiare — incita alla frode. Il successo di coloro che frodano costituisce un cattivo esempio per gli alunni rimasti onesti;
- gli esami danno una idea falsa del lavoro adulto dove, più l'argomento è complesso e difficile, più si ricorre a consigli e a opere di consultazione.

Per il professore:

- insegna in funzione dell'esame, perfino delle esigenze particolari dei membri di una commissione esterna. Vedremo che in buone condizioni, lo spirito dell'insegnamento e dell'esame deve essere unico. Qui si allude ad esami estranei agli obiettivi educativi;
  - è soffocato nella sua metodologia, se lo spirito dell'esame è estraneo agli obiettivi pedagogici che gli sembrano essenziali.
- Come praticare una pedagogia della scoperta, dell'esplorazione personale — processi lenti, ma generatori di apprendimenti profondi — se l'esame imposto è di carattere enciclopedico?
- schiavo del programma, non lascia procedere gli alunni secondo il loro ritmo e quindi digerire la materia;
  - tende a dare troppa importanza alle attitudini e alle nozioni utili all'esame.

Su questa base, si è battezzato *effetto di contraccolpo* il fenomeno per il quale i professori modificano il metodo ed il contenuto del loro insegnamento in funzione dell'evoluzione degli esami imposti dall'esterno.

Importanti aspetti educativi possono così essere trascurati e spesso lo sono stati.

Riassumendo, gli esami tradizionali presentano generalmente gravi difetti di costruzione. La loro validità è discutibile. La valutazione dei lavori è, da parte sua, aggravata da grossi difetti. Inoltre, gli esami possono nuocere alla salute fisica e mentale degli alunni. Aggiungeremo infine che, in alcuni casi, i professori svolgono il loro corso in funzione di un esame e non il contrario...

Ma è tempo di dare la parola alla difesa.

## II

### Difesa del voto soggettivo e dell'esame

Nessuno ci contesterà che si è abusato degli esami al punto da viziare l'azione educativa della scuola. Sembra anche fuor di dubbio che molti concorsi furono soprattutto sinistre lotterie o imbrogli belli e buoni. È ugualmente chiaro che certi punteggi si sono visti attribuire un significato che certamente non avevano.

Da questo a concludere che esami e concorsi debbano essere proscritti per sempre per non lasciar sussistere che serie di valutazioni occasionali e le relazioni che le sintetizzano, sembra utopistico.

Nella prima parte, abbiamo visto che i diversi modi di valutazione corrispondono a funzioni particolari. Basta riprenderne l'elenco per constatare che essi non possono essere realizzati con un solo procedimento.

Del resto, tutti i sistemi di valutazione implicano l'aver stabilito un punteggio o un voto che ci si augura perfettamente obiettivi o, meglio, perfettamente controllati: intendiamo con questo che riconosciamo agli insegnanti il diritto di adattare molte delle loro valutazioni, sia in funzione dell'alunno che della materia, a condizione che essi sappiano esattamente ciò che fanno.

Prima di esaminare alcuni vantaggi riconosciuti spesso agli esami, s'impone una necessaria discussione di carattere negativo.

### La misura rigorosa è forse impossibile

Cercare di stabilire la perfetta validità degli esami e fedeltà delle valutazioni poggia su due postulati richiamati dal Guillaumin<sup>1</sup>:

- (1) « Le produzioni degli alunni sono di loro natura misurabili, quantificabili ».
- (2) « È possibile ridurre le differenze che possiamo trovare tra le misure praticate dai classificatori ».

Se questi due postulati sono falsi, la docimologia classica crolla. Persone spiritose constatano che, benché i suoi principi fondamentali siano definiti da parecchi decenni, essa non è quasi penetrata nella scuola. Questo insuccesso sarebbe dovuto, per una buona parte, ad una incompatibilità tra la natura psicologica della situazione d'insegnamento ed il carattere matematico o fisico dei procedimenti di misura proposti.

Il contrasto non è nuovo, e la soluzione sta nel compromesso. I considerevoli progressi della psicologia e della pedagogia contemporanea sono largamente dovuti alla oggettivazione della osservazione, dunque alla misura. Tuttavia, i ricercatori, nell'ambito delle scienze umane, hanno oggi una coscienza abbastanza chiara dei loro limiti e, in particolare, dell'impossibilità quasi generale di utilizzare scale di misura matematicamente perfette. Sanno che noi per lo più non misuriamo che comportamenti isolati dell'uomo nel suo complesso. Come valutare rigorosamente entità ipotetiche come lo spirito critico, lo spirito d'osservazione, ecc., se non attraverso particolari performance che si suppongono rappresentative dell'insieme considerato?

Come non si è mai giunti, fino ad ora, a misurare con precisione il rendimento dell'insegnamento, tanto sono numerosi i punti di vista da considerare in momenti diversi, così

<sup>1</sup> J. Guillaumin, *L'aspect interpersonnel de la notation scolaire: de la docimologie à la doxologie* cit., p. 270.

non si può render conto analiticamente del valore reale di una performance scolastica complessa. Anche se, nel calcolo, il risultato non può essere che vero o falso, il processo del pensiero e lo sforzo prodotto possono variare in maniera considerevole da un soggetto all'altro e quindi può essere difficile determinarli obiettivamente in cifre. Quanto alla composizione, considerata come opera d'arte, essa sfugge alla valutazione partecellare<sup>2</sup>. Anche il diritto alla soggettività, alla reazione globale, conta, si capisce, ardenti difensori.

Se la risposta da dare a domande di geografia, di storia, di scienze naturali costringe l'alunno ad esprimersi, oralmente o per iscritto, bisogna forse fare sempre astrazione dall'eleganza del linguaggio, dal modo di esprimere il pensiero, dal rigore del ragionamento?

In breve, appena si decide di non commisurare più il successo o l'insuccesso a criteri rigidi e spesso schematici, le produzioni degli studenti sfuggono alla quantificazione automatica, impersonale.

Questo argomento non basta certamente per rinunciare ad oggettivizzare una parte della classificazione degli alunni, ma solamente una parte. Il giudizio globale dell'insegnante nella sua finezza, la sua sensibilità sia alla performance materiale sia allo sforzo di superamento e alla debolezza umana, devono conservare il loro posto.

La ricchezza dell'insegnamento consiste innanzitutto nella qualità della relazione umana che esso crea e la valutazione è uno degli aspetti di questa relazione. Se essa si impoverisce al punto da non lasciar sussistere che una comunicazione impersonale, l'insegnante può essere vantaggiosamente sostituito da una macchina per insegnare.

<sup>2</sup> Tuttavia segnaliamo che Ellis Page è riuscita a costruire un programma per calcolatore che permette di valutare automaticamente le composizioni di lingua materna. La correlazione con i voti tradizionali è elevata. Nondimeno bisognerebbe sapere se, in questo caso, il ricercatore non ha fatto semplicemente adottare dal calcolatore i procedimenti imperfetti dei classificatori.

E. Page e D. Paulus, *The Analysis of Essays by Computer*, Washington, U.S. Office of H.E.W., Project 6/1318, 1968.



### Gli insegnanti giudicano bene i loro alunni

Se è vero che gli insegnanti tendono a riportare i loro giudizi al livello medio del gruppo, la loro classificazione all'interno di questo possiede nondimeno una elevata validità.

Basterebbe dunque applicare misure che permettano di rendere questi giudizi confrontabili tra scuole (vedremo come nel capitolo sulla « moderazione ») per disporre di importanti punti di riferimento.

La sicurezza di giudizio degli insegnanti si può spiegare in base ai seguenti fattori:

- (1) essi basano il loro giudizio su una osservazione lunga e continua;
- (2) considerano un numero di fattori (particolarmente di personalità) molto maggiore di quello considerato nell'esame;
- (3) possono, in particolare, tener conto di comportamenti eccezionali (in una tale occasione, il tale alunno ha dato prova di una lucidità, di una originalità poco comuni) che probabilmente non appariranno all'esame.

Il valore predittivo dei giudizi degli insegnanti (adattati per renderli confrontabili) è stato messo bene in luce da una ricerca fatta dalla National Foundation for Educational Research in England and Wales<sup>3</sup>.

Nel 1951 e nel 1952, sono stati esaminati circa 1200 alunni della regione di Twickenham al momento della loro selezione all'entrata nella scuola secondaria (Grammar School). I loro risultati sono stati studiati due anni dopo e poi tre anni dopo.

Prove amministrare all'inizio:

#### Test

- (1) test d'intelligenza verbale (V);
- (2) test standardizzato di conoscenze di lingua materna (E1) (quesiti a scelta multipla);

<sup>3</sup> Vedere F. Yates e D. Pidgeon, *Admission to Grammar School* cit., p. 57 ss.

- (3) test standardizzato di conoscenze di aritmetica (A) (calcoli e problemi);
- (4) test d'intelligenza non verbale (N/V);
- (5) test d'attitudine spaziale (Sp.1) (spazio a due dimensioni);
- (6) test d'attitudine spaziale (Sp.2) (spazio a tre dimensioni);
- (7) test di conoscenze di lingua materna (E2) (quesiti più aperti che in E1).

Giudizio di un maestro di sesta elementare:

- (1) predice la riuscita negli studi alla Grammar School su una scala di 15 livelli (M);
- (2) classifica i suoi alunni secondo l'ordine della loro riuscita alla Grammar School. Questi giudizi sono adattati dai ricercatori in funzione di un test d'intelligenza verbale (vedere la descrizione del sistema nel capitolo sulla « moderazione ») (F).

Prove utilizzate per misurare il successo nell'insegnamento secondario: ordine di merito fornito dal prefetto degli studi, adattato in funzione dei risultati ad una batteria di test obiettivi di attitudini e di conoscenze, amministrati immediatamente dopo l'elaborazione della classificazione (SHA).

La tabella qui sotto mostra, per il gruppo del 1951, la correlazione tra ogni prova di predizione ed i risultati durante il corso degli studi:

Predittore	Gruppo del 1951: 2 anni dopo	Gruppo del 1951: 3 anni dopo
Giudizio del maestro (F)	.821	.748
(M)	.796	.722
Test verbale (V)	.789	.704
E2	.749	.623
A	.734	.659
E1	.729	.622
N/V	.648	.535
Sp.1	.565	.453
Sp.2	.491	.361

Il giudizio adattato dei maestri si rivela il migliore predittore. È anche confermato l'elevato valore prognostico del semplice test verbale.

Queste osservazioni, è necessario insistervi, non sono valide che nella situazione considerata: la riuscita nelle Grammar Schools. Non sappiamo ciò che accadrebbe in altre forme d'insegnamento. Tuttavia qui disponiamo di una osservazione strettamente controllata di cui non si ha nessun motivo di credere che non si possa applicare ad altri casi.

### Validità limitata ma reale degli esami tradizionali

Quali che fossero le imperfezioni, gli esami tradizionali non hanno impedito alla nostra civiltà di raggiungere un livello scientifico, mai uguagliato nella storia dell'umanità. Il vaglio che essi hanno operato si è dunque rivelato, almeno parzialmente, valido.

È evidente. Con la restrizione fatta nel capitolo precedente, si può affermare che gli esami tradizionali hanno permesso d'identificare gli alunni meglio dotati e i meno dotati, scelta capitale nella forma di civiltà che abbiamo conosciuto negli ultimi cento anni.

Che il procedimento sia stato grossolano (nel senso che ha ignorato o sprecato dei talenti, a detrimento soprattutto degli strati sociali svantaggiati) e ingiusto verso la massa degli alunni di livello medio, messi brutalmente tutti nello stesso fascio, non toglie niente al fatto che, in una democrazia poco evoluta, gli esami hanno avuto pure una loro funzione.

### « Temprarsi » per la vita

Secondo i behavioristi un comportamento non si apprende se non quando esso si sia prodotto effettivamente. Sembra dunque augurabile che, periodicamente, lo studente sia condotto a fare uno sforzo eccezionale, a tendere le sue energie,

ad affrontare l'ordalie degli esami, con tutte le loro imperfezioni o la loro ingiustizia. Perché la vita non gli risparmierebbe simili prove, ed è bene che vi sia preparato.

Del resto, la concorrenza, la competizione, sono tratti fondamentali della nostra civiltà. L'esame ed il concorso esistono, e si può anche desiderare che s'impongano sempre più dove ci sono più candidati che posti o impieghi disponibili, altrimenti hanno libero gioco il favoritismo in tutte le sue forme e l'inefficacia che lo accompagna.

### Definirsi in rapporto agli altri

Se una classificazione sfavorevole in una prova può dare uno choc, permette anche di confrontarsi con gli altri (non in assoluto, ma nelle condizioni dell'esame!). Colui che ha fatto del suo meglio può così meglio adattare le sue ambizioni; colui che non lo ha fatto ha l'occasione di scoprire, forse prima che sia troppo tardi, le conseguenze delle proprie deficienze.

In generale il successo giova più dell'insuccesso, ma una certa frustrazione costituisce per alcuni un utile sprone.

### Ampia sintesi e integrazione di conoscenze

È innegabile che gli esami che si basano su vasti complessi di conoscenze obbligano lo studente a costruire delle sintesi in occasione delle quali percepisce, talora per la prima volta, l'economia di tutto l'edificio, le relazioni tra le parti ed, eventualmente, i punti comuni tra diverse discipline.

### L'esame esterno controlla il professore

Un esame regionale o nazionale costituisce un mezzo di controllo del lavoro dei professori.

Se è escluso che si possano asservire gli insegnanti a programmi sovrabbondanti, fissati nei particolari, è tuttavia necessario che l'insegnamento si svolga secondo un piano di studi preciso nei suoi obiettivi e definito nelle sue materie principali.

Senza dubbio verrà un giorno in cui ogni alunno lavorerà secondo un programma individualizzato o semi-individualizzato, ma, anche in questo caso, gli scopi da raggiungere dovranno essere definiti prima di intraprendere l'azione pedagogica.

I genitori e le autorità organizzatrici avranno sempre il diritto di verificare se le finalità didattiche sono state realizzate.

L'esistenza di un controllo relativo a tutti gli elementi importanti del piano di studi sembra anche una sana salvaguardia contro il lassismo e la trascuratezza. Dedicare tutto il tempo necessario all'insegnamento è lodevole; perderne troppo è condannabile. Mi ricordo di un professore di lingue che in due anni, non aveva oltrepassato l'introduzione al corso... Evidentemente, nessun ispettore si accorse del fatto.

Infine, l'esame permette di stabilire, almeno in una certa misura, se il professore concentra troppo i suoi sforzi su alcuni alunni, a scapito di altri.

### L'esame esterno, « feedback » per il professore

Sia per la nostra salute mentale che per la miglior efficacia della nostra azione, abbiamo bisogno di essere informati della validità dei nostri comportamenti.

L'esame ben concepito permette al professore di giudicare il valore di alcuni aspetti del suo insegnamento attraverso gli apprendimenti realizzati da alcuni allievi.

Sarebbe sbagliato valutare gli insegnanti unicamente sui risultati ottenuti a breve scadenza. Tuttavia, attualmente, solo questo genere di valutazione sembra potersi effettuare con qualche rigore scientifico. Tutti quelli che esercitano una professione attinente alle scienze umane risentono, a dire il vero

a livelli molto diversi, della mancanza d'informazione sicura sulla pertinenza della loro azione. È la preoccupazione del magistrato come dell'assistente sociale, del professore come del prete.

Quando un maestro può avere la certezza di aver ben assolto la sua missione?

Riassumendo, noi non siamo sostenitori né della completa soppressione degli esami, né della intera rinuncia alla classificazione soggettiva. L'importante è adottare una maniera di fare che attinga da ogni procedimento quanto esso ha di meglio e di più sicuro. Cerchiamo di vedere come un simile scopo potrebbe essere raggiunto.

III

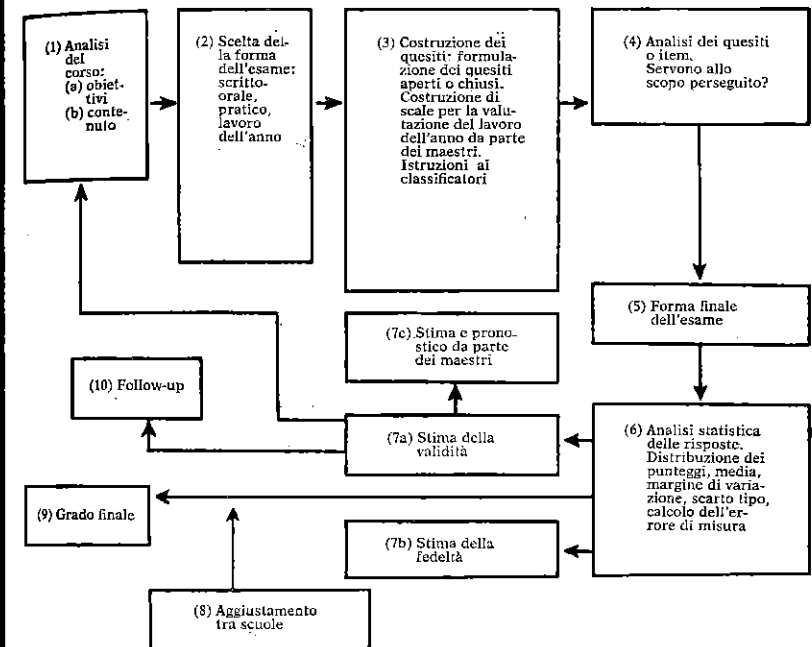
Costruzione dell'esame

Come i test standardizzati, gli esami preparati dai maestri dovrebbero essere costruiti per tappe ben definite. Certo non è possibile, per una semplice interrogazione di routine, spiegare gli stessi sforzi che per una prova destinata a centinaia di alunni. Tuttavia le esigenze di base restano le stesse.

Le grandi fasi della costruzione di un esame sono:

- (1) definire l'oggetto e gli obiettivi;
- (2) formulare i quesiti;
- (3) standardizzare la presentazione, l'amministrazione e la correzione;
- (4) tarare;
- (5) mettere a prova la fedeltà dell'esame;
- (6) stabilire la validità.

Lo schema seguente mostra la via che percorreremo.



## I

### L'oggetto e gli obiettivi

La tappa più decisiva nella costruzione di una prova è probabilmente vedere chiaramente la ragion d'essere dell'esame che si prepara e definire, con la maggior precisione possibile, i tipi d'apprendimento da valutare. Ne deriva non solamente la sua validità, ma anche il modo di valutazione da adottare!

#### L'oggetto

Tre oggetti sono possibili: il *pronostico*, l'*inventario* e la *diagnosi*. È raro che una prova scolastica si possa situare esclusivamente in una di queste tre categorie. Più spesso, gli esami mirano a fare il bilancio, l'inventario delle acquisizioni, ma hanno nello stesso tempo una funzione prognostica a breve o a medio termine (l'alunno è pronto ad affrontare un nuovo capitolo del corso? l'alunno potrà seguire con successo i corsi della classe superiore?), e una funzione diagnostica (dove si trova in difficoltà l'alunno?).

Del resto, importa sapere se l'esame si inserisce nell'azione educativa quotidiana — caso in cui l'alunno sarà confrontato con se stesso e col gruppo che lo circonda (più spesso, la sua classe) — o se serve ad informare l'individuo del suo valore relativo nella sua comunità o nel suo paese o a classificarlo in un concorso — caso in cui dovranno essere utilizzate norme adeguate.

## (A) Il pronostico

A lunga scadenza (più di un anno, almeno), il pronostico formulato a partire dai risultati di un esame scolastico è spesso deludente. I cambiamenti di insegnante, di metodo, di spirito di un corso, possono provocare nell'alunno una trasformazione profonda dell'atteggiamento. Inoltre, soprattutto nei soggetti giovani, gli interessi manifestati vanno incontro a importanti fluttuazioni. Bisogna anche tener conto dell'evoluzione psicologica degli alunni: per esempio, i professori di scuola secondaria conoscono bene l'incidenza della crisi d'adolescenza sul rendimento scolastico.

Si è dimostrato molte volte che, per il pronostico a lungo termine, i test d'intelligenza sono almeno altrettanto buoni predittori dei test di conoscenze e sono molto più sicuri dei risultati scolastici<sup>1</sup>. Questi test esplorano attitudini che, soprattutto dopo gli 8 o 9 anni, hanno una notevole stabilità<sup>2</sup>; essi riguardano il ragionamento, i fattori verbale, spaziale, numerico, ecc.

Inoltre, se le dominanti profonde della personalità, ivi comprese le zone d'interesse (e non degli interessi particolari, manifestati occasionalmente) vengono identificate, la predizione del successo scolastico e accademico può raggiungere un notevole livello di sicurezza.

A breve termine, il pronostico pedagogico si stabilisce in tre modi.

## (1) Test di maturità specifica (« readiness »)

Essi si basano sulle forme di ragionamento, sulle attitudini che intervengono in un determinato apprendimento, per esempio nell'apprendimento della lettura. La costruzione e spesso anche l'utilizzazione di questi strumenti complessi sono

<sup>1</sup> Vedere F. Hotyat (1962), W. McClelland (1945), E. Peel e D. Rutter (1951), P. Vernon (1957), F. Bacher (1965), ecc.

<sup>2</sup> Tuttavia possono ancora prodursi alcune modificazioni. Ora sappiamo che il *quoziente intellettuale* non è fissato una volta per tutte.

di competenza dello specialista. Alcuni test che permettono un primo *dépistage* sono tuttavia concepiti per i maestri, ad esempio il *test d'Inizan* per l'apprendimento della lettura.

Nei paesi di lingua francese, sarebbe augurabile che i centri di ricerca aumentassero considerevolmente i loro sforzi nel campo della maturità specifica. Abbiamo dimostrato altrove come siamo mal attrezzati per la maggior parte delle materie e a quasi tutti i livelli<sup>3</sup>.

## (2) Verifica di conoscenze-chiave o nozioni critiche necessarie alle nuove acquisizioni

Prove di questo genere possono essere costruite abbastanza facilmente dai maestri, nella misura in cui essi abbiano compilato sistematicamente la lista degli errori commessi durante una attività di insegnamento precedente<sup>4</sup>.

Anche qui, schede di ricerca che precisino i prerequisiti per affrontare i punti importanti della materia sarebbero per i maestri di utilità inestimabile.

## (3) Prova

Per determinare se un alunno è pronto ad affrontare una nuova materia, gli si presentano i primi elementi di questa materia e si osservano sistematicamente le reazioni. Una lezione programmata di tipo misto (Skinner-Crowder) costituisce spesso una felice mescolanza d'insegnamento e di controllo analitico che permette un pronostico a breve scadenza.

## (B) L'inventario (prove di rendimento)

Esso ha come primo scopo quello di misurare gli apprendimenti realizzati durante un periodo più o meno lungo. Salvo nei casi di concorsi o di *surveys* che preparano riforme, gli

<sup>3</sup> Vedere G. De Landsheere, *Les tests de connaissances*, Bruxelles, Éditest, 1965.

<sup>4</sup> Ogni nozione critica deve figurare almeno in tre esercizi; su questo argomento ritorneremo.

inventari assolvono anche, nella scuola, una funzione prognostica e diagnostica.

In realtà, sembra contrario a qualunque sano atteggiamento educativo che un maestro constatando gravi debolezze non si cerchi di localizzarne le cause per porvi rimedio e non si preoccupi delle probabili difficoltà che le debolezze stesse fanno presagire per gli ulteriori apprendimenti.

Specialmente per questo motivo, è importante distinguere accuratamente in ogni inventario le nozioni critiche, indispensabili per alcuni apprendimenti ulteriori, dalle nozioni marginali.

### (C) La diagnosi

L'oggetto delle prove diagnostiche è quello di scoprire e di spiegare le debolezze e le abitudini difettose in tutti i campi dell'apprendimento scolastico.

Senza avventurarsi nella psicologia dell'insuccesso (esso in certi casi è di competenza dello psicanalista), noteremo, con Durrell<sup>5</sup>, che le cause più comuni delle difficoltà scolastiche sono:

- (1) la mancanza di attitudini per eseguire un compito;
- (2) la imperfetta conoscenza degli elementi di base;
- (3) un insegnamento inadeguato al livello delle attitudini del bambino e alla sua prontezza d'apprendimento;
- (4) l'acquisizione di cattive abitudini che rallentino il progresso;
- (5) l'inattitudine a trasferire, a scoprire il « sistema », a generalizzare le tecniche di soluzione;
- (6) la mancanza di slancio, nell'affrontare problemi, che deriva da ripetuti insuccessi e da mancanza d'interesse.

Se sorgono difficoltà scolastiche, lo stato fisico e la personalità richiedono anch'essi la più grande attenzione. Tuttavia

<sup>5</sup> D. Durrell, *Analysis of Reading Difficulties*, New York 1955.

non si dimenticherà che molti dei problemi emozionali dell'alunno derivano dall'insuccesso anziché esserne la causa.

Nella pratica quotidiana della scuola, la diagnosi minuziosa dovrebbe essere una delle preoccupazioni dominanti dei professori e dunque occupare un tempo importante. Contrariamente a quanto una specie di mito della curva di Gauss ha talora fatto credere, l'insegnamento ideale sarebbe quello che, almeno per quanto riguarda le nozioni chiave, conducesse tutti gli alunni di un gruppo ad una performance perfetta.

Alcuni buoni test diagnostici sono stati pubblicati e i maestri non possono ignorarli. Ma più spesso ancora essi potrebbero costruire da sé strumenti analitici semplici ed efficaci, particolarmente adattati alle loro necessità.

È facile immaginare l'utilità di un quadro d'insieme nel quale, per ciascuno degli alunni di una classe, si vede apparire mediante le notazioni + o - se la risposta è stata o no corretta per tre esercizi che si basano su uno degli argomenti della materia<sup>6</sup>. Basta colorare in rosso le caselle che contengono due o tre segni *meno* per ottenere una prima impressione della situazione<sup>7</sup>. Qualche anno fa, abbiamo preparato e tenuto, con la collaborazione degli alunni, un quadro di questo genere per la pronuncia di una lingua straniera. I risultati sono stati positivi.

Perché la prova possa avere la sua funzione diagnostica, bisogna evidentemente che gli esercizi si basino su un aspetto molto preciso della materia e che siano concepiti, se possibile, in modo da non chiamare in causa che una difficoltà alla volta.

Poiché l'elaborazione di simili controlli diventa facilmente gravosa, è spesso auspicabile la collaborazione di parecchi professori.

<sup>6</sup> In una prova diagnostica si verifica almeno tre volte la conoscenza della stessa nozione, in diversi punti del test, in modo da ridurre seriamente la eventuale funzione assolta da una momentanea distrazione, da una risposta corretta data per caso, ecc.

<sup>7</sup> Per un esempio di applicazione di questo sistema vedere Bongrain, e altri, *Épreuves analytiques d'arithmétique (fin du primaire)*, Morlanwelz, Institut Supérieur de Pédagogie, 1961.



## Gli obiettivi

È strano osservare che quando si domanda alle persone che hanno l'incarico di educare i bambini di precisare i fini che perseguono, le si getta spesso nella perplessità.

P. Osterrieth, *Faire des adultes*.

Affermare che i maestri devono non solamente *istruire*, ma anche *educare*, è divenuto un truismo. È importante altresì non contentarsi di pie promesse, di vaghe preoccupazioni. Certo, si insegna ciò che si è: la nostra personalità, il nostro modo di pensare e di agire influenzano direttamente, e quasi nostro malgrado, i nostri alunni.

Ma se noi vogliamo organizzare sistematicamente la nostra azione, diventa necessaria una definizione precisa degli scopi perseguiti.

Educare, è mettere in opera i mezzi adatti a sviluppare l'intelligenza e la personalità nel senso voluto dall'ambiente culturale; senza dimenticare che anche la salute fisica esige l'attenzione dell'educatore.

Istruire significa mettere qualcuno in possesso di nuove conoscenze.

Gli *obiettivi generali* derivano dall'educazione, gli *obiettivi speciali*, dall'istruzione. Ma entrambi sono inseparabili, o almeno dovrebbero esserlo.

Evidentemente è *prima dell'inizio dell'anno* scolastico che il professore deve interrogarsi sugli obiettivi da raggiungere (altrimenti agisce alla cieca) e *quindi anche definire la materia degli esami*.

Certo, i piani di studio e i programmi scolastici ufficiali danno una parte delle risposte. Ma è chiaro che esse non appariranno che dopo un minuzioso studio e una lunga meditazione di questi documenti di base. Del resto non è raro, soprattutto nell'insegnamento secondario, che ogni materia diventi l'oggetto di un programma separato e che siano esplicitamente trattati solo gli obiettivi privilegiati. Spetta allora al professore aggiungere gli obiettivi più generali che egli

persegue in comune con i suoi colleghi delle altre materie. Se il maestro non ha una coscienza viva degli obiettivi e non li ha fatti suoi, ha ben poche possibilità di raggiungerli.

Preparare un quadro degli obiettivi da perseguire è di grande aiuto al professore nel momento in cui prepara il suo programma annuale. Questo quadro (vedere esempio a pagina 69) conterrà tante colonne quanti sono gli obiettivi generali da perseguire. Per ogni capitolo od ogni tappa del programma si indicano gli obiettivi che si desidera raggiungere e la relativa importanza da riservare loro (scala a 3 o 5 gradi).

La valutazione dell'importanza relativa dei differenti punti del programma sfugge alla rigorosa quantificazione: si tratta di una materia di base che bisogna necessariamente conoscere per affrontarne altre? si tratta di conoscenze o di capacità indispensabili per l'esercizio di una professione o per l'inserimento nella vita?

Per quanto possa essere soggettiva la valutazione dell'importanza, lo sforzo di riflessione che essa esige tuttavia porta con sé, quasi sempre, una utile chiarificazione.

### (A) Gli obiettivi generali

In fondo, l'obiettivo dell'insegnamento è unico: fare un adulto. Osterrieth ha dedicato a questo tema pagine penetranti<sup>8</sup>.

Più particolarmente i maestri devono, da una parte, almeno non nuocere alla salute fisica e mentale dei loro alunni e, per quanto possibile, aiutare a svilupparla, e, dall'altra, perseguire sistematicamente obiettivi conoscitivi ed affettivi che ora cercheremo di precisare.

Solo teoricamente si possono separare i campi conoscitivo (pensare), affettivo (essere soddisfatti o irritati; amare o rifiutare) e inerente alla volontà (volere, desiderare). Si può pensare senza provare sentimenti, agire senza pensare? Sap-

<sup>8</sup> P. A. Osterrieth, *Faire des adultes*, Bruxelles, Dessart, 1964. Vedere specialmente il primo capitolo: « Le but de l'éducation ».

priamo bene che è sempre l'organismo nella sua totalità che risponde ad una stimolazione e, per ogni comportamento, una sottile analisi dovrebbe rivelare l'importanza relativa delle componenti conoscitiva, affettiva e psicomotoria.

Il nostro proposito è ben più modesto. Attualmente siamo alla ricerca dei punti di riferimento che serviranno a delimitare insegnamento ed esami. È per questo che stiamo per separare arbitrariamente il campo conoscitivo dal campo affettivo e per dare a quest'ultima parola una accezione molto ampia.

### Gli obiettivi cognitivi

Si devono a due ricercatori americani, Bloom<sup>9</sup> e Guilford<sup>10</sup>, classificazioni gerarchizzate degli obiettivi cognitivi che, malgrado alcuni punti deboli<sup>11</sup>, si rivelano strumenti utili.

#### (a) La tassonomia di Bloom

Eccone innanzitutto le grandi articolazioni: (1) sapere a memoria; (2) capire; (3) applicare; (4) analizzare; (5) sintetizzare; (6) valutare.

Ogni dimensione deve essere intesa nel senso fissato dagli autori della tassonomia<sup>12</sup>. Del resto, anche se sapere a memoria e giudicare appaiono come il meno e il più nobile tra i comportamenti conoscitivi, questo non implica assolutamente che

<sup>9</sup> B. Bloom e altri, *Taxonomie des objectifs pédagogiques*, vol. I: *Domaine cognitif*, Montréal, «Éducation Nouvelle», 1969.

<sup>10</sup> J. P. Guilford, *Modèle tridimensionnel de l'intellect*. Una quarantina di monografie sono state pubblicate su questo argomento da Guilford e la sua équipe (University of South California, Los Angeles).

<sup>11</sup> Vedere particolarmente la critica di J. Carroll.

<sup>12</sup> Per esempio, applicare richiede un'analisi più o meno minuziosa. Non bisognerebbe dunque porre l'analisi prima dell'applicazione? Bloom e i suoi collaboratori hanno ristretto il senso di *applicazione* per sfuggire a questa difficoltà. Del resto essi indicano in molti punti della loro opera la sovrapposizione tra alcuni gradi ed analizzano, con molta accuratezza, le difficoltà di utilizzazione della *tassonomia*. Il nostro scopo non è quello di ricostruirne tutte le sfumature, bensì quello di attirare l'attenzione sullo strumento.

la memorizzazione debba sparire dal nostro insegnamento. Almeno teoricamente, bisogna aver superato ciascun grado per accedere al livello superiore. La tassonomia è stata creata per *aiutare l'educatore a non saltare* alcune tappe e per incitarlo ad elevare gradualmente il livello del suo insegnamento.

#### (1) Sapere a memoria

- (1.1) Dati particolari: denominazioni, fatti, date, simboli.
- (1.2) Modi di trattare dati particolari (senza applicarli): convenzioni, classificazioni, criteri, metodi.
- (1.3) Dati universali: principi, leggi, teorie, ecc.

La differenza fra le tre sottocategorie è più quantitativa che qualitativa: sapere a memoria, e basta, le date di nascita dei re di Francia o la tale esposizione della filosofia di Kant rende necessari soprattutto sforzi di memorizzazione differenti. Il livello di pensiero resta, nei due casi, molto basso. Quante domande d'esame, dall'insegnamento primario all'insegnamento superiore, si pongono a questo livello (Che cosa sapete di ...? Quali sono le clausole di ...? Come si prepara il ...?)!

#### (2) Capire

Si tratta qui della comprensione al livello più basso. Per una nuova formulazione del dato, l'individuo mostra che ha superato lo psittacismo, che il messaggio ha per lui un significato. Tuttavia, ancora non si tratta di applicare, quindi percepire i rapporti del dato con altri materiali, altre situazioni.

Bloom distingue due gradi:

##### (2.1) Tradurre, trasportare.

Il contenuto della comunicazione è conservato senza che il suo ordine sia modificato, ma la forma è cambiata.

Esempi: parafrasare una narrazione, una proposizione: «una *tassonomia* è una *classificazione*».

Esprimere verbalmente dei simboli matematici:

A > B significa che A è maggiore di B.

## (2.2) Interpretare.

Spiegare o riassumere una comunicazione.

L'interpretazione implica un nuovo adattamento, una nuova veduta del materiale. Essa suppone dunque la capacità di riconoscere e di cogliere le idee principali di una comunicazione e di comprendere i rapporti che esistono tra di loro. « Sotto questo aspetto, l'interpretazione diventa sinonimo dell'analisi e possiede alcune caratteristiche della valutazione »<sup>13</sup>.

Esempio: interpretare dei dati rappresentati sotto forma di tabelle o di grafici traendone deduzioni che tengono conto di relazioni tra dati, o del loro significato d'insieme.

Superare i dati e le informazioni fornite: estensione di tendenze, generalizzazione.

## (2.3) Estrapolare.

## (3) Applicare

L'applicazione suppone che il soggetto distingua i tratti comuni a due situazioni, a due problemi; si è dunque prodotta una astrazione.

Esempio: uso di metodi sperimentali per risolvere problemi di lavoro domestico.

## (4) Analizzare

## (4.1) Ricercare elementi.

Esempio: distinguere i fatti dalle ipotesi in una comunicazione.

## (4.2) Ricercare relazioni.

Esempio: sono logiche le ipotesi in rapporto alle informazioni di cui si dispone?

## (4.3) Ricercare principi di organizzazione.

Esempio: Identificare le tecniche di propaganda utilizzate in volantini.

<sup>13</sup> B. Bloom e altri, op. cit., p. 104.

## (5) Sintetizzare

## (5.1) Produrre un'opera personale.

Esempio: Narrazione avvincente di un'esperienza vissuta.

## (5.2) Elaborare un piano d'azione rispondente alle esigenze fissate.

## (5.3) Derivare un insieme di relazioni astratte. Indurre una regola.

## (6) Valutare

Giudizi qualitativi o quantitativi stabiliscono in quale misura il materiale e i metodi rispondono ai criteri.

Esempi: Scoprire il sofisma in una discussione. Apprezzare un lavoro confrontandolo con un modello.

## (b) Il modello di J. P. Guilford

Meno utilizzato, fino ad ora, della tassonomia di Bloom per la costruzione degli esami, il modello di Guilford offre forse ancor maggiori possibilità per il suo rigore.

Ecco innanzitutto come Guilford e Marrifels definiscono le tre dimensioni dell'intelletto e le loro componenti<sup>14</sup>:

## Le operazioni

Sono le attività o i principali processi intellettuali; è ciò che fa l'organismo a partire dalla materia prima delle informazioni, a partire da ciò che esso sceglie.

(1) *Cognizione*. Coscienza, apprendimento, scoperta o riscoperta, riconoscimento, comprensione d'informazioni sotto diverse forme.(2) *Memoria*. Ritenzione d'informazioni.(3) *Produzione convergente*. Generazione d'informazioni uniche, convenzionalmente accettate, a partire da un dato. Sono rispettati l'uso, la consuetudine, la regola.

<sup>14</sup> Vedere i *Reports from the Psychological Laboratory of the University of Southern California*, pubblicati dopo il giugno 1950.

- (4) *Produzione divergente*. Generazione d'informazioni varie a partire da uno stesso dato. Originalità, creatività.
- (5) *Valutazione*. Decisioni o formulazione di giudizi concernenti l'esattezza, l'adeguatezza, la desiderabilità, ... in conformità a criteri, a ideali, a obiettivi dati.

#### I contenuti

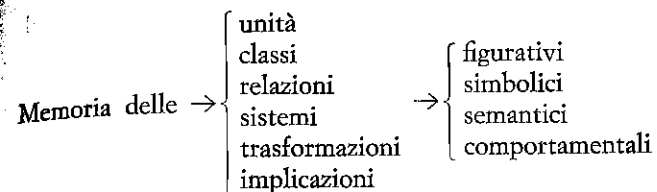
- (1) *Figurativi*. Informazione nella sua forma concreta, percepita o ricordata in immagini. È necessario un minimo di organizzazione, di strutturazione. Intelligenza pratica.
- (2) *Simbolici*. Informazioni sotto forma di segni sprovvisti di per sé di significato: lettere, numeri, note musicali. Intelligenza teorica.
- (3) *Semantici*. Informazioni sotto forma di significati legati a parole. Intelligenza verbale.
- (4) *Comportamentali*. Informazioni, essenzialmente non verbali, che intervengono nelle interazioni umane, dove hanno una funzione la percezione d'atteggiamenti, di bisogni, di desideri, d'intenzioni, di pensieri su altri e su se stessi. Intelligenza sociale.

#### I prodotti

Sono il risultato della elaborazione delle informazioni da parte dell'organismo.

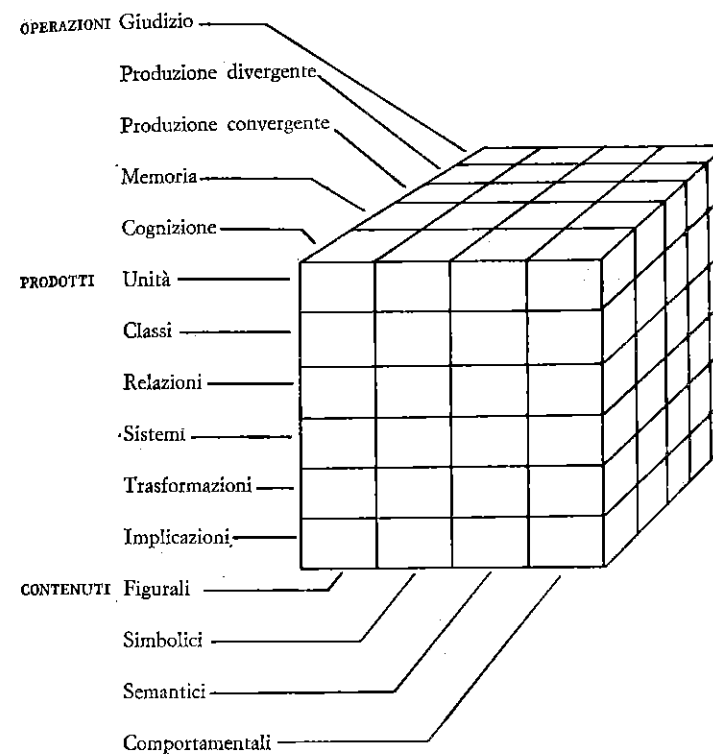
- (1) *Unità*. Parti di informazione relativamente isolate o circoscritte.
- (2) *Classi*. Unità raggruppate in ragione di loro proprietà comuni.
- (3) *Relazioni*. Connessioni riconosciute tra unità.
- (4) *Sistemi*. Raggruppamenti di unità organizzate o strutturate; complessi di parti che si trovano in interrelazione o in interazione.
- (5) *Trasformazioni*. Cambiamenti apportati in seno ad informazioni o nella loro utilizzazione.

- (6) *Implicazioni*. Estrapolazione d'informazioni: predizione, conseguenze, antecedenti. Ciascuna delle componenti delle tre dimensioni si combina con tutte le altre. Così:



ossia 24 combinazioni.

Siccome si distinguono cinque tipi di operazioni, esistono quindi, in tutto, centoventi combinazioni nel modello.



L'esperienza dimostra che è difficile tradurre tutte le combinazioni in termini utilizzabili per l'insegnamento.

In un primo tempo a richiamare l'attenzione saranno soprattutto i quattro tipi di operazioni. Idealmente, e fin dall'inizio della scolarità, il maestro dovrà sforzarsi di introdurle in tutte le attività.

- (1) La ripetizione a memoria attesta la fissazione della nozione.
- (2) La produzione convergente attesta la capacità di applicare le acquisizioni in conformità con le regole, gli usi;
- (3) La produzione divergente attesta la capacità di scoprire soluzioni o usi nuovi.
- (4) Infine la valutazione attesta la capacità di criticare, di formulare un giudizio in funzione di criteri ben definiti e di agire di conseguenza.

Troppo spesso l'attività scolastica e gli esami che la sanciscono si limitano alle prime due operazioni.

Sforzarsi di conservare un equilibrio tra le quattro categorie dei contenuti è importante tanto nell'insegnamento generale che in quello tecnico od artistico.

### Ambito affettivo

Schematizzando la sottile analisi di Osterrieth<sup>15</sup>, diremo che un individuo accede pienamente alla condizione di adulto:

- (1a) se il suo comportamento ha trovato la sua logica, la sua coerenza e se supera la volubilità;
- (1b) se ha, inoltre, acquisito una sana tolleranza al cambiamento, alla contraddizione, all'insuccesso;

<sup>15</sup> P. Osterrieth, *Faire des adultes* cit.

- (2a) se ha conquistato l'indipendenza, l'autonomia intellettuale ed affettiva;
- (2b) se è perciò capace di fare dono di sé, di restare fedele ai suoi impegni ed ai suoi sentimenti.

L'educazione è una lunga preparazione verso quest'ultimo equilibrio.

Proponiamo qui sotto un adattamento ed una interpretazione della tassonomia degli obiettivi di Kratwohl e Bloom<sup>16</sup>. Si ottengono di nuovo cinque gradi, cinque tappe che avviano dal comportamento più passivo al più attivo.

- (1) L'individuo risponde ad una stimolazione esterna

(1) È semplicemente recettivo. È una specie di stato affettivo amorfo nel quale il soggetto percepisce la bellezza o la bruttezza, i diversi sentimenti, senza reagire, un po' come uno specchio che non rimandi l'immagine.

Questo comportamento del resto è difficile da distinguere dalla semplice cognizione che precede il mandare a memoria. Si può osservare solo un certo risveglio dell'attenzione. Esempio: ascolta la musica, ascolta parlare gli altri.

(2) Riceve e reagisce. L'individuo reagisce nettamente sia obbedendo, sia manifestando piacere, con la parola, col gesto o l'atteggiamento. A questo stadio non si osserva ancora un rifiuto esplicito che dimostrerebbe una scelta deliberata.

Per il professore di letteratura, è il momento in cui gli alunni, posti di fronte a belle opere, cominciano a sentirne la grandezza, pur non avendo ancora il gusto abbastanza formato per fare una scelta personale né la sensibilità abbastanza raffinata che permetta loro di scoprirle da soli.

- (3) Riceve e reagisce accettando o rifiutando. Ora, l'indi-

<sup>16</sup> D. R. Kratwohl, B. S. Bloom, B. Masia, *Taxonomie des objectifs pédagogiques*, vol. II: *Domaine affectif*, Montréal, « Éducation nouvelle », 1970.

viduo sa ciò che vuole o ciò che ama, a condizione di essere messo in presenza di persone o di cose; egli s'impegna.

(2) L'individuo prende l'iniziativa

(4) *Prova spontaneamente a comprendere, a giudicare, ad « avvertire ».* L'individuo prova abbastanza interesse, curiosità, per istruirsi senza esservi invitato, abbastanza sensibilità per prendere una iniziativa sentimentale o, ancora, egli ha scoperto sufficientemente il senso dei valori per scegliersi una filosofia o una religione.

(5) *Agisce secondo le sue scelte.* È lo stadio psicologicamente adulto, come lo ha definito Osterrieth.

Per esempio, l'individuo vive in funzione delle sue scelte morali, sentimentali, estetiche, ma è anche capace di cambiare condotta alla luce di prove, di argomenti convincenti.

Questa ultima tappa dell'ascensione affettiva corrisponde alla valutazione nel campo conoscitivo.

## (B) Gli obiettivi speciali

Teoricamente tutte le materie, tutti i punti dei programmi scolastici offrono l'occasione di avvicinarsi agli obiettivi generali e di verificare se essi sono stati raggiunti.

Tuttavia si impongono tre osservazioni:

- (1) La pratica dimostra che una stessa prova non può riferirsi che ad un numero limitato di obiettivi.
- (2) Determinate materie si prestano meglio di altre a perseguire certi obiettivi.
- (3) Il grave problema del transfert degli apprendimenti continua a porsi in modo acuto. Per esempio, coltivare la divergenza in occasione delle attività artistiche non garantisce affatto che questa caratteristica si manifesterà nel campo scientifico o nella vita pratica.

Comunque, il primo procedimento mirante all'identificazione degli obiettivi speciali resta fundamentalmente sempre lo stesso e consiste nell'elaborazione di una tavola a doppia entrata. In alto si segnano gli obiettivi generali; di lato, si scrivono le materie del corso. Ciascuna intersezione di colonna e di riga indica, in linea di principio, un obiettivo speciale.

Esempio: Chimica.

	Conoscenza dei fatti, metodi e tecniche	Applicazione	Valutazione
Equilibri ionici			
(1) Generalità			
	(1) Grado di ionizzazione o frazione ionizzata		
	(2) Misure del grado di ionizzazione		
	(3) Applicazione della legge d'azione delle masse agli equilibri ionici		
	(4) Legge della diluizione d'Ostwald		
(2) Prodotto di solubilità			
	(1) Definizione		
	(2) Calcolo del prodotto di solubilità in funzione della solubilità		
	(3) Diminuzione della solubilità di un elettrolito poco solubile		
	(4) Precipitazione di un elettrolito per aggiunta, alla sua soluzione satura, di un elettrolito apportante un ione comune		
	(5) Dissoluzione dei precipitati		
Appendice: Applicazione del prodotto di solubilità alla precipitazione dei solfuri			

(continua)

	Conoscenza dei fatti, metodi e tecniche	Applicazione	Valutazione
(3) Prodotto ionico dell'acqua e pH			
(1) Prodotto ionico dell'acqua			
(2) Il pH			
(a) La notazione pH			
(b) pH dell'acqua pura			
(c) La scala dei pH			
(d) Determinazione del pH			
(e) Gli indicatori colorati			
(f) Metodi sperimentali di determinazione del pH			
(g) Distinzione tra elettroliti (acidi e basi) forti e deboli, con determinazione del pH di soluzioni diluite di concentrazioni conosciute			
(h) Calcolo del pH delle soluzioni di acidi e di basi a partire dalle loro concentrazioni molari e — eventualmente — dalle loro costanti di ionizzazione			
(4) Soluzioni tamponate			
(1) Introduzione sperimentale			
(2) Calcolo del pH di un miscuglio-tampone			
(5) Idrolisi			
(1) Generalità			
(2) Caso dei sali derivanti da un acido forte e da una base forte			
(3) Caso dei sali derivanti da un acido forte e da una base debole			
(4) Caso dei sali derivanti da un acido debole e da una base forte			
(5) Caso dei sali derivanti da un acido debole e da una base debole			

(continua)

	Conoscenza dei fatti, metodi e tecniche	Applicazione	Valutazione
(6) Fattori che influenzano l'idrolisi			
(6) Reciproca neutralizzazione degli acidi e delle basi			
(1) Calore di neutralizzazione			
(2) Curve di neutralizzazione			
(a) Acido forte - base forte			
(b) Acido debole - base forte			
(c) Base debole - acido forte			
(7) I metodi d'analisi quantitativa per via chimica			
(1) Metodi gravimetrici			
(2) Metodi volumetrici			
Acidi-basi			
Scelta degli indicatori			
Titolazioni redox			

Da Berger e Dighaye, *Chimie IV*, Liège, « Sciences et Lettres », 1967.

Gli argomenti di chimica che figurano nella tabella qui sopra costituiscono uno dei quattro capitoli di un corso destinato all'ultimo anno degli studi secondari. Tra molti altri obiettivi generali possibili non ne abbiamo preso in considerazione che tre.

Un semplice colpo d'occhio sull'insieme rivela che anche in queste condizioni molto semplificate il professore si trova davanti ad una scelta difficile. Su quali punti dovrà basare l'esame per ottenere una campionatura sufficiente della materia? Quale importanza relativa riserverà agli argomenti prescelti e ai diversi obiettivi speciali che vi corrispondono? In che cosa consiste esattamente la valutazione in un corso come questo?

È chiaro che il problema è tutt'altro che semplice. Supera

quasi sempre la competenza di una sola persona e sfocia in problemi che restano insoluti. È auspicabile che si creino delle commissioni di definizione di obiettivi nelle quali insegnanti, psicologi e specialisti della ricerca pedagogica uniscano i loro sforzi.

Ecco ora una serie d'item di scienze che illustrano i principali livelli della gerarchia di Bloom.

### Tassonomia di Bloom

Illustrazione: Scienze naturali. Adattato da D. G. Lewis, *Objectives in the Teaching of Science*, in « Educational Research », giugno 1965, pp. 186-199.

#### (1) Conoscenza

##### (a) Conoscenza di dati specifici.

I glucidi e i lipidi sono utilizzati dall'organismo:

- (A) come alimenti energetici,
- (B) per formare del nuovo citoplasma,
- (C) come fonte di sali minerali,
- (D) come alimenti plastici,
- (E) come fonte di vitamine.

##### (b) Conoscenza della maniera di trattare informazioni specifiche.

Equilibrate l'equazione:  $\text{H}_2\text{S} + \text{O}_2 \rightarrow \text{H}_2\text{O} + \text{SO}_2$  e contate il numero di molecole per l'insieme.

- (A) 4
- (B) 18
- (C) 5
- (D) 9
- (E) un numero diverso dai precedenti.

##### (c) Conoscenza dei principi generali.

Una macchina semplice non può:

- (A) cambiare la direzione di una forza applicata,

- (B) sostituire una grande forza ad una piccola forza,
- (C) offrire una resa del 100% tra le forze messe in gioco,
- (D) vincere una resistenza maggiore della forza applicata,
- (E) aumentare l'energia del sistema.

#### (2) Comprensione

L'alunno conosce e sa fare uso della materia che gli è stata comunicata senza poter necessariamente stabilire delle relazioni con altre materie.

Se con una temperatura freddissima, un bambino poggia la mano su un palo d'acciaio, può darsi che la mano resti attaccata, cosa che non succederebbe se il palo fosse di legno.

È perché l'acciaio:

- (A) ha un punto di fusione più elevato di quello del legno,
- (B) irradia il calore più del legno,
- (C) ha una massa specifica più elevata del legno,
- (D) è miglior conduttore di calore del legno,
- (E) possiede qualità magnetiche mentre il legno no.

#### (3) Applicazione

Astrarre da una situazione particolare e applicare l'astrazione in un altro modo.

Quando il magnesio brucia per formare ossido di magnesio,  $\text{MgO}$ , due grammi di ossigeno si combinano a tre grammi di magnesio. Se si bruciano 6 grammi di magnesio in un recipiente chiuso che contiene 3 grammi d'ossigeno:

- (A) non vi sarà reazione,
- (B) si formerà  $\text{Mg}_2\text{O}$ ,
- (C) si formerà  $\text{MgO}_2$ ,
- (D) una certa quantità di magnesio avanzerà,
- (E) l'ossido di magnesio conterrà una minor percentuale d'ossigeno.



## (4) Analisi

Se due oggetti X e Y hanno lo stesso momento (momentum), l'energia cinetica di Y non può essere superiore a quella di X che se:

- (A) Y è più pesante di X,
- (B) Y si sposta più velocemente di X,
- (C) Y ha lo stesso peso di X,
- (D) Y si sposta più lentamente di X,
- (E) Y è posto ad una distanza maggiore dall'asse di X.

## (5) Sintesi

Combinare gli elementi per formare un tutto strutturato (autore di questo item è H. Guillaume).

Un certo numero di problemi che hanno fissato l'attenzione degli uomini di scienza possono essere classificati sotto il titolo « struttura della materia ». In questo ambito, le prime ricerche si riferiscono alla struttura discontinua della materia, ipotesi emessa da Dalton nel 1805.

Un altro gruppo di ricerche ha per oggetto la teoria cinetica dei gas. Essa era apparsa già nel XVII secolo, poi fu ripresa in maniera molto più completa da Maxwell (1859) e Boltzmann (1896).

Dimostrate come la teoria cinetica dei gas costituì un argomento favorevole all'ipotesi della rappresentazione atomistica della materia.

La vostra discussione deve comprendere:

- (A) una esposizione delle ipotesi fondamentali di questa teoria;
- (B) alcune importanti deduzioni di questa teoria, applicata al caso dei gas perfetti;
- (C) l'interpretazione delle leggi sperimentali di Gay-Lussac e dell'ipotesi di Avogadro.

## (6) Valutazione

Giudizi sul valore degli argomenti che servono ad una dimostrazione o su dei criteri.

La teoria atomica di Dalton è:

- (A) un buon esempio di verità assoluta,
- (B) un buon esempio di concezione statica della scienza,
- (C) esatta ed ha permesso altre scoperte scientifiche,
- (D) parzialmente falsa, ma ha permesso altre scoperte scientifiche,
- (E) né corretta, né scorretta, e nessuno se n'è più occupato.

Ecco infine alcuni esempi di domande che si riferiscono a tre aspetti caratteristici del modello di Guilford:

Materia	Produzione convergente	Produzione divergente	Valutazione
Scienze	Spiegate perché non vi può essere vita su Mercurio	In che cosa la vita su Marte potrebbe essere diversa dalla nostra?	Pensate che vi sia una vita su Marte?
Geografia	In che cosa lo stretto di Bering ha influenzato il popolamento dell'America del Nord?	Cosa sarebbe successo se lo stretto di Bering non fosse esistito?	Attualmente, qual è, secondo voi, la principale funzione dello stretto di Bering?
Storia	Spiegate l'importanza della scoperta dell'America da parte di Colombo sulla vita economica dell'Europa.	Cosa sarebbe successo se Colombo avesse scoperto la via delle Indie invece dell'America?	Quali sono secondo voi le due conseguenze più importanti del viaggio di Colombo?
Lingua materna	Spiegate perché la novella si è sviluppata molto più negli Stati Uniti che in Europa.	Ecco l'inizio di una novella. Immaginate il maggior numero di svolgimenti possibili.	Che cosa è più importante in una novella: i caratteri o la trama?

Da J.-R. Verduin, *Conceptual Models in Teacher Education*, Washington, AACRC, 1967, p. 93.

## Formulazione dei quesiti

### Osservazioni generali

Per essere imparziale e valido, un esame deve quasi sempre comportare un gran numero di quesiti. Se il campo da coprire è vasto, la cosa appare evidente. Non interrogare che su una parte, talvolta minima, della materia, significa rispondere ad una esigenza difficilmente giustificabile, soprattutto nel nostro sistema attuale d'insegnamento, non individualizzato<sup>1</sup>: si presume che tutti gli alunni abbiano appreso tutto e si suppone omogenea la qualità di tutti i loro apprendimenti. Domande che si riferiscono arbitrariamente ad una parte del tutto permetterebbero quindi di formulare un giudizio valido.

In realtà, se, per sfortuna, il professore interroga su una delle rare parti che uno studente non ha studiato, un ventesimo della totalità, per esempio, è la catastrofe. Di frequente anche una semplice distrazione, una stanchezza passeggera, spiegano una risposta sbagliata in un campo che può anche essere ben conosciuto.

Abbiamo già visto che, per prevenire questo pericolo, i formulatori di test cercano, in genere, di porre lo stesso quesito almeno tre volte, sotto diverse forme, cosa che non succede quasi mai negli esami tradizionali.

Del resto, la molteplicità degli obiettivi da perseguire, così

<sup>1</sup> Vedere la parte v: « Il mito della curva di Gauss ».

come ci ha rivelato il modello di Guilford, rende difficilmente concepibile un piccolo numero di quesiti.

Tuttavia, non sarebbe realistico preconizzare una elaborata costruzione per il minimo esercizio di controllo o di diagnosi.

Il saper porre domande è probabilmente la capacità più necessaria ad un professore. Ma è un'arte difficile, ne sono prova gli errori nella scelta del livello di linguaggio e la poca chiarezza frequentemente riscontrate nella stesura degli esercizi di applicazione e nei quesiti d'esame.

### (A) Quesiti comprensibili

Molte risposte errate non sono dovute tanto all'ignoranza della materia, quanto a malintesi, ad una errata comprensione dei quesiti. Il rischio appare specialmente nell'ambito della matematica perché facilmente vi si trovano combinate tre difficoltà: la difficoltà inerente al problema, la difficoltà del linguaggio astratto della matematica e le debolezze linguistiche degli stessi matematici.

Ma, si sa, l'astrusità non risparmia nemmeno gli storici, i geografi, i professori di scienze ... o di lingua.

Non assicurandosi della chiarezza dei quesiti, gli insegnanti rischiano, in particolare, di commettere una ingiustizia sociale. La ricerca contemporanea conferma che molti bambini usciti da ambienti socialmente svantaggiati soffrono di gravi handicap nel campo del linguaggio. Davis e Haggard, per esempio, hanno dimostrato che basta cambiare la forma di un problema, senza cambiarne il senso, perché la differenza di riuscita tra bambini che provengono da ambienti socio-economici avvantaggiati o no passi dal 12 al 32%<sup>2</sup>.

### (B) Tener conto del livello d'informazione

Per verificare la capacità di trasferire gli apprendimenti, spesso è necessario porre i problemi in un contesto non ancora

<sup>2</sup> A. Davis, *Education for the Conservation of Human Resources*, in « Progressive Education » 27 1950, pp. 221-224.

rammentato in classe. Inoltre bisogna che questo contesto abbia un senso per l'alunno.

Certi piccoli problemi, trovati recentemente in un manuale, presupponevano la conoscenza delle regole del gioco del tennis o del modo di riempire schedine del totocalcio. Che significato avrebbero qui le insufficienze se gli alunni non hanno la possibilità di informarsi del senso reale dei problemi prima di cercare di rispondere?

#### (C) Saggiare o provare in anticipo i quesiti

Nei limiti del possibile occorre provare i quesiti prima di utilizzarli in un esame. Si scoprono così le imprecisioni, i difetti di formulazione, gli errori materiali e, anche, il livello di difficoltà.

Tale prova è difficilmente realizzabile da parte di un insegnante isolato; ancora una volta, il lavoro di gruppo offre in quest'ambito molte possibilità, essendo l'esame preparato molto tempo prima.

#### (D) Calcolo della facilità dei quesiti

Se è possibile tarare in anticipo i test su numerosi soggetti, questo calcolo si farà prima dell'esame. Se questo non è possibile sarà sempre utile procedere all'operazione, sia per capire meglio la fisionomia delle risposte, sia ai fini di un uso ulteriore.

La *percentuale di riuscita* è l'indice più semplice. Ricordiamo, tuttavia, che non è corretto calcolare percentuali a partire da piccoli numeri. Normalmente, il numero di soggetti dovrà essere superiore a cento.

Una volta noto l'indice di facilità, il professore sa meglio come dosare l'esame.

#### (E) Calcolo dell'efficacia e potere discriminante

L'indice di facilità da solo può indurre in errore, perché a volte risulta da fattori accidentali. È quindi prudente accom-

pagarlo con l'*indice di efficacia*, che rivela in quale misura un determinato quesito *discrimina* gli alunni meglio preparati dagli alunni scadenti.

#### (1) Metodo semplice.

Pidgeon e Yates propongono un metodo di calcolo semplice, molto utile per gli insegnanti:

- dividere la classe in tre gruppi: superiore, medio ed inferiore, sulla base dei punteggi totali dell'esame considerato;
- per ciascun item, vedere quale percentuale di ciascuno dei tre gruppi è riuscita;
- per ciascun item, la differenza tra la percentuale di riuscita del gruppo superiore e quella del gruppo inferiore dà una buona stima dell'efficacia.

Se la differenza è debole o nulla, significa che il quesito non distingue bene gli alunni più preparati da quelli più scadenti.

Più tardi vedremo che più la discriminazione è sottile più è ampia la distribuzione dei risultati (la curva di Gauss si appiattisce).

#### (2) Metodo più sofisticato

Ecco un altro metodo, utilizzato per l'esame di fine del ciclo secondario in Inghilterra<sup>3</sup>.

Vengono costituiti quattro gruppi A, B, C, D<sup>4</sup>.

Si danno le seguenti consegne:

- (1) determinare, per ciascuno degli item e in ognuno dei quattro gruppi, il numero di alunni che rispondono correttamente;
- (2) calcolare la media ottenuta da ognuno dei quattro gruppi per ciascun item o quesito;

<sup>3</sup> D. Mather e altri, *The CSE. A Handbook for Moderators*, London, Collins, 1965, p. 108.

<sup>4</sup> A partire da un insieme di 80 compiti, D. Mather costituisce quattro gruppi di 20: A = i primi 20; B = i 20 seguenti; ecc.

(3) riunire tutti questi dati in una tabella d'insieme e confrontare le medie per ciascun item nei quattro gruppi.

Se le quattro medie si ordinano analogamente ai punteggi medi per l'intero test, nei quattro gruppi, si può affermare che l'item considerato contribuisce alla discriminazione totale del test. Altrimenti, l'item è sospetto. Esempio:

Quesito	1	2	3	4	5	6	7	Totale
Massimo	6	6	6	6	10	8	8	50
Gruppo A	4,9	5,1	5,3	5,2	4,3	5,8	4,3	34,9
Gruppo B	3,2	4,9	5,8	4,1	4,7	5,9	3,8	32,3
Gruppo C	3,4	4,6	5,4	3	3,4	3,6	3,9	27,3
Gruppo D	2	3,4	5	2,8	2,3	4,1	2,3	21,8
Ordine dei punteggi medi	A	A	B	A	B	B	A	A
	C	B	C	B	A	A	C	B
	B	C	A	C	C	D	B	C
	D	D	D	D	D	C	D	D
Discriminatività	No	Sì	No	Sì	No	No	No	Sì

La risposta così ottenuta manca di finezza. Per gli item considerati come accettabili, non possiamo dire se essi discriminano finemente o se, al contrario, sono appena accettabili. Inoltre, non sappiamo abbastanza come migliorare la formulazione del quesito, o il modo di valutarlo.

Per ottenere informazioni più fini, gli statistici utilizzano tecniche complicate<sup>5</sup>. La seguente soluzione grafica è semplice e porta le sfumature desiderate.

Consideriamo i risultati individuali di ciascuno dei 20 alunni di ogni gruppo. Prenderemo per esempio i quesiti 2, 3 e 4. Invece di scrivere i risultati in cifre, li spuntiamo in una tabella.

Se, su tavole simili (vedi le seguenti), indichiamo il quartile

<sup>5</sup> Sebbene l'utilizzazione del calcolatore li metta oggi alla portata di tutti.

Quesito 2 — Massimo 6					Quesito 3 — Massimo 6				
Punteggio ottenuto	Gruppi				Punteggio ottenuto	Gruppi			
	A	B	C	D		A	B	C	D
6					6				
5					5				
4					4				
3					3				
2					2				
1					1				
0					0				

Quesito 4 — Massimo 6				
Punteggio ottenuto	Gruppi			
	A	B	C	D
6				
5				
4				
3				
2				
1				
0				

superiore, il medio e il quartile inferiore<sup>6</sup>, appare chiaramente il comportamento degli alunni.

Idealmente per ognuna delle misure, dovremmo osservare

<sup>6</sup> La mediana: voto di mezzo in una serie di voti ordinati; se il numero dei voti è pari si calcola la media aritmetica tra i due voti mediani 2 4 8 -9- 10 11 13.

In una scala di misura, la mediana è il punto al di sotto del quale si trovano esattamente la metà dei casi.

Il Quartile superiore o  $Q_3$  è il voto di mezzo della metà superiore della serie.

Il Quartile inferiore o  $Q_1$  è il voto di mezzo della metà inferiore.

Quesito 2					Quesito 3				
Punti	A	B	C	D	Punti	A	B	C	D
6	QS	QS			6	QS	QS	QS	QS
5	M	M	QS	QS	5	M	M	M	M
4	QI		M		4	QI	QI	QI	QI
3		QI	QI	M	3				
2				QI	2				
1					1				
0					0				

Quesito 4				
Punti	A	B	C	D
6	QS	QS		QS
5	M	M	QS	
4	QI		M	
3		QI		M
2			QI	
1				QI
0				QI

una discesa continua da A in D. La discesa è più o meno rivelata, secondo il potere discriminativo. Una salita indica che un gruppo inferiore (dopo il risultato complessivo dell'esame) ottiene migliori punteggi del gruppo superiore. Si osserva:

#### Quesito 2:

- Non vi è nessuna risalita ma una tendenza alla discesa, sia pure non molto netta.  
Item debolmente discriminativo.

#### Quesito 3:

- Nessuna discesa per il quartile superiore.
- Risalita per il medio.
- Risalita per il quartile inferiore.
- Item non discriminativo. Da rifiutare.

#### Quesito 4:

- Nessuna risalita.
- Discesa ben marcata per il medio.
- Discesa molto ben marcata per il quartile inferiore.

Buona discriminazione. Tuttavia, una riserva per il QS del gruppo D: i correttori sono stati troppo generosi per gli alunni migliori del gruppo più scadente (D).

#### Risposte aperte o chiuse?

A un quesito a *risposta aperta*, l'alunno risponde spontaneamente, utilizzando il proprio vocabolario. È detta *risposta chiusa* se il soggetto è tenuto a fare una scelta tra parecchie risposte proposte.

Tradizionalmente, la scuola si è attenuta al primo tipo di quesiti; trascurava così uno strumento di grande utilità.

#### (A) Risposte aperte (orali o scritte)

Sono le domande più naturali, quelle che noi poniamo ad ogni istante della vita.

Esse convengono particolarmente, sia come prove di controllo fatte rapidamente, in tutta spontaneità dai maestri, nel corso dell'insegnamento, sia come verifica di apprendimenti talmente complessi che sfuggono all'analisi rigorosa.

La valutazione delle capacità superiori (creatività, giudizio, spirito critico, ecc.) sembra di loro competenza. Non esclusivamente, tuttavia. In primo luogo, come vedremo fra poco, i quesiti a scelta multipla ben costruiti permettono esplorazioni molto più sottili di quanto non sembri. Inoltre perché, per

scrupolo d'obiettività e di rigore, i classificatori tendono, forse inconsciamente, a non prendere in considerazione che gli elementi più concreti, i più fattuali che esse contengono. La divergenza delle valutazioni non sarà funzione della « sottigliezza » dell'oggetto cui esse si riferiscono?

« Dobbiamo ammettere — scrive Vernon <sup>7</sup> — che gli alunni eccellenti hanno occasione di dimostrare alcune qualità eccezionali nell'esame tradizionale, e che un classificatore perspicace può, a volte, accorgersene, mentre gli altri possono benissimo penalizzare la risposta per il suo anticonformismo ». La discussione è ben lontana dall'essere chiusa.

Costruendo interamente la risposta, l'alunno tenta di provare due cose: la sua conoscenza della materia e la sua capacità di esprimerla verbalmente.

Si è creduto a lungo che le due cose andassero di pari passo. Il vecchio adagio: « Ciò che si è capito bene lo si espone anche chiaramente, e le parole per dirlo si trovano facilmente », fu verità accettata da molti maestri fino ad oggi. Tuttavia le cose non sono così semplici. La traduzione verbale del pensiero non è che una forma di espressione fra altre, disprezzate durante secoli dalle classi sociali privilegiate alle quali ripugnava qualunque lavoro manuale, qualunque manifestazione del reale.

Rinunciare interamente ai quesiti a risposta aperta sarebbe un errore in una cultura in cui domina ancora la comunicazione verbale. Ma talvolta è bene isolare, almeno parzialmente, le conoscenze rispetto alle capacità d'espressione verbale, ciò che è reso possibile dalle risposte chiuse. I professori non devono esitare a ricorrervi perché, come rilevano Pidgeon e Yates, non è stato mai dimostrato che la loro utilizzazione, anche frequente, nuocesse allo sviluppo verbale.

Una risposta formulata in tutta libertà presenta un altro grave inconveniente: il suo carattere unico mal si presta alla valutazione per comparazione con le risposte di altri individui. È questa, certamente, una delle origini principali dei disaccordi tra correttori, così spesso denunciati dalla docimologia.

<sup>7</sup> « Bull. » 4, p. 7.

Risposte chiuse. Quesiti (item) a scelta multipla

Utilità

L'esame intensivo mira a verificare in particolare la qualità delle acquisizioni; esso ha spesso uno scopo diagnostico. L'esame estensivo si basa su una materia vasta.

Nei due casi, sono necessarie numerose verifiche. Da cui il ricorso a quesiti a « risposte chiuse », soprattutto del tipo « a scelta multipla ».

Esempio: il primo test d'intelligenza utilizzabile nella pratica corrente è stato costruito da: (a) Binet, (b) Galton, (c) Goddard, (d) Spearman, (e) Terman.

(2) Costituire una scorta di quesiti

Riuscire a formulare diverse decine di quesiti precisi, per un solo esame, mette a dura prova l'immaginazione. Costituire uno schedario d'item che si arricchisce ad ogni occasione, facilita considerevolmente il compito.

Si scrive un solo quesito per scheda per poter segnare anche i « distrattori », cioè le risposte false, ma verosimili, via via che si scoprono. Questi distrattori sono forniti particolarmente dagli errori commessi di frequente dagli alunni; si è così sicuri di rispettare la loro « logica » ed è garantita la verosimiglianza.

I professori di una stessa materia possono facilmente unire i loro sforzi in questo lavoro.

Così sembra che ci si avvii verso la creazione di banche d'item, uffici centrali che possono mettere a disposizione dei maestri diverse centinaia di quesiti a scelta multipla, ben messi a punto per determinate popolazioni.

(3) Sfruttare la gamma delle possibilità logiche

Rispondere ad un item a scelta multipla non è necessariamente un semplice esercizio di memoria, tutt'altro <sup>8</sup>.

Le ricerche intraprese dai servizi dell'insegnamento superio-

<sup>8</sup> La tecnica è stata anche applicata con successo all'analisi letteraria, vedi B. Choppin e A. Purves, *A Comparison of Open-ended and Multiple Choice Items Dealing with Literary Understanding*, in « Research in the Teaching of English » 1 1969, pp. 15-24.

re in vista della riforma degli esami di medicina, in Francia, hanno ben messo in luce la ricchezza del sistema<sup>o</sup>. Scegliamo di proposito degli esempi nell'ambito degli esami universitari, nella speranza di convincere i maestri che, qualunque sia la materia che essi insegnano, il ricorso all'esame obiettivo non è escluso a priori.

Nel presente esame, si distinguono otto tipi di quesiti.

(a) *Quesito a complemento semplice*. L'item si presenta, nel suo spirito se non sempre nella sua forma, come una frase da completare.

Esempio: tra i caratteri seguenti, quello che si applica a tutti gli enzimi è:

- (A) Contengono sempre un coenzima dissociabile.
- (B) Sono termostabili.
- (C) Contengono sempre l'azoto nella loro molecola.
- (D) Contengono sempre del fosforo nella loro molecola.
- (E) Sono dializzabili.

Osserviamo che non sempre è necessario che la risposta corretta figuri tra le scelte proposte. C avrebbe potuto prendere il posto di E e diventare: « Nessuna delle risposte precedenti ».

L'item può prendere la forma negativa:

Esempio: un idrosolubile metallico ha tutti i caratteri seguenti meno uno. Indicate quale.

A — B — C — D — E

Per rispondere, il candidato deve conoscere tutti i caratteri dell'idrosol. Tuttavia, la forma negativa costringe ad una ginnastica intellettuale che si sovrappone alla difficoltà inerente alla materia.

(b) *Associazione semplice*. Essa serve a verificare la conoscenza « di un certo numero di entità che possono essere o non essere in relazione ».

Qui bisogna far corrispondere un elemento preceduto da una

<sup>o</sup> Paris, Ministère de l'Éducation Nationale. *Enseignement supérieur, examens et concours (section médicale) 1961*. Tutti gli esempi medici che seguono sono tratti da questa pubblicazione.

lettera ad un elemento preceduto da una cifra (associazione o accoppiamento).

Esempio: ramo posteriore del nervo radiale.

- (A) Segmento d'origine.
- (B) Gruppo delle diramazioni posteriori.
- (C) Gruppo delle diramazioni anteriori.
- (D) Nervo interosseo posteriore.
- (5) È applicato sulla faccia posteriore del legamento interosseo.
- (6) Contorna il collo del radio.
- (7) Innerva i muscoli dello strato superficiale della regione antibrachiale posteriore.
- (8) Passa tra i due capi del corto supinatore.
- (9) Innerva i muscoli dello strato profondo della regione antibrachiale posteriore.

(c) *Associazione composta*. È una semplice variante della precedente.

Esempio:

- (A) Malaria da *Plasmodium vivax*.
- (B) Malaria da *Plasmodium falciparum*.
- (C) I due contemporaneamente (A e B).
- (D) Nessuno dei due.

#### Quesito 1

- (1) Lo sviluppo clinico ha ogni probabilità di essere meno grave in un uomo di razza nera che in un uomo di razza bianca (risposta B).
- (2) Una associazione di primachina e di cloroquina è il trattamento elettivo per un attacco acuto (risposta A).
- (3) Gli episodi clinici sono soppressi con l'ingestione di cloroquina una volta alla settimana in zona endemica (risposta C).
- (4) Guarisce in maniera definitiva col trattamento con la cloroquina (risposta B).
- (5) La contaminazione è evitata con l'ingestione di cloroquina una volta alla settimana (risposta D).

Osserviamo che il numero di quesiti può essere qui aumentato o diminuito secondo l'importanza attribuita alla materia.

(d) *Associazione a termine escluso.* Nell'item seguente, 4 dei 5 fenomeni numerati sono comuni ad una delle turbe A, B, C. Bisogna indicare la turba A e il fenomeno che non vi corrisponde (2).

Esempio:

- (A) eosinofilia d'importanza diagnostica,
- (B) plasmocitosi d'importanza diagnostica,
- (C) linfocitosi d'importanza diagnostica,
- (1) trichinosi,
- (2) mieloma multiplo,
- (3) sindrome di Loeffler,
- (4) malattia di Hodgkin,
- (5) scistosomiiasi.

(e) *Analisi di relazione di causa ed effetto.* Esempio: l'articolazione radio-cubitale superiore permette movimenti di rotazione limitati PERCHÉ la testa del radio è circondata dal legamento anulare.

- (A) La constatazione e il motivo proposto sono tutti e due veri e tra i due esiste una relazione di causa ed effetto.
- (B) La constatazione ed il motivo proposto sono tutti e due veri e tra loro non vi è relazione di causa ed effetto.
- (C) La relazione è vera, ma il motivo proposto è falso.
- (D) La constatazione è falsa, ma il motivo proposto è un fatto o un principio accettato.
- (E) La constatazione e il motivo proposto sono tutti e due falsi.

(f) *Analisi d'osservazione.* Il complesso insieme che segue mette il candidato in una situazione confrontabile all'esperienza reale.

*Descrizione del caso:* Il malato è un uomo di 21 anni che si lamenta di malesseri, di una tosse e di febbre.

La malattia è iniziata dieci giorni prima dell'ammissione all'ospedale con un malessere ed una tosse senza espettorazione, seguiti nelle 24 ore da una temperatura che varia tra 37° 8' e 38° 3' che ha persistito fino al momento dell'ammissione.

Il quarto giorno della malattia, la tosse si accentua producendo piccole quantità di espettorazione bianca viscosa.

Tre giorni prima dell'ammissione, erano cominciati degli accessi di tosse parossistica, a volte seguiti da vomito. Furono notate delle sensazioni di brividi, ma non dei veri brividi con fremiti. Dopo il quinto giorno di malattia si manifesta un dolore parasternale che precede la tosse.

All'esame fisico, la temperatura è a 38° 3', il polso a 110, il ritmo respiratorio 32, la tensione massima 10½, minima 8.

Il malato è ben sviluppato, non magro, la sua malattia sembra acuta, è dispnoico, ma non cianotico.

L'esame fisico della gabbia toracica rivela fremiti vocali alla palpazione ed all'auscultazione. Il mormorio vescicolare è normale. Nell'ascella sinistra, si sentono alcuni rantoli sottili e la qualità bronchiale del suono è aumentata, benché di intensità normale.

La formula sanguigna è la seguente: globuli bianchi 3400 (polinucleari 30%, linfociti 62%, monociti 5%, eosinofili 3%).

La radiografia del torace rivela un aumento della densità della regione periilare con delle aree non ben definite di densità ineguale, nebulose alle due basi e in un'area polmonare superiore sinistra.

Quesiti

- (1) Qual è la diagnosi più probabile?
  - (a) tubercolosi,
  - (b) polmonite da pneumococco,
  - (c) polmonite (primaria atipica) da virus,
  - (d) coccidiomicosi,
  - (e) bronco-polmonite.
- (2) Qual è il sintomo fisico che probabilmente vi si aggiunge?
  - (a) splenomegalia,



- (b) sintomo di sofferenza meningea,  
 (c) rumore di sfregamento pleurico,  
 (d) frequenti cambiamenti nella distribuzione dei sintomi toracici,  
 (e) sintomo di condensazione lobale sinistra.
- (3) Quali dei seguenti esami di laboratorio va d'accordo con la diagnosi?
- (a) l'elevazione e l'aumento degli agglutini freddi,  
 (b) emocoltura positiva,  
 (c) marcata leucocitosi all'inizio della convalescenza,  
 (d) esame degli espettorati,  
 (e) cuti: reazione positiva.
- (4) Quale terapia dovrà essere usata?
- (a) riposo a letto e streptomina,  
 (b) riposo a letto e penicillina,  
 (c) streptomina e acido para-amino salicilico,  
 (d) riposo a letto e aureomicina,  
 (e) psicoterapia e rieducazione fisica.
- (5) Qual è l'esito probabile di questa malattia senza terapia?
- (a) la febbre sparirà completamente con una crisi terminale,  
 (b) la convalescenza sarà progressiva con una prevedibile ricaduta,  
 (c) si svilupperà un empiema residuo,  
 (d) dopo la guarigione apparirà una fibrosi residua,  
 (e) può apparire una caverna polmonare.
- (g) *Confronti quantitativi.* Confrontate X a Y:  
 X - pressione meccanica nel capillare venoso,  
 Y - pressione oncotica nel capillare venoso.
- E dite se:
- (A) X è maggiore di Y,  
 (B) Y è maggior di X,  
 (C) X è uguale ad Y.

- (b) *Relazioni.* Sia:
- (1) Il bilancio circolatorio ET.  
 (2) La quantità di calore perduta nell'unità di tempo.  
 (A) L'aumento del primo è accompagnato da un aumento del secondo, o: la diminuzione del primo è accompagnata da una diminuzione del secondo.  
 (B) L'aumento del primo è accompagnato da una diminuzione del secondo, o: la diminuzione del primo è accompagnata da un aumento del secondo.  
 (C) Le variazioni del secondo sono indipendenti dalle variazioni del primo.

(i) *Complementi riuniti.* Procedimento utilizzato quando un quesito può avere più di una risposta corretta.

Esempio: Cinque coscritti misurano: m 1,65, m 1,67, m 1,69, m 1,63 e m 1,61.

- (1) La media delle stature del campione è m 1,65.  
 (2) Lo scarto tipo è vicino a 8.  
 (3) Lo scarto tipo è vicino a 2,8.  
 (4) Il campione ha buone possibilità di informare sulla distribuzione delle stature degli individui in generale.
- (A) 1, 2 e 3 sono corrette.  
 (B) 1 e 3 sono corrette.  
 (C) 2 e 4 sono corrette.  
 (D) 4 è corretta.  
 (E) Una sola delle proposizioni 1, 2 o 3 è corretta.

Si troverà in appendice un confronto tra un esame condotto secondo il metodo tradizionale e uno secondo il metodo a scelta multipla.

(4) *Calcolo dell'efficacia dei distrattori*

In un quesito a scelta multipla, è importante verificare se i distrattori assolvono bene alla loro funzione. A questo scopo, si calcola per ogni quesito quale percentuale di scelte ha raccolto ogni possibilità di risposta.

(a) Situazione ideale

A	B	C	D	E
10%	10%	10%	60%	10%

D è la risposta buona.

Le altre scelte hanno avuto uguale attrattiva.

(b) Situazione da correggere

A	B	C	D	E
3%	0%	35%	60%	2%

Solamente il distrattore C ha realmente avuto la sua funzione. Gli altri si rivelano senza influenza.

(5) Critiche e parziale confutazione

Per quanto il suo valore sia da molto tempo provato, l'esame con quesiti a scelte multiple continua a suscitare vive critiche. Alcune possono essere confutate facilmente; altre sembrano giustificate; altre ancora non possono essere né infirmate né confermate, per mancanza di criteri scientifici.

(a) *Una obiettività ingannevole.* La classificazione dei quesiti a scelta multipla è indiscutibilmente obiettiva: le risposte corrette sono definite da prima; l'alunno o le ha trovate o no. I vantaggi di questo metodo sono evidenti.

Tuttavia la soggettività è lungi dall'essere completamente eliminata: essa sussiste, almeno in parte, nella formulazione dei quesiti e nella decisione concernente la risposta da considerare come corretta.

Prima di tutto nella formulazione dei quesiti. Essi sono frutto della riflessione, dell'invenzione degli esaminatori che, in ultima analisi, agiscono soggettivamente. Accanto ai quesiti ai quali essi hanno pensato, quanti altri, forse più validi, si sarebbero potuti immaginare?

Poi nella formulazione delle risposte considerate come ideali. Raramente si ha a che fare con delle risposte così evidentemente corrette come  $2 \times 2 = 4$ . Nel corso della costruzione di test con i nostri studenti, abbiamo più di una volta tro-

delle risposte presentate come corrette, passate vittoriosamente al vaglio dell'analisi matematica, e che tuttavia non sono, al massimo, che grossolane approssimazioni espresse in linguaggio incerto.

Come Vernon osserva, non senza malizia, la superiorità indiscutibile degli esami « obiettivi » sulle prove tradizionali è senza dubbio più spesso dovuta alla preparazione molto curata e all'unione degli sforzi e della competenza dei numerosi insegnanti e psicopedagoghi, che alla natura stessa dell'esame<sup>10</sup>.

Si dimentica anche che tra l'utopia dell'obiettività e la reale soggettività, esistono molti stadi intermedi.

(b) *Scelte « corrette » contestabili.* Anche questo punto è stato discusso in maniera eccellente da Vernon<sup>11</sup> che scrive: « Accade di frequente che persone profondamente educate accolgano i quesiti a scelta multipla con critiche quali: "Questo item è stupido" o pretendano che le scelte, considerate come "false" dai compilatori di test, siano altrettanto ammissibili, se non di più, che la risposta detta "buona". Questa critica ha origine principalmente dal fatto che l'item obiettivo non si basa sulle stesse capacità che i quesiti tradizionali. Evidentemente esami costruiti da dilettanti possono benissimo contenere cattivi item. Se ne trovano anche nelle prove costruite da professionisti, ma più raramente, perché gli item non soddisfacenti sono o rifiutati al momento del controllo preliminare, dagli specialisti della disciplina a cui gli item si riferiscono, oppure eliminati al momento dell'analisi matematica delle risposte (...). I critici rischiano anche di dimenticare che le loro reazioni sofisticate possono essere molto diverse da quelle degli alunni sui quali gli item sono stati provati. Infine, i critici possono leggere nei quesiti cose che non vengono in mente nemmeno agli alunni intelligenti ».

<sup>10</sup> Vernon, *The CSE. An Introduction to Objective-type Examinations*, London, HMSO, 1964, pp. 4-5.

<sup>11</sup> Vernon, op. cit., p. 6.

(c) *Un gioco d'azzardo*. In un quesito chiuso con risposta a due scelte, una vera e l'altra falsa, la probabilità di riuscita dell'alunno che lavora sulla base del puro caso è teoricamente del 50%. I costruttori di test lo sanno da molto tempo.

Oggi, si è ricorsi più spesso ai quesiti a cinque scelte, di cui l'una è corretta e le altre solamente verosimili (distrattori)<sup>12</sup>. In questo modo il gioco del caso è considerevolmente ridotto. Il procedimento degli alunni varierà, del resto, secondo il loro livello di conoscenza della materia. In virtù della legge della probabilità, lo studente che non sa niente non ha che una possibilità minima di far centro alla cieca sulla soluzione corretta. Invece, una conoscenza parziale permette di eliminare scientemente un certo numero di distrattori; la probabilità di riuscita nella scelta cieca tra le possibilità che restano è allora molto maggiore che nel caso precedente, ed è giusto che sia così.

Si può del resto ridurre largamente la funzione della probabilità correggendo i risultati ottenuti con l'aiuto di una formula semplice<sup>13</sup>, la quale invero penalizza generalmente in modo esagerato, poiché l'incidenza della scelta cieca varia a seconda degli alunni.

(d) *Acrobazia mentale*. Si rimprovera agli esami tradizionali di dare troppa importanza alla facilità d'espressione. Ci si può tuttavia domandare se gli item a scelta multipla un po' complicati (ne abbiamo appena trovato degli esempi) non presentino un inconveniente almeno altrettanto grave, aggiungendo alla difficoltà inerente alla materia del quesito l'obbligo di districare dubbie negazioni, di afferrare sottigliezze logiche o, più generalmente, concedendo il premio ad una certa attitudine all'astrazione a partire da dati verbali. Vernon con-

<sup>12</sup> A volte sono anche convenienti diverse scelte, ma a livelli diversi. Bisogna allora scegliere la più adeguata.

<sup>13</sup> Numero di risposte corrette —  $\frac{\text{Numero di risposte non corrette}}{\text{Numero di scelte} - 1}$

sta in ogni caso che la comprensione della lettura ha una importante funzione nella riuscita di queste prove<sup>14</sup>.

Non è importante conoscere solamente il livello di sviluppo mentale degli alunni, ma anche sapere chiaramente ciò che si vuole: verificare la conoscenza della materia, la capacità di comprensione della lettura e di ragionamento, o le varie capacità combinate insieme.

Non si vede come potrebbe procedere il sondaggio delle qualità intellettuali superiori senza un accrescimento della complessità degli item. Per questo parecchi autori pensano che i quesiti aperti continueranno ad avere un'importante funzione negli esami di livello superiore.

In caso di ricorso alla correzione automatica, si aggiunge ancora un'altra difficoltà: l'utilizzazione di schede di risposta in cui non figurano generalmente che i numeri dei quesiti e le lettere A, B, C, D, E, che rappresentano le cinque scelte. Questo sistema esige un supplemento d'attenzione da parte dell'alunno, in particolare quando non può rispondere ad alcuni quesiti e decide di saltarli. Se egli dimentica di saltare anche la linea corrispondente sulla scheda di risposta, le conseguenze possono essere disastrose. La difficoltà non deve tuttavia essere esagerata: Remmers, Gage e Rummel hanno dimostrato sperimentalmente che il sistema è utilizzabile con bambini dai 9 o 10 anni<sup>15</sup> in su.

In tutti i modi, gli alunni devono essere accuratamente allenati alla tecnica degli esami a scelta multipla prima di subire una prova decisiva. Mediante questa precauzione ed una costruzione rigorosa, i quesiti a scelta multipla danno buoni risultati; ne è prova sufficiente la loro utilizzazione intensiva nei paesi anglosassoni da numerosi anni, e l'utilizzazione sempre più diffusa altrove.

(e) *Inconvenienti « incerti »*. Qualifichiamo i rilievi seguenti come « incerti » perché, a quanto ne sappiamo, nessuna

<sup>14</sup> P. Vernon, *The Determinants of Reading Comprehension*, in « Educational Psychological Measurement » 22 1962, pp. 269-286.

<sup>15</sup> H. Remmers, N. Gage e G. Rummel, *Educational Measurement and Evaluation*, New York, Harper, 1955, p. 246.

rigorosa ricerca scientifica ne ha stabilito il fondamento. Sembrava che alcune critiche non si giustificino che nella misura in cui l'esame è mal costruito.

- Scegliere la risposta buona tra le altre è più facile di costruirla. La memoria assume il ruolo principale. Rischiamo di instaurarci una certa pigrizia intellettuale, una ripugnanza allo sforzo necessario alla formulazione chiara del pensiero.
- Sapendo che per subire l'esame in lingua materna basta sottolineare o indicare la risposta scelta con una crocetta, i maestri trascureranno gli esercizi di espressione.
- La preferenza andrà a numerosi piccoli esercizi da fare in un tempo molto breve; si trascureranno così i problemi che esigono una riflessione molto lunga. Ora, gli studenti superiori e la vita reale non ci risparmiano questo genere di prove.

(C) Come conclusione: un compromesso

Ci sembra che si imponga un compromesso almeno a due livelli.

Poiché quesiti aperti e quesiti chiusi sembrano possedere ciascuno particolari vantaggi e sembrano mettere in causa diversi apprendimenti, non vi è alcuna ragione perché uno dei due abbia il monopolio. Possono benissimo coesistere. Tuttavia i migliori docimologi reputano che non sia auspicabile mescolare in una stessa prova i due tipi di quesiti.

Del resto una formula intermedia tra quesiti aperti e chiusi esiste e forse s'imporrà sempre più, secondo i progressi della ricerca pedagogica.

L'esame consisterebbe in un numero molto elevato di quesiti *aperti*, di portata relativamente limitata, che siano stati anche provati praticamente. La gamma delle probabili risposte sarebbe quindi conosciuta prima, ciò che permetterebbe di proporre uno schema di classificazione che si avvicina moltissimo al rigore del quesito a risposta chiusa.

Questa possibilità è già confermata da diversi lavori spe-

imentali. Così, correlazioni quasi perfette (.98 e .99) si sono ottenute tra sette gruppi di classificatori che rappresentano ciascuno una commissione d'esami in Gran Bretagna. L'esperienza è stata fatta su prove di storia di livello superiore del GCE (fine del livello secondario)<sup>16</sup>. Senza andare così lontano, dei professori possono intendersi perfettamente su dei punti da esigere nelle risposte a quesiti aperti.

Del resto, la maniera di formulare i quesiti può chiudere parzialmente delle risposte e quindi diminuire il gioco della soggettività. Un quesito come: Scrivete un inizio appropriato alla frase: «... quando comincio a piovere», è, in qualche modo, a mezza strada tra la prova soggettiva e la prova oggettiva. Perché qui saranno rari i casi di disaccordo tra classificatori invitati a pronunciarsi sulla correttezza della risposta.

<sup>16</sup> Joint Matriculation Board, *The Marking of Scripts in Advanced Level History*, Universities of Manchester, Liverpool, Leeds, Sheffield and Birmingham, 1964.

## Un preambolo indispensabile: la curva di Gauss

È impossibile discutere sulla valutazione dei risultati senza valersi di alcune nozioni matematiche. Non dimentichiamo la promessa fatta all'inizio di quest'opera: sarà sufficiente l'aritmetica elementare!

## (A) La curva di Gauss, immagine della probabilità

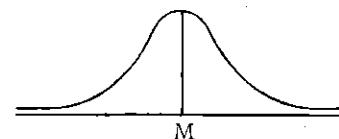
In un sacco vi sono due palline, simili in tutto e per tutto, ma una è rossa e l'altra bianca.

In queste condizioni ogni pallina ha una possibilità su due di uscire ad ogni estrazione alla cieca.

Se noi facciamo questo gioco un piccolissimo numero di volte, è possibile che estraiamo lo stesso colore più volte consecutive. Man mano che aumentiamo il numero delle estrazioni, questa probabilità diminuisce.

Immaginiamo 100 estrazioni consecutive. È molto probabile che il rosso uscirà pressappoco tante volte quante il bianco. Invece, è molto poco probabile che si estrarrà 100 volte consecutivamente la pallina rossa o la pallina bianca. È già lievemente più probabile che si estrarrà 99 volte la rossa ed una volta la bianca, ancora più probabile 98 volte la rossa e 2 volte la bianca, ecc. In breve, la probabilità aumenterà fino a 50 volte la rossa e 50 volte la bianca, per poi diminuire progressivamente per arrivare a 1 volta la rossa e 99 volte la bianca.

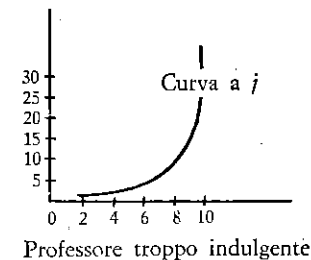
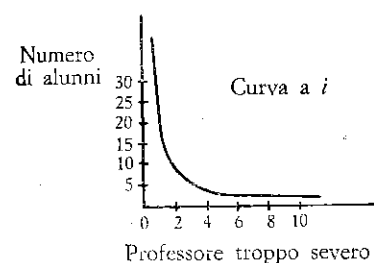
In un gran numero di estrazioni, questo movimento ascendente e discendente corrisponde ad una curva che assume la forma di una campana: è la famosa curva di Gauss, le cui due metà sono simmetriche in rapporto alla media aritmetica, e le cui estremità non toccano mai la linea dello zero, dato che la probabilità nulla non esiste che all'infinito.



Questa *distribuzione*, detta *normale*, è fatta a immagine di molte qualità umane, così come esse si distribuiscono nei gruppi numerosi, *presi a caso*. Così tra gli uomini di una grande città sono più numerosi gli individui di statura media, mentre i giganti e i nani sono molto rari; tra questi due estremi la popolazione si distribuisce secondo la curva di Gauss.

Anche l'errore spesso si sottopone alla legge normale: se si fa un grandissimo numero di misure, si vedrà molto probabilmente apparire un errore di grandezza media, ed una distribuzione che va dall'errore minimo all'errore massimo.

Appena il caso non interviene più, la distribuzione si modifica. Un professore può, per esempio, dare in grandissima maggioranza cattivi voti e pochissimi voti buoni, o il contrario. È possibile che si ottenga allora una delle due curve caratteristiche seguenti:



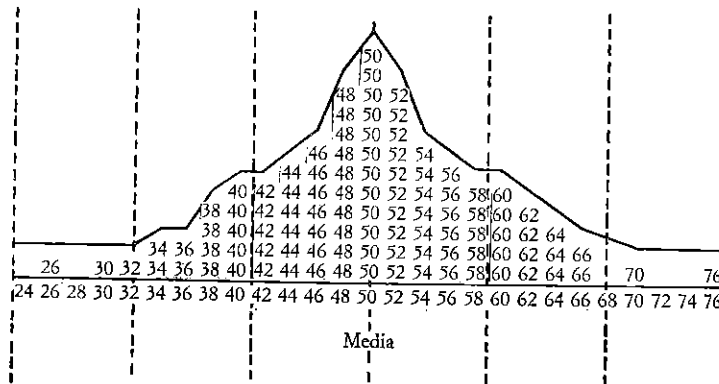
(B) La curva di Gauss, rappresentazione dei risultati dell'insegnamento non individualizzato

Un professore che insegna in modo non individualizzato in una classe dove gli alunni non sono particolarmente selezionati, tiene normalmente la sua lezione ad un livello di difficoltà adeguato alla maggioranza del gruppo. Se l'adattamento è corretto, vi saranno dunque molti risultati medi, pochi ottimi e pochi pessimi. La distribuzione di questi risultati si avvicinerà alla curva di Gauss.

Uguale fenomeno si produce, ancora più spontaneamente, negli esercizi che sfuggono alla quantificazione rigorosa perché mettono in gioco un complesso insieme di fattori.

Qui sotto sono stati disposti 100 risultati (immaginari) ottenuti ad un esame, esso stesso classificato in centesimi. Si constata che 12 alunni hanno ottenuto 50 su 100 (voto medio), mentre uno solo ottiene il voto più basso (26) e uno solo il più alto (76).

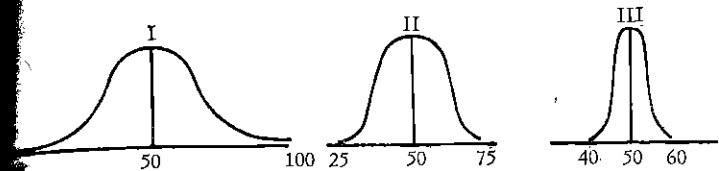
Si osserva anche che la linea corrispondente alla ripartizione dei risultati non è una curva a campana perfetta, ma solo il suo abbozzo. In un caso simile (daremo presto alcune precisazioni in proposito) si suppone che se il numero di soggetti fosse stato più elevato la curva sarebbe venuta più smussata e si considera la ripartizione come « normale ».



Lo scarto tipo o sigma, indice prezioso

(1) Significato

La variazione dei 100 risultati che abbiamo or ora esaminato avrebbe potuto essere maggiore, per esempio da 0 a 100, o minore, per esempio da 40 a 60, pur distribuendosi sempre nella curva di Gauss. Avremmo potuto avere:



Nei tre casi, la media è 50. Tuttavia si tratta di situazioni molto diverse.

Il margine di variazione dei risultati è:

- curva I :  $100 - 0 = 100$
- curva II :  $75 - 25 = 50$
- curva III:  $60 - 40 = 20$

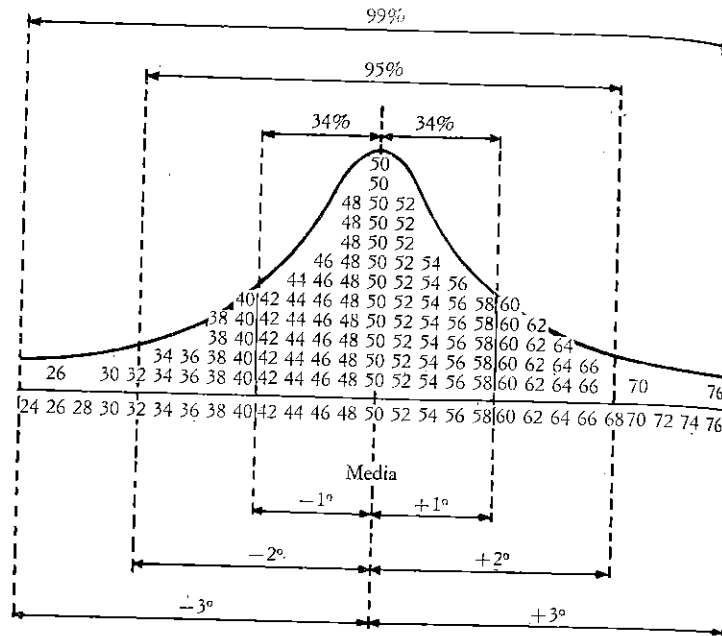
*Sigma* o  $\sigma$  è un indice facile da calcolare (vedere più giù il metodo di calcolo), che ci indica immediatamente come variano i risultati in rapporto alla media aritmetica dei voti ottenuti.

Per i 100 voti presi come esempio,  $\sigma = 9$ . Se si prende un  $\sigma$  prima e dopo la media, si ottiene: 41 e 59. Contate il numero di risultati che cadono tra questi due limiti, otterrete 68 voti, ossia qui il 68% di voti di tutto il gruppo.

È sempre così in una *distribuzione normale*; il  $\sigma$  indica sempre la stessa proporzione dei risultati in rapporto all'insieme:

- 1  $\sigma$  da una parte e dall'altra della media = 68% dei voti
- 2  $\sigma$  da una parte e dall'altra della media = 95% dei voti
- 3  $\sigma$  da una parte e dall'altra della media = 99% dei voti.

Vediamo ora la distribuzione dei risultati con nuovi occhi



Siccome ad un dato sigma corrisponde sempre la stessa superficie dell'area determinata dalla curva, cioè la stessa proporzione dei risultati, questo indice ci fornisce il modo di confrontare i risultati attribuiti a professori diversi, in quanto questi risultati si ripartiscono normalmente. Mostriamo subito come.

(2) Rapida stima della media e del sigma

(a) *Problema.* Il calcolo rigoroso della media e dello scarto tipo (sigma) è relativamente faticoso. Inoltre quando la distribuzione non è normale la media aritmetica dà un'idea falsa della tendenza centrale.

Un procedimento semplice permette di evitare lunghi calcoli. Esso non fornisce che risultati *approssimativi*. Tuttavia

essi sono sufficienti in quasi tutti i casi della pratica corrente nella scuola<sup>1</sup>.

Si considerino due quesiti d'esame ai quali 10 alunni hanno risposto. Ecco i voti attribuiti, le medie e gli scarti tipo calcolati col metodo classico.

Alunno	Quesito 1	Quesito 2
1	39	32
2	33	28
3	25	32
4	22	28
5	26	27
6	18	31
7	23	33
8	13	27
9	57	35
10	45	36
	301	309
	Media aritmetica: 30	Media aritmetica: 31
	Scarto tipo: $\sigma = \sqrt{\frac{\sum x^2}{N}} = \pm 13$	Scarto tipo: $\sigma = \pm 3$

(b) *Metodo semplice di calcolo.* Rappresentazione grafica dei risultati.

Margini di variazione:

quesito 1:  $57 - 13 = 44$

quesito 2:  $36 - 27 = 9$

I voti del quesito 1 sono molto più dispersi. Essi influenzano più la classificazione finale che quella del quesito 2. Spesso è utile ignorare i due voti estremi.

In questo caso:

margine: quesito 1 = 27; quesito 2 = 8;

mediana = voti del mezzo.

<sup>1</sup> Secondo D. R. Mather e altri, op. cit.

Voti	Quesito 1	Quesito 2
60		
58	.	
56		
54		
52		
50		
48		
46	.	
44		
42		
40		
38	.	:
36		
34	.	.
32		..
30		
28		:
26	:	:
24		
22	:	
20		
18	.	
16		
14	.	
12		
10		

Qui, cifre pari, dunque media tra 5° e 6°.

$$\text{Quesito 1: } (25 + 26) : 2 = 25,5.$$

$$\text{Quesito 2: } (31 + 32) : 2 = 31,5.$$

Per il quesito 1, la differenza abbastanza netta tra la media aritmetica e la mediana (30 e 25,5) indica una distribuzione asimmetrica dei voti (qui: la maggioranza sotto la mediana).

Per il quesito 2, la simmetria è buona (media 31 e mediana 31,5).

Osserviamo che, in una distribuzione perfettamente normale, la media e la mediana coincidono: vedere il nostro esempio.

Stima dello scarto tipo ( $\sigma$ ).

$\sigma = \frac{3}{4}$  dello scarto interquartile.

Dunque:

(a) Cercare il quartile superiore, cioè la fascia dei voti al di sopra della mediana.

$$\text{Quesito 1} = 39$$

$$\text{Quesito 2} = 33$$

(b) Cercare il quartile inferiore.

$$\text{Quesito 1} = 22$$

$$\text{Quesito 2} = 28$$

(c) Scarto interquartile.

$$\text{Quesito 1} = 39 - 22 = 17$$

$$\text{Quesito 2} = 33 - 28 = 5$$

(d) Stima del  $\sigma$ :

$$\text{Quesito 1} = 17 \times \frac{3}{4} = 12\frac{3}{4}$$

$$\text{Quesito 2} = 5 \times \frac{3}{4} = 3\frac{3}{4}$$

(D) La concentrazione dei risultati attorno alla media

Rifacciamoci alla distribuzione dei 100 voti presi come esempio. Osserviamo che, più ci avviciniamo alla media, più i voti sono numerosi:

12 voti su 100 alla media esatta,

22 voti su 100 tra 48 e 52,

68 voti su 100 tra 41 ( $-1\sigma$ ) e 59 ( $+1\sigma$ ).

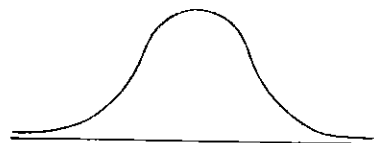
Immaginiamo che si tratti di composizioni francesi. Nonostante il professore possa dare voti da 0 a 100, quasi un quarto dei voti si situano tra 48 e 52, ossia a quattro punti di distanza.

Che cosa significa una differenza di quattro punti su 100 in composizione francese? È quasi certo che, se facciamo giudicare lo stesso lavoro da dieci professori diversi, gli scarti



saranno quasi sempre superiori a 2. Ora, se è la media che costituisce il limite dell'insuccesso (qui essa coincide con la metà della scala totale), si può dire senza timore di sbagliarsi che, per il 22% del gruppo di alunni presi come esempio, l'insuccesso o la riuscita dipenderà dal puro caso.

Ecco, in sfumato, una rappresentazione della densità di popolazione in rapporto alla media:



(E) Curva di Gauss voluta dagli insegnanti

La curva di Gauss è un eccellente strumento di classificazione poiché permette di identificare i migliori (per esempio tra 25 e 35) e i più scadenti.

Nel caso di un concorso in cui numerosi candidati aspirano ad un piccolo numero di posti, si può, per esempio, scegliere tra quelli che si situano tra  $+1\sigma$  e  $+3\sigma$  (parte che sfugge alla forte concentrazione attorno alla media). Analogamente, se si desidera classificare gli alunni — ciò che si è fatto nel passato troppo esclusivamente, ma che resta pur sempre utile per permettere a ciascuno di situarsi al proprio posto —, la distribuzione *normale* è certamente utile.

Dato che questa volta non si tratta più di gruppi scelti a caso ma di soggetti particolarmente allenati, il professore *crea artificialmente* le condizioni necessarie ad una ripartizione gaussiana. Come? Dosando degli item secondo i loro indici di difficoltà (e di efficacia).

(F) Come sapere se una distribuzione è normale?

Tutto quanto abbiamo or ora detto a proposito della curva di Gauss e del sigma vale solo nel caso che i risultati si distribuiscano *normalmente*. Prima di iniziare qualunque ope-

razione è necessario dunque verificare se questa condizione è stata soddisfatta.

Esistono rigorosi processi matematici per raggiungere questo scopo, ma, ancora una volta, ci contenteremo qui di una semplice approssimazione grafica: l'*istogramma*.

Il procedimento è semplice.

(1) *Ordinare i risultati*. Li si può ordinare dal più grande al più piccolo, o viceversa.

Ecco ciò che si ottiene classificando i 100 risultati che abbiamo già preso come esempi:

Voti	Numero degli alunni che hanno ottenuto questi voti (frequenze) (f)
26	1
30	1
32	1
34	2
36	2
38	4
40	5
42	5
44	6
46	7
48	10
50	12
52	10
54	7
56	6
58	5
60	5
62	4
64	3
66	2
70	1
76	1

N = 100

Ma, se i voti sono numerosi, questo processo non è né rapido né pratico. È molto più agevole costruire delle classi:

- Calcolare il margine di variazione tra i due voti estremi:  
 $76 - 26 = 50$ ;
- dividere questo margine per 15\*:  $50 : 15 = 3,33$ ;
- scegliere come intervallo di classe uno dei due numeri dispari più vicini: 3 o 5. Qui è stato scelto 5, visto il piccolo numero di voti;
- porre il voto superiore a metà dell'intervallo superiore: voto superiore: 76; la classe superiore è dunque: 74-75-76-77-78.

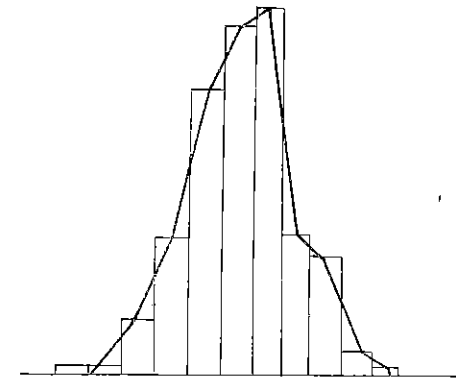
Classi		Frequenze (f)	
1	74-78		1
2	69-73		1
3	64-68		5
4	59-63		9
5	54-58		18
6	49-53		22
7	44-48		23
8	39-43		10
9	34-38		8
10	29-33		2
11	24-28		1

11 classi    Intervallo (i) = 5    Numero dei voti (N) = 100

(2) *Disegnare l'istogramma.* Nella parte « frequenze » della tabella qui sopra, il controllo è stato fatto con molta cura. Basta tenere la pagina orizzontalmente per veder apparire lo scheletro del grafico chiamato istogramma: non rimane che tracciare i rettangoli.

\* Questo numero è arbitrario. La pratica dimostra che, nella maggior parte dei casi, porta ad una buona ripartizione.

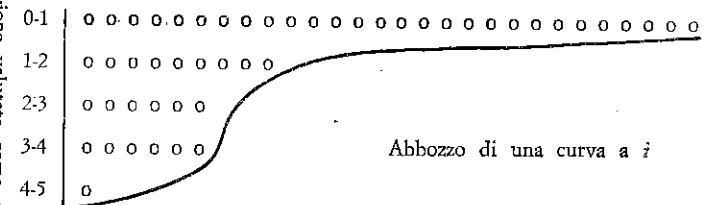
Se si congiungono i punti medi delle basi superiori, si ottiene una linea (poligono di frequenze) che qui è un abbozzo di cui ci si accontenterà per considerare che la distribuzione è molto probabilmente normale. Perché? Perché il tracciato

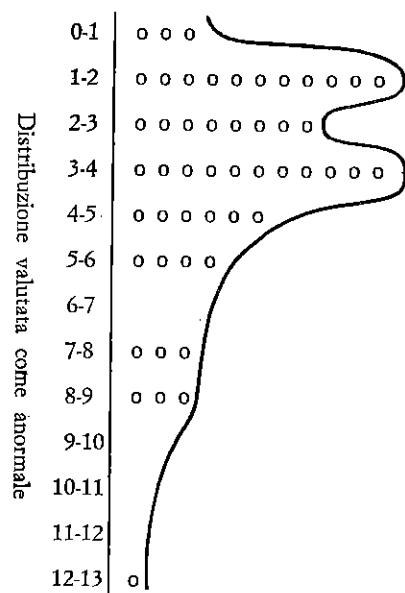


non ha che un *vertice* pressappoco *centrato*, a partire dal quale le quantità *decregono abbastanza regolarmente da una parte e dall'altra.*

Invece le due distribuzioni qua sotto non sarebbero considerate come *normali.*

Distribuzione valutata come anormale





Abbozzo di una curva de-centrata e a due vertici (curva bimodale)

## La classificazione soggettiva: la scala di valutazione

### (A) Introduzione

Per classificazione soggettiva intendiamo ogni valutazione fatta dai maestri, in maniera globale, sia immediatamente al momento dell'osservazione della performance degli alunni, sia dopo aver avuto una visione d'insieme su punteggi particolari. Più in generale, è soggettivo qualunque voto al quale si arriva con un procedimento non automatizzabile.

Il grado di soggettività varia considerevolmente. Ora procede quasi senza limiti (per esempio: «bene, benissimo...»); ora cede quasi il passo al procedimento obiettivo (esempio: scala i cui gradini sono definiti con una precisione che non lascia praticamente più posto all'interpretazione personale).

In ogni caso la valutazione soggettiva si fa tenendo presenti dei criteri esterni o interni. Il classificatore si pronuncia per

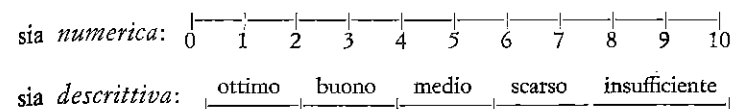
l'uguaglianza o l'inuguaglianza in rapporto al criterio. Opera quindi una classificazione.

Per questo possiamo limitare questa parte allo studio delle scale di valutazione.

La scala di valutazione aiuta a *ordinare* individui od oggetti in rapporto ad una qualità che essi possiedono ad un grado più o meno elevato.

Essa rappresenta un continuum che può teoricamente andare da un minimum assoluto ad un maximum assoluto, ma che, praticamente, quasi sempre, non è altro che una frazione, un segmento di questo assoluto. Dall'eccellenza... all'insufficienza, dal piccolissimo... al grandissimo, ecc.

La scala di valutazione, grafica o no, è:



La scala di valutazione è lo strumento usato più di frequente a scuola... e quello peggio conosciuto. Si può dire senza timore di sbagliarsi che essa ha letteralmente viziato una gran parte degli esami, a tutti i livelli della scolarità, e probabilmente crea guai ad un numero considerevole di individui.

Per questo è necessario assolutamente conoscerne la vera natura e soprattutto le gravissime limitazioni.

### (B) Natura e manchevolezze delle scale di valutazione

La scala di valutazione è *una scala ordinale*.

(1) *Essa non ha né zero né maximum naturali o assoluti*; comincia e finisce non importa dove, a discrezione di chi la utilizza. Di conseguenza, coloro che giudicano, anche se valutano uno stesso fenomeno, dallo stesso punto di vista, non hanno mai scale rigorosamente uguali; le differenze possono essere considerevoli, e spesso lo sono.

Esempio: *Conoscenza dell'inglese come seconda lingua.*

Teoricamente lo zero potrebbe essere il momento in cui l'alunno non conosce ancora una sola parola della lingua. Tuttavia, a scuola si giudica raramente a partire da questo punto. Quanto alla conoscenza totale, assoluta, essa evidentemente non esiste.

Immaginiamo due professori d'inglese che giudicano, senza mettersi d'accordo, due classi in cui ogni alunno ha fatto una breve esposizione orale. Molto probabilmente, i due estremi delle scale di valutazione, utilizzate indipendentemente dai due professori, saranno meglio definite dall'alunno giudicato più preparato e dall'alunno giudicato più scadente, nel suo gruppo, che da una valutazione astratta della quantità d'inglese che dovrebbe essere conosciuta al livello pedagogico considerato.

(2) *I gradi non sono alla stessa distanza all'interno di una stessa scala.*

Per esempio, è impossibile dimostrare che la distanza che separa la buona dalla ottima conoscenza dell'inglese è uguale alla distanza che separa la conoscenza media dalla buona.

Se i gradi sono: A, B, C, D, E, non si può dunque dire che  $A - B = B - C = 0$ .

(3) *Gradi corrispondenti da una scala all'altra (per una stessa materia) non sono della stessa grandezza.*

Buono in inglese, per un professore, non è quantitativamente uguale a *buono* in inglese per un altro: niente può stabilire tale uguaglianza in maniera matematica.

(4) *Scale che vertono su materie diverse non sono confrontabili.*

Per gli alunni di una stessa classe della scuola elementare in cui un solo maestro insegna tutte le materie, avere *buono* in matematica non ha lo stesso significato che avere *buono* in francese. I gradi non sono certo confrontabili sul piano quantitativo.

Da tutto questo risulta che le scale di valutazione non permettono nessuna operazione aritmetica. Se si vota la compo-

nizione francese e l'inglese seconda lingua da 0 a 10, niente permette di sommare i due risultati per valutare un sapere totale. Tuttavia è ciò che si fa, da decenni, nelle nostre scuole. E anche quanto meno sospetto, utilizzare scale numeriche diverse per le diverse materie in funzione di una ponderazione empirica: da 0 a 120 per il francese, da 0 a 80 per le scienze, ecc.

Se il sistema ha potuto funzionare, è perché, comunque, ha permesso di identificare i migliori (quelli che si classificano tra i primi *in tutto*) e i più scadenti (ultimi *in tutto*).

### (C) Utilità

Malgrado le sue limitazioni, la scala di valutazione è, in molti casi, il solo strumento di cui disponiamo per concretizzare il nostro giudizio su comportamenti umani complessi.

Infatti essa permette di portare ad un dossier scolastico elementi *qualitativi*, per quanto possibile obiettivi, *complementi indispensabili* dei risultati *quantitativi* apportati dagli esami obiettivi o dai test.

### (D) Costruzione

Molti educatori si lasciano trarre in inganno dalla apparente facilità con cui si costruisce una scala.

In realtà non è raro che la messa a punto di tale strumento esiga lunghi mesi di lavoro, necessari alla chiarificazione teorica, e richieda l'utilizzazione di tecniche perfezionate come l'analisi fattoriale. Questa aiuta particolarmente a determinare in quale misura parecchie scale valutano le stesse cose sotto denominazioni diverse.

Anche in questo campo, è importante dunque che, il più rapidamente possibile, gli insegnanti possano avere un appoggio da parte di servizi di ricerca ben forniti di specialisti e di calcolatori.

Questo peraltro non significa che gli insegnanti non possano far niente nel frattempo.

## (1) Quanti gradi?

Teoricamente una scala di valutazione può contare una infinita quantità di gradi. In pratica, chi giudica classifica con una certa sicurezza solo se si utilizza un piccolo numero di gradi: 3, 5, 7 o 9.

Riferendoci alla letteratura sperimentale e alle osservazioni che abbiamo potuto fare direttamente, consigliamo di attenersi il più spesso a 5 gradi. Tali scale permettono classificazioni relativamente sicure e fedeli, a condizione che ogni grado sia molto chiaramente definito e i classificatori ben formati.

Praticamente, si classifica dapprima in tre categorie: i migliori, i meno preparati, gli altri; si dividono poi gli « altri » in buoni, medi e scadenti.

eccellente	ottimo	buono	sufficiente	scadente o negativo
A	B	C	D	E

Osserviamo che il grado C è « a cavallo » sul centro della scala, cioè sulla media aritmetica. Come abbiamo visto a proposito della curva di Gauss, si produce spesso una forte concentrazione attorno a questo punto. Scegliere un numero pari di gradi creerebbe nella scala un centro intorno al quale parecchi alunni oscillerebbero a caso.

Nell'evenienza di una distribuzione normale, non è escluso che questo giudizio del caso colpisca il 20, il 30 e perfino il 40% degli alunni.

## (2) Definire l'oggetto della valutazione

Chiedere, per esempio, d'apprezzare il « coraggio » dei membri di un dato gruppo — senza altre precisazioni — condurrebbe a risposte quasi completamente prive d'interesse. Quale o qual altro senso avranno dato gli osservatori alla parola « coraggio »: ardore, volontà, zelo, perseveranza, bravura, fermezza, stoicismo?

E anche se precisiamo che per « coraggio » intendiamo la fermezza davanti al pericolo, faremo facilmente la distinzione tra l'essere intrepido e l'essere temerario?

Per sopperire alle imprecisioni di classificazioni quali ottimo, medio, ecc., è necessario aggiungere alla scala una descrizione il più precisa possibile dell'aspetto o dell'oggetto da valutare, e di illustrare la definizione con situazioni tipo. Definire non con formule astratte, ma con comportamenti precisi, è una condizione sine qua non di validità.

Ecco come Schonell<sup>2</sup> presenta l'atteggiamento « Fiducia in sé »:

Fiducia in sé				
Estremamente fiducioso di sé. Quasi troppo sicuro di sé.	Molto fiducioso nelle sue forze.	Fiducioso	Mancanza di fiducia. Timido.	Estrema mancanza di fiducia. Dipende dagli altri. Declina le responsabilità.

*Descrizione dell'atteggiamento:* « Nella sua forma positiva, questa qualità è rivelata dalle seguenti manifestazioni: l'individuo conta su se stesso, è capace di far fronte alle difficoltà, ha sicurezza, è indipendente e pronto ad assumersi responsabilità. Il bambino che ha fiducia in se stesso prova ad andare avanti col minimo d'assistenza; colui che manca di fiducia deve essere aiutato costantemente. Il primo ama vedere ciò che è capace di costruire e di produrre quando ha ricevuto istruzioni chiare; il secondo vuole la pappa pronta, e che lo si aiuti durante tutta la fase di realizzazione ».

*Situazioni tipo:* Ha paura dell'oscurità? Può aver cura di se stesso e delle sue cose o bisogna che qualcuno sia sempre al suo fianco? Viaggia solo in tram o in autobus [per bambini di più di nove anni]? Parla liberamente a visitatori sconosciuti? È bravo nei giochi? Sa nuotare? Si trova a suo

<sup>2</sup> F. J. Schonell, *Backwardness in the Backward Subjects*, citato da F. Wasburne, *Measurement of Personality*, in « Educational Research », novembre 1961, p. 9.

agio e risponde con sicurezza agli esami orali? Legge bene, drammatizza bene un testo davanti alla classe? Si applica bene a compiti nuovi o pone continuamente domande ai suoi compagni e ai suoi maestri?

*Definire i gradi della scala.* Esempio: Organizzazione dell'insegnamento della lettura.

0	1	2	3	4	5
Mediocre	Abbastanza bene	Medio	Ottimo	Eccellente	
In lettura, tutti gli alunni seguono la stessa progressione. Niente lavoro di gruppi.	Cfr. 1. Ma a volte un alunno molto lento riceve un po' meno lavoro degli altri.	Costituzione di due o tre gruppi, secondo le attitudini in lettura. Poca flessibilità nei gruppi.	Raggruppamento secondo le attitudini. Flessibilità.	Raggruppamento dopo studio approfondito delle attitudini e delle difficoltà incontrate. Grande flessibilità.	

Si sarà notato che quest'ultimo esempio combina le scale grafica, numerica e descrittiva.

Il pericolo di una definizione insufficiente dei cinque gradi è stato mostrato sperimentalmente da Weiss<sup>3</sup>. Adottando le scale descrittive imposte dal ministero austriaco dell'educazione nazionale, Weiss ha invitato 92 maestri a classificare due composizioni di alunni del livello al quale essi insegnavano (4<sup>a</sup> elementare) e 272 maestri a classificare due compiti di aritmetica (problemi; 4° e 5° anno).

I diagrammi seguenti mostrano come le classificazioni sono divergenti, per uno stesso compito (dispersione dei voti su tutta la lunghezza della scala).

È chiaro che non basta adottare una nuova scala in sostituzione del sistema di voti tradizionale perché siano risolti tutti i problemi. Tutt'altro.

<sup>3</sup> R. Weiss, op. cit.

## CLASSIFICAZIONE DI DUE COMPOSIZIONI

		Classificazione di due temi			Voto medio	Sigma
		25%	50%	75%		
I	Ortografia				2,89	1,04
	Stile				2,29	0,53
	Contenuto				2,08	0,79
	Voto globale				2,45	0,80
		25%	50%	75%	Voto medio	Sigma
II	Ortografia				2,99	0,96
	Stile				2,12	1,05
	Contenuto				1,78	0,85
	Voto globale				2,54	0,80

## Classificazione di due compiti d'aritmetica

		25%	50%	75%	Voto medio	Sigma
I					2,56	0,77
II					3,55	0,80

## (E) Utilizzazione

## (1) Quanti alunni in ogni livello?

*Primo caso:* Alunno confrontato con se stesso.

I livelli segnano i gradi da superare nell'ascesa verso un potere o un sapere. Il maestro ha definito un obiettivo da raggiungere. In un insegnamento non individualizzato, questo obiettivo è comune alla maggioranza, se non alla totalità, degli alunni della classe. Esso, evidentemente, deve essere scelto in maniera che *tutti* gli alunni possano superarlo. Sarà dunque un obiettivo ampio, ciò che rischia di nuocere alla sua precisione al punto da renderlo inoperante.

È chiaro che noi agiamo attualmente come se all'inizio dell'anno scolastico tutti gli alunni si trovassero approssimativa-

mente ad uno stesso punto della scala del sapere. Si suppone dunque che l'anno precedente tutti abbiano raggiunto il grado superiore della scala precedente.

Ora, ce ne rendiamo ben conto, la realtà è diversa, e la ricerca conferma la nostra opinione.

(a) *Margini di variazione dell'età mentale.* Al momento della revisione del test d'intelligenza di Binet-Simon da parte di Terman e Merrill (1937), è stato constatato negli Stati Uniti<sup>4</sup>:

- che nel primo anno di scuola primaria, l'età mentale variava da 4 a 8 anni;
- che nel sesto anno di scuola primaria, l'età mentale andava da 8 a 16 anni;
- che negli studi secondari, non era eccezionale un margine di variazione da 8 a 10 anni.

(b) *Margini di variazione del rendimento scolastico.* Basta osservare come le curve di distribuzione dei test nozionistici, standardizzati per anno scolastico, si accavallino, per prendere coscienza dell'ampiezza dei margini di variazione.

Appoggiandosi alle sue ricerche e a quelle di Lindquist, Cornell, Learned e Wood, Cook<sup>5</sup> fornisce le seguenti indicazioni:

- Per la comprensione della lettura, per il vocabolario, le scienze, la geografia e la storia, il margine di variazione del rendimento è: in 1<sup>a</sup> elementare: da 3 a 4 anni; in 4<sup>a</sup> elementare: da 5 a 6 anni; in 6<sup>a</sup> elementare: da 7 a 8 anni.
- Per il ragionamento aritmetico ed il calcolo: in 6<sup>a</sup> elementare: da 6 a 7 anni.
- In cultura generale (scienze, letteratura, arte, storia, geografia e lingua materna), al livello universitario: il 28%

<sup>4</sup> Cfr. Q. McNemar, *The Revision of the Stanford-Binet Scale*, Boston 1942.

<sup>5</sup> W. W. Cook, *The Functions of Measurement in the Facilitation of Learning*, in E. F. Lindquist, op. cit., pp. 3-47.

degli studenti di 4° anno sono inferiori alla media degli studenti di 2° anno; il 10% degli studenti di 4° anno sono inferiori alla media degli alunni dell'ultima classe degli studi secondari.

Tali osservazioni — che non bisogna prendere alla lettera, di cui bisogna ricordare la tendenza — potrebbero essere moltiplicate a piacere.

Esse provano che sino a quando le nostre scuole funzionano secondo il sistema delle classi rigide, il fatto che vi sia una valutazione continua che proceda di pari passo con la preoccupazione di permettere a ciascun alunno di progredire secondo un suo proprio ritmo, di superarsi sempre, di andare lontano quanto gli sia possibile, sarà un'illusione.

Tuttavia, siccome per motivi tecnologici ed economici, l'insegnamento individualizzato non si generalizzerà prima di parecchi anni, una soluzione realistica, immediatamente applicabile, ci viene offerta solo dall'insegnamento semi-individualizzato. In una scuola semi-individualizzata, si sostituiscono alla classe gruppi omogenei, secondo le materie principali. Così un alunno dotato in lingua straniera può lavorare ad un livello A, per questa materia, ma unirsi al livello D per i problemi di aritmetica<sup>6</sup>.

Per non appesantire questo testo, abbiamo riportata in appendice una descrizione particolareggiata di una scuola funzionante da cinquanta anni secondo questo principio.

*Secondo caso:* Alunni confrontati tra loro.

Anche se si cerca di condurre ciascuno ad un livello elevato, sussisteranno probabilmente differenze sufficienti per fare apparire nuovamente una distribuzione gaussiana (o, almeno, il suo abbozzo) dove la media si sarà semplicemente spostata verso l'alto.

<sup>6</sup> Non è per caso che parliamo di *problemi d'aritmetica* e non di aritmetica in generale. In una materia complessa come questa, un alunno può essere ben dotato in calcolo, ma scadente in problemi. La ricerca contemporanea dimostra che se non si usano prove di abilità raggruppate con molta cura, i margini di variazione tra alunni restano considerevoli.

Appare dunque giustificato cercare di popolare ogni livello secondo percentuali che si avvicinano alla distribuzione normale.

In una classe tradizionale, è evidente che non è possibile rispettare rigorosamente questa proporzione, dato il ristretto numero di alunni (20). Avremo, per esempio:

			La distribuzione normale sarebbe:
Eccellenti:	1	5%	2,5%
Ottimi:	4	20%	13,5%
Buoni (medi):	10	50%	68 %
Scadenti:	4	20%	13,5%
Molto scadenti:	1	5%	2,5%
	20	100%	100%

Bisogna insistere sul fatto che sforzarsi di rispettare tale ripartizione (*distribuzione forzata*) non implica necessariamente un insuccesso per un certo numero di alunni. Tutti i voti sono *relativi* gli uni agli altri: « molto scadente » può, in alcuni gruppi ben preparati, porsi al di là del voto di insuccesso.

È per questo, d'altra parte, che è *meglio adottare per la classificazione lettere* anziché le formule tradizionali.

## (2) Lottare contro la contaminazione e la tendenza centrale

Oltre ai pericoli di stereotipia e di effetto di alone descritti a p. 28, un altro fenomeno di contaminazione, di natura più meccanica, attende al varco coloro che giudicano. Se essi utilizzano consecutivamente per uno stesso alunno una serie di scale orientate nello stesso senso (per esempio dal positivo al negativo o dal migliore al più scadente) e forse anche rappresentate su una stessa pagina, le classificazioni tendono a raggrupparsi su uno stesso lato.

Si consiglia di presentare ciascuna scala su una pagina di-

... e inoltre di estrarre a sorte l'ordine di presentazione di ciascuna.

Infine è bene disporre un intervallo abbastanza lungo dopo la valutazione di ogni qualità di uno stesso soggetto e far valutare una stessa qualità dal maggior numero possibile di giudici.

Molti giudici hanno anche tendenza a raggruppare i loro voti verso il centro della scala: per timore di sopravvalutare o di sottovalutare un alunno, timore di assumersi la responsabilità di un insuccesso. La regola della distribuzione forzata è qui di grande utilità. Ma per essere realmente efficace, deve potersi basare su una descrizione precisa e, insieme sfumata dei diversi livelli.

Quando le valutazioni compiute da parte di diversi giudici devono essere combinate, i professori che si lasciano sedurre dalla soluzione di comodo che è quella del voto centrale, devono sapere che rinunciano ad una larga parte della loro influenza, in favore dei loro colleghi che osano utilizzare la scala completa.

## (F) Come sintetizzare le valutazioni

Al termine di un periodo scolastico, è necessario sintetizzare le valutazioni (che si augura siano il più numerose possibile), senza di che maestri ed alunni non sapranno interpretarle.

Tornando al semplice calcolo della media si rischierebbe di distruggere gli sforzi di rigore precedenti. Perché valutazioni ordinali, almeno teoricamente, non permettono né addizioni né sottrazioni.

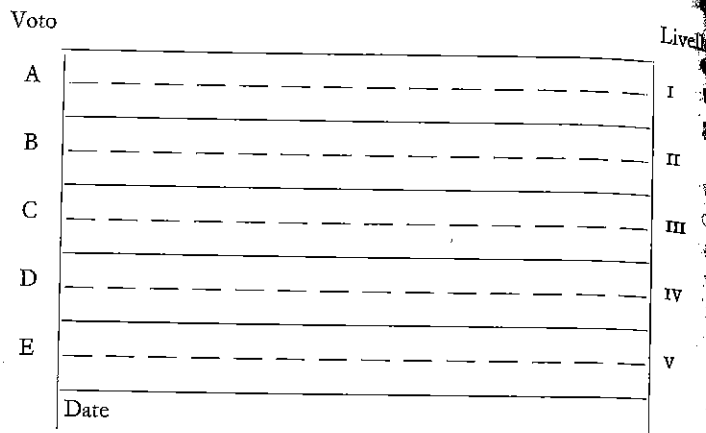
Ecco un metodo semplice che permette di situare un alunno ad un determinato livello in ogni materia, poi per l'insieme delle materie (vedi la figura seguente).

Un professore utilizza tante tabelle uguali per quanti aspetti desidera distinguere nel suo corso: dettato, grammatica, ortografia, componimento, ecc.

Man mano che si procede nelle valutazioni, si segna una croce nella tabella, indicando la data in basso.

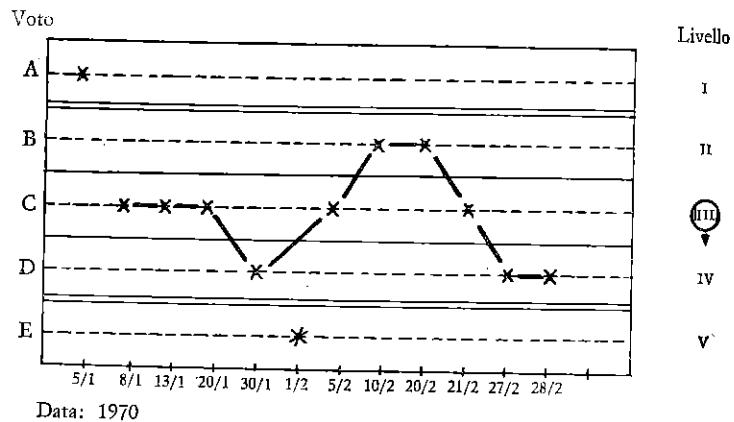


All'inizio, si dispone della seguente tabella:



Al momento di fare la sintesi, si comincia con lo sbarrare (senza cancellarla!) la croce corrispondente al voto più alto e quella che rappresenta il voto più basso, per eliminare voti che potrebbero essere accidentali.

Supponiamo che si ottenga la tabella qui sotto.



interpretazione

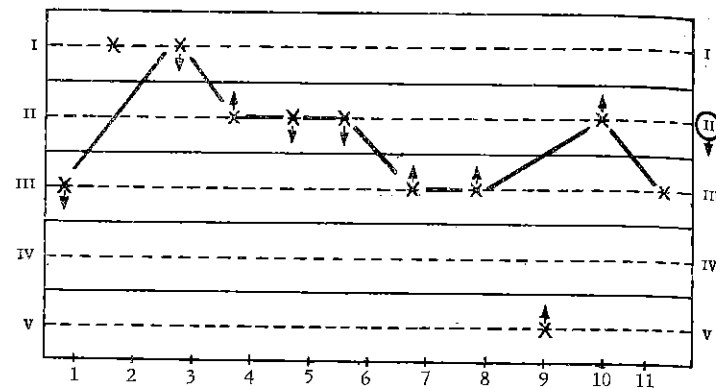
- (1) I risultati si concentrano su tre fasce: II, III, IV. Non sono dunque né eccellenti, né scadenti o insufficienti. La zona di concentrazione è segnata con delle linee doppie.
- (2) Il maggior numero di voti (il valore dominante) è situato nella fascia III. È la tendenza dominante.
- (3) Siccome la fascia dove sono state segnate il maggior numero di crocette, dopo la III, è la IV, sarà posta una freccia, in questa direzione, sotto la III.
- (4) Riunendo le croci si ottiene un profilo che rivela:
  - (a) Una omogeneità abbastanza buona dei risultati. Se l'alunno si muovesse anarchicamente sulle cinque fasce, sarebbe il caso di vedere se si trova da solo in questa situazione (causa da determinare a livello dell'alunno), o se la maggior parte della classe si trova nella stessa situazione (vedere se l'opera del professore non sia in qualche modo inadeguata).
  - (b) Una discesa alla fine del periodo. Essa non si manifesta di frequente: può essere spiegata come causata da stanchezza, da una malattia benigna, ecc.

È chiaro, non basta constatare una situazione denunciata dal profilo, bisogna cercare di scoprirne il perché. Così, i commenti, orali o scritti, dei maestri, chiariscono e guidano l'alunno e i suoi genitori.

Rileviamo, del resto, che questo sistema non può funzionare se il professore si limita a due o tre valutazioni; ancor meno con una sola! In questo caso sarà l'introduzione di una nuova forma di valutazione a migliorare felicemente il comportamento di alcuni insegnanti frastornati dalla quantità di programma da svolgere ... o allergici al lungo lavoro di classificazione dei lavori.

## TABELLA PER L'INSIEME DELLE MATERIE

PERIODO DAL ..... AL ..... 19..



1 = lingua materna; 2 = matematica; 3 = biologia; 4 = storia;  
 5 = geografia; 6 = inglese; 7 = tedesco; 8 = olandese; 9 = fisica;  
 10 = chimica; 11 = educazione fisica.

## Interpretazione

Quattro voti d'insieme si trovano nella fascia II e quattro nella fascia III. È la tendenza del profilo (qui verso l'alto) che deciderà quale sia la fascia prevalente: qui II, con freccia verso il basso.

In caso di uguaglianza pratica tra due livelli, si può sia decidere che il livello superiore prevale (soprattutto se si tratta di due livelli contigui), sia circondare due cifre e riunire i due cerchi con una linea.

## Osservazioni

- (1) Questo sistema di sintesi delle valutazioni non ha di per sé alcun potere magico! Se si commettono alcuni errori segnalati nel presente capitolo, il voto di sintesi sarà probabilmente privo di senso. Un fenomeno di concentrazione nelle fasce mediane stupisce spesso gli utilizzatori. Essi dimenticano che, nel sistema tradizionale, che finisce in una percentuale, i voti si concentravano

anch'essi nella stessa zona: tra 65 e 75 e 80%, mentre tutto il resto della scala (di 100 gradi!) rimaneva molto spesso quasi vuoto.

In realtà il fenomeno di concentrazione è più sorprendente poiché la scala a cinque gradi è molto meglio visualizzabile della vecchia.

- (2) Niente impedisce di equilibrare i voti, cioè di accordare, per esempio, importanza doppia al lavoro svolto in classe rispetto a quello svolto a casa, a un aspetto di una materia piuttosto che ad un'altra, ecc.
- (3) Si troveranno nella parte riservata alla « moderazione » indicazioni supplementari per utilizzare i voti, in particolare per determinare i risultati e le classificazioni alla fine dell'anno (cfr. p. 181).

## (G) Un caso particolare: la classificazione della composizione francese

Questo problema merita alcune considerazioni particolari, poiché è considerato tra i più difficili.

Se esso fosse risolto, la docimologia riporterebbe forse la più bella vittoria. Ahimè! o, piuttosto, fortunatamente per l'uomo, la più nobile delle sue attività, la valutazione del bello, del vero e del bene, sfuggirà senza dubbio sempre alla quantificazione obiettiva, e quindi automatizzabile.

## (1) Quattro metodi di valutazione

I principali metodi si riducono a quattro: il metodo dell'impressione generale, la scala di specimen, il metodo analitico, e il conteggio di frequenze.

## (a) Il metodo dell'impressione generale

È il più soggettivo. Gli si rimprovera la mancanza di fedeltà da parte di uno stesso correttore, il disaccordo tra parecchi classificatori esperti o no, e la sua frequente inefficacia: molti professori si sono logorati a classificare con cura centinaia di lavori per anni senza ottenere risultati convincenti.

Ci sarebbe molto da dire sulla metodologia della composizione francese, e soprattutto sulla necessità di una riforma profonda delle tradizioni scolastiche che trattano ogni alunno come se fosse uno scrittore in potenza, ma non è nostro compito occuparci di questo. Teniamo presente solo che l'unione delle due manchevolezze spesso denunciate, una docimologica, l'altra metodologica, grava l'esercizio di composizione di una ipoteca così pesante che alcuni dichiarano inutile questo esercizio. Noi non siamo di questo avviso.

Malgrado le sue manchevolezze, il metodo dell'impressione generale è il migliore, in armonia con la complessità essenziale della composizione. Utilizzando una scala corta, a tre o al massimo a cinque gradi, può essere ottenuto un accordo molto elevato tra classificatori esperti, di uguale formazione ed appartenenti ad uno stesso ambiente scolastico. Vedremo più in là come decidere di un insuccesso.

Una composizione giudicata con questo metodo non deve, crediamo, avere l'ambizione di migliorare molto lo stile o la semplice correttezza del linguaggio scritto: per raggiungere questo scopo ci sono altri esercizi.

La composizione tradizionale deve essere una prestazione d'eccezione, destinata a reperire talenti o a controllarne la maturazione, ed anche ad identificare debolezze, carenze, goffaggini che si combatteranno per mezzo di esercizi sistematici di arricchimento e di recupero.

Se si decide di attenersi alla sola valutazione globale, la tecnica della valutazione in un tempo prestabilito merita di essere presa in considerazione.

Essa comporta minori varianti. È caratteristico il metodo proposto da Wiseman<sup>7</sup>. Quattro classificatori valutano indipendentemente dei lavori fatti in trenta minuti, ad un ritmo di circa cinquanta copie all'ora (tasso imposto per obbligarli a decidersi molto rapidamente). Si calcola la media aritmetica dei voti.

<sup>7</sup> S. Wiseman, *The Marking of English Composition in Grammar School Selection*, in «British Journal of Educational Psychology», novembre 1949, p. 208.

Questo metodo è probabilmente il più efficace per i lavori creativi o descrittivi. Per le «dissertazioni», parola adoperata per tutte le composizioni in cui il soggetto deve argomentare, la superiorità del metodo è più discussa. Tuttavia essa sembra valida anche in quest'ambito, se i classificatori sono ad alto livello intellettuale.

La validità generale del metodo e la sua applicabilità alla «dissertazione» è confermata da due ricerche sperimentali di J. Britton.

*Prima esperienza*<sup>8</sup>. Un insieme di composizioni scritte da 68 alunni di 15 anni è stato dapprima classificato, secondo il metodo analitico, da un esaminatore esperto. Gli stessi lavori sono stati poi valutati indipendentemente, secondo il metodo dell'impressione generale a tempo prestabilito, da otto classificatori.

Per disporre di un criterio, altri cinque lavori di ogni alunno sono stati votati da due esaminatori.

I voti analitici e il voto medio di tre impressioni rapide tirate a sorte tra le otto di cui si disponeva per ogni studente, sono state allora confrontate ai voti di criterio: correlazione tra voti analitici e criterio: .71; correlazione tra voti rapidi e criterio: .76.

Inoltre, la fedeltà della classificazione rapida è stata calcolata confrontando i voti di due gruppi di tre esaminatori, scelti a caso tra gli otto. Si ottiene un coefficiente molto elevato: .87.

*Seconda esperienza*<sup>9</sup>. Questa seconda esperienza, molto più sottile della precedente, si è basata su un campione di 500 alunni di 17 anni, stratificato secondo il sesso, il tipo di scuola secondaria e il luogo di abitazione (Londra, centro e sobborghi, grandi città e piccole città). Erano anche rappresentati tutti i tipi di soggetti di composizione, dalla semplice narrazione alla dissertazione filosofica ed alla creazione poetica.

<sup>8</sup> J. Britton, *Experimental Marking of English Composition Written by Fifteen-Year-Olds*, in «Educational Review» 1 1963, pp. 17-23.

<sup>9</sup> J. Britton, N. Martin e H. Rosen, *Multiple Marking of English Composition*, London, HMSO, 1966.

Ancora una volta, la classificazione multipla secondo l'impressione rapida si è dimostrata più fedele e più valida della classificazione analitica.

(b) Le scale di specimen

Un piccolo numero di composizioni, spesso cinque, sono scelte in modo da costituire modelli rappresentativi dei differenti livelli di qualità: dalla mediocrità all'eccellenza.

Il classificatore valuta i lavori confrontandoli ai cinque testi-tipo della scala.

Dopo avere conosciuto un successo considerevole in alcuni paesi<sup>10</sup>, le scale di specimen sono state abbandonate per parecchi motivi<sup>11</sup>: è raro che la composizione da votare rassomigli strettamente ai testi-tipo; i classificatori tendono a identificare le caratteristiche comuni che li interessano particolarmente e trascurano le altre; sono necessarie scale diverse a seconda del tipo di soggetto e a seconda del livello degli alunni.

Pidgeon ci ha segnalato che una vasta esperienza, condotta recentemente in Gran Bretagna, con una scala a 5 gradi di specimen, si è chiusa con uno scarto medio di 2 gradi tra i correttori. Il risultato non è molto migliore dunque che in condizioni puramente soggettive.

Osserviamo tuttavia che, almeno dopo il 1903, il metodo delle scale di specimen ritrova periodicamente il favore dei ricercatori.

(c) Il metodo analitico

Due o tre classificatori accordano un certo numero di punti per ognuna delle diverse qualità che si è convenuto di osservare. Si totalizza e si calcola la media tra classificatori.

Questo metodo è il più gravoso dei quattro; spesso viene criticato perché esso analizza un tutto che sembra sfuggire ad

<sup>10</sup> Per uno studio d'insieme, vedere E. Hinton, *Study of the Qualities of Style and Rhetoric Found in English Compositions*, New York 1940.

<sup>11</sup> R. Braddock e altri, *Research in Written Composition*, Champaign (Ill.), NCTE, 1963, p. 12.

una analisi sistematica. Wiseman rileva d'altra parte che, in un insieme di lavori classificati analiticamente, il lavoro considerato all'unanimità come il migliore dagli esperti, raramente si colloca in testa.

Tuttavia la ricerca dimostra che, se le qualità da osservare e l'importanza da accordare loro sono definite con precisione, il metodo analitico è il più sicuro, se ci si deve affidare ad un solo correttore. Il metodo dell'impressione generale dà, l'abbiamo già visto, altrettanto buoni risultati, ma esige la partecipazione di parecchi classificatori<sup>12</sup>.

Perché sia efficace e praticabile il metodo analitico sembra dover rispondere almeno a due esigenze: i correttori devono accettare il piano d'analisi; l'ideale sarebbe che essi partecipassero alla sua elaborazione; il piano non deve essere troppo dettagliato: così, il metodo molto sottile utilizzato da Buxton<sup>13</sup> offre interesse soltanto per la ricerca, a causa del tempo molto prolungato che esige.

Che qualità osservare? Remondino<sup>14</sup> ha dimostrato che gli atteggiamenti fondamentali dei correttori colti (professori o no) non hanno fra loro differenze essenziali: le stesse qualità impegnano la loro attenzione. Ma l'importanza data a ciascuna di esse nella valutazione d'insieme varia di molto, e questa varietà è in realtà l'origine delle divergenze.

Elaborare una tabella molto dettagliata di analisi del contenuto è un'illusione, poiché, anche se esistono sfumature distinte, i classificatori inconsciamente le rifondono. L'analisi fattoriale mette in evidenza questo processo. Il lavoro di Remondino porta a questo proposito interessanti risultati.

Remondino ha dapprima interrogato a lungo venti profes-

<sup>12</sup> Si troverà anche un buono studio comparativo in B. Cast, *The Efficiency of Different Methods of Marking English Composition*, in « British Journal of Educational Psychology », novembre 1939, pp. 257-269, e febbraio 1940, pp. 49-60, citato da R. Braddock, op. cit., p. 13.

<sup>13</sup> E. Buxton, *An Experiment to Test the Effects of Writing Frequency*, in « Alberta Journal of Educational Research », giugno 1959, pp. 91-99.

<sup>14</sup> *Étude factorielle sur la notation des compositions scolaires portant sur la langue maternelle*, in « Le travail humain », gennaio-giugno 1959, pp. 27-40.

sori di materie letterarie (a livello di scuole secondarie) ha anche compilato la lista delle qualità rilevate in composizioni scolastiche. Egli giunge a catalogarle come segue:

(1) Leggibilità	Qualità d'una scrittura che si presta a una lettura facile, rapida, senza equivoco.
(2) Estetica	Linea armoniosa e gradevole delle lettere e buon gusto nella impaginazione.
(3) Presentazione	Decoro, cura, buona presentazione dell'elaborato.
(4) Esattezza ortografica	Densità degli errori d'ortografia.
(5) Esattezza morfologica	Densità degli errori morfologici.
(6) Esattezza sintattica	Densità degli errori di sintassi.
(7) Struttura della esercitazione	Qualità di una esposizione fatta con ordine, nelle volute proporzioni, e secondo un piano.
(8) Ricchezza di idee	Quantità di idee; risorse utilizzate.
(9) Pertinenza delle idee	Qualità per cui le idee esposte sono in giusto rapporto col tema trattato.
(10) Precisione d'informazione	Veracità ed esattezza obiettiva delle affermazioni e dei fatti esposti.
(11) Esaustività	Qualità che consiste nel non lasciar da parte niente di tutto ciò che si doveva dire.
(12) Concisione	Qualità per cui le cose sono dette col minimo indispensabile di termini senza ripetizioni, ridondanze o tortuosità.
(13) Proprietà del linguaggio	Giusto uso dei termini.
(14) Stile	Facilità, esattezza, padronanza della lingua dal punto di vista della « costruzione delle frasi ».
(15) Originalità	Qualità per cui, attraverso il lavoro, traspare e si afferma qualcosa della personalità.
(16) Maturità	Capacità di giudizio, profondità critica, acutezza delle argomentazioni esposte.
(17) Immaginazione	Capacità di creazione, di trasfigurazione, di « proiezione » rivelata dal lavoro.

Una analisi fattoriale ha poi dimostrato che queste 17 qualità derivano da quattro gruppi solamente: presentazione grafica (1, 2 e 3); uso della lingua (4, 5, 6, 13, 14); contenuto ed organizzazione dell'esposizione (7, 8, 9, 10, 11); aspetti personali del contenuto (15, 16 e 17).

Le conclusioni di Remondino confermano largamente quelle degli altri ricercatori. Le scale seguenti costruite dall'Educational Testing Service (ETS) di Princeton<sup>15</sup>, e leggermente rimaneggiate da Page<sup>16</sup>, testimoniano della analogia di vedute.

Traduciamo il testo completo perché esso dà uno dei migliori esempi che si possano dare attualmente per indicare la via da seguire.

### Criteri per classificare le composizioni

Definizione delle caratteristiche da valutare

- (A) Idee o contenuto: la quantità e la qualità del materiale utilizzato per trattare del soggetto.
- (B) Organizzazione: la relazione tra le parti della composizione e l'insieme.
- (C) Stile: utilizzazione del linguaggio al di là della semplice correttezza grammaticale.
- (D) Meccanica: ortografia, grammatica, punteggiatura.
- (E) Creatività.

Guida per la valutazione delle cinque caratteristiche

- (A) Idee o contenuto

*Livello elevato.* Lo studente ha idee buone, originali e numerose, egli comprende bene il tema e utilizza delle definizioni chiare. È capace di considerare il tema in una prospettiva più ampia di quella degli altri alunni della classe. In altre parole, manifesta un'esperienza più ricca.

<sup>15</sup> ETS, *Definitions of Ratings on the ETS Compositions Scale*, citato da E. Page, op. cit., pp. 76-77.

<sup>16</sup> Op. cit., pp. 78-80.

*Livello medio.* Le idee sono appropriate, ma convenzionali e poco numerose. Alcuni aspetti del tema sono trascurati. L'alunno non sembra avere una notevole ricchezza intellettuale.

*Livello basso.* Lo studente omette molti aspetti importanti dell'argomento. Sembra che non disponga di una riserva di conoscenze relative al tema e, di conseguenza, ripete continuamente alcune idee semplici.

#### (B) Organizzazione

*Livello elevato.* Lo studente segue un piano definito. Se fa conoscere il pro e il contro, presenta dei motivi pertinenti, in un ordine efficace. Se descrive qualcosa, lo fa in modo ordinato (dall'alto verso il basso, in ordine d'importanza, in ordine di complessità, ecc.). Se lo studente spiega un concetto o un processo, utilizza un piano coerente d'analisi, di definizione o d'illustrazione. Lo studente sa bene ciò che si riferisce al suo piano ed evita ripetizioni. Manifesta senso della misura nel trattare le diverse parti del suo lavoro.

*Livello medio.* Lo studente non si attiene al suo piano oppure introduce delle idee che non hanno rapporto col tema. Dedicava troppo tempo a cose poco importanti o si ripete. Tratta l'argomento con accostamenti liberi (che cosa mi viene in mente quando penso alle Hawaii?) anziché perseguendo un obiettivo ben definito.

*Livello basso.* Sembra che lo studente non si sia domandato ciò che stava per fare prima di mettersi a scrivere. Egli non segue alcun piano. Il lavoro prende una direzione, poi cambia, cambia ancora ed ancora, fino a che il lettore non riesce più a seguire. I punti principali non sono chiaramente separati uno dall'altro e il loro ordine di presentazione è lasciato al caso.

#### (C) Stile

Parecchi aspetti dello stile possono interferire nella valutazione: individualità, vivacità, eleganza, ecc. Tuttavia noi qui ci interessiamo a tre aspetti stilistici solamente: chiarezza, variazione e gamma di risorse linguistiche.

*Livello elevato.* Lo studente utilizza un linguaggio che rende agevole la comprensione del lavoro. Utilizza parole adeguate, nel loro senso abituale. Le parole sono presentate in ordine normale. I passaggi sono ben condotti. L'alunno evita le ambiguità e non delude l'attesa del lettore. Nello stesso tempo, lo studente evita la ripetizione monotona di parole, di complementi o di strutture di frasi. Infine, manifesta la conoscenza di una vasta gamma di risorse linguistiche. Il suo vocabolario è buono. Utilizza strutture parallele o un uso ingegnoso della subordinazione.

*Livello medio.* Lo studente a volte fuorvia il lettore utilizzando una parola non appropriata o una costruzione bizzarra; oppure utilizzando una metafora poco chiara o spostando in maniera inopportuna un complemento o una subordinata, oppure facendo dei bruschi passaggi. La ripetizione di parole, di costruzioni e di strutture di frasi diventa monotona. Le risorse linguistiche sono limitate. L'alunno utilizza volentieri dei clichés e delle costruzioni banali.

*Livello basso.* Le parole sono utilizzate in modo generico. Forme ambigue, costruzioni zoppicanti, vocabolario e strutture di frasi infantili.

#### (D) Meccanica

*Livello elevato.* La struttura delle frasi è abitualmente corretta, anche quando si tratta di modelli vari e complicati. Le regole d'ortografia sono rispettate, anche le parole difficili sono generalmente scritte senza errori. Nessun grave errore di punteggiatura, di maiuscole, di abbreviazioni, di scrittura di numeri.

*Livello medio.* Difetti occasionali di sintassi. Le parole difficili hanno a volte errori di ortografia. Qualche errore di punteggiatura, ecc.

*Livello basso.* Gran numero di errori.

#### (E) Creatività

*Livello elevato.* Lo studente sorprende per i modi nuovi ed efficaci di considerare il problema. Introduce idee nuove nella sua trattazione dell'argomento. Trova modi nuovi ed in-

teressanti di utilizzare il linguaggio per mettere in rilievo le sue idee.

*Livello medio.* Lo studente pensa a ciò che ci si aspetta che egli pensi. Tratta le cose come quasi tutti gli altri. Utilizza espressioni e strutture di frasi fatte.

*Livello basso.* Lo studente utilizza dei clichés di pensiero e di espressione. Svolge il tema in maniera superficiale. Ripete delle formule senza capirne veramente il significato.

(d) Il metodo dei conteggi di frequenze

Per quanto esso non sia tra le valutazioni soggettive, lo segnaliamo qui, per avere un quadro completo dei metodi.

Esso consiste nel definire tipi di errori (dopo uno studio diagnostico), nel contarli, nella composizione e nell'esprimere il numero ottenuto per 100 o per 1000.

Invece di adottare un punto di vista negativo, si può anche fare l'inventario (relativamente esaustivo) dei diversi procedimenti utilizzati nella composizione. Il migliore esempio di tale metodo è fornito dagli *Eléments pour servir à l'étude d'analyses littéraires* di Purves<sup>17</sup>.

Un ultimo metodo positivo consiste nell'isolare alcuni aspetti che si rivelano predittivi, ad un determinato livello, del voto normalmente attribuito da correttori qualificati. Un calcolatore programmato adeguatamente può allora incaricarsi della classificazione.

I conteggi di frequenze offrono probabilmente la sola speranza di classificazione obiettiva delle composizioni. Essi sembrano tuttavia dover restare di dominio della ricerca estranea alla pratica scolastica, a meno che il calcolatore non diventi talmente accessibile che lo studio delle frequenze si faccia in maniera economica e costituisca uno degli aspetti della valutazione. Il che non è escluso.

*Diversi soggetti a scelta?* Britton osserva con ragione che il tema d'esame su di un unico argomento implica spesso, nel

<sup>17</sup> A. Purves e V. Ripper, *Eléments pour servir à l'étude d'analyses littéraires*, in « *Scientia Paedagogica Experimentalis* » 2 1969, pp. 229-334.

professore, l'ipotesi che un soggetto come « Nuvole » rivelerà l'attitudine di uno studente bene quanto « Esponete il vostro punto di vista sulla pena capitale ».

Si è dimostrato sperimentalmente che non è così, torneremo su questo argomento. Tuttavia non è certo che la giustizia sia molto meglio rispettata se si offrono tre argomenti a scelta quando i voti d'esame sono assegnati per un solo tema. Bisognerebbe dapprima provare che con tre, perfino cinque argomenti, si copre tutto lo spettro degli stimoli necessari per dare a ciascuno la stessa possibilità; e se anche questo fosse stabilito, sarebbe importante inoltre che il professore selezionasse sempre gli argomenti coscientemente.

L'esperienza, dicevamo, ci dimostra che il rendimento dell'alunno varia secondo il tema proposto. Finlayson<sup>18</sup> ha stabilito, per la fine della scuola elementare, che se ad una settimana d'intervallo, gli alunni sono invitati a fare un tema a cominciare da una stessa scelta di quattro argomenti, quelli che cambiano argomento (nell'esperienza più della metà) ottengono voti notevolmente diversi.

Wiseman e Wrigley<sup>19</sup> giungono alla stessa conclusione (scelta nello stesso insieme di cinque soggetti ad un intervallo di quattro mesi), pur provando che la differenza di voto non è imputabile alla mancanza di fedeltà del professore. Vernon e Millican<sup>20</sup> confermano la conclusione al livello universitario.

Che fare? Se si vuole verificare la capacità di esprimersi per iscritto in una prospettiva nettamente definita (per esempio: capacità di fare un rapporto scientifico o di scrivere una lettera commerciale), la sensibilità particolare dei candidati non deve quasi essere presa in considerazione ed è sufficiente un solo argomento.

<sup>18</sup> D. S. Finlayson, *The Reliability of the Marking of Essays*, in « *British Journal of Educational Psychology* » 2 1951, citato da Britton (1966).

<sup>19</sup> S. Wiseman e J. Wrigley, *Essay-Reliability. The Effect of Choice of Essay-Title*, in « *Educational and Psychological Measurement* » 1 1958.

<sup>20</sup> P. Vernon e G. Millican, *A Further Study of the Reliability of English Essays*, in « *British Journal of Statistical Psychology* » 2 1954.

Se, invece, si tratta di scoprire una attitudine all'espressione, ovunque essa sia, o di condurre gli alunni a discutere problemi in cui la forma mentis ed il livello d'informazione possono avere una funzione notevole, non è quasi possibile dare un giudizio basandosi su una sola composizione. L'ideale è allora di riferirsi al lavoro fatto durante il corso dell'anno; se, ogni volta, l'alunno ha potuto scegliere tra diversi soggetti (per esempio cinque), in qualche mese avrà reagito a diverse decine di temi differenti e si può supporre veramente che ciascuno avrà avuto la sua possibilità. Sarebbe allettante lasciar inventare il soggetto da trattare, se ci si potesse premunire contro la preparazione fraudolenta. È molto difficile.

*Conclusioni.* La sorte riservata alla composizione francese dipenderà dallo scopo perseguito. Che cosa si vuole? Correggere un difetto particolare? In questo caso, è meglio non esaminare che quello nei lavori e combatterlo con esercizi speciali.

Conoscere la capacità degli alunni a correggersi, ad identificare la prosa conforme al buon uso? In questo caso, un test obiettivo, con risposte a scelta multipla, potrà andar bene.

Far acquistare la capacità di esprimersi chiaramente? In questo caso, temi molto limitati permettono di concentrare l'attenzione sull'argomento del lavoro. Perché bandire soggetti come: « Descrivete una bicicletta da corsa »?

Scoprire talenti di scrittori? Dunque bisogna lasciar libero corso all'espressione dell'alunno, invitandolo spesso a scegliere lui stesso il soggetto.

Ma non ci si inganni, si metteranno così alla prova differenti capacità. Nella seconda esperienza di Britton, descritta più su, l'autore ha dimostrato che la correlazione tra i risultati nella composizione creativa e un riassunto varia da .30 a .40; che la correlazione della composizione con un test di comprensione del testo è un po' più elevata: da .35 a .45.

Gli scarti sono dunque considerevoli.

Come abbiamo già detto, non pensiamo che sia necessario bandire la composizione tradizionale. In alcuni paesi dove era stata soppressa per la difficoltà di valutazione rigorosa, è stata peraltro reintrodotta, particolarmente negli Stati Uniti

dove la commissione degli esami d'ammissione all'insegnamento superiore ha rinunciato al test obiettivo di composizione — ed in Inghilterra, alle due prove chiave di selezione: l'11+ e il 16+ Examination.

Tuttavia, se la composizione continua ad avere una funzione importante negli esami, la maniera, ancora troppo diffusa, di classificare deve essere modificata. Facciamo nostra la proposta di Braddock: « Se dobbiamo valutare un gran numero di studenti per decidere della loro riuscita o del loro insuccesso, è necessario permettere a coloro che falliscono di fare un secondo lavoro. Se si intende attribuire dei gradi, sono necessari almeno tre lavori; si prenderanno in considerazione i due migliori tra cui si farà la media. I lavori devono essere classificati da almeno due correttori, che usano un sistema di correzione ben compreso ed accettato, nel quale sono stati ben esercitati »<sup>21</sup>.

### La classificazione obiettiva

La classificazione obiettiva non lascia intervenire l'opinione personale dei correttori. L'esempio più semplice è fornito da un quesito che comporta dieci moltiplicazioni in ragione di un punto per ogni risultato esatto. Allo stesso modo, la classificazione di un test composto d'item a scelta multipla si fa per semplice conteggio del numero di scelte corrette.

Più generalmente, si chiama classificazione obiettiva « l'assegnazione di valori numerici a campioni comportamentali sufficientemente limitati, definiti e controllati per permettere un accordo generale tra i giudici o classificatori »<sup>22</sup>.

Come abbiamo già osservato, l'obiettività della classificazione non garantisce affatto l'obiettività dell'esame nel suo insieme. La scelta dei quesiti rimane, in ultima analisi, sempre soggettiva e si può immaginare facilmente un esame di

<sup>21</sup> R. Braddock e altri, op. cit., p. 45.

<sup>22</sup> J. C. Flanagan, *Units, Scores and Norms*, in E. Lindquist, *Educational Measurement*, Washington, ACE, 1961<sup>4</sup>.



cui nessun quesito « campioni » realmente la capacità che pretende di valutare. Per obiettiva che possa essere la correzione, la prova, in tal caso, non avrebbe alcuna validità. Sono moltissimi gli esempi di disavventure di questo tipo.

### La taratura o misura della posizione relativa

L'interpretazione di ogni voto solleva un problema di comparabilità.

Il tarare consiste nel definire *norme*, cioè distribuzioni statistiche che permettono di situare le performance individuali.

Secondo lo scopo perseguito, le norme sono stabilite per un gruppo ristretto o per insiemi più vasti che possono comprendere fino alla popolazione di un paese. Praticamente la loro validità si limita sempre ad un insieme determinato e, per utilizzarle, è necessario sapere quale esso sia.

Per lungo tempo la percentuale è stata l'alfa e l'omega della norma scolastica. Per tarare alcuni test, ci si accontentava a volte anche di esprimere il numero di risposte corrette in percentuale. Si conoscono tutti i punti deboli di tale sistema.

Oggi ha una funzione determinante la differenza tra individui e non la situazione su una scala del sapere considerata come assoluta<sup>23</sup>. La media o la mediana<sup>24</sup> costituiscono i punti di riferimento a partire dai quali si costruiscono scale « universali », che mediante alcune condizioni sulle quali torneremo, permettono di esprimere punteggi grezzi, ottenuti da differenti prove, in unità comparabili.

I sistemi di verifica più conosciuti e più accessibili sono i seguenti: (1) il centilaggio; (2) i voti standard o punti Z; (3) la scala normalizzata a 5 classi; (4) la scala normalizzata a 9 classi (Stanines).

Esaminiamoli uno dopo l'altro.

<sup>23</sup> Vedere natura e debolezza delle scale di valutazione. Calcolare la percentuale equivale a considerare una scala di cento gradi.

<sup>24</sup> Ricordiamo che la mediana è il punto che divide una serie di voti ordinati in due parti uguali: 1-2-3-4-5 o 1-2-3-4-5-6.

### Il centilaggio

#### Definizione

Il centile, o posto occupato nella scala a cento gradi, gode ancora di un certo favore a causa della sua rassomiglianza superiore con la percentuale, dalla quale tuttavia esso differisce in modo fondamentale. La percentuale indica quale porzione del totale dei punti attribuiti ad un esame l'alunno ha ottenuto. Il centile indica quanti alunni si classificherebbero dopo un alunno dato se la classe contasse cento alunni. Dunque un punteggio che equivale al 90° centile è superiore al 90% della popolazione considerata.

La mediana corrisponde al 50° centile.

#### 2) Calcolo

In linea di principio, non si calcolano i centili, a partire da meno di cento voti. L'esempio seguente utilizza i cento voti che ci sono già serviti al momento dello studio della curva di Gauss. Procedimento generale:

- preparare il quadro delle classi (cominciando dalla classe inferiore);
- indicare le frequenze ( $f$ );
- calcolare le frequenze cumulate ( $f_c$ ): ogni frequenza è sommata a quella che la precede;
- applicare la formula:

$$C_x = I + \frac{(N.C)/100 - f_c}{f_i} \times i.$$

$$C_x = \text{Valore del voto } (x) \text{ in centile.}$$

$(N.C)/100$  = Il grado del voto corrispondente ad un centile qualunque.

Esempio: il 50° centile corrisponde al voto medio, sia  $N/2$ . Per gli altri si tratta di una semplice regola del tre. Il 75° centile sarà dunque:  $(100 \times 75) : 100 = 75$  (le cifre sono le stesse perché, nel nostro esempio, vi sono precisamente 100 voti).

- $I$  = La media tra il limite inferiore della classe in cui si trova il centile cercato ed il limite superiore della classe che la precede.  
 $f_c$  = frequenza cumulata della classe che precede quella in cui si trova il centile cercato.  
 $f_s$  = la frequenza semplice dell'intervallo in cui si trova il centile cercato.  
 $i$  = intervallo di classe.

Esempio:

Classi	Frequenze ( $f$ )	Frequenze cumulate ( $f_c$ )
24-28	1	1
29-33	2	3
34-38	8	11
39-43	10	21
44-48	23	44
49-53	22	66
54-58	18	84
59-63	9	93
64-68	5	98
69-73	1	99
74-78	1	100

	Calcolo del 50° centile	Calcolo del 16° centile	Calcolo dell'84° centile
$\frac{N.C}{100}$	$\frac{100 \times 50}{100}$	$\frac{100 \times 16}{100}$	$\frac{100 \times 84}{100}$
$I$	$\frac{49+48}{2} = 48,5$	$\frac{39+38}{2} = 38,5$	$\frac{54+53}{2} = 53,5$
$f_c$	44	11	66
$f_s$	22	10	18
$C_x$	$48,5 + \frac{50-44}{22} \times 5 = 49,9$	$38,5 + \frac{16-11}{10} \times 5 = 41$	$53,5 + \frac{84-66}{18} \times 5 = 58,5$

Scarto semi-interquartile:  $(75^\circ C - 25^\circ C)/2$ .  
 Decilaggio: il 10° centile = il 1° decile, ecc.

## Critica

La scala in centili non è, l'abbiamo già detto, che un'ammessa scala di valutazione costituita da 100 gradi. Essa informa sulla posizione occupata, ma non sulla distanza tra le posizioni. Se la distribuzione è normale, questa distanza diminuisce man mano che ci si avvicina alla mediana (uguale alla media nei casi di normalità perfetta).

La tabellina seguente mostra chiaramente questo fenomeno.

Centili	$-2\sigma$ 2°	$-1\sigma$ 16°	0 50°	$+1\sigma$ 84°	$+2\sigma$ 98°
---------	------------------	-------------------	----------	-------------------	-------------------

Si constata che 68 centili sono raggruppati attorno alla mediana, cosa che determina una classificazione troppo approssimativa per la maggioranza del gruppo. Si rischia, per esempio, di considerare che una differenza importante separi il 75° dal 25° centile. La vecchia norma della percentuale sussiste tenacemente nelle menti ed inclina a pensare che il 75° centile vale tre volte il 25°. La tabella di cui sopra ci mostra che si tratta, in realtà, di due risultati relativamente vicini e medi.

In realtà, se le norme in centili continuano ad essere utilizzate, particolarmente per la selezione in alcune università americane, è perché queste non reclutano che nella fascia dei dieci o venti centili superiori, zona in cui la discriminazione è soddisfacente.

## (B) I voti standard o punti Z

Un alunno ha ottenuto i seguenti voti grezzi<sup>25</sup>. Calcolo: 22 su 25; lettura 72 su 100; scienze 26 su 50.

L'esame di questi voti non ci dice gran che. Se il professore è severo in lettura e molto meno in calcolo, può essere che il 72/100 in lettura sia più meritorio che il 22/25

<sup>25</sup> Adattato da R. Thomas, op. cit.

in calcolo. Inoltre, per esempio, dei voti grezzi non permettono il confronto con altri alunni dello stesso livello educativo.

Per rendere possibile il confronto, si esprimono questi voti in funzione degli scarti tipo, ciò che permette di situarli su una stessa curva (voti normalizzati o punteggi standard).

Supponiamo che si ottengano i risultati seguenti:

	Media	Scarto tipo
Calcolo	15	2,5
Lettura	50	10
Scienze	29	5

Calcolo dei voti standard:  $z = (x - M)/\sigma$

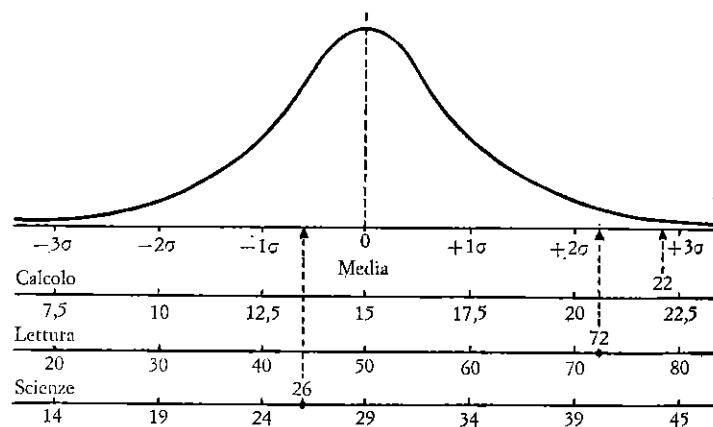
Esempio:  $(22 - 15) : 2,5 = 2,8\sigma$ .

Conoscendo  $z$ , si può ritrovare il voto grezzo:

$$x = M + z\sigma$$

$$x = 15 + (2,8 \times 2,5) = 22.$$

Rappresentazione grafica:



Applicazione. Gli studenti I e II sono stati sottoposti ad una serie di test.

Confrontare i risultati\*.

Test	Media M	Scarto tipo $\sigma$	X Punteggi grezzi		x Deviazioni (M-X)		Punteggi standard z	
			I	II	I	II	I	II
Calcolo	155,7	26,4	195	162	+39,3	+ 6,3	+1,49	+0,24
Lettura	33,7	8,2	20	54	-13,7	+20,3	-1,67	+2,48
Scienze	54,5	9,2	39	72	-15,5	+17,5	-1,67	+1,88
Attitudini psicologiche	87,1	25,8	139	84	+51,9	- 3,1	+2,01	-0,12
Intelligenza	24,8	6,8	41	25	+16,2	+ 0,2	+2,38	+0,03
Totale			434	397			+2,54	+4,51
Media							+0,51	+0,90

\* Cifre tratte da Guilford, *Fundamental Statistics* cit., p. 513.

Se ci si attenesse ai semplici totali (punteggi grezzi I e II), lo studente I sarebbe considerato superiore allo studente II. Il calcolo delle percentuali confermerebbe questa impressione.

Se attraverso il calcolo dei punteggi Z riconduciamo tutti i risultati alla stessa media e alla stessa unità di deviazione in rapporto a questa, si constata che in media, lo studente II (0,90) ottiene un risultato favorevole quasi doppio di quello dello studente I (0,51); che lo studente II ottiene i risultati più omogenei.

Critica. La trasformazione in punteggi Z si basa sull'ipotesi che, per il gruppo considerato, le attitudini sulle quali si basano i test si distribuiscano alla stessa maniera e conducano alle stesse medie e dispersioni. Questa ipotesi è praticamente impossibile da verificare, ma, come osserva Guilford, si può essere certi che questo sistema è senza dubbio migliore del calcolo delle percentuali.



## Controllo della fedeltà dell'esame

Molte delle misure che abbiamo esaminato fino ad ora tendono ad assicurare una miglior giustizia scolastica.

Il controllo della fedeltà deriva dalla stessa preoccupazione: idealmente, uno stesso esame (scritto od orale), superato più volte senza che l'alunno abbia il tempo di apprendere cose nuove, dovrebbe condurre sempre allo stesso risultato. Ci si augura dunque che sia *esatto* come un metro che, ben utilizzato, misuri sempre la stessa lunghezza, approssimata ad un errore minimo.

Evidentemente, far subire la stessa prova a diverse riprese non permette un buon controllo della fedeltà. Il primo controllo provoca una riflessione generatrice di apprendimento. In seguito, l'alunno continua a pensare al problema, verifica l'esattezza delle sue risposte, prende informazioni complementari, ecc. In breve, bisogna ricorrere ad un altro mezzo o, più esattamente, ad un insieme di misure che favoriscano la rilevazione dalla fedeltà.

Supponendo acquisita la fedeltà della correzione resta da garantire la fedeltà delle risposte.

### Evitare ogni ambiguità nei quesiti

Un quesito che manchi di chiarezza al punto da prestarsi a diverse interpretazioni differenti compromette a priori la fedeltà dell'esame. Niente permette, infatti, di prevedere con

qualche certezza quale interpretazione dei significati sarebbe stata data da uno stesso individuo o da uno stesso gruppo, in diverse occasioni.

Abbiamo preso in considerazione, a p. 77, un certo numero di mezzi atti ad assicurare la chiarezza dei quesiti.

### Quesiti in numero sufficiente

Se si pone soltanto un limitatissimo numero di quesiti, si lasciano vaste zone di materie inesplorate. Si può dare il caso che l'alunno non abbia compreso tutti gli argomenti o abbia trascurato parti del corso speculando sulla probabilità o sulla preferenza rivelata dal professore. A seconda del punto su cui andranno a cadere i quesiti, si registrerà un brillante risultato o un insuccesso.

Il solo mezzo per evitare tale evenienza è di coprire *tutta* la materia. Cioè, l'esame deve *campionare bene* l'insieme.

In linea di massima, più si aggiungono quesiti, più la fedeltà aumenta.

### Un controllo matematico

Per esami importanti è auspicabile assicurarsi garanzie ancora maggiori.

(a) *Il metodo pari-dispari*. Se il numero di quesiti è abbastanza elevato, il procedimento seguente simula una situazione in cui uno stesso alunno subirebbe, praticamente nello stesso tempo, due volte lo stesso esame.

- (a) Essendo i quesiti numerati, si dividono in due gruppi: pari e dispari.
- (b) Si compilano due tabelle parallele dei risultati e si calcola la loro correlazione ( $r$ ); in altre parole, si valuta numericamente la relazione che esiste tra i due gruppi<sup>1</sup>.

<sup>1</sup> Il metodo del calcolo della correlazione è spiegato in tutti i manuali elementari di statistica.

- (c) I due esami creati artificialmente sono di lunghezza mezzata rispetto all'originale. In virtù di quanto abbiamo visto al paragrafo 2, sono quindi meno fedeli e la correlazione sottovaluta dunque la situazione reale.

Una formula semplice permette di operare la necessaria correzione: l'indice di fedeltà è finalmente ottenuto con  $\frac{2r}{1+r}$ .

Più il risultato sarà vicino a 1, maggiore sarà la fedeltà. Si esige generalmente almeno 0,80.

Quando si prepara un esame per mezzo di un pre-test, l'indossare sufficiente fedeltà si combatte in pratica aumentando il numero dei quesiti.

Se il controllo di fedeltà si effettua dopo l'esame vero e proprio, una troppo scarsa fedeltà spingerà gli insegnanti ad un supplemento di prudenza nell'interpretazione dei risultati e ad una maggiore cura nella costruzione dei futuri esami.

(b) *Due forme parallele.* Il controllo matematico è lo stesso che nel metodo *pari-dispari*, ma qui gli esaminatori preparano due serie complete di quesiti, giudicati equivalenti all'unanimità, a partire dalla stessa rete di obiettivi.

Se la popolazione è molto numerosa, una metà, scelta a caso (per esempio, secondo l'ordine alfabetico dei nomi), riceve la prima forma e l'altra, la seconda. Per popolazioni poco numerose, si invitano gli alunni a rispondere alle due serie di quesiti dopo qualche giorno. Questa seconda soluzione è evidentemente molto meno rigorosa.

### Ripetizione della classificazione

La fedeltà delle risposte non basta. Essa deve essere accompagnata dalla fedeltà della classificazione. Questa si controlla sia invitando uno stesso professore a classificare due volte gli stessi lavori a un intervallo di alcuni giorni o di alcune settimane, sia confrontando i voti assegnati da parecchi correttori allo stesso lavoro.

## Controllo della validità

La fedeltà di un esame non garantisce affatto la sua validità. Un metro mal costruito che misurasse cinque centimetri di troppo, indicherebbe la stessa lunghezza ad ogni misura, ma condurrebbe tuttavia ad una conclusione falsa.

In breve, validare un esame, è provare che esso misura effettivamente ciò per cui viene proposto. Acquistare questa certezza è fondamentale, poiché se la conclusione è negativa tutti gli sforzi compiuti per l'organizzazione ed il passaggio degli esami sono vani o, perlomeno, inadeguati al loro oggetto.

La validazione non è cosa facile. Sommando tutto quanto sappiamo sulla validità, Cox concluse nel 1969 che la ricerca ha appena toccato le questioni fondamentali<sup>1</sup>. A maggior ragione il problema merita l'attenzione e la prudenza dei professori.

Secondo la natura retrospettiva o prospettiva dell'esame, si distingue la validità del contenuto e la validità predittiva.

### La validità del contenuto

Interessa principalmente gli esami destinati a fare il bilancio delle acquisizioni e quindi anche di un insegnamento. Que-

<sup>1</sup> R. Cox, *Reliability and Validity of Examinations*, in J. Lauwerys e D. Scanlon, *Examinations* cit., p. 43.

sto lo scopo che era stato fissato, questi gli obiettivi scelti. Sono stati raggiunti?

Negli esami che vertono su lunghe tappe di apprendimento, è evidente che non si può interrogare in particolare tutta la materia. Non solo, dunque, è importante localizzare i punti principali, ma bisogna inoltre tener conto della loro importanza relativa. Se si soddisfa a questi due criteri, i quesiti potranno suscitare risposte *rappresentative* dell'insieme delle acquisizioni, della competenza totale.

In altri termini, la validità del contenuto dipende dalla qualità e dall'adeguatezza del campionamento delle conoscenze e delle capacità.

Abbiamo visto che, in un insegnamento ben concepito, gli obiettivi da raggiungere, e dunque da controllare al momento dell'esame, sono stati definiti all'inizio dell'anno.

Ma allora si trattava d'intenzioni che le circostanze possono aver modificato. Prima di cimentarsi a fare il bilancio del lavoro dei suoi alunni, il professore dunque farà innanzi tutto il suo. In mancanza di questo ripensamento sul proprio lavoro, gli esami non hanno che una validità illusoria o, peggio ancora, una validità di facciata destinata a scaricare le responsabilità del professore agli occhi dei suoi superiori. Così si spiega che alcuni quesiti si riferiscano ad argomenti appena sfiorati, ma che figurano nel programma ufficiale, a capacità intellettuali elevate che si è omesso di impiantare pazientemente e sistematicamente durante l'anno. Quanti insuccessi scolastici sono dovuti ai bei quesiti « d'intelligenza », posti al termine di un insegnamento che non l'ha coltivata affatto?

Se la definizione precisa degli obiettivi permette una convalida di contenuto molto migliore che nel passato, questa nondimeno resta ancora molto limitata allo stato attuale delle nostre conoscenze. In primo luogo perché è molto difficile acquistare la certezza che quel comportamento reale è proprio rappresentativo di quell'aspetto che si desiderava coltivare. In secondo luogo, perché anche se la visione è corretta in partenza le vicende della classificazione possono sempre tradire le intenzioni.

Un professore di storia o geografia, scrive Vernon<sup>2</sup>, che, in un esame scritto, cerca di giudicare la comprensione profonda di queste discipline, attribuisce spesso una parte dei punti per la ripetizione espositiva esatta di fatti partecipi, la ricostruzione delle sue teorie favorite, la lunghezza delle risposte, la loro intelligibilità e l'eleganza dello stile col quale esse sono espresse. In un esame orale, quanti sono i punti che non dipendono dall'intelligenza sociale, dal controllo di sé e specialmente dall'attitudine dell'alunno a creare con la parola e l'atteggiamento un buon rapporto con l'esaminatore?

Cosa fare in pratica?

In primo luogo esigere che per ogni quesito d'esame che si formula, il professore indichi l'obiettivo perseguito su una speciale scheda.

Poi, sottomettere i quesiti e lo schema di classificazione a diversi educatori esperti (di cui uno almeno apparterrà ad un gruppo di materie diverse dal settore cui si riferisce) che giudicheranno indipendentemente dalla validità del contenuto dei quesiti. In caso di grave disaccordo il quesito deve essere respinto<sup>3</sup>.

L'esperienza dimostra che attualmente bisogna attenersi ad un piccolo numero di obiettivi generali concepiti in maniera abbastanza ampia. Altrimenti i giudici non riuscirebbero più a mettersi d'accordo. L'abbiamo detto prima, le nostre conoscenze in materia di convalida sono ancora imperfette.

Per la stessa ragione, in alcuni casi nessun accordo sarà raggiunto sulle questioni fondamentali. Senza dubbio questo non è un motivo sufficiente per cessare d'insegnare l'argomento o la materia incriminata, ma l'impossibilità di accordo deve allora essere chiaramente indicata.

Si rischia per esempio di ritrovare tale difficoltà a propo-

<sup>2</sup> P. Vernon, *Types of Examinations*, in J. Lauwerys e D. Scanlon, *Examinations*, p. 43.

<sup>3</sup> Un coefficiente di concordanza tra le opinioni può essere calcolato con la formula di Kendall ed una soglia d'accettazione scelta di comune accordo.

sito del controllo sistematico della realizzazione degli obiettivi assegnati alle lingue antiche, ad alcuni argomenti della matematica, ecc.

Di nuovo, constatiamo che nel cercare di accrescere il rigore scientifico degli esami, alcuni insegnamenti saranno messi in causa. Il feticismo pedagogico non scomparirà che a questo prezzo.

### La validità predittiva

Essa sembra, dal di fuori, la più semplice da controllare, poiché non esige comprensione profonda dei fenomeni. Anche se non si spiega il rapporto che esiste tra la riuscita in un dato test ed un brillante successo in un ambito determinato, il fatto può essere osservato agevolmente.

Che si tratti di un esame che dà accesso ad un nuovo ciclo di studi o che autorizzi l'esercizio di una professione, basta seguire gli individui per un certo periodo di tempo per sapere se il pronostico formulato a partire dai voti attribuiti si verifica o no.

Si constata con meraviglia che tale verifica è raramente positiva. Passeron vi scorge il segno che l'esame è destinato meno a misurare obiettivamente la capacità che a servire tutti i bisogni *sociali* delle classi privilegiate<sup>4</sup>. Questa affermazione merita certo delle riserve, ma abbiamo visto nella prima parte che essa non è interamente priva di fondamento.

Perché non sia un'illusione, la verifica sistematica della validità predittiva esige il controllo rigoroso delle variabili; in mancanza di tale controllo le variabili nascoste falseranno le conclusioni. Per esempio, recentemente si traeva ancora argomento in favore del latino facendo la statistica del numero di studenti universitari brillanti che avevano fatto studi secondari classici. Ora, questa osservazione non prova niente, se non che, tradizionalmente nei nostri paesi, gli alunni più in-

<sup>4</sup> J.-C. Passeron, *Sociologie des examens* cit., p. 7.

elligenti sono orientati fin dai dodici anni verso questo tipo di studi.

A mo' di sintesi, attingiamo la seguente ricapitolazione dall'« Examination Bulletin » 3<sup>5</sup>.

### Fattori di validità di un esame

- (1) Identificazione adeguata degli obiettivi la cui realizzazione deve essere verificata dall'esame scritto, la valutazione del lavoro dell'anno, la prova pratica, ecc.
- (2) Scelta degli obiettivi sui quali si concentrerà la verifica, nella gamma di quelli così identificati.
- (3) Valutazione efficace dell'adeguamento del contenuto e della struttura dell'esame allo scopo perseguito.
- (4) Relazione chiara tra ogni quesito e gli obiettivi dell'insegnamento.
- (5) Elaborazione di uno schema di classificazione e formulazione di direttive ai correttori in funzione degli obiettivi. Deve essere assicurato un sufficiente livello di fedeltà.
- (6) Buona conoscenza delle capacità dei candidati.
- (7) Disposizione dei classificatori a tener conto dei giudizi indipendenti dal loro e dei dati obiettivi che saranno loro forniti.
- (8) Confronto con prove anteriori la cui validità è stata provata.

<sup>5</sup> London, HMSO, 1964, pp. 19-20.



IV

Le procedure di moderazione

## posizione del problema

### Definizione

La *moderazione*, in docimologia, ha dapprima avuto come scopo di moderare gli eccessi di severità o di generosità in alcuni esaminatori.

In senso lato, oggi essa indica l'insieme delle misure prese per rendere confrontabili i voti d'esame interni e dunque per unificare il loro significato al livello delle diverse classi dello stesso tipo in una scuola, in un gruppo di scuole, in istituti simili di una regione o di un paese.

Le ricerche più raffinate sulla moderazione sono certamente dovute ai docimologi britannici. In un sistema completamente decentralizzato come quello della Gran Bretagna, i programmi ed i metodi delle scuole primarie o secondarie possono essere considerevolmente diversi. Così, per lungo tempo esami esterni, cioè organizzati al di fuori delle scuole, con commissioni speciali, hanno dovuto sanzionare gli studi nei momenti cruciali della scolarità. Conosciamo il famoso Eleven-plus-Examination, esame di ammissione alle secondarie, che decide in particolare dell'ammissione all'insegnamento classico, e il CSE (Certificate of Secondary Education), prima dell'ingresso all'insegnamento superiore. Quest'ultimo esame, per esempio, fornisce ogni alunno di un documento che indica il livello raggiunto in ciascuna materia, in rapporto a norme nazionali.

Per ammettere un alunno in una determinata sezione, le università dichiarano le loro particolari esigenze: questa ri-

chiede voti eccellenti (A) in matematica, in biologia e in lingua materna; quella accetta anche la votazione « molto bene » (B), ecc.

Questo sistema di esami esterni è criticato, a volte molto severamente, perché limita gravemente la libertà pedagogica dei maestri e delle autorità locali.

Così, dopo qualche anno, si tenta di sostituirli con esami interni i cui risultati sono resi confrontabili con procedure di moderazione locale, regionale, poi nazionale, messe a punto sperimentalmente.

In un paese come il Belgio, dove un unico programma di studi è imposto da un'autorità centrale, è, se vogliamo, molto più facile moderare gli esami. Prima di vedere come, forse non è inutile ricordare perché è auspicabile una moderazione.

Per sommi capi, si può parlare di un doppio motivo. Il primo riguarda gli individui, alunni e genitori, che hanno il diritto di conoscere il livello « reale » delle performance scolastiche, prima di decidere l'ulteriore orientamento nel campo degli studi o della professione. Chi si crederrebbe qualificato per i giochi olimpici perché ha vinto una corsa organizzata tra alcuni amici? D'altra parte un verbale sfavorevole è più una diagnosi di debolezza o di malattia che una condanna senza appello. È necessario sapere in tempo che un problema si pone per cercarne le cause, poi i rimedi, se essi esistono.

Il secondo motivo si riferisce alla comunità. Come un consumatore non può essere imbrogliato sulla merce, così la società non può essere tenuta o a pagare molto cari gli insuccessi di uno studente entrato all'università sulla fiducia in un certificato non valido, o ad affidare una funzione a qualcuno che non la merita. Per esempio, il giorno in cui gli esami nei corsi di preparazione degli insegnanti saranno rigorosamente controllati, si può sperare di non veder più affidare generazioni di bambini a certi maestri la cui conoscenza della lingua materna, particolarmente dell'ortografia<sup>1</sup>, è insufficiente.

<sup>1</sup> Non abbiamo affatto il feticismo dell'ortografia, ma riteniamo evidente che debba conoscerla colui che la deve insegnare.

Moderare non è sottoporre ad un regime autoritario

Si potrebbe temere che la volontà di rendere confrontabili i risultati scolastici provochi il risorgere di vecchi contrasti. Ci si ricorda di quell'ispettore, di tradizione napoleonica, che consultando il suo orologio, credeva di poter dire: « In questo momento, si insegna tale lezione in tutte le quinte elementari di Francia ». Conosciamo i punti deboli di un insegnamento autoritario.

Ma la necessità di una plasticità pedagogica non impedisce che, al di là degli adattamenti contingenti, ogni ciclo di studi persegua degli obiettivi fondamentali, comuni a tutti: acquisizione di tecniche di base o di conoscenze e di capacità giudicate essenziali. Maestri e moderatori devono interpellarsi l'un l'altro ed accordarsi su questi apprendimenti fondamentali.

Moderazione volontaria o imposta?

Normalizzare tutti gli esami a ciascuna tappa del corso di studi non sembra né auspicabile, né del resto possibile. Poco auspicabile, poiché gli alunni progrediscono secondo ritmi a volte molto diversi e mal si adattano alla rigorosa suddivisione in periodi ed in anni scolastici. Impossibile perché la moderazione, anche con i metodi economici che ora saranno presentati, resta sempre un lavoro pesante, nella sua preparazione e nella sua esecuzione.

La moderazione deve essere imposta nei momenti decisivi, innanzi tutto al momento dell'attribuzione di diplomi o di certificati di fine ciclo.

Per il resto, la decisione dovrebbe essere lasciata sia ai maestri individualmente, sia ai capi di istituto o alle autorità locali.

La moderazione comincia all'inizio dell'anno scolastico

Come abbiamo ora accennato, è importante che professori e moderatori abbiano potuto trovare un accordo sugli

obiettivi dell'insegnamento e su alcuni grandi principi cui informino le classificazioni perché gli esami siano confrontabili.

Affermazioni vaghe quali: « L'alunno deve potersi esprimere correttamente per iscritto » sono inutili. La capacità desiderata deve essere tradotta in termini di comportamenti concreti, che si avrà del resto interesse ad inserire nelle scale di valutazione descrittive. Se si tratta dell'espressione scritta, si pongono molti problemi: come graduare le esigenze in funzione dell'età degli alunni e del tipo di scuola? l'originalità avrà un riconoscimento nella valutazione? in che modo? come sarà identificata? quale importanza attribuire all'ortografia? sarà considerata la lunghezza dei lavori di composizione? si esigerà una lunghezza minima, al di sotto della quale saranno scalati dei voti? quanti? quale importanza dare alla ricchezza di vocabolario? come definirla? quali saranno le esigenze sintattiche?

Anche i matematici possono accordarsi fra di loro. Se, per limitarci ad un solo esempio, alcuni professori considerano gravi gli errori di operazioni e perfino gli errori di ortografia nella soluzione di problemi, mentre altri loro colleghi reputano questi aspetti come secondari, voti uguali rischiano di nascondere realtà molto diverse.

I programmi ufficiali possono anche avere un'importante funzione precisando al massimo gli scopi da raggiungere ad ogni grado. L'efficacia di questi dati di base è tanto maggiore perché il sistema scolastico è centralizzato.

È inutile farsi illusioni: gli obiettivi comportamentali delle diverse materie non si definiscono in qualche ora di riflessione.

Commissioni in cui insegnanti, ispettori e ricercatori uniranno i loro sforzi, dovranno lavorare a lungo prima di arrivare ad un risultato soddisfacente. Peraltro molte domande rischiano di restare provvisoriamente senza risposta.

La definizione degli obiettivi formerà l'oggetto di una ricerca permanente, non soltanto perché essi potranno precisarsi in funzione del procedere della psicologia dell'apprendimento, ma anche perché gli stessi obiettivi si evolvono.

Si preciseranno parallelamente le istruzioni relative agli esami. A fianco delle griglie di obiettivi generali e speciali, i

programmi scolastici dell'avvenire determineranno nettamente la quantità e i tipi di quesiti da porre agli esami e forniranno indicazioni sulle linee di forza degli schemi di classificazione.

Non è possibile una confrontabilità senza una fedeltà elevata

Non è concepibile confrontare tra loro risultati d'esame che, presi da soli, fossero eminentemente fluttuanti.

Il problema della fedeltà è stato discusso nella parte dedicata alla costruzione dell'esame. Non ci torneremo.

Si può fare affidamento sui test?

Nei sistemi di moderazione che stiamo per esaminare, i test d'intelligenza oppure di conoscenze occupano un posto importante. In alcuni casi essi danno i punti di riferimento sui quali gli altri risultati vanno adattati; in altri casi si attribuisce ad essi soli un valore pari a quello del lavoro di un intero anno e dell'esame finale.

Ora che abbiamo visto con quale rigore e quali precauzioni i buoni test di conoscenze sono costruiti, non ci deve più stupire la fiducia accordata a questi strumenti. Del resto, fin dalla prima parte di questo libro, abbiamo visto che se i maestri giudicano bene i loro alunni, un semplice test verbale, amministrato in meno di un'ora, permette una predizione quasi altrettanto sicura:

	Correlazione con risultati dell'alunno:	
	2 anni dopo	3 anni dopo
Pronostico dell'insegnante	.821	.748
Test verbale	.796	.722

Ecco ancora alcuni risultati di ricerche che confermano queste osservazioni.

A medio termine, un test d'intelligenza o meglio ancora, una combinazione di punteggi in diversi test d'intelligenza, sono migliori predittori del successo scolastico che i risultati d'esami. Questa superiorità è stata dimostrata a più riprese, soprattutto da ricercatori anglosassoni:

- (1) Emmet<sup>2</sup> dimostra che un test d'intelligenza verbale permette di prevedere meglio i risultati, dopo due o tre anni d'insegnamento secondario generale, che gli esami di lingua materna e d'aritmetica valutati dai maestri.
- (2) Emmet e Wilmut<sup>3</sup> hanno, in seguito, fatto una dimostrazione altrettanto convincente del valore predittivo del test d'intelligenza, cinque anni dopo.
- (3) Wrigley<sup>4</sup> ha confermato questi risultati e dimostrato che la predizione può essere ancora migliore se possono essere combinati i risultati ad un test d'intelligenza e a test di conoscenza standardizzati.
- (4) In Belgio, tutti gli psicologi conoscono il valore predittivo dei punteggi verbali e di ragionamento del test PMA di Thurstone, il famoso 2V+R, per gli studi secondari generali.

Non disponiamo di ricerche simili per gli studi tecnici o per strutture d'insegnamento forse meglio adatte alla civiltà dell'anno 2000. È molto probabile che dovrebbero essere utilizzati i punteggi di altri tipi di test.

<sup>2</sup> W. Emmet, *An Inquiry into the Prediction of Secondary School Success*, London, University of London Press, 1942.

<sup>3</sup> W. Emmet e F. Wilmut, *The Prediction of School Certificates Performance in Specific Subjects*, in «British Journal of Educational Psychology» 22 1952, pp. 52-62.

<sup>4</sup> J. Wrigley, *The Relative Efficiency of Intelligence and Attainment Tests as Predictors of Success in Grammar Schools*, in «British Journal of Educational Psychology» 25 1955, pp. 107-116.

## Alcuni sistemi di moderazione degli esami

La formula più liberale: il sistema svedese di moderazione per materia a partire da un test di conoscenze

Dalla fine della seconda guerra mondiale, la Svezia ha adottato un sistema di moderazione, semplice e facoltativo, applicabile da ogni professore nella sua classe, per tutta la durata della scolarità.

Tutti i voti dei professori sono assegnati secondo una scala di 7 gradi<sup>1</sup>.

Si suggerisce che la percentuale di alunni che ricevono un voto determinato sia, grosso modo, la seguente<sup>2</sup>, restando il professore libero di tener conto delle caratteristiche della sua classe (per esempio: una classe che ha una fisionomia molto precisa, o alunni quasi tutti mediocri, ecc.): voti: 1, 2, 3, 4, 5, 6, 7; percentuale: 1, 6, 24, 38, 24, 6, 1.

Per altro, una batteria di test di conoscenze basati sulle materie principali del programma è elaborata ogni anno da un ufficio centrale a cui collaborano specialisti e professori esperti.

Questa batteria è tarata su un campione nazionale rappresentativo; i risultati sono divisi in 7 classi, nella proporzione consigliata per i voti dei professori.

<sup>1</sup> Ci riferiamo all'esposizione molto chiara di S. Henrysson, *The Swedish System of Equalising Marks*, in «Educational Research» 2 1964, pp. 156-160.

<sup>2</sup> Distribuzione normale (vedere curva di Gauss).

Se il professore amministra il test nazionale nella sua classe, dispone dunque di norme direttamente confrontabili ai voti che ha assegnato. Gli è dunque più facile adattare questi ultimi.

Ecco un esempio concreto proposto da S. Henrysson:

24 alunni			
Voti	Distribuzione dei voti preliminari assegnati dal professore	Distribuzione dei voti ottenuti amministrando il test	Distribuzione dei voti dopo l'adattamento
7	—	1	—
6	—	4	4
5	4	4	4
4	4	9	10
3	14	5	4
2	2	1	2
1	—	—	—
Media	3,42	4,33	4,17

Il professore scrive dapprima, nella seconda colonna, i voti che ha assegnato. La media (3,42) fornisce una prima indicazione, se la media nazionale è conosciuta. (In Svezia, essa è di 4 per la 6<sup>a</sup> elementare. Sembra dunque che il professore sia qui troppo severo).

Ecco ora i risultati ottenuti amministrando il test con gli stessi alunni:

Punteggio grezzo	Voto corrispondente	Numero di alunni che ottengono questo voto
94 - 100	7	1
84 - 93	6	4
67 - 83	5	4
46 - 66	4	9
30 - 45	3	5
22 - 29	2	1
0 - 21	1	—

Questi risultati sono riportati nella prima tabella, colonna 3. La media è 4,33, ciò che conferma la prima impressione di severità.

Senza cambiare l'ordine della classificazione iniziale, il professore rettifica i suoi voti (colonna 4).

Il sistema svedese seduce per molte ragioni.

Gli insegnanti conservano interamente la loro libertà di far passare o no il test (lo fanno quasi tutti), di tener conto dei risultati, di comunicare i risultati dei test agli alunni, agli altri professori, al direttore, ai genitori. Così il test diventa semplicemente uno strumento messo a disposizione degli esperti. La tentazione di una « preparazione approssimativa » in vista di un esame è quindi ridotta al minimo.

Gli elaboratori dei test stanno attenti a coprire una vasta gamma di obiettivi campionando bene tutto il programma.

È evidente che i professori prestano una particolare attenzione alle materie trattate nei test. Si dispone così di un mezzo efficace per sensibilizzare l'insegnante a talune innovazioni.

Rileviamo che, dopo il 1965, un sistema molto simile è sperimentato in classi ginevrine di 5<sup>a</sup> e di 6<sup>a</sup> elementare (10-12 anni) per l'ortografia e l'aritmetica<sup>3</sup>.

Tra le critiche formulate contro questo sistema<sup>4</sup>, ricordiamo:

- (1) Se un professore ha giudicato uno degli alunni della sua classe con una severità eccessiva, il riferimento alla prova nazionale non ripara all'ingiustizia.
- (2) La prova di riferimento può non misurare le stesse dimensioni misurate dai voti.
- (3) La costruzione della prova di riferimento può anche essere difettosa: fedeltà troppo bassa, ecc.

<sup>3</sup> Vedere S. Roller, *Le problème de l'attribution des notes scolaires. Essai de solution*, in *Docimologie et éducation*, numero speciale della rivista « Les sciences de l'éducation », aprile-settembre 1969, p. 66 ss.

<sup>4</sup> W. Angoff, *Can Usual General Purpose Equivalency Tables Be Prepared for Different College Admission Tests?*, in A. Anastasi (ed.), *Testing Problems in Perspective*, Washington, ACE, 1966, pp. 251-264. F. Bacher, *La normalisation des notes* cit., p. 63.

### Sistema imposto di moderazione per materia a partire da un test di conoscenze

A partire dagli stessi dati di base di quelli della Svezia (esami interni e test di conoscenze, per le materie principali, tarati nazionalmente o regionalmente), il sistema organizzativo seguente può dare buoni risultati.

Una commissione nazionale o regionale di moderazione si incarica di un certo numero di scuole.

Queste scuole mandano ad essa i risultati delle prove liberamente preparate dai professori ed i punteggi ottenuti amministrando i test di conoscenze.

In caso di differenza, in più o in meno, uguale o superiore a due scarti tipo (per esempio) tra le medie riportate agli esami e ai test, la scuola riceve la visita dei moderatori.

Il loro compito non è quello di dire al direttore come deve condurre la sua scuola, né al professore come svolgere il suo corso, bensì di attrarre l'attenzione su un fatto e di cercare di trovare, in piena collaborazione con la scuola, la spiegazione e, se è possibile, il rimedio.

Se il disaccordo sussiste, può essere dato alle autorità regionali il diritto di adattare d'ufficio i voti; da parte sua la scuola ha il diritto di interporre appello.

### Moderazione volontaria con ricorso ad una banca di item

Nel caso in cui i maestri utilizzino quesiti sotto forma di item a scelta multipla, il seguente procedimento, attualmente sperimentato in Gran Bretagna sembra pieno di promesse<sup>5</sup>. Tuttavia, esso richiede l'esistenza di un servizio di ricerca, necessità di cui oggi un educatore avveduto, del resto, non può più dubitare.

<sup>5</sup> Vedere D. Pidgeon e A. Yates, op. cit.

Gli insegnanti indicano su una griglia gli obiettivi perseguiti. Essi inviano questa griglia al servizio di ricerca, assieme agli item che hanno redatto e, eventualmente, già provati in pratica in loco.

Il servizio esamina questi item e, secondo le possibilità, ne prova alcuni; vengono calcolati la loro difficoltà e il loro potere discriminativo.

Il servizio rimanda all'insegnante gli item esaminati e degli item complementari i cui indici di facilità e di efficacia sono conosciuti per una data popolazione: città, cantone, paese, ecc.

Questi item forniranno un termine di confronto a partire dal quale potrà essere adattato il risultato registrato per gli altri.

È interessante notare che dopo essersi strettamente fissati alla redazione di item a risposta chiusa, con scelta multipla, gli inglesi cominciano a proporre dei quesiti semi-aperti. Si tratta infatti di quesiti aperti formulati molto accuratamente, per i quali si sono identificati sperimentalmente i tipi fondamentali di risposte, in funzione delle quali sono proposte delle tabelle di correzione.

L'organizzazione di una banca di item è gravosa durante i primi anni. In seguito, la scorta di item tarati diviene tale che il lavoro si alleggerisce e il sistema acquista una grande agilità e una grande rapidità di funzionamento.

### Un sistema semplice di moderazione globale al servizio della selezione

Abbiamo già osservato che in generale, gli insegnanti giudicano bene i loro alunni, ma che tendono a condizionare i loro giudizi in rapporto al livello globale della classe.

Il sistema seguente, messo a punto dalla Fondazione nazionale inglese per la ricerca educativa<sup>6</sup> per l'ammissione all'insegnamento generale classico, permette di *selezionare* tenendo

<sup>6</sup> Vedere D. Pidgeon e A. Yates, op. cit.





prende un margine di sicurezza di tre volte l'errore standard che una formula semplice permette di calcolare a partire dal coefficiente di fedeltà del test utilizzato<sup>7</sup>.

### In Inghilterra, un sistema di moderazione completo

Il sistema che ora descriveremo si riferisce, anch'esso, alla Gran Bretagna. Che noi sappiamo, non vi è niente di più completo: si basa sulla preparazione dell'esame in collaborazione con le scuole, la classificazione, l'adattamento dei voti e dei livelli finali.

Lo scopo perseguito è di *perfezionare gli esami interni*, fino a poter affidare loro la funzione che fino ad ora era sostenuta dalle grandi prove esterne.

Per giungere ad una moderazione nazionale, si procede per gradi: su scala locale si raggruppano un certo numero di scuole che perseguono una stessa finalità; poi su scala regionale, secondo una tecnica che non differisce fondamentalmente da quella applicata localmente; per il passaggio al livello nazionale, si aggiunge alla tecnica di moderazione propriamente detta, una campionatura molto elaborata il cui studio tecnico sarebbe qui fuori luogo.

Suggeriamo, per paesi come il Belgio, che alcune scuole animate dallo stesso spirito comincino col fare un'esperienza volontaria del sistema. È per questo che concentriamo la nostra presentazione sul processo di base.

La moderazione inter-scuole non sarà fatta che al momento cruciale della scolarità e per alcune materie principali. Niente impedisce, tuttavia, ai professori di una stessa scuola, che insegnano le stesse materie, di utilizzare spontaneamente lo stesso metodo di correzione.

<sup>7</sup> Errore standard =  $\sigma \sqrt{1-r}$ , ove  $\sigma$  = scarto tipo dei punteggi;  $r$  = coefficiente di fedeltà. Questo coefficiente è indicato generalmente nel manuale che accompagna il test.

### A) Preliminari

Abbiamo già alluso agli accordi da prendere sugli obiettivi e alla costruzione dell'esame. Non vi ritorneremo.

Ogni scuola sceglierà, tra i suoi, un professore che avrà due funzioni: coordinatore degli esami nella sua scuola e membro della commissione di moderazione inter-scuole.

Un ritratto ideale del moderatore non esiste. Sembra auspicabile che possieda le qualità seguenti<sup>8</sup>: maturità generale e buona esperienza pedagogica; contatti frequenti con alunni del livello e del tipo di scuola cui l'esame si riferisce; capacità di formulare chiaramente i suoi criteri e i suoi giudizi; bontà ma fermezza; capacità di discutere senza passione, senza creare tensione; attitudine a comprendere alcune tecniche d'analisi statistica.

In ciascuna scuola, i quesiti d'esame sono formulati con una grande libertà di movimento: sono da rispettare, secondo l'accordo preso, soltanto la griglia dei principali obiettivi, il numero e l'estensione dei quesiti.

All'inizio, sarebbe auspicabile sottoporre i quesiti alla commissione di moderazione che si assicuri dell'unità generale.

(B) I professori correggono i loro esami

(C) Nuova correzione di un campione di compiti da parte dei moderatori

Il lavoro è improntato a tre principi.

- (1) L'intervento delle scuole negli esami deve essere il più discreto possibile. Solo i casi di divergenze importanti meritano attenzione.
- (2) I correttori devono essere, in genere, ugualmente severi, apprezzare le stesse qualità ed essere d'accordo sul significato dei voti finali.
- (3) I campioni da correggere devono essere piccoli, e il lavoro statistico deve essere il più semplice possibile.

<sup>8</sup> D. Mather e altri, op. cit., p. 67.

In ogni scuola e per uno stesso tipo di esame, si prelevano, a caso, 20 compiti corretti.

Prendiamo il caso di 12 licei<sup>9</sup>. Ciascuno ha delegato un moderatore alla commissione, la quale riceve dunque 12 pacchi di 20 compiti.

La prima operazione mira a determinare l'equivalenza dei moderatori da tre punti di vista:

- (1) severità: il suo grado è rivelato dalla media;
- (2) discriminazione: votazione troppo disinvolta o troppo prudente: rivelata dalla dispersione o margine di variazione dei voti;
- (3) conformità: uno stesso alunno è classificato allo stesso modo da tutti i correttori: rivelata dalla correlazione tra due serie di voti.

Per verificare l'accordo tra moderatori, da questi tre punti di vista, i dodici ricorreggono dapprima uno stesso pacchetto di 20 compiti.

I 20 compiti vengono divisi, sempre a caso, in cinque gruppi di quattro. In ogni gruppo gli alunni sono classificati in ordine alfabetico.

Dopo si fanno le semplici operazioni che seguono, nell'ordine in cui esse figurano nella tabella a fronte.

#### Controlli

Le seguenti regole non hanno niente di magico. Per formularle, Peaker si è ispirato alle tecniche di controllo di qualità utilizzate nell'industria. Esperienze molto avanzate nell'ambito degli esami hanno permesso di adattarle.

Il controllo così realizzato è ragionevole, spedito, efficace, ma non molto acuto. Inoltre, se, nonostante la grossolanità dei criteri, un aspetto della correzione sembra inaccettabile, vi è certamente un problema!

<sup>9</sup> Esempio preso dall'« Examinations Bulletin » 5, London, HMSO, 1965.

Studio dell'accordo tra i moderatori: voti assegnati da 12 moderatori.

Nome del candidato	Moderatori												Media
	a	b	c	d	e	f	g	h	i	j	k	l	
A	4	4	4	3	4	4	3	1	4	4	4	4	4
B	2	2	2	2	3	2	1	1	2	2	2	2	2
C	6	5	5	5	5	5	5	5	5	5	5	5	5
D	2	1	1	2	2	3	3	2	2	3	2	3	2
Totale	14	12	12	12	14	14	12	9	13	14	13	14	13
Margine di variazione	4	4	4	3	3	3	4	4	3	3	3	4	3
E	3	3	3	4	3	5	3	3	3	3	3	3	3
F	1	1	1	1	1	1	1	1	1	1	1	1	1
G	5	6	6	6	5	3	6	6	6	6	6	6	6
H	3	4	4	4	3	4	3	4	4	2	4	4	4
Totale	12	14	14	15	12	15	13	14	14	12	14	14	14
Margine di variazione	4	5	5	5	4	4	5	5	5	5	5	5	5
I	2	2	2	2	4	3	2	3	3	3	3	3	3
J	2	1	2	1	3	2	2	2	2	2	2	2	2
K	5	5	5	6	5	5	5	6	5	5	5	4	5
L	1	2	2	1	2	2	2	3	2	2	3	2	2
Totale	10	10	11	10	14	12	11	14	12	12	13	11	12
Margine di variazione	4	4	3	5	3	3	3	4	3	3	3	2	3
M	4	2	3	4	3	4	5	5	5	4	5	4	4
N	4	3	4	4	4	5	4	4	4	4	4	4	4
O	5	4	4	3	4	3	4	4	4	5	4	5	4
P	5	6	5	5	5	5	5	5	5	5	5	5	5
Totale	18	15	16	16	16	17	18	18	18	18	18	18	17
Margine di variazione	1	4	2	2	2	2	1	1	1	1	1	1	1
Q	1	2	1	1	1	1	1	1	1	1	2	1	1
R	1	1	2	1	1	2	1	1	1	1	1	1	1
S	3	3	2	3	3	2	2	3	3	3	3	3	3
T	2	1	2	1	2	2	1	2	2	2	2	2	2
Totale	7	7	7	6	7	7	5	7	7	7	8	7	7
Margine di variazione	2	2	1	2	2	1	1	2	2	2	2	2	2
Totale generale	61 <sup>2</sup>	58	60	59	63	65	59	62	64	63	66	64	63 <sup>1</sup>
Somma dei margini	15 <sup>4</sup>	19	15	17	14	13	14	16	14	14	14	13	14 <sup>1</sup>

#### (1) Severità

L'esperienza rivela che gli sforzi di armonizzazione devono basarsi soprattutto su questo punto. Vedremo ora che, in caso di problema, la soluzione per fortuna è facile.

Differenze tra i voti assegnati da ciascun moderatore e la media dei voti assegnati da tutti i moderatori ad uno stesso candidato.

Nome del candidato	Moderatori											
	a	b	c	d	e	f	g	h	i	j	k	l
A	0	0	0	-1	0	0	-1	-3	0	0	0	0
B	0	0	0	0	1	0	-1	-1	0	0	0	0
C	1	0	0	0	0	0	0	0	0	0	0	0
D	0	-1	-1	0	0	1	1	0	0	1	0	1
Margine di variazione	1	1	1	1	1	1	2	3	0	1	0	1
E	0	0	0	1	0	2	0	0	0	0	0	0
F	0	0	0	0	0	0	0	0	0	0	0	0
G	-1	0	0	0	-1	-1	0	0	0	0	0	0
H	-1	0	0	0	-1	0	-1	0	0	-2	0	0
Margine di variazione	1	0	0	1	1	3	1	0	0	2	0	0
I	-1	-1	-1	-1	1	0	-1	0	0	0	0	0
J	0	-1	0	-1	1	0	0	0	0	0	0	0
K	0	0	0	1	0	0	0	1	0	0	0	-1
L	-1	0	0	-1	0	0	0	1	0	0	1	0
Margine di variazione	1	1	1	2	1	0	1	1	0	0	1	1
M	0	-2	-1	0	-1	0	1	1	1	0	1	0
N	0	-1	0	0	0	1	0	0	0	0	0	0
O	1	0	0	-1	0	-1	0	0	0	1	0	1
P	0	1	0	0	0	0	0	0	0	0	0	0
Margine di variazione	1	3	1	1	1	2	1	1	1	1	1	1
Q	0	1	0	0	0	0	0	0	0	0	1	0
R	0	0	1	0	0	1	0	0	0	0	0	0
S	0	0	-1	0	0	-1	-1	0	0	0	0	0
T	0	-1	0	-1	0	0	-1	0	0	0	0	0
Margine di variazione	0	2	2	1	0	2	1	0	0	0	1	0
Somma dei margini	4	7	5	6	4	8	6	5	1	4	3	3

Regola: (a) partire dalla media<sup>10</sup> dei totali generali (in questo caso 63); (b) non tollerare scarti al di là di un margine centrale di 10 punti<sup>10</sup> in rapporto a questa media, cioè 5 punti in più o in meno (in questo caso: 58-68).

<sup>10</sup> I classificatori attribuiscono i voti a 20 composizioni da 1 a 5. Dieci punti equivalgono a una differenza media di un mezzo punto a compito per l'insieme dei 20 compiti.

Constatazione: tutti i totali generali sono compresi in questo margine; tutti i correttori sono dunque di una severità accettabile.

Soluzione in caso di eccessi di severità o di generosità. Supponiamo che ci si trovi davanti ai seguenti totali: 75-72-68-66-65-64-60-49-46-44-41-40. Si constata che se si riunisce il più severo e il più generoso, poi il secondo per severità ed il secondo per generosità, ecc., le medie delle coppie sono molto vicine alla media generale.

Sia:				Media
(1)	40	75		57,5
(2)	41	72		56,5
(3)	44	68		56
(4)	46	66		56
(5)	49	65		57
(6)	60	64		62
	Media generale			57,5
	Margine di 10 punti			52,5-62,5

Soltanto i due correttori della coppia (6) si pongono tra 52,5 e 62,5. Questi due correttori saranno autorizzati a correggere da soli. Negli altri casi i compiti dovranno essere visti dai due membri di uno dei gruppi costituiti, e la media dei punti sarà fatta.

### (2) Discriminazione

Regola: il totale dei margini medi (3) non può essere superiore al doppio del totale dei margini di un correttore (4) e viceversa.

Constatazione: nessun problema.

### (3) Conformità

Regola: il totale dei margini delle differenze (5) tra i voti assegnati da un moderatore e la media dei voti assegnati dal-

l'insieme dei moderatori assolutamente non deve essere superiore a 12<sup>11</sup>.

Costatazione: nessun problema.

Conclusione: nel caso in questione, tutti i moderatori hanno superato le tre prove di controllo. Essi potranno dunque lavorare da soli.

A che punto siamo? Di dodici campioni di 20 compiti, un campione è ora corretto definitivamente (l'hanno visto dodici classificatori!).

Siccome ogni moderatore può lavorare da solo, il resto dell'operazione procederà speditamente.

#### (D) Nuova correzione dei campioni che restano e controllo

Le operazioni sono praticamente le stesse che per il controllo dei moderatori. Diamo, tuttavia, un esempio particolareggiato perché la presentazione più concisa (un solo moderatore e un solo professore) dà una migliore visione complessiva.

#### Controlli

(1) *Severità dei correttori*. La differenza tra totale (1) e totale (2) non deve essere superiore a 10. Qui,  $66 - 59 = 7$ : accettabile.

(2) *Discriminazione*. Il totale (3) non può essere superiore al doppio del totale (4) e viceversa. Qui, 15 e 17: accettabile.

(3) *Conformità*. Il totale dei margini delle differenze (5) non deve essere superiore a 12. Qui, = 7: accettabile<sup>12</sup>.

<sup>11</sup> Dimostriamo a p. 177 che con questo riferimento empirico si ottiene una buona valutazione della correlazione tra due serie di voti.

<sup>12</sup> Questo modo semplice di procedere evita il lungo calcolo della correlazione tra le due serie di voti.

Si vedrà del resto che si ottiene un coefficiente di correlazione molto vicino al coefficiente calcolato col classico metodo facendo l'operazione:

$$\frac{15^2 + 17^2 - 7^2}{2 \times 15 \times 17} = .90$$

oli per la comparazione tra un professore ed un moderatore (1 è il voto migliore; 5 il peggiore).

gruppi di 4	Alunni scelti a caso	Voti assegnati da				Differenze (moderatore-professore)	
		Moderatore	Professore				
gruppo	Enrico	Min. 5	5	Min.	0		
	Giovanni	4	4		0	Min.	
	Paolo	Max. 2	1	Max.	1		
	Pietro	3	2		1	Max.	
	Margine: Min.-Max.	3	14	12	4	Margine delle differenze = 1	
gruppo	Andrea		2	1	Max.	1	Max.
	Edoardo	Max. 1	1		0	Min.	
	Giulio	1	1	Min.	0		
	Renato	Min. 4	4		0		
	Margine: Min.-Max.	3	8	7	3	Margine delle differenze = 1	
gruppo	Antonio		3	3	Max.	0	Max.
	Camillo		5	5		0	
	Eugenio	Max. 2	3		-1*	Min.	
	Gerolamo	Min. 6	6	Min.	0		
	Margine: Min.-Max.	4	16	17	3	Margine delle differenze = 1	
gruppo	Giacomo	Min. 5	5	Min.	0		
	Lorenzo	Max. 1	1	Max.	0	Min.	
	Martino	4	3		1	Max.	
	Vittorio	4	3		1		
	Margine: Min.-Max.	4	14	12	4	Margine delle differenze = 1	
v gruppo	Bruno	Max. 3	2	Max.	1		
	Ugo	Min. 4	2		2	Max.	
	Leone	4	5	Min.	-1	Min.	
	Simone	3	2		1		
	Margine: Min.-Max.	1	14	11	3	Margine delle differenze = 3	
I cinque gruppi riuniti	Totale dei gradi		66 (1)	59 (2)		Totali dei margini di differenze = 7 (5)	
	Margini riuniti	15 (3)	(4)	17			

\* Si considera che (0) sia maggiore di (-1).

(4) *Conclusioni.* In questo caso nessun problema sembra porsi. I voti dei professori sono accettati e le operazioni sono terminate.

Se i voti del professore non sembrano accettabili, l'ideale è di intavolare una discussione con lui per trovare un accordo.

Questo non sempre è possibile, soprattutto per mancanza di tempo.

La soluzione seguente è più rapida. Se un moderatore accetta un problema, invita quattro dei suoi colleghi a ricorreggere indipendentemente lo stesso campione di venti compiti. Si calcola la media tra i cinque moderatori, ed il professore è invitato ad adattare tutti i suoi voti in funzione di quelli dei moderatori.

Secondo Peaker, i casi che necessitano di tale correzione diventano sempre più rari, man mano che il sistema di moderazione funziona, in ragione dei fruttuosi scambi di opinioni che ogni volta essi provocano.

(E) Come adattare voti discordanti?

Si tratta, sia di un semplice adattamento della media (o della mediana) sollevando o abbassando voti di tutti gli alunni (severità o generosità eccessive), sia anche di una modificazione della distribuzione generale, per migliorare la discriminazione<sup>13</sup>.

*Adattamento della mediana.* Per uno stesso esame, la situazione è la seguente:

	Quartile superiore	Mediana	Quartile inferiore	Scarto tipo approssimativo
Moderatore	16	13	11	$\sigma = 4$
Professore	11	8	6	$\sigma = 4$

<sup>13</sup> Adattato da D. Mather e altri, op. cit., p. 183.

Il professore è più severo del moderatore, ma discrimina altrettanto bene. Se il professore aggiunge 5 punti ad ogni alunno, il parallelismo è ristabilito.

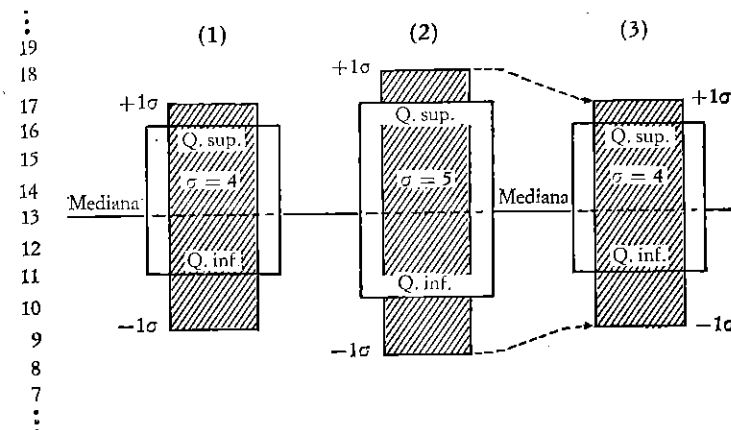
*Adattamento della mediana e dello scarto tipo.* Situazione di partenza:

	Quartile superiore	Mediana	Quartile inferiore	Scarto tipo approssimativo
Moderatore	16	13	11	$\sigma = 4$
Professore	16	12	9	$\sigma = 5$

Per allinearsi al moderatore, si corregge dapprima la mediana; viene aggiunto un punto a tutti gli alunni.

Supponiamo che i voti di partenza del professore siano:  
... 19 18 17 16 15 14 13 12 (mediana) 11 10 9 8 7 6 5 ...

Mostreremo innanzitutto per mezzo del disegno in che cosa consiste l'adattamento dello scarto tipo.



(1) Voti del moderatore.

(2) Voti del professore aumentati di un punto. La mediana è divenuta la stessa di quella del moderatore. Resta da adattare  $\sigma$ .

(3) Voti del professore diminuiti una seconda volta per mettere lo scarto tipo sulla stessa linea di quello del moderatore.

Per l'operazione (3), l'adattamento in funzione del nuovo scarto tipo è un po' più complicato che per la mediana. La preparazione di una tavola di conversione faciliterà le operazioni.

Per trovare il nuovo voto corrispondente a  $+1\sigma$ , basta aggiungere il nuovo sigma alla nuova mediana, sia  $13 + 4 = 17$ . Il nuovo voto corrispondente a  $-1\sigma = 13 - 4 = 9$ , ecc.

Per i voti corrispondenti a frazioni di sigma, si calcola dapprima lo scarto in rapporto alla mediana, si moltiplica per  $\frac{4}{5}$  e si aggiunge il risultato alla mediana. L'esempio seguente chiarirà questa fase.

	Voti di partenza del professore	Voti adattati in funzione della nuova mediana	Secondo adattamento in funzione del nuovo scarto tipo ( $\sigma$ )
	.	.	.
	19	.	.
	18	.	.
+ 1 $\sigma$	17	18	17*
	16	17	16**
	15	16	15***
	14	15	15****
	13	14	14
Mediana	12	13	13
	11	12	12
	10	11	11
	9	10	11
	8	9	10
- 1 $\sigma$	7	8	9
	6	.	.
	5	.	.
	.	.	.
	.	.	.

Come sono state trovate queste cifre?

$$\left. \begin{array}{l} * \text{Nuova mediana: } 13 \\ \text{Nuovo scarto tipo: } 4 \end{array} \right\} 13 + 4 = 17$$

\*\* Nella seconda colonna, 17 è a 4 punti dalla mediana.

$$\text{Moltiplicare } 4 \text{ per } \frac{4}{5} = 3,2.$$

$$\text{Mediana} + 3,2 = 16,2 \text{ arrotondata a } 16.$$

\*\*\*  $16 - 13 = 3$ ;  $3 \times \frac{4}{5} = 2,4$ ;  $13 + 2,4 = 15,4$  arrotondata a 15.

\*\*\*\*  $15 - 13 = 2$ ;  $2 \times \frac{4}{5} = 1,6$ ;  $13 + 1,6 = 14,6$  arrotondata a 15.

Il voto di fine d'anno:

lavoro dell'anno + lavori pratici + test

*Problema*<sup>14</sup>. Si vogliono classificare gli alunni in cinque gruppi di merito alla fine dei loro studi. *A* sarà il grado superiore e varrà 1, *E* sarà il grado inferiore e varrà 5.

Si vuol tener conto di tre elementi, valutati ciascuno secondo le stesse scale a cinque gradi:

- lavoro dell'anno = L.A.
- lavori pratici = L.P.
- test regionale = T.R.

Il test regionale si vede attribuire la stessa importanza che gli altri due elementi. Da cui la ponderazione:

$$\text{L.A.} = 1/4$$

$$\text{L.P.} = 1/4$$

$$\text{T.R.} = 1/2$$

Ecco la tabella dei gradi e la tabella dei risultati dopo la ponderazione e l'arrotondamento finale.

<sup>14</sup> Da D. Mather e altri, op. cit., pp. 149-154.

Numero alunni	Test regionale (T.R.)		Valutazione da parte della scuola		Ponderazione			Totale	Classifi- cazione finale
	Pun- teggio	Voto	L.A.	L.P.	T.R. $\times \frac{1}{2}$	L.A. $\times \frac{1}{4}$	L.P. $\times \frac{1}{4}$		
1	59	2	3	3	1	1	$\frac{3}{4}$	$2\frac{3}{4}$	3
2	77	1	1	1	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{4}$	1	1
3	51	3	4	3	$1\frac{1}{2}$	1	$\frac{3}{4}$	$3\frac{1}{4}$	3
4	12	6	5	5	3	$1\frac{1}{4}$	$1\frac{1}{4}$	$5\frac{1}{2}$	6
5	53	3	3	4	$1\frac{1}{2}$	1	1	$3\frac{1}{2}$	4
6	40	4	2	5	2	$\frac{3}{4}$	$1\frac{1}{4}$	4	4
7	66	1	2	2	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$1\frac{1}{2}$	1
8	60	2	3	4	1	$\frac{3}{4}$	1	$2\frac{3}{4}$	3
9	38	5	4	5	$2\frac{1}{2}$	$1\frac{1}{4}$	$1\frac{1}{4}$	5	5
10	70	1	2	2	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$1\frac{1}{2}$	1
11	56	2	2	3	1	$\frac{1}{2}$	$\frac{3}{4}$	$2\frac{1}{4}$	2
12	69	1	2	2	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$1\frac{1}{2}$	1
13	44	4	4	5	2	1	$1\frac{1}{4}$	$4\frac{1}{4}$	4
14	64	2	1	2	1	$\frac{1}{4}$	$\frac{1}{2}$	$1\frac{3}{4}$	2
15	19	6	5	5	3	$1\frac{1}{4}$	$1\frac{1}{4}$	$5\frac{1}{2}$	6
16	49	3	2	4	$1\frac{1}{2}$	$\frac{3}{4}$	1	$3\frac{1}{4}$	3
17	54	3	2	1	$1\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{4}$	$2\frac{1}{4}$	2
18	47	4	1	2	2	$\frac{1}{4}$	$\frac{1}{2}$	$2\frac{3}{4}$	3
19	52	3	3	1	$1\frac{1}{2}$	$\frac{3}{4}$	$\frac{1}{4}$	$2\frac{1}{2}$	2
20	48	3	3	3	$1\frac{1}{2}$	$\frac{3}{4}$	$\frac{3}{4}$	3	3
21	50	3	1	2	$1\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{2}$	$2\frac{1}{4}$	2
22	24	6	3	4	3	$\frac{3}{4}$	1	$4\frac{3}{4}$	5
23	61	2	4	2	1	1	$\frac{1}{2}$	$2\frac{1}{2}$	2
24	57	2	1	3	1	$\frac{1}{4}$	$\frac{3}{4}$	2	2
25	42	4	3	5	2	$\frac{3}{4}$	$1\frac{1}{4}$	4	4
26	35	5	5	2	$2\frac{1}{2}$	$1\frac{1}{4}$	$\frac{3}{4}$	$4\frac{1}{2}$	4
27	45	4	4	1	2	1	$\frac{1}{4}$	$3\frac{1}{4}$	3
28	41	4	2	4	2	$\frac{3}{4}$	1	$3\frac{3}{4}$	4
29	27	5	5	5	$2\frac{1}{2}$	$1\frac{1}{4}$	$1\frac{1}{4}$	5	5
30	43	4	3	2	2	1	$\frac{1}{2}$	$3\frac{1}{2}$	4
31	67	1	1	1	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{4}$	1	1
32	31	5	5	3	$2\frac{1}{2}$	$1\frac{1}{4}$	1	$4\frac{3}{4}$	5
33	72	1	2	2	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$1\frac{1}{2}$	1

I voti ponderati in corsivo saranno adattati in funzione del test regionale (T.R.).

La lettura della tabella ci mostra che l'alunno 1 ottiene i tre voti equilibrati: 1, 1 e  $\frac{3}{4}$ . Cioè in totale  $2\frac{3}{4}$ ; classificazione finale 3. L'alunno 3:  $1\frac{1}{2} + 1\frac{3}{4} = 3\frac{1}{4}$ ; classificazione

finale 3. [Le cifre in *corsivo* nella tabella indicano che il moderatore ha, in accordo con la scuola, adattato il voto del L.A. per armonizzarlo con il livello medio della regione].

Perché un ultimo adattamento prima di assegnare la classificazione finale? Perché la somma delle classificazioni provoca una nuova concentrazione verso la media; il restringimento è tale che, in una scala finale a cinque gradi, tutta la popolazione rischia di trovarsi al centro. Il seguente esempio fittizio illustra questo fenomeno.

Alunni	Classificazione		
	Test regionale	Lavoro dell'anno	Media
Pietro	1	5	3
Paolo	2	4	3
Giovanni	3	3	3
Ruggero	4	2	3
Enrico	5	1	3
Dispersione	4	4	0

Se ci si riporta alla tabella precedente, si osserva che, tra le classificazioni ponderate, si ritrovano molti 1 (17 in tutto), mentre se ne trovano soltanto due nella colonna del totale. I voti di eccellenza sono stati « annegati » dalla somma delle classificazioni.

Il processo assicura una miglior giustizia distributiva tra gli alunni di una stessa classe ed un miglior allineamento sul livello regionale rivelato dal test regionale, essendo questo riconosciuto come il campione più sicuro.

*Operazioni* (cfr. tabella alla pagina seguente). Segnare nella colonna (2) il numero degli alunni e fare il totale. Cominciare a segnare nella colonna (4) e fermarsi quando il totale uguaglia quello della colonna (2). Fermarsi a questo punto. Perché alcuni alunni non siano svantaggiati, si oltrepassa, se occorre, il numero della colonna (3), in modo da esaurire il livello a cui ci si è fermati. Si vedrà, per esempio, nella tabella successiva, che si assegna la classificazione 2 a 7 alunni e non a

Test regionale		Totale delle classificazioni ponderate		Classificazione finale	
Qualificazione (1)	Scolari che hanno ottenuto questa classificazione (2)	Qualificazione (3)	Scolari che hanno ottenuto questa classificazione (4)	Scolari ai quali è stata attribuita (5)	Classificazione (9)
1		1			1
		1¼			
		1½			
		1¾			
2		2			2
		2¼			
		2½			
		2¾			
3		3			3
		3¼			
		3½			
		3¾			
4		4			4
		4¼			
		4½			
		4¾			
5		5			5
Al di là		Al di là			Non classificato

6, perché 2 tra loro hanno ottenuto 2½. È chiaro che non si poteva assegnare una classificazione 2 all'uno e una classificazione 3 all'altro.

*Conclusione.* Nessuno dei sistemi descritti è perfetto, ma tutti permettono di migliorare il nostro sistema tradizionale di votazione. Una prima scelta verrà operata in funzione dello scopo perseguito: selezione o comparazione. Questo secondo aspetto ha soprattutto fermato la nostra attenzione.

Noi non ci pronunciamo in favore di un sistema particolare. La decisione appartiene alle autorità politiche e pedagogiche e agli insegnanti.

Test regionale		Totale delle classificazioni ponderate		Classificazione finale	
Classificazione (1)	Scolari che hanno ottenuto questa classificazione (2)	Classificazione (3)	Scolari che hanno ottenuto questa classificazione (4)	Scolari cui è stata attribuita (5)	Classificazione (6)
1	6	1	// 2	6	1
		1¼	// 1		
		1½	//// 4		
		1¾	/ 1		
2	6	2	/// 3	7	2
		2¼	// 2		
		2½	/// 3		
		2¾	/ 1		
3	7	3	/// 3	7	3
		3¼	// 2		
		3½	/ 1		
		3¾	// 2	7	4
4	7	4	/ 1		
		4¼	// 2		
		4½	/ 1		
		4¾	/ 1		
5	4	5	// 2	4	5
Al di là	3	Al di là		2	Non classificato

Un intenso sforzo di sperimentazione deve essere fatto a tutti i livelli: piccoli gruppi di professori, ricerche regionali e nazionali in collaborazione coi centri psico-medico-sociali ed i laboratori di pedagogia sperimentale delle università.

Man mano che questi lavori andranno avanti, si vedrà probabilmente emergere un nuovo sistema che tiene conto delle nostre tradizioni e anche degli obiettivi particolari del nostro paese. Tuttavia, quasi immediatamente, la necessità di efficienti centri di ricerca pedagogica regionale si farà nuovamente sentire. Ne ripareremo.



v  
il mito della curva di Gauss

## Il pericoloso mito della curva di Gauss

Nelle scienze umane, la curva a campana di Gauss ha una considerevole funzione, perché essa è l'immagine stessa della ripartizione di molte attitudini e qualità: gli individui medi abbondano, ma i geni e gli idioti, i giganti e i nani sono rari.

La curva di Gauss è sia il riflesso della legge del caso che presiede alla nostra nascita, sia la risultante dell'influenza di un gran numero di fattori che agiscono in maniera più o meno indipendente su di un individuo o un oggetto.

Poiché i test misurano spesso attitudini, tratti della personalità o performance di vaste popolazioni, è naturale che essi siano campionati secondo la ripartizione gaussiana: grosso modo, il 70% medi, il 13% buoni, il 13% mediocri, il 2% eccellenti, il 2% molto scadenti.

Nel corso della costruzione di tali test, si eliminano particolarmente i quesiti che venissero risolti positivamente da troppi o da troppo pochi soggetti. Lo scopo perseguito è quello di classificare ciascuno, di attribuirgli il posto che gli spetta in un gruppo numeroso. In breve, si tratta di organizzare una specie di concorso, in cui il più forte occuperà necessariamente il primo posto.

È per questo che molti test di attitudine o di conoscenza sono eccellenti *strumenti di selezione*.

Nella sua classe, l'insegnante persegue uno scopo totalmente diverso. Il suo ideale è che *tutti* gli alunni apprendano a

leggere, a calcolare e, in linea generale, a padroneggiare perfettamente tutte le conoscenze giudicate necessarie o utili dalla società. *Istruire non è selezionare.* Al contrario. È sforzarsi di far riuscire *tutti*. È quindi lottare contro la curva di Gauss presa come modello di selezione.

Le conseguenze pedagogiche di questa osservazione sono particolarmente importanti.

## I

## Evoluzione della curva delle conoscenze

Quando, nel nostro sistema di classi rigide, un maestro riceve, il giorno della riapertura delle scuole, un gruppo di alunni che non conosce, egli si trova normalmente davanti due curve: una rappresenta la distribuzione delle attitudini e l'altra quella del sapere.

## La curva delle attitudini

Nella sua abituale accezione — che provvisoriamente conserviamo — la parola *attitudine* indica caratteristiche, innate o acquisite, considerate come sintomatiche della capacità di un individuo ad acquistare un livello di competenza più o meno elevato, in un determinato campo.

Nell'insegnamento non strettamente specializzato (che per buona parte dei nostri alunni dura fino oltre i vent'anni), la stessa ampiezza della gamma delle conoscenze e delle capacità da fare acquisire (matematica, lingue, scienze naturali, arti, ecc.) rende impossibile la selezione molto rigorosa secondo una attitudine particolare.

Così, fino ad un livello molto avanzato della scolarità, le attitudini degli alunni restano distribuite a caso.

In queste condizioni, un professore di matematica, per esempio, che misurasse le speciali attitudini per la sua materia, otterrebbe normalmente una distribuzione gaussiana, o — vi-

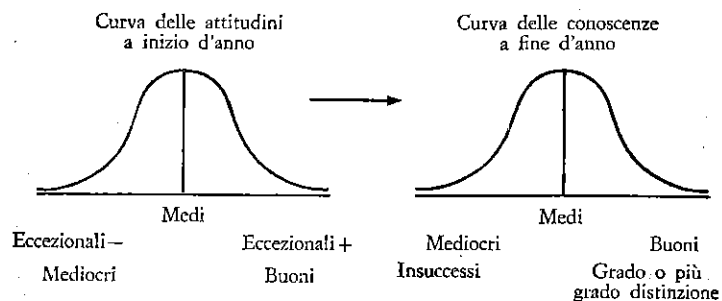
sto il numero probabilmente poco elevato di alunni — un abbozzo di questa.

I professori del resto non hanno quasi bisogno di ricorrere a test per conoscere questa situazione; essi sanno per esperienza che i medi sono la maggioranza e che gli eccellenti sono rari.

Tuttavia, generalmente si produce nelle menti una singolare distorsione. Si considera questa ripartizione delle attitudini come pronostico dei risultati scolastici di fine anno e si fissa il livello dell'insegnamento in modo tale che questo pronostico si verifica realmente: l'insegnamento cioè sarà « mediamente » difficile pur permettendo ai migliori di realizzare pienamente le loro possibilità e lasciando una minima « chance » ai mediocri.

Insistiamo su questo, la difficoltà « media » è determinata dalla media delle attitudini del gruppo considerato e non da una media delle difficoltà *obiettive* delle nozioni da insegnare<sup>1</sup>.

In queste condizioni, se il professore fa *lo stesso corso* a tutta la classe, è normale che la *curva delle conoscenze* acquisite alla fine dell'anno rispetti, a sua volta, la distribuzione gaussiana.



Ma adempie così l'insegnamento a quella che dovrebbe essere la sua vera funzione?

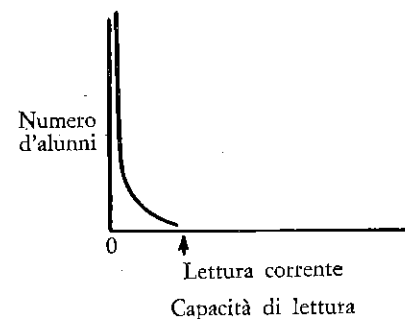
<sup>1</sup> Così si spiegano le considerevoli differenze del livello medio secondo le classi e le regioni, di cui abbiamo parlato nella prima parte di questo libro.

## La curva delle conoscenze

Torniamo al primo giorno dell'anno scolastico e, anziché considerare le specifiche attitudini degli alunni, esaminiamo le loro conoscenze.

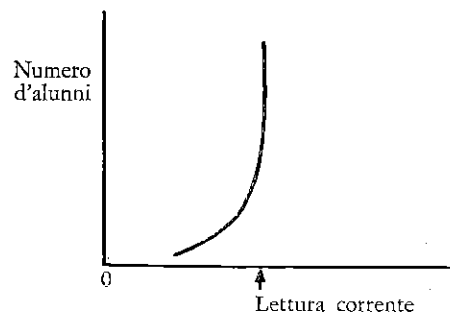
Sul piano logico, la funzione del professore è quella di suscitare l'apprendimento di conoscenze *nuove*. Certo è un'utopia immaginare che tutti gli individui che formano una classe possiedano esattamente la stessa quantità e qualità di conoscenze. Tuttavia il sistema di classi fisse che noi pratichiamo si basa sull'ipotesi che *tutti* si trovino approssimativamente allo stesso livello. Altrimenti, come oseremmo ancora dispensare lo stesso insegnamento a tutti?

E in realtà? Prendiamo il caso dell'entrata in prima elementare. La maggioranza dei bambini non sa leggere; alcuni sono sulla buona strada; due o tre leggono già correntemente. In questo momento, la curva della capacità in lettura assume, grosso modo, la forma della lettera *i*.

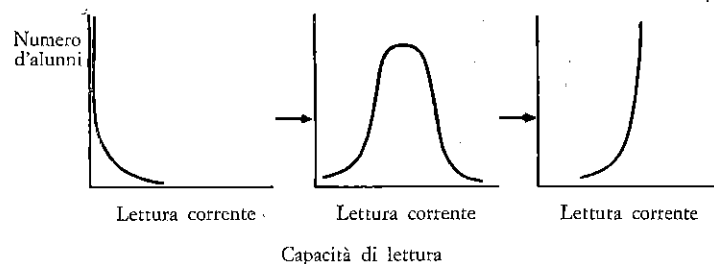


Ora, per quanto la curva delle attitudini specifiche alla lettura sia molto probabilmente gaussiana, il maestro non ammette di primo acchito che soltanto una parte della classe potrà imparare a leggere; al contrario. Per quanto non si sia portati al perfezionismo, si può dire che dopo un anno o due,

nella grande maggioranza gli alunni sapranno leggere correntemente. La curva delle conoscenze avrà completamente cambiato forma: assomiglierà ad una  $j$ .



Tra la curva a  $i$  e la curva a  $j$ , è probabile che ci sia stato un momento in cui i meglio dotati hanno avanzato più in fretta, dove i meno dotati sono rimasti indietro e i medi si sono situati tra i due. L'evoluzione è dunque schematicamente la seguente:



Nel caso della lettura non si concepisce che i genitori o i responsabili dell'insegnamento possano accontentarsi di una evoluzione diversa.

Ma, a partire da quale livello della scolarità, tale esigenza non è più sentita?

## II

## Una pedagogia di curva a J

Un esame in cui i risultati si distribuiscono secondo la curva a campana di Gauss si presta bene alla selezione, al concorso.

Lo spirito dei concorsi ha impregnato il nostro insegnamento per secoli per tre motivi storici principali.

- (1) Per motivi socio-economici: solamente una parte della popolazione scolarizzabile aveva accesso alla scuola. Cent'anni or sono, le famiglie modeste facevano la prima selezione non lasciando finire la scuola elementare che ai più dotati dei loro figli, cioè a quelli che risultavano migliori agli esami. Le borse di studio, distribuite con parsimonia, si guadagnavano attraverso concorsi.
- (2) Fino a questi ultimi tempi — diciamo prima dell'era del calcolatore —, tutti i paesi industrializzati disponevano di un enorme surplus di materia grigia. Si coltivava quindi quella che si offriva a miglior prezzo e con maggior facilità, cioè ci si preoccupava soprattutto di identificare i meglio dotati. Anche per i bambini agiati, l'insegnamento secondario aveva una funzione selettiva.
- (3) Le conoscenze psicologiche e pedagogiche erano rudimentali. I maestri dunque non erano in grado di applicare trattamenti speciali agli alunni che provavano difficoltà di apprendimento. Del resto, ancor oggi, spesso si fa ripetere un anno all'alunno che non è stato capace del minimo di performance richieste per il passaggio. In altre parole, anziché applicare rimedi particolari, ci

si contenta di porre nuovamente l'alunno nelle condizioni (stesso professore, stesso metodo) nelle quali si è prodotto l'insuccesso.

Essendo ogni anno scolastico come un filtro per il successivo, una certa quantità di insuccessi sembrerebbero dunque normali. In breve, la ripartizione gaussiana sembra soddisfacente.

Ma la costante preoccupazione della promozione dei più idonei ha causato progressivamente una deformazione pedagogica più sottile. Si è arrivati a pensare che le conoscenze « che contano veramente », le conoscenze « approfondite », il « vero gioco delle idee astratte », non sono assimilabili che da una minoranza che ha attitudini specifiche ad un grado elevato. E sulla base di tale concezione sono stati creati i metodi d'insegnamento.

Gli alunni e i loro genitori accettano del resto questo stato di cose senza troppe discussioni. Ci si iscrive raramente nella sezione « latino - matematica » senza sentirsi particolarmente dotati in questa direzione.

Certamente, non si può ambire a fare indifferentemente di ognuno un virtuoso della matematica, del pianoforte o della letteratura. Ma dove si pone la delimitazione? A partire da quale grado l'accesso ad un settore del sapere è impossibile a coloro che possiedono attitudini medie o perfino mediocri?

Per natura, ciascuno di noi è mediocrementemente dotato in molti campi. Non mancano, per esempio, intellettuali più o meno refrattari alla matematica. Tuttavia si osserva spesso che a forza di volontà, di perseveranza, di sentir ripetere le spiegazioni, di cambiare manuale o maestro per trovare una forma di insegnamento che convenga, nozioni di matematica dapprima considerate inaccessibili sono veramente apprese con sicurezza.

Per quanto vi consacrino il tempo necessario, gli alunni medi (cioè, vista la selezione già operata dagli studi anteriori, probabilmente più dell'80% della popolazione scolastica di un

dato livello) possono andare molto più lontano di quanto si immagini.

In questa prospettiva la nuova definizione dell'attitudine proposta da Carrol acquista tutto il suo valore: « L'attitudine è la quantità di tempo richiesta da chi apprende per padroneggiare una materia »<sup>1</sup>.

L'implicazione di questa concezione è considerevole: se si concede loro il tempo necessario tutti gli alunni che normalmente si trovano in una classe potrebbero giungere ad un buono e perfino ottimo risultato.

<sup>1</sup> J. Carrol, *A Model of School Learning*, in « Teachers College Record » 1963-64, pp. 723-733.

## La teoria della valutazione formativa

L'espressione *valutazione formativa* — in opposizione a *valutazione complessiva* — è stata creata da Michael Scriven<sup>1</sup>.

La *valutazione complessiva* (Roller parla in questo caso di « norma psicotecnica ») ci è familiare. Per interpretare il punteggio ottenuto in un test classico di conoscenze o d'intelligenza, lo si pone in una distribuzione statistica: la performance di un individuo è giudicata in rapporto a quella di altri. Allo stesso modo, si classificano spesso ancora gli alunni tra loro secondo l'ordine crescente o decrescente dei loro risultati scolastici, ed è secondo il posto così occupato che molti genitori apprezzano il lavoro dei loro bambini.

Nei due casi, il risultato è sostanzialmente relativo. Sia che il gruppo di riferimento vari nella sua composizione sia che il risultato appaia sotto un'altra luce.

Una semplice differenza di età può anche cambiare considerevolmente l'aspetto delle cose. In molte norme di test di conoscenza d'uso corrente, un anno d'età o un anno scolastico in più o in meno sono sufficienti perché una stessa performance sia considerata come mediocre o buona.

<sup>1</sup> M. Scriven, *The Methodology of Evaluation*, in R. Staker (a cura di), *Perspectives of Curriculum Evaluation*, Chicago, Rand McNally, 1967. Vedi anche B. S. Bloom, *Learning for Mastery*, in « *Evaluation Comment* » 2 1968. La nostra esposizione del metodo della valutazione formativa si basa direttamente su questa pubblicazione. Parecchi risultati di esperienze in cui la teoria del *Mastery Learning* è stata applicata sono pubblicati da J. Block, B. Bloom, Hastings (1970); trad. it. *Mastery Learning*, Torino, Loescher, 1972.

La prima preoccupazione di chi ha costruito questi test non è stata dunque quella che la conoscenza sia acquisita o no, ma piuttosto quella della velocità alla quale questa acquisizione si è realizzata. Riportandoci alla definizione di Carrol, essi hanno valutato dunque l'attitudine anziché valutare il contenuto dell'apprendimento.

Coloro che parteggiano per la *valutazione formativa* rovesciano questa concezione.

Poiché l'alunno viene a scuola per apprendere, l'importante non è di determinare a che punto si trovi nell'acquisizione del sapere?

Immaginiamo che una rigorosa analisi scientifica ci riveli che, con tutte le sue sfumature e le sue complicazioni, l'accordo del participio passato usato con l'ausiliare avere, presenti cento casi diversi. Per ciascuno di essi si possono definire criteri di padronanza.

Può dunque essere fissato il numero dei casi da padroneggiare, secondo il livello scolastico. In questo contesto, la natura della valutazione scolastica cambia.

L'alunno è classificato in funzione di un criterio obiettivo: la strada percorsa nell'acquisizione.

Inoltre, se si desidera condurre tutti gli alunni fino ad un livello minimo di conoscenze, se non fino alla totale padronanza della nozione o della tecnica, è rimesso in discussione lo stesso modo di insegnare: non è più possibile applicare indifferentemente a tutti un solo e identico metodo durante uno stesso periodo.

In questa prospettiva cosa diventerà l'insegnamento?

Si pongono tre problemi cruciali:

- (1) come segnare i momenti del processo dell'acquisizione del sapere?
- (2) come guidare l'alunno?
- (3) dove si pone il limite pratico di questa pedagogia della curva a *j*?

Che io sappia, non esistono ancora risposte complete a queste domande. Tale imprecisione del sapere sembra normale

in scienze naturali o in medicina. Dobbiamo imparare ad accettarla anche nelle scienze dell'educazione. La nostra discussione dunque molte volte termina su ricerche da intraprendere o da continuare.

### Tracciare l'itinerario conoscitivo

In viaggio, per determinare a quale distanza ci si trova dalla meta, devono essere soddisfatte due condizioni: da una parte, sapere dove ci si trova e dove si va, e dall'altra, disporre di una carta che indichi chiaramente la strada.

Allo stesso modo, nel campo dell'educazione, dobbiamo definire gli obiettivi da raggiungere e determinare con precisione gli apprendimenti particolari che vi condurranno. Il problema varia a seconda che si abbia a che fare o meno con apprendimenti di base. Per questi ultimi bisogna scoprire il concatenamento « critico » delle materie, cioè quello per cui una non è accessibile se non è stata assimilata la precedente. Per questo motivo, l'apprendimento delle conoscenze e delle tecniche di base deve, in ultima analisi, essere lineare, mentre le acquisizioni e le applicazioni che vanno al di là vedono aprirsi davanti a loro vie sempre più numerose.

Per esempio, qualunque sia il metodo d'insegnamento, non è possibile applicare completamente una regola del tre, senza avere — in particolare — la nozione della moltiplicazione e della divisione. Ad un livello superiore, come fare del calcolo integrale senza sapere che cosa è una funzione? Ma tali proposizioni sono ancora troppo vaghe. Quali sono esattamente gli apprendimenti necessari e sufficienti per poter assimilare la regola del tre? E, tra essi, quali sono critici in rapporto agli altri?

Il problema non è facile. E sarebbe ingenuo credere che la riflessione e i tentativi che hanno presieduto all'elaborazione dei nostri migliori manuali scolastici l'abbiano risolta. Certo il buon concatenamento dei capitoli costituisce un avvio alla soluzione, ma basta provare a programmare la più piccola nozione, non fosse che un sistema skinneriano, per scoprire

tutte le ignoranze, tutta l'imprecisione dei metodi d'insegnamento attuali.

Tuttavia, è precisamente negli sforzi di programmazione intrapresi da parte di ricercatori e d'insegnanti sempre più numerosi che risiede una delle grandi sorgenti di progresso.

Sicuramente, le scienze deduttive sono il terreno d'elezione del metodo che si va abbozzando. Anche se ci limitiamo ad esse non avremo compiuto opera trascurabile. Ma la linea di demarcazione non appare così presto. Là dove l'induzione domina nell'elaborazione del sapere, la deduzione riappare per la sua applicazione. Del resto, anche l'apprendimento induttivo può essere guidato sistematicamente.

### Guidare l'alunno

Una volta chiaramente definito il cammino che conduce ad un apprendimento, bisogna impegnarvi l'alunno. Per le acquisizioni fondamentali, abbiamo, inoltre, deciso che *tutti* dovrebbero, in linea di principio, arrivare allo scopo.

In questo caso la frequente valutazione dei progressi è essenziale. Da cui la necessità di *esercizi di padronanza* e di *test diagnostici* che si basano su parti di materia molto limitate e sono utilizzati dai maestri stessi.

Questi strumenti — che solo la collaborazione degli insegnanti e dei ricercatori permetterà di costruire in numero sufficiente — indicano il *livello a cui l'alunno è giunto* (ciò che determina il suo voto) e *dove trova delle difficoltà*.

I *rimedi* sono molteplici e devono entrare nella nostra quotidiana pratica pedagogica.

- (a) Indicare in maniera precisa la parte del corso da ristudiare.
- (b) Il lavoro da parte dei sottogruppi: se un gruppo di più di tre alunni trova la o le stesse difficoltà, il professore trova vantaggioso dedicare una parte del tempo della lezione a questo gruppo, mentre gli altri lavorano in modo indipendente.



- (c) Se il gruppo conta al massimo tre alunni, Bloom suggerisce che ciascuno di essi sia integrato in un gruppo eterogeneo di due o tre. Siccome gli alunni non sono più in concorrenza, ma lottano per impadronirsi di una materia, entra in gioco l'aiuto reciproco.
- (d) Nelle scuole che contano parecchi professori di una stessa materia, un alunno dovrebbe avere la facoltà di domandare una spiegazione, persino una breve lezione particolare ad un professore diverso dal suo.
- (e) Cambiare la presentazione: mettere a disposizione degli alunni dei manuali che spiegano la nozione in maniere diverse; dare possibilità di ricorrere all'insegnamento programmato, ai mezzi audiovisivi; variare il livello d'astrazione.
- (f) In generale, non fissare a tutti lo stesso tempo per le stesse cose.

È una procedura che funziona. Bloom riferisce un'esperienza condotta su questa linea in cui più dell'80% degli alunni hanno ottenuto il voto migliore a fine d'anno. È il trionfo della pedagogia della curva ad *j*.

Può essere sempre così? Niente permette di affermarlo.

### Limite pratico del sistema

Anche se è vero che tutti gli alunni delle nostre scuole sono capaci di raggiungere un alto grado di padronanza nei limiti in cui gli si lasci un tempo sufficiente, è evidente che oltre un certo limite la durata diventa proibitiva.

Come permetterete di trascorrere vent'anni anziché sei nella scuola secondaria? Questa considerazione, tuttavia, non deve rimettere in discussione il sistema. L'essenziale è orientare gli alunni in funzione delle loro reali attitudini. Il ciclo d'osservazione e d'orientamento, introdotto progressivamente nell'insegnamento della maggior parte dei paesi, lo rende possibile.

Ma quando dopo questo smistamento lo studente è ammesso in una classe, normalmente egli deve poter apprendere l'essenziale di *tutto* quello che vi si insegna.

### Il sistema d'insegnamento messo in discussione

La pedagogia della curva a *j* richiede un livellamento verso l'alto; sarebbe tuttavia chimerico immaginare che essa permetta un livellamento a livello di... genialità.

Abbiamo visto che meno del cinque per cento della popolazione possiede attitudini *eccezionali*. Esse devono essere sviluppate al massimo, sia per rispetto verso la persona che nell'interesse della comunità.

Un nuovo sistema d'insegnamento permette contemporaneamente di istruire ognuno ad un ritmo conveniente e di lasciare che si realizzino in pieno i superdotati: è la scuola senza classi, in cui si lavora con gruppi resi omogenei secondo le attitudini specifiche *per una sola materia*<sup>2</sup>.

Nelle condizioni attuali, la classe è una costrizione. Perché un alunno di sei anni che sa leggere quando entra in prima elementare non potrebbe partecipare agli esercizi di lettura del secondo anno? Perché un tale alunno del quarto anno dell'insegnamento secondario non potrebbe seguire un corso di matematica del livello del terzo o del secondo?

Nelle nostre vecchie scuole elementari di campagna, dove tutte e sei le classi erano a volte affidate ad un solo maestro, era facile vedere tali spostamenti. In alcune scuole secondarie che attualmente contano parecchie classi con alunni della stessa età è possibile lavorare a quattro o cinque livelli d'attitu-

<sup>2</sup> Sistema da non confondere con lo *streaming*, che consiste nel costituire classi omogenee con l'aiuto di test di attitudini generali o, peggio ancora, in funzione dei risultati scolastici globali. Si commette così un doppio errore scientifico: nella maggior parte dei casi uno stesso soggetto è dotato in maniera diversa per le varie materie del programma; inoltre omogeneizzare globalmente fa abbassare il rendimento dell'insieme. In questo sistema, del resto, si giunge spesso ad una segregazione sociale. Tutti questi inconvenienti sono stati messi in evidenza così bene che lo *streaming* in Svezia è proibito dalla legge.

dine differenti almeno in alcune materie principali. Tale sistema funziona perfettamente da numerosi anni in scuole come la New Trier High School, che noi abbiamo descritto dettagliatamente fin dal 1961<sup>a</sup>.

Preveniamo subito l'obiezione finanziaria. Una scuola a gruppi omogenei specifici non costa più di un'altra, al contrario. Se aggiungiamo il guadagno di tempo, di ingegno e il rarefarsi degli insuccessi che il sistema permette, si può affermare senza paura di sbagliare che il beneficio è considerevole.

<sup>a</sup> *L'individualisation de l'enseignement dans une école multilatérale*, in « Education », maggio 1961.

## Conclusioni e raccomandazioni

Man mano che andavamo avanti in questo studio già lungo, si sono imposte le conclusioni parziali. Tuttavia siamo lontani dall'aver esaurito l'argomento e ancor più, dall'aver dato soluzioni definitive.

Uscire dalle routines meccaniche e ambire a tradurre le grandi teorie pedagogiche nella pratica quotidiana dell'educazione, rivela tanti scogli, tante condizioni da superare che si rischia di essere assaliti dal dubbio e dallo scoraggiamento!

A rischio di lasciare i nostri lettori insoddisfatti, non abbiamo voluto mascherare le difficoltà e meno ancora lasciar credere all'esistenza di una docimologia compiuta, capace di condurre ad una valutazione perfetta se gliene vengono forniti i mezzi. È verosimile, ed è senza dubbio una fortuna, che la misura rigorosa dei tratti umani più sottili resterà sempre impossibile: la sorte ci salvaguarda dalla macchina per pesare le anime!

Il fatto che la valutazione continua si inserisca funzionalmente nel processo d'insegnamento e d'apprendimento da cui è indissociabile, non sopprime, tuttavia, la necessità di una docimastica, cioè di una tecnica d'esami e di votazioni, in altre parole di una tecnica di osservazione sistematica del rendimento, immediato o posticipato.

Del resto, l'esigenza di valutazione comparata, che l'educatore può rimpiangere per motivi ideali, va oltre l'ambito scolastico: essa corrisponde ad un carattere fondamentale della nostra civiltà.

Può darsi che un giorno le cose andranno diversamente. Tuttavia non vediamo, se la nostra cultura continua ad intellettualizzarsi, e quindi a crescere in complessità, come potrebbe rinunciare completamente agli esami e ai concorsi.

La nostra posizione docimologica è decisamente eclettica nel suo desiderio di conciliare i sicuri vantaggi di nuove tecniche, di natura soprattutto statistica, e l'apporto innegabile e fondamentale dell'analisi qualitativa.

Per limitarci ad un solo esempio di questo eclettismo, noi non possiamo concepire che le decisioni riguardanti il passaggio alla fine di ciascun anno di studi universitari siano prese solamente in funzione dei risultati ottenuti da un test a scelta multipla che si presta alla votazione automatica. Senza dubbio, l'introduzione di questo tipo di prove è altamente auspicabile per il controllo obiettivo ed approfondito delle conoscenze, ma non basta. Deve intervenire la valutazione del lavoro dell'anno e l'incontro conclusivo del maestro e del suo allievo deve fornire tutte le sfumature che l'approccio quantitativo ha lasciato nell'ombra.

Come l'introduzione delle macchine nell'industria ha permesso all'intelligenza di prendere il sopravvento sulla forza muscolare e sulla routine avvilente, così il controllo automatico della conoscenza dei fatti, dei metodi e delle tecniche, rende possibile un esame finale realmente basato sui processi mentali e i tratti più nobili della personalità. Quale rispettiva importanza si accorderà ai tre tipi di valutazione? In questo caso, la decisione spetta alla comunità accademica, illuminata dalle ricerche docimologiche.

Sembrano imporsi un certo numero di *raccomandazioni*.

Proprio a causa della nostra ignoranza e della grande complessità della valutazione, innanzitutto è urgente che una rete di centri di ricerche pedagogiche si estenda nell'intero paese e sia messa al servizio del sistema scolastico. Le funzioni di questi centri sono apparse nel corso della nostra analisi: ricerche che si basano sui problemi fondamentali e sullo sviluppo dei metodi, delle tecniche e degli strumenti necessari. Questi centri non devono essere innestati come corpi estranei nel

complesso delle nostre istituzioni educative, bensì lavorare in stretta collaborazione con esse.

Più particolarmente, ogni università dovrebbe creare al suo interno un ufficio di studi sui problemi dell'insegnamento e degli esami. Non si può concepire che l'istituzione scientifica per eccellenza escluda dal rigore della sua analisi una delle sue due ragion d'essere: l'educazione.

La definizione degli obiettivi generali e specifici, fondamento di tutto l'edificio dell'insegnamento e della valutazione, richiede il lavoro in comune di un gran numero di ricercatori e di pratici di tutte le discipline. Commissioni specializzate permetteranno la cooperazione necessaria e la comunicazione — nei due sensi — tra la base e il vertice della gerarchia pedagogica.

La riforma degli esami e dei modi di valutazione determinerà senza dubbio la sparizione definitiva del calcolo della percentuale, diffuso quasi universalmente nelle nostre scuole fino a questi ultimi tempi. Quali che fossero le sue debolezze, la scala delle percentuali offriva tuttavia il vantaggio di essere familiare a tutti, e ciascuno aveva almeno l'illusione di capirla. Attualmente, nelle nostre scuole sono provate le scale più diverse. *Una unificazione del sistema di votazione e una campagna d'informazione* sono necessarie al più presto, altrimenti diventerà difficile, se non impossibile, la comunicazione tra scuola e famiglia, tra le diverse scuole, tra le scuole e le autorità. Tuttavia, prima di prendere una decisione la cui importanza è evidente, dovranno essere ascoltati gli esperti in docimologia e, più generalmente, in misura nelle scienze umane.

Sarà anche intrapreso un ampio sforzo di informazione presso insegnanti chiamati a cambiare profondamente le loro abitudini di valutazione. Il mezzo più efficace sembra risiedere nell'*organizzazione di esperienze locali* che riguardino la preparazione degli esami, la votazione e la moderazione. Queste esperienze sarebbero, per esempio, inserite nei lavori necessari all'introduzione dell'insegnamento secondario rinnovato, in Belgio, e nella riforma della formazione dei maestri; progressivamente, vi parteciperanno tutte le scuole.

Saranno costituite a poco a poco *commissioni di moderazione locali e regionali*.

Un'azione troppo affrettata rischia di mettere in pericolo un rinnovamento estremamente lodevole. Ci si guarderà, in particolare, dal deludere o turbare gli insegnanti tentando di porli sotto l'autorità di docimologi improvvisati.

Opera umana per eccellenza, l'educazione non guadagnerà in qualità che nella misura in cui gli educatori accetteranno profondamente le nuove idee e tenteranno di superare se stessi.

Appendici

Un quesito d'esame secondo il metodo tradizionale e secondo il metodo con quesiti a scelta multipla

#### Formulazione del quesito<sup>1</sup> (metodo tradizionale)

(1) (a) Elencare le cause d'ipoglicemia; (b) descrivere le lesioni anatomico-patologiche risultanti dall'ipoglicemia.

La griglia seguente è stata costruita insieme da tutti i correttori, per permettere ad essi un certo grado di uniformità nella valutazione:

(A) Uguale valore per le due parti del quesito.

(B) Per ottenere un voto di 75/100, il candidato deve aver indicato nella lista delle cause: (1) eccesso d'insulina; (2) tumore funzionale delle isole di Langerhans; (3) influenza endocrina con ipofunzionamento dell'ipofisi e delle surrenali; (4) malattia del fegato che impedisce di accumulare una riserva di glicogeno (necrosi acuta) o la liberazione del glicogeno (malattia di von Gierke).

(2) Da che cosa si riconosce la differenza delle lesioni causate da una sola crisi acuta d'ipoglicemia da quelle causate da una ipoglicemia cronica?

(3) Descrivere le modificazioni visibili al livello del cervello.

(C) Per ottenere un voto tra 75 e 90/100, il candidato deve aver indicato: o una lista che dimostri che ha compreso

<sup>1</sup> Robert A. Moore, *Methods of Examining Students in Medicine*, in « Journal of Medical Education », gennaio 1954.

il meccanismo d'azione di ognuna delle cause, o una lista che comprenda altre cause, ma che comunque indichi che capisce che i meccanismi di omeostasi hanno abitualmente tendenza a correggere l'ipoglicemia, nei casi come: (1) diminuzione del cibo ingerito; (2) assorbimento perturbato; (3) metabolismo aumentato come nell'ipertiroidismo; (4) infezione grave; (5) surmenage fisico; (6) glicosuria renale.

- (D) Per ottenere un voto superiore a 90/100, il candidato deve aver indicato: una lista logica delle quattro cause maggiori e delle sei cause minori, oltre ad un tipo idiopatico; una differenziazione dell'ipoglicemia acuta e cronica con una descrizione dettagliata delle lesioni cerebrali.

#### Formulazione del quesito (metodo con quesiti a scelta multipla)

Ecco come uno stesso quesito potrebbe essere presentato secondo il nuovo metodo.

(1) Si può determinare se il candidato è capace di riconoscere le quattro cause più importanti e differenziarle dalle meno importanti nel quesito seguente.

*Istruzioni.* Quattro delle cinque frasi numerate sono comuni a una delle tre turbe funzionali della lista alfabetica (A), (B), (C). Indicate quale è l'eccezione e la turba funzionale comune alle quattro frasi che rimangono.

(1) Eccesso d'insulina; (2) tumore funzionale delle isole di Langerhans; (3) glicosuria renale; (4) ipofunzionamento ipofisario; (5) malattia di von Gierke.

(A) Ipoglicemia clinicamente accertabile; (B) iperglicemia clinicamente accertabile; (C) glicosuria clinicamente accertabile.

Se il candidato risponde correttamente, dimostra che sa che (1), (2), (4), e (5) possono produrre una ipoglicemia clinicamente accertabile, che non è il caso (3), e che nessuna combinazione dei quattro casi tra i cinque può essere associata

né all'iperglicemia, né alla glicosuria. In altri termini, è provato il possesso di conoscenze positive e negative.

(2) Se si vuole sapere ciò che sa il candidato sui fenomeni che dominano l'equilibrio del livello dello zucchero sanguigno e dell'insulina, si può porre il quesito seguente.

*Istruzioni.* Ciascuna delle frasi seguenti è composta da due parti: l'esposizione di un fatto e la causa di questo fatto. Indicate per ciascuna delle frasi numerate la lettera A, B, C, D o E, a seconda che (A) l'esposizione del fatto e il suo motivo sono veri e hanno una relazione di causa ed effetto; (B) l'esposizione del fatto e il suo motivo sono veri, ma non hanno relazione di causa ed effetto; (C) l'esposizione del fatto è vera, ma la relazione è falsa; (D) l'esposizione del fatto è falsa, ma il motivo è un fatto o un principio accettato; (E) l'esposizione del fatto e il motivo sono falsi:

- (A) (1) il tasso di zucchero nel sangue cade di colpo dopo epatectomia perché il glicogeno contenuto nel fegato è la sorgente principale dello zucchero sanguigno;
- (E) (2) la diagnosi anatomo-patologica di un adenoma delle isole di Langerhans del pancreas implica che la malattia era iperinsulinica perché tutti gli adenomi delle isole sono funzionali e secernono insulina;
- (C) (3) i malati colpiti d'ipertiroidismo hanno molte possibilità di avere una ipoglicemia perché esiste un iperinsulinismo associato;
- (C) (4) i malati affetti dalla malattia di von Gierke presentano una diminuzione del tasso di zucchero nel sangue perché, in questa malattia, il glicogeno non viene accumulato dal fegato.
- (E) (5) una ipoglicemia durante e dopo parecchi mesi non è seguita da postumi perché le modificazioni cellulari prodotte dall'ipoglicemia sono reversibili.

Se il candidato risponde correttamente a questa serie, dimostra di sapere: (a) che il glicogeno epatico è la fonte principale che permette il mantenimento del tasso dello zucchero nel sangue; (b) che non tutti i tumori delle isole di Langerhans

sono funzionali; (c) che i malati colpiti da ipertiroidismo hanno una ipoglicemia, ma che la causa non è un iperinsulinismo associato; (d) che i malati colpiti dalla malattia di von Gierke hanno una ipoglicemia, ma che essa non è dovuta al fatto che non vi è glicogeno nel fegato; (e) che vi sono dei postumi in seguito a ipoglicemia cronica e che le modificazioni cellulari risultanti non sono reversibili.

(3) Si può determinare se il candidato ha qualche nozione semplice riguardante le modificazioni al livello del cervello col seguente quesito.

*Istruzioni.* Ogni formulazione incompleta (numerata) è seguita da cinque complementi a scelta. Indicate in ogni caso il complemento più conveniente.

- (1) Le modificazioni anatomiche secondarie a una ipoglicemia cronica sono messe in evidenza più spesso in (a) la milza, (b) il rene, (c) le surrenali, (d) il cervello, (e) la tiroide.
- (2) Tra le modificazioni provocate dall'ipoglicemia cronica al livello del cervello, le più importanti sono su (a) i neuroni, (b) gli astrociti, (c) le cellule dell'ependima, (d) le cellule oligodendroglie, (e) le cellule microglie.
- (3) Tra le alterazioni delle cellule nervose provocate dall'ipoglicemia acuta, la più evidente è (a) lo spostamento del nucleo, (b) la sparizione della parete cellulare, (c) l'alterazione della sostanza di Nissl, (d) lo scoppiare del nucleo della cellula, (e) la fusione di mitocondri.
- (4) Tra le modificazioni provocate al livello del cervello dall'ipoglicemia cronica, la più significativa è (a) l'idrocefalia interna, (b) l'ispessimento fibroso dell'aracnoide, (c) la distruzione dei neuroni, (d) la proliferazione degli astrociti, (e) la proliferazione delle cellule dell'ependima.

Se il candidato risponde correttamente a questa serie, dimostra di sapere che le alterazioni principali dell'ipoglicemia cronica colpiscono i neuroni del cervello, che l'ipoglicemia acuta provoca una alterazione della sostanza di Nissl dei neuroni e che l'ipoglicemia cronica provoca la distruzione delle cellule nervose.

(4) Si può determinare se il candidato comprende il meccanismo delle turbe, con il quesito seguente.

*Istruzioni.* Nella lista alfabetica sono indicati cinque meccanismi diversi che possono condurre all'ipoglicemia. Scrivere la lettera appropriata dopo ciascuna delle formulazioni numerate associando al disturbo il meccanismo responsabile: (A) aumento dell'attività metabolica; (B) iperinsulinismo; (C) accumulo di un glicogeno anormale nel fegato; (D) assenza di accumulo di glicogeno nel fegato; (E) ipofunzionamento dell'ipofisi o delle surrenali.

- (B) (1) Adenomi delle isole di Langerhans.  
 (A) (2) Esercizio fisico violento.  
 (A) (3) Iperitiroidismo.  
 (E) (4) Morbo di Simmonds.  
 (C) (5) Morbo di von Gierke.  
 (D) (6) Epatite epidemica.  
 (C) (7) Morbo di Addison.

Se il candidato risponde correttamente a questa serie, questo dimostra che egli comprende le basi del metabolismo dei glucidi e conosce i fattori che influenzano questo metabolismo.

#### Riepilogo

Così, per mezzo di questi diciassette « quesiti obiettivi », abbiamo messo in evidenza tutte le conoscenze richieste per poter dare un voto superiore ai 90/100, e cioè: (1) lista delle quattro cause principali; (2) differenza tra gli effetti dell'ipoglicemia cronica ed acuta; (3) descrizione delle alterazioni al livello del cervello; (4) comprensione dei meccanismi; (5) lista delle cause minori; (6) differenza fra le lesioni dell'ipoglicemia acuta e cronica.

## II

### Composizione in lingua materna

Inghilterra: General Certificate of Education 1967  
Scuola secondaria livello « O »<sup>1</sup>

#### Prova I: due quesiti, un'ora e mezza

- (1) Scegliete uno dei seguenti argomenti. Dedicategli un'ora circa.
  - (a) Una notte nebbiosa.
  - (b) Un negoziante compra e rivende, lo stesso giorno, un oggetto d'occasione. Descrivete le due scene.
  - (c) Quale sarebbe la vostra politica se dirigeste i programmi radiofonici o televisivi?
  - (d) Una grande folla si disperde. Descrivete la scena.
  - (e) Piacere della fotografia o del disegno o della danza o del ciclismo.
  - (f) Pensate che ragazzi e ragazze abbiano le stesse possibilità di carriera?
  - (g) Quali sono le vostre reazioni davanti al progresso e alle realizzazioni dell'esplorazione dello spazio?
- (2) Scegliete uno dei seguenti argomenti. Dedicategli mezz'ora circa.
  - (a) Descrivete un allenamento destinato a migliorare le vostre performance in uno sport di vostra scelta.

<sup>1</sup> Si tratta cioè di una prova a livello « ordinario », che si può sostenere anche prima del biennio finale (sixth form) dell'istruzione secondaria, durante il quale si preparano quasi esclusivamente le due o tre materie in cui si sosterranno le prove di livello A (avanzato) del General Certificate of Education, necessario per l'ammissione all'Università [NdT].

- (b) Dopo aver visitato un'azienda, fate un rapporto ai vostri compagni sulle condizioni di lavoro e le prospettive di carriera che essa offre.
- (c) Descrivete chiaramente *uno* degli apparecchi seguenti e spiegate come funziona: un walkie-talkie, un asciugacapelli, un aspiratore, un frullatore.
- (d) Nella vostra città si vuol creare un centro commerciale nel quale sarà proibita la circolazione ai veicoli. Scrivete una lettera ad un giornale locale per esporre il vostro punto di vista sul progetto.

#### Prova II: quattro quesiti, un'ora e 3/4

- (1) Riassumete il brano seguente in buona prosa corrente e con un massimo di 110 parole. Alla fine, indicate quante parole avete utilizzato. Il brano è costituito da 314 parole [segue un testo sulla eccessiva fiducia nella scienza e nella tecnologia].
- (2) Leggete il seguente passaggio; poi, rispondete ai quesiti [il testo descrive due grandi categorie di gitanti — coloro che seguono una guida e coloro che partono all'avventura — e sottolinea l'interesse delle escursioni geologiche]:
  - (a) Spiegate con parole vostre la differenza tra le due categorie di gitanti.
  - (b) Quale consiglio dà l'autore a proposito delle grotte? Formulatelo con parole vostre.
  - (c) Spiegate brevemente perché l'autore crede che la geologia sia: un hobby divertente - un hobby istruttivo.
  - (d) L'autore scrive che utilizza l'espressione « scavare un fossato » in senso letterale. Spiegare perché il senso è qui letterale.
  - (e) Spiegate le espressioni seguenti: precauzioni prescritte; essere perfettamente coscienti della natura dei propri atti.



- (f) Scegliete quattro delle parole seguenti. Sostituitele con dei sinonimi o delle perifrasi che potrebbero essere utilizzate nel testo senza cambiarne il senso [...].
- (3) Rispondete a scelta ad uno dei due quesiti seguenti:
- (a) Scegliete tre parole fra le seguenti. Costruite delle frasi (6 in tutto) dimostrando che queste parole possono essere usate con due diversi significati [...].
- (b) Definite in una frase tre delle parole seguenti: monopolio, interludio, prefazione, microscopio, antidoto.
- (4) Rispondete a scelta ad uno dei due quesiti seguenti:
- (a) Spiegate chiaramente, ma brevemente, la differenza di senso in ciascun paio di frasi (si basa su *could-should; can-may; will-shall; might-must*).
- (b) Riscrivete correttamente il seguente passaggio, rispettando tutti i concetti. Potete cambiare l'espressione, l'ordine delle parole e dei concetti, l'ortografia e la punteggiatura [segue un testo difettoso, di un centinaio di parole]: « Per esempio se si progettasse di fare passare una strada attraverso una città ma una casa storica fosse lungo il suo tracciato allora il piano dovrebbe essere cambiato, comportando delle considerevoli spese, per aggirare la costruzione creando una curva per la strada e rendendola anche dannosa per le auto. A causa della deviazione di questa strada non solamente vi saranno delle spese ma il prezzo delle diverse cose aumenta se la deviazione è considerevole, se la distanza aggiunta è, diciamo, dieci chilometri e un camion che porta taluni articoli percorre la strada, il camion consumerà in tal caso più benzina impiegherà più tempo per arrivare e avrà una corrispondente usura aggiuntiva ».

forme per la correzione

Prova I: massimo 50 punti

Lo schema della votazione che figura qui sotto non costituisce che una guida preliminare. Potranno essere apportate aggiunte e correzioni al momento della riunione degli esaminatori che sarà convocata dopo una prima lettura dei lavori.

Quesito 1 (massimo 35 punti)

Ci si aspetta un minimo di 400 parole, ma le composizioni non devono essere classificate principalmente in funzione della lunghezza. Tener conto del soggetto scelto e del modo con cui è trattato. Anche se è breve, una composizione la cui argomentazione è serrata ed è espressa in buona forma deve ottenere più punti che una narrazione lunga ma espressa in una forma scadente.

L'esaminatore dovrebbe avere una idea chiara di ciò che è una composizione che riceve appena appena il voto di passaggio (16 punti). Tale composizione deve contenere idee ragionevoli, ma non molto originali. L'espressione deve essere chiara, ma senza particolare distinzione. Nel lavoro non si dovrebbe trovare che qualche errore meccanico. I candidati che hanno superato questo livello generale devono essere ricompensati e quelli che hanno lavorato al di sotto di questo livello devono essere penalizzati.

Si troveranno nel documento allegato note dettagliate sulle qualità da osservare nelle composizioni. Si ritiene tuttavia che gli esaminatori classificheranno in funzione della loro impressione generale: non assegneranno quindi una proporzione fissa dei punti per i diversi aspetti. Se il soggetto si presta alla controversia, le idee e la loro connessione possono essere più importanti che in una composizione descrittiva in cui potrebbe avere un posto maggiore il vocabolario. Noi facciamo affidamento sul discernimento degli esaminatori.

Riservando 35 punti a questo primo quesito, si è voluta

sottolineare la sua maggiore importanza nell'insieme degli esami. Molti candidati sono mediocri. Tuttavia, se l'esaminatore non disperde largamente i suoi voti, questo primo quesito non avrà un peso determinante nell'insieme dei risultati.

Gruppo A (29-35 punti): la composizione è di qualità eccezionale.

Gruppo B (22-28 punti): la qualità del lavoro è al di sopra della media.

Gruppo C (14-21 punti): il lavoro è di qualità media.

Gruppo D (7-13 punti): il lavoro è ad un livello inferiore a quanto consideriamo soddisfacente.

Gruppo E (0-6 punti): il candidato è incapace di presentare le sue idee con coerenza.

Questa prova è un esame di lingua materna. Nella prova 1, esaminiamo la capacità del candidato ad esprimere le sue opinioni, le sue esperienze, le sue impressioni, i suoi sentimenti e i suoi interessi. Non si tratta né di un test di conoscenze generali, né di una valutazione delle attitudini del candidato. Se una ragazza descrive chiaramente una scena in inglese, deve ottenere un voto favorevole se i materiali utilizzati sono pertinenti, anche se l'esaminatore pensa che la ragazza in questione affronta il soggetto in modo troppo passionale. Per altro, se dietro una massa di prosa incoerente l'esaminatore ha l'impressione che si trovino emozioni profonde e scelte morali elevate, non deve occuparsi di quest'ultimo aspetto: conta solo ciò che è scritto. Per riuscire nella prova, sono essenziali la chiarezza dell'espressione e la precisione dello stile.

Questito 1 (massimo 35 punti)

- (a) Una narrazione o una descrizione sono accettabili. In una narrazione, la notte nebbiosa deve avere una funzione essenziale.
- (b) Per sommi capi, le due scene devono essere equilibrate. Ricompensare la vivacità di narrazione, di dialogo, di descrizione e il contrasto.

- (c) Si attende una definizione chiara della politica; di essa parecchi punti devono essere sviluppati. Il candidato deve avere scelto la radio o la televisione; non può averle scelte entrambe.
- (d) Il soggetto del lavoro è la dispersione; autorizzare tuttavia una breve introduzione.
- (e) Una risposta coerente, chiara, di media lunghezza, dice più che lunghe divagazioni e ripetizioni.
- (f) Questo soggetto non è facile; ricompensare generosamente la disposizione degli argomenti e gli esempi ben scelti.
- (g) Uno svolgimento puramente narrativo non può essere accettato, ma alcuni esempi concreti dei progressi realizzati possono essere necessari per spiegare le reazioni.

Quesito 2 (massimo 15 punti)

- (a) Si attende una descrizione chiara e logica.
- (b) Il rapporto deve trattare i tre aspetti del quesito. Si insiste soprattutto sull'informazione riferita; quindi non si esigerà una forma di rapporto particolare.
- (c) Si esige nello stesso tempo una descrizione dell'oggetto e una spiegazione del suo funzionamento. Non punite severamente gli errori materiali; votate semplicemente la chiarezza dell'espressione.
- (d) Si possono considerare numerosi aspetti. Quindi, penalizzate qualunque idea inadeguata.

Detraete:

- 1 punto per una brutta redazione dell'indirizzo;
- 1 punto per una mancanza di coerenza tra il modo in cui ci si rivolge al destinatario all'inizio e ai saluti;
- 1 punto se l'alunno ha firmato « Sig. John Smith » o « Signorina Jeannette Smith »;
- 1/2 punto per altri errori di disposizione o di punteggiatura, o di ortografia in parole essenziali.

Tutti questi soggetti forniscono un materiale sufficiente per una mezz'ora di lavoro. Ci si aspetta un minimo di 200 parole. La scelta è ricca.

### Prova II: massimo 50 punti

#### Quesito 1 (massimo 16 punti)

I punti attribuiti a questo quesito sono abitualmente molto più bassi che per gli altri. Gli esaminatori sono pregati di non considerare 11 come il massimo attribuibile al riassunto.

(A) Attribuire un massimo di 2 punti per ognuno degli aspetti seguenti. Per ottenere 2 punti, il candidato deve avere compreso chiaramente l'idea ed averla espressa correttamente. Sfumate i vostri voti con 1 e 1/2, 1 o 1/2 punto. Alla fine del lavoro, calcolate un'altra volta l'insieme in funzione della fluidità e della coerenza del riassunto completo. Se un passaggio è incoerente, sopprimete al massimo 1/4 dei punti attribuiti; se esso è piuttosto aspro sottraete 1/8. Questa sottrazione dovrebbe apparire sulla composizione scrivendo per esempio  $10 - 1 = 9$ . Ogni sottrazione di punti destinata a penalizzare un testo più lungo del limite fissato o una proposizione non solamente inesatta ma assurda, deve apparire separatamente. Il totale finale deve essere nel margine circondato da un cerchio.

- (1) L'uomo della strada accetta oggi le scoperte scientifiche,
- (2) senza dubitare della loro origine, della loro validità, o dei loro effetti.
- (3 e 4) La domanda di nuovi progressi destinati ad elevare il livello di vita non rallenta mai.
- (5) Gli uomini hanno fiducia nell'uomo di scienza e nel suo lavoro,
- (6) e credono che non si possa arrestare il progresso.
- (7) Pur ammettendo che gli uomini di scienza non sono sempre d'accordo sulla sicurezza apportata dalle nuove scoperte,

(8) il pubblico è convinto che gli uomini di scienza finiranno per trovare l'unanimità o per lo meno un largo accordo.

(B) Il limite di 120 parole dà un margine generoso. Deducete un punto per ogni gruppo di 5 parole che superano il massimo. Non sono da considerare nel conteggio parole introduttive quali « In questo passaggio l'autore spiega che... ». Ricontate le parole. Non accettate semplicemente la cifra scritta dal candidato.

#### Quesito 2 (massimo 20 punti)

Classificando questo quesito, accordate il massimo in ogni sezione al candidato che espone chiaramente il punto.

- (a) (1) Coloro cui piace che si preveda tutto al loro posto (1)
- (2) e coloro che preferiscono disporre semplicemente di un piano generale che permetta di seguire l'ispirazione del momento (spirito d'esplorazione). (2)
- (b) Visitare grotte. (1)  
A meno che non siate guidati da qualcuno che conosca molto bene il terreno. (1)  
Rispettate tutte le regole di sicurezza. (2)
- (c) (1) Non occorre quasi un equipaggiamento speciale. (2)  
(2) Il geologo fa continuamente nuove scoperte (1 e 1/2)  
su piccola scala. (1/2)
- (d) Quando il geologo rompe una pietra, in realtà compie un atto unico, perché nessuno dopo di lui potrà rompere nuovamente questa stessa pietra. (2)
- (e) (1) Misure di sicurezza (1)  
imposte o raccomandate. (1)

- (2) Per comprendere esattamente le conseguenze di ciò che si è fatto. (1)
- (f) Sotterraneo: (1)
- che è sotto il suolo, sotto la superficie della terra; (1)
- sotto terra. (1)
- Speleologi: (1)
- persone che studiano le grotte scientificamente; esploratori di grotte. (1)
- Coscienziosamente: (1)
- di buona lena, senza lasciarsi distrarre, con molta applicazione (1)
- onestamente (1/2)
- fedelmente (1/2)
- Eccetera.

Per i sottoquesiti (a), (b) ed (e), non attribuite nessun punto agli alunni che si limitano a ricopiare una parte del testo.

### Quesito 3 (massimo 6 punti)

- (a) Assegnate un punto per ciascuna frase costruita correttamente. Non assegnate nessun punto in caso di costruzione scorretta.
- (b) Concedete un punto per ogni definizione esatta espressa in una frase corretta (massimo 3 punti). Concedete un punto per ogni frase in cui il senso della parola che comincia con lo stesso prefisso risalti chiaramente (massimo 3 punti). 1/2 punto per una definizione espressa in una frase scorretta. 1/2 punto per una definizione fornita in frase incompleta.

Nessun punto se la parola che comincia con lo stesso prefisso non è presentata in una frase.

Un *monopolio* è una proprietà esclusiva detenuta da una ditta (o)

Un *monopolio* è il nome dato ad una ditta che detiene dei diritti commerciali esclusivi.

Un *interludio* è un intervallo disposto nel corso della rappresentazione di un lavoro teatrale (o un intervallo nello svolgimento di un avvenimento).

Eccetera.

### Quesito 2 (massimo 8 punti)

- (a) Assegnate un punto per frase. Verificare la presenza delle seguenti idee... Non assegnate nessun punto se non si vede chiaramente a quale frase si riferisce il candidato.
- (b) I candidati risponderanno in maniere diverse. Detraete 1 punto a causa di ogni errore di ortografia, 2 punti per ogni costruzione zoppicante, 1 punto per una espressione scorretta, 1/2 punto per l'omissione o l'uso sbagliato di una virgola essenziale, 1 punto per ogni idea omessa. Rivedete il voto totale in funzione dell'impressione generale del brano. Per i quesiti 1 e 2, se il candidato ottiene la metà dei punti, arrotondate all'unità superiore (esempio:  $6\frac{1}{2} = 7$ ). Se i due quesiti danno un voto che comprende 1/2 punto, arrotondate uno per eccesso e l'altro per difetto. Per i quesiti 3 e 4 le stesse osservazioni.

## Esempio d'insegnamento semi-individualizzato<sup>1</sup>

### La New Trier Township High School di Winnetka

La New Trier Township High School è una grande scuola di livello secondario superiore in cui l'insegnamento è semi-individualizzato e dove allo studente è data l'opportunità di correggere il suo orientamento fino al termine dell'adolescenza.

Essa accoglie gli alunni da 14 anni<sup>2</sup> in poi e gode di una grande reputazione negli Stati Uniti e all'estero.

3740 studenti seguono regolarmente i corsi<sup>3</sup>. Il corpo insegnante conta 255 membri di cui: 40 professori d'inglese, 31 di matematica, 30 di educazione fisica, 29 di lingue straniere, 27 di *social studies* (storia, geografia, educazione civica, sociologia ed economia), 17 di scienze, 9 di musica, 7 di corsi tecnici, 7 di commercio, 6 d'arte drammatica, 6 di pit-

<sup>1</sup> Estratto d'un articolo che abbiamo pubblicato in « Education » nel maggio 1961.

<sup>2</sup> Si sa che l'organizzazione dell'insegnamento degli Stati Uniti non è uniforme. La New Trier High School deriva dal sistema NK 8-4: un anno di *nursery school*, un anno di giardino d'infanzia, otto anni di primaria e quattro di secondaria. Gli altri due sistemi più frequenti sono: NK 6-3-3 e NK 6-6. Si trova anche NK 7-5, NK 6-2-4 ed NK 6-4-4.

<sup>3</sup> Queste cifre si riferiscono all'anno scolastico 1959-60. Sono tratte da « Information for College Admission Officers », New Trier Township High School, novembre 1959, o sono stati raccolte sul posto.

tura e disegno, 6 d'automobile, 4 di economia domestica, 2 d'igiene, 1 d'insegnamento speciale (ritardati mentali educabili).

Si conta un insegnante ogni 15-16 studenti e le classi riuniscono generalmente circa 25 alunni.

Le prestazioni di un professore che non ricopre speciali responsabilità (quali presidente di un dipartimento, per es.) comportano 24 periodi di 40 minuti a settimana: 4 corsi di 5 periodi e 4 periodi di *counseling*.

Il livello intellettuale degli studenti, giudicati sulla base dei testi classici, è elevato. Circa l'80% di coloro che terminano i loro studi a New Trier ottengono risultati superiori alla media nazionale nei test d'attitudine e di conoscenze (SCAT, STEP, *National Merit*). Il 92% dei diplomati continuano studi superiori.

### L'individualizzazione dei programmi

Il principio fondamentale dell'azione pedagogica della New Trier High School è definito nella prima fase del suo programma di corso: « Nessuno può credere alla dignità degli uomini senza provare un profondo piacere allo spettacolo della loro infinita varietà »<sup>4</sup>.

Restando quindi così impregnati dello spirito del Piano di Winnetka, gli studi sono organizzati in maniera da fornire a ciascuno la possibilità di svilupparsi col ritmo che gli è proprio secondo le sue capacità e le sue tendenze. Questo non significa affatto che siano permesse la sterile fantasia e le soluzioni facilistiche.

Secondo il sistema diffuso in tutto il paese, lo studente può ottenere un diploma di compimento di studi solo se nel

<sup>4</sup> New Trier Township High School, *Curriculum Guide*, dicembre 1959, p. 1.

corso di questi egli ha ottenuto un numero totale di punti o di *credits* fissato dal consiglio d'amministrazione della scuola che tiene conto a sua volta di norme generali. Questi *credits* esprimono in una unità convenzionale l'importanza qualitativa e quantitativa dei diversi corsi, per un semestre (5 *credits* corrispondono, per esempio, ad un semestre di corso « minore », in ragione di cinque periodi a settimana). I *credits* sono acquisiti solo se lo studente raggiunge un voto superiore ad un limite minimo fissato.

Per ottenere il certificato di studi della New Trier High School, è necessario avervi guadagnato 350 *credits* in 4 anni.

Il gioco dei *credits* è contemporaneamente origine di elasticità e di sicurezza, poiché permette di delimitare esattamente il campo di libertà dell'alunno.

In linea di principio, questi stabilisce il suo programma di studi come si compone il menù di un pasto, scegliendo sulla carta ciò che attrae e conviene di più.

Ecco, per esempio, la lista dei corsi proposti per il primo anno, dove sono citati i *credits* che essi fruttano se sono seguiti con successo per un semestre<sup>5</sup>:

Corsi « maggiori » (che rendono 10 *credits* per semestre; salvo indicazione contraria, essi devono essere seguiti durante 2 semestri consecutivi): inglese, algebra, alimentazione (un semestre), educazione civica, commercio, disegno industriale, elettricità (un semestre), storia universale, lingue straniere, matematica generale, pittura e disegno (due periodi al giorno), fotografia, radio amatore, lavorazione del legno, lavori manuali, abbigliamento (un semestre).

Corsi « minori » (il numero di *credits* è tra parentesi): arte drammatica (6), canto corale (3), dattilografia (5), corso d'armonia (6), disegno industriale (un periodo al giorno) (5), di-

<sup>5</sup> Cfr. New Trier Township High School, *Registration Bulletin for Freshmen 1959-60. Planning a Course of Study*, p. 2.

zione (6), armonia (5), orchestra sinfonica (5), lavori manuali (un periodo al giorno) (5).

Tuttavia, parecchie restrizioni influenzano la scelta dell'alunno. In primo luogo, un certo numero di corsi, giudicati indispensabili alla cultura di base di tutti i membri della nazione, sono obbligatori: lingua materna per 4 anni (*credits* 80), matematica per 2 anni (40), *social studies* per 2 anni<sup>6</sup> (40), scienze per un anno (20), educazione fisica per 4 anni (16), automobile per un semestre (3). Totale: 199 *credits*.

Circa 2/3 del totale dei *credits* richiesti sono dunque forniti da corsi obbligatori. Tuttavia, come vedremo, si può soddisfare a queste esigenze molto liberamente; così, abbiamo rilevato una decina di possibilità diverse che permettono di portare a termine validamente l'anno di scienze richiesto.

Un secondo fattore importante guida lo studente nell'elaborazione del suo programma: la professione o gli studi superiori ai quali aspira. In particolare, ogni università determina le sue condizioni di ammissione e specifica particolarmente quanti *credits* lo studente deve avere acquistato nelle scuole secondarie, per determinate materie.

Infine i genitori, i maestri, i consiglieri pedagogici e gli psicologi si sforzano di guidare lo studente utilizzando nel modo migliore i suoi interessi, badando che egli tragga il miglior profitto possibile dalle sue capacità.

La settimana scolastica che conta 5 giorni di 8 periodi di 40 minuti — escluse le riunioni quotidiane con il consigliere pedagogico (20') e la durata della colazione fatta sempre a scuola (25') — ha suggerito generalmente all'alunno di costruirsi un programma non eccessivamente pesante, facendo buon uso del tempo per lo studio personale.

<sup>6</sup> Di cui un anno dedicato obbligatoriamente alla storia degli Stati Uniti.

Ecco quattro esempi tipo per il primo anno <sup>7</sup>:

I		II		III		IV	
Inglese	5	Inglese	5	Inglese	5	Inglese	5
Algebra	5	Algebra	5	Algebra	5	Algebra	5
Latino	5	Scienze	5	Storia	5	Storia	5
Pittura e disegno	10	Commercio	5	Alimentazione	7	Scienze	5
Ginnastica (ragazze)	4	Ginnastica (ragazzi)	5	Canto	3	Lingua straniera	5
		Musica strumentale	5	Ginnastica (ragazze)	4	Ginnastica (ragazzi)	5
	29		30		29		30
Studio personale	11	Studio personale	10	Studio personale	11	Studio personale	10
	40		40		40		40

Si rileverà come questi piani di lavoro contrastino con la dispersione dello sforzo così frequente da noi. Praticamente tutti i corsi accademici possono essere concentrati nella mattinata.

Circa un quarto del tempo passato a scuola è riservato agli studi personali. Le ore così lasciate libere saranno spesso dedicate a lavoro di ricerca nella biblioteca scolastica che, negli Stati Uniti, ha una funzione incomparabilmente più importante che in Belgio.

### L'individualizzazione dell'insegnamento

Non solo l'alunno sceglie le materie che gli convengono maggiormente, ma inoltre l'insegnamento di ognuna di queste sarà adattato alle sue possibilità.

<sup>7</sup>New Trier Township High School, *Courses for Freshmen*, aprile 1959, p. 4.

In realtà, tutte le materie importanti possono essere studiate a cinque diversi livelli di attitudini: inferiore, medio debole, normale, accelerato, avanzato. In questo modo, lo sforzo richiesto allo studente, che egli possieda un'intelligenza superiore o sia poco dotato, che sia ferrato in un campo e in ritardo in un altro, è sempre proporzionato alle sue possibilità.

Siccome l'alunno segue raramente tutti i suoi corsi ad uno stesso livello, il danno di una segregazione generale, secondo le attitudini, sembra minimo: del resto la scuola sta molto attenta a questo problema, raggruppando sistematicamente tutti gli studenti in occasione di alcune attività. Si fa di tutto per inculcare una vera tolleranza nei confronti del maggiore o minore talento dei compagni di studio e per incoraggiare ciascuno a superarsi: « Nello sport, non tutti sono capaci di giocare in modo eccellente; lo stesso caso si verifica negli studi. Ma si attende da ciascuno il meglio di se stesso » <sup>8</sup>.

Il sistema d'insegnamento a diversi livelli è praticato a New Trier, dalla fine della prima guerra mondiale, con la più grande soddisfazione di tutti. In media gli studenti si dividono secondo le percentuali seguenti: superiori 15-20%, normali (medio superiori) 40-55%, medi deboli 36-40%, limitati 5-8%, avanzati  $\pm 10\%$ .

Quando gli alunni entrano alla New Trier High School, i gruppi sono provvisoriamente determinati sulla base dei risultati scolastici precedenti e di altre informazioni riunite secondo un sistema che studieremo più avanti. In seguito i risultati ottenuti nella scuola stessa correggeranno e guideranno le assegnazioni, non essendo queste mai definitive, qualunque sia la materia.

Tale flessibilità conduce certamente ad una organizzazione complessa, ma non bisogna tuttavia esagerare la difficoltà.

Nella tabella seguente, facciamo apparire tutte le possibilità offerte durante il secondo semestre dell'anno scolastico

<sup>8</sup>New Trier Township High School, *Guide Book to New Trier*, Winnetka 1959, p. 39.

1959-60. Alcune note marginali precisano lo spirito degli studi<sup>9</sup>.

Ogni corso è indicato da un numero di tre cifre:

- (a) la cifra delle centinaia indica in quale anno si dà: si avrà dunque 1, 2, 3 e 4;
- (b) la cifra delle decine indica il semestre: 1 = 1° semestre; 2 = 2° semestre; 0 = può essere seguito al primo o al secondo semestre; 3 = corso di vacanze estive;
- (c) la cifra delle unità indica il livello d'attitudine: 1 = livello inferiore; 2 = medio debole; 3 = normale; 4 = insegnamento accelerato; 5 = avanzato; 6 = seminari; 9 = tutti i livelli riuniti.

Materie e livelli di corso	Osservazioni
LINGUA MATERNA	
125	Per studenti superlativamente intelligenti, eccellenti in lingua materna (integra la lingua materna, la biologia e la storia).
124	Corso arricchito, fondato sulle opere letterarie di livello universale.
123	Corso medio.
122	Corso medio inferiore; si rivolge particolarmente agli alunni che non hanno ancora acquistato un metodo di lavoro razionale.
122 R	R = <i>remedial</i> . Destinato agli alunni identificati dall'ufficio del <i>testing</i> come aventi una notevole deficienza in lettura (comprensione e rapidità) e in ortografia.
121	Accoglie gli alunni che presentano una debolezza generale in tutti i settori della lingua materna.

<sup>9</sup> Sono stati utilizzati i due documenti di base seguenti: New Trier Township High School, *Curriculum Guide 1959* e *Program of Classes 1959-60*.

Materie e livelli di corso	Osservazioni
224, 223, 222, 221 324, 323, 322, 321 424, 424 (+), 423, 422, 421	Le differenze di livello nel 2°, 3° e 4° anno sono, <i>mutatis mutandis</i> , parallele a quelle che indichiamo qui sopra per il primo. 323 = giornalismo. 424 (+) = <i>Great Books</i> .

## MATEMATICA

Algebra A (accelerato)	124	La differenza tra i livelli 2, 3 e 4 si trova meno nell'accelerazione che nel metodo d'insegnamento e nell'approfondimento. <i>Al livello 2</i> le spiegazioni sono dettagliate; molte applicazioni; nessuna incursione nei campi vicini. <i>Al livello 3</i> lo studente deve soprattutto lavorare da sé: teoria più rigorosa, arricchimento. <i>Al livello 4</i> i concetti sono trattati rapidamente; studio più approfondito; numerose incursioni nei campi vicini. Lo stesso manuale è usato per questi tre livelli e tutti gli alunni studiano lo stesso capitolo nello stesso tempo, ciò che permette, ad ogni momento, il passaggio da un livello all'altro.
Algebra —	124	
Algebra E (sperimentale)	124	
Algebra —	123	
Algebra E	123	
Algebra —	122	
Algebra E	122	
Matematica di base	121	
Matematica approfondita (20 credits)	225	
Geometria	224	
Geometria	223	
Geometria	222	
Matematica	222	
Matematica di base	221	
Matematica approfondita (20 credits)	315	Quattro semestri di matematica sono obbligatori; tuttavia, due di questi possono essere dedicati alla contabilità o all'aritmetica commerciale.
Matematica	324	
Algebra	323	
Algebra	322	
Geometria	303	
Matematica approfondita	425	
Matematica	424	
Algebra (livello universitario)	403	
Trigonometria	403	
Uso del regolo calcolatore (2 credits; va con l'algebra)	323 322	

## « SOCIAL STUDIES »

Educazione civica	123, 122, 121	Quattro semestri di <i>social studies</i> sono obbligatori; due di questi devono essere dedicati alla storia degli Stati Uniti.
Storia universale	125, 124, 123	
Storia: tempi moderni	224	
Storia: antichità	204	



Materie e livelli di corso	Osservazioni
Storia: medioevo 202, 203, 204	
Storia: universale 221	
Geografia 222, 223	
Storia: tempi moderni 323	
Storia: Gran Bretagna 303	
Storia: USA 324, 323, 322, 321	
Storia: USA (approfondita) 425	
Storia: contemporanea (Europa) 425	
Storia: Gran Bretagna (approfondita) 404	
Storia: America latina 403	
Storia: Estremo Oriente 403	
Storia: universale del xx secolo 402	
Storia: Stati Uniti del xx secolo 402	
Educazione civica 403	
Sociologia 403	
Scienze economiche 403	
<b>SCIENZE</b>	
Biologia 123, 124, 125	Gli alunni che desiderano continuare gli studi nelle facoltà di scienze delle università sono invitati a seguire sei semestri di scienze e otto di matematica.
Radio amatore 123	
Biologia 224, 223, 222	
Chimica 225	
Biologia 323, 322, 321	
Elettronica 323	
Fisica 325	
Chimica 324, 323	
Chimica (seminario) 426	
Fisica 425, 424, 423	
Scienze 409	
<b>LINGUE</b>	
Latino 124, 123, 122	Nessun corso di lingue è obbligatorio. Gli studenti sono tuttavia incoraggiati a seguirli se si orientano verso studi superiori. Si pensa che, molto presto, le università esigeranno all'ammissione che lo studente abbia studiato almeno una lingua straniera durante tre anni.
224, 223, 222	
324, 323	
424, 423	
124, 123	
Tedesco 224, 223, 223C	
324, 323	
426	

Materie e livelli di corso	Osservazioni
Russo 124	
224	
324	
Francese 125, 124, 123	
224C, 224, 223C, 223	
324C, 324, 323C, 323, 324	
424C, 423C	C indica che si tratta della continuazione di un corso precedente. = corso di conversazione.
Spagnolo 125, 124, 123, 122	
224, 224C, 223, 223C, 222	
324, 324 = 323, 322	
424, 423	
<b>ARTI</b>	
Disegno e pittura 129	Tutti questi corsi possono essere seguiti in ragione di due periodi al giorno (= corso « maggiore »: 10 <i>credits</i> ) o di un periodo al giorno (= corso « minore »: 5 <i>credits</i> ). Tuttavia, storia dell'arte 329 è sempre un « maggiore ».
Lavoro manuale 121	
Disegno e pittura 229C, 229	
Lavoro manuale 221	
Disegno e pittura 329	
Ceramica 329	
Gioielleria 329	
Storia dell'arte 329	
Ceramica 429	
Gioielleria 429	
Pittura e disegno 429	
<b>COMMERCIO</b>	
Commercio generale 122, 123	Obiettivi perseguiti con questo corso:
Commercio generale 222	(1) preparazione agli studi superiori di scienze commerciali;
Contabilità 223	(2) formazione generale;
Il punto di vista del consumatore 323	(3) preparazione per gli studenti che vogliono lavorare a mezzo tempo nel commercio durante i loro studi universitari;
Vendita 302	(4) preparazione professionale per gli alunni che non faranno studi superiori;
Pubblicità 303	(5) possibilità per gli studenti che cercano la loro strada, di ve-
Organizzazione commerciale 303	
Stenografia 323	
Dattilografia 303	
Pratica di ufficio 322	
Segretariato 423	
Stenografia 423	

Materie e livelli di corso	Osservazioni
Diritto commerciale Pratica commerciale	423, 403 422 dere se la carriera commerciale potrà interessarli. Il 422 = una riunione di 30' ogni mattina scuola; il resto consiste in stages pratici (20 h settimanali).
AUTOMOBILE	
Teoria e pratica (guida)	209 Una legge, approvata nel 1955, sta diventando obbligatoria questo corso in tutte le <i>High Schools</i> dell'Illinois (3 <i>credits</i> ).
ECONOMIA DOMESTICA	
Alimentazione Abbigliamento	109 109 Destinato principalmente alle ragazze; tuttavia i ragazzi che si interessano al restauro o all'industria alberghiera possono seguire i corsi d'alimentazione; e i ragazzi che si interessano all'architettura possono seguire i corsi di decorazione di interni.
Alimentazione Abbigliamento	209 209
Decorazione di interni	329
CORSI TECNICI	
Disegno industriale Legno Elettricità	123, 122 123, 121 103 Lo scopo perseguito non è di dare una formazione professionale, ma di fare dello studente « un consumatore intelligente dei prodotti dell'industria ».
Disegno (architettura) Legno Metallurgia generale	223 222 223 Ciascuno di questi corsi frutta 5 o 10 <i>credits</i> secondo il tempo che gli si dedica.
Disegno (meccanica) Metallurgia generale Nozioni di architettura Lettura di piani	323 323 323 302 Le università riconoscono questi <i>credits</i> ; sempre più esse sperano che gli studenti acquistino una formazione tecnica.
MUSICA	
Canto Orchestra	129 129 Si tratta sia di una iniziazione per l'amatore, sia di studi approfonditi.

Materie e livelli di corso	Osservazioni
Corso d'armonia	129 diti che possono preparare alle carriere musicali.
Apprezzamento musicale	129
Canto	229
Piccolo gruppo vocale	229
Apprezzamento musicale	229
Composizione	329
Madrigali (gruppo vocale)	429
Composizione	429
Corale (selezione di 32 studi)	029
Opera	029
Grande corale	029
Armonia (principianti, gruppo d'uscita, gruppo d'onore)	029
Orchestra (avanzati)	029
Insieme strumentale	029
Piano	029
EDUCAZIONE FISICA, SPORTS, DANZA	
ARTE ORATORIA, ARTE DRAMMATICA	
FOTOGRAFIA	
PRONTO SOCCORSO	

### Lo studio di ciascun alunno prima del suo ingresso a New Trier

Abbiamo detto che la scuola raccoglie informazioni partecolareggiate sui suoi futuri alunni affinché questi possano immediatamente trovare un posto, almeno provvisorio, nel sistema complesso che abbiamo descritto. Ecco come si deve procedere.

La maggioranza degli studenti che frequentano New Trier provengono da sei scuole pubbliche e da sette scuole confessionali dei dintorni.

Fin dal mese di gennaio, i servizi di *testing* della High School forniscono alle sei scuole pubbliche test d'intelligenza, d'attitudini e di conoscenze da amministrare ai bambini

che hanno manifestato l'intenzione di iscriversi ai corsi di New Trier. Gli alunni delle scuole confessionali verranno a Winnetka, in aprile, per essere sottoposti a prove simili.

Ogni scuola inoltre deve riempire una scheda individuale che ha, oltre ai risultati dei test citati, una tabella che riassume l'apprezzamento generale dei professori<sup>10</sup>:

#### Valutazioni globali

Intelligenza	1	2	3	4	5
Applicazione	1	2	3	4	5
Senso di responsabilità	1	2	3	4	5
Condotta a scuola	1	2	3	4	5
Qualità di capo	1	2	3	4	5

#### Valutazioni speciali

Inglese	1	2	3	4	5
Matematica:					
ragionamento	1	2	3	4	5
conoscenze di base	1	2	3	4	5
<i>Social studies</i>	1	2	3	4	5
Scienze	1	2	3	4	5

Codice: 1 = superiore; 2 = al di sopra della media; 3 = medio; 4 = al di sotto della media; 5 = scadente. La cifra riportata viene contrassegnata inquadrandola.

Dopo la amministrazione dei test, il direttore del servizio di *counseling* di New Trier si incontra con i consiglieri delle diverse scuole inferiori per raccogliere informazioni complementari che riguardano non solamente l'istruzione, ma anche la storia, la famiglia, la salute dei futuri studenti.

Poi, ciascuno di questi è intervistato nella sua scuola. Il colloquio verte dapprima sulle abilità (arte, sport, lavoro manuale) e sulle preferenze (materie preferite), poi sulla famiglia, mentre l'intervistatore cerca di identificare gli eventuali problemi. Un quesito scioccante termina l'intervista.

<sup>10</sup> Cfr. New Trier High School, *Test and Personal Data Card*.

Nel mese di maggio, genitori ed alunni — che hanno già potuto studiare il programma dei corsi — sono invitati ad assistere ad una riunione d'informazione alla High School. Vengono raccolte le iscrizioni ai diversi corsi.

La direzione di New Trier è in possesso allora di sufficienti elementi per costituire gruppi di circa 30 studenti che, durante tutti i loro studi, saranno affidati allo stesso consigliere pedagogico e lo incontreranno ogni mattina. In questi gruppi, si cerca sistematicamente di avere l'eterogeneità sia per quanto riguarda le attitudini che le origini socio-economiche. Tuttavia, i gruppi sono costituiti da studenti dello stesso sesso, e i consiglieri sono dello stesso sesso degli studenti. Inoltre, si cerca di dare una personalità equilibrata ad ogni gruppo evitando particolarmente che si trovino riuniti troppi adolescenti « con problemi ».

Il giorno dell'inizio del corso, non solo sono organizzati i corsi e gli orari, ma ogni studente sa esattamente ciò che deve fare: sa, per esempio, che per tale materia, egli deve unirsi al gruppo dei medi e per tal'altra, al gruppo superiore; sa che un consigliere lo aspetta; sa anche quali libri, quali quaderni, quale equipaggiamento sportivo deve comprare ed è stato informato prima del prezzo esatto di ciascuno di questi articoli.

Terminiamo queste indicazioni che si riferiscono agli inizi scolastici segnalando che circa 15 giorni dopo la riapertura delle scuole, i genitori sono invitati ad una prima riunione in cui incontreranno i consiglieri dei loro bambini e nel corso della quale essi saranno informati sui motivi della classificazione ai diversi livelli, sul significato delle manifestazioni extra-scolastiche, ecc.

La settimana successiva, padri e madri vanno a scuola con i loro figli e seguono una giornata di scuola completa al loro fianco (la giornata, in questa occasione, è ripartita tra il pomeriggio e la sera). Si indovina facilmente come tale esperienza faciliti la comprensione tra la scuola e la famiglia.

Cominciati i suoi studi secondari in queste eccellenti condizioni, lo studente sarà seguito, giorno dopo giorno, dal suo consigliere.

## I consiglieri pedagogici

Grazie ai suoi consiglieri pedagogici, la New Trier High School riesce in questo vero tour de force che consiste non soltanto nel guidare efficacemente ognuno dei suoi 3740 studenti nel dedalo dei programmi, ma anche nell'aiutarli sia sul piano psicologico che sul piano sociale e sanitario.

Il *counseling* è assicurato da 126 consiglieri pedagogici a mezzo tempo, professori che dedicano 1/5 del loro tempo a questa missione; 10 consiglieri a pieno tempo che dirigono i precedenti; 6 psicologi a tempo pieno.

Nel corso della riunione quotidiana del mattino, ogni gruppo di 30 studenti di cui abbiamo indicato la costituzione incontra il suo consigliere per discutere di problemi di disciplina, di carriera e, in generale, di tutto quanto può interessare la maggioranza. L'atmosfera distesa di questi incontri costituisce un eccellente passaggio verso il difficile lavoro che sta per cominciare.

Lo studente può anche consultare individualmente il suo consigliere, ogni volta che lo desidera.

Inoltre, durante l'anno, il consigliere farà almeno una visita personale alla famiglia di ogni alunno che appartiene al suo gruppo.

Naturalmente una missione così delicata non s'improvvisa: essa richiede uomini aperti ai problemi della gioventù e particolarmente formati per aiutarla.

Al momento dell'assunzione dei professori, il consiglio di amministrazione della scuola attribuisce un'importanza tutta particolare alle qualità intellettuali e morali indispensabili a esercitare tale funzione nei riguardi degli studenti e dà priorità a coloro che le possiedono.

Durante il suo primo anno di attività, il consigliere deve assistere a 50 riunioni di formazione di circa 1/2 ora ciascuna e dopo ogni ciclo di 4 anni, è sottoposto ad un nuovo addestramento.

La preparazione di base si svolge nella maniera seguente.

## Prima dell'inizio dell'anno scolastico

Un certo numero di sedute sono dapprima dedicate allo studio dell'amministrazione della scuola. Poi, al consigliere vengono comunicate tutte le informazioni raccolte sui suoi futuri alunni.

Infine qualche giorno prima della fine delle vacanze, il direttore delle attività di studio del primo anno riunisce i « freshman helpers », cioè gli studenti dell'ultimo anno che aiuteranno il consigliere nel suo compito. Durante le prime nove settimane, l'assistente parteciperà a tutte le riunioni mattutine. Durante le nove settimane seguenti, non darà che due prestazioni settimanali; in seguito, non interverrà altro che su richiesta del consigliere.

## Durante l'anno scolastico

Sotto la direzione di un direttore specializzato, la formazione prosegue durante le amichevoli conversazioni fatte dai direttori dei settori (inglese, matematica, scienze, ecc.), dai capi dei diversi servizi (direttori del testing, psicologo, bibliotecario, medico-assistente sociale, delegato all'associazione dei genitori) e degli adulti responsabili dei club studenteschi.

Il direttore dei consiglieri convoca anche alcune riunioni speciali per esaminare nuovi problemi che si riferiscono all'ammissione di alunni, alla revisione dei programmi, ecc.

Infine, in aprile, una relazione di chiusura permette di fare il bilancio dell'attività dell'annata che sta per finire e di stabilire le prime previsioni per l'anno seguente.

## Bibliografia

- A. Agazzi, *Les aspects pédagogiques des examens*, Strasbourg, Conseil de l'Europe, 1967.
- F. Bacher, *La normalisation de la notation*, in *Docimologie et éducation*, numero speciale di « Les sciences de l'éducation » 2-3 1969, pp. 131-156.
- *L'évaluation des résultats scolaires au niveau de l'école moyenne*, in « Le travail humain » 28 1965, pp. 219-230.
- J. Block, *Mastery Learning*, New York, Holt, Rinehart and Winston, 1970.
- B. Bloom, Hasting e Madaus, *Formative and Summative Evaluation of Student Learning*, New York, McGraw-Hill, 1970.
- R. Bonnardel, *Application de la méthode d'analyse factorielle de Thurstone à l'étude de la notation des copies d'examens*, in « Le travail humain » 8 1946, pp. 130-139.
- J. Britton, *Experimental Marking of English Composition Written by Fifteen-Year-Olds*, in « Educational Review » 1 1963, pp. 17-23.
- J. Britton, N. Martin e H. Rosen, *Multiple Marking of English Composition, An Account of an Experiment*, London, HMSO, 1966.
- L. Brunelle, *Pourquoi des examens?* Paris, Société des Éditions Rationalistes, 1968.
- B. Choppin e A. Purves, *A Comparison of Open-ended and Multiple Choice Items Dealing with Literary Understanding*, in « Research in the Teaching of English » 1 1969, pp. 15-24.
- M. Demangeon e S. Larcabeau, *Une expérience de correction multiple*, in « BINOP » 14 1958, pp. 131-156.
- Docimologie et Éducation*, numero speciale della rivista « Les sciences de l'Éducation » 2-3 1969, pp. 166.
- F. V. Edgeworth, *The Statistics of Examinations*, in « Journal of the Royal Statistical Society » 51 1888, pp. 599-635.
- Examen des examens*, Numero speciale dei « Cahiers de pédagogie » 92, settembre 1970.

« Examinations Bulletins » London, HMSO:

1. *The Certificate of Secondary Education: Some Suggestions for Teachers and Examiners*, 1963.
  2. *The CSE: Experimental Examinations: Mathematics 1*, 1964.
  3. *The CSE: An Introduction to Some Techniques of Examining*, 1964.
  4. *The CSE: An Introduction to Objective-types Examinations*, 1964.
  5. *The CSE: School-based Examinations*, 1965.
  6. *The CSE: Experimental Examinations: Technical Drawing*, 1965.
  7. *The CSE: Experimental Examinations: Mathematics 2*, 1965.
  8. *The CSE: Experimental Examinations: Science*, 1965.
  9. *The CSE: Trial Examinations: Home Economics*, 1966.
  10. *The CSE: Experimental Examinations: Music*, 1966.
  11. *The CSE: Trial Examinations: Oral English*, 1966.
  12. *Multiple Marking of English Compositions*, 1966.
  13. *The CSE: Trial Examinations: Handicraft*, 1966.
  14. *The CSE: Trial Examinations: Geography*, 1966.
  15. *Teachers' Experience of School Based Examining: English and Physics*, 1967.
  16. *The CSE: Trial Examinations: Written English*, 1967.
  17. *The CSE: Trial Examinations: Religious Knowledge*, 1967.
  18. *The CSE: The Place of the Personal Topic: History*, 1968.
  19. *The CSE: Practical Work in Science*, 1969.
- P. Hartog e E. C. Rhodes, *An Examination of Examinations*, London, Macmillan, 1936.
- P. Hartog, *The Marking of English Essays*, London, Macmillan, 1941.
- E. M. Hinton, *An Analytical Study of the Qualities of Style and Rhetoric Found in English Compositions*, New York 1940.
- F. Hotyat, *Les examens*, Paris, Bourrelier, 1962.
- H. Laugier e E. Schreider, *Recherche docimologique sur un examen de l'enseignement supérieur*, in « Biotypologie » 2 1958, pp. 61-72.
- J. A. Laurwerys e D. G. Scanlon, *Examinations*, The World Year Book of Education 1969, London, Evans, 1969.
- D. Mather, N. France e G. Sare, *The CSE. A Handbook for Moderators*, London, Collins, 1965.
- D. McIntosh, D. Walker e D. McKay, *The Scaling of Teachers' Marks and Estimates*, Edinburgh, Oliver and Boyd, 1962<sup>2</sup>.
- Joint Matriculation Board, *The Marking of Scripts in Advanced Level History*, Universities of Manchester, Liverpool, Leeds, Sheffield and Birmingham, 1964.

- H. S. Otter, *A Functional Language Examination*, Oxford, Oxford University Press, 1968.
- J. C. Passeron, *Sociologie des examens*, in « Éducation et gestion » 2 1970, pp. 6-16.
- D. Pidgeon e A. Yates, *An Introduction to Educational Measurement*, London, Routledge & Kegan Paul, 1968.
- H. Piéron, *Examens et docimologie*, Paris, PUF, 1963, trad. it. Roma, Armando.
- J. B. Piobetta, *Examens et concours*, Paris, PUF, 1943.
- C. Remondino, *Recherche sur les systèmes numériques d'évaluation scolaire*, in « Le travail humain » 3-4 1965, pp. 263-265.
- Reports of the Secondary School Examinations Council*, London, HMSO, 1947-64.
- M. Reuchlin, *L'orientation pendant la période scolaire*, Strasbourg, Conseil de l'Europe, 1964.
- M. Reuchlin e F. Bacher, *L'appréciation des élèves par leurs professeurs*, in « Revue française de pédagogie » 2 1968, pp. 19-25.
- N. Rot e Z. Bujas, *Les distributions de notes scolaires comparées aux distributions des résultats obtenus dans les tests de connaissances*, in « Le travail humain » 22 1959, pp. 19-26.
- R. L. Thorndike, *Marks and Marking Systems*, in R. L. Ebel, *Encyclopedia of Educational Research*, London, Macmillan, 1969, pp. 759-766.
- Symposium sur la docimologie*, XIII<sup>e</sup> Congrès de l'Association Internationale de Psychologie (Roma 1958), in « Le travail humain » 1-2 1959.
- C. W. Valentine, *The Reliability of Examinations*, London, University of London Press, 1932.
- P. Vernon, *Secondary School Selection*, London, Methuen, 1957.
- A. S. Walker, *Pupils' School Records*, Newnes, Educational Publishing, 1955.
- S. Wiseman, *The Marking of English Composition in Grammar School Selection*, in « British Journal of Educational Psychology » 1949, pp. 200-209.
- *Examinations and English Education*, Manchester, University Press, 1961.
- J. Wrigley, *The Relative Efficiency of Intelligence and Attainment Tests as Predictors of Success in Grammar Schools*, in « British Journal of Educational Psychology » 25 1955, pp. 107-116.
- A. Yates e P. Pidgeon, *Admission to Grammar Schools*, London, NFER, 1957.

Finito di stampare  
dalla Grafica Toscana - s.a.s.  
Via Mannelli, 29 r - Firenze  
Settembre 1973

## La nuova scuola media

1. Emma Castelnuovo, *Didattica della matematica*, pp. iv-208. 6ª ristampa.
2. Montagu V. C. Jeffreys, *L'insegnamento della storia secondo il metodo delle « linee di sviluppo »*. Traduzione e prefazione di Lydia Tornatore, pp. xviii-110. 3ª ristampa.
3. Mary L. Northway, *Sociometria scolastica. Breve guida per gli educatori*. Traduzione, prefazione e appendice di Maria Corda Costa, pp. xii-100. 2ª ristampa.
4. Georges Zadou-Naïsky, *L'insegnamento delle scienze fisico-matematiche*. Traduzione di Salvo D'Agostino. Prefazione di Aldo Visalberghi, pp. viii-124. 2ª ristampa.
5. Amleto Bassi, *L'insuccesso e il ritardo nella scuola media*, pp. iv-252. 4ª ristampa accresciuta.
6. Guido Petter, *Problemi psicologici della preadolescenza e dell'adolescenza*, pp. vi-278. 4ª ristampa.
7. Nelson Brooks, *L'apprendimento delle lingue straniere*. Traduzione di Vittorio Giglio. Prefazione di Aldo Visalberghi, pp. xvi-312. 1ª ristampa.
8. Boris Vasil'evic Beljaev, *Saggi di psicologia dell'insegnamento delle lingue straniere*. Traduzione di Alberto Carpitella. Prefazione di Maria Corda Costa, pp. vi-234. 1ª ristampa.
9. Ornella Andreani Dentici, *Abilità mentale e rendimento scolastico*, pp. xii-340. 2ª edizione.
10. Luisa Ceccarelli, Maria Ferretti, Maria Teresa Monari, *Didattica delle osservazioni scientifiche*, pp. vi-242. 2ª edizione.



11. James J. Gallagher, *L'educazione dei ragazzi dotati*. Traduzione di Mirella Manciola Billi. Prefazione di Egle Becchi, pp. xviii-350.
12. Winfred F. Hill, *L'apprendimento. Interpretazioni psicologiche*. Traduzione di Luigi Borelli, pp. vi-262. 1ª ristampa.
13. Bruno Betta, *Educazione civica e storia nella scuola media*, pp. xx-112.
14. C. Burleigh Wellington e Jean Wellington, « *Cattivo rendimento* ». *Problemi e linee di azione*. Traduzione di Alberto Semeraro. Presentazione di Maria Corda Costa, pp. xii-144.
15. Gilbert De Landsheere, *Elementi di docimologia. Valutazione continua ed esami*. Traduzione di Anna Corda, pp. xii-248.

la nuova scuola media **15**

Lire 3000 (2830)

407500F  
La Nuova Italia