COME FORMULARE GIUDIZI DI VALUTAZIONE DIDATTICA
La costruzione di scale descrittive

L'esigenza di validità, oggettività e precisione è sempre più pressante nel contesto educativo. In questo libro gli autori propongono modalità indispensabili per meglio operare in classe una valutazione continua, analitica, centrata sull'apprendimento.
EDUCAZIONE NUOVA

Collana diretta da Giuseppe Lisciani

J. Beckers, R. De Bal, G. De Landsheere
Come formulare giudizi di valutazione didattica
COME FORMULARE GIUDIZI DI VALUTAZIONE DIDATTICA
La costruzione di scale descrittive

Giunti & Lisciani Editori
Premessa

Benedetto Vertecchi

Le norme sulla valutazione degli allievi introdotte dalla legge n. 517 del 1977 non sono più una novità, né per gli insegnanti, né per gli allievi e per le loro famiglie. Da una parte e dall'altra, dopo un primo momento di incertezza, a volte di sbandoamento, si è constatato che la normativa introdotta poteva essere osservata in modi non traumatici. In effetti, chi si attendeva un mutamento radicale nei modi della valutazione per effetto della sostituzione dei voti numerici con giudizi espressi verbalmente è stato disilluso. Non occorre una particolare competenza per sibilare una semplice tavola di equivalenze tra voti e giudizi verbali; chi è appena un po' esercitato nella lettura delle schede di valutazione riesce non solo a tradurre i giudizi in voti, ma anche a cogliere quelle sfumature che un tempo corrispondevano all'aggiunta di mezzi punti, o a correggere il voto facendo seguire da un "più" o da un "meno". Dobbiamo dunque concludere che nulla è cambiato? Hanno ragione quanti sostengono che formulare giudizi verbali è solo una perdita di tempo, e che i voti sono più adatti ad esprimere valutazioni perché meglio e più rapidamente comprensibili?

Per non giungere a conclusioni affrettate è bene soffermarsi su alcuni aspetti del problema valutativo che sono stati alla base del dibattito che ha condotto alla revisione delle norme operata con la legge n. 517. Innanzi tutto, si trattava di individuare modalità valutative che fossero coerenti con i traguardi propri della formazione di base. Poiché tale formazione non ha intenti selettivi, ma costituisce un diritto-dovere per tutti, non aveva significato ricorrere ad una modalità valutativa attraverso la quale venivano esaltate le differenze tra gli allievi, invece di tendere a consentire a ciascuno di acquisire quel
corredo di capacità e di conoscenze che si considera necessario per una partecipazione consapevole alla vita sociale. Su questa scelta a carattere generale si inserivano considerazioni più specifiche. Sotto l'apparente neutralità del voto si sovrapponevano due differenti giudizi, uno relativo all'apprendimento, l'altro collegabile all'approccio di aspetti socio-affettivi e morali del comportamento degli allievi. In altre parole, un voto positivo corrispondeva non solo alla constatazione di un livello desiderabile delle prestazioni di un allievo, ma anche all'apprezzamento di tratti della sua personalità, come l'impiego, il comportamento di relazione ecc. Nello stesso modo, un "+" in italiano significava sia non sa scrivere correttamente" sia "non si impara nelle attività che gli vengono proposte.

L'ambiguità dei voti si è accresciuta attraverso il tempo in corrispondenza col crescere del sistema scolastico. La parte socio-affettiva e morale del voto poteva infatti considerarsi neutralizzata dall'esistenza di componenti dinamiche sostanzialmente omogenee negli allievi e condivise dalle famiglie: in altre parole, in presenza di un quadro scolastico a base sociale ristretta, l'essere inseriti in un processo formale di istruzione comportava una distinzione di status tale da far accettare il compito di apprendimento. Gli allievi che mostravano di non aderire al modello prevalente di valori potevano essere rapidamente esclusi dalla scuola. Ma una volta che il principio del diritto-dovere all'istruzione sia stato adottato in pratica, e quindi che la totalità o quasi della popolazione in una certa fascia d'età sia presente dentro la scuola, le componenti dinamiche menzionate cessano di costituire un riferimento comune. Se andare a scuola costituisce un aspetto caratterizzante dell'esperienza dei bambini e dei ragazzi, non ci si può attendere che essi attribuiscono ad esso un particolare valore affettivo. L'accettazione del compito di apprendimento non è quindi più un presupposto del lavoro didattico, ma costituisce un risultato da conseguirsi attraverso di esso. I voti sono stati spesso considerati un espediente per incentivare l'affettività, attribuendo ad essi valore di premio o castigo, di stimolo all'impegno o di determinare per la sua mancanza o limitatezza.

D'altrò canto lo sviluppo della ricerca didattica e le esperienze di innovazione sottolineavano l'importanza che assume nel processo di formazione la pratica valutativa. Ad una espressione soltanto terminale dei giudizi, tesa a evidenziare in termini comparativi i livelli di apprendimento raggiunti dagli allievi, si contrapponeva una strategia più articolata, per la quale, senza disconoscere la rilevanza dei giudizi finali, si affermava la necessità di disporre con continuità di elementi di informazione circa il modo in cui gli allievi accedono all'attività di formazione e procedono attraverso di essa. La nuova esigenza poneva problemi di revisione dei fondamenti metrici della valutazione. Un esempio chiarisce il problema: se l'allievo A ottiene "6", l'allievo B "7" e l'allievo C "5", in una prova di matematica, sarebbe del tutto improprio interpretare tali voti affermando che l'allievo B ha raggiunto una buona conoscenza del calcolo algebrico, l'allievo A una conoscenza appena sufficiente e l'allievo C una scarsa. In realtà posso solo affermare che la prestazione fornita dall'allievo B è complessivamente migliore di quella fornita dall'allievo A, che a sua volta è migliore di quella fornita dall'allievo C. Se la prova fosse consistita nell'impostare e risolvere una equazione di primo grado, non sarei in condizione di affermare sulla base dei voti ottenuti dagli allievi quale dei tre abbia raggiunto gli obiettivi che ci si attendevano, in termini di capacità di usare simboli, di effettuare operazioni algebriche, di tradurre un problema formulato verbalmente in una equazione ecc. Eppure, da un punto di vista didattico, sono proprio queste le informazioni che interessano. Se si deve intervenire per indirizzare l'attività di apprendimento, c'è bisogno di sapere che cosa ciascun allievo ha già appreso e che cosa deve ancora apprendere, mentre non ha rilevanza sapere che un allievo ha appreso meglio di un altro. In breve, tutto ciò presuppone un cambiamento radicale nella valutazione, per il quale ad apprezzamenti d'insieme si sostituiscono, o almeno si affianchino, descrizioni accurate dei risultati che si ottengono nelle prove di apprendimento, e descrizioni altrettanto accurate dell'insieme degli elementi che emergono nel corso delle attività didattiche, e che sono in grado di incidere su di esse.

Il cambiamento delle norme relative alla valutazione avrebbe dovuto consentire di affrontare in modo più consapevole i problemi posti dal rinnovamento didattico. Non a caso la stessa legge che ha innovato in materia di valutazione ha anche stabilito che le scuole dovessero programmare la loro attività, operando scelte sensibili alle esigenze degli allievi. Va allora rivisto il senso delle domande che ci ponevano all'inizio.

Dobbiamo chiederci: perché non si è avuto uno sviluppo nella direzione che il dibattito precedente all'approvazione della legge n. 517 lasciava intravedere? Perché si è avuto un accomodamento per riduzione al consueto della pratica valutativa? La risposta è per certi versi semplice, e per altri complessa. È semplice rispondere che un cambiamento effettivo delle modalità di valutazione non richiede solo consapevolezza di questioni generali, ma una precisa competenza tecnica. È più complessa l'articolazione della seconda parte della risposta, perché dovremmo collegare molti fattori, che vanno dai riti della cultura pedagogico-didattica nel nostro paese alla mancanza di una politica di interventi per il sostegno dell'attività professionale degli insegnanti, alla rilevanza dell'inadeguatezza dei modelli di divisione del lavoro all'interno delle scuole.

Può essere interessante notare che certe difficoltà non sono esclusive della scuola italiana, e che l'innovazione introdotta con la legge del 1977 corrisponde a linee di intervento che si riscontrano anche in altri paesi. La lettura di questo libro offre un quadro puntuale dei
problemi che il rinnovamento delle modalità di valutazione ha posto in Belgio. Il lettore non avrà difficoltà a constatare che si tratta di problemi assai simili a quelli che hanno dovuto affrontare gli insegnanti italiani. Ma se i problemi sono simili, almeno in parte possono essere anche le soluzioni: quanto propone De Landsheere può costituire la base per una revisione della pratica valutativa che rappresenti una interpretazione impegnativa della legge n. 517.

Introduzione

Il rinnovamento dell’insegnamento secondario in Belgio, presto esteso all’insegnamento primario, è stato accompagnato immediatamente da una riforma dei metodi di valutazione. I principali obiettivi di una tale riforma sono i seguenti:

— concedere maggiore importanza alla funzione educativa della valutazione, utilizzata troppo a lungo nel passato per fini essenzialmente selettivi;
— allargare sistematicamente lo spettro valutativo, non solo per coprire tutto il settore cognitivo, ma anche per dare spazio al settore affettivo;
— introdurre nella vita scolastica metodi e tecniche di valutazione che offrano maggiori garanzie scientifiche rispetto ai procedimenti usuali.

Un rinnovamento così profondo del sistema educativo non può verificarsi all’improvviso. Da prima, riguarda decine di migliaia di insegnanti e centinaia di migliaia di alunni. Poi, chiama in causa le abitudini, gli atteggiamenti e le convinzioni acquisite dagli insegnanti durante i loro studi e rafforzate e approfondite in una carriera talvolta lunga; d’altra parte la riforma richiede un ampio sforzo di aggiornamento che si prolungherà nel tempo. Infine, non è sufficiente decidere di cambiare metodi e strumenti di valutazione; è necessario che i nuovi strumenti esistano e che i destinatari li accettino e sappiano servirsene correttamente. Diciamo subito che, per soddisfare queste due condizioni, gli insegnanti devono partecipare direttamente alle decisioni di principio e alla costruzione di nuovi strumenti.

Certamente la valutazione non si riduce ad un problema di strumenti. Tutta l’intelligenza e la sensibilità dell’educatore intervin- gono nei giudizi e negli apprezzamenti espressi nel corso delle molteplici interazioni tra l’insegnante e i suoi alunni. Per quanto bana-le possa sembrare la constatazione, il ricordare che la pedagogia è arte quanto scienza, si impone allorquando si tratta di valutazione in materia di educazione.

Quali sono allora i metodi fondamentali della valutazione? Essa è analitica o globale, matematica o soggettiva, matematica o affetti-
va. I due aspetti non si escludono necessariamente: al contrario si completano felicemente. Si può contare il numero degli errori di ortografia commessi da un alunno e poi formulare un giudizio relativo (performance) migliore della precedente, buona qualità della performance, in considerazione dell'età del soggetto...).

Nell'esempio appena proposto, il calcolo di un punteggio (score) è seguito da una valutazione (appréciation) globale. Al contrario, una performance valutata dapprima globalmente, può in seguito vedersi attribuita una valutazione numerica per stabilire una gerarchia tra persone o tra produzioni successive di uno stesso individuo.

Questa osservazione non ci ha condotto al nucleo del problema. Evitiamo gli esempi scolastici, che potrebbero risultare condizionati emotivamente: soffermiamoci un istante nel campo dello sport e confrontiamo il modo di valutare una corsa di velocità con pattini a rotelle con una prova di pattinaggio artistico. Entrambi possono comportare valutazioni affettive, ma soltanto la corsa si presta ad una misurazione esatta, oggettiva. È questo per un motivo: nella corsa risulta importante un solo criterio, valutabile con il semplice conteggio, il tempo. Si conta (con la precisione di cui sappiamo) il numero dei secondi, dei minuti... In tal caso si misura una sola modalità del comportamento. Per facilità di comprensione, si prenda come esempio una performance semplice (il che non significa che è facile o che costituiscia il risultato di un processo privo di complessità).

Al contrario, il pattinaggio artistico non può essere misurato in questo modo. Si possono immaginare dispositivi per la valutazione del ritmo, il ricorso ad un inventario di comportamenti per giudicare le originalità delle figure: ma cosa accadrà quando si dovranno valutare gli aspetti più o meno graziosi delle evoluzioni, la bellezza della performance, la fluidità dei passaggi? I giudici sportivi separano, saggiantemente, gli aspetti tecnici (suscettibili di valutazione, con il conteggio dei punti) dagli aspetti artistici che, in ultima analisi, diventano globali. In questo caso si tiene presente un insieme di modalità e la performance viene definita complessa.

La complessità varia naturalmente a seconda dei casi. Ma non bastano i molti fattori suscettibili di interazioni per pervenire a tutti quelli non riducibili: l'insieme supera i comportamenti isolati (?).

Torniamo ora su questi argomenti, per altre vie, nel contesto del presente lavoro, ma è sembrato opportuno chiarire subito il problema una prima volta: la scuola, al fine di portare a termine il suo compito educativo, ha bisogno di strumenti di misura dei comportamenti semplici (tests oggettivi) e di scale di valutazione per i comportamenti complessi.

Ebbene, nel momento in cui il Belgio ha stabilito di rinnovare il suo insegnamento fondamentale, le scuole disponevano di pochi test di conoscenza e si può ritenere che praticamente non esistessero scale di valutazione abbastanza perfezionate. Eppure, si avverte un'urgenza necessità di test di previsioni e di analisi, e, soprattutto, di scale di valutazione da usare quasi quotidianamente.

Si sa quanto è accaduto: i test sono sempre pressoché integralmente carenti, è stato imposto l'uso di semplici scale verbalizzate a cinque o sette gradi. Sarebbe ingiusto accusare le autorità scolastiche di incapacità per non aver atteso la disponibilità di strumenti necessari prima di iniziare la riforma. In una crisi generale come quella che attraversiamo, alcuni cambiamenti istituzionali non possono essere differiti nel tempo e, anche se è interessante tentare di comprendere il perché, fino ad ora, la formazione degli insegnanti nel nostro paese ha riservato un posto limitato alla pedagogia scientifica, ciò non toglie che uno dei problemi più urgenti consista nel costruire quanto ci manca.

Ben presto si è avvertito che l'adozione di scale verbalizzate di valutazione a cinque o sette gradi, come quelle proposte, provocava, in linea generale, un'utile riflessione tra gli insegnanti, ma non era ancora lontani da una rigorosa valutazione. Per comprendere meglio la situazione, fin dal 1971 abbiamo soppotto ad una inchiesta tutti gli insegnanti di lingua materna e di matematica della rinnovata scuola secondaria. Ne è risultato al tempo stesso quanto grande fosse la disponibilità degli insegnanti interessati e quanto scarsi fossero i progressi apportati dalle nuove scale di rapporto al sistema tradizionale.

Evidentemente non è sufficiente sostituire il voto 8/10 con la nota A, per promuovere un maggior rigore nella valutazione! L'importante, come mostreremo nel presente lavoro, è definire, in maniera sia univoca che possibile, le condizioni in cui tale nota di giudizio verrà attribuita.

Così si è manifestata l'urgenza necessità di perfezionare il sistema inizialmente proposto dai riformatori: i quali — rendiamo loro giustizia — sono stati i primi ad intendere la loro proposta: il che comporta fortunatamente un cambiamento rispetto al vecchio rapporto autoritario (ch c decide e chi esegue).

Su loro richiesta, abbiamo provato a rispondere a molti quesiti:

- è possibile costruire scale di valutazione, se non perfette, che garantiscono un'elevata precisione?
- è possibile rilevare esercizi tipo frequenti nella pratica scolastica e costruire, per valutarli, scale suscettibili di essere utilizzate regolarmente o facilmente limitate?
- che lavoro comporta la costruzione di una batteria di scale ritenute soddisfacenti?
Occorre intraprendere una specifica azione perché le scale costruite siano utilizzate correttamente o è sufficiente metterle a disposizione delle scuole?

La struttura di questo libro prevede, dopo le nozioni essenziali di docimologia, la presentazione delle principali conclusioni derivate dall'inchiesta preliminare. Si passa poi alla costruzione di scale di valutazione, spiegandone inizialmente la teoria.

Sono state costruite batterie di scale per la lingua materna (esercizio di espressione scritta e riassunto del testo, seguito dall'analisi critica) e le scienze naturali (resoconto di un'esperienza, metodo di misurazione, ...).

I criteri di valutazione sono stati scelti in riferimento ai risultati di ricerche sperimentali, laddove esistevano, e su suggerimento di insegnanti esperti. Così, con la cooperazione di commissioni composte da insegnanti e da ispettori, è stata realizzata la costruzione e gli strumenti sono stati sperimentati sul campo.

La ricerca, di cui stiamo per leggere il resoconto ed i risultati, intende dimostrare quale grado di precisione può essere raggiunto nel valutare le performance complesse, mettendo in atto tutte le conoscenze e gli strumenti disponibili (cosa che, evidentemente, non è alla portata dell'insegnante impegnato, quotidianamente, nel suo mestiere di educatore). Procedendo in tal modo si è inteso mostrare anche i limiti del sistema. La garanzia di una maggior precisione e giustizia nella valutazione delle performances complesse può condurre alla completa unanimità di giudizio soltanto in via eccezionale. Perché non esistono bilance che possano consentire giudizi di valore quando si tratta di valutare la bellezza o il bene.

Ritenuti validi ed accettati questi limiti, ne derivano quattro conclusioni supplementari:

1. Si possono costruire scale precise di valutazione per esercizi ritenuti difficilmente valutabili in maniera oggettiva;

2. La precisione delle scale porta ad una elevata concordanza di giudizi tra gli insegnanti che valutano in maniera autonoma;

3. Tuttavia, per giungere ad un accordo, gli insegnanti devono accettare talvolta, se non spesso, di rivedere la propria scala implicita di valutazione, seguendo le definizioni operative proposte. Si devono radicare nuove abitudini mentali, specialmente per rispettare in modo rigoroso le disposizioni valutative standardizzate. Una prima esperienza in campo scolastico ha mostrato, in effetti, che senza questo cambiamento individuale, le deformazioni e le omissioni riducono gran parte dei vantaggi legati al sistema rigorosamente costruito. Intendiamo che sia stato proposto l'impiego graduale delle scale valutative, le cui modalità d'impiego sono state in qualche misura programmate, per ridurre al massimo le difficoltà;

4. La messa a punto delle scale proposte in questo volume ha richiesto un tempo considerevole. Molto più di quello che singoli insegnanti, per quanto zelanti fossero, ne potrebbero investire in una attività che costituisce per essi soltanto una parte del loro compito.

In tal modo si conferma, ulteriormente, la necessità di instaurare una collaborazione permanente tra i centri di ricerca pedagogica e gruppi di insegnanti.

Se tentativi come questo potranno continuare, finiremo per disporre di un'ampla scelta di scale, studiate per i diversi settori e per i principali tipi di compiti che le richiedono. Intanto si istaureranno nuove abitudini valutative da parte degli educatori e degli alunni (in effetti, l'autovalutazione, può seguire strade simili a queste). Si può pertanto sperare che, fra qualche anno, quando oggi è lavoro pesante divenga normale attività.

Sappiamo bene che gli strumenti proposti possono sembrare complicati per gli insegnanti che li utilizzarono per la prima volta. A costoro chiediamo di provare pazientemente le scale, di discutere con i colleghi, di procedere, in piccoli gruppi, ad esperienze di valutazione comparativa per alcuni lavori. Costoro renderanno un gran servizio a colleghi e ricercatori informandoli delle loro critiche, dei risultati dei loro tentativi, dei loro suggerimenti, sia per correggere gli strumenti proposti che per costruirne altri.

NOTE ALL'INTRODUZIONE

(1) Si preferisce lasciare il termine "performance", in quanto entrato ormai abitualmente nella nostra letteratura pedagogica. Del resto, anche quello di "prestazione", che potrebbe costituirne la traduzione più fedele, sarebbe riduttivo, o comunque insoddisfacente, in taluni contesti (n.d.c.).

(2) Si osserverà pure che, per misurare il tempo di una corsa, è sufficiente un solo cronometro elettronico, mentre il punteggio artistico, in un campionato di pattinaggio, è la risultante della media di undici valutazioni formulate simultaneamente.
PARTE PRIMA
INDAGINE SULLA VALUTAZIONE

1. Per una valutazione formativa
2. Risultati di un'indagine
Capitolo primo

Per una valutazione formativa

Premessa

Una valutazione, che intenda giocare efficacemente il proprio ruolo, deve essere innanzitutto valida, deve cioè misurare realmente quel che pretende di misurare. Una simile affermazione può sembrare banale, se non ingenua. Eppure, nella pratica scolastica, come nella vita quotidiana, quante volte si confonde la parte col tutto (ad esempio, si valuta la memoria pensando di valutare l’intelligenza) o si misura “a vanvera” (ad esempio, si considerano indiscutibilmente equivalenti il contenuto di un programma scolastico e le esigenze della vita pratica)? Del resto, lo stesso concetto di validità non è assoluto, ma relativo: esso è inscindibile dalla finalità perseguita.

Anche quando non si commettono errori circa l’oggetto della valutazione, quest’ultima può talvolta variare sotto il profilo qualitativo: l’ideale (quasi mai conseguito) è quello di raggiungere una precisione e un’oggettività assolute.

Infine, in un contesto educativo, la valutazione deve presentare caratteristiche tali da essere utilizzate nel modo migliore da parte degli insegnanti e alunni nell’esercizio dei rispettivi compiti.

Le scale di valutazione ben costruite devono rispondere a tutte queste esigenze primarie. Vedremo come si è provato a soddisfarle.

La valutazione deve essere valida

Al di là delle molteplici sfumature, si distinguono tre tipi di validità: predittiva, di costrutto e di contenuto.
Supponiamo di poter predire le possibilità di riuscita di un adolescente nell’insegnamento secondario generale, partendo dalla combinazione dei punteggi (scores) ottenuti in un gioco di parole crociate, in un gioco di ragionamento astratto e in un puzzle difficile. Evidentemente misuriamo comportamenti cognitivi e affettivi che non sappiamo precisamente quali. Talvolta conta poco per noi che le predizioni si realizzino in misura soddisfacente. In casi simili si dice che la prova possiede una buona validità predittiva.

Questo modo di procedere privilegia l’efficienza rispetto al rigo-re scientifico, infatti si tralascia ogni sforzo specifico per analizza-re le processi e scoprire le cause. È noto, tuttavia, che tanto più si co-Ma, nelle scienze umane, l’ignoranza in settori fondamentali è ben più estesa del sapere.

Accade il contrario quando ci si preoccupa di conoscere ciò che si misura. Di rado in materia di educazione, la qualità o l’oggetto cui siamo interessati vengono valutati direttamente. Ciò accade, in-vece, quando si misura, ad esempio, la velocità di lettura, la ric-la valutazione cade sui tratti psicologici (attenzione, interessi, atteg-giamenti, ...) o sui attitudini (intelligenza verbale, creatività, ...). Se muni ad un numero più o meno grande di oggetti che abbiamo di- per i tratti psicologici o per le attitudini.


Per Thordike e Hagen la validità di costrutto di ununo strumento di misurazione è solida se da essa derivano informazioni significative sugli individui misurati (1). Va da sé che la validità di costrutto dell’entità considerata: in teoria, un test di intelligenza verbale deve misurare tutti gli aspetti di questo tipo di intelligenza.

Prima di preoccuparsi di valutazione esasativa, l’insegnante do-vrebbe preoccuparsi di educazione globale, nel senso che, almeno in teoria, sarebbe auspicabile che si prendesse cura di tutte le attitu-dini dell’alluno. Sorge però un duplice ostacolo: il primo luogo, le attitudini di un individuo normale sono talmente numerose da non poter essere sviluppate tutte in maniera omogenea; secondaria-mente, molte attitudini, cui ci dedichiamo in modo particolare, si sviluppano molto lentamente. La scuola oggi agisce pertanto sag-giamente, interessandosi più delle attitudini degli alunni e conse-pendo la sua azione a lungo termine, invece di investire il meglio delle proprie energie in conoscenze momentanee.

La validità di contenuto è il terzo approccio delle realtà educati-ve e per lungo tempo è stata concepita esclusivamente in termini di materie di insegnamento.

Ad esempio, si riteneva che un insegnamento o una valutazione soddisfassero alle esigenze sulla validità di contenuto se tutti i punti importanti del programma o del piano di studi ufficiale venivano applicati. Per costruire un test di aritmetica, che controlli se gli alunni hanno ben appreso quanto previsto dal programma, biso-gna procedere ad una catalogazione degli argomenti menzionati nei documenti ufficiali e non chiedersi cosa significhi essere ben educa-to nel settore dell’aritmetica (interrogativo che rientra nella validi-tà di costrutto).

E questo un modo ristretto di concepire la validità di contenuto. Gli argomenti non sono che materiali inerti; prendono significato soltanto se ne conseguono comportamenti tali da renderli operanti. La validità di contenuto implica anche un inventario di comportamenti da apprendere e da perfezionare. Tassonomie come quella di Bloom (2) sono utili all’educazione proprio in questo settore: invitano ad un superamento cognitivo o affettivo continuo e mettono in guardia permanentemente contro la stagnazione dei fatti a livello di semplice memorizzazione e contro la passività.

Riguardo alla validità di contenuto si pone anche il problema della corrispondenza tra insegnamento e valutazione. Certamente si valutano le conoscenze reali o possibili di un alunno senza preco-cuparsi delle fonti del suo sapere o delle sue potenzialità. Ma in una situazione di insegnamento è bene sapere quanto è stata efficace l’azione dell’educatore ... e non è ammissibile punire un alunno incapace di eseguire compiti cui non è stato preparato. In breve, e a meno che non si decida esplicitamente in altro modo, deve esistere un parallelismo rigoroso tra i contenuti dell’insegnamento e quelli della valutazione. Non si può, ad esempio, rimpiangere la memoria degli alunni durante la spiegazione e porre loro domande di concetto nella valutazione: il che significherebbe invitare l’usare ad un livello cognitivo in cui non hanno imparato a muoversi. Comportamen-to complessi richiedono maggiori incomprensioni educative. Pre-tendere che gli apprendimenti più elevati richiedano processi talmente sottili da non poter essere coltivati organicamente nè valuta-ti oggettivamente significa, semplicemente, eludere le difficoltà... e le proprie responsabilità di educatore.

Il necessario parallelismo tra insegnamento e valutazione com-porta una partecipazione indispensabile degli insegnanti alla co-

18
La valutazione deve essere oggettiva

Una valutazione è ritenuta oggettiva se fornisce la misura imparziale e fedele di un oggetto. Nelle scienze umane non si consegna mai un elevato livello di obiettività, per la complessità degli oggetti valutati e per l'inelastica interazione che si genera tra valutatore e soggetto valutato. Così si ricorre spesso ad un sottotipo: non accordo tra esaminatori concordanza sulla qualità della misurazione, ci si sente delle tuite da uno stesso individuo fedeltà. Si è in diritto di pensare che gli esaminatori si comportino in modo simile e che il loro giudizio sia espressione di una oggettività unanime piuttosto che di una oggettività assoluta. Ma che fare?

Dapprima definire, in modo per quanto preciso possibile, l'oggetto della valutazione, condizione per la validità di costrutto. Poi, cercare di scoprire le ragioni delle discordanze tra esaminatori. Ci limitiamo a ricordarne alcune: il lettore particolarmente interessato a questo problema potra consultare i principali trattati di didattologia.

In assenza di una concertazione complessiva sugli obiettivi perseguiti, coloro che sono chiamati alla valutazione, basandosi sulla propria esperienza e sulla possibilità degli alunni, seguono criteri differenti a volte poco precisi. Una stessa performance può pertanto essere stimata eccellente in un gruppo e scadente in altri.

È vero che si può raggiungere un accordo sui principali criteri valutativi, ma spesso essa cambia l'importanza ad essi accordata dai diversi valutatori. Ad una performance complessa dell'alunno, corrispondono scarsi sempre più accentuati. Inoltre, il giudizio può essere influenzato dall'intervento di fattori estranei alla performance valutata: simpatia dell'alunno, irritazione per una calligrafia illeggibile, pregiudizi di chi valuta,... ecc.

Con alcuni procedimenti si può rimediare, almeno in parte, a molte carenze: correzione anonima, correzione in serie di tutte le risposte ad una stessa domanda, variazione dell'ordine di correzione delle risposte, ecc. Ma se di lì a questi antidoti si impone un modo di procedere ben diverso. Esso consiste essenzialmente nella definizione comune e operativa degli obiettivi del compito proposto, in somma alla definizione dei criteri (qualitativi e quantitativi) della valutazione. Gli esperimenti presentati in quest'opera rappresentano il risultato di un simile sforzo di precisione e di concentrazione.

In una situazione di apprendimento, la valutazione serve essenzialmente a informare l'alunno sulla qualità delle sue prestazioni, dei suoi progressi in direzione dell'obiettivo perseguito. Già da tempo si conosce l'importante ruolo svolto da una simile informazione. Grazie ai lavori teorici e sperimentali di B.F. Skinner, sappiamo quanto sia determinante questa funzione (%).

Va ulteriormente sottolineato che la valutazione è molto più esatta se l'obiettivo da perseguire è definito con precisione. Naturalmente, l'alunno deve conoscere questo obiettivo, comprenderlo e averlo fatto suo. La valutazione è utile all'insegnante: senza di essa egli non potrebbe conoscere in quale misura la sua azione è stata efficace, né sapere come dovrà modificarla nel seguito per aiutare meglio l'alunno.

Per assolvere efficacemente a queste funzioni la valutazione deve soddisfare un certo numero di esigenze, oltre a quelle già ricordate.

La valutazione deve essere formativa

È necessario un ulteriore approfondimento di questo importante aspetto, cui si è già fatto cenno nei suoi caratteri generali. La va-
luzione si è ridotta troppo spesso, in passato, ad una classificazione che trasformava la vita scolastica in una specie di gara continua. Due preoccupazioni dominavano questo modo di fare: scoprire i migliori e se possibile l'elitè e stimolare ciascuno per mezzo della competizione. Sfortunatamente, la concorrenza ha incoraggiato solo praticamente i primi (che quasi sicuramente sarebbero stati tali comunque) e scoraggiato gli altri. Inoltre, un tal genere di valutazione è pressoché inutile per l'apprendimento. Quali informazioni si intende puramente quanto manca per raggiungere l'obiettivo, e quali difficoltà gli impediscono di procedere? Questo secondo modo di intendere la valutazione può veramente aiutare l'alunno a progredire: riguarda direttamente l'educazione, la formazione (di chi il termine di valutarsi in rapporto con gli altri. Certamente anche questo tipo di incolologicamente opportuno ripetere ad un alunno, ogni giorno, quale

E opportuno che in una situazione di apprendimento il termine insuccesso (spesso vissuto dramaticamente da alunni e genitori) venga sostituito dal termine difficoltà da superare.

Nel momento in cui ci circondiamo di consigli altrui per meglio svolgere un lavoro, attendiamo suggerimenti, indicazioni, critiche relative al compito e siamo disposti ad accettare la misura in cui ci aiutano a progredire. Per contro, in tale circostanza non accettiamo la personalità dell'intelligenza. La valutazione si deve dirigere sull'oggetto dell'apprendimento da conseguire e non sull'individuo. Ciò non si verifica perché è stato fatto del lavoro di per se stesso scadente.

Gli studi documentali sulle performances complessive hanospesso dimostrato che il voto scaturito da un procedimento analitico non è necessariamente diverso dal voto globale, accordato in tempi brevi. Può darsi che le cose non siano così semplici, ma sicuramente la laboriosità del procedimento analitico non è sempre una ragione per rinunciare, una volta per tutte, al metodo analitico? Non ha mai detto di un alunno che il suo lavoro è «buono» o «scarso» se non gli si mostra perché è così. L'allievo non potrà superare le difficoltà incontrate se le stesse non vengono definite e circoscritte per essere affrontate sistematicamente.

In breve, la valutazione formativa è essenzialmente analitica, diagnostica. Il criterio di scelta tra vote analitico e vote globale non può consistere nella rapidità e convenienza per colui che correge, ma nell'obiettivo perseguito.

Un pedagogista inglese affermava recentemente con malizia, che quando si sa dove andare, sussistono maggiori possibilità di giungere alla meta (4).

Pertanto, gli obiettivi educativi devono essere dunque definiti. Inoltre, è anche opportuno ordinare, non potendo perseguirli tutti contemporaneamente. In alcuni casi, e dopo aver precisato l'obiettivo generale (ad esempio, imparare a risolvere alcuni problemi che si svolgono all'interno del calcolo algebrico), l'ordinamento degli obiettivi specifici è in larga parte dettato dalla struttura della materia (bisogna conoscere la regola dei segni per poter ricorrere al calcolo algebrico). Altre volte, l'ordine non si imponi con tale forza.

Man mano che ci si allontana dagli apprendimenti di base e ci si immerge nella vita intellettuale e affettiva maggiormente si articolano i percorsi dell'apprendimento.

Uno dei compiti essenziali dell'insegnante è di scoprire quali sono i nodi, i «centri», i punti critici e di individuare le strategie che ottimizzano l'acquisizione di queste conoscenze o di questi poteri basilari per la maggior parte degli alunni, in teoria per tutti.

In questa prospettiva, una valutazione ha per criterio la performance di altri è priva di senso. Quel che conta è informare l'alunno dei suoi progressi rispetto all'obiettivo perseguito. In tal caso (e opportuno ribadirlo) la valutazione può pienevolentrievi il suo ruolo di controllo e di rinforzo perché è parte integrante del processo di apprendimento e trova l'oggettività del suo criterio nell'esito più o meno felice dell'impresa. Si pensa a quanto è diverso un siffatto modo di procedere dall'abitudine abbia ancora frequente, di svalutare una performance a causa del gioco di classificazioni comparative, non necessariamente per la sua cattiva qualità, ma per essere inferiore, a volte di poco, a quella di altri. L'insegnamento è talmente impregnato di preoccupazione selettiva, di artificiale emulazione, che molti professori ben difficilmente accettano l'idea di attribuire a tutti gli alunni un buon voto (5). Eppure, un'insenazione basilare riuscito non è quello che conduce tutti gli alunni a buoni risultati?

Una valutazione continua, analitica, centratà sull'apprendimento da realizzare: ecco le tre qualità indispensabili per guidare l'alunno nel suo progresso quotidiano e per procurare all'insegnante le informazioni necessarie per meglio adattare la propria azione. Priva di queste finalità, la valutazione non assumesi al suo ruolo formativo.

L'esigenza di validità, oggettività e precisione, era presente anche nelle direttive ufficiali. Risponde a queste esigenze cruciali la scala di valutazione a cinque livelli, come era inizialmente nell'insegnamento secondario rinnovato? La risposta a questa domanda è, in genere, l'esame delle reazioni degli insegnanti dopo un primo periodo di applicazione dei nuovi metodi valutativi, è stato l'obiettivo di una ricerca preliminare.


(³) Ricordando che B.F. Skinner definisce col termine *rinforzo* tutti quei fatti che, verificandosi come conseguenza di una risposta, ne aumentano la probabilità di trasmissione.


Specialmente nell’insegnamento superiore, si riscontrano ancora casi in cui il voto minimo per superare un esame è fissato soltanto dopo la correzione dei lavori. L’esaminatore può pertanto continuare a distribuire i voti in maniera tradizionale.

---

**Capitolo secondo**

**Risultati di un’indagine**

---

**Ambito dell’indagine**

Il nuovo sistema di valutazione dell’insegnamento secondario rinnovato è stato presentato e giustificato in numerosi documenti ufficiali (¹). Il rinnovamento si caratterizza innanzitutto per l’abbandono degli esami tradizionali a vantaggio di un controllo delle conoscenze che fanno parte integrante dell’apprendimento. L’obiettivo è di «consentire al professore di adattare costantemente il lavoro della classe alla situazione reale e di tendere ad una migliore individuazione dell’insegnamento» (1969). Si tratta perciò di una valutazione formativa.

La seconda caratteristica è l’abbandono del voto in cifre, delle somme di punti e delle classificazioni. Si rimprovera al sistema dei punteggi (1970) di cristallizzare una situazione momentanea al punto tale che la somma di questi punti deforma la realtà e comporta decisioni in cui il passato dell’alluno grava continuamente sul suo avvenire. Questo modo di agire si oppone agli obiettivi dei cicli di osservazione e orientamento, che dovrebbero consentire di seguire l’evoluzione degli alunni e «assicurare una guida migliore durante il corso degli studi, allo scopo di preservarli da errori di orientamento e dalle conseguenti perdite di tempo». È stato evidenziato il carattere soggettivo della valutazione in cifre; e in particolare, la mancanza di concordanza tra le valutazioni dei diversi professori e il pericolo di stereotipie nei giudizi formulati da uno stesso insegnante sui suoi alunni. Ha sostituito il sistema tradizionale una scala qualitativa suddivisa in cinque livelli: *Molto bene, Bene, Sufficiente, Debole, Insufficiente*. Questo sistema di classificazione permette di valutare ogni esercizio o attività dell’alluno durante l’anno scolastico (²).

Le valutazioni trimestrali, riportate sulla pagella, riguardano
tre operazioni mentali essenziali: le conoscenze, la comprensione e l'utilizzazione dell'esperienza. Una voce supplementare «Osservazione e consigli educativi» è riservata alla valutazione di altri processi, quali la capacità di analisi, di sintesi, di inventiva, il senso critico, assieme ad altre annotazioni sul comportamento psicosociale.

L'osservazione qualitativa è confermata da controlli di natura quantitativa, condotti in conformità dei principi della statistica. Ad esempio, è stato previsto che gli insegnanti della stessa disciplina propongono, due o tre volte l'anno, a tutte le classi di egual livello un esercizio di controllo tarato a cento. Il voto medio può servire allora come asse di distribuzione degli alunni appartenenti al livello esaminato. Questo modo di procedere aggiunge un riferimento normativo alla valutazione continua, di tipo formativo.

«In un sistema che abbia abolito punti e classificazioni, il consiglio di classe acquista maggior importanza. Esso deve tender più che mai a cogliere l'immagine evolutiva di ciascun ragazzo e a comprendere quanto sono motivi del debolezza o delle difficoltà connotate». Questo profilò si concreta in due documenti diretti agli alunni e ai genitori (la pagella) e al corpo docente (dossier scolastico) (?).

Fin dall'apparizione delle istruzioni ufficiali, numerosi insegnanti, incoraggiati dai ispettori e da capi di istituto, si sono messi decisamente al lavoro per provare in maniera adeguata un sistema di cui approvarono i presupposti tecnici. A ben vedere, il nuovo sistema si differenziava notevolmente dall'antico: nella presentazione si faceva spesso riferimento a nozioni di statistica descrittiva poco familiari agli insegnanti e altre difficoltà erano in aggiunto. Ma non anticipiamo: è importante sottolineare il carattere positivo dell'accettazione iniziale e la buona volontà legata ai primi tentativi.

**Obiettivi dell'indagine**

A due anni dall'introduzione del nuovo sistema si ritenne opportuno raccogliere due tipi di informazioni:

- **informazioni oggettive**: che tipo di scala utilizzano gli insegnanti? Con quale frequenza e in che modo controllano l'apprendimento degli alunni? Come riassumono le valutazioni trimestrali? Come individuano i risultati degli alunni?...

- **informazioni soggettive**: essenzialmente, l'opinione degli insegnanti sulle difficoltà, sugli aspetti positivi e gli inconvenienti del nuovo sistema.

Dal tempo dell'indagine sono trascorsi diversi anni. I suoi risul-

tati sono ancora sufficientemente attuali per essere ricordati. E lo sono ancor più dal settembre 1976, da quando molti insegnanti, rimasti fino ad allora estranei ai nuovi metodi valutativi, hanno adottato il nuovo sistema.

**Popolazione e discipline interessate**

Agli insegnanti di lingua materna (francese) e di matematica fu inviato un questionario. La scelta era motivata dalle seguenti ragioni:

1. l'importanza dei due settori prescelti;
2. l'impossibilità materiale di estendere l'indagine a tutti gli insegnanti;
3. la necessità di disporre, per un'ulteriore fase della ricerca, di dati di riferimento attualmente esistenti per quelle due discipline soltanto.

Degli 89 istituti di istruzione secondaria che avevano iniziato il rinnovamento al momento dell'indagine ne hanno partecipato 80. Complessivamente hanno risposto 391 professori, di cui 197 di francese e 194 di matematica.

**Questionario**

Sono stati presi in considerazione i seguenti aspetti:

- Modalità di controllo
- Frequenza dei controlli
- Livelli tassonomici dei controlli
- Sintesi delle valutazioni trimestrali
- Coordinamento dei controlli
- Riferimenti utilizzati
- Difficoltà legate al nuovo sistema di valutazione

Le domande, prima di assumere la forma definitiva, sono state provate con un numero limitato di insegnanti.
Spoglio dei dati

Lo spoglio dei dati si è svolto in tre fasi:
1. Codificazione delle risposte a domande chiuse.
2. Analisi del contenuto per estrarre i temi più frequenti nelle risposte a domande aperte. Calcolo della frequenza di apparizione di questi temi.
3. Traduzione grafica delle informazioni.

Risultati

1. Modalità di controllo

1.a.1. **Utilizzare la scala a cinque livelli**: MB = B - S - D - I? (\*)
   
   - Tutti i professori utilizzano la scala a cinque livelli imposta dal Ministero.

1.a.2. **Alcuni professori hanno elaborato un loro personale sistema di notazione. Se è il vostro caso, descrivetelo.**
   
   35 professori di matematica e 29 di francese hanno ampliato la scala a cinque livelli:
   - aggiungendo due gradi estremi (ad es.: Perfetto o Eccellente, Nulla);
   - frazionando uno o più gradi (Creazione di sottocategorie. Ad es.: S + , S = , S);
   - utilizzando il voto tra due gradi nella pagella;
   - utilizzando segni di colore diverso per ogni domanda:
     - **verde** = lavoro globalmente valido
     - **bleu** = lavoro incompleto o parzialmente valido o medio
     - **arancione** = lavoro molto incompleto o non soddisfacente.

La scala non è soddisfacente per tutti i professori perché manca di sfumature. Si vedrà alla domanda 6.a.1. ("Il nuovo sistema di valutazione vi sembra meno preciso del vecchio?") che i commenti si dirigono in questo senso.

(*) **Legenda**: MB = Molto Bene; B = Bene; S = Sufficiente; D = Debole; I = Insufficiente (N.d.t.).

1.b.1. **Valutare il lavoro di un alunno con un voto numerico per trasferirlo nella nuova scala?**

Il 62% dei professori di matematica e il 50% di francese hanno risposto positivamente (cfr. il grafico seguente).

Percentuali così elevate meritano attenzione. Sarebbe semplicistico ritenerle una manifestazione di resistenza al cambiamento, anche se l'abitudine svolge sicuramente un certo ruolo. Pensiamo piuttosto che, in un processo di ordinamento, il numero è per molti uno strumento di pensiero più significativo della lettera.

Gli esercizi corretti secondo una rigorosa scomposizione (con l'adozione di un punteggio, score), vanno esaminati a parte. Lo stesso esercizio può, in effetti, essere trattato simultaneamente con voto numerico e con una nota di valutazione.
Accanto alla procedura finale che comporta la trasformazione dal numero in un giudizio globale, è importante esaminare su quali criteri si fonda la prima valutazione. Se questi criteri non vengono precisati in modo che ciascuno possa adottarli e ricondursi ad essi con le stesse modalità, si rischia di trovare tutta la gamma delle suddivisioni settoriali e circostanze, spesso evidenziate dalla docimologia.

A questo proposito, un'indicazione chiara proviene dai limiti numericci dei diversi gradi, segnalati da alcuni professori.

- Il limite inferiore del livello MB va da 7,5 a 9
- Il limite superiore del livello B va da 7,5 a 9
  — inferiore del livello B va da 6 a 7,5
- Il limite superiore del livello S va da 6 a 8
  — inferiore del livello S va da 4 a 7
- Il limite superiore del livello D va da 4 a 6,5
  — inferiore del livello D va da 2 a 5
- Il limite superiore del livello I va da 2 a 5
  — inferiore del livello I va da 0 a 2

Il grafico mostra chiaramente la sovrapposizione dei voti:

![Grafico mostra sovrapposizione dei voti](image)

Così, secondo gli insegnanti, si constata che:
- 8/10 si traduce con MB, B o S.
- 6/10 si traduce con B, S o D.
- 4/10 si traduce con S, D o I.

Appare evidente la necessità di una scala descrittiva e di un procedimento di armonizzazione dei giudizi.

Per le disposizioni attualmente impartite ai docenti, il significato dei voti risulta poco chiaro. È impossibile pronunciarsi sull'omogeneità di comportamenti rapportati ad uno stesso grado della scala.

1.b.2. Valutate il lavoro dell'alunno con una valutazione globale della performance?

![Numero dei professori in %](image)

Il grafico mostra una classificazione dei professori:
- professori di matematica
- professori di lingua

Il 48% dei professori di matematica e il 63% di francesi hanno risposto positivamente a questa domanda. La ripartizione delle percentuali è inversa, in rapporto alla domanda precedente (1.b.1.).

Potrebbe sembrare che molti professori utilizzino contemporaneamente il voto numerico ed il giudizio globale. La comparazione tra i due grafici lascia supporre che per alcuni insegnanti i due metodi di valutazione si escludono.

- Parte dei professori di matematica utilizza esclusivamente il voto numerico.
- Parte dei professori di francesi utilizza soltanto il giudizio globale.
1. b. 3. Valutate lo studio di un alunno con un voto che sintetizza diversi criteri ben definiti?

Hanno risposto positivamente il 30% dei professori di matematica ed il 55% di quelli di lettere. I criteri più frequentemente menzionati sono ripartiti in due categorie.

**Lingua francese**

1. Criteri specifici di un compito (specialmente la stesura).
   Es.: Contenuto, forma, ortografia, struttura, ricchezza di vocabolario, mantenersi in argomento, procedimento logico delle idee, chiarezza di espressione, ecc.
   52 professori

2. Criteri più generali, non specifici di un compito.
   Es.: Applicazione, spirito critico, impegno, lavoro in classe, conoscenze acquisite, personalità, progressi, iniziativa, "efficienza", pulizia, cura, chiarezza, esattezza, rendimento, ecc.
   57 professori

**Matematica**

I principali criteri proposti dai professori di matematica sono i seguenti:

1. Metodo, esattezza, precisione, presentazione. 12 professori
2. Capacità di applicazione, attitudine al ragionamento, memorizzazione, ordine, cura, attività, rapidità, ecc. ... 60 professori

Per molti professori esiste una sola matematica e ciò limita il numero di criteri specifici.

2. **Frequenza dei controlli**

2.a. Frequenza dei controlli scritti, per periodo e per alunno?

Per una gran parte dei professori il numero dei controlli scritti non va oltre l'uno per settimana.

- Matematica: la frequenza dei controlli oscilla tra due e sei per alunno e per periodo (un periodo = 6 settimane);
- Francese: si verifica una ripartizione più ampia (da 1 a 6). La frequenza maggiore è da 2 a 3 controlli per periodo.
2.a.1. Quando vi sembra necessario un controllo scritto?

Le risposte si possono suddividere in tre categorie; la categoria pedagogica comprende cinque sottocategorie:

1. Motivi disciplinari
   - Per costringere gli alunni a studiare.
   - Considerata la lentezza generale.
   - In presenza di un rilassamento.

2. Motivi amministrativi
   - Per la pagella, per l’omologazione.
   - Alla fine di un periodo, di una settimana.

3. Motivi pedagogici
   a. I controlli sono determinati sistematicamente dalla progressione logica dell’insegnamento della materia. Hanno luogo:
      - dopo l’insegnamento di ogni nuovo argomento;
      - dopo aver esaminato un aspetto del programma;
      - alla fine di un capitolo;
      - dopo una serie di esercizi dello stesso tipo.
   b. I controlli si effettuano in funzione dei punti trattati del programma e ritenuti fondamentali. Hanno luogo:
      - dopo lo studio di un argomento importante per il proseguimento degli apprendimenti;
      - dopo una fase di sintesi;
      - dopo un insieme di esercizi o l’apprendimento di argomenti che formano un tutto.
   c. Lo scopo dei controlli consiste nel verificare che gli obiettivi perseguiti siano stati raggiunti. Si valuta essenzialmente:
      - la conoscenza teorica;
      - la comprensione della materia;
      - la capacità di applicazione.
   d. Ci si preoccupa d’osservare alcuni comportamenti particolari:
      - la frequenza di apparizione di alcuni tipi di errori;
      - l’acquisizione di una nozione.
   e. I controlli si inscrivono all’interno di un quadro di esercizi abituali:
      - ortografia, grammatica, stesura (risposte tipiche degli insegnanti di francese);

La necessità di un controllo con motivazioni pedagogiche è ritenuta opportuna dall’80% dei professori di matematica e dal 37% di francese.

2.a.2. Idealmente l’insegnamento, e perciò il controllo che l’accompagna, conduce progressivamente all’acquisizione di conoscenze alla loro applicazione, alla ricerca originale (invenzione), per culminare, infine, nell’uso dello spirito critico. Quale posto è occupato dai quattro obiettivi nel vostro insegnamento?

---

<table>
<thead>
<tr>
<th>Numero dei professori in %</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Conoscenza</strong></td>
</tr>
<tr>
<td><strong>Utilizzazione</strong></td>
</tr>
<tr>
<td><strong>Invenzione</strong></td>
</tr>
<tr>
<td><strong>Spirito critico</strong></td>
</tr>
<tr>
<td><strong>Obiettivi</strong></td>
</tr>
</tbody>
</table>

---

Nella gerarchia ideale degli obiettivi su cui verte la valutazione, la maggioranza dei professori di matematica e di lingua francese (circa il 65%), accorda il primo posto all’applicazione delle conoscenze. Segue la conoscenza della disciplina (matematica: 30%; francese: 20%). La percentuale riservata all’invenzione è trascurabile in matematica (1%). D’altra parte non è poi troppo elevata in lingua (11%), tenendo presente che qui le occasioni sono molto più numerose.
2.b. Frequenza dei controlli orali, per periodo e per alunno?

2.b.1. Quali sono i vantaggi dell'interrogazione orale in rapporto al controllo scritto?

Le osservazioni degli insegnanti si possono raggruppare in tre categorie:
1. **Vantaggi per l'alunno**
   1.1. Valutazione immediata
   1.2. Consente di evitare nuovi errori, di rimettere in carreggiata
   1.3. Incoraggiamento per gli alunni più deboli
   1.4. Adatta per gli alunni che si emozionano nelle prove scritte, o per coloro che si esprimono con difficoltà per iscritto
   1.5. Domande individualizzate a seconda delle capacità
   1.6. Obbliga gli alunni a studiare regolarmente
   1.7. L'alunno si sente seguito con regolarità
   1.8. Valorizza l'alunno davanti alla classe
   1.9. Offre all'alunno l'occasione per esercitare il suo senso critico
   1.10. Offre all'alunno l'occasione per esprimersi

2. **Vantaggi per l'insegnante**
2.a. Nel suo ruolo di insegnante generale
   2.a.1. Consente di adeguare il suo insegnamento (il controllo è immediato)
   2.a.2. Maggior facilità di contatto (dialogo)

2.b. **In particolare**
2.b.1. È un'occasione per esercitare l'elocuzione
2.b.2. È un'occasione per modificare gli errori di linguaggio
2.b.3. È un'occasione per rinforzare la memoria della classe (ricchi o sintesi delle lezioni precedenti)

2.c. **Come controllare**
2.c.1. Controllo diretto della materia
2.c.2. Rapidity maggiore dello scritto
2.c.3. Offre una più vasta gamma di tipi di valutazione
2.c.4. Controllo più approfondito, più sfumato, più oggettivo dello scritto. E una conferma dello scritto.

3. **L'orale è più efficace dello scritto per valutare alunni aspetti:**
3.1. La sintesi, la comprensione e l'applicazione di una nozione
3.2. La rapidità nel risolvere un problema
3.3. La spontaneità, la prontezza, la creatività
3.4. Altri aspetti:
3.4.a. Si possono osservare le reazioni dell'alunno di fronte ad una difficoltà
3.4.b. Migliore conoscenza dell'alunno
3.4.c. Consente di osservare la personalità dell'alunno

3. **Sintesi delle valutazioni al termine di un periodo**
3.a. **Come sintetizzare le valutazioni alla fine di un periodo?**
   Dopo le osservazioni dei professori si è proceduto alla classificazione delle risposte, che possono essere raggruppate in due grandi categorie:
   A. Una categoria generale che raggruppa risposte sul piano qualitativo:
      • Utilizzazione dei grafici
      • Osservazione della tendenza dominante (con l'eventuale considerazione dei giudizi estremi).
   B. Una categoria che raggruppa risposte a livello quantitativo: calcolo della media dei voti numerici e sua trasposizione nella nuova scala.
   Sono state introdotte inoltre tre sottocategorie:
   a. Riferimento al gruppo classe (calcolo della media generale della classe e aggiustamento delle medie individuali).
b. Considerazione degli sforzi, dei progressi e delle capacità dell'alunno
c. Menzione di valutazioni parziali rivolte alle differenti componenti di una disciplina (risultati periodici in ortografia, composizione, ecc.) anteriori alla sistemazione dei risultati medi del periodo.

La risposta di un professore può essere classificata in più categorie, a seconda degli elementi contenuti.
Infine, si è distinta una terza grande categoria (C) in cui sono riunite le risposte di natura prettamente statistica (principalmente il calcolo dello scarto-tipo e la curva di Gauss).

Tendenza generale
Numerosi professori (80 di matematica e 83 di lingua) prendono in considerazione gli sforzi, i progressi e le capacità degli alunni per sistemarli in un grado della scala a cinque livelli (categorie A, B, C, D, E). Questa tendenza sarà confermata dall’esame della domanda 3.c.1. (le qualità umane dell’alunno intervengono nel voto comportando una modificazione del livello raggiunto nella scala).

Ricorso alla statistica (calcolo del sigma, curva di Gauss, media aritmetica).
I professori di matematica vi ricorreranno più frequentemente dei colleghi di lingua.
Già la domanda 1.b.1. verteva sull’utilizzazione dei voti numerici: vi ricorreranno il 62% dei professori di matematica e il 50% dei professori di francese. Nelle sintesi di fine periodo, lo scarto dei risultati tra le due categorie di professori aumenta notevolmente: 91% tra i primi e 44% tra i secondi. Sembra che i professori di lingua che utilizzano voti numerici nel momento di valutare una performance isolata, non vi ricorrano più nella sintesi generale.

3.b. Quali difficoltà incontrate?
Le principali cause di difficoltà risiedono:
1. Nella sintesi dei risultati e nella traduzione in un livello della scala: 32 professori di matematica e 40 di francese.
2. Negli intervalli poco numerosi della scala a cinque gradi, nella mancanza di sfumature: 33 professori di matematica e 30 di lingua. Molti insegnanti non avvertono difficoltà. Altri esprimono alcuni motivi di lagnanza; essi sono stati suddivisi in cinque categorie:
Affermano di modificare il loro voto, in funzione delle qualità umane dell’alunno, il 59% dei professori di matematica e il 76% di lingua.
Il 94% di entrambi le inseriscono sotto forma di commento aggiunto alla valutazione.

4. Coordinamento dei controlli
4.4.1. I colleghi della stessa disciplina utilizzano controlli uguali ai vostri?

Ha risposto positivamente il 70% dei professori di matematica e di lingua.
4.4.2. I tipi di controllo utilizzati sono stati elaborati in collaborazione:
   a. con insegnanti dello stesso ramo?
   b. con insegnanti di altri rami disciplinari?

4.4.3. a più istituti?

4.4.4. Pensate che sia possibile adottare criteri di valutazione comuni: a. all’insieme degli insegnanti di lingua e di matematica di uno stesso istituto;

---

4.b. Pensate che sia possibile adottare criteri di valutazione comuni:
   a. all’insieme degli insegnanti di lingua e di matematica di uno stesso istituto;

---

4.4.5. pensate che sia possibile adottare criteri di valutazione comuni: a. all’insieme degli insegnanti di lingua e di matematica di uno stesso istituto;

---

4.b. Pensate che sia possibile adottare criteri di valutazione comuni:
   a. all’insieme degli insegnanti di lingua e di matematica di uno stesso istituto;

---

4.4.6. pensate che sia possibile adottare criteri di valutazione comuni: a. all’insieme degli insegnanti di lingua e di matematica di uno stesso istituto;

---

4.b. Pensate che sia possibile adottare criteri di valutazione comuni:
   a. all’insieme degli insegnanti di lingua e di matematica di uno stesso istituto;
5. Norme
5.a. I vostri giudizi tendono a situare l'alunno:
   5.a.1. in rapporto a se stesso;
   5.a.2. in rapporto alla classe.

5.a.3. in rapporto ad una popolazione più ampia (esempio: insieme di classi prime con insegnamento rinnovato; taratura a livello regionale, provinciale...).

---

- **Situazione dell'alunno in rapporto a se stesso:**
  - 69% dei professori di matematica
  - 86% dei professori di lingua

- **Situazione dell'alunno in rapporto alla classe:**
  - 79% dei professori di matematica
  - 73% dei professori di lingua

- **Situazione dell'alunno in rapporto ad una popolazione più ampia:**
  - 21% dei professori di matematica
  - 14% dei professori di lingua

5.b. Come procedete?

---

Si distinguono sei tipi di risposte:

A. **In rapporto all'alunno stesso:**
   a. confrontando il risultato attuale con i precedenti;
   b. leggendo il grafico dei risultati ottenuti nel corso del periodo.

B. **In rapporto alla classe:**
   a. stabilendo dapprima la media della classe e in seguito le medie individuali;
   b. cercando la «moda» e la posizione del gruppo.

C. Curva di Gauss (partendo dall'istogramma).

D. Calcolo dello scarto tipo.

**Analisi dei risultati**

1. Gli insegnanti di matematica rispondono alla domanda in modo più preciso dei colleghi di lingua. Questi ultimi indicano generalmente lo strumento usato per valutare un gruppo di alunni (Esempi: esercizio comune proposto alle prime classi; test collettivo elaborato dai professori, ecc. ...). Tuttavia, non viene indicato il proce-
dimento matematico che permette di rapportare l'alunno a se stesso e alla sua classe.

2. I procedimenti generalmente usati per rapportare l'alunno a se stesso e alla classe sono, nell'ordine:
- la comparazione del risultato attuale con i precedenti;
- l'aggiustamento delle medie individuali in funzione della media della classe.

3. Raramente si menziona il ricorso alla statistica (Curva di Gauss e calcolo dello scarto tipo)

Pertanto, dall'analisi dei procedimenti utilizzati per situare un alunno in rapporto a se stesso e alla classe, risulta che l'eterogeneità dei metodi è ancora grande.

6. Difficoltà incontrate

6.a.1. Il nuovo sistema di controllo vi sembra meno preciso del precedente? In che cosa?

Il 38% dei professori di matematica e il 33% di quelli di lingua ritengono il nuovo sistema meno preciso del precedente, adducendo come motivo quanto abbiamo osservato nella questione 1.a.2. Molti di essi, rispettivamente nella misura del 58 e del 52%, lamentano la carenza di ulteriori differenziazioni nella scala a cinque livelli.

Una buona metà degli insegnanti (il 55% di matematica e il 54% di lingua) ritiene il nuovo sistema di valutazione altrettanto valido quanto il precedente.

6.a.2. Il nuovo sistema di valutazione vi sembra meno oggettivo del precedente?

6.b. Il nuovo sistema di valutazione richiede maggior tempo di quello tradizionale?

6.c. Il nuovo sistema di valutazione è più soddisfacente del precedente? In che cosa?

Distinguiamo subito gli insegnanti soddisfatti dagli altri:

I. Aspetti positivi del nuovo sistema:
   a. è più umano
   b. assicura maggiormente l'alunno
   c. permette di seguire meglio l'alunno
   d. permette di rapportare meglio l'alunno a se stesso
e. permette la comparazione dei risultati di una classe con quelli di un'altra
f. permette una valutazione più articolata
g. è più razionale
h. è meno rigido, meno arido
i. è più preciso, più oggettivo e consente una migliore conoscenza dell'alunno.

II. I limiti sono i seguenti:
a. è meno preciso
b. si equivale col sistema precedente
c. la non differenziazione nel voto demotiva l'alunno
d. è troppo soggettivo: varia da un professore all'altro, da una scuola all'altra
e. trascura l'emulazione
f. l'alunno studia meno

I principali motivi di soddisfazione sono, nell'ordine:
1. Il nuovo sistema è più umano
2. Il nuovo sistema è più oggettivo, permette una valutazione più articolata ed una migliore conoscenza dell'alunno.

Il principale limite del nuovo sistema viene individuato dagli insegnanti di matematica nella mancanza di precisione.

7. Osservazioni complementari fatte dalla maggior parte dei professori

Ecco, in ordine di frequenza, tre osservazioni fatte da un elevato numero di insegnanti:
1. Il nuovo sistema di valutazione accresce la parte burocratica del lavoro (consigli di classe, rapporti, commenti aggiuntivi al voto, doppi pagella);
2. La scala non è sufficientemente articolata: ad un medesimo livello si fanno corrispondere punteggi troppo differenti.
3. Il voto in cifre influenzava maggiormente il rendimento degli allevili ed era più comprensibile da parte dei genitori.

Ed inoltre si fa osservare:
• bisognerebbe informare meglio i genitori;
• riformare la secondaria, va bene! Ma anche la primaria!
• gli alunni si concedono un'eccessiva libertà di lavoro;

si doveva effettuare il passaggio dal vecchio al nuovo sistema di valutazione con maggiore gradualità;
la votazione varia a seconda dei professori (personalità) e delle classi (livello generale della classe).

Come aspetti positivi vengono sottolineate la maggior giustizia ed umanità del nuovo sistema di valutazione.
Tra le osservazioni isolate, si nota:

A. Osservazioni negative
• l'elevato numero di alunni per classe non consente di adottare efficacemente il nuovo sistema;
• il nuovo sistema non consente all'insegnante di sanzionare la pigriola di un alunno abbassando il voto;
• permette all'alunno scadente di dissipar le sue debolezze (si ritiene più forte di quanto non sia in realtà);
• perché non adottare un metodo di valutazione continua mediante l'osservazione quotidiana?
• va bene la soppressione degli esami! No alla promozione garantita!
• con quali criteri si stabilisce un livello? La scala non è sufficientemente descrittiva;
• la materia è troppo ampia e difficile;
• ponendo in luce l'impegno, l'attività e la condotta si rischia di falsare l'opinione sulle reali attitudini dell'alunno.

B. Osservazioni positive
• il nuovo sistema rassicura maggiormente l'alunno;
• costringe l'alunno a studiare regolarmente;
• smorza le rivalità nella classe;
• è più articolato.

Conclusioni

La valutazione è una delle funzioni essenziali per la maggior parte dei professori cui ci siamo rivolti. Generalmente, il nuovo sistema è ritenuto più umano e rassicurante per l'alunno. Tuttavia non risponde completamente alle esigenze di una valutazione rigorosa, oggettiva e direttamente fruibile da parte dell'alunno.

Le principali difficoltà derivano sicuramente dall'assenza di
una chiara definizione delle finalità dei nuovi metodi di valutazione e dalla mancanza di strumenti adeguati. Cosa si vuol valutare? Con che cosa e con chi? Non abbiamo ottenuto risposte alle seguenti domande:

— Il voto numerico deve essere abbandonato completamente?
— La scala di valutazione ordinale si addice a tutti i tipi di performance? Quale strutturazione potrebbe rendere le scale più oggettive e articolate?

Un tentativo di risposta alle ultime due questioni viene dal presente lavoro.

**NOTE AL SECONDO CAPITOLO**


(2) Quando i cinque livelli vengono rappresentati con le lettere A, B, C, D, E, si parla di scala letterale.


**PARTE SECONDA**

**LE SCALE DESCRITTIVE**

1. Le scale di valutazione
2. Tappe e condizioni della ricerca
3. Metodologia della costruzione
4. Analisi dei risultati
Capitolo primo

Le scale di valutazione

Due tipi di misurazione

Al di là delle differenziazioni, esistono due tipi di misure: quelle che si riferiscono ad una unità convenzionale, talvolta generalmente accettata, sempre la stessa per quanti l’utilizzano, e quelle che si limitano a classificare se l’individuo o la cosa considerati posseggano più o meno un certo attributo, una determinata proprietà o caratteristica.

Per ben comprendere la fondamentale differenza esistente tra i due tipi di valutazione, confrontiamo la misura della lunghezza di una strada, usando il sistema metrico (misura I), con la valutazione dell’eleganza linguistica di un gruppo di alunni (misura II).

<table>
<thead>
<tr>
<th>MISURA I</th>
<th>MISURA II</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Critierio universale</td>
<td>1. Non esiste un critierio universale</td>
</tr>
<tr>
<td>Dal momento che sono nel mondo, qualunque sia la mia cultura o la mia personalità, se decido di misurare una lunghezza usando il sistema metrico, mi riferirò sempre ad uno stesso metro-campione (stabilito in funzione delle radiazioni di krypton).</td>
<td>Due colleghi di identica formazione, che insegnano la lingua materna allo stesso livello, nella stessa scuola, considerano l’eleganza del linguaggio in modo diverso. Al massimo si accontentano di alcuni coefficienti di eleganza. Cosa c’è di comune tra ciò che un Bantu ed un Latino considerano come linguaggio elegante?</td>
</tr>
</tbody>
</table>
2. Gli intervalli della scala sono tutti uguali

Se un tratto di strada dell’Amazzonia è lunga 200 km, e se l’autostrada Bruxelles-Liegi misura 100 km, sappiamo immediatamente che il secondo tratto è metà del primo. Inoltre sappiamo che il tragitto Bruxelles-Liegi e ritorno è uguale al tratto di strada dell’Amazzonia.

3. Lo zero della scala è assoluto

Non esiste una tratta di lunghezza zero.

Mentre la misura I si effettua grazie ad una scala di rapporti (1), la misura II ricorre ad una semplice scala ordinale che indica soltanto una graduatoria. Sapere che un alunno è il primo della classe in inglese, il più alto di statura, il più veloce nel calcio mentale, non comporta alcuna reale precisazione su una sua conoscenza dell’inglese, sulla sua altezza o sulla sua attitudine ai numeri: può essere il più alto dei nani o il più piccolo dei giganti.

Nella vita scolastica numerose valutazioni hanno valore relativo: si classificano i valori, le performances, i comportamenti, le attitudini dal migliore al peggiore, dal più grande al più piccolo, dal più favorito al più svantaggiato, ecc... Le scale di valutazione a cinque livelli, adottate nell’insegnamento secondario rinnovato, sono di questo tipo.

In breve, la scala di valutazione aiuta ad ordinare individui ed oggetti in rapporto ad una qualità, caratteristica, o proprietà posseduta ad un livello più o meno elevato. E una scala ordinale.

2. Gli intervalli non sono ben definiti

Giudicando l’eleganza dello stile, un insegnante classifica le composizioni dei suoi allievi in cinque categorie: Molto Bene, Bene, Sufficiente, Debole, Insufficiente. E impossibile dimostrare che la “quantità” di eleganza che separa Sufficiente da Bene equivale a quella che intercorre tra Bene e Molto Bene. Di più, due composizioni giudicate Bene e scritte l’una di seguito all’altra non divenivano una composizione che vale Molto Bene. Se due insegnati di identica formazione, docente presso la stessa scuola e in classi parallele, danno a Molto Bene un valore diverso, che dire di ugual giudizio formulato in Giappone e in Francia?

3. Non esiste lo zero assoluto

Per quanto povera, è impossibile dimostrare che una composizione manchi totalmente di eleganza. Quel che oggi può sembrare sciocchezza, un giorno potrà diventare una trovata. Al massimo si può dire che una composizione classificata ultima in un determinato momento, occupa questa posizione in rapporto ad altre, secondo l’opinione soggettiva del professore.

Tipi di scale

1. La scala grafica

La forma più elementare è data da una semplice linea, di lunghezza variabile, che rappresenta tutti i gradi di una caratteristica posseduta da persone o cose. Chi valuta traduce il suo giudizio con un semplice segno.

Esempio:
Conoscenza dell’inglese parlato: + x

2. La scala verbale

Esempio 1: Conoscenza dell’inglese parlato:

- Molto Bene
- Bene
- Sufficiente
- Debole
- Insufficiente

Esempio 2: L’insegnante:

- tiene sempre conto delle osservazioni degli alunni;
- x - tiene qualche volta conto delle osservazioni degli alunni;
- non tiene affatto conto delle osservazioni degli alunni.

3. La scala numerica

Conoscenza dell’inglese parlato: 1 - 2 - 3 - 4 - 5

4. La scala letterale

Conoscenza dell’inglese parlato: A - B - C - D - E

5. La scala concreta

In queste scale i gradi vengono stabiliti per mezzo di persone o cose assunte come modello.

Esempio: Invitiamo 130 alunni a disegnare lo stesso oggetto. L’inse-
gnante sceglier le disegni migliore, il peggiore, e quello medio. Tutti gli altri disegni vengono suddivisi in tre categorie, comparabili ai tre individuati.

6. La scala combinata
Spesso la scala grafica si combina con la scala numerica, letterale o verbale. Nell'esempio che segue sono riunite tutte le possibilità:

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Insufficiente</td>
<td>Debole</td>
<td>Sufficiente</td>
<td>Bene</td>
<td>Molto Bene</td>
</tr>
</tbody>
</table>

7. La scala a scelta obbligatoria
Il valutatore deve scegliere tra una serie di tratti quello che più corrisponde all'individuo valutato.

Esempio: Scegliere tra le seguenti proposizioni quella che descrive meglio l'insegnante che avete osservato:

A. Prende sempre in considerazione le osservazioni degli alunni
B. Prende in considerazione soltanto le osservazioni inerenti a quanto ha esposto
C. Persegue la sua idea senza preoccuparsi delle reazioni della classe.

In teoria, in una scala a scelta obbligata le proposizioni vanno redatte in modo che non sia possibile notare, da parte di chi giudica, qual è la più adeguata ad una scelta pedagogica o, più in generale, ad un criterio definito. Si spera, in tal modo, di porre tutte le proposizioni sullo stesso livello di interesse e di ridurre al massimo il gioco della accettazione sociale.

8. Scala stabilita da chi valuta (Self-anchoring rating scale)
In tutte le scale precedenti, la definizione del tratto da valutare ed i differenti gradi soni imposti a chi valuta effettivamente. Ebbene, nel valutare, ad esempio, la qualità di un insegnante può accadere che si abbia un concetto assai diverso di questa qualità rispetto a chi ha costruito la scala; o addirittura, che chi la utilizza non comprenda cosa ha voluto dire chi ha costruito la scala.

In alcuni casi, è opportuno conoscere la scala di valori del valutatore e constatare come inserisce l'oggetto dalla sua osservazione (eventualmente se stesso) su questa scala. F.P. Kilpatrick e H. Cantril (5) propongono un procedimento che si può così schematizzare:

1. Si chiede al soggetto di descrivere, ad esempio, come egli concepisce l'insegnante ideale. L'intervista (registrata) si prolunga finché il soggetto non ha più niente da aggiungere;
2. Si procede in modo analogo per definire l'insegnante peggiore;
3. Si traccia quindi una scala numerica e si invita il soggetto a situare in essa se stesso o un insegnante che è stato osservato.

Le ultime scale presentate stanno a dimostrare un primo sforzo nel tentativo di cogliere con precisione il tratto valutato. Torneremo sul problema di definire i criteri di misura a proposito della valutazione delle performances scolastiche.

Quando utilizzare le scale di valutazione?

Tenendo presente la natura degli apprendimenti o delle performances da valutare, a volte si richiede un calcolo o un punteggio (score), a volte una valutazione ordinale. Infatti, sono pochi i casi in cui una performance scolastica non può essere oggetto di una valutazione ordinale. In un dettaglio, ad esempio, si intreca spesso il punteggio con una nota del tipo «bene», «c'è un miglioramento», ecc.

La distinzione tra sistema di categorie e sistema di segni (5) ci aiuterà a chiarire i concetti.

- In un sistema di categorie si conteggia la quantità degli eventi: ad esempio, quante volte interviene il professore per incoraggiare l'alluno.
- In un sistema di segni, l'evento è contato una sola volta, ma è oggetto di valutazione. Ad esempio, si valuta in quanti segni un professore ha parlato chiaramente durante la lezione.

Il conteggio, proprio del sistema di categorie, ha un carattere prettamente oggettivo; la valutazione inerente al sistema di segni è invece più soggettiva. Comunque, man mano che si precisano i criteri di una scala di valutazione, un sistema di segni tende a trasformarsi in un sistema di categorie.

Quand'è necessario ricorrere alle scale di valutazione?
L'alunno, al quale è stata posta una domanda, può essere invitato:

- sia a costruire interamente la sua risposta,
- in tutta libertà;
- o rispettando alcune limitazioni, ad esempio la lunghezza.
— sia a selezionare la risposta ritenuta migliore tra quelle proposte.

Esempi:
A. Selezioni di risposte
1. Carlo V è nato nel 1500               VERO - FALSO
2. La capitale dell'Olanda è: Amsterdam Rotterdam L'Aia
3. L'articolazione radio-cubitale consente movimenti di rotazione limitati perché la testa del radio è circondata dal legamento annulaire.
   A. La constatazione e la spiegazione sono vere e tra le due esiste un rapporto di causa-effetto.
   B. La constatazione e la spiegazione sono vere, ma non esiste, tra le due, un rapporto di causa-effetto.
   C. La constatazione è vera, la spiegazione è falsa.
   D. La constatazione è falsa, ma la spiegazione è un fatto o un principio accettato.
   E. Sia la constatazione che la spiegazione sono false.
4. Esempio in cui più risposte sono esatte:
   Cinque soldati di leva sono alti: 1,65m - 1,67m - 1,69m - 1,63m - 1,61m.
   1. La media delle stature del campione è di m. 1,65
   2. Il sigma (σ) è prossimo a 8 (°)
   3. Il σ è prossimo a 2,8
   4. Il campione offre adeguate informazioni sulla statura in generale.
   A. Le risposte 1, 2, 3 sono esatte.
   B. Le risposte 1 e 3 sono esatte.
   C. Le risposte 2 e 4 sono esatte.
   D. La risposta 4 è esatta.
   E. È esatta una sola risposta: la 1, la 2 o la 3.
B. Costruzione di risposte.
1. Risposte completamente aperte:
   • «Ascoltate il Prélude à l'après-midi d'un faune, di Debussy ed esprimete le vostre impressioni».
   2. Risposte soggette a limiti:
   • Risposta semplice:
     — Es.: «La formula chimica del sale da cucina si scrive ...»
   • Ricercare i giudizi di valore in un testo.
   • Nello stesso spirito dell'esempio precedente, e in modo più generale, gli esercizi di analisi e di interpretazione possono consistere:
     — nello scoprire relazioni;
     — nel decidere se certe conclusioni sono valide oppure no;
     — nello scoprire i casi di appropriata applicazione di alcuni principi;
     — nell'interpretare mappe, grafici, ecc. ...
Valutazione delle risposte
   Nella selezione delle risposte, allorché si tratta di risposte costruite puramente fattuali o dipendenti da limitate scelte, la correzione è agevoluta; nel conteggio di elementi completamente definiti la correzione può essere automatizzata. In tal modo, alla rapidità generale di risposta si aggiunge una corrispondente rapidità di correzione.
   La realtà è ben diversa quando l'alluno risponde in completa libertà. Al limite, egli crea un'opera propria, unica, la cui valutazione gli sia possibile. È proprio qui che la scala di valutazione svolge appieno il suo ruolo, perché, come abbiamo già visto, essa riguarda essenzialmente le performances complesse.
Livello cognitivo
   Dagli esempi precedenti notiamo che le risposte a scelta multipla non si limitano alla semplice memoria dei fatti, ma possono chiamare in gioco i processi intellettivi superiori.
   Non sarebbe comunque corretto strutturare tutti gli esercizi sotto forma di domande a scelta multipla. È altrettanto importante che l'alluno sappia esprimersi correttamente, vale a dire creativamente (°). In questo tipo di risposte, non si pone l'accento soltanto sulle capacità particolari, ma anche, se non soprattutto, sul modo in cui l'alluno le collega, le articolare, per produrre un tutto integrato. La difficoltà della valutazione dipende dalla complessità di queste risposte ed è accresciuta dal fatto che gli obiettivi tendono ad esprimersi in misura tanto più genere quanto più la performance da essi implicata è complessa (per es.: sviluppare lo spirito critico).
   Comunque sia, le domande con risposte liberamente costruite (°)
Ostacoli all'oggettività e validità delle valutazioni

Abbiamo già constatato che se gli obiettivi della valutazione non sono adeguatamente definiti i valutatori adottano criteri di valutazione differenti e instabili. Nella letteratura sperimentale si trovano descrizioni di fenomeni più specifici.

1. Effetto di alone

L'aspetto esteriore dell'alunno, il linguaggio ricercato, il modo di presentare i propri lavori ... sono fatti che possono influenzare il giudizio dell’insegnante.

Ad esempio: si tende a sopravalutare l'intelligenza di un alunno vivace.

2. Effetto di contaminazione

Alcuni insegnanti sono portati a sopravalutare gli alunni che negli anni precedenti hanno superato gli esami con «una votazione brillante».

3. Effetto di stereotipia

È il caso di una particolare contaminazione. A dispetto dei loro sforzi, molti alunni hanno visto confermato per tutto l'anno il voto ottenuto nel primo esercizio.

Quando si invita a valutare caratteristiche o tratti differenti con l'aiuto di una batteria di scale tra loro concatenate, si osserva talvolta la tendenza a collocare la maggior parte delle valutazioni in un unico settore.

4. Eccesso di severità o di indulgenza

5. La convergenza al centro (la tendenza centrale)

Spesso la valutazione si concentra intorno ai “valori medi”, sia perché gli alunni di medio valore sono in generale i più numerosi, sia per evitare di assumere eccessive responsabilità. In questo caso, gli insegnanti che osano utilizzare i gradi estremi della scala pesano in maniera autonomi sulle decisioni di successo o fallimento.

6. Effetto di contrasto

Si è portati, ad esempio, a sottovalutare un alunno interrogato immediatamente dopo un compagno assai brillante.

7. Distribuzione forzata

A dispetto di spiacerevole confusioni intercorse tra ripartizione normale delle attitudini e risultati positivi di un insegnamento, tra selezione ed educazione fondamentale, alcuni hanno affermato che i risultati scolastici si distribuiscono secondo la curva di Gauss. Si è previsto, ad esempio, che su 20 alunni, 1 sarà classificato con Ottimo, 5 con Buono, 8 con Medio, 5 con Mediocre, 1 con Debole. Si prevedono così i risultati di un insegnamento che, in teoria, dovrebbe portare tutti gli alunni al livello più alto.

Numerose ricerche, soprattutto nel settore della composizione scritta, illustrano l'inquinamento dei voti da parte di fattori esterni. Ecco alcuni esempi:

John C. Marshall ha evidenziato il condizionamento esercitato da alcune variabili sulla valutazione di un esaminatore. È stato scelto un compito di scienze sociali giudicato dagli esperti di valore medio. Ne sono state prepare tredici copie diverse con errori di ortografia o di grammatica, ma senza modificare il contenuto:

- 3 copie con 6, 12 o 18 errori di ortografia
- 3 copie con 6, 12 o 18 errori di grammatica
- 3 copie con 6, 12 o 18 errori di punteggiatura
- 3 copie con 2, 4 o 6 errori di vari tipi
- 1 copia di controllo senza alcun errore.

Per valutare queste composizioni sono stati chiamati settecento insegnanti scelti in funzione della loro esperienza dovuta agli anni di insegnamento e della loro specializzazione. Ognuno è stato invitato a giudicare una delle tredici copie, utilizzando una scala a nove livelli e badando soltanto al contenuto della risposta.

Dall'esperimento sono derivate le seguenti osservazioni:

1. Pur avendo chiesto di valutare soltanto il contenuto della risposta, gli insegnanti restano condizionati dall'aspetto formale della copia: si nota un rapporto inverso tra quantità di errori in ortografia o in grammatica e voto assegnato alla composizione. Il voto è determinato essenzialmente da questo tipo di errori, più che dagli errori di punteggiatura. L'influenza esercitata dalla loro quantità non
è uniforme. Le copie con 12 o 18 errori sono sottovalutate rispetto a quelle che ne contengono 6 o ancor meno.
2. Gli insegnanti con maggior esperienza si dimostrano più severi.

In un'esperienza dello stesso genere, C. Chase (1) ha anche dimostrato come influenzino gli esami il modo di presentarsi all'esame. In questo caso si è fatto variare la scrittura e l'ortografia. Alcuni insegnanti utilizzano una griglia di correzione, altri ne sono privi. Le conclusioni da trarre sono le seguenti:

1. La qualità della scrittura condiziona significativamente il voto finale: la buona grafia induce ad una valutazione generosa;
2. Si verifica un effetto di voto negativo (sottovalutazione) per gli elaborati scritti male, valutati subito dopo altri di buona qualità grafica, e viceversa;
3. Gli errori di ortografia non influenzano il voto finale (2);

**Necessità di precisazioni operative**

Anche se tecniche specifiche (2) consentono di attenuare l'effetto di fattori esterni, perché la fedeltà delle valutazioni di una persona o la concordanza tra più persone nel valutare una stessa performance risulta tanto bassa finché si utilizzano le scale classiche di valutazione?

Immanzi tutto, perché molte di esse presentano troppi livelli. Pur non essendo una regola precisa, si nota che un insegnante riesce a valutare con una certa precisione se resta all'interno di uniche categorie. Le scale a cinque livelli si rilevano esatte, ma come ha confermato la nostra inchiesta preliminare, danno spesso l'impressione di non articolare a sufficienza le valutazioni espresse. L'esperienza dimostra che è preferibile iniziare con le cinque categorie ed aggiungere in seguito, secondo le necessità, altre suddivisioni verso l'alto o il basso. Nella tradizione accademica belga, ad esempio, in cui è prevista una classificazione che va da altamente distinto all'insuccesso, passando attraverso grande distinzione, distinto, soddisfacente, si sente spesso parlare, nella deliberazione della commissione, di "bella distinzione" o, al contrario, di "piccola distinzione".

Come è stato evidenziato da alcuni autori (3), non conviene utilizzare scale con un numero dispari di livelli per la tendenza espressa da chi valuta a soffermarsi sul livello medio (effetto di convergenza al centro). W. McColl e R. Remstad hanno sperimentato la relativa efficacia di scale non descrittive a quattro o sei livelli nel valutare i componenti. Tra i due strumenti non esistono grandi diversità sul piano della fedeltà e sensibilità. L'applicazione più rapida è l'unico vantaggio della scala a quattro livelli.

Infatti, le controversie sul numero dei livelli hanno valore soltanto con le scale non descrittive. Il motivo sostanziale della scarsa esattezza delle scale, quali che siano, risiede altrove e precisamente nell'indeterminatezza dei criteri di classificazione. Suddividere gli alunni in cinque categorie, A, B, C, D, E o 1, 2, 3, 4, 5, o Ottimo, Buono, ecc., significa lasciare a ciascun insegnante l'incombenza di interpretare questi simboli o votazioni a seconda della ispirazione del momento. Ed il fatto che alcuni insegnanti abili riescano ad accordarsi abbastanza facilmente sulle valutazioni non deve far dimenticare i disaccordi anche gravi che si producono tra loro.

C'è una soluzione soltanto a questo problema: accordarsi su una definizione operativa (4) dei criteri di valutazione e su ogni livello della scala relativo a ciascun criterio.

Senza questo sforzo di operatività, gli insegnanti adottano criteri differenti ed i fattori esterni ed interni, sopra descritti, esercitano la loro influenza perturbatrice.

**NOTE AL CAPITOLO PRIMO**


(4) Il sigma, o deviazione quadratica media, o scarto tipo, è dato dalla "radice quadrata della media dei quadrati degli scarti della media" (G. De Landsheere, *Introduzione alla ricerca in educazione*, cit.) (n.d.l.).
Capitolo secondo

Tappe e condizioni della ricerca

Le tappe

Abbiamo già insistito sulla necessità di far partecipare gli insegnanti alla costruzione di strumenti di valutazione. Riuscire in commissioni, ispettori e insegnanti svolgono il ruolo di esperti da un duplice punto di vista. Da una parte, soltanto essi hanno una conoscenza abbastanza approfondita della materia per scegliere i criteri di valutazione e per prevederne tutte le implicazioni. Dall’altra, la loro esperienza consente di stabilire le attività e le esigenze più rilevanti ai livelli pedagogici prestabiliti. Nella presente ricerca, le commissioni hanno operato a Liegi e nella regione del Lussemburgo.

L’obiettivo consisteva nella messa a punto di una metodologia di costruzione di scale descrittive e nel verificarne efficacia e limiti. È stata realizzata in dettaglio soltanto un’esperienza condotta nel primo ciclo dell’insaggio secondario rinnovato; i commenti riservati, tuttavia, un carattere più generale.

Le tappe per la messa a punto di scale nel primo anno dell’insaggio rinnovato sono state le seguenti:

- Stesure e pre-test;
- Sperimentazione su un campione statisticamente valido;
- Interviste per verificare le ipotesi suggerite dai risultati dell’esperimento;
- Revisione degli strumenti e nuova sperimentazione.
Discipline interessate

La scala descrittiva era uno strumento adatto per valutare performances complesse; pertanto, si è dovuto scegliere, in ciascun settore, un esercizio che si prestasse ad un’ampia discriminazione delle valutazioni e che fosse regolarmente utilizzato nelle classi. Per la lingua materna, la composizione risponde perfettamente a queste esigenze.

In matematica, al contrario, è stato impossibile trovare l’esercizio adatto tra i più utilizzati nelle classi (1). Abbiamo deciso di cercare l’esercizio appropriato in un’altra disciplina di tipo scientifico. La scelta è caduta sulla relazione di un esperimento inserito nel quadro di un corso di attività scientifiche sperimentali.

Esercizi scelti

*Lingua materna* - Si è preferito, alla tradizionale stesura, il testo incompleto nella parte centrale. Si presentano all’alunno gli elementi di un romanzo poliziesco: un riassunto dell’inizio, poi due paragrafi in cui deve inserirsi la composizione dell’alunno.

È stato preferito questo esercizio (2) perché le costrizioni imposte dal riassunto e dai paragrafi di orientamento forniscono punti di riferimento oggettivi per la valutazione, lasciando contemporaneamente sufficiente spazio alla spontaneità dell’alunno. Abbiamo incluso, nell’appendice al volume, il testo presentato agli alunni.

*Scienze* - Alla presenza degli alunni, il professore conduce un esperimento sulla durezza dell’acqua, fornendo soltanto le indicazioni sui prodotti utilizzati. Si chiede all’alunno di riferire, per iscritto, i fenomeni osservati, di rappresentarli con uno schema, di interpretarli e di inventare un metodo per misurare la durezza dell’acqua. Anche questo esperimento è inserito nell’appendice al volume.

A titolo indicativo, presenteremo gli esercizi scelti per il secondo ciclo dell’insegnamento secondario rinnovato. Per la lingua gli alunni sono stati invitati a riassumere e commentare un brano tratta da «Proposte per l’educazione» di Alain. Per le scienze, è stato proposto, nuovamente, di osservare un esperimento (spostamento di alogeni con altri alogeni), di interpretarlo e di trovare un metodo d’analisi.

(1) Ricerche bibliografiche e numerose sedute di lavoro tra ispettori e insegnanti hanno condotto all’elaborazione di un esercizio di matematica che sollecita l’alunno a performances complesse. Presentava, tuttavia, il grave inconveniente di essere completamente estraneo al contesto scolastico abituale.

(2) Proposto da G. Leroy.
Capitolo terzo

Metodo di costruzione di scale descrittive di valutazione*

Costruzione iniziale dello strumento

Le scale di valutazione (rating scales) sono state al centro di numerose ricerche sperimentali, rivolte essenzialmente alla loro utilizzazione. Al contrario, la letteratura sulla metodologia della loro costruzione è estremamente povera. Le regole o, almeno, le raccomandazioni o i consigli che verranno proposti derivano dalle difficoltà incontrate nell’impiego e nella sperimentazione delle scale esistenti.

La metodologia di costruzione, tratteggiata in seguito, è la risultante di numerose prove e tentativi. Non ha la pretesa di essere esaustiva e, in linea di massima, si suddivide in cinque tappe:

- selezione e definizione dei criteri di valutazione;
- selezione e definizione degli aspetti interni a ciascun criterio;
- determinazione dei livelli richiesti per ciascun aspetto;
- descrizione di vari livelli;
- presentazione della scala e stesura delle norme per l’utilizzazione.

* Prima di affrontare il presente capitolo, relativo alla metodologia della costruzione di scale descrittive, è opportuno leggere uno dei dosiers dell’appendice. In effetti, l’esposizione metodologica è illustrata con esempi estratti da questi protocolli che risulteranno più chiari se inseriti nel contesto generale. Convinti del valore dell’esercizio individuale, abbiamo aggiunto ai dosiers del primo ciclo la risposta di un alluno. Così il lettore ha immediatamente l’occasione di esercitarsi. Per molti, il procedimento proposto rompe con le tradizionali abitudini professionali e può sembrare, a prima vista, assai laborioso. L’esperienza mostra che, dopo alcune correzioni, si instaurano nuove abitudini tendenti ad alleggerire notevolmente il lavoro.

Tutte le tappe sono descritte e spiegate con esempi tratti da dosier allegati in appendice: uno di lingua materna e l’altro di scienze (primo ciclo dell’insegnamento rinnovato).

Prima tappa: Selezione e definizione dei criteri di valutazione

A. Un piccolo numero di criteri

Si potrebbe credere che la valutazione di una performance complessa obblighi a prendere in considerazione numerosi criteri. Non è consigliabile agire in tal modo, perché la precisione desiderata e le distinzioni sottili che l’accompagnano sono spesso illusorie.

Numerose esperienze dimostrano che correttore non avvertiti credono spesso di tener presenti un gran numero di criteri, mentre esaminano, senza accorgersene, sempre gli stessi, chiaramondi per con nomi diversi. La ricerca di C. Remondino (1) illustra chiaramente questo fenomeno. Remondino ha invitato venti insegnanti di lingua materna a stilare la lista delle caratteristiche tenute presenti nel correggere le composizioni (2). Dall’analisi fattoriale si è visto che le diciassette caratteristiche evidenziate si riducono di fatto a quattro: presentazione grafica, uso della lingua, contenuto e ordinizzazione delle idee, originalità.

In ogni classificazione, di cui la valutazione è uno degli aspetti, le categorie scelte devono rispondere all’esigenza fondamentale dell’esclusione reciproca (una stessa osservazione non deve poter entrare in due categorie contemporaneamente). È difficile rispettare questa esigenza e non è altrettanto facile rilevare le trasgressioni. Così, in una scala proposta dall’Educational Testing Service (3), sulla votazione dei compimenti in lingua materna, appaiono alcune sovrapposizioni nelle definizione di due categorie.

La descrizione del livello medio della scala “Idee o contenuto” penalizza la presenza di idee convenzionali: “Le idee sono appropriate, ma convenzionali e scarse...”. Peraltro, in una scala di valutazione della creatività, il livello medio è così definito: “Lo studente affronta gli argomenti pressappoco come fanno tutti...”.

Pretendere di valutare tutto negli esercizi svolti dagli alunni sarebbe al tempo stesso poco economico e poco conveniente: poco economico perché il professore dovrebbe impiegare un tempo eccessivo nella correzione dei lavori; poco conveniente perché un alunno non può dedicarsi contemporaneamente a colmare numerose lacune.

Molti studi evidenziano inoltre che i correttori, valutando gli elaborati degli studenti senza averne precedentemente specificati
gli obiettivi, prendono in considerazione quasi sempre gli stessi criteri (1). Talune caratteristiche rischiano così di non essere mai prese in considerazione.

Ma allora, come scegliere i criteri opportuni per valutare un lavoro? Il metodo da noi scelto consiste nel consultare dei professori che insegnano la disciplina ad un livello pedagogico stabilito. Costoro, riuniti in più gruppi, lavorano simultaneamente e discutono sul criteri da seguire. Questo metodo consente di trovare rapidamente gli aspetti essenziali da prendere in considerazione.

Il metodo proposto da Abeles adotta questo procedimento (4). Abeles invita gli insegnanti a descrivere nel modo più preciso possibile una performance dell'alunno, che viene loro proposta. Le descrizioni così ottenute sono sottoposte ad un'analisi del contenuto, che spesso fornisce una lunga lista di criteri. Per ciascuno di questi criteri gli insegnanti sono invitati a valutare i lavori degli studenti utilizzando una scala a cinque livelli. I risultati vengono sottoposti ad una analisi fattoriale per scoprire i principali criteri di valutazione.

Pur ridotte alla loro più semplice espressione, le caratteristiche del lavoro di un alunno non rivestono tutta la stessa importanza. Ad esempio, la presentazione grafica di un componimento è meno importante del suo contenuto. Si può privilegiare un processo intellettivo rispetto ad altri, accordando, ad esempio, più attenzione alla capacità di sintetizzare le idee che non alla semplice accumulazione di fatti dispersi attinti nella memoria. Si può sottolineare l'importanza di un particolare procedimento dedicando ad esso un maggior numero di scale di valutazione rispetto ad altri. Nel dossier allegato si vedrà che per valutare il rapporto scientifico richiesto nel primo ciclo dell'insegnamento secondario, è stato dedicato un usuale numero di scale ai processi mnemonici (relazione dell'esperimento con il testo e lo schema) che ai processi superiori (interpretazione e invenzione). Invece, al livello di secondo ciclo, uno spazio maggiore viene riservato ai processi superiori; ad essi vengono dedicate due scale, mentre una sola alla relazione dell'esperimento.

B. Come definire i criteri?

Con questo metodo da noi seguito la definizione di un criterio, daprima piuttosto generale, si precisa progressivamente con la scelta dei vari aspetti e con la determinazione dei livelli richiesti. Guardiamo i seguenti esempi:

1. **Lingua materna: definizione dei criteri di valutazione**
   a. **Pertinenza semantica**: la composizione è pertinente dal punto di vista semantico se l'alunno sviluppa idee appropriate al contesto, vale a dire compatibili con il riassunto e con i paragrafi 1 e 3.
   b. **Pertinenza sintattica**: rispetto delle regole sintattiche richieste nei paragrafi 1 e 3.

2. **Vocabolario**: insieme di termini usati dall'alunno. Non si considera l'ortografia.
3. **Strutture grammaticali**: il criterio è qui dettato dalla sintassi del paragrafo redatto dall'alunno, indipendentemente dalle regole imposte nei paragrafi 1 e 3 (già valutate in b).
4. **Organizzazione delle idee nel paragrafo 2**: si valuta la struttura del lavoro, indipendentemente dalle regole richieste nei paragrafi 1 e 3 (già valutate in a).

---

**Seconda tappa: Selezione e definizione degli aspetti di ciascun criterio**

Illustreremo questa tappa con la **pertinenza semantica** in lingua materna, e con la **descrizione dell'esperimento** in scienze.

---

**Lingua materna: pertinenza semantica**

Possiamo valutare la pertinenza semantica con due domande:

- L'alunno ha introdotto idee che contraddicono le informazioni contenute nei testi messi a sua disposizione? Non è necessario utilizzare tutte le informazioni fornite; è importante non contraddirele.
- L'alunno integra le informazioni in un contesto più ampio?

Descriviamo adesso i due aspetti da valutare.

---

**I. Inserimento di idee non pertinenti**

Si ritengono non pertinenti le idee che contraddicono le informazioni addotte dal testo. La lettura di un numero rilevante di elaborati degli alunni ha consentito di definire la lista che servirà a valutare questo primo aspetto della pertinenza semantica.

- L'inizio del paragrafo 2, non colloca i fatti nel laboratorio (Ad esempio: spingere la porta, entrare, ..., sono azioni non pertinenti; sono già espresse nel paragrafo 1).
- Non c'è alcun riferimento allo "spettacolo insolito". Tale riferimento può limitarsi ad una esclamatione, ad una sola parola.
c. Nel paragrafo 2 il commissario è presente o è atteso.
d. L’aspetto e il contenuto del messaggio sono modificati (ad esempio, il pezzo di carta diventa una lettera, lunga e precisa).
e. Il ragazzo menziona ferite gravi, sinistri, catastrofi,...

2. Integrazione delle informazioni

La sola considerazione del numero di elementi non pertinenti non basta render conto del valore del lavoro dell’alunno. Si attesta questa considerazione valutando il modo in cui l’alunno riferisce i fatti integrandoli in un contesto più ampio.

a. Si potranno rinvenire indicazioni di questa integrazione:— nella evocazione della personalità dello zio;— dai riferimenti a personaggi secondari;— nell’espressione dei sentimenti di Luca...
(b. Invece l’alunno si limita a riferire i fatti collocati tra l’entrata di Luca nel laboratorio (fine del primo paragrafo) e la risposta di Luca al commissario (inizio del terzo paragrafo) l’integrazione delle informazioni non si è avverata.

c. In alcuni elaborati sono espresse idee che, pur non essendo in contraddizione con i paragrafi 1 e 3, non hanno tuttavia alcun rapporto con la situazione.

— Scienze: descrizione dell’esperimento
Tre aspetti consentono di valutare la relazione scritta dell’esperimento:
— La completezza dell’elaborato
— La precisione delle informazioni
— Il rispetto dell’ordine cronologico.

A. Una descrizione è completa quando comporta i seguenti elementi:
A.1. Quantità uguali di volumi di acqua distillata in partenza
A.2. Quantità diverse di sale di calcio aggiunto ai due volumi di acqua.
A.4. Dosi differenti di saponina che provocano l’apparizione di una schiuma persistente.

B. Per essere precisa la descrizione deve contenere le seguenti informazioni:
B.1. Valori relativi delle quantità di sale Ca. (Es.: semplice-doppio, 1 dose - 2 dosi, 1 gr. - 2 gr.)
B.2. Valori relativi delle dosi di saponina (Es.: 6 misure - 13 misure).

C. L’ordine cronologico dei fenomeni è rispettato quando l’alunno presenta i fatti secondo il loro svolgimento.

Generalmente, la scelta definitiva degli aspetti di ciascun criterio dipende da precisazioni successive, associate alla correzione di numerosi elaborati. In un primo momento, per valutare la pertinenza semantica, i professori valorizzano l’introduzione dei maggior numero di riferimenti ai paragrafi di inquadramento. Ma la correzione degli elaborati ha evidenziato che la loro qualità non è necessariamente proporzionale alla quantità di riferimenti introdotti. È più significativo allora rimarcare la comparsa di idee non pertinenti.

Mentre i criteri debbono essere indipendenti, gli aspetti di un criterio sono necessariamente in relazione.

Nelle scale presentate, la valutazione secondo i diversi criteri risulta sia dall’intervento di diversi criteri, che agiscono come filtri successivi, sia dall’aggiunta di elementi giustapposti. Ad esempio, per la “pertinenza semantica” si valuta l’inserimento dei soli elementi pertinenti dell’elaborato. È possibile così articolare la valutazione condotta su elaborati equivalenti per il primo aspetto. Nelle scienze, lo stesso rapporto unisce gli aspetti della “completezza” e della “precisione”: dapprima si individuano gli elementi significativi della descrizione, poi se ne valuta la precisione con cui tali aspetti sono riferiti.

Il rispetto della cronologia dei fenomeni, al contrario, è un criterio aggiunto ai precedenti, indipendentemente dalla valutazione condotta su questi.

Per consentire un elevato grado di accordo, addirittura l’unanimità tra quanti valutano, la definizione degli aspetti deve risultare praticabile oltre che possibile.

Fin qui ci siamo limitati a presentare i criteri già stabiliti. Questo storico non è sufficiente da solo ad ottimizzare la valutazione. Bisogna specificare ulteriormente i comportamenti che più chiaramente traducono questi criteri.

Come riconoscere i comportamenti-criteri

I comportamenti che concretizzano gli aspetti da valutare entrano in un universo finito o non finito.

Nel primo caso (universo finito), i limiti imposti dagli obiettivi dell’esercizio e precisati con la formulazione della domanda permettono la stesura di un elenco esaustivo dei comportamenti possibili. Ad esempio: Elementi non pertinenti in lingua materna = Completezza e precisione nelle scienze.

Nel secondo caso (universo non finito), è impossibile preparare l’elenco completo dei comportamenti desiderati, non potendo prevedere cosa produrrà l’alunno. Lo stesso dica, ad esempio, per valutare la correzione, precisione, originalità del vocabolario e delle strutture grammaticali.

Spetterà pertanto al professore elencare, per ciascuna produzione da valutare, le manifestazioni caratteristiche dell’aspetto pre-
so in esame. Questa situazione è evidentemente meno favorevole della prima ai fini di una oggettivazione completa dei criteri e rende inutile sperare in una unanimità dei giudici. La presentazione di un elenco, pur non esaustivo, con esempi di comportamenti ritenuti adeguati, può costituire un punto di riferimento preciso e favorire maggiori possibilità di accordo. È quanto abbiamo fatto nei dossiers allegati a questa pubblicazione.

**Terza tappa: Determinazione dei livelli richiesti per ciascun aspetto**

Questo procedimento comporta due fasi:

- Classificazione dei possibili comportamenti in categorie ordinate gerarchicamente;
- Scelta dei livelli adeguati ad una età prestabilita.

**Come classificare i comportamenti possibili in categorie gerarchizzate**

Negli universi comportamentali finiti, le caratteristiche della materia conducono ad una scelta. O si considerano tutti i comportamenti come unità equivalenti, o si accorda loro un'importanza differenziata.

Nel primo caso, le distinzione tra i gradi sono di natura quantitativa e la loro gerarchia dipende dalla frequenza della comparsa. Per esempio, l'aspetto «Inserimento di elementi non pertinenti», nella scala «Pertinenza semantica», rientra in questa categoria. In effetti, con le regole imposte dai parametri di inquadramento si è potuto stabilire un elenco esaustivo dei possibili elementi non pertinenti. Essi sono stati considerati di uguale importanza.

Nel secondo caso, con comportamenti valutati differenziamente, le distinzioni tra i gradi assumono carattere qualitativo e la distinzione gerarchica si fonda sull'importanza specifica dei comportamenti rispetto agli obiettivi del lavoro. La definizione della completezza nella scala «Relazione dell'esperienza mediante una descrizione scritta» (Scienze) rientra in questa categoria. Le caratteristiche dell'esperienza proposta hanno consentito di individuare una lista di quattro elementi che vanno poi considerati nella relazione. Si accorda maggiore importanza alla resa di elementi codificati A.2 e A.4. La descrizione dei gradi evidenzia chiaramente questo tipo di valutazione.

Negli universi comportamentali non finiti, le unità elencate sono spesso considerate come equivalenti. I gradi si differenziano sul piano quantitativo e la loro gerarchia si fonda sulla frequenza di apparizione. Ad esempio, il giudizio sulla correttezza e precisione del vocabolario passa attraverso l'identificazione del numero di errori e di imprecisioni osservati nell'elaborato.

**Osservazione**

Per evitare una enumerazione fastidiosa, abbiamo ridotto talvolta il numero delle categorie a semplici dicotomie o tricotomie, vicine alla parafrasi. Non è comunque una soluzione molto soddisfacente: se ne vedranno gli inconvenienti nel capitolo dedicato all'analisi. Gli esempi che seguono sono stati ripresi dalla colonna 4 della tabella.

1. **Universo finito:** Tutti i comportamenti hanno uguale importanza (Scienze — 2° ciclo — Relazione dell'esperimento).
   - Per il primo aspetto, «Relazione dei fenomeni essenziali», si considerano cinque livelli:
     1. Nessun errore o omissione
     2. Un errore o una omissione
     3. Due errori o due omissioni
     4. Più di due errori o omissioni
     5. Nessun riferimento all'una delle due fasi.

2. **Universo finito:** Valutazione differenziata a seconda dei comportamenti (Scienze — 1° ciclo — Relazione dell'esperimento con una descrizione scritta).
   - Per l'aspetto «completezza» si considerano tre livelli:
     2. Sono presenti almeno le informazioni A2 e A4.

3. **Universo non finito:** Uguaglianza accordata a tutti i comportamenti (Lingua materna — 1° ciclo — Vocabolario).
   - I livelli considerati per l'aspetto «correttezza» sono:
     1. Nessun errore.
     2. Uno o due errori.
     3. Più di due errori.

4. **Universo non finito:** Dicotomie o tricotomie (Lingua materna — 1° ciclo Pertinenza semantica)
   - Per l'aspetto «integrazione» si considerano tre livelli:
     1. Integrazione.
     2. Nessuna integrazione.
     3. Idea senza alcun rapporto con il contesto.
### Lingua materna

<table>
<thead>
<tr>
<th>Universo finito</th>
<th>Universo non finito</th>
</tr>
</thead>
<tbody>
<tr>
<td>Uguale importanza di tutti i comportamenti; Valutazione differenziata a seconda dei comportamenti; Livelli differenziati per quantità tra i gradi</td>
<td>Uguale importanza di tutti i comportamenti; Vedere osservazione a p. 73</td>
</tr>
<tr>
<td>Livelli differenziati per quantità tra i gradi</td>
<td>Livelli differenziati per quantità</td>
</tr>
</tbody>
</table>

#### 1° ciclo

<table>
<thead>
<tr>
<th>Pertinenza semantica</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inserimento di elementi non pertinenti</td>
<td></td>
<td></td>
<td>Integrazione delle informazioni</td>
<td></td>
</tr>
<tr>
<td>Pertinenza sintattica</td>
<td></td>
<td></td>
<td>Rispetto dei tempi; Rispetto delle persone in grammatica</td>
<td></td>
</tr>
<tr>
<td>Vocabolario</td>
<td></td>
<td></td>
<td>Correttezza; Precisione</td>
<td></td>
</tr>
<tr>
<td>Strutture grammaticali</td>
<td></td>
<td></td>
<td>Originalità</td>
<td></td>
</tr>
<tr>
<td>Struttura del paragrafo 2</td>
<td></td>
<td></td>
<td>— Coerenza semantica; — Collegamento di idee; — Progressione del racconto</td>
<td></td>
</tr>
</tbody>
</table>

#### 2° ciclo

<table>
<thead>
<tr>
<th>Riassunto</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ridondanza; idee secondarie</td>
<td>Comprensione delle idee principali</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Critica</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Quantity e correttezza delle idee riprese; Validità e tipo di argomenti</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Scienze

<table>
<thead>
<tr>
<th>Universo finito</th>
<th>Universo non finito</th>
</tr>
</thead>
<tbody>
<tr>
<td>Uguale importanza di tutti i comportamenti; Valutazione differenziata a seconda dei comportamenti; Livelli differenziati per quantità</td>
<td>Uguale importanza di tutti i comportamenti; Vedere osservazione a p. 73</td>
</tr>
<tr>
<td>Livelli differenziati per quantità</td>
<td>Livelli differenziati per quantità</td>
</tr>
</tbody>
</table>

#### 1° ciclo

<table>
<thead>
<tr>
<th>Relazione dell’esperimento con una descrizione scritta</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Precisione</td>
<td>Completezza</td>
<td></td>
<td>Rispetto dell’ordine cronologico</td>
<td></td>
</tr>
<tr>
<td>Relazione dell’esperimento con uno schema</td>
<td></td>
<td>Completezza</td>
<td></td>
<td>Dinamicità</td>
</tr>
<tr>
<td>Interpretazione</td>
<td>Completezza</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Invenzione di un metodo di misura</td>
<td>Invenzione</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

#### 2° ciclo

| Relazionare l’esperimento | | | |
| | — Relazione dei fenomeni essenziali; — Distinzione • Localizzazione • Ripetizione |
| Interpretazione dei fenomeni osservati | | | |
| | — Classificazione degli allogeni — Giustificazione della classificazione |
| Inventare un metodo di analisi | Invenzione | | |
Come scegliere i livelli

Per selezionare le richieste corrispondenti ad un determinato livello pedagogico si esaminano numerosi lavori di alunni che rientrano sia nel livello pedagogico prescelto, sia in quelli immediatamente inferiore e superiore (\(^1\)). Così, riguardo all’esperienza condotta nel primo anno di insegnamento rinnovato, abbiamo raccolto i lavori del sesto anno di scuola primaria e del secondo anno dell’insegnamento secondario rinnovato.

Le informazioni raccolte permettono:

— di precisare, nella gamma dei comportamenti caratteristici di un aspetto, le richieste minimali e massimali per un livello pedagogico prestabilito;
— di riunire eventualmente i comportamenti compresi entro questi limiti.

I due procedimenti combinati assicureranno un’andatura gaussiana nella distribuzione delle valutazioni, come illustrato dallo schema:

![Schema di confronto tra 6° A. Prim., 1° A. Insegnamento sec., 2° A. Ins. sec.](image)

Si sceglie un criterio di valutazione per la sua corrispondenza ad un obiettivo educativo ritenuto importante. Con apprendimenti complessi, quali la capacità di esprimersi chiaramente per iscritto, l’obiettivo aumenta sempre più di complessità, nella misura in cui progresdicono gli studi e si innalza parallelamente il livello richiesto dal valutatore.

La figura a p. 76 schematizza l’evolversi della valutazione. Il quadro a tratto continuo indica i comportamenti considerati nel primo anno di insegnamento rinnovato.

Un adattamento erroneo delle richieste al livello degli alunni porterebbe ad una scala scarsamente discriminativa. Questa risulterebbe al tempo stesso troppo facile (infatti, la maggioranza sarebbe concentrata ai livelli superiori), e troppo difficile (concentrandosi la maggior parte dei comportamenti ai gradi inferiori).

Quarta tappa: Descrizione dei livelli

Nell’intento di precisare meglio le componenti della valutazione, abbiamo adottato fino ad ora un procedimento analitico: dopo aver selezionato i criteri, si è cercato di determinare gli aspetti esenziali di ciascuno di essi e di precisare, in ogni caso, i tipi di comportamento da prendere in considerazione. Si è anche osservato che i differenti aspetti di uno stesso criterio si completano a vicenda: dalla loro combinazione deriva una votazione significativa.

La quantità di aspetti ed i livelli richiesti determinano l’elevato numero di combinazioni. Si può tuttavia decidere di considerare equivalenti alcune combinazioni, giungendo così ad una medesima votazione. Questa decisione presuppone tuttavia che sia stata definita l’importanza relativa dei diversi aspetti.

Se invece si stabilisce di accordare eguale peso a tutti gli aspetti, tutti i livelli di ciascuno di essi sarebbero compresi nella scala ingrandendosi perfettamente. È il caso dei primi tre livelli della scala «Organizzazione delle idee» (1° ciclo).

<table>
<thead>
<tr>
<th>Coerenza semantica</th>
<th>Collegamento di idea</th>
<th>Progressione del racconto</th>
<th>Livello della scala</th>
</tr>
</thead>
<tbody>
<tr>
<td>— Racconto coerente</td>
<td>— Collegamento adeguato</td>
<td>— Racconto ridondante...</td>
<td>MB</td>
</tr>
<tr>
<td></td>
<td>— Collegamento non adeguato</td>
<td>— Progressione rallentata</td>
<td>P</td>
</tr>
<tr>
<td>— Contraddizioni limitate</td>
<td>— Collegamento adeguato</td>
<td>— Racconto ridondante...</td>
<td>B</td>
</tr>
<tr>
<td></td>
<td>— Collegamento non adeguato</td>
<td>— Progressione rallentata</td>
<td>S</td>
</tr>
<tr>
<td>— Racconto non coerente</td>
<td>— Collegamento adeguato</td>
<td>— Racconto non ridondante...</td>
<td>D</td>
</tr>
<tr>
<td></td>
<td>— Collegamento non adeguato</td>
<td>— Progressione rallentata</td>
<td>1</td>
</tr>
<tr>
<td>Da osservare che il voto B (Bene) viene dato in tre casi:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>— Racconto coerente</td>
<td>— Collegamento adeguato</td>
<td>— Progressione rallentata</td>
<td>1</td>
</tr>
<tr>
<td>— Racconto coerente</td>
<td>— Collegamento non adeguato</td>
<td>— Racconto non ridondante</td>
<td>1</td>
</tr>
<tr>
<td>— Contraddizioni limitate</td>
<td>— Collegamento adeguato</td>
<td>— Racconto non ridondante</td>
<td>1</td>
</tr>
</tbody>
</table>
In tal modo si penalizza una lacuna nello stesso modo in ciascuno dei tre aspetti. Qualora si voglia attribuire minor importanza ad un aspetto, si costruirà la scala in modo tale che secol aspetto interverrà solo per meglio esprimere i primi livelli di un altro aspetto.

Esempio: scala «Descrizione scritta dell'esperimento» (1° ciclo)

<table>
<thead>
<tr>
<th>Completezza</th>
<th>Precisione</th>
<th>Rispetto della cronologia</th>
</tr>
</thead>
<tbody>
<tr>
<td>— B1. o B2.</td>
<td>— No .......................... B</td>
<td></td>
</tr>
<tr>
<td>— No .......................... S</td>
<td></td>
<td></td>
</tr>
<tr>
<td>— No .......................... D</td>
<td></td>
<td></td>
</tr>
<tr>
<td>— B1. e B2.</td>
<td>— Si .......................... B</td>
<td></td>
</tr>
<tr>
<td>— No .......................... S</td>
<td></td>
<td></td>
</tr>
<tr>
<td>— Si .......................... D</td>
<td></td>
<td></td>
</tr>
<tr>
<td>— No .......................... I</td>
<td></td>
<td></td>
</tr>
<tr>
<td>— B1. e B2.</td>
<td>— Si .......................... B</td>
<td></td>
</tr>
<tr>
<td>— No .......................... S</td>
<td></td>
<td></td>
</tr>
<tr>
<td>— Si .......................... D</td>
<td></td>
<td></td>
</tr>
<tr>
<td>— No .......................... I</td>
<td></td>
<td></td>
</tr>
<tr>
<td>— B1. e B2.</td>
<td>— Indipendentemente dalla precisione e dal l'ordine cronologico</td>
<td></td>
</tr>
</tbody>
</table>

L'esame applicato al solo livello inferiore (idee non rapportate al contesto) potrebbe estendersi agli altri livelli, penalizzando così le lacune del secondo aspetto.

**Quinta tappa:** Presentazione della scala e procedimento per utilizzarla.

La tappa precedente ha consentito di enunciare le regole che presiedono alla combinazione dei criteri stabiliti per ciascun aspetto della valutazione. Non resta che presentare l'insieme dei criteri nella forma più conveniente. Sono stati presi in considerazione due tipi di procedimenti.

Inizialmente ci si attiene al procedimento cronologico adottato da chi valuta. In una prima parte del dossier si affronta sul piano teorico, le informazioni relative alle tappe fin qui descritte e si giustificano i criteri. Per agevolare la valutazione, vengono riconsiderate sistematicamente le informazioni essenziali delle quali il professore dovrà tener conto per analizzare la performance dell'alunno: per ciascun aspetto, si esamina il livello richiesto, dal più elevato al più basso; si procede poi alla sintesi delle caratteristiche rinvenute. Esempi relativi a questo procedimento corredano le scale riportate negli allegati.

Nella seconda parte, priva di informazioni nuove, i criteri sono organizzati in modo differente, per evidenziare la gerarchia di comportamenti relativi a ciascuna dimensione. Le caratteristiche, che portano ad una stessa votazione, sono state riunite in una tabella, permettendo maggior completezza di informazione all'alunno sul valore della performance: indica le tappe, ancora da percorrere, per dominare comportamenti individuati.

La scala di valutazione è operativa se evita due scoglitori:
— la parafrasia pura e semplice delle categorie o la definizione dei criteri in termini esclusivamente mentali;
— la ricerca di uno strumento totalmente oggettivo, ma impraticabile nella vita scolastica quotidiana.

Illustriamo la prima difficoltà ricorrendo ad un esempio già citato, la scala di valutazione delle composizioni costruita dall'Educational Testing Service di Princeton (C).

Ecco i livelli considerati per valutare le idee contenute nelle composizioni:

**Idee (contenuto)**
— Livello elevato: L'argomento della composizione o il piano di lavoro sono affrontati in tutti gli aspetti. L'alunno comprende bene l'argomento e utilizza definizioni chiare. Sa considerare l'argomento in una prospettiva più
Revisione dello strumento

Ricerca dei vizi di costruzione e di utilizzazione

L'applicazione del metodo qui sopra descritto ha condotto ad una prima stesura degli strumenti. Ne è seguito un controllo sulla validità, chiedendosi: «Le scale descrittive permettono una valutazione più oggettiva e articolata delle altre usate abitualmente?». Una risposta esatta può derivare soltanto da un esperimento effettuato in condizioni rigidamente controllate.

Inizialmente sono stati chiamati a rispondere alcuni professori, per evitare, prima della sperimentazione massiccia, le ambiguità più gravi e gli errori più evidenti. Lasciando al prossimo capitolo i risultati dettagliati dell'esperimento, facciamo presente, per ora, che essi sono been longi dal riflettere una situazione perfetta. La valutazione discorde tra quanti valutano può dipendere da due ipotesi: errori di lettura e di utilizzazione, scale costruite erroneamente.

Metodo

In questa fase della ricerca, si ricorre al metodo clinico per scoprire il procedimento di valutazione adottato, in funzione dei criteri stabiliti.

Generalmente torna utile la tecnica della riflessoiva parlata che viene così definita da G. De Landsheere: «Il soggetto è invitato a formulare ad alta voce i procedimenti del suo pensiero mentre sta per risolvere un problema; in questo modo è possibile studiare i processi mentali nel vivo del loro applicarsi e, quindi, identificare le cause dei successi e degli errori» (7).

La richiesta di informazioni di carattere diagnostico, derivate da questo procedimento analitico, in verità assai laborioso, offre il vantaggio di un minor numero di professori da intervistare. In un primo tempo, ciascun professore, che si apprestava all'esperimento, è stato invitato a valutare il lavoro presentato, esprimendo ad alta voce i dubbi, le scelte e le motivazioni connesse.

Rivolgendoci soltanto ai professori del gruppo sperimentale, abbiamo superato le difficoltà di integrazione di tutti i criteri proposti nel dossier.

In un secondo tempo, si è proceduto a interviste più strutturate nell'intento di ottenere una risposta esaustiva alle domande suggerite dall'analisi dei risultati. Riportiamo alcuni esempi di domande poste (scienze):

— Quale o quali frasi nell'elaborato dell'alunno, vi fanno eventualmente pensare che uno o più metodi di valutazione sono stati ben ideati?
— Alcuni insegnanti (44%) hanno dato come voto B. Perché, secondo voi? Qual è, secondo voi, la differenza tra MB e B?
— Da che cosa deducete che l'alunno è o non è capace di distinguere le due fasti dell'esperimento?

Per l'attribuzione del giudizio globale la domanda è di tipo aper-
to. Il professore è invitato semplicemente ad effettuare la valutazio-
ne; non gli viene suggerito di riferirsi alle valutazioni analitiche
precedenti. Al termine della valutazione complessiva deve, tuttavia,
spiegare il metodo utilizzato:
1. Quale procedimento utilizzate per passare dalle valutazioni analitiche a
quella globale?
2. Attribuite egual importanza ai due criteri? SI-NO.
3. In caso di risposta negativa, la vostra decisione è subordinata:
   — all'importanza di ciascun criterio in sé?
   — al grado di difficoltà che i diversi aspetti considerati presentano per gli
     alunni del livello scolastico nel quale insegnate?

Spoglio

La schedatura delle informazioni raccolte prevede due tipi di osservazioni:
— I comportamenti del professore di fronte allo strumento di valu-
tazione;
— I difetti di costruzione nelle scale.

Risultati delle interviste

1° Errori di utilizzazione

a. Errori di manipolazione

Il comportamento di alcuni insegnanti evidenzia il desiderio di liberarsi
dalle regole proposte, per valutare in modo più autonomo:
— assegnare un voto, per ciascun criterio, senza utilizzare la scala;
— per ciascun criterio, assegnare empiricamente un voto, giustificandolo a
   posteriori con quella descrizione della scala che più si addice;
— affidarsi alla descrizione degli aspetti, e non alla valutazione degli stessi
   secondo le regole stabilite, per farsi un'opinione dell'elaborato;
— procedere alla valutazione di un criterio tenendo presente un solo aspet-
to di esso;
— utilizzare la scala senza aver letto la descrizione degli aspetti;
— introdurre criteri estranei al dossier;
— introdurre criteri che differenziano il voto in funzione del periodo scolastico
   in cui è stato proposto il lavoro agli alunni, o del livello della classe
   cui l'insegnante si rivolge.

— introdurre criteri relativi alla materia: o aggiungendo criteri nuovi (orto-
grafia, punteggiatura), o aspecti nuovi nei criteri già dati (originalità, nel-
la pertinenza semantica);

b. Errori di lettura

Oggetti errori riguardano tanto la comprensione dei singoli criteri, che la
loro combinazione nella scala.

Ad esempio, nella scala «Pertinenza semantica», alcuni insegnanti inter-
pretano male le indicazioni relative all'introduzione di elementi non per-
tenenti. Segnano l'assenza di alcuni elementi presentati nei paragrafi di ri-
ferrimento e non rivelano le contraddizioni riportate nell'elenco.

Oppure, alcuni professori di scienze evidenziano nella «Relazione
dell'esperimento con una descrizione scritta», le caratteristiche A2 e A4, B1 e
B2, e considerano sostanzialmente rispettata la cronologia. Così attribui-
scono la valutazione B (Buono), mentre la combinazione delle predette qua-
lità dovrebbe condurre a S (Sufficiente).

L'aver rilevato errori nell'applicazione e nella lettura non giusti-
ifica un atteggiamento pessimistico. Bisogna pur considerare le con-
dizioni in cui si è svolta la ricerca.

Almeno nella prima versione, i dossiers erano abbastanza comp-
plessi per chi non aveva partecipato all'elaborazione. In effetti, i
criteri proposti sono numerosi e i professori non li utilizzano che
una volta.

Questa situazione artificiosa può indurli a ritenere che lo stru-
mento è di applicazione troppo laboriosa per la valutazione dell'in-
sieme degli elaborati di una classe. In più, bisognerebbe precisare
ulteriormente gli obiettivi perseguiti con questo esperimento. Si
tratterebbe di sottoporre a test, con un gruppo di insegnanti, l'effi-
cacia dei criteri stabiliti.

Da parte loro, i docenti cui è chiesta la collaborazione, dovreb-bero conformarsi rigidamente alle indicazioni proposte per la cor-
rezione, anche se queste divergono dalle loro scelte. Anzi, cosi
torso dovrebbero precisare quali altri criteri intendono fissare per valu-
tare la performance.

2° Difetti di costruzione delle scale

Si potevano evitare molti difetti, qui appresso rilevati, precisan-
dodo rigidamente e fin dall'inizio il metodo descritto. Come già detto,
le regole essenziali da rispettare e le precauzioni da prendere sono
conseguenziali ai risultati dell'esperimento e alla loro analisi. Tra i
difetti di interesse metodologico, si nota:
— In alcuni casi le differenti categorie non erano esclusive. La prima ver-
sezione della scala «Vocabolario» si presentava come indicato nella tabella
della pagina seguente.
<table>
<thead>
<tr>
<th>Correttezza</th>
<th>Precisione</th>
<th>Originalità</th>
</tr>
</thead>
<tbody>
<tr>
<td>MB</td>
<td>corretto</td>
<td>preciso</td>
</tr>
<tr>
<td>B</td>
<td>il più delle volte corretto</td>
<td>preciso</td>
</tr>
<tr>
<td>S</td>
<td>generalmente corretto</td>
<td>impreciso</td>
</tr>
<tr>
<td>D</td>
<td>generalmente corretto</td>
<td>generalmente preciso</td>
</tr>
<tr>
<td>I</td>
<td>spesso non corretto</td>
<td>impreciso</td>
</tr>
</tbody>
</table>

Si noterà che non è prevista la voce: Vocabolario corretto, impreciso e originale.

Altre formulazioni erano ambigue. Nella scala «Invenzione di un metodo di misura», la distinzione tra i livelli S e I non era sufficientemente chiara.

S: Il metodo è appropriato al contesto, ma presentato in modo impreciso e incoerente

I: Nessun suggerimento, oppure il metodo suggerito è inadeguato.

L'ambiguità è evidenziata dagli istogrammi di frequenza dei voti all'incirca l'80% dei professori distribuisce le valutazioni in maniera quasi uguale tra i due livelli.

La scala «Pertinenza semantica» presenta un altro caso di ambiguità. La definizione dell'aspetto «Integrazione» è stato preciso con due esempi. Molto suggerimenti li hanno considerati come elementi di una lista esaustiva, rilevandone la presenza o assenza negli elaborati degli alunni. In altri termini, l'esempio illustrativo del tratto da valutare ha assunto valore di criterio.

— Infine, si osserva una scala i cui livelli (sottocategorie) non si escludevano reciprocamente. Nella scala «Invenzione di un metodo di misura» due livelli non sono totalmente indipendenti:

S: Si indica un metodo appropriato al contesto, ma è presentato in modo impreciso e incoerente.

D: L'alunno indica il procedimento dell'evaporazione.

Alcuni lavori degli allievi hanno abbozzato un metodo appropriato al contesto la costruzione è impossibile precisare se le risposte di questo tipo vanno collocate in S o in D.

Come ridurre le discordanze tra i giudici

I difetti di costruzione rinvenuti nelle scale potranno essere evitati applicando il metodo proposto nelle pagine precedenti. Si cercherà di definire ciascun criterio in termini di comportamento richiesto all'alle (in particolare occorre stabilire esattamente i livelli) e in pari tempo sarà eliminata l'ambiguità dei termini troppo generici.

Si aggiungeranno regole sistematiche relative alla combinazione dei criteri precisi. Verranno così rispettate le esigenze fondamentali per l'analisi del contenuto: la completezza e l'indipendenza delle categorie.

Per evitare gli errori di lettura, si è cercato di:

1. precisare le disposizioni definendo sistematicamente criteri e livelli richiesti;
2. programmare le fasi della valutazione. L’insegnante è invitato a pronunciarsi su tutti gli aspetti; è obbligato, dalle scelte multiple, a tener presenti tutte le possibilità prospettate. L'obbligo di ripetere questo procedimento per tutti gli aspetti annula il rischio di tralasciarne alcuni.

L'insegnante deve combinare successivamente i suoi giudizi, prima di formulare una valutazione relativa a ciascun criterio. Anche questo procedimento è stato semplificato perché viene presentato sotto forma di schema.

Il carattere analitico di questo procedimento consente, tra l'altro, di ritrovare facilmente le cause delle discordanze tra coloro che giudicano. Da queste informazioni dovrebbero scaturire ulteriori miglioramenti.

**NOTE AL CAPITOLO TERZ**


(2) La ricerca, condotta su 20 insegnanti ha evidenziato le seguenti componenti: leggibilità; estetica; presentazione; correttezza ortografica; correttezza morfologica; correttezza sintattica; organicità di esposizione; ricchezza di idee; adeguatezza di idee; precisione dell'informazione; completezza; concisione; proprietà di linguaggio; stile; originalità; maturità; fantasia. In seguito 10 correttori, utilizzando queste componenti per valutare un saggio unico, hanno notato che le stesse tendevano a raggrupparsi attorno ai quattro fattori (citati in questa pagina) su cui è stata impostata l'analisi fattoriale. La ricerca si trova in C. Remondino, *Studio fattoriale sulla valutazione di composizioni scolastiche in lingua materna*, in «Bollettino di Psicologia e Sociologia applicata» 1958, n. 25-30, pp. 205-218 (n.d.t.).


(6) La validità dell'analisi è garantita dal campione rappresentativo e sufficiente della popolazione scolastica.

(7) E.T.S., *Definition of Rating Son the ETS Composition Scale*, op. cit., p. 3.


(9) G. De Landshere, *Introduzione alla ricerca in educazione*, cit., pag. 95.
Capitolo quarto

Risultati sperimentali

Il costruire uno strumento di misura applicando un metodo sperimentale lascia sperare in una soddisfacente validità. La speranza va comunque confermata dai fatti. Vale a dire, gli strumenti costruiti devono essere sperimentati. In questo capitolo si confrontano le principali indicazioni emerse da due esperimenti successivi. Nell’appendice vengono riportati i risultati dettagliati della sperimentazione.

Prima sperimentazione

Schema

Alcuni insegnanti sono stati invitati a valutare un solo elaborato dell’alunno. La situazione può sembrare artificiosa perché, nella pratica quotidiana, un insegnante valuta quasi sempre un gran numero di copie ed esprime un giudizio in parte relativo. Qui si è voluto concentrare al massimo l’attenzione sulla materia.

Sono stati contattati tutti gli insegnanti di lingua e scienze impegnati nel rinnovamento al momento della sperimentazione. Essi sono stati divisi in tre gruppi:

- Il primo gruppo ha valutato complessivamente l’elaborato dell’alunno servendosi di una semplice scala verbale a cinque gradi;
- Il secondo gruppo ha valutato lo stesso esercizio:
  - dapprima in modo analitico, con scale verbali riferite a più criteri precedentemente stabiliti;
  - poi globalmente, con una scala verbale a cinque gradi;
- Il terzo gruppo ha valutato lo stesso esercizio:
  - dapprima in modo analitico, con scale descrittive riferite agli stessi criteri;
  - poi globalmente, con una scala verbale a cinque gradi.

<table>
<thead>
<tr>
<th>Gruppo I (Gruppo di controllo)</th>
<th>Gruppo II (Gruppo di controllo)</th>
<th>Gruppo III (Gruppo di controllo)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Valutazioni analitiche</td>
<td>Valutazioni analitiche</td>
<td>Valutazioni analitiche</td>
</tr>
<tr>
<td>Criterio 1: MBBS DII</td>
<td>Criterio 1: MBBS DII</td>
<td>Criterio 1: MBBS DII</td>
</tr>
<tr>
<td>Criterio 2: MBBS DII</td>
<td>Criterio 2: MBBS DII</td>
<td>Criterio 2: MBBS DII</td>
</tr>
<tr>
<td>Criterio 3: MBBS DII</td>
<td>Criterio 3: MBBS DII</td>
<td>Criterio 3: MBBS DII</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ciascun grado è definito.</td>
</tr>
</tbody>
</table>

Valutazione globale MBBS DII  Valutazione globale MBBS DII  Valutazione globale MBBS DII

L’esperimento tende a chiarire tre questioni:

1. Il passaggio attraverso le valutazioni analitiche accresce la concordanza delle valutazioni globali condotte su una performance complessa? Il confronto dei voti globali espressi nei gruppi I e II fornirà elementi di risposta a questo interrogativo.

2. La definizione precisa di ciascun criterio accresce la concordanza delle valutazioni analitiche? Il confronto dei voti attribuiti nei gruppi II e III consentirà di rispondere al quesito.

3. Le giustificazioni dei voti sono più precise e rispondenti agli obiettivi se gli insegnanti utilizzano scale descrittive di valutazione?

I professori dei gruppi I e II sono stati invitati a giustificare le proprie valutazioni analitiche e/o globali. Attraverso l’analisi di queste giustificazioni, desideriamo provare l’ipotesi che la specificità dei criteri di valutazione utilizzati dall’insegnante, da una parte, e della specificità dell’informazione data all’alunno sul valore della sua performance, dall’altra.

Gli insegnanti del III gruppo, usando scale descrittive, non sono tenuti a giustificare le proprie valutazioni analitiche. In effetti, la scelta di un livello della scala rende conto di per sé con sufficiente precisione delle qualità del lavoro dell’alunno. Qui va giustificata solo la valutazione globale.
Campione

1. Criteri di campionamento

Si è inteso facilitare lo svolgimento della ricerca rinunciando i professori di uno stesso istituto nello stesso gruppo. Il campionamento si è svolto per scuole. Sono stati scelti criteri di campionamento dei quali si poteva supporre che influenzassero le modalità di valutazione:

— indirizzo di insegnamento seguito dalla scuola prima del rinnovamento: generale o tecnico;
— scuole di sei anni o scuola per il ciclo inferiore (questa caratteristica è legata alla proporzione dei professori o dei laureati che insegnano nel ciclo inferiore dell'istituto);
— grandezza dell'istituto (numero di alunni nella scuola).

Successivamente, sono stati considerati questi tre fattori per stabilire il campione. La stratificazione dei livelli si può rappresentare schematicamente così come segue:

<table>
<thead>
<tr>
<th>Insegnamento generale</th>
<th>Insegnamento tecnico</th>
</tr>
</thead>
<tbody>
<tr>
<td>Con il ciclo superiore</td>
<td>Con il solo ciclo inferiore</td>
</tr>
<tr>
<td>le scuole da 1 a n sono classificate per ordine di grandezza</td>
<td></td>
</tr>
<tr>
<td>n</td>
<td>n</td>
</tr>
</tbody>
</table>

In tal modo i tre gruppi così formati sono dunque equivalenti per quanto riguarda i tre criteri di campionamento presi in considerazione.

2. Grandezza del campione

Per il primo esperimento abbiamo preso contatto con 700 professori di lingua e 1050 di scienze (1), ripartiti tra le 132 scuole che avevano iniziato il rinnovamento al momento della ricerca. Per questo tipo di impresa la percentuale di risposte è stata elevata: 36% dei professori di lingua materna e 33% dei professori di scienze.

Seconda sperimentazione

Sulla base dei risultati ottenuti nel corso della prima sperimentazione si è proceduto ad alcune revisioni e sono stati modificati i criteri metodologici di costruzione dello strumento.

Gli strumenti rimaneggiati sono stati oggetto della seconda sperimentazione che, fortunatamente, è stata realizzata tardiamente nel corso dell'anno scolastico: le informazioni che ne sono scaturite inducono talvolta a delle riserve. Non potendo ottenere dall'analisi risultati convalidati da osservazioni condotte su un gran numero di soggetti, ci siamo limitati in molti casi a delineare le tendenze.

Per non appesantire ulteriormente una presentazione già gravosa, riportiamo soltanto i risultati relativi a ipotesi che vanno verificate.

Ipotesi da verificare

Sono state prese in esame tre ipotesi. Da criteri oggettivi di valutazione derivano:

— una valutazione più oggettiva (1) (ipotesi 1);
— una valutazione più valida (ipotesi 2);
— una valutazione meglio utilizzabile sia da parte del professore che dell'alunno (ipotesi 3).

Gli effetti delle scale descrittive saranno valutati attraverso due tipi di informazioni:

a. il procedimento analitico espresso dai giudizi relativi ai differenti criteri; verranno pure esaminate le osservazioni degli insegnanti che corredano il giudizio sintetico.

b. il giudizio sintetico attribuito all'elaborato.

Spoglio dei dati

Per poter procedere all'analisi statistica dei dati, abbiamo tradotto in cifre le valutazioni formulate:
Dopo aver valutato:
- i professori dei gruppi I e III devono giustificare la votazione complessiva;
- i professori del gruppo II devono solo giustificare la votazione complessiva, ma anche ciascuna valutazione analitica.

Le giustificazioni così ottenute sono state fatte oggetto di un'analisi del contenuto.
Per le giustificazioni delle valutazioni globali usiamo dapprima le categorie riportate nel dossier: pertinenza semantica, pertinenza sintattica, ... in lingua; per le scienze, le categorie sono: relazione dell'esperimento con una descrizione scritta o uno schema, interpretazione ...

Le giustificazioni che sfuggono a queste categorie, ma che tornano frequentemente, saranno ugualmente considerate. Ad esempio, per la lingua, si è tenuto conto dell'ortografia, della puntiggiatura, della coniugazione, dello stile... e, per le scienze, della qualità della respiration scritta, dell'osservazione, della comprensione dell'oggetto dell'esperimento ...

Per le valutazioni analitiche, sono stati considerati dapprima gli aspetti previsti nelle scale: «Numero di elementi non pertinenti» e «Integrazione», per la pertinenza semantica; «Completezza, precisione e cronologia della descrizione», per le scienze. Non abbiamo comunque dimenticato le altre giustificazioni proposte dai professori.

Infine, per ciascun professore che ha partecipato alla ricerca, sono state raccolte le seguenti informazioni:
- professore o laureato;
- tipo di scuola: ginnasio e liceo; scuola media o istituto tecnico;
- ubicazione della scuola;
- disciplina insegnata: lingua o scienze;
- specializzazione dei professori di scienze: matematica, fisica, chimica, geografia, biologia.

Risultati

Ipotesi I: La scala descrittiva permette una valutazione più oggettiva
Come verificare questa ipotesi?
Come è stato già sottolineato in diverse occasioni, si ritiene spesso, come indice di obiettività, la concordanza di valutazioni dedotte attraverso risultati differenti.
È stato chiesto ai professori che hanno collaborato alla ricerca di valutare le performances con una scala a cinque gradi ('). Se manca un criterio omogeneo di valutazione l'insieme dei voti si distribuiranno probabilmente secondo la curva di Gauss (o pressappoco). In caso contrario, se i giudici adottano gli stessi criteri di valutazione, può verificarsi una concentrazione di voti su uno stesso livello. Sfortunatamente, una simile concordanza si trova di rado nella realtà; tuttavia l'obiettivo di avvicinarsi ad essa è il più possibile.

L'indice statistico che testimonia una maggiore o minore dispersione di voti è lo scarto-tipo (σ). Per accertare l'efficacia dei criteri valutativi si confrontano dunque gli indici di dispersione dei giudizi analitici provenienti dai gruppi II e III. A titolo indicativo, viene pure confrontata la distribuzione dei voti globali nei tre gruppi.

Ma non è sufficiente evidenziare le differenze nella distribuzione dei voti: bisogna anche determinare se la differente distribuzione si spiega con oscillazioni dovute al caso o, al contrario, con variazioni significative delle condizioni in cui ha operato ciascun gruppo.

Inoltre, le informazioni ottenute dall'analisi degli indici di dispersione vanno complete con uno studio più approfondito della ripartizione dei voti. In effetti, i giudici potrebbero accordarsi su livelli diversi da quelli ai quali i criteri dovrebbero condurre. Così sono stati fissati istogrammi di frequenza per ogni tratto valutato, al fine di visualizzare meglio questa ripartizione.

Dopo il confronto degli indici di dispersione per ciascun criterio, si confronta la serie, lo schema (pattern) dei voti analitici espressi da un giudice col profilo richiesto dalla scala descrittiva.

PRIMA SPERIMENTAZIONE (')

1° Concordanza di voti parziali, considerati isolatamente

a. Comparazione delle valutazioni analitiche formulate dai gruppi II e III

Prima di affrontare l'interpretazione dei risultati, bisogna precisare che le condizioni sperimentali in cui sono stati ottenuti sono tali da approssimare l'omogeneità desiderata. I professori hanno preso coscienza del dossier soltanto attraverso una sola lettura individuale. Tale procedura non garantisce una riflessione approfondita dei criteri di valutazione ed un rispetto rigoroso delle disposizioni per l'uso di essi.
Per la lingua, tre scale si sono rivelate di immediata efficacia:

1. La scala relativa alle strutture grammaticali

Nel gruppo II — che procede nella valutazione senza poter fare affidamento su una definizione di ciascun livello — il 61% dei voti si raggruppa al centro della scala (tendenza centrale; media prossima a 3). Invece, nel gruppo III (livelli definiti), il 65% delle valutazioni si concentra sul livello B (media 3,7).

2. Pertinenza semantica

Stessi risultati della scala precedente.

3. Pertinenza sintattica

La precisione nei confronti dei criteri si svolge di nuovo nello stesso senso, ma in modo un po' meno netto. La maggioranza dei professori del gruppo II ha scelto il livello B, mentre nel gruppo III la scala si sposta su MB.

Le altre due scale deludono.

1. Organizzazione delle idee

I voti si concentrano meglio nel gruppo III, ma sembra che le ambiguità nella definizione dei livelli abbiano ridotto la chiarezza dei risultati.

2. Vocabolario

Qui non si è registrato alcun progresso sensibile. È difficile definire con precisione l'adeguatezza e la ricchezza del vocabolario. Coloro che hanno costruito le scale e gli insegnanti che li hanno aiutati lo sapevano. Nondimeno hanno voluto tentare l'esperimento, che è fallito.

Per le scienze i risultati sono meno soddisfacenti che per la lingua. Generalmente nel gruppo II si assiste ad una valutazione migliore che nel gruppo III.

b. Valutazioni finali comparate nei tre gruppi

Ricordiamo che, nei tre gruppi, la valutazione finale è stata formulata liberamente senza alcun criterio di supporto. In lingua, la differenza di dispersione dei voti è nettamente più accentuata tra il gruppo I e gli altri, che non tra i gruppi II e III.

In scienze, al contrario, si osserva un migliore accordo nel gruppo I.

SECONDA SPERIMENTAZIONE

Le scale sono state rimaneeggiate sulla base delle osservazioni già effettuate.

a. Comparazione delle valutazioni analitiche formulate dai gruppi II e III

In lingua, nel gruppo III sono espresse valutazioni altamente concordanti, ad eccezione che per la scala «vocabolario». Per quest'ultima, la comparazione con un vocabolario di base costituirrebbe il criterio obiettivo più accessibile.

Per le scienze, risultano, per ora, piuttosto soddisfacenti due scale: «Relazione scritta dell'esperimento» e «Descrizione dell'esperimento con uno schema». Nelle altre due scale («Interpretazione» e «Invenzione»), si registra minore progresso. Il fenomeno sembra comunque accidentale: nel lavoro dell'alluno da giudicare le risposte non sono date secondo l'ordine delle domande; dall'intervista a più professori risulta che si sono create così delle confusioni.

b. Comparazione delle valutazioni globali

Sono confermati i risultati della prima sperimentazione.

2ª Concordanza tra i profili

Fino ad ora ciascuna valutazione è stata considerata isolatamente. È altrettanto interessante constatare se, di fronte ad una valutazione analitica, i professori pensano a profili di valutazioni che si possano aggiungere.

Se si considerano tutti i criteri di valutazione, i risultati possono sembrare imperfetti. In realtà questo fenomeno era previsto: si pensi, in effetti, a quale grado di perfezione bisognerebbe giungere nella costruzione dello strumento perché tutti i voti si armonizzino perfettamente per tutti gli aspetti considerati.

E estremamente interessante constatare che, tralasciando la scala «Vocabolario» in lingua, la seconda taratura comporta progressi importanti.

<table>
<thead>
<tr>
<th>1ª Sperimentazione</th>
<th>2ª Sperimentazione</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gruppo II</td>
<td>Gruppo III</td>
</tr>
<tr>
<td>Lingua</td>
<td>34</td>
</tr>
<tr>
<td>Scienze</td>
<td>24</td>
</tr>
</tbody>
</table>

Percentuali (arrotondate) di concordanza con il profilo ideale

Prime conclusioni

Se ci si pone come obiettivo di accrescere la concordanza delle valutazioni globali dei giudici riguardo ad una performance complessa, da analisi precedenti risulta che il procedimento più conveniente sta nel dare agli stessi giudici la definizione dei criteri da prendere in considerazione. La specificazione dei gradi per ciascuno di questi criteri comporta minori benefici.
Volendo rendere più oggettive e articolate le valutazioni analitiche delle performances complesse, resta come punto fondamentale la descrizione dei livelli. L’ideale è di dare una valutazione che sia contemporaneamente globale e analitica; l’alunno viene così informato sull’importanza specifica accordata ai diversi aspetti della valutazione.

3ª Condizionamento dei fattori esterni

a) Le scale descrittive contrastano il fenomeno della «tendenza centrale»?

1. Valutazione analitica

In lingua, gli istogrammi mostrano che la descrizione di ciascun gruppo (gruppo III) contrasta efficacemente la tendenza alla convergenza al centro (ad eccezione che per il «vocabolario»).

Per le scienze, i risultati — seppure più facilmente acquisiti per le caratteristiche del lavoro dell’alunno — sono altrettanto soddisfacenti.

2. Valutazione globale

Sia per il francese che per le scienze, la media dei voti globali si allontana sempre più dal voto centrale man mano che si passa dal I al II gruppo e dal II al III gruppo. L’esperimento è pertanto positivo.

3. Profili

In lingua, il gruppo III si allontana maggiormente dal voto centrale.

b) Le scale descrittive contrastano il fenomeno della contaminazione?

Analisi delle correlazioni

Lingua materna

L’analisi delle correlazioni tra le valutazioni analitiche porta a queste osservazioni:

1. In generale i professori del gruppo II non differenziano adeguatamente il criterio «Organizzazione delle idee» dagli altri. Si constata anche una forte convergenza tra «Vocabolario» e «Pertinenza semantica». Al contrario, la «Struttura grammaticale» è indipendente da tutti gli altri.

2. Nel gruppo III, nessun criterio è fortemente correlato agli altri. Pertanto il tentativo di oggettivazione ha retto. In particolare, la «Pertinenza semantica» svolge il ruolo specifico che le è assegnato. «Pertinenza sintattica» e «Strutture grammaticali» restano collegate (.28); ciò sembra normale perché le due variabili si completano.

Scienze

Nei due gruppi c’è una correlazione significativa tra «Relazione scritta di un esperimento» e «Relazione dell’esperimento con uno schema». Sembra che non si tratti di un effetto di contaminazione: le due variabili sono effettivamente vicine tra loro. Anche l’«Interpretazione» e l’«Invenzione» sono collegate; fenomeno questo altrettanto normale, trattandosi di processi intellettivi superiori di cui conosciamo la saturazione in fattori d’intelligenza generale.

Analisi fattoriale

La tecnica permette di determinare il grado di analogia tra due o più variabili. I risultati che seguono vanno acolti con riserva, perché le valutazioni erano riferite ad un solo elaborato di un alunno. La relativa dei risultati può dipendere tanto dalle caratteristiche specifiche del lavoro che dai criteri di valutazione.

Lingua materna

Nel gruppo II domina una constatazione: «L’organizzazione delle idee» riempie i tre fattori rivelati dall’analisi. Si conferma così un’ipotesi già formulata: valutando con questo criterio, i professori considerano anche gli aspetti già esaminati altrove.

Si osserva, inoltre, una analogia fattoriale tra, da una parte, la pertinenza semantica ed il vocabolario e, dall’altra, la pertinenza sintattica e l’organizzazione delle idee.

La situazione si chiarisce notevolmente nel gruppo III. Si osserva per prima cosa la presenza di un fattore generale che ben riflette l’omogeneità dei voti, alla quale tende l’applicazione corretta dei criteri valutativi proposti.

Le dimensioni relative al contenuto («Pertinenza semantica» e «Organizzazione delle idee») si rivelano come un altro fattore specifico.

Scienze

La chiarificazione è netta: per il gruppo II e per il gruppo III compaiono due fattori, di cui il primo raggruppa i due criteri descrittivi (descrizione scritta e descrizione per schemi) ed il secondo le due operazioni cognitive superiori (interpretazione e invenzione).

4ª Condizionamento dei fattori interni

Abbiamo visto che due principali cause di discordanza nella valutazione risiedono nell’intervento dei fattori esterni al procedimento appena esaminato o nella scelta dei criteri diversi, da parte dei giudici. In quest’ultimo caso, si tratta allora di fattori interni.

Ricordiamo ancora che, definendo l’oggetto e i gradi delle scale, si intende unificare i criteri valutativi adottati dagli insegnanti. Si sa pure che, in un primo tempo, i differenti criteri di valutazione sono stati scelti e definiti, per quanto possibile, rigorosamente. I
I gruppi II e III hanno ricevuto a questo proposito le stesse informazioni.

I giudici hanno inserito altri criteri nella valutazione, oltre quelli proposti?

Soltanto un colloquio approfondito con ciascuno insegnante implicato nell'esperimento permetterebbe una risposta completa a questa domanda. Non potendo procedere ad un siffatto esame, ci siamo proposti di conoscere, con l'analisi statistica, quale importanza è riservata ai voti analitici nell'attribuire il voto di sintesi (1). I risultati non sono molto probanti, specialmente in scienze. Sembra che nella valutazione globale i professori si sentano liberi di dare un nome ai vincitori analitici facendo intervenire criteri personali, più o meno adeguati a quelli proposti.

Ipotesi 2: La scala descrittiva permette una valutazione più valida

Come verificare questa ipotesi?

Per costruire le scale abbiamo consultato specialisti dell'insegnamento nei settori che interessavano i diversi aspetti della valutazione. Pur avendo considerato tutti gli aspetti importanti, non si può dire per scontato che gli insegnanti terranno conto di tutto nel formulare la valutazione globale. In particolare, si possono porre le seguenti domande.

— Il voto globale è significativamente correlato a ciascuna valutazione analitica (analisi delle correlazioni semplici)?
— Qual è l'influenza esercitata da ciascuno dei singoli giudizi analitici sulla valutazione globale (analisi di regressione multipia, analisi per segmentazione, analisi fattoriale)?

1° Correlazione tra voto globale e ciascuna valutazione analitica

Lingua materna

Rispetto al gruppo II, in cui alcuni criteri («Pertinenza semantica» «Vocabolario», «Organizzazione delle idee») sono correlati con il voto globale, a scapito della pertinenza sintattica e delle strutture grammaticali, tutti i criteri utilizzati nel gruppo III sono correlati significativamente con la valutazione globale.

Scienze

I coefficienti di correlazione tra valutazioni analitiche e globali sono più elevati nel gruppo II che nel gruppo III, specialmente per la relazione scritta. Questo risultato innesto richiede un'interpretazione simile a quella proposta per la percentuale a varianza esplicita, nei due gruppi, per i voti analitici.

2° Influenza dei singoli giudizi analitici sulla valutazione globale

La questione è stata affrontata con diverse tecniche statistiche. Qui ci limitiamo a presentare la parte essenziale delle conclusioni derivate da queste analisi.

Analisi di regressione

Nelle due discipline, la parte di varianza annessa, più importante nel II che nel III gruppo, gioca in favore di una maggiore indipendenza dei voti del gruppo sperimentale.

Analisi fattoriale

Nel gruppo III, la presenza del voto globale nel primo fattore, assai saturato in tutte le variabili, evidenzia ulteriormente il contributo proporzionato dei voti analitici al voto globale. In scienze, le valutazioni analitiche svolgono un ruolo ben diverso nei confronti del voto globale. Nel gruppo di controllo, le due variabili appartenenti allo stesso fattore del voto globale sono in relazione alla «de- scrizione» (scritta e per schemi) dei fenomeni osservati. Nel gruppo sperimentale, invece, il voto globale deriva dallo stesso fattore delle variabili «Interpretazione» e «Invenzione».

Analisi per segmentazione

In lingua materna, la «Pertinenza semantica» svolge un ruolo preponderante nei due universi. Al di là di questo criterio, i due gruppi si comportano diversamente. In realtà, le altre variabili svolgono un ruolo nettamente più significativo nel gruppo III piuttosto che nel II. Ciò conferma, ulteriormente, il contributo più proporzionato delle valutazioni analitiche al voto globale.

In scienze, mentre numerosi insegnanti del gruppo II fondano la valutazione globale soltanto sull'«Interpretazione» e la «Descrizione», i loro colleghi del gruppo III si affidano maggiormente alla valutazione analitica e in particolare, all'«Invenzione».

Ipotesi 3: La scala descrittiva permette una valutazione che insegnante e alunno possono utilizzare in modo migliore

Come verificare questa ipotesi?

La verifica migliore proviene da come gli insegnanti giustificano il voto globale.

La valutazione formativa presuppone che gli obiettivi perseguiti
siano chiaramente definiti e comunicati agli alunni (*). Quel che importa è che, al termine del lavoro, la giustificazione del voto attribuito sia in rapporto diretto con questi obiettivi. Inoltre, la formulazione precisa e, per quanto possibile, di questa giustificazione deve aiutare l’alunno nel proseguire l’apprendimento.

1^a La quantità di giustificazioni differenzia i gruppi tra loro?

Le giustificazioni sono più numerose nel gruppo I, poiché rappresentano quelli che in passato non erano in grado di gestire l’ambiente di lavoro. Negli altri gruppi sono molto meno numerose, essendo esso spesso spiegato nel gruppo finale del precedente svolgimento analitico. Questo fenomeno si accentua nel gruppo II, in cui i professori sono stati invitati a giustificare ciascuna valutazione analitica.

2^a Le giustificazioni sono rapportate agli obiettivi del lavoro proposto?

In lingua, il rapporto tra giustificazioni principali e secondarie è particolarmente importante nel gruppo III. Questa suddivisione dipende dal valore che ortografia e punteggiatura assumono nel gruppo I. Inoltre, si assiste ad un migliore equilibrio tra le giustificazioni principali nel gruppo III.

In scienze, le domande poste all’alunno indicano più chiaramente i criteri di valutazione. Ciò spiega perché gli insegnanti del gruppo I rinforzano frequentemente le valutazioni globali con giustificazioni principali. Tuttavia i professori del gruppo III concedono minor importanza all’invenzione, forse per i requisiti assunti a questo livello pedagogico.

3^a L’alunno riconosce facilmente i criteri dei commenti che lo riguardano?

Diverse analisi, evidenziano che poche giustificazioni, tranne se possono classificare in una data categoria, sono sufficientemente specifiche da favorire l’apprendimento successivo dell’alunno. Alcune si riferiscono soltanto a dettagli. È, così, relativamente elevata la percentuale di giustificazioni non classificabili.

In siffatte condizioni, quale vantaggio può trarre l’alunno da un solo commento presentato a sostegno di un voto globale? A questo punto, appare evidente il vantaggio delle scale descrittive: ciascuna valutazione analitica è giustificata, automaticamente e senza ambiguità, dall’attribuzione del grado nella scala. Ora gli strumenti permettono un duplice riferimento. Descrivendo i requisiti che determinano l’attribuzione di ciascun grado della scala, si stabiliscono i punti critici per sviluppare un’abilità e si permette all’alunno di rapportare la sua performance all’obiettivo prefissato. Inoltre, adattando i livelli alle capacità mentali degli alunni, si dà alla valutazione un riferimento normativo. Se vengono apportate queste precisazioni, la giustificazione della valutazione globale può essere riassunta in un breve spiegazione su come direzionare principalmente gli apprendimenti da realizzare per progredire e, eventualmente, in un commento di natura affettiva.

4^a I professori evidenziano nei commenti gli aspetti positivi o negativi del lavoro?

Per entrambe le discipline, la proporzione delle giustificazioni positive è di gran lunga più importante nel gruppo II che negli altri due, in cui le giustificazioni secondarie sono spesso negative.

Inoltre, i professori del gruppo III si accordano meglio, in lingua, sugli aspetti positivi o negativi di una particolare caratteristica del lavoro. Nel gruppo I, la stessa caratteristica è spesso segnata come positiva da alcuni e negativa da altri.

Conclusioni

Al termine di questa ricerca si impongono necessariamente alcune osservazioni. È utopico pensare di eliminare ogni divergenza nella valutazione di performance complesse. È tuttavia possibile ridurre notevolmente le divergenze tra i correttori.

Lo strumento costruito non riguarda soltanto i criteri pertinenti alla valutazione degli esercizi scelti. Intende proporli, innanzitutto, come occasione di riflessione; ciascuno potrà eventualmente procedere a modificare secondo le sue necessità, le sue esperienze personali, le sue critiche.

In ultima analisi, è più importante il metodo di costruzione della scala che non l’esempio scelto a dimostrazione. Si può ragionevolmente ritenere che la metodologia proposta, correttamente applicata, garantisca la validità delle scale di valutazione.

Generalmente, l’insegnante da solo non dispone del tempo necessario per costruire scale descrittive operative e per convalidarle sperimentalmente. Perciò si dovrebbero costituire delle equipe di lavoro ovunque sia possibile. Si creerebbe così, a condizione che l’informazione circoli correttamente, una gamma sempre più ampia di strumenti di valutazione, disponibili per tutti.

Bisogna anche aggiungere che strumenti complessi, come le scale descrittive, non sono fatti per essere usati in qualsiasi momento.
Utilizzate ad intervalli regolari, possono costituire dei punti di riferimento oggettivi e validi, tali da completare le osservazioni, peraltro effettuate dagli insegnanti.

Un po’ paradossalmente, le principali conclusioni sono già presenti al termine... dell’introduzione:
— la possibilità di accresce la precisione nel valutare le performance complesse e, da qui, l’omogeneità delle valutazioni;
— la necessità di modificare, nel valutare, alcuni comportamenti praticati da lungo tempo;
— la laboriosità nel costruire strumenti di misura un po’ più perfezionati;
— la necessità di una collaborazione sempre più stretta tra insegnanti e centri di ricerca.

Si è già detto che le scale proposte e le considerazioni sulla loro costruzione sono, in apparente, complicate. In effetti, la realtà è ben più semplice della descrizione verbale, fenomeno ben conosciuto da noi tutti che, ogni tanto, ci perdiamo nel labirinto delle spiegazioni di strumenti così facili da maneggiare!

Sostenere che tutto è semplice, che la pratica dell’insegnamento e della valutazione rinnovati alleggerisce i compiti degli insegnanti, sarebbe altrettanto errato. Il lavoro degli insegnanti, che intendono svolgere adeguatamente le loro mansioni, è sempre più complesso e duro; in particolare la valutazione formativa continua richiede un notevole impegno intellettuale e materiale. Meno che mai le soluzioni semplicistiche e di comodo possono albergare nella scuola. Questo non è il più piccolo paradosso presente nel nostro mestiere: più l’educazione guadagna in libertà, spontaneità, creatività, più il nostro compito si fa arduo. La società ha il dovere di riconoscerlo.

**NOTE AL CAPITOLO QUARTO**

(5) Questi professori hanno una formazione, di insegnamento o di laurea nelle seguenti materie: matematica, fisica, chimica, geografia, biologia. Tutti possono ottenere l’incarico in lavori scientifici sperimentali.

(6) Si affronta l’oggettività della valutazione, prima della validità, per presentare più direttamente i risultati sperimentali.

(7) Pur trattandosi di una scala ordinale, utilizziamo nell’esame delle tesi tecniche statistiche generalmente riservate ai dati emersi in scala di intervallo e di rapporto. Permettendoci di farla continuo (precisamente, Guilford e Kerlinger) hanno dimostrato che, rendendo la misurazione più rigorosa con alcuni tentativi, possiamo giustificare questa libertà grazie alla validità dei risultati ottenuti.

(8) Nell’allegato B sono riportati i risultati dettagliati delle analisi statistiche. In questo capitolo si riferisce soltanto degli aspetti dominanti.
Allegati

Introduzione

A. Le scale di valutazione

1. Lingua materna (insegnamento secondario rinnovato: I ciclo)

2. Lingua materna (insegnamento secondario rinnovato: II ciclo)

3. Un esperimento scientifico (insegnamento secondario rinnovato: I ciclo)

4. Scienze (insegnamento secondario rinnovato: II ciclo)

B. Analisi statistica dei risultati
Introduzione

La valutazione fa parte integrante del processo di insegnamento di cui è una componente essenziale. Va pertanto garantita la validità degli strumenti di misura utilizzati. Uno strumento valido è rappresentativo della materia di cui vogliamo verificare l'apprendimento: le domande esplorano contemporaneamente tutti i contenuti e tutti i comportamenti applicabili a questi contenuti. Una simile procedura necessita di studi teorici che verifichino gli elementi fondamentali per acquisire una informazione. La valutazione consente allora di informare ciascun alunno sui suoi progressi verso l'obiettivo determinato. Inoltre, la sua performance può essere comparata a quelle di un universo di riferimento (taratura normativa). Siamo purtroppo ancora lontani dal disporre studi e strumenti necessari...

Inoltre, la valutazione deve essere oggettiva e articolata per assolvere al suo ruolo nel confronto dell'alunno. Il livello di concordanza tra i diversi giudici è un indice importante della oggettività di una valutazione. Sappiamo da numerose ricerche docimologiche che, praticamente per tutte le discipline, i giudici sono piuttosto discordi nel valutare una performance. Di più, l'alunno può trarre un profitto reale dalla valutazione solo se le informazioni che l'accompagnano sono analitiche, complete e precise. Perché il professore possa determinare un tale feed-back nell'alunno, anche il procedimento sottostante all'attribuzione del voto deve essere lo stesso analitico.

Le difficoltà di valutazione sono tanto più grandi quanto più la performance richiesta all'alunno necessita della messa in opera di conoscenze e attitudini numerose. Le scale descrittive qui presentate sono state costruite per la valutazione di performances complesse. Alcuni professori hanno messo insieme la loro esperienza per individuare criteri valutativi, per quanto possibile, sicuri e validi. Questi criteri sono stati applicati a numerosi lavori.

Nella speranza di assicurare l'omogeneità di voti, sono state proposte e giustificate delle dettagliate disposizioni di correzione.

Con le scale descrittive, l'alunno riceve due tipi di informazioni:
— viene a conoscere i punti forti e le lacune del suo lavoro;
— dispone di punti di riferimento per rapporcare la sua performance a quella dei compagni di classe.

In ragione del suo carattere analitico, la valutazione condotta con l'aiuto di scale descrittive porta via molto tempo: tanti è vero che non si potrebbe applicare quotidianamente in una classe. La si dovrebbe riservare ad alcuni momenti cruciali dell'apprendimento, o più semplicemente, alle verifiche periodiche. Invece di servirsi di batterie complete di prove, che gli vengono proposte, il professore ha sempre la possibilità di utilizzarne una o due scale, secondo le necessità.
Si può sperare che, a poco a poco, vengano create banche di strumenti di valutazione a cui gli insegnanti potranno far ricorso con benefici per se stessi e per i loro alunni.

Note per la valutazione dei lavori:

1. I criteri di valutazione proposti nelle scale che seguono sono stati definiti e provati dalla commissione di professori, in molte riunioni di lavoro. Può darsi che differiscano da quelli che usate abitualmente per valutare. Per provare le scale, vi dovete comunque attenere scrupolosamente alle disposizioni proposte per la correzione. Se intendete modificare alcuni criteri, le vostre osservazioni e critiche saranno ben accette;

2. Per ottenere una valutazione agevole ed efficace è opportuno giudicare tutti gli elaborati dei vostri alunni con la prima scala, poi potrete procedere alle successive nello stesso modo;

3. Quasi per allestimento, viene proposto nell'allegato un lavoro corretto.

A. LE SCALE DI VALUTAZIONE

1. Lingua materna / I ciclo, insegnamento secondario

Completamento di un racconto*

Descrizione dell'esercizio

Per non ricorrere alla stesura di un genere usuale, si è proposto un racconto privo della parte centrale. Alcuni alunni del primo anno dell'insegnamento secondario rinnovato sono stati invitati a ricostruire la parte mancante, dopo una lettura individuale del riassunto e dei paragrafi di inquadramento.

Perché si è scelto questo esercizio

L'esercizio richiede agli alunni una performance complessa. La sua valutazione non può limitarsi al semplice conteggio di punti sulla base degli errori commessi; necessita della messa a punto di uno strumento più adeguato.

E preferito questo tipo di esercizio a quelli abituali perché le limitazioni imposte dal riassunto e dai paragrafi di inquadramento forniscono dei punti di riferimento oggettivi per la valutazione.

Scelta dei criteri di valutazione

I primi due criteri — pertinenza semantica e pertinenza sintattica — sono specifici per questo genere di esercizio. Gli altri tre criteri — vocabolario, strutture grammaticali e organizzazione delle idee nel paragrafo 2 — investono un carattere più generale.

Naturamente, la valutazione condotta secondo questi cinque criteri non sarà esaustiva. Nondimeno, viste su quegli aspetti che numerose ricerche fattoriali mettono maggiormente in evidenza, non si inteso costruire una scala speciale per apprezzare l'atmosfera o lo stile del racconto. Può darsi che gli insegnanti ne abbiano in un qualche modo tenuto conto nel mentre operavano sulle scale proposte.

Non vengono valutati gli errori di ortografia e di grammatica: l'esercizio proposto non era specificamente destinato a misurare l'apprendimento in questi settori.

Organizzazione dei dossiers

Si individuano per ciascun criterio:
— la definizione del criterio;
— la definizione degli aspetti o delle caratteristiche interni al criterio. (Tutti gli aspetti considerati sono ugualmente importanti per la valutazione).

— la determinazione dei livelli richiesti (i.e. richiesti, per ciascun grado della scala, non sono arbitrari; esse sono state determinate dopo approfondite analisi di molti lavori di alunni iscritti al primo anno di insegnamento secondario rinnovato);
— il procedimento della valutazione;
— la presentazione della scala (Non comporta alcuna informazione complementare rispetto al procedimento. Tuttavia offre il vantaggio di dare una descrizione gerarchica della performance).

UN'AVVENTURA DI LUCA*

Riassunto dei capitoli precedenti: Luca, un giovane studente belga di quindici anni, è appena arrivato a Pirna, in Germania, per trascorrere, come ogni estate, un mese di vacanze dallo zio Vincent, un vecchio sapato alquanto originale, che, in un antico castello isolato, si dedica alla sua passione: la chimica. Era stato stabilito che Georges, il fedele servitore dello zio, sarebbe andato a ricever Luca alla stazione ferroviaria, per condurlo in macchina al castello, distante quasi 20 chilometri. Ma all'appuntamento non ci è riuscito. Luca, un po' smarrito, cerca di telefonare allo zio; non risponde nessuno. Si reca dunque alla abitazione in auto, da cui to la porta d'ingresso leggermente apest, la dimora sembra apparentemente deserta. Luca inutilmente cerca nelle stanze del pianterreno e del primo piano, inutilmente chiama dappertutto, nessuno si fa vivo. Sembra che lo zio Vincent, il vecchio Georges, e la cuoca Anna, siano volatilizzati. Inaspettatamente di questo inspiegabile silenzio, Luca decide di proseguire le sue ricerche nello scantinato dell'edificio...

§ 1. — Scendendo i primi scalini della stretta scala che conduceva nello scantinato, Luca sentiva il cuore battere più velocemente. Mai, fino ad allora, egli si era avventurato in questa parte del castello. Una ventina di volte, da quando soggiornava a Pirna, aveva avuto la tentazione di spingerla la piccola porta scolpita, oltre la quale non si era deciso a passare. Adesso, invece, si trovava di nuovo in un luogo che lo faceva sentire a disagio. Le pareti erano oscurate e sembravano nascere da una realtà sconosciuta. Luca si domanda se l'abbandono di zio Vincent e di Anna possa avere un significato. Con paura e curiosità, Luca inizia a esplorare l'edificio, cercando di capire il motivo di questo misterioso silenzio.

§ 2. — Paragrafo scritto dal signor Croisic.
Oggetti mancanti, ribadi dietro, che continuerà ogni giorno di
problemi, un po' di fretta... ecc... il tetto sarà un grosso lavoro,
no da lavare. "Oh, che disgrazia!" esclamò Luca. "Se quel che odoro
non potrò più resister. Cerco il laboratorio cercando di
rapidamente aprire la piccola finestra, ma la finestra è chiusa.
Che fare? Non era tanto tranquillo. Sentiva il suo cuore battere
molto in fretta. "Non lo so, ne' García, ne' Gena. Luca si avvicina
per guardare. All'improvviso si decise di sfilare
no di leggeri rumori. Luca si decise ad alzare i rumori
erano sempre più forti. Gli sembrava di non trovare la
riva tra la molla non è di zio che prevede il rumore. Ebbene!
Siamo entrati tra quei rumori erano di origine. Luca
 disse fra sé: 'che fare, non è vero? Ora telefonare in
non si conoscevo alcun 'di che (al castello di zio
so qui dov'essere mai... ma non..."
(Si è cercato di riprodurre fedelmente il testo dell'autore, compresi gli errori. N.d.T.)

§ 3. — "Ma no, signor Croisic, vi asserisco che non c'era.
Esplorato ogni angolo: non c'era nessuna cosa. Cosa ha fatto? Eh! Dappi, ha aperto la piccola finestra in fondo. Capirete... l'odore era insopportabile. Ed è attraverso il laboratorio che si è
trovato il pezzo di carosello di cui vi ho parlato. Se è la scrivania di zio Vincent? Certo,
ono sicuro. Certo che conoscere la sua scrittura: da quando so leggere, mi ha scritto tutti i mesi... Ma... Aspettiamo un attimo... E veramente curioso! Sì, adesso me ne accorgo. Lo zio Vincent ha commesso un errore di ortografia! Ha scritto... "Questi sono entrati..." al posto di "...entrati..." Sapete come egli è maniaco in fatto di ortografia? Proprio lui, che è un autore di errori! E' impossibile! Certamente non era in condizioni normali quando ha scritto in quel modo. Cosa dite? Si, Bene. Sarà qui tra una ventina di minuti. No, non arriverà a
suo padre del suo arrivo... A presto, signor Croisic.

Dopo aver riempito il retro del telefonino, Luca andò in cucina per lavarsi le mani con abbondante acqua, per cercare di far scomparire quella maleodorante sostanza che gli pizzicava la pelle. "E più la paura che non il dolore...", pensava constatando che il suo dito non sanguinava più e che il taglio non era profondo. Prese un bicchiere di succo di frutta, per rimettere le idee a posto, mentre ripensava a quanto era successo, e ritornò nel salone ancor più deserto e silenzioso. Mise davanti a sé, sul tavolo, il curioso messaggio, lo rilese attentamente per la terza volta, si avvicinò per osservarlo meglio, poi, con il mento nel cavo della mano, si mise a riflettere intensamente. Gli occhi fissi sulla piccola porta scolpita, che inspiegabilmente aveva trovato chiusa al ritorno dalla sua spedizione, attenuto al minimo rumore, aspettava...
I. Pertinenza semantica

Definizione del criterio

Il lavoro è pertinente dal punto di vista semantico se l'alunno sviluppa le idee appropriate al contesto, vale a dire il riassunto ed i paragrafi 1 e 3.

Definizione degli aspetti che intervengono nel criterio

Due domande permetteranno di valutare la pertinenza semantica:
— L'alunno ha introdotto elementi che contraddicono le informazioni fornite?
L'elaborato non deve comportare necessariamente la presenza di tutte queste informazioni; è importante che non le contraddica.
— Le informazioni utilizzate vengono integrate in un contesto più ampio?
Questi due aspetti della valutazione sono illustrati qui di seguito.

1° aspetto: Inserimento di idee non pertinenti.
Si ritengono non pertinenti quelli elementi che contraddicono le informazioni date dal testo. La lettura di numerose copie di elaborati ha consentito di stabilire l'elenco di questi elementi non pertinenti, che servirà a valutare questo primo aspetto della pertinenza semantica.
1. L'inizio del paragrafo 2 non collocato gli avvenimenti nel laboratorio (esempio: spinse la porta, entrare, ... sono azioni non pertinenti; sono già espresse nel paragrafo 1).
2. Non si fa alcuna allusione allo «spettacolo insolito». Questa allusione può limitarsi ad una esclamazione, ad una sola parola.
3. Nel paragrafo 2 il commissario è presente o atteso.
4. L'aspetto e il contenuto del messaggio sono modificati (esempio: il pezzo di carta diventa una lettera lunga e accurata).
5. Il ragazzo menziona ferite gravi, sinistri, catastrofi, ...

2° aspetto: Integrazione delle informazioni

Soltanto l'individuazione di elementi non pertinenti non è sufficiente per dar conto del valore del lavoro dell'alunno. Sarà pure apprezzata la maniera con cui l'alunno integra in un contesto più ampio i fatti che riferisce.
a) Si cercheranno gli indici di questa integrazione:
— quando viene ricordata la personalità dello zio;
— dai riferimenti a personaggi secondari;

— quando vengono espressi i sentimenti di Luca, ...
(Questo elenco non è esaustivo).
b) Si giudicherà l'alunno che non ha integrato le informazioni se si limita a riferire strettamente i fatti che si svolgono dopo l'entrata di Luca nel laboratorio (fine del primo paragrafo) e prima della sua risposta al commissario (inizio del terzo paragrafo).
c) Alcuni elaborati sviluppano soltanto idee che, senza essere necessariamente in contraddizione con i paragrafi 1 e 3, non hanno rapporto con la situazione.

Determinazione dei livelli richiesti

1. Per l'aspetto «Inserimento di elementi non pertinenti», saranno considerati quattro livelli:
1. assenza di elementi non pertinenti;
2. uno o due elementi non pertinenti;
3. tre o quattro elementi non pertinenti;
4. più di quattro elementi non pertinenti.
2. Per l'aspetto «Integrazione», saranno considerati tre livelli:
1. integrazione;
2. nessuna integrazione;
3. niente, al di là di idee non rapportate al contesto.

Procedimento della valutazione

1. Valutate il primo aspetto, «Inserimento di elementi non pertinenti»
   — assenza di elementi non pertinenti
   — 1 o 2 elementi non pertinenti
   — 3 o 4 elementi non pertinenti
   — più di 4 elementi non pertinenti

2. Valutate il secondo aspetto «Integrazione»
   L'alunno:
   — integra le informazioni in un contesto più ampio
   — non integra le informazioni
   — sviluppa soltanto idee senza rapporto con la situazione

3. Rapportandovi allo schema che segue, scegliete il livello conforme alle due caratteristiche individuate nell'elaborato dell'alunno:

<table>
<thead>
<tr>
<th>Inserimento di elementi non pertinenti</th>
<th>Integrazione</th>
</tr>
</thead>
<tbody>
<tr>
<td>Assenza di elementi non pertinenti</td>
<td>Integrazione</td>
</tr>
<tr>
<td>1 o 2 elementi non pertinenti</td>
<td>Nessuna integrazione</td>
</tr>
<tr>
<td>3 o 4 elementi non pertinenti</td>
<td>Idee non integrate alla situazione</td>
</tr>
<tr>
<td>più di 4 elementi non pertinenti</td>
<td></td>
</tr>
</tbody>
</table>
Presentazione della scala «Pertinenza semantica»

<table>
<thead>
<tr>
<th>Inserimento di elementi non pertinenti</th>
<th>Integrazione delle informazioni ricevute</th>
</tr>
</thead>
<tbody>
<tr>
<td>MB</td>
<td>Assenza di elementi non pertinenti</td>
</tr>
<tr>
<td></td>
<td>Integrazione dei fatti del paragrafo 2 in un contesto più ampio</td>
</tr>
<tr>
<td>B</td>
<td>Assenza di elementi non pertinenti</td>
</tr>
<tr>
<td></td>
<td>Nessuna integrazione</td>
</tr>
<tr>
<td></td>
<td>1 o 2 elementi non pertinenti</td>
</tr>
<tr>
<td></td>
<td>Integrazione dei fatti del paragrafo 2 in un contesto più ampio</td>
</tr>
<tr>
<td>S</td>
<td>1 o 2 elementi non pertinenti</td>
</tr>
<tr>
<td></td>
<td>Nessuna integrazione</td>
</tr>
<tr>
<td>D</td>
<td>3 o 4 elementi non pertinenti</td>
</tr>
<tr>
<td></td>
<td>Con o senza integrazione</td>
</tr>
<tr>
<td>I</td>
<td>Più di 4 elementi non pertinenti</td>
</tr>
<tr>
<td></td>
<td>Con o senza integrazione</td>
</tr>
<tr>
<td></td>
<td>Soltanto idee senza rapporto con la situazione, indipendentemente dal numero di elementi non pertinenti.</td>
</tr>
</tbody>
</table>

II. Pertinenza sintattica

**Definizione del criterio**

Rispetto delle regole sintattiche imposte dai paragrafi 1 e 3.

**Definizione degli aspetti**

L'elaborato dell'alunno è pertinente dal punto di vista sintattico se concorda con i tempi e le persone grammaticali utilizzati dall'autore.

1° aspetto: **Concordanza dei tempi**

L'autore utilizza il passato remoto, l'imperfetto e il piuaccheperfetto, oltre al presente nel discorso diretto.

**Osservazioni:**
1. Non verranno seguiti gli errori nella coniugazione dei tempi (es.: Allora, Luca uscirà...)
2. Non vengono seguiti gli errori nella concordanza di tempi e di modi nella scelta dei tempi.

2° aspetto: **La persona grammaticale**

L'alunno deve usare la terza persona singolare (la prima singolare nel caso del discorso diretto).

**Determinazione dei livelli richiesti**

a. Per l'aspetto «Rispetto dei tempi», si considereranno tre livelli:
1. Nessun errore;
2. Uno o due errori;
3. Più di due errori.

b. Anche per l'aspetto «Rispetto della persona» saranno considerati tre livelli:
1. Nessun errore;
2. Uno o due errori;
3. Più di due errori.

**Procedimento della valutazione**

1. Valutare il primo aspetto «Concordanza dei tempi».
   - Nessun errore
   - Uno o due errori
   - Più di due errori

2. Valutare il secondo aspetto «Corretto uso della persona grammaticale».
   - Nessun errore
   - Uno o due errori
   - Più di due errori

3. Rapportandovi allo schema che segue, scegliete il livello conforme alle due caratteristiche individuate nell'elaborato dell'alunno:

**Concordanza dei tempi**

- Nessun errore
- Uno o due errori
- Più di due errori

**Rispetto della persona grammaticale**

- Nessun errore
- Uno o due errori
- Più di due errori

(Segue)
III. Vocabolario

Definizione del criterio
Si valuta l’insieme dei vocaboli utilizzati dall’alunno nel testo. Non si tiene conto dell’ortografia.

Definizione degli aspetti che intervengono nel criterio
Si valuta il vocabolario secondo tre punti di vista: correttezza, precisione, originalità.

1° aspetto: Correttezza
È difficile dare una definizione precisa della correttezza del vocabolario. Si cerca di valutarla individuando le principali inesattezze, come:
- Plenonasmi (salire in alto — prevedere in anticipo, ...);

2° aspetto: Precisione
- Si sottolinea l’uso di:
  - Vocaboli non appropriati (es.: Apri la finestra per far evaporare quella peste);
  - Vocaboli passe-partout (es.: tubo invece di proietta; mettere, fare, ...); si sottolinea pure la ripetizione insopportabile di un vocabolo.

3° aspetto: Originalità
Per mezzo dei primi due aspetti si rilevano gli errori nell’elaborato senza valutare le qualità.
Per articolare la valutazione, si considera l’originalità dei vocaboli scelti. In tal modo si separano gli elaborati che presentano una stessa quantità di errori e di imprecisioni.

Determinazione dei livelli richiesti
a. Per l’aspetto «Correttezza» si considerano tre livelli:
1. Nessun errore;
2. Uno o due errori;
3. Più di due errori.

b. Per l’aspetto «Precisione» due sono i livelli considerati:
1. Al massimo tre imprecisioni;
2. Più di tre imprecisioni.

c. Per l’aspetto «Originalità» i livelli considerati saranno due:
1. SI;
2. No.

Procedimento della valutazione
1. Valutate il primo aspetto: «Correttezza»
   - Nessun errore
   - Uno o due errori
   - Più di due errori

2. Valutate il secondo aspetto: «Precisione»
   - Al massimo tre imprecisioni
   - Più di tre imprecisioni

3. Valutate il terzo aspetto: «Originalità»
   SI
   NO
4. Rapportandoci allo schema che segue, scegliete il livello che corrisponde alle caratteristiche individuate nell'elaborato dell'allievo.

<table>
<thead>
<tr>
<th>Correttaza</th>
<th>Precisione</th>
<th>Originalità</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nessun errore</td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Al massimo tre imprecisioni</td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Sì</td>
<td>MB</td>
<td></td>
</tr>
<tr>
<td>- No</td>
<td>B</td>
<td></td>
</tr>
</tbody>
</table>

| Uno o due errori |
| - Al massimo tre imprecisioni |
| - Sì | B |
| - No | S |
| - Si |

| Più di due errori |
| - Al massimo tre imprecisioni |
| - Sì | S |
| - No | D |
| - Si |

Voto per il vocabolario:

### Definizione degli aspetti

Si valuta la sintassi grazie a due aspetti: correttezza e varietà delle costruzioni impiegate.

1° aspetto: Correttezza

Si intende soprattutto valutare, a questo livello, la correttezza sintattica _interna_ della frase presa come unità. Molti errori della «sintassi tra le frasi», vanno attribuiti, in effetti, ad un erroso controllo semantico (ordine nelle idee) e saranno esaminati nella scala «Organizzazione delle idee».

E arduo definire i criteri per correggere una frase e redigere un elenco esaustivo degli errori. A titolo esemplificativo si ricordano le carenze più frequenti:
- Uso equivoco di pronomi personali e aggettivi possessivi, e nella costruzione delle relative;
- Collocazione all'inizio di frase, di aggettivi o participi che non concordano con il soggetto;
- Frasi incomplete (mancano i termini essenziali per la comprensione);
- Errori nel concordare tempi e modi, e nella scelta dei tempi;
- Errori dovuti a forme di automatismo linguistico: errore uso dei pronomi (esempio: il libro che parlo).

2° aspetto: Varietà

Si possono osservare diversi moduli di varietà:
- Varietà nelle strutture: differenze nell'iniziare le frasi (inversioni trasferimenti fraseggiamenti ellissi ...);
- Varietà nelle forme usate: affermativa, negativa, esclamativa, interrogativa, impiego del discorso diretto.
- Varietà nel presentare le possibili relazioni:
  - all'interno di una stessa frase: giustapposizione, coordinazione, inclusione (impiego di trattini, di parentesi), subordinazione;
  - tra le frasi: non ricorrere eccessivamente a elementi ridondanti o intercalari (e poi, e allora, e dopo, ...);
- Varietà nell'espansione del sintagma nominale: epiteto, apposizione, complemento di determinazione, proposizioni relative.

Si considera che l'alunno varia le frasi all'interno di un determinato modulo se fa riferimento alle diverse possibilità previste nella descrizione del modulo stesso.

### Determinazione dei livelli richiesti

a. Per l'aspetto «Strutture corrette», si considerano tre livelli:
1. Nessun errore
2. Uno o due errori
3. Più di due errori

b. Per l'aspetto «Varietà di strutture» si considerano tre livelli:
1. Almeno tre tipi di varietà
2. Due tipi di varietà

### IV. Strutture grammaticali

#### Definizione del criterio

Diversamente dalla «Pertinenza sintattica», qui si valuta la sintassi del paragrafo redatta dall'alunno, indipendentemente dalle regole presenti nei paragrafi 1 e 3.
3. Un tipo soltanto
4. Nessun tipo di varietà

**Procedimento della valutazione del criterio «Strutture grammaticali»**

1. Valutate il primo aspetto: «Correttezza»
   - Nessun errore
   - Uno o due errori
   - Più di due errori

2. Valutate il secondo aspetto: «Varietà»
   - Almeno tre registri di varietà
   - Due registri
   - Un registro soltanto
   - Nessun registro di varietà

3. Rapportandovi allo schema che segue, scegliete il livello che corrisponde alle caratteristiche individuate nell’elaborato dell’allievo.

<table>
<thead>
<tr>
<th>Correttezza</th>
<th>Varietà</th>
</tr>
</thead>
<tbody>
<tr>
<td>MB</td>
<td>Nessun errore</td>
</tr>
<tr>
<td>B</td>
<td>Nessun errore</td>
</tr>
<tr>
<td>S</td>
<td>Nessun errore</td>
</tr>
<tr>
<td>D</td>
<td>Nessun errore</td>
</tr>
<tr>
<td>I</td>
<td>Nessun errore</td>
</tr>
</tbody>
</table>

**V. Organizzazione delle idee nel paragrafo 2**

**Definizione del criterio**
Diversamente dalla dimensione «Pertinenza semantica», si valuta la struttura dell’elaborato, indipendentemente dalle regole imposte dai paragrafi 1 e 3.

**Descrizione degli aspetti**
Oggi valutazione verta su tre aspetti: coerenza semantica, collegamento di idee e progressione del racconto.

1. **Coerenza semantica**
   Le idee espresse dall’allievo non devono contrastarsi.

2. **Collegamento delle idee**
   Si valuta come l’allievo collega le idee le une alle altre, sia utilizzando elementi di raccordo (congiunzioni di coordinazione, locuzioni avverbiale, ...) sia con la punteggiatura.
   Viene corretto l’uso inadeguato degli elementi di raccordo. Esempi:
   - Assenza di un elemento di collegamento necessario: «Lo raccolse e si tagliò qualcosa» era scritto ».
   - Scolta inadeguata di questo elemento: «Andò nell’altra stanza che non conosceva ed era la cucina». — «Sentii odori di sostanze chimiche perché c’era disordine».
   Anche gli errori evidenti di punteggiatura verranno corretti nella misura in cui determinano collegamenti non adeguati di idee.

3. **Progressione del racconto**
   Si valuta il modo in cui l’allievo dà ritmo sostenuto al suo lavoro.
   Si corregge la presenza di elementi ridondanti che rallentano la progressione del
racconto. Esempio: «Aprendo la pesante porta, si ferì leggermente al dito. Dopo aver aperto la pesante porta, Luca salì le scale».

Determinazione dei livelli richiesti
a. Per l'aspetto «Coerenza semantica», si considerano tre livelli:
   1. Racconto coerente
   2. Limitate contraddizioni
   3. Racconto non coerente
b. Per l'aspetto «Collegamento di idee», si considerano due livelli:
   1. Collegamento generalmente adeguato
   2. Collegamento non adeguato
c. Per l'aspetto «Progressione del racconto», si considerano i livelli:
   1. Racconto non ridondante
   2. Progressione rallentata

Procedimento della valutazione del criterio «Organizzazione delle idee nel paragrafo 2»
1. Valutare il primo aspetto: «Coerenza semantica»
   - Racconto coerente
   - Limitate contraddizioni
   - Racconto non coerente
2. Valutare il secondo aspetto: «Collegamento delle idee»
   - Collegamento generalmente adeguato
   - Collegamento non adeguato
3. Valutare il terzo aspetto: «Progressione del racconto»
   - Racconto non ridondante
   - Progressione rallentata
4. Rapportandovi allo schema che segue, scegliete il livello che corrisponde alle caratteristiche individuate nell'elaborato.

<table>
<thead>
<tr>
<th>Coerenza semantica</th>
<th>Collegamento delle idee</th>
<th>Progressione del racconto</th>
</tr>
</thead>
<tbody>
<tr>
<td>Racconto coerente</td>
<td>Collegamento adeguato</td>
<td>MB</td>
</tr>
<tr>
<td></td>
<td>Collezionato non adeguato</td>
<td>B</td>
</tr>
<tr>
<td>Limitate contraddizioni</td>
<td>Collegamento adeguato</td>
<td>B</td>
</tr>
<tr>
<td></td>
<td>Collezionato non adeguato</td>
<td>B</td>
</tr>
<tr>
<td>Racconto non coerente</td>
<td>Collegamento adeguato</td>
<td>D</td>
</tr>
<tr>
<td></td>
<td>Collezionato non adeguato</td>
<td>I</td>
</tr>
</tbody>
</table>

Voto per l'organizzazione delle idee del paragrafo 1: [ ]

Voto di sinesi: [ ]

Presentazione della scala «Organizzazione delle idee nel paragrafo 2»

<table>
<thead>
<tr>
<th>Coerenza semantica</th>
<th>Collegamento delle idee</th>
<th>Progressione del racconto</th>
</tr>
</thead>
<tbody>
<tr>
<td>MB</td>
<td>Racconto coerente</td>
<td>Racconto non ridondante</td>
</tr>
<tr>
<td></td>
<td>Collegamento adeguato</td>
<td>Racconto non ridondante</td>
</tr>
<tr>
<td></td>
<td>Collezionato non adeguato</td>
<td>Racconto non ridondante</td>
</tr>
<tr>
<td>B</td>
<td>Racconto coerente</td>
<td>Racconto non ridondante</td>
</tr>
<tr>
<td></td>
<td>Collegamento adeguato</td>
<td>Racconto non ridondante</td>
</tr>
<tr>
<td></td>
<td>Collezionato non adeguato</td>
<td>Racconto non ridondante</td>
</tr>
<tr>
<td>S</td>
<td>Limitate contraddizioni</td>
<td>Racconto non ridondante</td>
</tr>
<tr>
<td></td>
<td>Collegamento adeguato</td>
<td>Racconto non ridondante</td>
</tr>
<tr>
<td></td>
<td>Collezionato non adeguato</td>
<td>Racconto non ridondante</td>
</tr>
<tr>
<td>D</td>
<td>Racconto incoerente</td>
<td>Racconto non ridondante</td>
</tr>
<tr>
<td></td>
<td>Collegamento adeguato</td>
<td>Racconto non ridondante</td>
</tr>
<tr>
<td></td>
<td>Collezionato non adeguato</td>
<td>Racconto non ridondante</td>
</tr>
<tr>
<td>I</td>
<td>Racconto incoerente</td>
<td>Racconto non ridondante</td>
</tr>
<tr>
<td></td>
<td>Collegamento adeguato</td>
<td>Racconto non ridondante</td>
</tr>
<tr>
<td></td>
<td>Collezionato non adeguato</td>
<td>Racconto non ridondante</td>
</tr>
</tbody>
</table>

Sintesi delle valutazioni

Alette appena valutato ciascun criterio. Adesso riportate i voci su queste scale:
1. Pertinenza semantica
   | MB | B | S | D | I |
2. Pertinenza sintattica
   | MB | B | S | D | I |
3. Vocabolario
   | MB | B | S | D | I |
4. Strutture grammaticali
   | MB | B | S | D | I |
5. Organizzazione delle idee
   | MB | B | S | D | I |

Voto di sinesi: [ ]
2. Lingua materna / II ciclo, insegnamento secondario

Riassunto e critica di un brano*

Descrizione dell'esercizio

All'inizio dell'anno scolastico, si è chiesto agli alunni del quarto anno dell'insegnamento secondario rinnovato di riassumere e criticare le idee esposte in un brano tratto da Proposizioni sull'educazione di Alain.

Il testo è stato letto individualmente. Per non condizionare la performance dell'alunno, il brano era privo di titolo. L'insegnante non ha fornito alcuna informazione complementare. I compiti richiesti erano i seguenti:

1. Riassumi le idee esposte nel testo il più brevemente possibile, impiegando parole tue appropriate;
2. Discuti le idee dell'autore con tue argomentazioni.

I due compiti vengono sullo stesso testo e sono stati presentati simultaneamente agli alunni. Richiedendo però due performances diverse abbiamo suddiviso il fascicolo in altrettante parti: la prima destinata alla valutazione del riassunto, la seconda alla critica.

«Intendo con insegnamento aristocratico, quel tipo di insegnamento finalizzato a separare quelli che sapranno e governeranno da quelli che ignorano e obbediranno.

Rivedo con la memoria il mio professore di matematica, mentre ammendava con la sua cronaca un po' pesante un mio compagno di scuola. Costui, pure quanto altri mai, era costretto a leggere con il naso attaccato alle righe. Quanto a vedere con un solo sguardo un triangolo intero, se lo poteva dimenticare. Penso che l'insegnante avrebbe dovuto farlo esercitare con piccole figure, non più larghe dell'estremità del suo naso, dalla scoperta del triangolo intero si poteva poi passare a cogliere i rapporti, a ragionare bene come gli altri.

L'addestramento dei suoi confronti era ben diverso: Lo si pressava continuamente. Era costretto a passare dal vertice di un triangolo all'altro, parlava a vuoto, diceva A per B, retto al posto di angolo; ti che lo rendeva perfettamente ridicolo, mentre noi ri-

Secondo me bisognerebbe agire del tutto diversamente: istruire tutto il popolo; piacere a tutti ognuno, alle menti più grossolane; pungolare la pigrizia; svelare ad ogni costo quelli che dormono; mostrare più entusiasmo per un contadino appena appena politico che per l'elegante scienziato abituato a volare sulle vette dell'Ecole Poly-

di quei pubblici dovrebbero impegnarsi a portar chiezza tra le masse del basso e dell'interno, e non soltanto far brillare alcune superbe vette, di origini ple-

(Alain)

Perché sì è scelto questo esercizio

La scala di valutazione è uno strumento destinato a valutare performances compiute per le quali risulta inadeguato un semplice conteggio di punti per errore commesso. Tra le performances compiute da considerare all'inizio dell'insegnamento secondario superiore, è opinione generale che il riassumere e criticare un testo sia un apprendimento fondamentale.

Organizzazione dei dossiers

Ciascun dossier comprende:

1. La definizione del criterio;
2. La descrizione degli aspetti che intervengono nella valutazione;
3. La determinazione dei livelli richiesti;
4. Il procedimento della valutazione;
5. La presentazione della scala.

Le richieste, in ciascun livello della scala, non sono arbitrarie; derivano da un'analisi approfondita di numerosi elaborati di alunni del quarto anno dell'insegnamento secondario rinnovato.

1. Riassunto

1. Definizione del criterio

Per riassunto si intende la ricostruzione, in forma sintetica, delle principali idee esposte nel testo.

2. Descrizione degli aspetti

La definizione indica i due aspetti da valutare:
1. La comprensione delle principali idee presenti nel testo;
2. La concisione nel presentarle.

I due aspetti assumono uguale importanza per l'attribuzione del voto. La valutazione non verrà sull'ortografia, il vocabolario, le strutture grammaticali, ecc., ma sull'esercizio proposto non è specificatamente destinato a misurare l'apprendimento in questi settori.

I. Compressione delle idee principali del testo

Nel testo di Alain sono contenute tre idee fondamentali:
A. L'insegnamento attuale è nobile, aristocratico.
B. Bisogna passare ad un insegnamento eguagiaritario.
C. Vi è relazione tra insegnamento e sistema politico e si fanno concessioni a questo sistema («dare apparenza di giustizia all'inuguaglianza»).

La comprensione è valutata da due punti di vista:
— E completa?
— È corretta?

a. Completa comprensione delle idee principali

L'idea A comprende questi elementi:
1. Si tratta di un insegnamento.
— E sufficiente che l'alunno menzioni una sola volta, nella sua risposta, che si tratta di insegnamento.
2. che favorisce.
— L'alunno deve evidenziare l'idea di un diverso trattamento che porta a favorire alcuni allievi e a trascurare altri.
3. una cifre (i futuri governanti) rispetto alla massa (destinata ad ignorare e ad obbedire).
— L'alunno deve liberarsi dall'esempio proposto da Alain (il miope; le preferenze per un allievo) e far notare che si parla di un sistema di insegnamento.

L'idea B comprende i seguenti elementi:
1. Bisogna passare.
— L'alunno deve evidenziare la necessità di un cambiamento. Si può esprimere questa idea anche con l'uso di un esemplare condizionale.
2. ad un insegnamento.
— L'alunno non è obbligatoriamente tenuto a ripetere questo elemento della risposta, se è già espresso nell'idea A.
3. che dia uguali possibilità a tutti.
— Se Alain difende un insegnamento, non identico per tutti, ma adattato ai bisogni di ciascuno e che si conterrà di minor precisione dall'alunno dell'insegnamento secondario e si considererà corretta l'idea di un insegnamento eguagliario.

Come si nota, nella scala si considerano ugualmente importanti le idee A e B.

L'idea C: L'analisi di un gran numero di elaborati mostra che la restituzione di questa idea è assai rara al livello di 2° ciclo dell'insegnamento rinnovato. É per questo che la sua comprensione non è presente nella scala; eventualmente, il professore può tuttavia valorizzare una buona performance con un commento appropriato. Se dal lavoro dell'alunno risulta l'incomprensione totale o parziale di questa idea, non se ne terrà conto.

Per ciascuna idea, l'alunno da prova di:
— comprensione completa, se nella risposta sono presenti tutti gli elementi menziona-
l'i qui sopra;
— comprensione parziale, se manca uno degli elementi;
— omissione, se mancano tutti gli elementi.

b. Compressione corretta delle idee principali

Si penalizzerà la presenza di idee falsate.
Una idea falsata è un'idea del testo a cui l'alunno attribuisce un significato diverso da quello di Alain (esempio: «Bisogna cambiare l'insegnamento per permettere a tutti di governare»).
Tutte le idee falsate devono riferirsi a specifici testi da Alain (esempio: «Bisogna cambiare l'insegnamento per permettere a tutti di governare»).

Riassumendo, bisogna determinare se ciascuna idea è:
— Completa e correttamente compresa;
— Parzialmente e correttamente compresa;
— Falsata;
— Omessa.

II. Concersione

La consegna impone all'alunno di riassumere «il più brevemente possibile» le principali idee del testo. Per questo motivo, viene penalizzato ogni ripetizione di idea (completa, parziale o falsata), sotto qualsiasi forma (ripresa direttamente dal testo o parzialmente) perché già espressa una prima volta.
Non si considerano errori le ripetizioni di termini o di gruppi di parole necessari ad evitare la ambiguità.

Esempi:
Alunno A: «Le mie idee venissere applicate, l'insegnamento da aristocratico diventerà democratico; vale a dire, che tutti avrebbero la stessa possibilità, sin il condadino che i figli di famiglie agiate: in questo modo, le differenze tra gli strati sociali scomparirebbero poco a poco, e il popolo uscirebbe dall'ignoranza in cui si trova tuttora e si innalzerà esso stesso verso le più alte scuole».
Fino a «famiglie agiate» il testo dell'alunno dimostra che sono state ben comprese le idee A e B. La parte seguente è ridondante.
Alunno B: «L'insegnamento aristocratico che umilia e annienta le masse dovrebbe essere sostituito con un insegnamento che collochi tutti sullo stesso piano».
I verbi umiliare e annientare, intesi come variazioni stilistiche, non vengono considerati errori.
3. Determinazione dei livelli richiesti

a. Per l'aspetto «Comprensione delle idee principali del testo», si considerano cinque livelli.

   L'elaborato comporta:

1. Due idee complete;
2. Un'idea completa e una parziale;
3. Un'idea completa e una falsata o omessa; oppure; due idee parziali;
4. Un'idea parziale e una falsata o omessa;
5. Due idee falsate o omesse.

b. Per l'aspetto «Concisione», si considerano due livelli soltanto:

1. Il testo è conciso;
2. Il testo non è conciso; presenta almeno un caso di ridondanza.

4. Procedimento per la valutazione del riassunto

1. Valutate il primo aspetto «Comprensione».

   - L'idea A è:
     * compresa in pieno     □
     * parzialmente compresa □
     * falsata                □
     * omessa                 □

   - L'idea B è:
     * compresa in pieno  □
     * parzialmente compresa □
     * falsata              □
     * omessa               □

   Scegliete il livello corrispondente alle caratteristiche che avete rinvenuto. Nel testo dell'alunno si rinvengono:

   - Due idee complete.
   - Un'idea completa e una parziale.
   - Un'idea completa e una falsata o omessa oppure; due idee parziali.
   - Un'idea parziale e una falsata o omessa.
   - Due idee falsate o omesse.

2. Valutate il secondo aspetto: «Concisione».

   - Il testo è conciso?       SI □    NO □

3. Rapportandovi allo schema che segue nella pagina a fianco, scegliete il livello che corrisponde alle caratteristiche individuate nell'elaborato.

1\o aspetto: Comprensione delle idee A e B.  2\o aspetto: Concisione  Livello della scala

<table>
<thead>
<tr>
<th>MB</th>
<th>Due idee complete</th>
<th>Si</th>
</tr>
</thead>
<tbody>
<tr>
<td>B</td>
<td>Due idee complete</td>
<td>No</td>
</tr>
<tr>
<td>S</td>
<td>Un'idea completa e una parziale</td>
<td>Si</td>
</tr>
<tr>
<td>S</td>
<td>Un'idea completa e una falsata o omessa oppure; Due idee parziali</td>
<td>No</td>
</tr>
<tr>
<td>D</td>
<td>Un'idea completa e una falsata o omessa oppure; Due idee parziali</td>
<td>No</td>
</tr>
<tr>
<td>I</td>
<td>Un'idea parziale e una falsata o omessa</td>
<td>Si</td>
</tr>
<tr>
<td>I</td>
<td>Due idee false o omesse</td>
<td>Quali che siano le caratteristiche della concisione</td>
</tr>
</tbody>
</table>

5. Presentazione della scala

Da questa scala non provengono informazioni supplementari al procedimento descritto precedentemente. La scala offre tuttavia il vantaggio di dare una descrizione gerarchica della performance.
II. Critica

1. Definizione del criterio
Criticare significa esaminare le idee presenti in un testo ed esprimere un parere su ciascuna di esse.

2. Descrizione degli aspetti
Si considerano due aspetti:
- Quantità e correttezza delle idee riportate nella critica;
- Tipo e qualità delle argomentazioni utilizzate.

Nella scala hanno maggior importanza il tipo e la validità delle argomentazioni che non la quantità e correttezza delle idee riportate. In effetti l’aspetto essenziale della critica è l’argomentazione.

Come per il riassunto, non si valutano gli aspetti formali (ortografia ecc.). Inoltre, non è stata prevista una scala specifica per valutare le argomentazioni. Saranno valutati assieme alla validità.

1. Quantità e correttezza delle idee riportate nella critica
Si valutano qui le idee discusse con almeno un argomento valido.

1. Verranno considerate soltanto le critiche alle idee esprese da Alain, non quelle relative allo stile, al vocabolario, al tono usati dall’autore.
2. L’alunno non critica
   - se riprende un’idea del testo dichiarandosi d’accordo o contrario, senza alcuna giustificazione;
   - se si limita a parafrasare l’autore senza recare elementi di novità.

II. Tipo e validità delle argomentazioni utilizzate
Questa valutazione sarà condotta in due fasi successive:

a) Si formerà subito un primo giudizio sulla validità dell’argomentazione.
   L’insegnante non può mettere in discussione la validità dell’argomentazione semplicemente perché non condivide l’opinione espressa. I criteri da tener presenti saranno di ordine logico.

b) Si riformula l’argomentazione che non presenta:
   - errori di ragionamento (1)
   Ad esempio, il ricorso:
   - ad argomentazioni non pertinenti (eleggere dalle idee che intendono sostenere);
   - a petizioni di principio (riportando nella dimostrazione gli argomenti da dimostrare);
   - a errori di causa (la relazione temporale intesa come relazione causale);
   - a generalizzazioni affrettate (la regola tratta da un solo esempio).

---

1) Questo elenco non riporta i tipi di errori di ragionamento contenuti nel testo di un alunno. D’altra parte, non è neppure esaustivo, ma aiuta il professore a determinare se un argomento è valido oppure no.

---

b) L’argomentazione valida viene ricondotta ad una di queste tre categorie:

Tipo A. Argomentazione oggettiva fondata su giudizi di valore d’ordine morale o su affermazioni non sostenute da fatti osservabili.
Tipo B. Argomentazione oggettiva fondata su precisi fatti osservabili. A differenza delle argomentazioni di tipo A, si richiede che l’alunno indichi i fatti sui quali si appoggia.
Tipo C. Argomentazione oggettiva fondata su implicazioni o anticipazioni.
- Implicazioni: l’alunno trae dalle idee di Alain delle conseguenze che Alain non poteva prevedere.
- Anticipazioni: l’alunno prevede possibili obiezioni e risponde.

3. Determinazione dei livelli richiesti.

a) Per l’aspetto «Quantità e correttezza delle idee riportate nella critica», si considereranno questi livelli:
   1. Critica ad almeno due idee principali complete;
   2. Critica ad almeno una idea principale;
   3. Critica ad almeno una idea parziale;
   4. Critica ad almeno una idea falsata (1);
   5. Nessuna critica.

b) Per l’aspetto «Tipo e validità degli argomenti utilizzati» si distingueranno tre livelli:
   1. Argomento(1) del tipo C (eventualmente accompagnati da argomenti di tipo A e/o B).
   2. Argomento(1) del tipo B (eventualmente accompagnati da argomenti di tipo A).
   3. Argomento(1) del tipo A solamente.

I livelli richiesti devono essere rapportati al lavoro dell’alunno nella sua totalità, non a ciascuna idea presa isolatamente. Per i giudizi da comunicare all’alunno è tuttavia preferibile ricordare ciascun argomento correlato alla sua idea.

4. Procedimento per valutare la critica.
La valutazione delle idee riportate nella critica (1° aspetto) dipende in parte dalle argomentazioni associate a queste idee (2° aspetto).
Pertanto, conerverà analizzare dapprima ciascuna idea sotto i due aspetti (Tavola I). In un secondo tempo, si procederà alla sintesi delle caratteristiche rinvenute nell’elaborato dell’alunno (Tavola II).

(1) La specificità della performance valutata con questo criterio, giustifica il quarto livello, che permette di considerare lo sforzo argomentativo dell’alunno.
### TAVOLA I/A

<table>
<thead>
<tr>
<th>1° aspetto: Idea criticata</th>
<th>2° aspetto: Tipo e validità degli argomenti utilizzati</th>
</tr>
</thead>
<tbody>
<tr>
<td>Validità dell'argomento</td>
<td>Tipo di argomento</td>
</tr>
<tr>
<td>Idea A. Idea completa criticata</td>
<td>con un argomento valido o meno</td>
</tr>
<tr>
<td>con un argomento non valido</td>
<td>A</td>
</tr>
<tr>
<td>non criticata</td>
<td>A</td>
</tr>
<tr>
<td>Idea parziale criticata</td>
<td>con un argomento valido o meno</td>
</tr>
<tr>
<td>con un argomento non valido</td>
<td>A</td>
</tr>
<tr>
<td>non criticata</td>
<td>A</td>
</tr>
<tr>
<td>Idea falsata criticata</td>
<td>con un argomento valido o meno</td>
</tr>
<tr>
<td>con un argomento non valido</td>
<td>A</td>
</tr>
<tr>
<td>non criticata</td>
<td>A</td>
</tr>
<tr>
<td>Idea omessa</td>
<td></td>
</tr>
</tbody>
</table>

### TAVOLA I/B

<table>
<thead>
<tr>
<th>1° aspetto: Idea criticata</th>
<th>2° aspetto: Tipo e validità degli argomenti utilizzati</th>
</tr>
</thead>
<tbody>
<tr>
<td>Validità dell'argomento</td>
<td>Tipo di argomento</td>
</tr>
<tr>
<td>Idea B. Idea completa criticata</td>
<td>con un argomento valido o meno</td>
</tr>
<tr>
<td>con un argomento non valido</td>
<td>A</td>
</tr>
<tr>
<td>non criticata</td>
<td>A</td>
</tr>
<tr>
<td>Idea parziale criticata</td>
<td>con un argomento valido o meno</td>
</tr>
<tr>
<td>con un argomento non valido</td>
<td>A</td>
</tr>
<tr>
<td>non criticata</td>
<td>A</td>
</tr>
<tr>
<td>Idea falsata criticata</td>
<td>con un argomento valido o meno</td>
</tr>
<tr>
<td>con un argomento non valido</td>
<td>A</td>
</tr>
<tr>
<td>non criticata</td>
<td>A</td>
</tr>
<tr>
<td>Idea omessa</td>
<td></td>
</tr>
</tbody>
</table>
**TAVOLA II**

1. Valutate il primo aspetto: «Quantità e correttezza delle idee».
   La critica contiene:
   - Almeno due idee principali complete e correttamente intese
   - Almeno un'idea principale e correttamente intesa
   - Almeno un'idea parziale e correttamente intesa
   - Soltanto idee false
   - Nessuna critica

2. Valutate il secondo aspetto: «Tipo e validità delle argomentazioni utilizzate».
   La critica contiene:
   - Uno o più argomenti del tipo C (accompagnati eventualmente da argomenti di tipo B e/o A)
   - Uno o più argomenti del tipo B (accompagnati eventualmente da argomenti del tipo A)
   - Uno o più argomenti del tipo A

3. Rapportandovi allo schema che segue, scegliete il livello che corrisponda alle due caratteristiche nel lavoro dell'alunno che avete valutato.

<table>
<thead>
<tr>
<th>Quantità e correttezza delle idee discusse</th>
<th>Tipo e validità degli argomenti</th>
</tr>
</thead>
<tbody>
<tr>
<td>Almeno due idee principali complete</td>
<td>Argomenti di tipo C (accompagnati eventualmente da argomenti B e/o A)</td>
</tr>
<tr>
<td>correttamente intese</td>
<td>Argomenti di tipo B (accompagnati eventualmente da argomenti A)</td>
</tr>
<tr>
<td>Almeno una idea principale compiuta</td>
<td>Argomenti di tipo A soltanto</td>
</tr>
<tr>
<td>e correttamente intesa</td>
<td>Argomenti di tipo B soltanto</td>
</tr>
<tr>
<td>Almeno una idea parziale e correttamente</td>
<td>Argomenti di tipo A soltanto</td>
</tr>
<tr>
<td>intesa</td>
<td>Argomenti di tipo B soltanto</td>
</tr>
<tr>
<td>Soltanto idee false</td>
<td>Argomenti di tipo B soltanto</td>
</tr>
<tr>
<td>Nessuna critica</td>
<td>Argomenti di tipo B o A</td>
</tr>
</tbody>
</table>

Voto per la critica:

<table>
<thead>
<tr>
<th>MB</th>
<th>Almeno due idee principali complete e correttamente intese</th>
<th>Argomenti di tipo C (accompagnati eventualmente da argomenti B e/o A)</th>
</tr>
</thead>
<tbody>
<tr>
<td>B</td>
<td>Soltanto argomenti del tipo B</td>
<td>Argomenti di tipo B (accompagnati eventualmente da argomenti A)</td>
</tr>
<tr>
<td>D</td>
<td>Almeno un'idea principale completa e correttamente intesa</td>
<td>Argomenti di tipo A soltanto</td>
</tr>
<tr>
<td></td>
<td>Soltanto idee false</td>
<td>Argomenti di tipo B (accompagnati eventualmente da argomenti A)</td>
</tr>
<tr>
<td>I</td>
<td>Almeno un'idea parziale e correttamente intesa</td>
<td>Argomenti di tipo B (accompagnati eventualmente da argomenti A)</td>
</tr>
<tr>
<td></td>
<td>Soltanto idee false</td>
<td>Argomenti di tipo B o A</td>
</tr>
<tr>
<td></td>
<td>Nessuna critica</td>
<td></td>
</tr>
</tbody>
</table>
Sintesi delle valutazioni

La sintesi di più valutazioni analitiche è giustificata quando una stessa performance è valutata in rapporto a differenti criteri. Per esempio, riguardo alla stesura, verranno sintetizzate le valutazioni relative alle idee esposte, alla loro organizzazione, alla scelta del vocabolario, ecc. ... Dal voto di sintesi, l'alluno vede quale importanza è accordata rispettivamente ad ogni criterio.

Per quanto concerne il lavoro presentato in questo dossier, il riassunto e la critica costituiscono due performances distinte. È inopportuno sintetizzarne le valutazioni.

3. Un esperimento scientifico / 1 ciclo, insegnamento secondario

R esoconto di un esperimento scientifico - Interpretazione e invenzione dei procedimenti di misura*

Descrizione dell'esercizio

1. Obiettivi

All'inizio dell'anno scolastico, gli alunni del primo anno dell'insegnamento secondario rinnovato sono stati invitati a:

- osservare un esperimento relativo alla durezza delle acque;
- interpretare i fenomeni osservati;
- pensare ad un metodo di misurazione.

2. Presentazione dell'esperimento

a. Introduzione

Partendo da osservazioni tratte dalla vita quotidiana si vuol dare una risposta alla domanda: «Si può usare qualsiasi tipo di acqua per qualsiasi uso?».

b. Esperimento

Il professore realizza un esperimento davanti a tutti gli alunni usando grandi recipienti. Le uniche informazioni fornite riguardano la natura dei prodotti utilizzati.

Materiale

- Tre provette graduate (A, B, C) da mezzo litro, contenenti ciascuna 490 cc. di acqua distillata
- Un recipiente con del sale solubile di calcio
- Una soluzione alcolica di sapone (Bourton-Bouled)
- Un misurino per il sale di calcio
- Un misurino per la soluzione di sapone

Svolgimento dell'esperimento

- Mettere una razione di sale di calcio nella provetta B
- Mettere due razioni di sale di calcio nella provetta C
- Aggiungere una razione di soluzione di sapone nella provetta A: la schiuma appare subito dopo aver agitato

* Hanno collaborato alla stesura di questo fascicolo i professori: Dighope, Ispettore scolastico; Lange, E.N.M.E.; Liege; Ruhwedel, E.N.M.E., Liege; Amel, A. R. Liége 12 Libourne A. R. Jupille; Pregaldien A. R., Vist. Ricercatore; J. Beckers e R. De Bal
• Aggiungere la stessa quantità di saponina nelle provette B e C: la schiuma non compare.
• Continuare a versare, contemporaneamente in B e in C, alcune razioni di soluzione di saponina finché non appare una schiuma persistente nelle due provette e si forma un precipitato.

**Compiti richiesti all’alunno**
1. Descrivete i fenomeni appena osservati
2. Illustrate l’esperimento con dei disegni
3. Spiegate i fenomeni osservati
4. Inventate uno o più procedimenti per misurare la quantità di sale di calcio contenuto nelle diverse acque.

**Perché si è scelto questo esercizio**
Questo esercizio richiede una performance complessa da parte degli alunni che hanno assistito allo svolgimento dell’esperimento.
La valutazione non può limitarsi al semplice conteggio di punti sulla base di errori commessi; bisogna mettere a punto uno strumento più adeguato.

**Dimensioni della valutazione**
La traccia dell’esercizio proposto indica questi criteri:
1) Relazione dell’esperimento con una descrizione scritta;
2) Relazione dell’esperimento per mezzo di schemi;
3) Interpretazione dei fenomeni osservati;
4) Inventione di un procedimento di misurazione.
Queste dimensioni si possono applicare a numerosi esercizi scientifici.

**Organizzazione del dossier**
Per ciascuna dimensione si presenta successivamente:
— La definizione del criterio
— La definizione degli aspetti interni a ciascun criterio.
Gli aspetti avranno sempre ugual peso nella valutazione.
— La precisazione dei livelli richiesti.
Le richieste, per ciascun grado della scala, non sono arbitrarie: sono state stabilite dopo approfondita analisi di molti lavori di alunni del primo anno dell’insegnamento secondario rinnovato.
— Il procedimento della valutazione.
— La presentazione della scala.
Questa informazione non apporta alcuna informazione complementare rispetto al procedimento. Tuttavia offre il vantaggio di dare una descrizione gerarchica della performance.

**PRESENTAZIONE DELLE SCALE**

**1. Relazione dell’esperimento con una descrizione scritta**

**Definizione del criterio**
Traduzione scritta dei fenomeni osservati.

**Definizione degli aspetti all’interno del criterio**
Si valutano successivamente tre aspetti: la completezza dell’esperimento, la precisione delle informazioni e il rispetto della cronologia.

**A. Un elaborato completo dovrà comportare queste informazioni:**
A1. Uguale quantità di acqua distillata in partenza (1).
A2. Quantità diverse di sale di calcio aggiunto ai due volumi di acqua.
A4. Dosii differenti di saponina che provocano l’apparizione di una schiuma persistente.

**B. Un elaborato preciso contiene le seguenti informazioni:**
B1. Valori relativi alle quantità di sale di Ca. (Esempio: semplice - doppio, 1 dose, 2 dose, 1 gr - 2 gr)

**C. La successione cronologica dei fenomeni è rispettata quando l’alunno presenta i fatti secondo l’ordine del loro svolgimento.**
Si esigerà almeno, che l’alunno presenti l’aggiunta di sale di calcio prima dell’aggiunta della saponina.

**Determinazione dei livelli richiesti**

**a. Per l’aspetto «Completezza», si considerano tre livelli:**
1. Sono presenti le informazioni A1, A2, A3, A4;
2. Sono presenti almeno le informazioni A2 e A4;

**b. Per l’aspetto «Precisione», si considerano tre livelli:**
1. Sono presenti le informazioni B1, B2;
2. È presente l’informazione B1 o B2;

**c. Per l’aspetto «Rispetto della cronologia**, si considerano due livelli:**
1. Sì
2. No.

**Procedimento per la valutazione della «Relazione scritta dell’esperimento»**
1. Valutate il primo aspetto: «Completezza»
   A1, A2, A3, e A4

2. Valutate il secondo aspetto: «Precisione»
   B1 e B2
   B1 o B2
   Né B1 né B2

3. Valutate il terzo aspetto: «Rispetto della cronologia»
   SI
   No

(1) Il presente la caratteristica A1 se l’alunno ricorda che si trattava di acqua distillata e che i volumi sono uguali.
Se manca o è errata una di queste due informazioni, il lavoro non comporta la caratteristica A1.
4. Rapportandovi allo schema che segue, scegliete il livello che corrisponde alle caratteristiche individuate nell'elaborato:

<table>
<thead>
<tr>
<th>Completezza</th>
<th>Precisione</th>
<th>Rispetto della cronologia</th>
</tr>
</thead>
<tbody>
<tr>
<td>- A1, A2, A3 e A4</td>
<td>- B1 e B2</td>
<td>- Si</td>
</tr>
<tr>
<td>- A1, A2, A3 e A4</td>
<td>- B1 o B2</td>
<td>- No</td>
</tr>
<tr>
<td>- Né B1 né B2</td>
<td>- Si</td>
<td>- No</td>
</tr>
<tr>
<td>- Né B1 né B2</td>
<td>- No</td>
<td>- No</td>
</tr>
</tbody>
</table>

- A2 e A4

<table>
<thead>
<tr>
<th>Completezza</th>
<th>Precisione</th>
</tr>
</thead>
<tbody>
<tr>
<td>- B1 e B2</td>
<td>- Si</td>
</tr>
<tr>
<td>- B1 o B2</td>
<td>- No</td>
</tr>
<tr>
<td>- Né B1 né B2</td>
<td>- Si</td>
</tr>
<tr>
<td>- Né B1 né B2</td>
<td>- No</td>
</tr>
</tbody>
</table>

- Manca A2 o A4

<table>
<thead>
<tr>
<th>Presentazione della scala «Relazione scritta sull'esperimento»</th>
</tr>
</thead>
<tbody>
<tr>
<td>Completezza</td>
</tr>
<tr>
<td>-------------</td>
</tr>
<tr>
<td>MB</td>
</tr>
<tr>
<td>B</td>
</tr>
<tr>
<td>A2 e A4</td>
</tr>
<tr>
<td>S</td>
</tr>
<tr>
<td>A2 e A4</td>
</tr>
<tr>
<td>D</td>
</tr>
<tr>
<td>A2 e A4</td>
</tr>
<tr>
<td>I</td>
</tr>
</tbody>
</table>

Voto per la descrizione: ☐

II. Relazione dell'esperimento per mezzo di schemi

Definizione del criterio

Si valutano sia gli schemi realizzati dall'alunno che le annotazioni di completa-
mento.
N.B.: L'annotazione di un elemento identico per le differenti parti dello schema
non deve essere necessariamente ripetuta.

Definizione degli aspetti interni al criterio

Si valutano successivamente tre aspetti dello schema all'alunno: la completezza,
la precisione, la dinamicità.

1. Completezza

Lo schema è completo quando, sotto forma di disegni o di annotazioni, dà le se-
guenti informazioni:
A1. All'inizio dell'esperimento, uguale acqua distillata nelle tre provette.
A2. Aggiunta di dosi diverse di sale di calcio nelle provette B e C.
A4. Apparizione della schiuma nelle tre provette dopo aver aggiunto sapore necessa-
rio (1).

A5. Diversa opacità della soluzione nelle provette B e C.

2. Precisione

Si possono completare le informazioni A2 e A3 specificando quantitativamente i
valori relativi...

B1. Alle quantità di sale di calcio (1 dose · 2 dosi);

Un'informazione importante — la necessità di rappresentare tre provette identi-
che — non è stata valutata: è presente in tutte le copie del campione preso in esame.

3. Dinamicità

I disegni dell'alunno sono dinamici se evidenziano le tappe dell'esperimento, con
numerosi disegni di schemi.

Per svolgimento completo dell'esperimento si intende la rappresentazione di que-
te tappe:
1. Tre provette identiche contenenti uguale volume di acqua distillata (stadio iniziale).
2. Le tre provette, con l'aggiunta di sale di calcio nelle provette B e C.
3. Le tre provette con aggiunta di sapore e apparizione della schiuma. (Per questa fa-
se dell'esperimento l'alunno può evidenziare che dopo l'aggiunta di ciascuna dose di
sapore si agita la provetta per osservare l'apparizione eventuale di schiuma persistente).
4. Tre provette con differente opacità (stadio finale)

L'alunno deve menzionare almeno i punti 2, 3 e 4. Si accetteranno raggruppamen-
ti di informazioni relative ai punti 2 e 3, o 3 e 4, in una stessa serie di disegni.

(1) Alcuni alunno usano il termine «sapone» per indicare la schiuma. In questo caso è presente l'informazione A4
e manca A3.
Determinazione dei livelli richiesti

a. Per l’aspetto «Completezza» si considerano tre livelli:

b. Per l’aspetto «Precisione» si considerano due livelli:

c. Per l’aspetto «Dinamicità» si considerano due livelli:
   1. Sì
   2. No.

Procedimento per la valutazione della «Relazione dell’esperimento con uno schema»

1. Valutare il primo aspetto: «Completezza»
   - Tutte le informazioni A
   - Almeno A2, A3 e A4
   - Manca A2, A3, o A4

2. Valutare il secondo aspetto: «Precisione»
   - B1 e B2
   - Manca B1 o B2

3. Valutare il terzo aspetto: «Dinamicità»
   - Sì
   - No

4. Rapportandovi allo schema che segue, scegliete il livello che corrisponde alle tre caratteristiche individuate nell’elaborato dell’allievo.

<table>
<thead>
<tr>
<th>Completezza</th>
<th>Precisione</th>
<th>Dinamicità</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tutte le informazioni A</td>
<td>B1 o B2</td>
<td>MB</td>
</tr>
<tr>
<td></td>
<td>Manca B1 o B2</td>
<td>B</td>
</tr>
<tr>
<td>Almeno A2, A3 e A4</td>
<td>B1 e B2</td>
<td>S</td>
</tr>
<tr>
<td>Manca A2, A3, o A4</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Voto per lo schema: □

III. Interpretazione dell'esperimento

Definizione del criterio

Si valutano i commenti dell'allievo che, oltre alla descrizione delle tappe dell'esperimento, sono finalizzati ad evidenziare i rapporti tra i fenomeni osservati per dare loro un significato.

Definizione degli aspetti interni al criterio

Si valutano successivamente due aspetti: completezza e precisione.

1. Completezza

L'interpretazione dell'allievo è completa se mette in rapporto:
A1. La presenza del sale di calcio e la quantità di sapone necessaria per far apparire la schiuma;
A2. La presenza del sale di calcio e la formazione di un precipitato.

2. Precisione

a) I fenomeni osservati possono essere semplicemente messi in relazione:
   - La presenza del sale di calcio richiede l'aggiunta di una maggior quantità di sapone;
   - La formazione di un precipitato dipende dalla presenza di un sale di calcio.
Senza necessariamente utilizzare una di queste formulazioni, l'alunno si limita, in questo caso, a esprimere la presenza dei due fenomeni.

b. Il rapporto può essere precisato sotto forma di funzione:
- Più sale di calcio c'è, più sapone è necessario per produrre la schiuma;
- L'importanza del precipitato è direttamente proporzionale alla quantità di sale di calcio.

Rispetto alla situazione precedente, l'alunno mette in relazione i gradi nell'intensità dei fenomeni osservati.

c. Il primo rapporto (sali di Ca. — sapone) può assumere la forma di una proporzione:
- Per una porzione di sale di Ca., occorrono 6 misure di sapone;
- Per 2 porzioni di sale di Ca., occorrono 12 misure di sapone.

Le quantità di sali di calcio e di sapone variano nel rapporto 1 a 6.

Qui i gradi sono quantificati con l'intervento delle unità di misura che permettono all'alunno di esprimere il rapporto di grandezza tra sali di calcio e sapone.

N.B. — Le condizioni sperimentali, basate su un solo esempio, rendono questo tipo di risposta poco probabile.

L'analisi di un gran numero di copie evidenzia del resto che questa formulazione è rara al primo anno dell'insegnamento secondario rinnovato. Pertanto non verrà considerata nella scala. Il professore può comunque valorizzare l'eventuale presenza di questa formulazione con un commento appropriato.

Nella scala, questa performance verrà classificata all'interno della categoria «Informazioni espresse almeno sotto forma di funzione».

Determinazione dei livelli richiesti
I due aspetti sopra presentati sono indissolubili. Pertanto si presenterà un livello soltanto (risultante dai due aspetti). Il procedimento da seguire viene semplificato. I livelli appaiono chiaramente nella scala.

Presentazione della scala «Interpretazione dell'esperimento»

<table>
<thead>
<tr>
<th>MB</th>
<th>Almeno A1 e A2 sotto forma di funzione</th>
</tr>
</thead>
<tbody>
<tr>
<td>B</td>
<td>Almeno A1 sotto forma di funzione e A2 sotto forma di relazione</td>
</tr>
<tr>
<td>S</td>
<td>Almeno A1 sotto forma di funzione</td>
</tr>
<tr>
<td></td>
<td>Almeno A2 sotto forma di relazione</td>
</tr>
<tr>
<td>D</td>
<td>Almeno A1 sotto forma di relazione</td>
</tr>
<tr>
<td></td>
<td>A2 sotto forma di relazione</td>
</tr>
<tr>
<td>I</td>
<td>Nessuna interpretazione</td>
</tr>
</tbody>
</table>

N.B. — Per i primi due criteri era riservata una pagina speciale nel procedimento di valutazione. Per conservare il criterio analitico alla valutazione, il professore sovrintende nella scala la caratteristica corrispondente al lavoro valutato.

Voto per l'interpretazione: □

IV. Invenzione di un metodo di misura

Definizione del criterio
Il procedimento di invenzione suppone, oltre l'osservazione e l'interpretazione dei fenomeni, la scelta di variabili pertinenti che permettano la messa a punto di uno schema generale.

L'invenzione di un metodo di misura presuppone un'interpretazione corretta dell'esperimento osservato. Perché una misura reale sia possibile occorre che l'alunno riesca a cogliere dall'osservazione la proporzione che lega la quantità di sale con il sapone.

Tuttavia non abbiamo già sottolineato che l'esperimento, così come è stato condotto, non consente di determinare questa proporzione. Però ci si contenterà che venga enunciato il principio generale: la quantità di sapone necessaria per produrre una schiuma è deducibile dalla quantità di sale presente nell'acqua.

L'applicazione di questo principio suppone la scelta di un'unità di misura e del procedimento da seguire.

Il solo metodo di misura possibile è dunque fondato sul rapporto tra sale e sale di calcio. Per contro, la valutazione dell'opacità del precipitato non può consistere che in una stima.

Il metodo fondato sull'evaporazione del liquido e la valutazione del residuo di sale di calcio rimasto al fondo della provetta è buono, ma inappropriato al contesto di questo esperimento.

Quando l'alunno propone più procedimenti, si valuterà quello più elaborato.

Presentazione della scala «Invenzione di un metodo di misura»
La valutazione di questa dimensione non comporta il ricorso a numerosi aspetti complementari. Così il procedimento da seguire è semplificato. I livelli richiesti appaiono direttamente nella scala.

<table>
<thead>
<tr>
<th>MB</th>
<th>Scoperta</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>del principio generale (deduzione della quantità di sale, partendo dalla quantità di sapone)</td>
</tr>
<tr>
<td></td>
<td>della necessità di un'unità di misura</td>
</tr>
<tr>
<td></td>
<td>del procedimento di utilizzazione di questa unità</td>
</tr>
<tr>
<td>B</td>
<td>Scoperta</td>
</tr>
<tr>
<td></td>
<td>del principio generale</td>
</tr>
<tr>
<td></td>
<td>della necessità di un'unità di misura</td>
</tr>
<tr>
<td>S</td>
<td>Scoperta del principio generale</td>
</tr>
<tr>
<td>D</td>
<td>Stima della quantità di sale di calcio per valutare l'opacità del precipitato</td>
</tr>
<tr>
<td>I</td>
<td>Ripetizione dell'interpretazione</td>
</tr>
</tbody>
</table>

N.B. — Come per l'interpretazione, sottolineare nella scala le caratteristiche corrispondenti al lavoro valutato.

Voto per l'invenzione di un metodo di misura: □
Sintesi delle valutazioni

Avete valutato alcune dimensioni. Riportate i voti attribuiti sulle scale che seguono:

1. Relazione dell'esperimento con una descrizione scritta

<table>
<thead>
<tr>
<th>MB</th>
<th>B</th>
<th>S</th>
<th>D</th>
<th>I</th>
</tr>
</thead>
</table>

2. Relazione dell'esperimento per mezzo di schemi

<table>
<thead>
<tr>
<th>MB</th>
<th>B</th>
<th>S</th>
<th>D</th>
<th>I</th>
</tr>
</thead>
</table>

3. Interpretazione dei fenomeni osservati

<table>
<thead>
<tr>
<th>MB</th>
<th>B</th>
<th>S</th>
<th>D</th>
<th>I</th>
</tr>
</thead>
</table>

4. Invenzione dei procedimenti di misura

<table>
<thead>
<tr>
<th>MB</th>
<th>B</th>
<th>S</th>
<th>D</th>
<th>I</th>
</tr>
</thead>
</table>

Voto finale

<table>
<thead>
<tr>
<th>MB</th>
<th>B</th>
<th>S</th>
<th>D</th>
<th>I</th>
</tr>
</thead>
</table>

Scieme : Relazione dell'esperimento

Gisca C.
I A.

Ci sono tre provette sul tavolo, vi si versa una quantità uguale di acqua distillata. Dopo ciò, si versa nella seconda provetta una piccola dose di sale di calcio e nella terza due dosi di sale di calcio. Tutte e tre le provette contengono 490 cc di acqua distillata.
La prima provetta, ancora piena di acqua pura, contiene anche 1 quantità da 10 cc. di sapore liquido. La stessa quantità è versata 6 volte nella seconda e 13 volte nella terza.
Abbiamo il seguente risultato:
La prima provetta è ancora trasparente e in superficie c'è la schiuma perché il sapore è statoagitato in tutti e tre i casi.
La seconda provetta graduata è di colore biancastro ma non trasparente e in superficie c'è un po' di schiuma; il volume è più alto rispetto alla prima provetta, per la maggior quantità di sapore immesso (60 cc).
La terza provetta è più bianca della seconda ed ha un volume più grande.
La schiuma è minore.
La schiuma

una dose di sale di calcio

6 dosi di sale di calcio

13 dosi di sale di calcio

Ecco il colore che risulta dopo aver miscelato il contenuto.

trasparente

biancastro

bianco

Lo si può intravedere nell'acqua distillata perché essa non lo contiene, per questo il sapone si mischia e fa la schiuma mentre nelle altre due provette, che contengono sale la schiuma si forma in minor quantità.

Si aggiunge sapone nelle differenti acque. Si agita. Se appare la schiuma, vuol dire che non c'è sale nell'acqua se non appare è perché c'è sale.

4. Scienze / II ciclo, insegnamento secondario

Osservazione di un esperimento, interpretazione, invenzione di un metodo di analisi*

Descrizione dell'esercizio

a. Obiettivi dell'esercizio

All'inizio dell'anno scolastico, gli alunni del 4° anno dell'insegnamento secondario rinnovato sono invitati a:

- osservare un esperimento;
- interpretare i fenomeni osservati;
- immaginare un metodo di analisi.

Si tratta di un esperimento di semplice sostituzione: sostituire alcuni alogeni con altri. Introduce alla nozione di graduazione sul potere ossidante degli elementi e sulle reazioni redox di semplice sostituzione.

L'esercizio prepara all'acquisizione di nozioni teoriche importanti che verranno sviluppate ulteriormente. Qui è necessario partire da una situazione che non richieda alcuna preliminare conoscenza del soggetto.

b. Presentazione schematica dell'esperimento

1. Presentazione dei prodotti

Nell'esperimento si utilizzano i seguenti prodotti:

- 3 reagenti semplici
  Iodio (I₂)
  Banchio (Br₂)
  Cloro (Cl₂)

- 3 reagenti composti:
  Ioduro di potassio (KI)
  Bromuro di potassio (KBr)
  Cloruro di potassio (KCl)

- 2 solventi:
  Acqua (H₂O)
  Tetracloruro di carbonio (CCl₄)

2. Preparativi

Perché l'alunno possa comprendere questo esperimento, è opportuno procedere a delle manipolazioni preliminari. I risultati di questi preparativi costituiscono dei riferimenti indispensabili per l'alunno.

- Presentazione dei due solventi (H₂O e CCl₄):

* Hanno collaborato alla stesura di questo fascicolo, i professori: Sig. Bigsby, Insegnante scolastico per l'insegnamento delle scienze; Aniel A. R. de Montegnée, Laige; E.M.E.; Laige; Bouschat, A. R. Jacbier; Dedoyard, A. R. Anglert; Guelle, A. R. Ahker; Otto, A. R. Laige II; Piratte, A. R. Liige I; Ricreato; J. Becker, R. De Bal.
— Presentazione dei due solventi con ciascun reattivo — reagenti semplici;
— Presentazione dei due solventi con ciascun reattivo — reagenti composti;

3. Esperimento
Si prendono i reattivi due a due (un reagente semplice ed uno composto):
— all'inizio in una soluzione acquosa;
— poi con l'aggiunta di CCl₄

c. Condizioni e svolgimento dell'esperimento
1. Requisiti preliminari
Prima di iniziare l'esercizio, l'alunno deve possedere queste nozioni:
— Reattivo — reazione o fenomeno chimico.
— Solvente — soluzione acquosa — insolubilità — miscibilità di due liquidi.
— Corpo semplice — corpo composto.
— Elemento.

2. Manipolazioni preliminari
Il professore prepara per tempo le viti provette (cfr. direttive) e le presenta agli
alunni. Durante le spiegazioni a proposito di ciascuna e completara' man mano una
tavola simile a quella che segue (Cfr. Tav. I).
La gamma dei colori e presentata in modo tale che venga soppressa ogni ambigui-
ta nei termini usati. Si raccomanda di apporre su ciascun tubo un'etichetta che in-
forma l'alunno del contenuto. Esempio: per la provetta 8: I₂, H₂O, CCl₄.

<table>
<thead>
<tr>
<th>Solventi</th>
<th>H₂O</th>
<th>CCl₄</th>
<th>H₂O + CCl₄</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reagenti</td>
<td>incolore</td>
<td>incolore</td>
<td>incolore (2)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>fase superiore</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>fase inferiore</td>
</tr>
<tr>
<td>Corpi semplici</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Br₂</td>
<td>giallo-arancio (3)</td>
<td>giallo-bruno (8)</td>
<td>giallo</td>
</tr>
<tr>
<td></td>
<td>incolore-giallo-</td>
<td>incolore (10)</td>
<td>incolore bruno</td>
</tr>
<tr>
<td></td>
<td>giallo</td>
<td></td>
<td>giallo-bruno</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>incolore giallastro</td>
</tr>
<tr>
<td>Cl₂</td>
<td>rosso (4)</td>
<td>viola (7)</td>
<td>rosso (5)</td>
</tr>
<tr>
<td></td>
<td>incolore (11)</td>
<td></td>
<td>incolore (11)</td>
</tr>
<tr>
<td>Corpi composti</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| KBr | incolore (limpido) | incolore + cri-
| | (12) | stalli (13) |
| KCl | incolore (limpido) | incolore + cri-
| | (15) | stalli (16) |
| | incolore (limpido) | incolore + cri-
| | (18) | stalli (19) |
| | | incolore | incolore (14) |
| | | | incolore (17) |
| | | | incolore (20) |

N.B. Nelle direttive che seguono verranno date informazioni relative alla preparazione di alcuni prodotti. I nu-
meri tra parentesi rinviano a queste direttive.

Osservazioni
1. A questo punto puo' essere utile far osservare agli alunni:
— che tutti i fenomeni incontrati fino ad ora sono fenomeni fisici; le sostanze incontrate sono rimaste inalterate;
— che le reazioni possono verificarsi nelle successive manipolazioni.

2. Va da se che, per i lavori pratici, suddivisi tra le varie scienze, saranno sollecitata la
partecipazione diretta degli alunni, per le manifestazioni precedenti e, soprattutto, per quelle che seguiranno.

Direttive per la preparazione dei prodotti
Si consiglia l'uso di prodotti specifici «per analisi».

(2) Presentare una provetta con 1 volume di acqua e una con 1 volume, per metà al-
meno, di CCl₄.
Aggiungere lentamente all'acqua il CCl₄; in modo che gli alunni lo vedano scende-
re in fondo alla provetta e condurlo sulla non miscibilità dei due solventi.
Agitare e far constatare che il CCl₄ discende nuovamente (1).
Se gli alunni osservano il formarsi di un "menisco" al di sopra della fase acquosa, far
presente che si tratta di un fenomeno fisico senza interesse per l'esperimento.
Tutte le soluzioni descrivute qui di seguito non vanno preparate troppo presto. Le
loro concentrazioni, verificate con prove preliminari, devono essere tali che i colori
ottenuti durante le reazioni siano nitidi e identificabili senza alcun dubbio.

(36-9) A sua discrezione, l'insegnante puo' mostrare lo iodio, il bromo ed il cloro allo
stato puro, per soddisfare la curiosita' scientifica degli alunni. Siccome i vari passag-
gi verranno fatti con le loro soluzioni acquose, per sicurezza, ci si può limitare a pre-
sentare e descrivere queste ultime.

(3) Presentare una soluzione satura di bromo.

(6) Aggiungere alcuni cristalli di ioduro di potassio.

(9) Usare una soluzione satura di cloro.

(457-8:10:11) Partire da una soluzione acquosa satura e aggiungere un volume di
CCl₄ Prelevare metà del CCl₄ (servendosi di un'ampolla di decantazione) per preparare le
soluzioni 4-7-10 e conservare il rimanente che forma le soluzioni 5-8-11.

(6) Aggiungere alcuni cristalli di ioduro di potassio, per aumentare la solubilità dello
iodio.

(9) Usare una soluzione satura di cloro.

(1) Questa osservazione vale per tutti i passaggi in cui si usa il CCl₄.
(12-15-18) Usare soluzioni al 5%.
(13-16-19) Alcuni cristalli in una stessa quantità di CCl₄.

3. Esperimento

Le questioni poste agli alunni verteranno esclusivamente su questa parte dell'esercizio. Saranno presentati nuovi esperimenti per ogni gruppo di tre alunni.

Ciascun esperimento sarà condotto in due tempi:
a. in soluzione acqua;
b. aggiungendo CCl₄ alla miscela delle due soluzioni acquose per riconoscere l'elemento semplice presente fin dall'inizio della reazione.

Prima di ogni esperimento il professore indica nella Tabella i prodotti che verranno utilizzati. Esempio:

\[ A_1 = \text{Sol KBr} + \text{acqua di cloro} \]

Dopo ogni esperimento, l'alunno annota le sue osservazioni, relative ai due momenti dell'esperimento.

<table>
<thead>
<tr>
<th>Gruppo A</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1 = Sol KBr + acqua di cloro</td>
</tr>
<tr>
<td>A2 = Sol KBr + acqua di bruno</td>
</tr>
<tr>
<td>A3 = Sol KBr + acqua di iodio (soluzione di iodio-iodurato)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Gruppo B</th>
</tr>
</thead>
<tbody>
<tr>
<td>B1 = Sol KI + acqua di cloro</td>
</tr>
<tr>
<td>B2 = Sol KI + acqua di bruno</td>
</tr>
<tr>
<td>B3 = Sol KI + acqua di iodio (soluzione di iodio-iodurato)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Gruppo C</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1 = Sol KCl + acqua di cloro</td>
</tr>
<tr>
<td>C2 = Sol KCl + acqua di bruno</td>
</tr>
<tr>
<td>C3 = Sol KCl + acqua di iodio (soluzione di iodio-iodurato)</td>
</tr>
</tbody>
</table>

La quantità delle due soluzioni di reagenti e del tetrachloruro di carbonio devono essere sostanzialmente uguali, in modo che al termine il volume di quest'ultimo sia approssimativamente un terzo del totale.

Il professore avrà cura di agitare le provette specialmente dopo aver aggiunto CCl₄.

Nelle condizioni descritte, l'esperimento deve portare ai seguenti risultati:

### TAVOLA II

<table>
<thead>
<tr>
<th>Reagenti composti</th>
<th>KBr</th>
<th>KI</th>
<th>KCl</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reagenti semplici</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>H₂O</td>
<td>CCl₄</td>
<td>H₂O</td>
</tr>
<tr>
<td>I₂</td>
<td>giallo-bruno</td>
<td>giallo-bruno</td>
<td>giallo-bruno</td>
</tr>
<tr>
<td></td>
<td>viola I₂</td>
<td>viola I₂</td>
<td>viola I₂</td>
</tr>
<tr>
<td>Br₂</td>
<td>giallo-bruno</td>
<td>giallo-arancio</td>
<td>giallo-bruno</td>
</tr>
<tr>
<td></td>
<td>rosso Br₂</td>
<td>viola I₂</td>
<td>viola I₂</td>
</tr>
<tr>
<td>Cl₂</td>
<td>giallo-bruno</td>
<td>giallo-arancio</td>
<td>giallo-bruno</td>
</tr>
<tr>
<td></td>
<td>rosso Br₂</td>
<td>viola I₂</td>
<td>viola I₂</td>
</tr>
</tbody>
</table>

Questa tavola è presentata a titolo indicativo per dare un'immagine completa dell'esperimento. Non viene comunicata agli alunni. Le presentazione delle osservazioni in una tabella a doppia entrata non è stata richiesta, d'altra parte, dagli insegnanti.

4. Domande poste agli alunni.

1. Cosa osservate nell'acqua e nel CCl₄?
L'alunno può scegliere il modo di presentare la risposta.

2. Per ciascun esperimento, spiegate le modificazioni di colore osservate.
Il professore precisionerà che non si tratta qui di scrivere le equazioni.

3. Scrivete le equazioni che spiegano i fenomeni chimici osservati.
Il professore scriverà alla lavagna, sotto forma di primo membro dell'equazione, i reagenti evidenziati.

A1 : KBr + Cl₂
A2 : KBr + Br₂
A3 : KBr + I₂
B1 : KI + Cl₂
B2 : KI + Br₂
B3 : KI + I₂
C1 : KCl + Cl₂
C2 : KCl + Br₂
C3 : KCl + I₂

(Segue)

(1) La disposizione scelta nella tavola rappresenta le due fasi osservate nella provetta: in alto, la fase acquea; in basso, la fase organica.
(2) Leggere variazioni di colore si possono osservare tra la sola fase acquea e la fase superiore dopo aver aggiunto CCl₄. Si spiegano per la minore concentrazione di reagenti nella fase superiore in CCl₄.
4. a) Quale atomo è accostato meno bene al potassio? Perché?
b) Quale è accostato meglio? Perché?
c) Stabilite una classificazione di tre elementi secondo la loro forza decrescente.

Il termine accostato è usato intenzionalmente: l'esercizio proposto è un approccio a nozioni teoriche importanti che saranno sviluppate ulteriormente.

5. Può darsi che una soluzione contenga ioduro; pensate ad un esperimento che ne permetta la verifica.

Giustificazione della scala di questo esercizio

La scala di valutazione è uno strumento destinato a valutare performances complesse per le quali risulta inadeguato un semplice conteggio numerico di punti.

Tra le performances complesse che si possono individuare nell'insegnamento secondario, il riferire ed interpretare una esperienza rappresenta un apprendimento fondamentale, a giudizio dei professori.

Scelta dei criteri di valutazione

1° Relazione dei fenomeni osservati
2° Interpretazione degli stessi
3° Invenzione di un metodo di analisi.

Naturalmente, una valutazione condotta secondo questi tre criteri non è esaustiva. Verte su ciò che è specifico per questo genere di esercizio e non sugli aspetti formalì (stile, ortografia, ...).

Organizzazione del dossier

Per ciascun criterio considerato si presenterà successivamente:
— La definizione;
— La definizione degli aspetti subentranti;
— La determinazione dei livelli richiesti.

Le richieste corrispondenti a ciascun grado della scala non sono arbitrarie: sono state stabilite dopo aver analizzato accuratamente molti lavori di alunni iscritti al quarto anno dell'insegnamento secondario rinnovato.
— Il procedimento della valutazione;
— La scala.

La scala non apporta informazioni complementari rispetto al procedimento già descritto. Tuttavia, la scala offre il vantaggio di dare una descrizione gerarchica della performance.

PRESENTAZIONE DELLE SCALE

1. Relazione dell’esperimento

Definizione del criterio

Relazione di un esperimento, vuol dire riferire, con un qualche mezzo (brano, schena, disegni, tabella a doppia entrata, ecc. ...) i fenomeni osservati.

Definizione degli aspetti interni al criterio

Si valuteranno due aspetti:
— La relazione dei fenomeni essenziali all’interpretazione dell’esperimento. Vengono valutate soltanto le informazioni minime, sufficienti a comprendere l’esperimento.
— Per le situazioni correttamente riferite, se ne valuterà la precisione. Si valORIZZANO in questo modo gli elementi del lavoro dell’alluno che, senza essere essenziali all’interpretazione dell’esperimento, evidenziano una più rigorosa osservazione.

I due aspetti avranno uguale importanza nella valutazione della relazione dell’esperimento.

1. Relazione dei fenomeni essenziali per l’interpretazione dell’esperimento

Un lavoro complesso comporterà la corretta descrizione dei nove esperimenti descritti precedentemente.

Nella valutazione, errore e omissione hanno la stessa importanza.

L’esperimento è complesso e correttamente riferito quando descrive i fenomeni osservati sia:
— nella soluzione acquosa (da sola o nella fase superiore dopo aver aggiunto CCl₄);
— sia il CCl₄ (fase inferiore).

Per valutare ciascun esperimento ci si rapporterà alla gamma di colori presentata nella Tavola II.

2. Precisione

La valutazione di questo aspetto verte sulle situazioni correttamente descritte.

a. Distinzione tra fase acquisa e fase organica

Un lavoro preciso differenziando la fase inferiore dalla fase superiore, dopo l’aggiunta di CCl₄. Inoltre, l’alluno deve localizzare almeno una volta il CCl₄ nella fase inferiore.

(1) Se in un esperimento, il lavoro presenta discordanza nell’identificazione dei colori in soluzione acquisa, e nella fase superiore in CCl₄, non si considererà errore, se una delle due osservazioni è errata. La discordanza verrà sottolineata altrove (cfr. in seguito: assenza di ripetizione).
b. Ripetizione dei fenomeni osservati in soluzione acquosa.

Sarà valorizzato ogni elaborato che, dopo aver descritto correttamente i fenomeni osservati esclusivamente in soluzione acquosa, riporta la fase superiore dopo l’aggiunta di CCl₄.

La ripetizione è possibile soltanto se l’alunno stabilisce precedentemente la distinzione tra fase acquosa e organica.

**Determinazione dei livelli richiesti**

a. Per il primo aspetto «Relazione dei fenomeni essenziali», saranno considerati cinque livelli:
- Nessun errore o omissione
- Un errore o omissione
- Due errori o omissioni
- Più di due errori o omissioni
- Nessun riferimento a una delle due fasi.

b. Per il secondo aspetto, «Precisione», saranno considerati i livelli:
- Distinzione: Si
- No
- Ripetizione: Si
- No

**Procedimento per la valutazione del criterio: «Relazione dei fenomeni osservati»**

1. Valutate il primo aspetto: Relazione dei fenomeni essenziali
- Nessun errore
- Un errore (?)
- Due errori (?)
- Più di due errori (?)
- Nessun riferimento a una delle due fasi

2. Valutate il secondo aspetto: Precisione.
   a. Distinzione
   b. Ripetizione

3. Rapportandovi allo schema che segue, scegliete il livello corrispondente alle tre caratteristiche valutate nell’elaborato dell’alunno.

**Presentazione della scala «Relazione dell’esperimento»**

<table>
<thead>
<tr>
<th>Relazione dei fenomeni essenziali</th>
<th>Precisione</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Distinzione</td>
</tr>
<tr>
<td><strong>MB</strong> Nessun errore</td>
<td>Si</td>
</tr>
<tr>
<td>B</td>
<td>Nessun errore</td>
</tr>
<tr>
<td>Un errore</td>
<td>Si</td>
</tr>
<tr>
<td>S</td>
<td>Nessun errore</td>
</tr>
<tr>
<td>Un errore</td>
<td>Si</td>
</tr>
<tr>
<td>D</td>
<td>Un errore</td>
</tr>
<tr>
<td>I</td>
<td>Due error</td>
</tr>
<tr>
<td></td>
<td>Due error</td>
</tr>
<tr>
<td></td>
<td>Più di due error, Nessun riferimento a una delle due fasi</td>
</tr>
</tbody>
</table>

(?) Indicate l’ipotesi delle situazioni indicate o non corrette.

---

II. Interpretazione dei fenomeni osservati

**Definizione del criterio**

Interpretare vuol dire mettere in relazione i fenomeni osservati per dar loro un significato.
Definizione degli aspetti interni al criterio
Le risposte alle domande 2, 3 e 4 si riferiscono allo stesso processo; saranno per tanto valutate su una medesima scala.

La valutazione verterà su tre aspetti:
- Identificazione dei corpi semplici presenti al termine dell'esperimento;
- La distinzione tra i fenomeni fisici e reazioni chimiche;
- Le conclusioni sulla forza specifica degli alogeni.

I Identificazione dei corpi semplici

L'identificazione corretta dei corpi semplici presenti al termine dell'esperimento presuppone che l'alluno abbia stabilito un rapporto preciso tra colore osservato nel CO₂ e natura del corpo semplice.

- In tre esperimenti (KI + Br₂ - KI + Cl₂ - KBr + Cl₂), il corpo da identificare si è formato a seguito di una reazione chimica (esperimento del tipo I).
- Per gli altri sei esperimenti, il corpo da identificare era stato già introdotto durante la manipolazione: non c'è stata reazione chimica (esperimento del tipo II).

Nella scala si darà un'importanza differente a queste due categorie.

II Distinzione tra fenomeni fisici e reazioni chimiche
L'alluno può stabilire la distinzione in due modi: con un testo scritto o con le equazioni.

a. Per iscritto
- Per ogni esperimento, l'alluno fa riferimento alla presenza o assenza del corpo identificato nei passaggi e conclude che c'è o non c'è reazione. Esempio:

<table>
<thead>
<tr>
<th>Identificazione</th>
<th>Distinzione</th>
</tr>
</thead>
<tbody>
<tr>
<td>KCl + Br₂</td>
<td>rosso</td>
</tr>
<tr>
<td>Br₂ → Br</td>
<td>Brosso</td>
</tr>
<tr>
<td>I₂ non introdotto nel passaggio</td>
<td>nessuna reazione</td>
</tr>
</tbody>
</table>

b. Per mezzo di equazioni.
L'alluno prova a stabilire, per mezzo delle equazioni, la distinzione tra reazioni chimiche e fenomeni fisici se, tra le 9 equazioni di cui gli è stato proposto il primo membro:
- non completa il secondo o ripete il primo per tutti gli esperimenti del II tipo,
- e la appare il corpo formato attraverso i tre esperimenti del I tipo.

III Conclusioni sulla forza specifica degli alogeni
Questo aspetto riguarda la valutazione delle domande 4a - 4b - 4c.

a. Classificazione
Si ritiene che l'alluno stabilisce correttamente la classificazione degli alogeni se risponde bene alle domande 4a, 4b, 4c. Tutti gli altri casi saranno valutati negativamente.

b. Giustificazione
Per essere corrette le giustificazioni devono comportare queste informazioni:
- I₂ non è bene accostato, perché lo iodio è stato posato dal Br₂ e dal Cl₂.
- Cl₂ è accostato meglio, perché il cloro non è stato posato dal Br₂.

L'errore in una delle due giustificazioni comporta una valutazione negativa del lavoro. Per contro, essendo equivalenti le due giustificazioni, si tollerà l'omissione dell'una o dell'altra.

Determinazione dei livelli richiesti

a. Per l'aspetto «Identificazione dei corpi semplici», saranno considerati tre livelli:
- I corpi sono correttamente identificati nei tre esperimenti del I tipo e in almeno cinque esperimenti del II tipo (si tollera perciò un errore per questo tipo di esperimento);
- I corpi sono correttamente identificati nei tre esperimenti del I tipo e in meno di cinque esperimenti del II tipo;
- I corpi non sono correttamente individuati nei tre esperimenti del I tipo.

b. Per l'aspetto «Distinzione tra fenomeni fisici e reazioni chimiche» saranno considerati due livelli:
- Si
- No

c. Per l'aspetto «Conclusioni sulla forza specifica degli alogeni», saranno considerati i livelli:
- Classificazione corretta / non corretta;
- Giustificazione corretta / non corretta.

Procedimento per la valutazione del criterio «Interpretazione dei fenomeni osservati»
— I corpi sono correttamente identificati:
  • nei tre esperimenti del tipo A e in almeno cinque esperimenti del tipo B
— I corpi sono correttamente identificati:
  • nei tre esperimenti del tipo A
  • in meno di cinque esperimenti di tipo B
— I corpi non sono identificati correttamente nei tre esperimenti di tipo A

2. Valutate il secondo aspetto: Distinzione tra fenomeni fisici e reazioni chimiche (col testo o con le equazioni).
— L'alunno opera la distinzione:
  — Si ☐
  — No ☐

3. Valutate il terzo aspetto: Conclusioni sulla forza specifica degli alogeni.
  a. Classificazione
     — corretta ☐
     — non corretta ☐
  b. Giustificazione: c'è almeno una giustificazione corretta
     — Si ☐
     — No ☐

4. Rapportandovi allo schema che segue, scegliete il livello corrispondente alle caratteristiche valutate nel lavoro dell'alunno.

**Forza specifica degli alogeni**

<table>
<thead>
<tr>
<th>Identificazione dei componenti</th>
<th>Distinzione</th>
<th>Classificazione</th>
<th>Giustificazione</th>
</tr>
</thead>
<tbody>
<tr>
<td>I corpi sono correttamente identificati nei tre esperimenti di tipo A e in almeno cinque esperimenti di tipo B</td>
<td>Effettuata</td>
<td>corretta</td>
<td>corretta</td>
</tr>
<tr>
<td>I corpi sono correttamente identificati nei tre esperimenti del tipo A e in meno di cinque esperimenti di tipo B</td>
<td>Non effettuata</td>
<td>non corretta</td>
<td>corretta</td>
</tr>
<tr>
<td>I corpi non sono identificati correttamente nei tre esperimenti di tipo A</td>
<td>Effettuata</td>
<td>corretta</td>
<td>corretta</td>
</tr>
<tr>
<td>I corpi sono correttamente identificati nei tre esperimenti del tipo A e in meno di cinque esperimenti di tipo B</td>
<td>Non effettuata</td>
<td>non corretta</td>
<td>non corretta</td>
</tr>
<tr>
<td>I corpi sono correttamente identificati nei tre esperimenti del tipo A e in meno di cinque esperimenti di tipo B</td>
<td>Effettuata</td>
<td>corretta</td>
<td>corretta</td>
</tr>
<tr>
<td>I corpi sono correttamente identificati nei tre esperimenti del tipo A e in meno di cinque esperimenti di tipo B</td>
<td>Non effettuata</td>
<td>non corretta</td>
<td>non corretta</td>
</tr>
<tr>
<td>I corpi sono correttamente identificati nei tre esperimenti del tipo A e in meno di cinque esperimenti di tipo B</td>
<td>Effettuata</td>
<td>corretta</td>
<td>corretta</td>
</tr>
<tr>
<td>I corpi sono correttamente identificati nei tre esperimenti del tipo A e in meno di cinque esperimenti di tipo B</td>
<td>Non effettuata</td>
<td>non corretta</td>
<td>non corretta</td>
</tr>
<tr>
<td>I corpi sono correttamente identificati nei tre esperimenti del tipo A e in meno di cinque esperimenti di tipo B</td>
<td>Effettuata</td>
<td>corretta</td>
<td>corretta</td>
</tr>
<tr>
<td>I corpi sono correttamente identificati nei tre esperimenti del tipo A e in meno di cinque esperimenti di tipo B</td>
<td>Non effettuata</td>
<td>non corretta</td>
<td>non corretta</td>
</tr>
</tbody>
</table>

III. Invenzione di un metodo di analisi

**Definizione del criterio**

Il procedimento d'invenzione richiede, oltre all'osservazione e all'interpretazione dei fenomeni, la selezione delle variabili pertinenti che permettano la sistemazione di uno schema generale.
Per l'esercizio proposto, l'esperimento immaginato dall'alunno dovrebbe presentare le seguenti caratteristiche:
1. Evidenziare la soluzione da identificare con un reattivo, sia il Br₂, sia il Cl₂;
2. Ricordare l'aggiunta di un solvente (CCL₄);
3. Precisare il colore che permette di trarre una conclusione dall'esperimento.
Queste caratteristiche possono apparire sia sotto forma di esercizio scritto che di equazioni.
   a. Per iscritto
   Esempio: Si mette in evidenza del cloro e una soluzione che contenga, ad esempio, uno ioduro.
   b. Con una equazione.
   Esempio: KI + Cl₂ (CCL₄)KCl + I₂ viola
Non è necessario che l'alunno faccia riferimento ad un nuovo corpo. E' sufficiente che utilizzi uno dei tre corpi semplici già presentati.

**B. Descrizione della scala «Invenzione di un metodo d'analisi»**

<table>
<thead>
<tr>
<th>MB</th>
<th>Il lavoro presenta le caratteristiche 1, 2, 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>B</td>
<td>Il lavoro presenta solo le caratteristiche 1 e 2</td>
</tr>
<tr>
<td>S</td>
<td>Il lavoro presenta solo la caratteristica 1</td>
</tr>
<tr>
<td>D</td>
<td>Il lavoro non presenta la caratteristica 1</td>
</tr>
<tr>
<td>I</td>
<td>Il lavoro non presenta alcuna caratteristica</td>
</tr>
</tbody>
</table>

Voto per il criterio «Invenzione»

Sintesi delle valutazioni

Fino ad ora avete valutato gli elaborati degli alunni riguardo alla relazione e interpretazione dei fenomeni osservati, così come per l'invenzione di un metodo di analisi.

Se l'attribuzione di un solo voto globale può smussare il beneficio di un procedimento analitico preliminare, la pratica scolastica impone, tuttavia, l'utilizzazione, in alcuni momenti, di un voto di sintesi complementare alle valutazioni analitiche.

Riguardo a questo voto di sintesi, l'insegnante può — secondo gli obiettivi perseguiti — accordare la stessa importanza a ciascun criterio, oppure dare maggior peso all'una o all'altra valutazione analitica. L'alunno dovrà essere comunque informato di questa decisione.

1. Relazione

<table>
<thead>
<tr>
<th>MB</th>
<th>B</th>
<th>S</th>
<th>D</th>
<th>I</th>
</tr>
</thead>
</table>
2. Interpretazione

<table>
<thead>
<tr>
<th>MB</th>
<th>B</th>
<th>S</th>
<th>D</th>
<th>I</th>
</tr>
</thead>
</table>
3. Invenzione

<table>
<thead>
<tr>
<th>MB</th>
<th>B</th>
<th>S</th>
<th>D</th>
<th>I</th>
</tr>
</thead>
</table>
4. Voto di sintesi

<table>
<thead>
<tr>
<th>MB</th>
<th>B</th>
<th>S</th>
<th>D</th>
<th>I</th>
</tr>
</thead>
</table>

Ipotesi 1: La scala descrittiva consente una valutazione più oggettiva

- 1. Concordanza dei voti considerati isolatamente
  - Tavole delle medie e degli scarti-tipo
  - Istogrammi di frequenza
Risultati

<table>
<thead>
<tr>
<th>GRUPPO I</th>
<th>GRUPPO II</th>
<th>GRUPPO III</th>
</tr>
</thead>
<tbody>
<tr>
<td>Media</td>
<td>Scarto tipo</td>
<td>Media</td>
</tr>
<tr>
<td>Perin semantica</td>
<td>3,258</td>
<td>0,827</td>
</tr>
<tr>
<td>Perin statistiche</td>
<td>3,558</td>
<td>0,933</td>
</tr>
<tr>
<td>Vocabolario</td>
<td>3,014</td>
<td>0,746</td>
</tr>
<tr>
<td>Strutture grammaticali</td>
<td>3,099</td>
<td>0,844</td>
</tr>
<tr>
<td>Organizzazione</td>
<td>3,944</td>
<td>0,928</td>
</tr>
</tbody>
</table>

Valutazione globale: 3,059, 0,785
Valutazione: 3,380, 0,641

(1) Si calcolà il livello significativo delle differenze tra le medie osservate nei gruppi II e III (F di Student) da una parte, e gli scarti tipici dei gruppi II e III (F di Student).
(2) Vedere avanti i risultati del gruppo II e III (F di Student).
(3) La differenza significativa tra le tre valutazioni globali e stata controllata con il test di Bartlett.

| N.S. | Descrizione | 3,811 | 0,956 | 3,695 | 0,886 | N.S. | S. da P. 05 | 4,852 | 0,316 | TS |
| Valutazione globale | 3,025 | 0,486 | Valutazione | 3,225 | 0,747 | 3,372 | 0,746 | 3,440 | 0,591 |
| Schema | 4,072 | 0,931 | 4,071 | 0,998 | N.S. | S. da P. 12 | 4,375 | 0,433 | TS |
| Interpretazione | 2,757 | 1,055 | 2,752 | 1,065 | N.S. | N.S. | 2,880 | 1,020 | TS |
| Invenzione | 1,685 | 0,914 | 1,973 | 1,030 | S. da P. 02 | S. da P. 05 | 0,976 | S. da P. 05 |

(1) Si calcolà il livello significativo delle differenze tra le medie osservate nei gruppi II e III (F di Student) da una parte, e gli scarti tipici dei gruppi II e III (F di Student).
(2) Vedere avanti i risultati del gruppo II e III (F di Student).
(3) La differenza significativa tra le tre valutazioni globali e stata controllata con il test di Bartlett.

1. Valutazioni globali
2. Valutazioni analitiche
1° sperimentazione

Istogrammi di frequenza:

A. LINGUA MATRERA

<table>
<thead>
<tr>
<th>1° gruppo</th>
<th>1° gruppo</th>
<th>1° gruppo</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
</tbody>
</table>

Pertinenza semantica

<table>
<thead>
<tr>
<th>1° gruppo</th>
<th>1° gruppo</th>
<th>1° gruppo</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
</tbody>
</table>

Pertinenza statistica

<table>
<thead>
<tr>
<th>1° gruppo</th>
<th>1° gruppo</th>
<th>1° gruppo</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
</tbody>
</table>

Pertinenza analitica

<table>
<thead>
<tr>
<th>1° gruppo</th>
<th>1° gruppo</th>
<th>1° gruppo</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
</tbody>
</table>

Il 1° gruppo e il 2° gruppo non hanno differenze significative.

<table>
<thead>
<tr>
<th>1° gruppo</th>
<th>1° gruppo</th>
<th>1° gruppo</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>1° gruppo</th>
<th>1° gruppo</th>
<th>1° gruppo</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>1° gruppo</th>
<th>1° gruppo</th>
<th>1° gruppo</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>1° gruppo</th>
<th>1° gruppo</th>
<th>1° gruppo</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>1° gruppo</th>
<th>1° gruppo</th>
<th>1° gruppo</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
</tbody>
</table>
1. Valutazioni globali

2. Valutazioni analitiche
2. Concordanza tra i profili

a. Concordanza con il profilo di base dei voti assegnati con i criteri di correzione

<table>
<thead>
<tr>
<th></th>
<th>1ª sperimentazione</th>
<th>2ª sperimentazione</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Gruppo II</strong></td>
<td><strong>Gruppo III</strong></td>
<td><strong>Gruppo III</strong></td>
</tr>
<tr>
<td>Lingua materna (5 voti compresi «Vocabolario»)</td>
<td>1,43 %</td>
<td>1,23 %</td>
</tr>
<tr>
<td>Lingua materna (4 voti senza «Vocabolario»)</td>
<td>1,43 %</td>
<td>10,11 %</td>
</tr>
<tr>
<td>Scienze</td>
<td>1,80 %</td>
<td>1,77 %</td>
</tr>
</tbody>
</table>

Percentuale d’accordo con il profilo di base

b. Concordanza con i profili prossimi al profilo di base

I profili considerati presentano, per uno o due criteri soltanto, valori prossimi, (una unità in più o in meno) al voto teorico.

<table>
<thead>
<tr>
<th></th>
<th>1ª sperimentazione</th>
<th>2ª sperimentazione</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Gruppo II</strong></td>
<td><strong>Gruppo III</strong></td>
<td><strong>Gruppo III</strong></td>
</tr>
<tr>
<td>Lingua materna (4 voti)</td>
<td>34,29 %</td>
<td>60,50 %</td>
</tr>
<tr>
<td>Scienze</td>
<td>24,33 %</td>
<td>34,51 %</td>
</tr>
</tbody>
</table>

3. Una valutazione oggettiva non è influenzata da fattori esterni

a. Le scale descrittive si oppongono alla tendenza centrale?
2. Analisi dei voti globali.

Comparazione globale

<table>
<thead>
<tr>
<th>Fattori di variazione</th>
<th>Summa dei quadrati</th>
<th>Grado di libertà</th>
<th>Quadrati medi</th>
<th>Valore di F</th>
<th>Livello di significatività</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intergruppo (1)</td>
<td>12,019</td>
<td>2</td>
<td>6,00</td>
<td>10,38</td>
<td>S da P 01</td>
</tr>
<tr>
<td>Intergruppo (2)</td>
<td>94,506</td>
<td>300</td>
<td>0,31</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Totale</td>
<td>106,525</td>
<td>308</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Lingua materna**

(1) La varianza globale intergruppo è determinata dai differenti metodi di valutazione utilizzati nei tre gruppi. Ha il compito di evidenziare le eventuali differenze dovute agli effetti sperimentali. Qui, l’introduzione di dimensioni analitiche (comparazione del 1 gruppo con i gruppi II e III) e l’utilizzazione di scale descrittive (comparazione dei gruppi II e III).
(2) La varianza intergruppo riguarda gli effetti dovuti al caso nel campionamento.
### b. Distribuzione dei voti estremi

#### Lingua Materna

<table>
<thead>
<tr>
<th>VG</th>
<th>PS</th>
<th>Pr</th>
<th>SG</th>
<th>Org.</th>
<th>Gruppo II</th>
<th>Gruppo III</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>3</td>
<td>5</td>
<td>3</td>
<td>3</td>
<td>2.86</td>
<td>14.60</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>-</td>
<td>2.86</td>
<td>13.58</td>
</tr>
<tr>
<td>2</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>2.47</td>
<td>0.04</td>
</tr>
<tr>
<td>5.72</td>
<td>30.86</td>
<td>10</td>
<td>23.46</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

#### Scienze

<table>
<thead>
<tr>
<th>VG</th>
<th>Descr.</th>
<th>Int.</th>
<th>Scop.</th>
<th>Inv.</th>
<th>Gruppo II</th>
<th>Gruppo III</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>3</td>
<td>3.60</td>
<td>13.27</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>-</td>
<td>7.21</td>
<td>4.42</td>
</tr>
<tr>
<td>2</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.90</td>
<td>0.88</td>
</tr>
<tr>
<td>5</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>11.71</td>
<td>12.57</td>
<td>10.81</td>
<td>6.19</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

#### Lingua Materna

<table>
<thead>
<tr>
<th>VG</th>
<th>Descr.</th>
<th>Int.</th>
<th>Scop.</th>
<th>Inv.</th>
<th>Gruppo II</th>
<th>Gruppo III</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>5</td>
<td>3</td>
<td>4</td>
<td>3</td>
<td>0.50</td>
<td>4.42</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td>4.50</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.90</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>3.40</td>
<td>4.42</td>
</tr>
</tbody>
</table>

Percentuali sull'apparizione dei profili che utilizzano voti estremi in scienze.
b. Le scale descrittive evitano l’effetto di contaminazione?

1. Analisi delle correlazioni tra valutazioni analitiche

### Lingua Materna

<table>
<thead>
<tr>
<th></th>
<th>Desc.</th>
<th>Sch.</th>
<th>Int.</th>
<th>Inv.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Descrizione</td>
<td>1.24</td>
<td>.14</td>
<td>.06</td>
<td></td>
</tr>
<tr>
<td>Schema</td>
<td>1.39</td>
<td>S da 0.005</td>
<td>.09</td>
<td>N.S.</td>
</tr>
<tr>
<td>Interpretazione</td>
<td>1.09</td>
<td>S da P. 01</td>
<td>.21</td>
<td>S da P. 02</td>
</tr>
<tr>
<td>Invenzione</td>
<td>1.07</td>
<td>N.S.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Gruppo III**

(n = 113)

2. Analisi fattoriale dei voti analitici

### Lingua materna

<table>
<thead>
<tr>
<th></th>
<th>Gruppo II</th>
<th>Gruppo III</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>F1</td>
<td>F2</td>
</tr>
<tr>
<td>% cumulativa dei valori propri</td>
<td>.35</td>
<td>.39</td>
</tr>
<tr>
<td>Saturazione di ciascuna variabile</td>
<td>.42</td>
<td>.71</td>
</tr>
<tr>
<td>Pertinenza semantica</td>
<td>.04</td>
<td>.83</td>
</tr>
<tr>
<td>Pertinenza sintattica</td>
<td>.02</td>
<td>.91</td>
</tr>
<tr>
<td>Pertinenza sinonimitica</td>
<td>.05</td>
<td>.80</td>
</tr>
<tr>
<td>Vocabolario</td>
<td>.95</td>
<td>.04</td>
</tr>
<tr>
<td>Struttura grammaticali</td>
<td>.27</td>
<td>.34</td>
</tr>
<tr>
<td>Organizzazione delle idee</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Scienze

<table>
<thead>
<tr>
<th></th>
<th>Gruppo II</th>
<th>Gruppo III</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>F1</td>
<td>F2</td>
</tr>
<tr>
<td>% cumulativa dei valori propri</td>
<td>.39</td>
<td>.69</td>
</tr>
<tr>
<td>Descrizione</td>
<td>.84</td>
<td>.01</td>
</tr>
<tr>
<td>Schema</td>
<td>.19</td>
<td>.73</td>
</tr>
<tr>
<td>Interpretazione</td>
<td>.01</td>
<td>.84</td>
</tr>
<tr>
<td>Invenzione</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

4. Una valutazione oggettiva non è influenzata da fattori interni

a. I giudici hanno utilizzato soltanto i criteri proposti?

La parte occupata dai voti analitici nella valutazione globale si individua con la percentuale di varianza e con il coefficiente di correlazione multiplica (I).

### Lingua materna

<table>
<thead>
<tr>
<th></th>
<th>Gruppo II</th>
<th>Gruppo III</th>
</tr>
</thead>
<tbody>
<tr>
<td>% di varianza esplicita</td>
<td>.46%</td>
<td>.681</td>
</tr>
<tr>
<td>Coefficiente di correlazione multiplica</td>
<td>.45%</td>
<td>.651</td>
</tr>
</tbody>
</table>

(1) La varianza inespresa dalle dimensioni analitiche può dipendere da due fattori alatori (fattori esterni) o da altri criteri introdotti volontariamente dall’intercettore (fattori interni). La tecnica non permette di distinguerne l’influenza rispettiva di questi due tipi di variabilità.
Ipotesi 2: La scala descrittiva consente una valutazione più valida

1. Il voto globale è rapportato significativamente alle valutazioni analitiche?
   a. Rapporto tra voto globale e valutazione analitica. Analisi delle correlazioni semplici.

<table>
<thead>
<tr>
<th>Lingua Materna</th>
<th>Gruppo II</th>
<th>Gruppo III</th>
</tr>
</thead>
<tbody>
<tr>
<td>FG</td>
<td>F1</td>
<td>F2</td>
</tr>
<tr>
<td>% cumulativa dei valori propri</td>
<td>.38</td>
<td>.59</td>
</tr>
<tr>
<td>Valutazione globale</td>
<td>.85</td>
<td>.03</td>
</tr>
<tr>
<td>Pertinenza semantica</td>
<td>.82</td>
<td>.02</td>
</tr>
<tr>
<td>Pertinenza sintattica</td>
<td>.03</td>
<td>.02</td>
</tr>
<tr>
<td>Vocabolario</td>
<td>.74</td>
<td>.08</td>
</tr>
<tr>
<td>Struttura Grammaticale</td>
<td>-.04</td>
<td>.06</td>
</tr>
<tr>
<td>Organizzazione delle idee</td>
<td>.31</td>
<td>.48</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Scienze</th>
<th>Gruppo II</th>
<th>Gruppo III</th>
</tr>
</thead>
<tbody>
<tr>
<td>FG</td>
<td>F1</td>
<td>F2</td>
</tr>
<tr>
<td>% cumulativa dei valori propri</td>
<td>.46</td>
<td>.70</td>
</tr>
<tr>
<td>Valutazione globale</td>
<td>.70</td>
<td>.57</td>
</tr>
<tr>
<td>Descrizione</td>
<td>.80</td>
<td>.05</td>
</tr>
<tr>
<td>Schema</td>
<td>.82</td>
<td>.01</td>
</tr>
<tr>
<td>Interpretazione</td>
<td>.19</td>
<td>.78</td>
</tr>
<tr>
<td>Invenzione</td>
<td>-.07</td>
<td>.83</td>
</tr>
</tbody>
</table>

b. Analisi fattoriale

Con l'introduzione del voto globale nell'analisi fattoriale, si cerca di determinare quelle variabili che spiegano al meglio la valutazione globale stessa.

174
Lingua materna — Gruppo II

Lingua materna — Gruppo III
Ipotesi 3: La scala descrittiva consente una valutazione meglio utilizzabile dal professore e dall’alunno.

La Tavola I presenta il modo in cui sono utilizzate le diverse giustificazioni. Per ciascun gruppo si leggerà il numero dei casi in cui non è data alcuna giustificazione (1ª colonna), quante volte la giustificazione è positiva (2ª colonna), quante volte è negativa (terza colonna).

Le giustificazioni sono ripartite in tre categorie:
1. Giustificazioni corrispondenti ai criteri che costituiscono l'oggetto delle scale (giustificazioni principali);
2. Giustificazioni presenti nei commenti dei professori (giustificazioni secondarie);
3. Giustificazioni non classificabili per diletto di precisione.

La Tavola II presenta le stesse informazioni tradotte in percentuali.

### TAVOLA I. Frequenze grezze

<table>
<thead>
<tr>
<th>Lingua materna</th>
<th>GRUPPO I (n=112)</th>
<th>GRUPPO II (n=111)</th>
<th>GRUPPO III (n=113)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>GIUSTIFICAZIONE</td>
<td>GIUSTIFICAZIONE</td>
<td>GIUSTIFICAZIONE</td>
</tr>
<tr>
<td></td>
<td>+ a b</td>
<td>+ a b</td>
<td>+ a b</td>
</tr>
<tr>
<td></td>
<td>tot. a+b</td>
<td>tot. a+b</td>
<td>tot. a+b</td>
</tr>
</tbody>
</table>

Le giustificazioni principali:
- Perfomance Semantica
- Sintassi
- Vocabolario
- Strutture grammaticali
- Organizzazione delle idee

Le giustificazioni secondarie:
- Ortografia
- Puntellatura
- Sintesi
- Immagazzinamento
- Condizionamento

Totale generale delle giustificazioni classificale

### TAVOLA II. Frequenze grezze

<table>
<thead>
<tr>
<th>Gruppo</th>
<th>+ a</th>
<th>b</th>
<th>a+b</th>
<th>- a+b</th>
<th>Tot. a+b</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**TAVOLA I. Frequenze grezze**

**TAVOLA II. Frequenze grezze**

<table>
<thead>
<tr>
<th><strong>Gruppo</strong></th>
<th><strong>Descrizione</strong></th>
<th><strong>Scheda</strong></th>
<th><strong>Interpretazione</strong></th>
<th><strong>Invenzione</strong></th>
<th><strong>Totale</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Gruppo I</strong></td>
<td>112</td>
<td>111</td>
<td>113</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Gruppo II</strong></td>
<td>111</td>
<td>111</td>
<td>113</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Gruppo III</strong></td>
<td>112</td>
<td>111</td>
<td>113</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Gruppo</strong></th>
<th><strong>Lezione</strong></th>
<th><strong>Osservazione</strong></th>
<th><strong>Conclusione</strong></th>
<th><strong>Comprensione</strong></th>
<th><strong>Distanzazione</strong></th>
<th><strong>Descrizione-interpretazione</strong></th>
<th><strong>Totale</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Gruppo I</strong></td>
<td>112</td>
<td>111</td>
<td>113</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Gruppo II</strong></td>
<td>111</td>
<td>111</td>
<td>113</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Gruppo III</strong></td>
<td>112</td>
<td>111</td>
<td>113</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Gruppo</strong></th>
<th><strong>Giustificazioni</strong></th>
<th><strong>Giustificazioni</strong></th>
<th><strong>Giustificazioni</strong></th>
<th><strong>Giustificazioni</strong></th>
<th><strong>Giustificazioni</strong></th>
<th><strong>Giustificazioni</strong></th>
<th><strong>Giustificazioni</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Gruppo I</strong></td>
<td>112</td>
<td>111</td>
<td>113</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Gruppo II</strong></td>
<td>111</td>
<td>111</td>
<td>113</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Gruppo III</strong></td>
<td>112</td>
<td>111</td>
<td>113</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### TAVOLA II. Percentuali

#### Lingua materna

<table>
<thead>
<tr>
<th>Giustificazioni</th>
<th>Positivo a</th>
<th>Positivo b</th>
<th>Negativo a</th>
<th>Negativo b</th>
<th>Nessuna</th>
</tr>
</thead>
<tbody>
<tr>
<td>GRUPPO I</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pertinenza</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Semantica</td>
<td>46</td>
<td>27</td>
<td>28</td>
<td>83</td>
<td>7</td>
</tr>
<tr>
<td>Sintattica</td>
<td>98</td>
<td>0</td>
<td>2</td>
<td>97</td>
<td>0</td>
</tr>
<tr>
<td>Vocabolario</td>
<td>74</td>
<td>9</td>
<td>17</td>
<td>81</td>
<td>0</td>
</tr>
<tr>
<td>Strutture</td>
<td>51</td>
<td>22</td>
<td>27</td>
<td>79</td>
<td>4</td>
</tr>
<tr>
<td>Grammaticali</td>
<td>69</td>
<td>16</td>
<td>15</td>
<td>90</td>
<td>9</td>
</tr>
<tr>
<td>Organizzazione</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>delle idee</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GRUPPO II</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Giustificazioni</td>
<td>69</td>
<td>16</td>
<td>15</td>
<td>90</td>
<td>9</td>
</tr>
<tr>
<td>Positivo a</td>
<td>26</td>
<td>14</td>
<td>17</td>
<td>79</td>
<td>4</td>
</tr>
<tr>
<td>Negativo a</td>
<td>11</td>
<td>10</td>
<td>7</td>
<td>20</td>
<td>0</td>
</tr>
<tr>
<td>Nessuna</td>
<td>32</td>
<td>22</td>
<td>1</td>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td>Nuovi dati</td>
<td>10</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Giustificazioni</td>
<td>90</td>
<td>5</td>
<td>5</td>
<td>97</td>
<td>1</td>
</tr>
<tr>
<td>Positivo a</td>
<td>46</td>
<td>27</td>
<td>28</td>
<td>83</td>
<td>7</td>
</tr>
<tr>
<td>Negativo a</td>
<td>98</td>
<td>0</td>
<td>2</td>
<td>97</td>
<td>0</td>
</tr>
<tr>
<td>Nessuna</td>
<td>74</td>
<td>18</td>
<td>9</td>
<td>76</td>
<td>2</td>
</tr>
<tr>
<td>Nuovi dati</td>
<td>31</td>
<td>0</td>
<td>69</td>
<td>76</td>
<td>1</td>
</tr>
<tr>
<td>Giustificazioni</td>
<td>45</td>
<td>3</td>
<td>52</td>
<td>90</td>
<td>0</td>
</tr>
<tr>
<td>Positivo a</td>
<td>38</td>
<td>2</td>
<td>89</td>
<td>79</td>
<td>1</td>
</tr>
<tr>
<td>Negativo a</td>
<td>11</td>
<td>10</td>
<td>7</td>
<td>19</td>
<td>2</td>
</tr>
<tr>
<td>Nessuna</td>
<td>11</td>
<td>10</td>
<td>0</td>
<td>7</td>
<td>2</td>
</tr>
<tr>
<td>Nuovi dati</td>
<td>76</td>
<td>11</td>
<td>13</td>
<td>83</td>
<td>0</td>
</tr>
<tr>
<td>Giustificazioni</td>
<td>90</td>
<td>5</td>
<td>5</td>
<td>97</td>
<td>1</td>
</tr>
<tr>
<td>Positivo a</td>
<td>69</td>
<td>16</td>
<td>15</td>
<td>90</td>
<td>9</td>
</tr>
<tr>
<td>Negativo a</td>
<td>23</td>
<td>22</td>
<td>19</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Nessuna</td>
<td>11</td>
<td>10</td>
<td>7</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Nuovi dati</td>
<td>69</td>
<td>20</td>
<td>9</td>
<td>2</td>
<td></td>
</tr>
</tbody>
</table>

#### Scienze

<table>
<thead>
<tr>
<th>Descrizione</th>
<th>Gruppo I</th>
<th>Gruppo II</th>
<th>Gruppo III</th>
</tr>
</thead>
<tbody>
<tr>
<td>Giustificazioni</td>
<td>61</td>
<td>33</td>
<td>6</td>
</tr>
<tr>
<td>Descrizione</td>
<td>66</td>
<td>33</td>
<td>6</td>
</tr>
<tr>
<td>Schema</td>
<td>56</td>
<td>41</td>
<td>2</td>
</tr>
<tr>
<td>Interpretazione</td>
<td>47</td>
<td>30</td>
<td>15</td>
</tr>
<tr>
<td>Invenzione</td>
<td>38</td>
<td>20</td>
<td>9</td>
</tr>
<tr>
<td>Letteratura</td>
<td>93</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>Osservazione</td>
<td>86</td>
<td>12</td>
<td>1</td>
</tr>
<tr>
<td>Conclusione</td>
<td>85</td>
<td>14</td>
<td>4</td>
</tr>
<tr>
<td>Compressione</td>
<td>76</td>
<td>14</td>
<td>1</td>
</tr>
<tr>
<td>Ogg della esperienza</td>
<td>90</td>
<td>1</td>
<td>96</td>
</tr>
<tr>
<td>Descrizione-</td>
<td>90</td>
<td>1</td>
<td>96</td>
</tr>
<tr>
<td>Interpretazione</td>
<td>90</td>
<td>1</td>
<td>96</td>
</tr>
</tbody>
</table>

### 1. La quantità di giustificazioni differisce da un gruppo all'altro?

#### a. Numero globale di giustificazioni

<table>
<thead>
<tr>
<th>Gruppo I</th>
<th>Gruppo II</th>
<th>Gruppo III</th>
</tr>
</thead>
<tbody>
<tr>
<td>339</td>
<td>94</td>
<td>175</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Numero degli insegnanti</th>
<th>101</th>
<th>70</th>
<th>81</th>
</tr>
</thead>
<tbody>
<tr>
<td>Numero medio per insegnante</td>
<td>3,3</td>
<td>1,3</td>
<td>2,1</td>
</tr>
</tbody>
</table>

### b. Suddivisione delle giustificazioni usate da ciascun professore

#### Lingua materna

La tabella seguente riprende il numero di giustificazioni usate da ciascun professore.

<table>
<thead>
<tr>
<th>Giustificazioni</th>
<th>I</th>
<th>%</th>
<th>II</th>
<th>%</th>
<th>III</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 giustificazioni</td>
<td>1</td>
<td>1</td>
<td>6</td>
<td>13</td>
<td>13</td>
<td>98</td>
</tr>
<tr>
<td>1 giustificazioni</td>
<td>-</td>
<td>-</td>
<td>3</td>
<td>5</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td>2 giustificazioni</td>
<td>-</td>
<td>-</td>
<td>10</td>
<td>19</td>
<td>10</td>
<td>19</td>
</tr>
<tr>
<td>3 giustificazioni</td>
<td>-</td>
<td>-</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>4 giustificazioni</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>5 giustificazioni</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>6 giustificazioni</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>7 giustificazioni</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

### 2. I professori utilizzano le giustificazioni in rapporto agli obiettivi del lavoro proposto?

Quali criteri utilizzano i professori nel valutare liberamente? Riprendiamo dalla tavola di pag. 182 e 183 i dati relativi al gruppo I.

#### Lingua Materna

<table>
<thead>
<tr>
<th>Giustificazioni principali</th>
<th>Frequenza grezze</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pertinenza semantica</td>
<td>55</td>
<td>15,2</td>
</tr>
<tr>
<td>Pertinenza sintattica</td>
<td>2</td>
<td>0,6</td>
</tr>
<tr>
<td>Vocabolario</td>
<td>26</td>
<td>7,7</td>
</tr>
<tr>
<td>Strutture grammaticali</td>
<td>49</td>
<td>14,5</td>
</tr>
<tr>
<td>Organizzazione delle idee</td>
<td>31</td>
<td>9,1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Giustificazioni secondarie</th>
<th>Frequenza grezze</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ortografia</td>
<td>70</td>
<td>20,6</td>
</tr>
<tr>
<td>Puntigliatura</td>
<td>56</td>
<td>16,5</td>
</tr>
<tr>
<td>Stile</td>
<td>16</td>
<td>4,7</td>
</tr>
<tr>
<td>Immaginazione</td>
<td>24</td>
<td>7,1</td>
</tr>
<tr>
<td>Composizione</td>
<td>10</td>
<td>2,9</td>
</tr>
</tbody>
</table>

| Totale giustificazioni classificate | 339 | 100% |

184
### Scienze

<table>
<thead>
<tr>
<th>Giustificazioni principali</th>
<th>Frequenze</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Descrizione</td>
<td>48</td>
<td>14,8</td>
</tr>
<tr>
<td>Schema</td>
<td>52</td>
<td>16</td>
</tr>
<tr>
<td>Interpretazione</td>
<td>65</td>
<td>20</td>
</tr>
<tr>
<td>Emissione</td>
<td>76</td>
<td>23,4</td>
</tr>
<tr>
<td>Letteratura</td>
<td>9</td>
<td>2,8</td>
</tr>
<tr>
<td>Osservazione</td>
<td>16</td>
<td>4,9</td>
</tr>
<tr>
<td>Conclusione</td>
<td>18</td>
<td>5,5</td>
</tr>
<tr>
<td>Comprensione dell'oggetto</td>
<td>29</td>
<td>8,9</td>
</tr>
<tr>
<td>dell'esperienza</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Distinzione</td>
<td>12</td>
<td>3,7</td>
</tr>
<tr>
<td>Descrizione e interpretazione</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Totale giustificazioni classificate</td>
<td>325</td>
<td>100,0%</td>
</tr>
</tbody>
</table>

### Lingua Materna

#### a. Rapporto tra giustificazioni principali e secondarie nei differenti gruppi.

Dall'insieme delle giustificazioni positive o negative si calcola la proporzione tra giustificazioni principali (in relazione agli obiettivi del lavoro) e secondarie.

#### I. Principali | II. Secondarie

| Principali | 163 (182) | 49 (50) | 113 (94) | 325 |
| Secondary  | 176 (157) | 45 (44) | 62 (81)  | 283 |

### Scienze

| Principali | 163 (182) | 49 (50) | 113 (94) | 325 |
| Secondary  | 176 (157) | 45 (44) | 62 (81)  | 283 |

**X² = 12,17 S da P.01**

---

1. Il test $X^2$ permette di controllare i differenti significati tra le frequenze.
2. I numeri tra parentesi rappresentano le frequenze teoriche.

---

186
**Scienze**

<table>
<thead>
<tr>
<th></th>
<th>I</th>
<th>II</th>
<th>III</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Freq. grezze</td>
<td>%</td>
<td>Freq. grezze</td>
</tr>
<tr>
<td>Nessuna giustificazione</td>
<td>9</td>
<td>7,38</td>
<td>26</td>
</tr>
<tr>
<td>Giustificazione secondaria soltanto</td>
<td>14</td>
<td>11,48</td>
<td>19</td>
</tr>
<tr>
<td>1 principale soltanto</td>
<td>28</td>
<td>22,9</td>
<td>10</td>
</tr>
<tr>
<td>2 principali soltanto</td>
<td>27</td>
<td>22,1</td>
<td>10</td>
</tr>
<tr>
<td>3 principali soltanto</td>
<td>19</td>
<td>15,5</td>
<td>4</td>
</tr>
<tr>
<td>4 principali soltanto</td>
<td>25</td>
<td>20,4</td>
<td>1</td>
</tr>
</tbody>
</table>

**b. Caratteristiche delle giustificazioni per professore**

**Lingua Materna**

<table>
<thead>
<tr>
<th></th>
<th>I</th>
<th>%</th>
<th>II</th>
<th>%</th>
<th>III</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Freq. grezze</td>
<td></td>
<td>Freq. grezze</td>
<td></td>
<td>Freq. grezze</td>
<td></td>
</tr>
<tr>
<td>Nessuna giustificazione</td>
<td>1</td>
<td>0,9</td>
<td>13</td>
<td>19,6</td>
<td>24</td>
<td>29,2</td>
</tr>
<tr>
<td>Soltanto 1 giustificazione secondaria</td>
<td>20</td>
<td>19,8</td>
<td>17</td>
<td>25,7</td>
<td>21</td>
<td>25,6</td>
</tr>
<tr>
<td>Soltanto 1 giustificazione principale, quale sia la quantità delle giustificazioni secondarie</td>
<td>51</td>
<td>50,5</td>
<td>12</td>
<td>18,1</td>
<td>27</td>
<td>32,9</td>
</tr>
<tr>
<td>2 giustificazioni principali</td>
<td>22</td>
<td>21,78</td>
<td>10</td>
<td>15,15</td>
<td>14</td>
<td>17,1</td>
</tr>
<tr>
<td>3 giustificazioni principali</td>
<td>7</td>
<td>6,93</td>
<td>2</td>
<td>3,03</td>
<td>3</td>
<td>3,7</td>
</tr>
<tr>
<td>4 giustificazioni principali</td>
<td>0</td>
<td>—</td>
<td>1</td>
<td>1,52</td>
<td>4</td>
<td>4,9</td>
</tr>
</tbody>
</table>

3. L' alunno identifica facilmente i commenti a lui destinati?

**a. Dalle giustificazioni l' alunno trae informazioni che lo aiutano concretamente a migliorare l'apprendimento?**

<table>
<thead>
<tr>
<th></th>
<th>Gruppo I</th>
<th>%</th>
<th>Gruppo II</th>
<th>%</th>
<th>Gruppo III</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>%</td>
<td>n</td>
<td>%</td>
<td>n</td>
<td>%</td>
</tr>
<tr>
<td>Descrizione</td>
<td>48</td>
<td>20</td>
<td>16</td>
<td>27</td>
<td>27</td>
<td>22</td>
</tr>
<tr>
<td>Schema</td>
<td>52</td>
<td>22</td>
<td>12</td>
<td>20</td>
<td>24</td>
<td>20</td>
</tr>
<tr>
<td>Interpretazione</td>
<td>65</td>
<td>27</td>
<td>19</td>
<td>32</td>
<td>29</td>
<td>22</td>
</tr>
<tr>
<td>Inversione</td>
<td>76</td>
<td>32</td>
<td>12</td>
<td>20</td>
<td>15</td>
<td>12</td>
</tr>
</tbody>
</table>

188
b. Ci sono molte giustificazioni sui dettagli e sugli aspetti marginali?

Questo tipo di giustificazioni compare soprattutto nella valutazione della pertinenza semantica. Per esempio, alcuni professori menzionano soltanto il dettaglio trascurato dall'alunno.

<table>
<thead>
<tr>
<th>Numero totale delle giustificazioni classificate</th>
<th>Gruppo I</th>
<th>Gruppo II</th>
<th>Gruppo III</th>
</tr>
</thead>
<tbody>
<tr>
<td>339</td>
<td>94</td>
<td>175</td>
<td></td>
</tr>
<tr>
<td>Numero di giustificazioni legate ai dettagli</td>
<td>12</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>Percentuale</td>
<td>7</td>
<td>1</td>
<td>2</td>
</tr>
</tbody>
</table>

c. Qual è la ripartizione delle giustificazioni «non classificabili»?

<table>
<thead>
<tr>
<th>Numero totale delle giustificazioni</th>
<th>Gruppo I</th>
<th>Gruppo II</th>
<th>Gruppo III</th>
</tr>
</thead>
<tbody>
<tr>
<td>423</td>
<td>139</td>
<td>237</td>
<td></td>
</tr>
<tr>
<td>Numero di giustificazioni inclassificabili</td>
<td>84</td>
<td>45</td>
<td>64</td>
</tr>
<tr>
<td>Percentuale</td>
<td>19.8</td>
<td>32.3</td>
<td>27</td>
</tr>
</tbody>
</table>

Tra le giustificazioni che si possono classificare in una categoria definita, poche sono sufficientemente specifiche perché l'alunno possa utilizzarle per la continua- zione del suo apprendimento. Alcune si riferiscono unicamente a dettagli senza mol- to significato. La percentuale delle giustificazioni non classificabili è relativamente elevata.

4. I professori rilevano nei loro commenti gli aspetti positivi o negativi del lavoro?

a. Analisi di tutte le giustificazioni

**Lingua Materna**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Giustific. principali</td>
<td>74</td>
<td>69</td>
<td>143</td>
<td>14</td>
<td>38</td>
<td>49</td>
<td>53</td>
<td>60</td>
</tr>
<tr>
<td>Percentuale</td>
<td>45.3</td>
<td>54.6</td>
<td>100</td>
<td>28.5</td>
<td>71.4</td>
<td>46.9</td>
<td>53.1</td>
<td>100</td>
</tr>
<tr>
<td>Giustific. secondarie</td>
<td>26</td>
<td>150</td>
<td>176</td>
<td>8</td>
<td>37</td>
<td>45</td>
<td>23</td>
<td>39</td>
</tr>
<tr>
<td>Percentuale</td>
<td>14.7</td>
<td>85.2</td>
<td>100</td>
<td>17.7</td>
<td>82.3</td>
<td>23.1</td>
<td>77.3</td>
<td>100</td>
</tr>
<tr>
<td>Tot. giustif. classificate</td>
<td>100</td>
<td>239</td>
<td>339</td>
<td>22</td>
<td>72</td>
<td>94</td>
<td>76</td>
<td>99</td>
</tr>
<tr>
<td>Percentuale</td>
<td>28.4</td>
<td>70.5</td>
<td>100</td>
<td>23.4</td>
<td>76.6</td>
<td>43.4</td>
<td>56.6</td>
<td>100</td>
</tr>
</tbody>
</table>

**Scienze**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Giustific. principali</td>
<td>98</td>
<td>42</td>
<td>143</td>
<td>32</td>
<td>54</td>
<td>86</td>
<td>48</td>
<td>50</td>
</tr>
<tr>
<td>Percentuale</td>
<td>90.6</td>
<td>59.3</td>
<td>100</td>
<td>54.2</td>
<td>45.7</td>
<td>50.5</td>
<td>49.5</td>
<td>100</td>
</tr>
<tr>
<td>Giustific. secondarie</td>
<td>26</td>
<td>58</td>
<td>84</td>
<td>25</td>
<td>25</td>
<td>50</td>
<td>29</td>
<td>25</td>
</tr>
<tr>
<td>Percentuale</td>
<td>31.0</td>
<td>69.0</td>
<td>100</td>
<td>50.0</td>
<td>50.0</td>
<td>53.7</td>
<td>46.3</td>
<td>100</td>
</tr>
<tr>
<td>Tot. giustif. classificate</td>
<td>124</td>
<td>201</td>
<td>325</td>
<td>57</td>
<td>52</td>
<td>109</td>
<td>77</td>
<td>72</td>
</tr>
<tr>
<td>Percentuale</td>
<td>38.2</td>
<td>61.8</td>
<td>100</td>
<td>52.3</td>
<td>47.7</td>
<td>51.7</td>
<td>46.3</td>
<td>100</td>
</tr>
</tbody>
</table>

b. Analisi delle giustificazioni principali

**Lingua Materna**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Totale delle giust. principali</td>
<td>74</td>
<td>45</td>
<td>119</td>
<td>14</td>
<td>29</td>
<td>35</td>
<td>51</td>
<td>49</td>
</tr>
<tr>
<td>P. semantica</td>
<td>27</td>
<td>36.4</td>
<td>28</td>
<td>31.4</td>
<td>5</td>
<td>35.7</td>
<td>7</td>
<td>20</td>
</tr>
<tr>
<td>P. sintattica</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Vocabolario</td>
<td>9</td>
<td>12.1</td>
<td>17</td>
<td>19.1</td>
<td>0</td>
<td>13</td>
<td>37</td>
<td></td>
</tr>
<tr>
<td>Strette gram-</td>
<td>22</td>
<td>29.7</td>
<td>27</td>
<td>30.3</td>
<td>12</td>
<td>34.2</td>
<td>17</td>
<td></td>
</tr>
<tr>
<td>maicali</td>
<td>17</td>
<td>22.4</td>
<td>18</td>
<td>20</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Organizzazione</td>
<td>16</td>
<td>21.6</td>
<td>15</td>
<td>16.8</td>
<td>6</td>
<td>42.8</td>
<td>2</td>
<td>2.8</td>
</tr>
<tr>
<td>delle idee</td>
<td>2</td>
<td>3</td>
<td>19</td>
<td>31.6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Scienze**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Tot. delle giust. principali</td>
<td>98</td>
<td>42</td>
<td>143</td>
<td>32</td>
<td>54</td>
<td>86</td>
<td>48</td>
<td>50</td>
</tr>
<tr>
<td>Descrizione</td>
<td>41</td>
<td>42</td>
<td>7</td>
<td>5</td>
<td>15</td>
<td>47</td>
<td>22</td>
<td></td>
</tr>
<tr>
<td>Scherna</td>
<td>50</td>
<td>51</td>
<td>2</td>
<td>1</td>
<td>11</td>
<td>34</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Interpretazione</td>
<td>4</td>
<td>4</td>
<td>6</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inventario</td>
<td>3</td>
<td>3</td>
<td>13</td>
<td>51</td>
<td>0</td>
<td>12</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

191
Indice

Premessa 5
Introduzione 9

Parte prima: Indagine sulla valutazione

Capitolo primo: Per una valutazione formativa
Premessa — La valutazione deve essere valida — La valutazione deve essere oggettiva — La valutazione deve essere formativa 17

Capitolo secondo: Risultati di un’indagine
Ambito dell’indagine — Obiettivi dell’indagine — Popolazione e discipline interessate — Questionario — Spoglio dei dati — Risultati — Conclusioni 25

Parte seconda: Le scale descrittive

Capitolo primo: Le scale di valutazione
Due tipi di misurazione — Tipi di scale — Quando utilizzare le scale di valutazione? — Ostacoli all’oggettività e validità delle valutazioni — Necessità di precisazioni operative 51

Capitolo secondo: Tappe e condizioni della ricerca
Le tappe — Condizioni 63

Capitolo terzo: Metodo di costruzione di scale descrittive di valutazione
Costruzione iniziale dello strumento 66

Capitolo quarto: Risultati sperimentali
Prima sperimentazione — Seconda sperimentazione — Ipotesi da verificare — Spoglio dei dati — Risultati — Ipotesi 1: La scala descrittiva permette una valutazione più oggettiva — Analisi delle correlazioni — Analisi fattoriale — Ipotesi 2: La scala descrittiva permette una valutazione più valida — Analisi delle correlazioni — Analisi fattoriale — Ipotesi 3: La scala descrittiva permette una valutazione che insegnante e alunno possono utilizzare in modo migliore — Conclusione 86

Allegati

Introduzione 105
A. Le scale di valutazione
B. Analisi statistica dei risultati 161

Stampato in Firenze, nello Stabilimento Grafico Aurora, 1985

Di Gilbert De Landsheere la nostra editrice ha pubblicato l'importante e suggestivo volume *I comportamenti non verbali dell'insegnante*, scritto in collaborazione con A. Delchambre; *Come si insegna. Analisi delle interazioni verbali in classe; La ricerca sperimentale nell'educazione*.

Jacqueline Beckers e Robert De Bal sono laureati in scienze dell'educazione e ricercatori nel Laboratorio di pedagogia sperimentale dell'Università di Liegi, diretto dal professor G. De Landsheere.