

VALUTAZIONE

Gli **OGGETTI** della valutazione sono **fondamentalmente 4**:

- ◆ **il progresso**, di cui fa parte tutto ciò che è a monte di un processo formativo. Occorre quindi considerare e valutare i bisogni formativi, il contesto socioeconomico in cui si opera, i condizionamenti positivi o negativi che hanno marcato gli allievi, le loro conoscenze e abilità pregresse, le risorse educative di cui si dispone;
- ◆
- ◆ **il processo**, cioè tutto l'insieme delle attività formative cui si è dato vita. E' necessario un controllo puntuale sull'itinerario avviato, per intervenire tempestivamente con i correttivi necessari;
- ◆
- ◆ **la performance**, cioè le prestazioni degli allievi, i loro comportamenti;
- ◆
- ◆ **il prodotto**, cioè gli "oggetti" prodotti dagli allievi, i compiti realizzati. Ecco lo schema.

IL PREGRESSO :
bisogni
risorse
abilità di ingresso

IL PROCESSO : sviluppo delle attività educative e didattiche

LA PERFORMANCE : conoscenze
abilità
competenze

IL PRODOTTO : oggetti prodotti dagli allievi.

FUNZIONI SPECIFICHE DELLA VALUTAZIONE.

Impiegando come criteri di classificazione il **TEMPO** in cui la valutazione va compiuta e gli **SCOPI** specifici della verifica, diviene agevole individuare le principali classi delle possibili funzioni valutative.

Si potranno distinguere allora i seguenti momenti e i corrispondenti obiettivi dell'accertamento delle conoscenze, per ognuno dei quali sono individuabili una o più funzioni valutative.

VALUTAZIONE INIZIALE. Gli scopi per cui viene compiuta sono essenzialmente due:

1. rilevare le conoscenze possedute dagli allievi all'ingresso di un anno scolastico o di un corso di studi, per giudicare se il patrimonio cognitivo è adeguato, ovvero per pronosticarne gli esiti (la funzione valutativa è quindi **PROGNOSTICA O PREDITTIVA**) dei probabili risultati;
2. rilevare il grado in cui sono posseduti dagli allievi i prerequisiti cognitivi, ma anche affettivo-motivazionali, in modo che dal conseguente giudizio derivi la decisione di revisionare la programmazione e di attivare procedure con cui garantire in tutti gli allievi il possesso delle abilità-prerequisito (la funzione valutativa corrispondente sarà perciò **DIAGNOSTICA**, ma anche **FORMATIVA**: alla diagnosi dei prerequisiti si accompagna un vero e proprio intervento formativo volto al loro consolidamento e/o recupero).

VALUTAZIONE PROCEDURALE. Si effettua durante lo svolgimento del processo formativo con lo scopo di rilevare analiticamente e con continuità le difficoltà e gli ostacoli che ciascun allievo incontra nell'apprendimento, e di individuare le cause almeno prossime che li hanno determinati, nonché gli eventuali itinerari compensativi che possano rimuoverle.

L'omologa funzione valutativa è **DIAGNOSTICO-FORMATIVA** poiché, alla diagnosi continua delle lacune, si accompagnano interventi finalizzati a migliorare la qualità della formazione attraverso una vera e propria individualizzazione del processo di insegnamento-apprendimento.

VALUTAZIONE INTERMEDIA. Va condotta dopo un periodo dell'attività didattica che ben può coincidere con la chiusura di un quadrimestre. Gli obiettivi che con essa generalmente si perseguono sono due e complementari tra loro:

1. compiere una sorta di bilancio di revisione parziale della programmazione per meglio calibrare i successivi interventi alle necessità e agli obiettivi;
2. esprimere giudizi valutativi che indichino la posizione di ciascun allievo lungo l'itinerario formativo ovvero rispetto agli obiettivi prefissati, i progressi compiuti e le direzioni verso cui sarebbe proficuo mobilitare le energie e l'applicazione per il pieno raggiungimento degli obiettivi. In questo caso la funzione prevalente è quella di migliorare la qualità della formazione che si attuerà in un tempo immediatamente prossimo al processo valutativo compiuto, quindi una **funzione FORMATIVA**, anche se sono presenti elementi propri della valutazione **SOMMATIVA**, di bilancio intermedio, e ovviamente di revisione, della programmazione.

VALUTAZIONE FINALE. Va condotta al termine di un cospicuo periodo di formazione, coincidente con un anno scolastico. Gli scopi che con essa si possono perseguire sono plurimi. Tra i più importanti vanno segnalati: l'espressione di un giudizio complessivo sugli apprendimenti conseguiti (la valutazione è detta anche **COMPLESSIVA** e la funzione sottesa è quella **SOMMATIVA**, nel senso che il giudizio finale deriva da una sorta di somma algebrica, non certo aritmetica, del lavoro compiuto), l'analisi complessiva della qualità dell'istruzione e quindi delle scelte didattiche compiute (una funzione di **BILANCIO CONSUNTIVO DELLA PROGRAMMAZIONE**), l'analisi della produttività qualitativa e quantitativa della scuola (in questo caso si può definire **COMPARATIVA** la funzione della valutazione), il pronostico

dei probabili risultati che ogni allievo potrà conseguire seguendo un determinato corso di studi. Si tratta della valutazione condotta sulla base degli esiti di un esame. **Le funzioni valutative saranno quella PROGNOTICA o PREDITTIVA e quella ORIENTATIVA.**

QUALI SONO I CRITERI DI CLASSIFICAZIONE DELLE PROVE DI VERIFICA?

Un criterio per discriminare e classificare le prove rimanda al tipo di SOLLECITAZIONE OFFERTA e di RISPOSTA RICHIESTA agli allievi affinché mostrino le proprie conoscenze. Infatti, a seconda delle modalità con cui lo STIMOLO viene presentato e la RISPOSTA determinata, e a seconda delle combinazioni dell'uno e dell'altra, deriveranno i diversi gradi di strutturazione dello strumento e i diversi modi di lettura delle prestazioni. In altri termini: dalla qualità dello stimolo e della risposta che uno strumento presenta derivano le principali caratteristiche formali delle prove di accertamento delle conoscenze.

Ai due estremi si avranno prove che presentano STIMOLI APERTI o CHIUSI e RISPOSTE APERTE o CHIUSE.

Lo stimolo può considerarsi aperto quando non presenta indicazioni univocamente determinabili circa la questione posta, lasciando così al soggetto ampi margini di libertà interpretativa della sollecitazione richiesta e quindi dei modi in cui può organizzare la risposta. Lo stimolo seguente: "Illustra le principali caratteristiche dell'età giolittiana" consente ampi margini interpretativi.

E' invece chiuso lo stimolo che pone in modo univoco una o più questioni. La sollecitazione: "Indica il nome della forma di governo in cui tutti i cittadini partecipano direttamente al potere politico e alla promulgazione delle leggi" è chiusa, poiché essa non dà adito a diverse interpretazioni.

A seconda delle caratteristiche dello stimolo e della risposta si avranno, schematizzando, prove che presentano:

- 1. stimolo aperto e risposta aperta;**
- 2. stimolo aperto e risposta chiusa;**
- 3. stimolo chiuso e risposta chiusa;**
- 4. stimolo chiuso e risposta aperta.**

◆ Le prove a stimolo e risposta aperti inglobano la quasi totalità degli strumenti tradizionali, in particolare il tema e l'interrogazione. Col loro impiego, infatti, gli stimoli risultano generalmente poco strutturati, cioè aperti; le risposte, a loro volta, non sono univocamente predeterminabili derivandone così una scarsa o nulla predefinitone delle modalità della loro lettura. Ciò comporta una forte influenza dei fattori soggettivi nell'interpretazione del quesito e nella formulazione della rispo-

sta da parte dell'allievo, e nella determinazione del giudizio da parte di chi corregge e valuta la prestazione.

- ◆ Le prove del secondo tipo possono considerarsi una sorta di non senso valutativo, in quanto l'apertura dello stimolo e la chiusura della risposta rimandano a forme di accertamento in cui risulta predominante lo spazio di chi pone e magari argomenta oltremodo i quesiti, e ridotto al minimo quello riservato a chi è chiamato a rispondere, magari con un sì o con un no, con un sono d'accordo e simili. Questi casi si verificano con docenti logorroici, che amano ripetere la lezione anche in fase di accertamento del profitto, o che impropriamente cercano di mettere a loro agio allievi timidi o poco preparati.
- ◆ Le prove che presentano stimoli chiusi e risposte aperte sono poco diffuse, ma costituiscono la tipologia più flessibile dell'intera strumentazione valutativa di cui si può disporre e quella cui può associarsi un affidabile livello di misurazione delle abilità sottoposte a controllo. Fanno parte di queste prove le domande strutturate, i saggi brevi, alcuni problemi ed esercizi, i riassunti con una chiave interpretativa determinata...
- ◆ Le prove a stimolo chiuso e risposta chiusa sono date dalle prove oggettive di verifica. Così definite perché il peso delle risposte a ciascuno dei quesiti presentati, proprio perché chiuse, può venire determinato prima della somministrazione dello strumento, in modo che la lettura dei risultati di una prova non può che essere la stessa chiunque sia il correttore.

In generale, nella elaborazione di un qualsiasi strumento di accertamento delle conoscenze, **occorre osservare i tre seguenti accorgimenti da cui derivano i principali requisiti che determinano le caratteristiche formali delle prove:**

- ◆ bisogna determinare con precisione, e subito dopo l'identificazione della funzione valutativa che si vuole svolgere, quali conoscenze si vogliono sottoporre a controllo. Si tratta di definire gli obiettivi che con quella prova si intendono verificare;
- ◆ bisogna costruire lo strumento omologo in modo che con il suo impiego gli stimoli offerti permettano agli allievi di manifestare quei comportamenti cognitivi che esprimano la padronanza degli obiettivi. Questo requisito garantisce che la rilevazione compiuta sia valida;
- ◆ bisogna costruire lo strumento in modo che con il suo impiego le risposte date dagli allievi siano, almeno tendenzialmente, univocamente interpretabili, e occorre definire le modalità e i criteri di lettura dei dati da rilevare. Entrambi questi aspetti consentono a uno o più esaminatori di cogliere gli stessi dati, anche quando la prova venga reiterata. I requisiti indicati garantiscono l'attendibilità delle rilevazioni compiute con lo strumento.

Qual è l'utilità del calcolo della media?

Una volta calcolata la media, è facile individuare la prestazione di ogni alunno rispetto alla classe e quindi vedere chi è al di sopra e chi al di sotto della media, sottraendo dalla media il punteggio GREZZO di ciascuno (punteggio grezzo = punteggio che non ha subito nessuna elaborazione).

Confrontando poi le posizioni rispetto alle medie di prove diverse e successive, si può anche calcolare l'andamento complessivo della classe e vederne registrato o il progresso, o il regresso, o la staticità dei risultati.

COME SI LEGGONO I DATI: MISURE DELLA VARIABILITA'

L'informazione data dalla media è ulteriormente sfruttabile. Una volta riportati i punteggi in ordine decrescente, è possibile calcolare la differenza tra punteggio massimo e punteggio minimo = la gamma. Essa costituisce la più semplice misura della variabilità di una distribuzione.

Quanto più alto è il numero che indica la gamma, tanto più alta è la dispersione dei voti rispetto alla media. In questo caso ci si trova di fronte a risultati poco omogenei, gli squilibri di prestazioni sono notevoli.

Possiamo fare ulteriori elaborazioni dei dati: calcolare la differenza di ciascuno rispetto alla media.

La **DEVIAZIONE STANDARD** si calcola in ogni caso in cui si vogliano informazioni più accurate di quelle date dalla media. **Come si calcola? Queste sono le operazioni da compiere:**

1. fare la somma degli alunni;
2. fare la somma dei voti;
3. trovare la media (totale dei voti diviso il numero degli alunni);
4. calcolare di quanto ogni voto è al di sopra della media (con +) o al di sotto (con -);
5. elevare al quadrato le deviazioni per eliminare i segni -;
6. fare la somma delle deviazioni al quadrato;
7. trovare la deviazione media;
8. fare la radice quadrata.

Il valore del SIGMA che misura la deviazione standard ci dice se i risultati si accentrano rispetto alla media e si distribuiscono allontanandosi.

Pertanto: più i dati si disperdono rispetto alla media, più il sigma aumenta, più si accentrano rispetto alla media, più il sigma diminuisce.

MISURE DELLA VARIABILITA' A LIVELLO DI SINGOLO PUNTEGGIO

Nelle misurazioni non interessa solo rilevare l'andamento di ciascun alunno rispetto alla classe (MEDIA E GAMMA), e l'andamento della classe nel suo complesso (MEDIA, GAMMA, SIGMA). Interessa anche, e molto, confrontare il rendimento di

ciascun alunno in prove diverse e confrontare i risultati ottenuti da alunni diversi in serie di prove.

Quando si fa la media dei voti relativi ai risultati ottenuti nella stessa disciplina, si tiene conto che le prove potrebbero avere avuto caratteristiche molto diverse tra di loro?

In termini ancor più problematici: gli insegnanti, quando calcolano la media complessiva dei risultati, si rendono conto che calcolano come se tutte le misurazioni avessero le stesse caratteristiche, mentre sono fatte per discipline spesso molto diverse tra loro?

In termini operativi: quale procedimento consente di confrontare punteggi ottenuti dal singolo alunno in misurazioni con distribuzione diversa di punteggio, indipendentemente dalla facilità maggiore o minore della prova e dalla variabilità del punteggio legata al numero dei quesiti?

Esiste un procedimento per confrontare i punteggi ottenuti da alunni in distribuzioni diverse. E' l'espressione di essi in UNITA' DI SIGMA. E cioè: se conosciamo la media e la deviazione standard, siamo in grado di calcolare altri punteggi, detti PUNTI Z o PUNTI STANDARD, punti tra loro confrontabili perché indipendentemente dalla cifra grezza per mezzo della quale i punteggi sono espressi.

I punti Z sono punteggi derivati, perché si ricavano dal punteggio grezzo modificato per mezzo di tecniche statistiche.

La formula dei PUNTI Z è la seguente:

$$Z = \frac{\text{ogni singolo punteggio grezzo} - \text{media aritmetica}}{\text{Deviazione standard}}$$

oppure, che è la stessa cosa, lo scarto dalla media di ogni singolo risultato
deviazione standard

Facciamo un esempio in cui si confrontano i punteggi di un singolo alunno in prove diverse.

Supponiamo che l'alunno A abbia ottenuto 130 punti in un test avente la media uguale a 120 e il sigma uguale a 10. Supponiamo poi che abbia ottenuto 150 punti in un altro test in cui la media sia uguale a 140 e il sigma a 20.

Applichiamo la formula che permette di passare dai punti grezzi ai punti zeta, o punti derivati.

$$\text{Prima misurazione: } Z = \frac{130 - 120}{10} \quad \text{Risultato} = + 1$$

$$\text{Seconda misurazione: } Z = \frac{150 - 140}{20} \quad \text{Risultato} = + 0,50$$

Come si vede, nonostante l'apparenza, il secondo risultato ha un valore notevolmente inferiore al primo.

Facciamo ora l'esempio di due alunni di cui confrontiamo i punteggi in prove diverse, precisamente in tre prove.

Supponiamo che i due alunni nelle tre misurazioni abbiano ottenuto complessivamente lo stesso punteggio grezzo uguale a 400 e supponiamo che i singoli risultati siano così distribuiti:

alunno A: prima misurazione = 130, seconda misurazione 150, terza misurazione 120

alunno B: prima misurazione = 140, seconda misurazione 130, terza misurazione 130.

Se la media della prima misurazione è 120 e il sigma 10, se la media della seconda misurazione è 140 e il sigma 20, se la media della terza misurazione è 100 e il sigma 15,

si ottiene:

$$\text{alunno A: } Z \text{ prima misurazione} = \frac{130 - 120}{10} = + 1$$

$$Z \text{ seconda misurazione} = \frac{150 - 140}{20} = + 0,50$$

$$Z \text{ terza misurazione} = \frac{120 - 100}{15} = + 1,33$$

TOTALE = 2,83

$$\text{alunno B: } Z \text{ prima misurazione} = \frac{140 - 120}{10} = + 2$$

$$Z \text{ seconda misurazione} = \frac{130 - 140}{20} = - 0,50$$

$$Z \text{ terza misurazione} = \frac{130 - 100}{15} = + 2$$

Il rendimento complessivo del secondo alunno essendo uguale a + 3,50 è superiore a quello del primo alunno.

L'utilizzazione dei punti Z consente una confrontabilità tra le misurazioni di diverse situazioni educative. E' possibile cioè con loro confrontare misurazioni tra discipline diverse.

Per concludere questa parte:

$Z = \text{PUNTEGGIO GREZZO} - \text{PUNTEGGIO MEDIO} / \text{DEVIAZIONE STANDARD}$

Esiste un limite di fruibilità dei punti Z; questo limite consiste nel fatto che essi assumono anche valori negativi. Tali valori possono rispecchiare anche situazioni abbastanza soddisfacenti se si considera che, ad esempio, la media è lo 0.

Per far fronte a tale problema e a tutto ciò che ne consegue in termini di comunicazione con un alunno e di conseguente motivazione, si possono rappresentare gli esiti di un confronto tra due prove anche con i PUNTI T. Attraverso i punti T lo 0 dei punti Z viene in pratica spostato a 50 attraverso la formula seguente:

$$T = Z \times 10 + 50$$

E' chiaro che tutti i valori al di sotto di 50 corrispondono a punti Z negativi.

Con un punteggio $Z = - 3$ si avrà:

$$T = 50 + 10(-3) = 20$$

LE VERIFICHE

La verifica è il momento del controllo di ogni fase significativa del lavoro svolto. Qual è la funzione delle verifiche? Raccogliere una fittissima rete di informazioni, tale da determinare l'opportuna differenziazione degli interventi didattici a seconda delle necessità di ciascuno.

Il problema è di sapere cosa deve essere verificato.

Il controllo scolastico si articola in tre momenti:

scelta dell'oggetto da sottoporre ad accertamento;

necessità di accertare il raggiungimento degli obiettivi;

interpretazione del risultato dell'accertamento.

PRIMA FASE: INDIVIDUAZIONE DELL'OGGETTO DEL CONTROLLO.

Si tratta di stabilire in modo preciso CHE COSA CI INTERESSA MISURARE CON UN CERTO STRUMENTO e di limitarci a prendere in considerazione pochi obiettivi per volta.

In termini didattici: la scelta dell'oggetto da controllare dipende direttamente dall'oggetto che nella programmazione, cioè nelle lezioni, è stato proposto e realizzato.

Se gli obiettivi sono stati tradotti prendendo in considerazione i verbi operatori delle capacità (dei livelli tassonomici) e se i contenuti sono stati precisati nei livelli lessicali-linguistici degli elementi di informazione,

allora **gli oggetti del controllo sono:**

- ◆ per gli obiettivi le capacità o livelli tassonomici,
- ◆ per i contenuti gli elementi di contenuto.

In concreto: se l'Unità didattica è così formulata: Conoscere e comprendere – del giornale – l'editoriale,

gli oggetti del controllo sono:

per gli obiettivi-capacità il conoscere e il comprendere,

per i contenuti l'editoriale.

SECONDA FASE: ACQUISIRE INFORMAZIONI INTORNO ALL'OGGETTO DEL CONTROLLO.

Si tratta di tutte le operazioni che costituiscono la MISURAZIONE EDUCATIVA. Questa fase è definita dal termine MISURAZIONE e comprende le operazioni concrete dell'ACCERTAMENTO.

Il concetto di misurazione può essere definito come il processo attraverso cui si assegnano dei numeri ad eventi, oggetti, persone, secondo determinate regole.

QUATTRO SONO IN GENERE I TIPI DI MISURAZIONE e vanno sotto il nome di SCALE, e sono: scala nominale, scala ordinale, scala a intervalli e scala di rapporti. Tralasciamo la scala di rapporti che nell'ambito delle scienze sociali è molto arduo e complicato adottare.

SCALA NOMINALE.

Usiamo scale nominali quando l'obiettivo è quello di verificare la presenza o l'assenza di una determinata qualità. Si chiamano nominali perché classificano le prestazioni dei ragazzi attribuendole a una classe-insieme piuttosto che a un'altra, definita attraverso un NOME.

Quindi la scala nominale si ha quando, considerando un certo numero di allievi, noi procediamo alla loro classificazione in riferimento a una determinata caratteristica. Noi assegniamo cioè tali allievi a classi o categorie contrassegnate da un nome o numero. Tale procedimento consente solo di stabilire UNA RELAZIONE DI EQUIVALENZA tra tutti i soggetti appartenenti ad una classe.

Nei fatti, ogni lavoro di osservazione sistematica sul comportamento dei ragazzi si traduce in una misurazione a scala nominale, che indica cioè la presenza o assenza dei comportamenti in osservazione.

Dal punto di vista delle verifiche la scala nominale consente di verificare se l'allievo sa o non sa.

Questa scala è molto usata per la VALUTAZIONE FORMATIVA.

Usando la scala nominale si possono fare alcune semplici operazioni:

1. determinare la FREQUENZA con cui compaiono determinati comportamenti o risultati;
2. si può individuare la MODA, cioè la frequenza più alta. La moda è quindi indicativa di una tendenza generale. La rappresentazione grafica dei dati per mezzo di ISTOGRAMMI consente poi una lettura rapida e facile.

SCALA ORDINALE.

In questo caso i dati vengono registrati utilizzando classi di valore secondo un ORDINE che va dal maggiore al minore o viceversa, in quanto tra le varie classi esiste una RELAZIONE DI GRANDEZZA. In altri termini: con la scala nominale SI COSTRUISCE UNA GRADUATORIA e a ciascun alunno viene assegnato un numero che indica il suo posto in graduatoria.

Questo tipo di misurazione permette di graduare comportamenti e prestazioni senza, però, che si possa stabilire la distanza esistente tra un livello e l'altro della graduatoria.

Nel linguaggio quotidiano noi usiamo con frequenza la scala ordinale servendoci di aggettivi o avverbi: quando confrontiamo un film con un altro, o un libro o un deter-sivo, dicendo che è migliore, più facile, più efficace di un altro.

Questo tipo di scala dà luogo a VALUTAZIONI SOMMATIVE O RIASSUNTIVE.

Con questa scala si può determinare la MEDIANA, cioè il risultato che occupa il posto centrale della graduatoria.

SCALA A INTERVALLI

Si ha quando disponiamo di una graduatoria ordinata sulla base di una UNITA' DI MISURA COSTANTE, che permette di misurare l'intervallo tra una posizione e l'altra della scala. Esempio tipico è quello della temperatura: rilevate le temperature in località diverse è possibile, confrontandole, stabilire quale delle temperature è la più alta, ma anche di quanto supera le altre.

La misurazione a scala di intervalli si usa per le VALUTAZIONI SOMMATIVE, cioè verificano unità ampie di apprendimento, costituite da più unità didattiche.

Con questa scala sono possibili altre operazioni oltre a quelle indicate per le scale precedenti:

1. l'addizione e la sottrazione dei risultati ottenuti;
2. il calcolo della MEDIA ARITMETICA e della DEVIAZIONE STANDARD della distribuzione dei dati.

La GAMMA è la differenza fra il valore più alto e il valore più basso.

La MEDIA indica il valore medio raggiunto dagli alunni che hanno partecipato a una prova.

La formula per calcolarla è la seguente:

MEDIA = somma di tutti i punteggi diviso per il numero degli allievi

Qual è l'utilità del calcolo della media? Una volta calcolata, è facile individuare la prestazione di ogni allievo rispetto alla classe, prova per prova, e quindi vedere chi è al di sopra e chi è al di sotto della media sottraendo dalla media il punteggio grezzo di ciascuno. Confrontando poi le posizioni rispetto alle medie di prove diverse e successive, si può anche calcolare l'andamento complessivo della classe e vederne registrato il progresso, o il regresso, o la staticità dei risultati.

Consideriamo la moda, la mediana e la media come parametri che misurano LA TENDENZA CENTRALE.

La MODA nella scala nominale dà la più alta frequenza di un risultato o comportamento. Utilizzata nella scala a intervalli, indica il valore-punteggio che compare più volte nella graduatoria.

La MEDIANA, a proposito della scala ordinale, è il risultato che occupa il posto centrale nella graduatoria. Applicata nella scala a intervalli, se i dati sono dispari di nu-

mero sarà mediana il numero che occupa il posto centrale. In questo caso la formula per trovare la mediana è la seguente:

$MEDIANA = \text{numero dei punteggi} + 1 \text{ diviso per } 2.$

Se invece i dati sono di numero pari, la mediana si calcola dividendo per 2 la somma dei due numeri centrali.

MISURE DELLA TENDENZA CENTRALE A LIVELLO DI CLASSE

Consideriamo la moda, la mediana, la media come parametri che ci consentono di avere un'idea più precisa dei risultati di un gruppo.

MISURE DELLA VARIABILITA' A LIVELLO DI CLASSE

Partiamo da una precisazione: le prove del singolo alunno acquistano una precisa collocazione se confrontate con i risultati delle prestazioni di tutti i compagni. Ciò significa che ogni punteggio attribuito al singolo alunno diviene leggibile solo se si tiene conto di altre informazioni, e cioè se si conosce:

il punteggio massimo teoricamente realizzabile;

il massimo dei punteggi effettivamente realizzati;

il rendimento medio del gruppo = media aritmetica;

la dispersione dei punteggi ottenuti dagli alunni attorno al punteggio medio = gamma.

Ciò significa che un punteggio di 24 acquista diverso valore se è rapportato a 38 o a 50 come massimo dei punteggi. Così pure: se la prova migliore di tutta la classe ha ottenuto 28 su 38, il 24 ha più valore che se il punteggio massimo ottenuto fosse stato il 38, e dieci alunni su 25 avessero superato i 32 punti. Proprio per questo **i punteggi ottenuti dalle misurazioni sono chiamati PUNTEGGI GREZZI. Sono punteggi che acquistano significato solo dopo opportune elaborazioni.**

L'informazione data dalla MEDIA è ulteriormente sfruttabile.

Una volta riportati i punteggi in ordine decrescente, è possibile calcolare la differenza tra punteggio massimo e punteggio minimo. Il numero ottenuto dà la GAMMA.

La gamma costituisce la più semplice misura della VARIABILITA' di una distribuzione.

Si può fare una ulteriore elaborazione dei dati: calcolare la differenza di ciascuno rispetto alla media. In questo caso l'importanza dell'uso della gamma sta nel fatto che ci possono essere casi di misurazione molto diverse, eppure con la stessa media.

Potremmo, però, trovarci di fronte a casi in cui media e gamma siano uguali in differenti misurazioni. Questo mostra la necessità di un'ulteriore elaborazione dei dati, per ottenere informazioni più accurate sulla VARIABILITA'. Questa elaborazione va sotto il nome di DEVIAZIONE STANDARD, che ha il suo simbolo nel SIGMA e che si legge come SCARTO QUADRATICO MEDIO.

La deviazione standard si calcola in ogni caso in cui si vogliono informazioni più accurate di quelle fornite dalla media.

Come si calcola? Illustriamo le operazioni con un esempio concreto.

Alunni	Voto	Deviazione dalla media	Deviazione al quadrato
A	15	+ 5	25
B	14	+ 4	16
C	11	+ 1	1
D	10	-	-
E	9	- 1	1
F	7	- 3	9
G	4	- 6	36
-----	-----	-----	-----
7	70	0	88

Media = 70 diviso 7 = 10.

- ◆ Prima operazione = fare la somma degli alunni = 7
- ◆ Seconda operazione = fare la somma dei voti = 70
- ◆ Terza operazione = trovare la media = 70 diviso 7 = 10
- ◆ Quarta operazione = calcolare di quanto ogni voto è al di sopra (+) o al di sotto (-) della media
- ◆ Quinta operazione = elevare al quadrato le deviazioni per eliminare i segni -
- ◆ Sesta operazione = fare la somma delle deviazioni al quadrato = 88
- ◆ Settima operazione = trovare la deviazione media = 88 diviso 7 = 12,57
- ◆ Ottava operazione = radice quadrata di 12,57 = 3,55 (deviazione standard).

Il valore del sigma ci dice se i risultati si accentrano rispetto alla media e si distribuiscono allontanandosi. Più i dati si disperdono più il sigma aumenta, più si accentrano rispetto alla media più il sigma diminuisce.

MISURE DELLA VARIABILITA' A LIVELLO DI SINGOLO PUNTEGGIO

Nelle misurazioni non interessa solo rilevare l'andamento di ciascun alunno rispetto alla classe (media e gamma), e l'andamento della classe nel suo complesso (media, gamma, sigma). Interessa anche, e molto:

1. confrontare il rendimento di ciascun alunno in prove diverse;
2. confrontare i risultati ottenuti da alunni diversi in serie di prove. Che significa? Quando si fa la media dei voti relativi a risultati ottenuti nella stessa disciplina, si tiene conto che le prove potrebbero avere avuto caratteristiche molto diverse tra di loro? In termini ancor più problematici: gli insegnanti quando calcolano la media complessiva di tutti i risultati, si rendono conto che calcolano come se tutte le misurazioni a-

vessero le stesse caratteristiche, mentre sono fatte per discipline spesso molto diverse tra loro?

In termini operativi: quale procedimento consente di confrontare punteggi ottenuti dal singolo alunno in misurazioni con distribuzione diversa di punteggio, indipendentemente dalla facilità maggiore o minore della prova e dalla variabilità del punteggio legata al numero dei quesiti?

Il procedimento esiste ed è l'espressione dei punteggi in UNITA' DI SIGMA.

Se conosciamo la media e la deviazione standard, siamo in grado di calcolare altri punteggi, detti PUNTI Z o PUNTI STANDARD, punti tra loro confrontabili perché indipendentemente dalla cifra grezza per mezzo della quale i punteggi sono espressi.

I punti Z sono punteggi DERIVATI, perché si ricavano dal punteggio grezzo modificato per mezzo di tecniche statistiche.

La formula dei punti Z è la seguente:

**punti Z = ciascuno dei punteggi ottenuti dagli studenti – la media aritmetica
diviso per la deviazione standard**

Facciamo un esempio. I risultati conseguiti da 10 studenti sono stati i seguenti:

6,6,6,16,16,16,16,20,20 Prima prova

18,21,23,30,37,45,63,63,70,75 Seconda prova

Nella prima prova lo studente ha ottenuto 16, nella seconda 30. Quello che interessa capire è se fra la prima e la seconda prova è possibile notare un miglioramento o un peggioramento. Per confrontare le due prove, se traduciamo i punteggi grezzi in punti Z otterremmo il seguente risultato:

prima prova = 0,39

seconda prova = - 0,66.

In base a tali risultati possiamo concludere che lo studente non è migliorato, anzi.

Mediamente, i punteggi grezzi che si conseguono in una prova a test risultano compresi fra i punti Z – 3 e + 3.

Il limite più evidente dei punti Z è che introducono punteggi negativi. Per ragionare soltanto con i punteggi positivi è sufficiente passare ai PUNTI T.

Per ottenere i punti T bisogna porre convenzionalmente la media aritmetica dei punteggi pari a 50 e la deviazione standard pari a 10, da cui la seguente formula:

$$T = 50 + \frac{10(\text{ciascuno dei punteggi ottenuti dagli studenti} + \text{media aritmetica})}{\text{deviazione standard}}$$

Se, invece, si desidera convertire i punti Z in punti T allora occorre semplicemente applicare la seguente formula:

$$T = 50 + 10 Z$$

L'utilizzazione dei punti Z consente una CONFRONTABILITA' tra le misurazioni di diverse situazioni. E' possibile con loro confrontare misurazioni tra discipline diverse.

L'insegnante, dopo aver costruito un test di profitto, si chiederà se tale strumento funzionerà bene, se cioè misurerà effettivamente ciò che egli vuole misurare, e se gli item sono stati scelti e formulati secondo le indicazioni normative. Sono state indicate le procedure più comunemente usate (coefficiente di correlazione di Pearson, formula di Kuder-Richardson, formula di Spearman-Brown) per calcolare l'attendibilità di un test, come pure i metodi che consentono di determinare, del test stesso, la validità. Tuttavia, nel caso specifico di un test di profitto costruito dall'insegnante per un numero limitato di allievi, è possibile verificare la bontà di tale strumento attraverso il calcolo di due indici, quello di facilità e quello di discriminazione, che consentono di rilevare il contributo dei singoli item alla fedeltà e alla validità del test.

INDICE DI FACILITA'. E' possibile che nel test siano stati inseriti quesiti che non servono allo scopo, perché troppo facili o troppo difficili.

Per il calcolo di tale indice si costruisce una matrice in cui siano riportati i risultati degli alunni negli items del test. Esaminando la tabella è possibile vedere quanto degli alunni hanno dato la risposta esatta per ciascuno degli items e quante risposte esatte ha dato ciascun allievo. Da queste informazioni è possibile calcolare l'indice di facilità utilizzando la seguente formula:

$$\text{Indice di facilità} = \frac{\text{frequenza delle risposte corrette}}{\text{Numero totale di allievi}}$$

Più si avvicina a 1 più l'item è facile, mentre più si avvicina a 0 più l'item è difficile. Gli item troppo facili o troppo difficili hanno scarso peso nell'economia di un test, in quanto non hanno la capacità di differenziare i soggetti. Da un punto di vista statistico l'ideale è l'item che ha un indice di facilità di 0,50. Comunque, gli insegnanti possono utilmente utilizzare quegli items il cui indice di facilità si trova nei valori compresi fra 0.30 e 0.70.

$$\text{FACILITA'} = \text{TOTALE RISPOSTE CORRETTE}/\text{NUMERO ALUNNI}$$

INDICE DI DISCRIMINAZIONE. Questa misura offre indicazioni sulla validità di un item in quanto ne indica il livello di OMOGENEITA' con il test nel suo insieme.

E' possibile, pertanto, calcolando tale indice, individuare ed escludere quegli items che non misurano la stessa cosa che misura il test completo.

Per il calcolo si procede inizialmente raggruppando il 25% oppure il 27% degli allievi che hanno ottenuto complessivamente il punteggio più alto, ed il 25% oppure il 27%

degli allievi che hanno ottenuto complessivamente il punteggio più basso. Si ottengono così un GRUPPO SUPERIORE (S) e un GRUPPO INFERIORE (I).

Successivamente si determina per ciascun item il numero degli allievi del gruppo superiore e il numero degli allievi del gruppo inferiore che hanno risposto correttamente. I dati ottenuti vengono sistemati in una tabella.

Si trasformano quindi in proporzioni i numeri ottenuti. Supponendo che i soggetti siano 30 per ciascun gruppo, dovremmo dividere ogni numero per 30. Risulta più agevole dividere 1 per 30 e quindi moltiplicare ognuno dei numeri per tale costante, che è 0.03. Per esempio: all'item 4 del gruppo superiore il punteggio è 22; allora noi moltiplichiamo $22 \times 0.03 = 0.66$ per il gruppo superiore. Allo stesso item del gruppo inferiore il punteggio è 9; allora noi moltiplichiamo $9 \times 0.03 = 0.28$ per il gruppo inferiore. La differenza fra 0.66 e 0.28, cioè 0.38 è l'indice di discriminazione per l'item 4.

Se tutti gli allievi del S rispondono correttamente ad un item e tutti gli allievi del I rispondono erroneamente, quell'item ha un alto valore discriminativo. E viceversa: se tutti gli allievi del I rispondono correttamente e tutti quelli del S erroneamente, il valore dell'indice di discriminazione è bassissimo. Se, infine, il numero degli allievi che rispondono correttamente a un item è uguale in ambedue i gruppi, il valore dell'indice sarà 0. Solitamente, nella scelta degli item, si accettano valori dell'indice di discriminazione intorno a 0.4° o superiori.

**DISCRIMINATIVITA' = RISPOSTE ESATTE ESTREMO SUPERIORE –
RISPOSTE ESATTE ESTREMO INFERIORE
DIVISO PER IL NUMERO DEGLI STUDENTI DI UN
ESTREMO**

L'indice di discriminatività viene detto tecnicamente CORRELAZIONE PUNTO BISERIALE.

Dobbiamo sapere che può assumere una certa rilevanza, nei test a scelta multipla, la

DISTRIBUZIONE DI FREQUENZA DELLE RISPOSTE ALTERNATIVE.

Quando la frequenza dei distrattori è stata calcolata, le osservazioni che l'insegnante può fare discutendone con gli allievi sono molteplici e di notevole interesse da un punto di vista didattico. Può risultare che il distrattore era troppo facile, e in questo caso non verrà utilizzato in test successivi. Può emergere che l'alternativa corretta era in qualche modo individuabile, e si tratta pertanto di procedere a una migliore formulazione.

Non è facile indicare norme precise in base alle quali decidere l'eliminazione delle alternative ambigue o che comunque non risultano idonee. Si indica tuttavia una procedura che può risultare utile. La formula è la seguente:

**LIVELLO DI FREQUENZA PERCENTUALE = 100 – LA PERCENTUALE
DELLE RISPOSTE ESATTE DIVISO 2 x LE ALTERNATIVE SBAGLIATE.**

Da questa formula si ottiene il valore del livello minimo di frequenza percentuale che un'alternativa sbagliata deve avere per essere mantenuta nell'item

STRUMENTI DELLE VERIFICHE

Perché la verifica sia possibile occorre che (si è già detto):

- ◆ si sappia con esattezza che cosa si vuol verificare: l'informazione realmente data;
- ◆ si sappia con esattezza a quale livello di acquisizione l'oggetto di informazione deve e può essere verificato: sono i livelli tassonomici.

IN SINTESI: VERIFICHIAMO ELEMENTI DI CONTENUTO E LIVELLI DI CAPACITA'.

A livello di INGRESSO le verifiche saranno:

di predisposizione = per accertare i prerequisiti, sia ad inizio di anno che ad inizio di ogni unità didattica o modulo;

A livello di RENDIMENTO le verifiche saranno:

diagnostiche = per accertare eventuali lacune;

di sondaggio = per misurare il livello massimo di profitto e quindi le differenze di rendimento tra gli allievi;

formative di padronanza = per misurare le abilità e i risultati d'apprendimento, e accertare il livello minimo di padronanza.

A livello di RISULTATI FINALI le verifiche saranno:

sommative, riassuntive, terminali = per misurare il successo e classificare gli allievi.

Per ognuno di questi tipi di accertamenti occorre un particolare tipo di verifica. Se noi usiamo i test, abbiamo allora più tipi di test, che chiamiamo TEST DI PROFITTO, e che si differenziano per funzioni e caratteristiche.

Tipo di test	Funzioni	Caratteristiche
Ingresso		
Livello	Misurano capacità che sono i Prerequisiti per apprendimenti Successivi	Per ogni nuovo apprendimento viene indicato un livello minimo di conoscenze che i soggetti Devono possedere. Gli items Del test sono semplici
Processo		
Formativi, di Padronanza	Misurano le abilità e i risultati di apprendimento che tutti Dovrebbero conseguire dopo Ogni sequenza d'istruzione. Danno informazioni sui Progressi degli alunni	Includono tutti gli obiettivi di un'unità di apprendimento. E' importante stabilire un criterio Di padronanza. Gli items del test sono di difficoltà diversa a seconda degli obiettivi e del

		Criterio stabilito
Diagnostici	Rilevano quali sono le Incapacità degli alunni e le Lacune in ciò che dovrebbe Essere stato appreso	Includono prove basate su comuni errori di apprendimento commessi. Gli items sono molto semplici
Sondaggio	Misurano il grado di Differenza nel profitto dei Diversi alunni	Includono obiettivi riferiti a particolari contenuti dei quali si vuole sondare il grado di Differenza nel profitto. E' posto L'accento sul profitto generale
Successo		
Sommativi	Misurano la prestazione alla Fine dell'istruzione, in genere Per una classificazione dei soggetti	Si seleziona un campione rappresentativo degli obiettivi del corso d'istruzione. Gli items del test devono avere diversi gradi di difficoltà

Prima di procedere alla definizione del test di verifica occorre:

- ◆ sapere con esattezza che cosa si vuol verificare (contenuti e capacità);
- ◆ sapere con esattezza a quale livello di acquisizione (livelli tassonomici);
- ◆ sapere con esattezza che tipo di verifica si vuol fare (di predisposizione? Di padronanza? Diagnostiche? Di sondaggio? Sommative?).

L'operazione successiva è quella di costruire un PIANO DEL TEST. Si tratta di decidere:

- ◆ quanto deve essere lungo un test (per un'unità didattica 20 items);
- ◆ quanti items vanno assegnati per ogni contenuto e per ciascun livello tassonomico (il 30% alla conoscenza, il 40% alla comprensione e il 30% ai livelli superiori).
Per un test di 40 items, per esempio, il numero di items relativo alla comprensione (40%), il cui peso specifico è 13, sarà: $13 \text{ per } 40 \text{ diviso } 100 = 5$.

Una volta stabilita la lunghezza del test e la quantità di items di ciascun livello per ciascun contenuto, la fase successiva è quella di scegliere e costruire gli items.

LE PROVE SEMISTRUTTURATE

Sono le prove costituite da una serie di QUESITI CHIUSI che richiedono ai soggetti cui si somministrano di formulare autonomamente le risposte rispettando però alcuni VINCOLI capaci di renderle confrontabili con criteri di correzione predeterminati. I vincoli possono riguardare la lunghezza massima consentita, la quantità e l'ordine di presentazione di elementi o concetti, il grado di generalizzazione da adottare...

Quali sono le prove semistrutturate?

DOMANDE STRUTTURATE.

Presentano una serie di quesiti scritti a STIMOLO CHIUSO, relativi a degli argomenti, contenuti, temi, quesiti che richiedono risposte scritte o grafiche APERTE ma molto brevi e circoscritte.

I quesiti sono preceduti da una INTRODUZIONE all'argomento. Ogni domanda può essere seguita da SOTTODOMANDE particolari a quella strettamente connesse; lo scopo è quello di evitare che un eventuale errore nelle risposte alle prime domande dia origine ad effetti a cascata.

Con tali quesiti si possono rilevare abilità convergenti e divergenti.

Si possono impiegare nell'introduzione materiali scritti, grafici, visivi, sonori, manipolativi e persino strumenti e apparecchi scientifici, e poter partire da questi per le richieste di prestazioni prevalentemente scritte, di interpretazione, ma anche di produzione di grafici, tabelle, diagrammi.

Chi elabora tale strumento deve contestualmente formulare per ogni domanda la relativa risposta. In questa occasione si determinerà il punteggio da assegnarsi ad ognuno dei singoli quesiti. Così, per una domanda strutturata del tipo: Indica quali sono le principali 4 cause dello scoppio della prima guerra mondiale, si potranno assegnare 5 punti alle risposte completamente esatte, ed un punto per ognuna delle cause elencate e descritte quando queste risultino meno di 4.

La somma dei punteggi parziali, relativi ad ognuno dei quesiti, originerà il punteggio complessivo.

SAGGI BREVI

I saggi sono composizioni scritte di lunghezza variabile. Non confondiamoli con i temi. Con i saggi l'articolazione dello stimolo è chiusa e tale da consentire la predefinita dei criteri di correzione; con i temi lo stimolo è aperto, così come aperta è la risposta, creando serie difficoltà di rilevare dati valutativi validi e di predefinire criteri univoci di correzione.

Quali sono le caratteristiche di un saggio breve?

IL SAGGIO COME PROBLEMA.

Scrivere un saggio è come risolvere un problema, descrivere e analizzare criticamente le risoluzioni di esso.

IL SAGGIO NON SI PUO' IMPROVVISARE.

Presuppone necessariamente un periodo di preparazione con letture, ricerche, documentazioni, discussioni, con un approccio mono o pluri disciplinare.

STIMOLO E STESURA DEL SAGGIO DEVONO POSSEDERE PARTICOLARI REQUISITI.

L'argomento o la questione sono totalmente conosciuti prima dell'elaborazione del saggio. Il titolo è, in altri termini, determinato per tempo, non costituisce uno stimolo improvviso e imprevisto. Né uno stimolo aggregato, complessivo e aperto, bensì una serie di stimoli integrati, anche se articolati, in rapporto ai quali siano predefinibili i criteri di congruenza e una scala delle qualità-requisiti delle risposte, capaci di configurarsi come veri e propri criteri oggettivi.

IL SAGGIO DEVE CONSENTIRE L'IMPIEGO DI CRITERI VALUTATIVI ANALITICI, OLTRE CHE COMPLESSIVI.

A seconda delle particolari abilità da sottoporre a controllo si determineranno **i criteri** corrispondenti. Tuttavia, quelli **che in generale caratterizzano molti saggi sono i seguenti:**

- ◆ **proprietà di linguaggio:** esattezza terminologica, padronanza dei termini che hanno significato specifico, uso pertinente del lessico specialistico;
- ◆ **efficacia linguistica:** congruenza della forma espressiva alla funzione comunicativa, uso appropriato dei registri linguistici;
- ◆ **rilevanza delle informazioni:** qualità e peso specifico dei dati e delle informazioni impiegate nella trattazione, nel sostenere o confutare una ipotesi o tesi, capacità di impiegare in modo gerarchico le varie informazioni;
- ◆ **pertinenza delle informazioni:** corrispondenza e congruenza delle informazioni con l'argomento, con il ragionamento svolto;
- ◆ **rigore logico:** stringatezza e sequenzialità del ragionamento, capacità di impiegare un modello di analisi, di spiegazione o di giustificazione di una ipotesi o tesi che non sia contraddittorio al proprio interno;
- ◆ **impalcatura teorica:** razionalità del modello complessivo di ragionamento costruito per sostenere o confutare una tesi;
- ◆ **criticità:** capacità di porre in relazione elementi, dati e informazioni apparentemente non connessi, capacità di impiegare tutta la strumentazione concettuale posseduta dallo studente nel processo di analisi e argomentazione, capacità di cogliere quanto un autore ha volutamente nascosto o taciuto nella giustificazione di certe asserzioni, di affermare consapevolmente l'impossibilità di sostenere certe tesi in assenza di ipotizzabili necessari elementi informativi e concettuali;
- ◆ **originalità:** capacità di aggiungere nuove informazioni, nuove relazioni, al quadro conoscitivo disponibile.

Ovviamente per ciascuno dei criteri indicati vanno determinate opportune scale descrittive e griglie di registrazione dei dati valutativi capaci di far distinguere i diversi livelli di possesso delle specifiche abilità e conoscenze, e quindi di far correggere e attribuire punteggi a ciascuna delle componenti della prova e all'intera prestazione.

RIASSUNTI

Il riassunto può definirsi come una prova di comprensione e di riscrittura condensata di un testo, in rapporto però a determinate chiavi di lettura e a vincoli di riscrittura indicati esplicitamente.

Le INDICAZIONI che è indispensabile offrire per iscritto agli alunni sono in generale le seguenti:

- ◆ specificazione della lunghezza massima del riassunto;
- ◆ indicazioni del genere e del tipo di testo che deve essere condensato (narrativo, descrittivo..., relativo ad argomenti o ambiti letterari, storici, scientifici, economici...);
- ◆ specificazione della chiave di lettura e di riscrittura: indicazione dei punti nodali da porre in evidenza (per esempio: la concatenazione degli avvenimenti, le relazioni tra personaggi ed eventi, la tesi principale sostenuta dall'autore, le idee fondamentali...);
- ◆ indicazione dell'interlocutore ipotetico, ovvero del destinatario del riassunto (riferimento ai codici, ai sottocodici, ai registri);
- ◆ specificazione del tipo di comunicazione che deve essere attivato (riferimento alla funzione comunicativa che in modo prevalente deve caratterizzare la riscrittura: da quella espressiva a quella persuasiva, da quella informativa a quella metalinguistica...).

I CRITERI di verifica saranno strettamente connessi alle specifiche indicazioni offerte come stimolo, compreso il testo, agli allievi.

Quelli generali impiegabili per la valutazione dei riassunti potrebbero essere i seguenti:

- ◆ quantità e qualità delle informazioni scelte come rappresentative del testo: livello di pertinenza ed essenzialità delle scelte rispetto al contenuto;
- ◆ funzionalità della struttura adottata nell'esposizione sintetica: capacità di sintetizzare in forma linguistica originale;
- ◆ coerenza del testo prodotto con quello originario: capacità del nuovo testo di rispecchiare i contenuti di quello originario, ma con una autonoma organizzazione interna del discorso;
- ◆ congruenza del testo prodotto con le chiavi di lettura imposte: scelta adeguata delle componenti interpretative del testo e indicate come vincolanti o come elementi da privilegiare nel lavoro di sintesi;
- ◆ congruenza del registro o dei codici adottati con l'interlocutore indicato: adeguatezza della forma linguistica alle caratteristiche del destinatario;

- ◆ congruenza della funzione comunicativa prevalentemente adottata con quella imposta: capacità del testo prodotto di assolvere la funzione comunicativa indicata come vincolante nelle istruzioni di lavoro.
- ◆ La correzione consisterà nel controllo della presenza in ogni prova delle componenti e dei punti predeterminati come essenziali per ognuna delle indicazioni date, senza farsi influenzare dallo stile espositivo di ciascun allievo.

COLLOQUIO ORALE

L'interazione diretta docente-allievo non può svolgersi su binari precostituiti, ma non può nemmeno costringere ad assumere comportamenti che verifichino solo abilità e conoscenze casualmente sollecitate.

E' indispensabile definire gli obiettivi della prova che dovranno essere controllati. Per evitare però il rischio che le risposte e le argomentazioni originali non vengano prese in considerazione e valorizzate, è necessario che vengano formalmente previste e preventivate divagazioni attorno al tema, e che il colloquio sia rigorosamente flessibile.

Una preventiva determinazione degli obiettivi permetterà di elaborare, prima del colloquio, una griglia che consenta una registrazione tempestiva della qualità delle risposte attraverso l'impiego di segni di codifica semplici e tali da non distrarre l'attenzione del docente e la concentrazione dell'allievo.

Se in una tabella si trascrivono in verticale gli obiettivi con i relativi criteri di accettabilità delle prestazioni e si pongono in orizzontale i nomi degli allievi, nella casella di incrocio obiettivo-allievo si potrà con un semplice segno registrare il possesso o meno dell'abilità corrispondente. Si lascerà uno spazio per eventuali osservazioni integrative di elementi non previsti.

In tal modo si eviterà:

- ◆ di rilevare più volte, nello stesso colloquio, una stessa abilità e di omettere la verifica di quelle repute più rappresentative;
- ◆ di non prendere in esame, anche per effetto dell'oblio, tutte le componenti essenziali della qualità della prestazione complessiva;
- ◆ di sottostimare o soprastimare l'intera prestazione in forza della maggiore o minore impressione destata dalla qualità di alcune risposte;
- ◆ di non offrire articolate informazioni di ritorno agli allievi sui punti deboli e su quelli forti della loro prestazione e quindi delle specifiche competenze e lacune;
- ◆ di far trascurare una più distaccata analisi interpretativa dei risultati di ciascuno, assai utile per modificare il proprio comportamento docente e per riorganizzare sulla base di dati tendenzialmente affidabili la didattica connessa alle funzioni valutative svolte;
- ◆ di differire troppo oltre nel tempo le opportune informazioni di ritorno agli alunni.

Quali sono le NORME GENERALI di conduzione del colloquio?

- ◆ chiarire preliminarmente il tema, gli scopi e gli obiettivi del colloquio;

- ◆ partire da una domanda generale, ampia, non specifica, ma nemmeno troppo generica da disorientare;
- ◆ ascoltare con disposizione amichevole, ma anche in modo intelligentemente critico;
- ◆ evitare di assumere verbalmente e gestualmente comportamenti autoritari, di dissenso netto anche se non si è d'accordo;
- ◆ evitare di dare risposte di dissenso con monosillabi o usando solo la comunicazione non verbale;
- ◆ non creare pause troppo lunghe se l'allievo si blocca o ha dato risposte ambigue, ma intervenire per sollecitare precisazioni, utilizzando non domande dirette e costringitive che potrebbero peggiorare la situazione, ma domande indirette e in forma di prosecuzione del discorso avviato;
- ◆ non bloccare l'allievo se compie delle divagazioni, ma dargli il tempo di accorgersi dell'uscita dal tema, e di tornare sull'argomento. Se non lo fa, intervenire con discrezione per ricondurre il discorso nei propri alvei;
- ◆ chiarire senza impazienza gli eventuali punti non compresi di una domanda, evitando di riformulare nello stesso modo la domanda con toni di derisione;
- ◆ evitare di insistere su una stessa questione oltre un certo limite;
- ◆ offrire il tempo necessario e condurre il colloquio in modo che venga evitato di far pensare agli alunni di aver concesso loro un tempo non sufficiente per esprimere le conoscenze possedute;
- ◆ offrire in modo tempestivo informazioni di ritorno agli allievi sulle specifiche prestazioni con suggerimenti, riflessioni e approfondimenti costruttivi.

LE PROVE TRADIZIONALI

Hanno limiti ampi e preoccupanti. Esaminiamone alcune.

DETTATO. Si misurano, oltre all'abilità di tradurre i suoni in simboli grafici, anche la comprensione uditiva, il senso del testo dettato, la comprensione del testo e delle singole parole, l'ortografia grammaticale.

I risultati sono influenzati dalla lettura preventiva o no del testo da dettare, dalla velocità della dettatura, dalla lunghezza delle unità di volta in volta dettate, dalla pronuncia di chi detta, dal suo tono di voce, dalla distanza da lui dei singoli alunni. Anche somministrazioni più sofisticate (il dettato in cuffia su testi preregistrati) presentano inconvenienti (impossibilità di leggere i movimenti delle labbra di chi detta).

INTERROGAZIONI. Si misurano anche le abilità verbali, la prontezza dei riflessi, la prontezza nel rispondere, la memoria verbale, l'emotività più o meno accentuata. L'ordine della presentazione delle domande influenza i risultati delle prove, la meccanica del loro svolgimento non consente che si abbia tempo per riflettere, o si possano consultare materiali.

Le domande poste sono spesso affidate all'improvvisazione del momento e non seguono criteri uniformi né all'interno della stessa giornata né, tantomeno, in giornate diverse; il grado di difficoltà delle domande e i contenuti sollecitati risultano inevitabilmente disomogenei. Il tempo richiesto per la valutazione di un gruppo di studenti è elevato e il numero di quesiti che è possibile porre nel tempo a disposizione è limitato e non consente di sottoporre a verifica un'ampia gamma di abilità e conoscenze. Il giudizio espresso in merito a un'interrogazione risente inevitabilmente del confronto con quelle precedenti; più in dettaglio, tale limite può essere definito dai seguenti effetti legati all'utilizzo di prove non oggettive: effetto di alone, di contrasto, di stereotipia. La scarsa trasparenza e obiettività dei criteri di giudizio porta con sé il rischio di una facile demotivazione degli alunni, che possono sentirsi valutati sulla base di criteri non tanto obiettivi e meritocratici, quanto piuttosto frutto di personali scale di misura elaborate dall'insegnante e, ciò che più importa, non esplicitate agli studenti.

LE PROVE SCRITTE. Ciò che influenza il giudizio dell'insegnante, tanto da renderlo ampiamente arbitrario, risulta essere:

la specifica situazione di valutazione: per esempio, come sono andati i compiti degli altri studenti rispetto a quello valutato, oppure se il compito è fra i primi o fra gli ultimi corretti;

le esperienze precedentemente acquisite: il giudizio che l'insegnante si è fatto una volta delle capacità di ciascun allievo tende spesso a irrigidirsi in uno stereotipo che consente, nel tempo, variazioni marginali;

il tipo di compito assegnato: temi, problemi, traduzioni, spesso mettono in gioco conoscenze e abilità diverse, che sono difficilmente valutabili in quanto sollecitate contemporaneamente.

PROBLEMI ED ESERCIZI. Le abilità stimolate si riducono spesso all'applicazione meccanica di regole, principi, formule.

Le situazioni proposte sono spesso artificiali e astratte, perciò difficilmente comprensibili agli alunni. Manuali, prontuari, libri di testo sono spesso sottratti alla consultazione, e si impone così un sovraccarico rilevante alla memorizzazione di regole e formule.

TEMA. Impone una vera e propria amplificazione della traccia, costringendo gli alunni a formulare con mille parole ciò che è possibile esprimere più efficacemente con dieci. Non impone di tener presenti interlocutori ben determinati con cui comunicare per ben definiti scopi. Non consente di valorizzare alcune caratteristiche peculiari del linguaggio scritto, prima fra tutte quella di permettere, a chi lo impieghi, di poter tornare su questioni già trattate da altri, di utilizzare elaborazioni e riflessioni compiute precedentemente da altri per suffragare le proprie ipotesi o tesi, per confutarne altre consolidate. Non permette di utilizzare, talvolta persino di capire che ci sono linguaggi differenziati, oltre che tecnici e settoriali, per trattare gli stessi argomenti a seconda delle sedi e degli interlocutori. Suggerisce comportamenti immorali, e certamente diseducativi, autorizzando ad esprimere, per compiacere l'insegnante, persino l'opposto di quello che si pensa si determinate questioni, nonché a fare asser-

zioni poco documentate e argomentate. Offre agli alunni stimoli poco precisi o comunque tali da venire interpretati soggettivamente. Non permette al docente di predefinire criteri omogenei di correzione o di dare ai criteri comunque utilizzati lo stesso peso nell'analisi delle diverse prestazioni, anche perché queste potrebbero non evidenziare abilità che, pur possedute, non vengono manifestate solo a causa della soggettiva interpretazione della traccia.

TRADUZIONE. Le abilità misurate sono artificialmente aggiuntive a quelle fondamentali dell'apprendimento linguistico: capire, parlare, leggere, scrivere. Le traduzioni fatte a scuola sono escogitate prevalentemente per misurare la padronanza di regole e di eccezioni (e sono per questa ragione fortemente artificiose), mentre le traduzioni utilizzate in contesti non artificiali sono giudicate in funzione di caratteristiche diverse, quali l'accuratezza delle informazioni (nella traduzione simultanea, nella traduzione degli articoli) e la resa estetica (nella traduzione di brani letterari). Le frasi e i brani proposti per le traduzioni, in quanto isolati da contesti più ampi, sono di solito privi di senso e di motivazione. L'esercizio della traduzione tende ad allontanare da contatti culturalmente più interessanti con i prodotti letterari stranieri. Grammatiche, vocabolari sono spesso sottratti alla consultazione, e su impone così in sovraccarico rilevante alla memorizzazione di regole ed eccezioni.

IN ESTREMA SINTESI SI PUO' DIRE CHE LE PROVE TRADIZIONALI SONO SOGGETTE AD UNA NOTEVOLE VARIABILITA' DELLE VALUTAZIONI SE E' VERO, COME E' STATO PIU' VOLTE RICONTRATO, CHE ADDIRITTURA LO STESSO INSEGNANTE, A DISTANZA DI TEMPO E POSTO IN CONDIZIONI DIVERSE, PUO' ESPRIMERE GIUDIZI DIVERSI IN MERITO ALLA STESSA PROVA. TALE VARIABILITA' ALLONTANA LE INTERROGAZIONI E LE PROVE SCRITTE NON STRUTTURATE DALLA COSTANZA DELLA MISURAZIONE, CIOE' DALLA PRIMA QUALITA' RICHIESTA PERCHE' LA MISURAZIONE STESSA RISULTI ATTENDIBILE.

UN ALTRO REQUISITO FONDAMENTALE CHE LE PROVE TRADIZIONALI DISATTENDONO E' QUELLO DELLA VALIDITA'. SI SA COME, NEL CASO DELLE PROVE DI VERIFICA TRADIZIONALI, NON SOLO DOCENTI DIVERSI ESPRIMANO GIUDIZI DIVERSI SU UNA STESSA PROVA, MA ANCHE COME LO STESSO INSEGNANTE, AL VARIARE DELLE CONDIZIONI IN CUI SI TROVA A VALUTARE LA STESSA PROVA, POSSA ESPRIMERE GIUDIZI ASSAI ETEROGENEI TRA LORO.

Una volta deciso quale tipo di item utilizzare, resta la **STESURA DEI QUESITI**. Si consiglia che sia fatta dagli insegnanti stessi. Per quali ragioni?

Perché la scala di valutazione o i test standardizzati e non, non corrispondono a ciò che gli insegnanti di fatto richiedono e generalmente non sono adatti a prove che si innestano su operazioni calzanti a bisogni specifici delle classi.

Così anche rimangono fuori le schede di misurazione-valutazione allegate ai libri di testo, perché non reggono alla sequenza dinamica dell'apprendimento realmente svolto e non sono in grado con ciò di misurare difficoltà e progressi degli alunni, né di corrispondere ad esigenze di misurazione nate dalle effettive differenze individuali di apprendimento, che sono poi quelle da prendersi in considerazione per il lavoro di intervento tempestivo e recupero.

Scelti gli items, questi **vanno montati** in una sequenza che ubbidisca a certe regole:

è bene raggruppare gli items secondo il tipo di abilità da misurare e mettere insieme tutti gli items che misurano la conoscenza, poi quelli che misurano la comprensione e poi quelli che misurano i processi intellettuali superiori;

è bene ordinare gli items di ogni livello di abilità, dal più facile al più difficile;

è bene raggruppare gli items secondo il loro tipo, nel caso che nello stesso test si ricorra a items di tipo diverso. I primi saranno sempre quelli vero-falso.

La fase di MONTAGGIO si completa con la preparazione delle **CONSEGNE**, cioè con la scrittura delle indicazioni che spiegano agli allievi come dovranno rispondere ai quesiti. E' sempre opportuno che le consegne contengano un esempio di risposte.

Ecco un esempio che indica i tre principali aspetti della consegna: lo scopo del test, il tipo di prove impiegate, come rispondere.

CONSEGNA.

Questa prova ha lo scopo di accertare ciò che tu hai appreso dell'unità didattica di geografia "Aspetti fisici del continente europeo". I risultati della prova saranno utilizzati per chiarire quali sono le difficoltà che tu incontri in questo argomento.

La prova è composta di 32 domande a scelta multipla.

Il tuo compito è quello di segnare, per ciascuna domanda, la risposta esatta individuandola tra le 4 risposte alternative. Ecco un esempio:

Qual è la capitale d'Italia?

- A. Bologna
- B. Firenze
- C. Roma
- D. Venezia.

L'ultimo accorgimento, in fase di montaggio, e in vista della correzione, è quello della **ASSEGNAZIONE DEI PUNTEGGI**. Al riguardo si terrà presente la tabella precedentemente riportata.

Per la SOMMINISTRAZIONE DEL TEST si raccomanda che:

- ◆ gli allievi sappiano in anticipo di qualche giorno la data della prova;
- ◆ prima della distribuzione dei fogli con i quesiti, si controlli che tutti abbiano l'occorrenza per scrivere;
- ◆ l'insegnante legga attentamente le istruzioni, le illustri con un esempio, si accerti che tutti abbiano capito bene, comunichi il tempo indicato alla prova;
- ◆ durante la prova l'insegnante riduca al minimo le interruzioni, eviti di dare ulteriori spiegazioni, si accerti che gli allievi non copino tra di loro, avverta quando mancano pochi minuti al tempo della consegna;
- ◆ a tempo scaduto gli allievi appoggino sul banco la penna e l'insegnante ritiri velocemente i fogli.

TERZO MOMENTO : LA VALUTAZIONE

La valutazione, come fase conclusiva del controllo, si identifica con quella serie di operazioni che, fondandosi sui dati della misurazione, attribuisce loro un significato dal punto di vista educativo.

Alle VERIFICHE DI INGRESSO corrispondono VALUTAZIONI DI INGRESSO.

Sulla base dei dati si valuta se iniziare subito con la nuova informazione o se procedere prima ad un'operazione di **RECUPERO PRECOCE**.

Alle VERIFICHE DEL PROCESSO corrispondono VALUTAZIONI DEL PRO-

CESSO. Sulla base dei dati raccolti dai controlli sulle fasi del processo, si valutano le condizioni di acquisizione dell'istruzione in rapporto agli obiettivi delle unità didattiche. Sulla base di tali valutazioni si giudica se iniziare subito la nuova unità didattica o procedere ad un'operazione di **RECUPERO TEMPESTIVO** usando correttivi adatti.

Alle VERIFICHE DEL SUCCESSO corrispondono VALUTAZIONI DEL SUC-

CESSO O RIASSUNTIVE. Sulla base di accertamenti fatti su ampi settori di informazioni, che possono essere scanditi per esempio quadrimestralmente, si valuta il raggiungimento degli obiettivi finali, terminali. Questa valutazione si presenta come conclusione di un ciclo e come punto di riferimento delle programmazioni future: chiude e insieme collega.

I CRITERI DI VALUTAZIONE che potremmo usare sui controlli del profitto, sono di due tipi: assoluto e relativo.

Esaminiamo prima i **CRITERI DI VALUTAZIONE IN ASSOLUTO**.

Il rendimento dell'alunno in rapporto all'obiettivo e non ai compagni.

Per una valutazione del genere occorre una prova che distingua chi sa da chi non sa. I test per il controllo del profitto sono quelli di **PADRONANZA** (misurare abilità e ri-

sultati di apprendimento e accertare il livello minimo di padronanza) e DIAGNOSTICI (per accertare le lacune sulla padronanza).

La SCALA di riferimento per la valutazione in questi casi è la NOMINALE.

Questo tipo di valutazione appartiene alla valutazione FORMATIVA, cioè alla valutazione dell'apprendimento.

Il rendimento dell'alunno in prove diverse e in rapporto al rendimento iniziale.

Ogni alunno deve essere valutato indipendentemente dalla facilità maggiore o minore della prova e dalla variabilità del punteggio legata al numero dei quesiti, ma sempre in rapporto agli obiettivi. In questo caso passeremo attraverso la MEDIA e la DEVIAZIONE STANDARD.

I test potrebbero essere di PADRONANZA e/o di SONDAGGIO, e gli items del tipo a scelta multipla, seppure non esclusivamente.

La SCALA di riferimento per la valutazione da usarsi in questo caso è quella AD INTERVALLI, che consente il calcolo della media e della deviazione standard. La formula per ottenere punteggi tra loro comparabili è quella dei PUNTI Z.

CRITERI DI VALUTAZIONE RELATIVI

Criteri di confronto con il gruppo.

E' chiaro che devono confrontarsi tutte le misurazioni di una prova e il risultato del singolo acquista un senso relativamente al gruppo.

I test possono essere DIAGNOSTICI, DI SONDAGGIO, DI PADRONANZA.

La SCALA di riferimento da usarsi è quella ORDINALE (per stabilire il posto che ciascuno occupa nella graduatoria) e AD INTERVALLI (che consente di stabilire sia la misura della tendenza centrale a livello di classe – moda, media aritmetica -, sia la misura della variabilità a livello di classe – gamma e sigma).

Criteri di confronto con i compagni singolarmente.

Possiamo utilizzare test DIAGNOSTICI, DI SONDAGGIO, DI PADRONANZA. La SCALA di riferimento è quella ORDINALE (per stabilire il posto che ciascuno dei compagni, messi a confronto, occupa nella graduatoria), e quella AD INTERVALLI (che consente il calcolo della media e del sigma).

La formula per ottenere punteggi che siano tra loro comparabili è quella dei PUNTI Z, che consente la misura della variabilità a livello di singolo punteggio o di singolo allievo.

Criteri di confronto con le condizioni sociali di partenza dell'alunno.

Una valutazione del genere presuppone un'attenta informazione iniziale e l'osservazione sistematica, per individuare ed esplicitare i modi, i luoghi, le circostanze in cui si manifestano gli svantaggi della condizione sociale.

Ciò significa che il campo di osservazione dovrebbe essere non solo cognitivo, ma sociale, affettivo.

In questi casi occorre mirare a porre gli svantaggi come problema da risolvere nella durata annuale e non in fondo, nella valutazione riassuntiva.

La scala di riferimento da usarsi è quella della presenza o assenza dei comportamenti che si presume derivati da svantaggi sociali. La valutazione in questi casi si traduce in resoconto delle osservazioni fatte, dei rimedi apportati, delle conversazioni fatte, degli interventi operati, delle decisioni collegialmente prese.

In termini di valutazione riassuntiva si dovrebbero riportare le condizioni di ingresso con quelle finali, indicando gli interventi e valutandoli nei risultati.

SINTESI OPERATIVA

1. Qual è l'oggetto del controllo?

CONTENUTI E CAPACITA'.

2. In che modo controllare l'apprendimento, che tipo di misurazione possiamo utilizzare?

Acquisiamo informazioni tramite 4 scale: SCALA NOMINALE, SCALA ORDINALE, SCALA A INTERVALLI, SCALA DI RAPPORTI.

3. Quali strumenti utilizziamo per le verifiche?

Le PROVE OGGETTIVE DI PROFITTO o PROVE STRUTTURATE : vero/falso, corrispondenze, integrazioni, confronti, completamenti, scelte multiple.

Le PROVE SEMISTRUTTURATE : domande strutturate, saggi brevi, riassunti, colloqui orali.

Le PROVE TRADIZIONALI : dettati, problemi ed esercizi, interrogazioni, temi, traduzioni.

4. Che tipo di verifica si vuol fare?

Di PREDISPOSIZIONE, DI PADRONANZA, SOMMATIVE...?

5. Quanto deve essere lungo un test?

20 items per unità didattica.

6. Quanti items vanno assegnati per ciascun livello tassonomico dei singoli contenuti?

ALLA CONOSCENZA IL 30%, ALLA COMPrensIONE IL 40%, AI LIVELLI SUPERIORI IL 30%.

7. Scegliere e costruire gli items. Quali dobbiamo usare, e quanti?

8. Stesura dei quesiti.

9. Montarli in una sequenza che ubbidisca a regole.

10. Assegnazione dei punteggi.

11. Somministrare il test.

12. Valutare.

PROVE OGGETTIVE

E' importante chiarire subito che l'oggettività si riferisce alla possibilità di predeterminare in modo chiaro l'esattezza delle risposte. E' evidente che tale possibilità tanto più aumenta quanto più precisi, chiari, circoscritti sono gli stimoli proposti dalle domande e quanto più viene limitato il numero delle risposte possibili.

Operativamente, l'univocità della risposta esatta può essere ottenuta proponendo un quesito a soluzione obbligata ("Il passaggio dallo stato liquido allo stato gassoso si chiama...", oppure: $16 \times 4 = \dots$), oppure proponendo un quesito con un numero predefinito di alternative tra cui scegliere la risposta esatta.

E' opportuno chiarire che l'obiettivo di approssimarsi il più possibile ad una valutazione coerente rispetto agli obiettivi didattici prefissati non significa certo "misurare" le prestazioni o le conoscenze degli alunni come un fisico misura la velocità di un corpo. Non si può affermare, dunque, che la prestazione di uno studente in una prova a test possa essere calcolata con strumenti "scientifici", così come non si tratta di ridurre i soggetti a oggetti quantitativamente misurabili.

Lo scopo della valutazione deve essere quello di verificare la corrispondenza o meno dei risultati raggiunti dagli alunni rispetto a obiettivi e criteri predeterminati, chiari, rigorosi, esplicitati. Non ha senso affermare che l'individuo è ridotto a una serie di numeri. Si deve più correttamente affermare che in un caso (prove di verifica tradizionale) la valutazione è effettuata sulla base di un'informazione incerta e destrutturata, nell'altro (prove oggettive) sulla base di un'informazione che si cerca di uniformare a scelte e termini di confronto esplicitamente individuati e perciò, in quanto tali, obiettivi.

Gli item di cui sono costituite le prove oggettive sono fundamentalmente le prove di tipo vero/falso, le prove di corrispondenza, le prove di completamento (cloze), le prove a scelta multipla.

QUESITI VERO/FALSO

Essi richiedono che la scelta del soggetto venga operata tra due soluzioni. Perciò, data un'affermazione, si richiede di indicare se essa sia vera oppure falsa.

Questo tipo di item verifica obiettivi modesti a livello tassonomico: l'alunno deve dimostrare di conoscere banalmente una legge, un principio elementare, un concetto... Un ulteriore elemento sfavorevole per questo tipo di item è costituito dal fatto che la possibilità di rispondere a caso in modo corretto sono elevate (il 50%).

Per ovviare agli inconvenienti di questo tipo di quesito si raccomanda di:

- A. allungare molto il test, proponendo non meno di 50-60 items;
- B. decidere, prima della somministrazione, di accettare punteggi alti come indici di profitto;
- C. sottrarre punti per le risposte sbagliate o omesse, usando la formula

$$P = E - S$$

in cui P = punteggio reale

E = numero risposte esatte

S = numero risposte sbagliate-omesse;

- D. richiedere di motivare in breve le ragioni della scelta adottata.

INDICAZIONI PER COSTRUIRE CORRETTAMENTE UN TEST CON QUESITI TIPO VERO/FALSO

- ◆ Le asserzioni non devono contenere negazioni.
- ◆ Le asserzioni non devono contenere più di una idea.
- ◆ Il linguaggio deve essere preciso e dettagliato.
- ◆ Il numero delle asserzioni esatte deve essere uguale a quello delle asserzioni sbagliate.
- ◆ Parole chiave come “sempre, mai, quasi, dovunque, spesso, a volte...” vanno evitate.
- ◆ Le asserzioni devono essere chiare ed esplicite, non caute e circospette.
- ◆ Le asserzioni devono essere brevi.

ITEM DI CONFRONTO (matching item).

Nella colonna 1 sono indicati i nomi di 4 mari europei. Nella colonna 2 sono trascritti i nomi di alcuni Stati europei. Scrivi accanto al nome di ciascun mare della prima colonna la lettera che precede il nome dello Stato che confina con quel mare.

Colonna 1

_____ La manica
_____ Mar del Nord
_____ Mare Adriatico
_____ Mar Nero

Colonna 2

A. Bulgaria
B. Danimarca
C. Finlandia
D. Francia
E. Islanda
F. Jugoslavia

ITEM CORRISPONDENZE o ITEM AD ASSOCIAZIONE.

A ciascuno dei tipi di impresa elencati a sinistra corrisponde un esempio nell'elenco a destra. Scrivi nello spazio indicato la lettera che indica l'esempio che ritieni corretto.

- | | |
|---|-------------------------------------|
| 1. piccola impresa commerciale _____ | a) allevamento di polli in batteria |
| 2. ente pubblico _____ | b) supermercato |
| 3. grande impresa industriale _____ | c) studio medico |
| 4. impresa individuale di servizi _____ | d) FIAT |
| 5. piccola impresa agricola _____ | e) negozio del fruttivendolo |
| 6. grande impresa agricola _____ | f) coltivatore diretto |
| 7. impresa artigiana _____ | g) ente per l'energia elettrica |
| 8. grande impresa commerciale _____ | h) sindacato dei lavoratori |
| | i) bottega del falegname |
| | l) confederazione degli industriali |

Individuare l'esatta corrispondenza tra gli artisti e le opere seguenti, abbinando il numero di ciascuna opera alla lettera dell'artista che ne è autore.

OPERE

1. Vergine delle rocce
2. Scuola di Atene
3. La nascita di Venere
4. La morte della Vergine
5. Battesimo di Cristo

AUTORI

- a. Piero della Francesca
- b. Leonardo da Vinci
- c. Raffaello Sanzio
- d. Caravaggio
- e. Botticelli

INDICAZIONI PER COSTRUIRE CORRETTAMENTE QUESITI DI CORRISPONDENZE.

- ◆ Numero ottimale delle corrispondenze da compiere: da 5 a 15.
- ◆ Ciascun elemento del primo elenco contenga il maggior numero possibile di parole. Ciascun elemento del secondo sia il più breve possibile. (In questo modo si ha maggior chiarezza intorno alle cose da fare e si perde meno tempo a cercare le risposte).

- ◆ A ciascun elemento del primo elenco ne corrisponde uno solo esatto del secondo.
- ◆ I componenti di ciascun elenco siano omogenei tra loro. Se l'omogeneità manca, la risoluzione può essere resa più facile – se si individuano eventuali insiemi di corrispondenze omogenee –, o più difficile – se tali insiemi non ci sono o non sono individuati.
- ◆ Nel secondo elenco si aggiungano a quelli esatti elementi sbagliati, al massimo fino a raddoppiare il numero di quelli esatti.
- ◆ L'elencazione degli elementi di entrambi gli elenchi sia fatta seguendo l'ordine alfabetico.

ITEM A COMPLETAMENTO O CLOZE O INTEGRAZIONE

Questo tipo di item richiede di completare un brano all'interno del quale sono stati eliminati alcuni termini. A conclusione del brano sono presenti, senza un ordine preciso, i termini eliminati ed altri che funzionano da distrattori. La verifica che questo item consente è abbastanza elementare. Per rendere più difficile questo tipo di item, il quesito viene strutturato nella forma a scelta multipla, in modo tale che siano predefinite le alternative possibili e, tra queste, ve ne sia una sola univocamente esatta.

Un'altra strategia a cui si consiglia di ricorrere per eliminare ogni ambiguità in tali quesiti è quella di inserire due lacune (in questo caso, anche ogni alternativa dovrà proporre due termini).

Facciamo un esempio.

“Le sue risposte erano solitamente... In quell'occasione, tuttavia, egli fu molto...”

- A. offensive – sintetico
- B. cortesi – gentile
- C. arroganti – superficiale
- D. concise – prolisso

INDICAZIONE PER COSTRUIRE CORRETTAMENTE I COMPLETAMENTI.

- ◆ Numero ottimale degli spazi vuoti: da 5 a 10.
- ◆ La contrapposizione tra le parole escluse e quelle a ciascuna di esse alternative sia reale. Se la contrapposizione non è reale, la scelta rischia di essere casuale.
- ◆ La eliminazione delle parole non deve impedire di comprendere la struttura del periodo di cui esse fanno parte.
- ◆ Se le integrazioni da fare riguardano dai quantitativi, siano dati i riferimenti per collocarli. Non si avrebbero altrimenti criteri per rispondere.
- ◆ Tra il testo da integrare e le parole escluse (e le loro alternative) ci sia continuità logica, grammaticale, sintattica.
- ◆ La larghezza degli spazi vuoti sia costante. Si può essere tentati di renderla proporzionale alla lunghezza delle parole mancanti, rendendone più facile l'individuazione.

- ◆ Per ogni vuoto ci sia, oltre alla risposta esatta, anche un'alternativa. Non è detto che anche tra le parole esatte ci sia alternativa. Se una parola esatta, posta in alternativa ad un'altra parola esatta, viene collocata in un vuoto, la collocazione dell'altra è quasi automatica.
- ◆ Elencare tutte le parole (esatte ed errate) in ordine alfabetico dopo il testo da integrare, e collocarle in due o più colonne. Si evitano inavvertite corrispondenze tra l'ordine della elencazione e l'ordine secondo cui le integrazioni sono presentate. Si facilita il reperimento delle parole.

ITEM A ORDINAMENTO

Si tratta di quesiti che richiedono di individuare l'ordinamento corretto di una serie di elementi proposti, sulla base di un preciso criterio che può essere temporale, logico, di misura...

Facciamo un esempio.

Individuare l'alternativa che presenta i filosofi ordinati cronologicamente:

- A. Aristotele – Boezio – Cartesio – Kant – Hume
- B. Kant – Fichte – Schopenhauer – Gentile – Wittgenstein
- C. S. Agostino – Pomponazzi – Leibniz – Hegel – Schelling
- D. Feuerbach – Hegel – Marx – Kierkegaard – Nietzsche

La loro applicazione si concentra soprattutto nei casi in cui si vogliono verificare conoscenze, ma anche abilità di confronto e di analisi.

SCELTA MULTIPLA

Costituiscono la modalità più flessibile dei quesiti che è possibile porre quando si vogliono utilizzare come strumento di verifica le prove oggettive o prove strutturate di conoscenza.

Sono costituiti, per lo più, da un'affermazione, dalla presentazione di una questione, cui fa seguito una serie di risposte o soluzioni alternative tra le quali va scelta quella che si considera corretta. Dal momento che l'affermazione iniziale può presentarsi anche in forma interrogativa – configurando così una questione piuttosto complessa -, la sua delineazione può avvenire anche attraverso l'uso di grafici, schemi, tabelle, illustrazioni, materiale scientifico, disegni, foto, diagrammi, figure...

Diventa così possibile rilevare una vasta gamma di obiettivi valutativi, dai più semplici ai più complessi.

Possono costruirsi scelte multiple anche a due soluzioni.

INDICAZIONI PER UNA CORRETTA COSTRUZIONE DELLE SCELTE MULTIPLE

- ◆ Nella parte generale presentare un problema, nelle alternative le soluzioni.
- ◆ Tutto quel che c'è di comune fra le alternative figure, nei limiti del possibile, nella parte generale. Si evitano appesantimenti nella lettura e perdite di tempo. Le alternative si differenziano in modo più netto.
- ◆ Numero ottimale delle alternative: 4 o 5.
- ◆ Le risposte esatte siano distribuite fra tutte le alternative. Ci si aspetta che sia così.
- ◆ Nella parte generale presentare un solo problema.
- ◆ Evitare, nei limiti del possibile, che le alternative presentino negazioni.
- ◆ Le parole usate nelle alternative abbiano un significato preciso, non generico.
- ◆ Le alternative devono escludersi a vicenda. Più di una risposta potrebbe altrimenti apparire esatta. Oppure le alternative sono individuate più facilmente come sbagliate.
- ◆ Tra parte generale e alternative ci deve essere sempre continuità sintattica e grammaticale.
- ◆ Per le risposte sbagliate usare asserzioni esatte, ma senza riferimento alla domanda proposta.
- ◆ Nelle alternative non devono figurare parole chiave come “sempre, mai, dovunque, spesso...”.
- ◆ Le alternative devono essere omogenee per lunghezza.
- ◆ Le alternative devono essere omogenee nel linguaggio utilizzato. Quella che usa un linguaggio più tecnico è probabilmente esatta.
- ◆ Nessuna o tutte le alternative devono riprendere concetti e termini del problema. Se non tutte le riprendono, tra quelle che lo fanno c'è probabilmente la risposta esatta.
- ◆ Le alternative devono essere omogenee nella base concettuale della loro costruzione.

INDICAZIONI PER LA COSTRUZIONE DELLE DOMANDE TIPICHE DI UN TEST, QUALUNQUE ESSO SIA

1. Il linguaggio non sia complicato, ma adeguato agli alunni. Altrimenti dovremmo misurare anche le abilità linguistiche.
2. Gli stimoli siano brevi ed essenziali, nella misura del possibile. Si perde meno tempo e si provoca minore stanchezza.
3. Non fare tranelli. Altrimenti misureremmo l'abilità di sfuggire ai tranelli.
4. Non chiedere cose banali o sciocche. Si misurerebbero conoscenze inutili.
5. Non fare domande alle quali si possa rispondere solo in base al buon senso o alla cultura generale.
6. Le risposte sbagliate non siano ingenuie né raffinate, ma adeguate alla preparazione degli alunni. Si misurerebbero acquisizioni estranee agli obiettivi e alcuni potrebbero trovarsi in difficoltà.
7. Citare un autore, se si fanno riferimenti a quanto egli abbia detto o scritto. Non si saprebbe altrimenti riconoscere come vere o false le asserzioni riportate.

8. La collocazione delle risposte esatte non deve essere preordinata. Qualcuno potrebbe scoprirla.
9. Non utilizzare mai pezzi presi tali e quali dai libri usati per lo studio. Qualcuno li potrebbe avere memorizzati, o ritrovarvi più facilmente la risposta esatta.
10. Ciascuna domanda sia indipendente dalle altre. Chi scoprisse i collegamenti potrebbe essere facilitato.

PUNTEGGI E DETRAZIONI PER DOMANDE DI TEST DI PROFITTO

	Risposta	
	Esatta	sbagliata
Vero-falso	+ 1	- 1
Scelta multipla (n = 4)	+ 3	- 1
Scelta multipla (n = 5)	+ 4	- 1
Integrazione-Completamento-Cloze	+ 1	- 1
Corrispondenza-Associazione	+ 1	- 1

PUNTEGGI E DETRAZIONI PER SCELTE MULTIPLE CON DUE RISPOSTE ESATTE

	Risposte date			
	Esatte		sbagliate	
	1	2	1	2
con 4 alternative	2	4	- 2	- 4
con 5 alternative	3	6	- 2	- 4

COME COSTRUIRE LE PROVE OGGETTIVE?

1. Delimitare l'ambito disciplinare oggetto di verifica, le caratteristiche degli alunni oggetto della somministrazione, le specifiche funzioni valutative da attuare.
2. Identificare l'estensione del segmento curricolare o delle competenze da esaminarsi.
3. Definizione, analisi e campionamento degli obiettivi specifici della verifica.
4. Indicazione della lunghezza e/o durata del test.
5. Selezione della tipologia più opportuna degli item da utilizzare.
6. Preparazione dei quesiti e definizione della durata della prova.
7. Definizione dei punteggi grezzi da assegnarsi a ogni tipo di quesiti a seconda che si registrino risposte esatte, sbagliate, omesse. Definire i punteggi ponderati da assegnarsi a ogni tipo di item che registri risposte esatte, sbagliate, omesse.

8. Correzione e attribuzione dei punteggi.
9. Analisi e interpretazione dei risultati e analisi della qualità dello strumento impiegato.
10. Una volta elaborato lo strumento lo si somministrerà.

LA SOMMINISTRAZIONE è necessario che avvenga nel rispetto di alcune regole fondamentali:

1. Bisogna scegliere un giorno adatto e soprattutto un'ora del giorno idonea e favorevole alla concentrazione degli alunni (prime ore del mattino, assenza di interrogazioni o compiti in classe).
2. Per prima cosa, si dovrebbe dire in che cosa consiste la prestazione richiesta, spiegando per iscritto in cima alla prova che cosa la prova chiede al soggetto (rispondi alle domande segnando quella che ritieni corretta con una croce / inserisci la risposta corretta nello spazio bianco...), leggendole poi ad alta voce prima dell'inizio della prova.
3. Dovrebbe essere scritto in calce anche il tempo concesso.
4. Si dovrebbe affermare e chiarire con correttezza il motivo per cui la prova si sta svolgendo (vogliamo verificare il seguente obiettivo:... / nell'ambito dell'indagine xy si è deciso di sottoporre...).
5. Si devono motivare gli studenti al compito presentandogli le difficoltà cui andranno incontro, ma contestualmente trasferendogli un senso di fiducia nelle loro capacità (la prova è complessa in quanto..., ma sono sicuro che sarete tutti in grado di portarla a termine adeguatamente e nel tempo stabilito).
6. Si devono comunicare gli elementi necessari ad una eventuale possibilità degli allievi di decidere le meta-strategie della prova (ricordate che è prevista una penalizzazione in caso di risposta sbagliata oppure omessa...).
7. Si deve controllare che le istruzioni siano state comprese da tutti e perciò vengano rispettate.
8. Durante la prova si dovrà comunicare il tempo residuo.
9. Si deve rispondere alle richieste di aiuto senza però fornire ulteriori informazioni soprattutto di carattere contenutistico o comunque non tecnico.
10. L'insegnante si accerti che gli allievi non copino.

ATTIVARE LE PROCEDURE

Vediamo quali sono le procedure da mettere in atto una volta che gli alunni hanno completato il test di profitto:

1. riscontro tra la risposta data da ogni allievo al singolo item con la risposta corretta;
2. attribuire i punteggi per ogni item di cui è costituita la prova;
3. predisporre una griglia-tabella;
4. calcolare l'indice di discriminatività;
5. calcolare l'indice di facilità;

item 8
item 9
item 10

totali

moda:___ media:___ mediana:___ gamma:___ deviazione standard:___

Dopo aver predisposto la griglia-tabella, **le prove di tutti gli alunni vanno corrette**, segnalando anche gli errori, cioè gli item cui si sono – per esempio – fornite più soluzioni e le omissioni. **Come procedere?**

1. Un metodo potrebbe consistere nel correggere una prova alla volta, **CERCHIANDO CON UN PENNARELLO COLORATO IL NUMERO D'ORDINE DI OGNI DOMANDA CUI SI E' FORNITA UNA RISPOSTA CORRETTA**, cosicchè il numero di cerchi possa corrispondere al punteggio conseguito dall'allievo, punteggio che si può poi riportare sulla parte in alto della prova.
2. Alcuni suggeriscono di compiere l'operazione **PREDISPONENDO DELLE GRIGLIE DI CORREZIONE, CIOE' UNA COPIA DEL TEST CON LE PAGINE FORATE IN CORRISPONDENZA DELLA LETTERA CHE CORRISPONDE PER OGNI QUESITO ALLA RISPOSTA GIUSTA**. Così, sovrapponendo ogni foglio della griglia alla corrispondente pagina di ogni test da correggere e colorando con un pennarello i vari fori, si identifica immediatamente, al solo togliere del foglio, l'esattezza o l'erroneità della risposta fornita.

Dopo aver corretto tutte le prove, queste si devono **ORDINARE PER PUNTEGGIO**, dal più basso al più alto, tabulando nella griglia tutte le risposte e registrando con la lettera corrispondente le risposte errate. **NON SI DEVE TABULARE NULLA NEL CASO DI RISPOSTA CORRETTA**: gli spazi vuoti, in tal modo, corrisponderanno al punteggio ottenuto.

IL PUNTEGGIO CHE SI OTTIENE IN QUESTA FASE VIENE DETTO GREZZO. Le risposte corrette vanno contate e poi segnate sulla colonna dei **TOTALI**, alla lettera corrispondente.

Devono, inoltre, **ESSERE CONTATE E REGISTRATE ANCHE LE RISPOSTE ERRATE ORDINATE SECONDO LA LETTERA CORRISPONDENTE**.

Tutti i dati totali che attengono alla risposta errata o corretta dovrebbero essere TABULATI IN FORMATO PERCENTUALE SECONDO LA FORMULA SEGUENTE:

Percentuale = $\frac{\text{totale delle risposte della singola alternativa}}{\text{numero degli Alunni}} \times 100$.

QUESTA PRIMA FASE VA CONSIDERATA PREPARATORIA, MA STRETTAMENTE NECESSARIA. LA PROCEDURA DI ITEM ANALISI VERA E PROPRIA INIZIERA' SOLO A CONCLUSIONE DI QUESTA FASE.

Ora, ipotizziamo che i 10 studenti si siano comportati nel modo seguente:

studente	errori	punteggio grezzo
numero 5	8	6
numero 7	7	9
numero 4	7	9
numero 1	6	12
numero 2	6	12
numero 3	6	12
numero 8	4	18
numero 6	3	21
numero 9	1	27
numero 10	0	30

Inseriamo tutti i dati nella tabella.

studenti	5	7	4	1	2	3	8	6	9	10	chiavi di	A	B	C	D	err.	omiss.	facil.	discrim.
	Risposta																		
item 1	D	A	A	A	B	B	A				C	40	20	30	10	0	0		
item 2	A	D	C	A	D	C					B	20	40	20	20	0	0		
item 3	A	A	B								C	20	10	70	0	0	0		
item 4	C	D		B							A	70	10	10	10	0	0		
item 5		C	B		c						A	70	10	10	0	10	0		
item 6	A	A	B		D	D	B	B			C	20	30	30	20	0	0		
item 7	A	B	C		B	C					D	10	20	20	50	0	0		
item 8	B		C	D	B						A	60	20	10	10	0	0		
item 9		D	A		B						C	10	10	70	10	0	0		
item 10	A	o	B	o	C	B	C				D	10	20	20	30	0	20		
totali	6	9	9	12	12	12	18	21	27	30									

moda: ___ media: ___ mediana: ___ gamma: ___ deviazione standard: ___

Calcoliamo ora gli INDICI DI FACILITA' e di DISCRIMINATIVITA'

L'INDICE DI FACILITA' ci consente di controllare quanto un certo quesito sia stato di difficile soluzione per gli studenti, mettendo in relazione il numero degli studenti con il numero complessivo di risposte corrette fornite secondo la formula

$$\text{facilità} = \frac{\text{totale risposte corrette}}{\text{numero alunni}}$$

Tale indice, che varia da 0, estrema difficoltà, ad 1, estrema facilità, dovrebbe, perché un item possa considerarsi adeguato, situarsi intorno ad una percentuale inferiore allo 0,5%. In concreto: si considerano poco significativi gli item che si trovano ad avere un indice inferiore allo 0,25%, perché considerati troppo complicati rispetto al gruppo di riferimento, e quelli con un indice superiore allo 0,75%, per la ragione inversa. Calcolato l'indice, lo si riporta nello spazio che gli si è riservato nella parte destra della tabella.

Item	Indice di facilità
1	0,3
2	0,4
3	0,7
4	0,7
5	0,7
6	0,3
7	0,5
8	0,6
9	0,7
10	0,3

In interazione diretta con l'indice di facilità lavora il secondo indice da prendere in considerazione: **l'INDICE DI DISCRIMINATIVITA'**.

Esso stima la capacità che un qualsiasi item ha di essere indicativo del risultato complessivo della prova da parte dello studente.

Si raccolgono dapprima i risultati ottenuti dagli studenti in due gruppi, corrispondenti ognuno a circa un terzo del totale, e denominati estremo inferiore ed estremo superiore (i più e i meno bravi, quelli che hanno conseguito i punteggi più alti e quelli che hanno conseguito i punteggi più bassi). Si fa poi la differenza tra la somma delle risposte esatte dei due estremi e la si divide per il numero degli studenti di ogni estremo.

Il risultato ottenuto ci dà l'indice di discriminatività di ogni singolo item.

La formula della discriminatività è la seguente:

$$\text{discriminatività} = \frac{\text{risp. Esatte estremo sup.} - \text{risp. Esatte estremo inf}}{\text{numero studenti di un estremo}}$$

Questo indice può andare da -1 a $+1$. Se le risposte corrette del gruppo di studenti più abili sono in quantità maggiore, l'indice sarà positivo, raggiungendo il valore massimo di $+1$, nella fattispecie in cui tutti gli alunni dell'estremo superiore abbiano risposto correttamente e tutti quelli dell'estremo inferiore abbiano risposto in modo errato. Nel caso contrario, l'indice sarà negativo e quel singolo item andrà scartato, dal momento che inverte il criterio di abilità della prova nel suo complesso. Quando le risposte corrette dei due estremi si eguagliano l'indice risulta pari allo 0; significa che quel determinato item non risulta discriminativo, cioè non è in grado di distinguere l'abilità misurata nella prova. Si devono considerare, nella pratica concreta, efficaci gli item che hanno un indice che si trova tra lo 0,30% e lo 0,60%.

Nel nostro caso abbiamo preso in considerazione due gruppi composti da tre elementi: il gruppo di alunni che ha conseguito i punteggi 6,9,9 e il gruppo di alunni che ha conseguito i punteggi di 21,27,30.

Da questo dato si parte per calcolare su ogni item l'indice di discriminatività. Il risultato è il seguente:

Item	Indice di discriminatività
1	0,66
2	1
3	0,66
4	0,66
5	0,33
6	0,33
7	1
8	0,33
9	0,33
10	0,66

In sintesi: POSSONO RITENERSI ACCETTABILI GLI INDICI DI DIFFICOLTA' E DI DISCRIMINATIVITA' COMPRESI FRA 0,3 E 0,6.

Consideriamo ora l'**INDICE DI ATTENDIBILITA'**. Esso fa riferimento alla coerenza o fedeltà dei punteggi ottenuti da uno stesso alunno quando questi venga sottoposto allo stesso test in occasioni diverse, o ad un insieme di prove equivalenti, o in diverse condizioni di somministrazione.

L'attendibilità di un test si esprime attraverso il **COEFFICIENTE DI CORRELAZIONE (r)** che dà la misura della relazione fra due gruppi di punteggi. La misura più usata è quella nota come **COEFFICIENTE DI CORRELAZIONE DI PEARSON**. Questo può essere di segno positivo o negativo e i valori vanno da -1 a $+1$. Più il coefficiente si avvicina al valore positivo di 1, maggiore risulta l'attendibilità di un test.

Per quanto concerne la formula per calcolare il coefficiente di correlazione, la più breve è la seguente:

$r = \frac{\text{la somma dei punti z del primo test con quelli del secondo test}}{\text{numero degli alunni}}$

La seconda caratteristica fondamentale di un test è la validità, cioè la capacità che ha il test di misurare effettivamente ciò che noi intendiamo misurare.

Circa la misurazione della validità, è possibile fare riferimento **all'ERRORE STANDARD DI PREVISIONE**, che indica il margine di errore che è ragionevole attendersi nel punteggio previsto del criterio per un determinato alunno, come risultato della imperfetta validità del test. La formula è la seguente:

$D.S.\text{prev.} = \text{deviazione standard dei punteggi del criterio} \times \text{la radice quadrata di } 1 - \text{il coefficiente di validità al quadrato}$

INDICI DI CENTRALITA'

Gli elementi minimi che abbiamo finora indicato, se possono darci indicazioni sull'organizzazione e strutturazione del test, non ci dicono nulla sull'ipotetico gruppo di alunni che abbiamo testato.

Per raccogliere qualche dato in questa direzione dobbiamo ricavare altri punteggi. Parleremo perciò ora delle cosiddette MISURE DI TENDENZA CENTRALE di una distribuzione: MODA, MEDIA, MEDIANA.

MODA. E' il punteggio ottenuto che ricorre con maggiore frequenza. Nel nostro test è il numero 12. Se il numero di moda è centrale rispetto ai punteggi estremi, ciò significa che la prova è adeguata; una moda bassa sta a indicare che la prova è stata difficile, un valore alto una eccessiva facilità. La moda è quindi indicativa di una tendenza generale.

Utilizzata come parametro in una misurazione con scala a intervalli, indica il valore o punteggio che compare più volte nella graduatoria. Rimane una misura poco utile e poco affidabile.

MEDIANA. E' il punteggio conseguito dallo studente che si trova al centro della distribuzione dei punteggi. Nel nostro caso, dal momento che il nostro gruppo risulta composto da 10 elementi, risulta opportuno considerare due studenti per ottenere il numero di mediana: il 5 e il 6. Il nostro numero di mediana risulta essere 12. La misura mediana è importante perché divide il gruppo testato in due parti, una metà alta e una metà bassa. Applicata alle misurazioni per scala a intervalli, se i dati sono dispari di numero, sarà mediana il numero che occupa il posto centrale; in questo caso la formula per trovare la mediana è

Mediana = numero dei dati più 1/2

Se invece i dati sono di numero pari, la mediana si calcola dividendo per 2 la somma dei due numeri centrali.

Anche in questo caso l'uso per una misurazione scolastica è limitato perché le informazioni che si ricavano non danno indicazioni apprezzabili.

La misura della mediana può essere di estrema utilità soprattutto se messa in rapporto con il punteggio della **MEDIA** che è il punteggio medio ottenuto dal gruppo testato. Questa misura si ottiene dividendo la somma di tutti i punteggi per il numero degli allievi. La formula è la seguente:

$$\text{media} = \text{somma dei punteggi} / \text{numero degli alunni}$$

Nel nostro test la moda è 12, la mediana è 12, la media è 15,6.

Facciamo una precisazione: la prova di un singolo alunno acquista una sua precisa collocazione se confrontata con i risultati delle prestazioni di tutti i suoi compagni. Ciò significa che **OGNI PUNTEGGIO ATTRIBUITO AL SINGOLO ALUNNO DIVIENE LEGGIBILE SOLO SE SI TIENE CONTO DI ALTRE INFORMAZIONI, E CIOE' SE SI CONOSCE:**

- ◆ il punteggio massimo teoricamente realizzabile;
- ◆ il massimo dei punteggi effettivamente realizzati;
- ◆ il rendimento medio del gruppo, cioè la media aritmetica;
- ◆ la dispersione dei punteggi ottenuti dagli alunni attorno al punteggio medio, cioè la GAMMA.

In generale si può sostenere che SE LA MEDIA E' PIU' BASSA DELLA MEDIANA LA PROVA, NEL SUO COMPLESSO, NON APPARE SIGNIFICATIVA PERCHE' TROPPO FACILE; NEL CASO OPPOSTO, ESSA E' TROPPO DIFFICILE. QUANDO LA MEDIA E LA MEDIANA COINCIDONO O SONO ASSAI PROSSIMALI, SI PUO' CONSIDERARE LA PROVA COME ADEGUATA.

INDICI DI DISPERSIONE

Gli strumenti proposti fino ad ora sono piuttosto autoreferenziali, cioè non fanno altro che prendere in considerazione il gruppo oggetto del test, non fornendoci elementi significativi di relazione tra essi e altri eventuali risultati misurativi conseguiti su altri gruppi; oltre a ciò, si deve dire che quanto esposto finora non ci fornisce alcun elemento di carattere diacronico rispetto al gruppo preso in considerazione. Ciò non può soddisfare chi deve poter assistere e pilotare l'evoluzione di un gruppo verso i risultati progettati all'inizio dell'anno scolastico in riferimento alla programmazione o alla realizzazione di un modulo.

La media, la mediana e la moda vedono riposare la loro funzionalità sui piccoli numeri, la singola prova, i piccoli gruppi.

Elementi di maggior interesse possono venirci dagli **INDICI DI DISPERSIONE**, ad esempio dalla **DEVIAZIONE STANDARD**, elemento che ci dice quanto i punteggi ottenuti sono dispersi rispetto alla media, qual è il livello di omogeneità di un gruppo. Questo indice si fonda sul principio di considerare gli scarti dei singoli punteggi ottenuti dagli studenti rispetto alla media. La formula per ottenere la deviazione standard è la seguente:

$\sigma = \text{radice quadrata della somma delle differenze dei punteggi dalla media} / \text{il numero delle prove, il tutto elevato al quadrato}$

Nel nostro test la deviazione standard è 7,6.

IN GENERALE: QUANTO PIU' LA DEVIAZIONE STANDARD SI AVVICINA ALLO 0, TANTO PIU' LA PROVA HA AVUTO UN RISULTATO OMOGENEO. SE LA DEVIAZIONE STANDARD E' ALTA (COME NEL NOSTRO TEST), CIO' SIGNIFICA CHE IL GRUPPO RISULTA ETEROGENEO, CON UNA NOTEVOLE DISPERSIONE DI PUNTEGGI RISPETTO ALLA MEDIA.

Esiste una misura ulteriore in grado di fornirci informazioni sull'omogeneità di un gruppo. Si tratta del **COEFFICIENTE DI VARIAZIONE**, ottenibile con un semplice calcolo che è il seguente:

$\text{c.d.v.} = \sigma / \text{media} \times 100$

Questa misura ci dà informazioni sulla percentuale di deviazione standard rispetto alla media di un punteggio. Quando il valore del coefficiente di variazione supera 20 siamo in presenza di una forte eterogeneità.

PUNTI STANDARDIZZATI

Il legame strutturale che esiste tra il punteggio e la prova rende difficile confrontare e interpretare l'andamento didattico di un alunno nel corso del tempo rispetto a due prove identiche per struttura, somministrate in momenti diversi, dal momento che, in via ipotetica, potrebbero cambiare la media, la deviazione standard e, ancor più, rispetto a prove somministrate in tempi diversi e lontane per struttura e natura dei quesiti che le costituiscono.

Un sistema per superare questo problema consiste nel **TRASFORMARE I PUNTEGGI GREZZI** (che, naturalmente, non possono essere confrontati tra loro per le ragioni già esposte) **IN PUNTI STANDARD**. Le fondamentali tipologie di punti standard utilizzabili in chiave didattica sono i **PUNTI Z** e i **PUNTI T**.

I punti Z si ottengono misurando in deviazioni standard la differenza tra il punteggio grezzo ottenuto e la media della distribuzione secondo la formula

$$Z = \text{punteggio grezzo} - \text{punteggio medio}/\text{deviazione standard}$$

I punti Z si riferiscono alla media e il loro incremento risulta determinato dalla deviazione standard. Perciò, in concreto, i punti Z possono variare da un minimo di -3 a un massimo di $+3$. Da tenere presente il fatto che i punti Z sono negativi quando i punteggi grezzi sono inferiori alla media e positivi quando sono superiori.

Quando $Z = 0$, ciò vuol dire che il punteggio è esattamente uguale alla media.

Un ulteriore aspetto significativo riguarda il fatto che attribuire ad un allievo un punteggio standardizzato Z , dal momento che esso si concorda con la deviazione standard, consente anche di inquadrarlo nella fascia di livello di appartenenza.

Esiste, comunque, un limite di fruibilità dei punti Z ; questo limite consiste nel fatto che essi assumono anche valori negativi. Tali valori possono rispecchiare anche situazioni abbastanza soddisfacenti se si considera che, ad esempio, la media è lo 0.

Per far fronte a tale problema e a tutto ciò che ne consegue in termini di comunicazione con un alunno e di conseguente motivazione, si possono rappresentare gli esiti di un confronto tra due prove anche con i **PUNTI T**. Attraverso i punti T lo 0 dei punti Z viene in pratica spostato a 50 attraverso la formula seguente:

$$T = Z \times 10 + 50$$

E' chiaro che tutti i valori al di sotto di 50 corrispondono a punti Z negativi.

SVILUPPO DELLE ATTIVITA' VALUTATIVE

Osservazione

Rilevazione

Prestazioni
Comportamenti
Attività

Processi
Atteggiamenti
Stili

Misurazione

Dati
quantitativi

Descrizione

Dati
qualitativi

Traguardi previsti

Percorsi individuali

Capacità, Abilità

Soglie di padronanza

Verifica

Interpretazione

Risultati
Livelli

Informazioni
significative

Valutazione