

La competenza valutativa degli insegnanti nella valutazione a supporto dell'apprendimento: il sistema **RTTI***

di *Marinka Drost, Petra Verra, Jaap Scheerens*

Le frasi in corsivo sono state volte in tondo rosso: secondo norme Carocci evidenziare in cvo solo i concetti chiave

7.1

Valutazione e *formative assessment*[†]: processi chiave

Tutte le forme di valutazione prevedono la raccolta sistematica di informazioni e l'espressione di giudizio sulla base di queste informazioni. Ci si aspetta, inoltre, che queste "informazioni di valore" siano usate per decisioni quotidiane nell'ambito delle attività dei sistemi di istruzione o per decisioni relative alla revisione e al cambiamento del sistema stesso. Il termine "monitoraggio" può essere considerato come un ulteriore aspetto del processo di valutazione, enfatizzando la connessione con la raccolta di informazioni così come con la gestione delle decisioni, la dipendenza dai dati di tipo amministrativo e una più forte preoccupazione rispetto alla descrizione di ciò che è "di valore" (Scheerens, 2004).

La restituzione delle informazioni valutative può rappresentare un importante punto di partenza per le eventuali azioni correttive e di miglioramento. La sequenza valutazione-**feedback**-azione, del resto, è un meccanismo centrale in qualsiasi processo di apprendimento, inclusi l'"apprendimento organizzativo" e il "*formative assessment*". L'idea di apprendere dalla valutazione è centrale nel concetto di *formative assessment*, che di solito è incluso nei modelli per la progettazione educativa

* Traduzione di Serafina Pastore.

[inserire attribuzione dei paragrafi?]

1. Per *formative assessment* (che può essere reso in italiano con "valutazione per l'apprendimento" o "valutazione a sostegno dell'apprendimento") si intende un processo di valutazione, attivo, sistematico e intenzionale, finalizzato alla promozione di ulteriore apprendimento negli studenti.

e lo sviluppo nell'istruzione. Valutare significa giudicare il valore di un oggetto, e valutazione, nel senso di un particolare tipo di indagine disciplinata, enfatizza come le azioni di "giudizio" e "attribuzione di valore" siano basate su un approccio di raccolta sistematica delle informazioni. Nel caso in cui la raccolta sistematica di informazioni sia organizzata sulla base dei criteri utilizzati per l'indagine scientifica, allora, il termine di ricerca valutativa è appropriato. Una terza componente importante della valutazione, simile all'attribuzione di valore e all'approccio sistematico alla raccolta di informazioni, è data dal contesto di applicazione: ci si aspetta cioè che i risultati della valutazione siano poi utilizzati in modo pertinente. Si tratta di una situazione tipica spesso correlata alla *policy-evaluation*, in cui i risultati della valutazione dovrebbero dare forma, o almeno dovrebbero avere un certo impatto rispetto alle decisioni politiche. Nella letteratura di settore, le tre componenti di base della valutazione (attribuzione di valore, indagine sistematica e utilizzo per il *decision-making*) assumono un'enfasi diversa. In tutte le possibili definizioni in cui il perseguimento di un obiettivo è posto come centrale, l'attribuzione di valore è preponderante (dal momento che gli obiettivi del programma sono o meno raggiunti, fornisce la base per giudicarlo come successo o insuccesso). Così R. Tyler definisce la valutazione come «il processo per stabilire il grado di effettiva realizzazione degli obiettivi educativi» (Tyler, in Nevo, 1995, p. 10). In estrema sintesi, allora, i processi chiave presenti in tutti i tipi di valutazione sono:

- raccolta di dati attraverso test validi stabiliti per il curriculum;
- interpretazione dei risultati dei test attraverso la creazione di *benchmarks*[c.vo] per i livelli di rendimento desiderati;
- interpretazione diagnostica e formativa dei risultati, e loro applicazione nella restituzione di *feedback* e nello sviluppo di *feedforward*².

Nel *formative assessment* c'è una particolare enfasi sulla fase tre che in realtà corrisponderebbe alla caratteristica specifica di ciò che si intende con l'espressione "*formative*". Allo stesso tempo, la raccolta dati e l'interpretazione dei risultati potrebbero essere problematiche perché relative ai criteri scientifici di una misurazione valida e affidabile e dell'associazione causale tra processi e risultati "modificabili". Per spiegare meglio questo aspetto, diamo un'occhiata più da vicino alle tre fasi.

2. Il *feedforward* indica i processi predittivi relativi al perseguimento dei risultati desiderati e le azioni che dovranno essere messe in campo per orientarsi al futuro.

Fase 1. *Raccolta dati*

La fase 1 dovrebbe prevedere lo sviluppo di test da parte degli insegnanti che dovrebbero sapere come costruire i test e i relativi *items*[c.vo]. Anche se i test sono già forniti, questa fase richiede conoscenze di base di psicometria (cosa è un p-value, un punteggio medio in un test, un coefficiente di correlazione, cosa sono validità e affidabilità).

Fase 2. *Definizione di norma, standards e benchmarks*

La fase 2 presuppone la conoscenza di cosa siano *norma, standards e benchmarks* come una base per l'interpretazione dei risultati dei test (ad esempio, comparandoli ai risultati dello scorso anno, o comparandoli ai punteggi normalizzati delle rilevazioni nazionali);

Fase 3. *Interpretazione valutativa*

La fase 3 comporta il ricavare delle conclusioni valutative per ciascuno studente e per gruppi di studenti (uso diagnostico), ma anche la definizione di possibili azioni di rimedio e di miglioramento; quest'ultima fase potrebbe essere particolarmente difficile.

7.2

Competenze richieste per realizzare i processi chiave
del *formative assessment*

Gli studi in valutazione (Visscher, Ehren, 2011) hanno puntualizzato come gli insegnanti per lo più non dispongano della conoscenza e delle abilità di base per interpretare gli esiti dei test e per costruire da sé i test. In misura ridotta questo aspetto vale anche per la terminologia relativa ai concetti di norma, standard e *benchmarks* e per l'utilizzo pratico di simili costrutti. Del resto, anche la fase 3 rappresenta un compito particolarmente impegnativo perché richiede:

- Abilità diagnostiche, capacità di analizzare gli errori, conoscenza sulle capacità generali di singoli studenti;
- Conoscenza su come procedere “dalla diagnosi alla terapia”, in questo caso pianificazione e adattamento del nuovo periodo di insegnamento (di solito indicato come *feedforward*).

In termini tecnici ciò significherebbe che gli insegnanti dovrebbero disporre di abilità di ricerca; in termini di professionalità docente ciò potrebbe essere espresso, invece, come il possesso della conoscenza didattica disciplinare. Per essere più chiari proviamo a considerare i requisiti tecnici e professionali richiesti. Assumendo la disponibilità di test o di dati valuta-

tivi che riportino gli esiti di apprendimento e i progressi fatti dagli studenti nel perseguire questi esiti, il prossimo aspetto da affrontare è di attribuire questi esiti a un “trattamento didattico”; in altre parole, i metodi che gli insegnanti hanno usato, l’approfondimento disciplinare, l’atmosfera durante le lezioni ecc. Quando queste azioni sono poste in un contesto di ricerca, sono considerate come particolarmente sfidanti in termini metodologici. Una domanda fondamentale è come ci si possa aspettare che gli insegnanti, solitamente non formati come ricercatori, debbano affrontare questi aspetti. Dopo aver risolto anche questo aspetto e aver attribuito credibilmente la performance degli studenti, buona o meno che sia, agli aspetti dell’insegnamento e dell’istruzione (in altre parole è stata fatta una “diagnosi”) la domanda successiva è come passare dalla diagnosi alla “terapia”.

Si possono considerare tre approcci:

1. fare appello all’esperienza e alle abilità professionali degli insegnanti e fidarsi del loro giudizio clinico su come procedere, possibilmente assistiti dai colleghi;
2. individuare “ciò che funziona” nella pratica, facendo degli esperimenti;
3. ricorrere alle soluzioni “basate sulle evidenze” riprese dalla ricerca educativa.

Queste richieste sulle competenze valutative dell’insegnante sono spesso incluse nell’espressione “insegnanti come ricercatori”. La sostenibilità di questa prospettiva dovrebbe essere considerata con molta cautela. È davvero ambizioso svolgere degli esperimenti e conoscere come usare la letteratura scientifica. Dal nostro punto di vista, sulla scorta della consapevolezza di quanto sia difficile, come ricercatori, condurre degli esperimenti in ambito educativo e sapendo quanto complessa, diversificata e contestata sia la base di conoscenza dell’efficacia educativa, queste ambizioni potrebbero apparire irrealistiche. Al riguardo però proponiamo due possibili direzioni risolutive:

1. sviluppare una “metodologia della ricerca semplificata” per gli insegnanti. Un esempio è il metodo del “*data-team*” sviluppato da Schildkamp e colleghi (Schildkamp *et al.*, 2014);
2. fornire strumenti alle scuole che guidino la pianificazione didattica basata sui test attraverso dei pacchetti multimediali. Nella parte centrale di questo capitolo presentiamo l’approccio di **RTTI** come un esempio di pacchetto multimediale in grado di facilitare il *formative assessment* e supportare gli insegnanti nel perseguire i compiti principali sopra richiamati in modo responsabile.

7.3

La descrizione di **RTTI**
L'approccio RTTI e le sue basi

In un senso molto generale **RTTI** può essere considerato come un metodo per migliorare l'insegnamento e rinforzare le performance degli studenti. Per tale ragione, **RTTI** rientra nell'ambito della ricerca sull'efficacia dell'istruzione e sul miglioramento scolastico. L'approccio **RTTI** è associato a una delle strategie di maggior successo, vale a dire "le strategie di riforma guidate dalla valutazione" (*evaluation driven reform strategies*). Tali strategie hanno le seguenti caratteristiche:

- focus sugli esiti di apprendimento degli studenti;
- azioni che impattano direttamente sull'apprendimento degli studenti;
- ruolo guida della valutazione;
- allineamento di valutazione e programmazione;
- applicazione ciclica della valutazione (diagnosi, **feedback**);
- condizioni organizzative della scuola in un ruolo secondario e di supporto.

La valutazione e il **feedback** possono essere considerati come determinanti nel guidare il miglioramento a livello di scuola e di classe. Diversi gli aspetti implicati: la chiarezza della finalità formative negli standard, l'esame dei **syllabi**[corretto?],[**elim**] ecc., la verifica di ciò che gli studenti hanno appreso, l'identificazione delle aree di forza e debolezza nei contenuti e nelle abilità che sono stati acquisiti, la restituzione del **feedback** e la diagnosi dei modelli di esito, la considerazione delle strategie correttive e la definizione di obiettivi concreti di miglioramento a livello di studenti, di classe e di scuola, in cooperazione con altri insegnanti, dirigenti e l'eventuale staff di supporto. Quest'ultima caratteristica potrebbe rendere valutazione-**feedback**-azione di correzione sistematica centrale per lo sviluppo professionale correlato ai compiti e alla cooperazione degli insegnanti.

7.4

Zoom sull'approccio centrato sulla valutazione
 per il miglioramento dell'istruzione
Un fondamento logico convincente

La ricerca sull'efficacia dell'istruzione ha una particolare relazione con la pratica didattica. Da un lato, "ciò che funziona", come definito dalla ricer-

ca, sembra poter essere direttamente applicabile. Dall'altro, però, l'approccio di ricerca che si concentra su fattori diversi, per necessità, non cattura del tutto la complessità della pratica didattica. Un approccio centrato sulla valutazione potrebbe essere teoricamente posizionato nel contesto della cibernetica intesa come approccio transdisciplinare per esplorare sistemi normativi. Il meccanismo-chiave consiste in una sequenza ciclica di valutazione, **feedback** e azione correttiva. La pratica di ricorrere alle informazioni valutative sul funzionamento organizzativo come una base per un'azione di correzione o di miglioramento è probabile abbia un andamento passo dopo passo, un orientamento incrementale in cui gli obiettivi o le attese svolgono la funzione di standard per interpretare le informazioni valutative. La discrepanza tra l'attuale livello di raggiungimento degli obiettivi e le attese crea le dinamiche che potrebbero eventualmente portare a una maggiore efficacia. "In teoria" l'approccio centrato sulla valutazione ha una serie di punti di forza:

- il focus chiaro sul rendimento scolastico garantisce di rimanere vicini al cuore dei "processi produttivi" nell'istruzione, in particolare ciò è che relativo all'apprendimento degli studenti;
- la possibilità di allineare le valutazioni, incluso il *formative assessment*, agli esami e alle valutazioni sommative di tipo *high-stakes*, può essere considerata come un potente meccanismo per aumentare "l'opportunità di apprendere" (come un buon incontro tra ciò che è insegnato e ciò che è valutato);
- il **feedback** basato sui risultati della valutazione è un importante stimolo sia per l'apprendimento degli studenti sia per l'insegnamento, con implicazioni di tipo cognitivo e motivazionale. Il **feedback** si riferisce a riportare i risultati della performance ai produttori, appunto studenti e insegnanti. Le informazioni su alcuni aspetti deficitari possono essere utilizzate in una successiva pianificazione didattica, e la ricerca ha mostrato come questo "**feedback** strumentale" stimoli l'apprendimento degli studenti (Kluger, DeNissi, 1996);
- quando si prende alla lettera la teoria del controllo è sufficiente un buon meccanismo di valutazione per consentire un efficace funzionamento organizzativo; in particolare nel caso di organizzazioni professionali come la scuola, tutto il resto potrebbe rimanere "deregolamentato";
- infine, le politiche di *accountability* a livello di sistema possono anche essere considerate come una parte della strategia basata sulla valutazione per l'efficacia dell'istruzione. Sebbene questa posizione non sia priva di contestazioni, l'*accountability* esterna può essere una condizione stimolan-

te per “l’*accountability* interna” (auto-valutazione della scuola e *formative assessment*) (Carnoy *et al.*, 2003[non in biblio]).

7.5

Evidenze di tipo misto

Tra le varie rassegne di studi e meta-analisi riguardo valutazione e *feedback* si possono indicare quelle di P. Black e D. Wiliam (1998), J. Hattie e H. Timperly (2007), V. J. Shute (2008), J. Hattie (2009), D. Wiliam (2011) L. S. Fuchs e D. Fuchs (1985)[1986 in biblio], S. E. Kim (2005), J. Hattie (2009)[ripetizione?], M. Burns e T. Symington (2002), R. L. Bangert *et al.* (1991), A. N. Kluger e A. DeNisi (1996), J. M. Faber e A. J. Visscher (2014), M. A. Hendriks *et al.* (2014). J. Scheerens (2016) fa una distinzione tra meta-analisi di studi relativi al *formative assessment*, valutazione e *testing*[c.vo] a livello di classe e gli studi condotti sul *feedback*. La differenza tra le categorie “*formative evaluation*” e “valutazione e *testing*” è relativa alla mancata indicazione, per gli studi condotti su quest’ultima categoria, dell’utilizzo dei test in termini *formative* o *summative*. I risultati dimostrano differenze consistenti tra le varie meta-analisi per ciascun concetto. Tra i tre diversi tipi di valutazione che sono stati discussi, i risultati sul *feedback* sono particolarmente alti, compresi tra .51 e 1.10. I risultati sul *formative assessment* variano tra .06 e .07 e quelli sulla valutazione e il *testing* a livello di classe tra .10 e .39. I risultati di J. Scheerens indicano, inoltre, che gli *effect-size* ottenuti in rassegne e meta-analisi più recenti tendono a essere considerevolmente più bassi di quelli ottenuti da studi più vecchi. Nel confrontare questi risultati delle meta-analisi su tempo di insegnamento e leadership scolastica, i risultati sulla valutazione, il *testing* e il *feedback* sono stati considerevolmente più alti (Scheerens, 2016). L’*effect-size* complessivo per il *feedback* probabilmente è dovuto al fatto che la maggior parte dei risultati sulle variabili del *feedback* sono stati calcolati in studi di livello micro, condotti a livello di classe.

7.6

Alto potenziale, implementazione sub-ottimale

Sebbene il quadro generale offerto dalle evidenze riportate nella sezione precedente paragoni in modo positivo i risultati delle meta-analisi su altre variabili di miglioramento dell’efficacia spesso studiate, queste posso-

no apparire non ottimali a fronte dei vantaggi teorici presentati in precedenza (Hendriks *et al.*, 2014; Scheerens *et al.*, 2007[non in biblio]; Scheerens, Bosker, 1997[non in biblio]). Per varie ragioni l'applicazione e l'implementazione di approcci basati sulla valutazione nelle scuole sono state spesso limitate da ragioni di carattere culturale, politico, tecnico e pratico:

- resistenze culturali. Nel mondo dell'istruzione, una prospettiva orientata al rendimento è stata tradizionalmente considerata con sospetto. E quando si è combinata con la valutazione dell'istruzione ha evocato una totale resistenza. Per la teoria organizzativa della scuola come “burocrazia professionale”, i professionisti altamente qualificati non amano la valutazione esterna del loro lavoro (Mintzberg, 1979). Anche quando si tratta di valutazioni per l'apprendimento degli studenti, è difficile liberarsi di questa cattiva immagine nelle scuole. E sebbene negli ultimi decenni questa situazione sia cambiata, la diffusione nelle scuole del processo di inclusione della valutazione e degli approcci basati sulla valutazione è ancora lenta a diffondersi;
- questioni di proprietà. In molti sistemi scolastici, l'autonomia della scuola e degli insegnanti è stata fortemente stimolata durante gli ultimi trent'anni. Gli strumenti per la valutazione sono stati spesso sviluppati esternamente e questo ha portato alla presenza di corpi estranei nella vita della scuola;
- applicazione incompleta. Ci sono numerosi esempi di scuole che hanno avuto considerevoli problemi nel raccogliere i dati valutativi che poi non hanno mai usati. Tabelle e report muoiono di morte tranquilla negli armadi, non sono argomento di discussione tra i docenti né sono usati per migliorare l'apprendimento degli studenti. Le fasi di **feedback**, interpretazione diagnostica dei dati e degli input per la successiva pianificazione didattica sono cruciali per il successo delle valutazioni;
- aspetti di carico amministrativo e burocratico. Raccogliere dati per finalità valutative richiede tempo, e registrare questi dati, forse ancora più tempo. A livello internazionale ci sono esempi alquanto sfortunati di scuole e insegnanti chiamati a un ulteriore carico di lavoro in termini di ore, anche quando non è chiara la rilevanza per migliorare insegnamento e apprendimento.

7.7

Considerazioni per un'applicazione di successo

I vincoli cui si è fatto riferimento nella sezione precedente possono essere considerati come sfide per migliorare la situazione. Si potrebbero considerare le seguenti soluzioni:

- ripensare la connessione tra *formative* e *summative assessment*. Concentrarsi sul *formative assessment* sembra essere la via più semplice per risolvere i problemi sopra indicati. Tuttavia, ci sarebbero più opportunità di migliorare l'efficacia dell'istruzione se il *formative assessment* fosse ben allineato con i test e le valutazioni di tipo *high-stakes*;
- massimizzare l'autonomia e la "proprietà" degli insegnanti. A seconda del margine di manovra che le scuole devono acquisire, al supporto esterno, i prodotti disponibili in commercio hanno il vantaggio di essere letteralmente "posseduti" dalla scuola. I seguenti pacchetti potrebbero fornire ampie opportunità alle scuole e agli insegnanti di usare il materiale nel modo più rilevante, dedicato a una particolare situazione; per esempio gli insegnanti potrebbero essere stimolati e supportati[sostenuti] a sviluppare dei test propri;
- stimolare e supportare l'applicazione ciclica del *testing*. Una sfida maggiore è quella di portare le scuole oltre la mera raccolta di dati e di arrivare all'applicazione e all'utilizzo concreto dei risultati dei test. Nel contesto olandese questo tipo di approccio si interseca perfettamente con l'idea di stimolare un "lavoro orientato ai risultati" (in olandese, *opbrengstgericht werken*). Non si tratta di una piccola cosa da realizzare e richiede supporto attraverso lo sviluppo professionale e dettagliate linee guida;
- applicare ICT di tipo user-friendly. Simili applicazioni sono importanti per mettere a disposizione il materiale e supportare i processi chiave di applicazione ciclica dei test, ma anche per facilitare la somministrazione agli studenti e la registrazione dei dati.

Le pubblicazioni sull'efficacia dell'istruzione e sul miglioramento scolastico tendono a trasmettere un messaggio positivo tipo "ricette per il successo". L'analisi delle evidenze di ricerca è un'esperienza che fa riflettere. Sebbene ci sia un considerevole accordo su quali siano le condizioni importanti dell'organizzazione scolastica e dell'insegnamento, l'evidenza su quanto queste condizioni funzionino è molto più frammentata. Inoltre gli *effect-sizes* sono spesso molto ridotti quando confrontati a standard definiti, e spesso sono anche inferiori rispetto a quelli delle condizioni "non-modificabili", come l'attitudine degli studenti, lo status socio-[elim]economico e le variabili aggregate a livello di scuola e di classe. Del resto, il miglioramento in "condizioni di processo" della scuola richiede spesso un tempo considerevole e uno sforzo per mostrarsi poi in piccoli incrementi nel miglioramento dell'apprendimento degli studenti. In altre parole, spesso si investe un chilo nel processo di miglioramento per ottenere un'oncia di successo negli esiti. Ciò che la letteratura scientifica dimostra è che alcu-

ne condizioni funzionano meglio di altre. Avere a che fare con le organizzazioni scolastiche e con l'ennesima moda nell'approccio della leadership scolastica non porterà alcun effetto, a meno che non ci sia una chiara connessione con le strategie che incidono sul processo primario dell'apprendimento degli studenti. Le strategie di miglioramento strettamente connesse all'apprendimento degli studenti e alla valutazione degli effetti di apprendimento hanno il potenziale per essere non solo efficaci, ma anche efficienti. L'approccio **RTTI**, come descritto più avanti, si presta a essere un approccio efficace e "agile" per migliorare l'apprendimento degli studenti. Nella sezione finale torneremo sulla questione della formazione docente affrontando due aspetti:

1. la presentazione di un metodo strutturato, qual è l'approccio **RTTI**, che aiuti gli insegnanti a compensare le lacune nella loro formazione e preparazione;
2. la necessità, per la formazione di base, di un certo grado di *scaffolding* quale quello offerto con il metodo **RTTI**.

7.8

L'approccio **RTTI** per migliorare l'apprendimento degli studenti

Quando abbiamo introdotto l'approccio **RTTI** come approccio basato sulla valutazione, gli insegnanti, a livello di scuola e di classe, hanno mosso le seguenti osservazioni rispetto all'attuale situazione scolastica nei Paesi Bassi. Gli insegnanti, a volte, non dispongono delle competenze relative al *testing* e non dispongono del *know-how[t.do]* richiesto per costruire da sé test e *items* che siano validi e affidabili. Per i test esistenti, la loro conoscenza di base in psicommetria non sembra essere sufficiente (ad esempio rispetto alla gestione di domande come: cosa sia un p-value, cosa sia la validità, cosa sia l'affidabilità). Non c'è stata sempre un'attenzione sufficiente nella formazione docente su come costruire e analizzare i test (Sanders, 2013). Inoltre, manca una letteratura utile sulla costruzione dei test perché le pubblicazioni disponibili tendono a concentrarsi principalmente sulla psicommetria e non invece sull'applicabilità per la pratica dell'insegnante. Gli insegnanti frequentemente usano i test che sono parte integrante dei metodi didattici e dei libri di testo, e poiché sono particolarmente diffusi, si pensa che questi test siano di buona qualità, cosa che però non è sempre vera. La

validità di contenuto di questi test è di solito eccellente, ma la validità di costruito e il modo in cui i test sono costruiti non lo sono. L'allineamento tra i vari test di solito non è adeguato: per molti soggetti nell'istruzione superiore c'è un divario tra i livelli di laurea di base e specialistica in quanto test e contenuti, specialmente negli anni inferiori, sono spesso focalizzati su un ordine di pensiero inferiore e inadatto per la valutazione per l'apprendimento. [cosa introduce l'elenco che segue?]

– Di solito i dati ricavati dalle rilevazioni esterne sono recepiti dal management della scuola e non sono condivisi con gli insegnanti. Inoltre, accade che quando i dati dei test sono resi disponibili agli insegnanti, non tutti gli insegnanti possono interpretare e utilizzare quelle informazioni. Esistono consistenti differenze rispetto alle abilità di ricerca degli insegnanti nell'effettuare ulteriori ricerche in risposta a questi dati. Molti insegnanti non sono ancora sufficientemente in grado di passare dal monitoraggio alla diagnosi e dalla diagnosi alla terapia (Datnow *et al.*, 2012);

– gli obiettivi di apprendimento non sono sempre indicati in modo chiaro dagli insegnanti. Cosa significa se uno studente deve studiare il capitolo 3 per il test? Alcuni studenti pensano di dover memorizzare il capitolo 3, altri pensano di dover esercitarsi un sacco con i compiti assegnati, altri ancora, convinti che le cose andranno bene perché hanno partecipato alle lezioni, e pensano di non aver bisogno di una grande preparazione. Manca la definizione degli obiettivi così come del contenuto (Folmer *et al.*, 2017, p. 87). Per tale ragione non è sempre chiaro agli studenti cosa ci si aspetta da loro in termini di preparazione. Di conseguenza, il **feedback** fornito dagli insegnanti non è sempre efficace;

– inoltre, nel fornire una base per l'attribuzione dei voti, i test sono sempre più utilizzati per dare informazioni sul processo di apprendimento a insegnanti e studenti. Gli insegnanti sono in linea di principio ben disposti verso la valutazione per l'apprendimento, a condizione che non dia loro troppe responsabilità extra amministrative;

– la cultura e gli aspetti organizzativi in alcune scuole non supportano a sufficienza la valutazione per l'apprendimento (Datnow *et al.*, 2012; Oprins, Andriessen, 2002). Questo, in parte, è imputabile alla mancanza di un processo di qualità (ciclo). Elementi del ciclo di qualità (formulazione degli obiettivi, costruzione dei momenti di valutazione, **feedback** e discussione sull'*accountability*), assicurano la qualità stessa. Gli insegnanti che hanno già fatto esperienza di una pressione lavorativa così alta, hanno poco tempo per pensare al miglioramento della qualità. Affrontare le normali questioni quotidiane impedisce la riflessione. Inoltre, per facilitare

gli aspetti strutturali, la cultura della qualità in una scuola rappresenta un importante fattore di successo, nel senso di una cultura in cui gli insegnanti gestiscono attivamente le informazioni, si confrontano l'uno con l'altro, lavorano assieme e hanno l'opportunità di imparare dagli altri (Onderwijsinspectie, 2015). La struttura e la cultura possono rinforzarsi l'una con l'altra così da creare una base sufficiente e un tempo e un focus sulla valutazione per l'apprendimento e sull'interpretazione valutativa (Datnow *et al.*, 2012).

Non è che gli insegnanti non siano desiderosi di migliorare le loro pratiche di valutazione, piuttosto non sanno come fare e il sistema non li **supporta[sostiene]** in questo (Lyons, 1998). È difficile cambiare e rinnovare, in modo fondamentale, routine ormai consolidate (Onstenk, 2004). Inoltre, la situazione scolastica immediata non lascia molto tempo per riflettere. Gli insegnanti prendono decisioni tempestive che si traducono in un'azione immediata (Eraut, 1994). Per portare avanti i processi chiave della valutazione per l'apprendimento è pertanto necessario integrare le competenze degli insegnanti e disporre di una cultura e di una struttura scolastica che siano di supporto in tal senso (Onderwijsinspectie, 2015). Questo implica un approccio con un meta-**[elim]**linguaggio efficace riguardo l'apprendimento, in modo che non ci sia una terminologia ambigua sull'apprendimento a scuola utilizzabile da tutti i soggetti, in tutti i gradi e in tutti i tipi di scuola. Questo approccio necessita di essere supportato con metodi pratici e con strumenti per la valutazione per l'apprendimento, idealmente attraverso applicazioni web-based, in cui il **feedback** concreto sulla costruzione del test rispetto ai valori psicometrici è pre-**[elim]**strutturato per gli insegnanti. Questo approccio richiede, inoltre, di affrontare la necessità di un ciclo di qualità facile da integrare nella struttura scolastica esistente combinata con una cultura orientata alla qualità.

7.9

Un meta-**[elim]**linguaggio per la valutazione a sostegno dell'apprendimento

La classificazione cognitiva **RTTI**, un meta-**[elim]**linguaggio sull'apprendimento per migliorare le performance degli studenti, è sviluppata a partire dalle pratiche didattiche e dalle classificazioni sistematiche esistenti rispetto alle tassonomie dell'apprendimento. Una delle più note tassonomie è

quella di B. Bloom (versione rivista). Inizialmente abbiamo provato a usare questa tassonomia come un meta-[elim]linguaggio, ma usare le 19 categorie della versione di Bloom non si è rivelato efficiente così come non è stato efficace il ricorso alle 6 categorie principali per la valutazione a supporto dell'apprendimento in quanto non descrivevano adeguatamente il processo di apprendimento in modo tale da fornire un **feedback** mirato efficace. Inoltre, la validità del percorso di apprendimento continuo non era assicurata poiché le 6 categorie non forniscono un sufficiente aggancio agli obiettivi di apprendimento. Il capitolo 17 del manuale di Bloom (versione rivista) indica diversi problemi che, una volta risolti, potrebbero rendere il quadro di riferimento molto più utile (Anderson, Krathwohl, 2001, p. 295). Nel prosieguo del capitolo andiamo oltre le soluzioni di B. Bloom e ci occupiamo dei seguenti aspetti: [elenco in tondo]

- il processo di integrazione di curriculum, insegnamento e valutazione;
- l'utilità del quadro di riferimento per gli studenti;
- la relazione tra il quadro di riferimento e una teoria dell'apprendimento ampiamente accettata, significativa e utile.

Questi aspetti sono stati affrontati nello sviluppo dell'approccio **RTTI** e dei relativi strumenti, assieme al già richiamato punto di partenza per lo svolgimento dei processi chiave della valutazione per l'apprendimento.

7.10

Dalla versione rivista di Bloom all'approccio **RTTI**

La prima categoria principale della tassonomia di Bloom nella versione rivista è *Ricordare*, che è simile alla definizione di *Ricordare* (R) di **RTTI**.

Le domande della categoria *Ricordare* (R) sono finalizzate a testare quanto gli studenti siano in grado di riconoscere, memorizzare e riprodurre informazioni fattuali cruciali e rilevanti. Il contenuto disciplinare per queste domande è indicato o segnato in modo esplicito.

Non ci si aspetta che gli studenti, nel rispondere alle domande di questa categoria, aggiungano qualcosa rispetto al contenuto.

La seconda categoria principale, *Comprendere* (U), è in pratica percepita da insegnanti e studenti come diffusa ma non come efficiente per la valutazione a supporto dell'apprendimento. Per un'efficiente *formative assessment*, le sette abilità del *Comprendere*, sono state suddivise in *Training* (T₁), *Transfer* (T₂), e *Insight-Innovation* (I). Le categorie della versione di Bloom *Fare esemplificazioni* e *Classificare* rientrano nella categoria *Training* dell'approccio

RTTI. *Interpretare, Comparare e Spiegare*, invece, nella categoria *Transfer*. *Sintetizzare e Fare inferenze*, infine, alla categoria dell'*Insight*.

Per la terza categoria principale della versione di Bloom, *Applicare*, sembra che la maggior parte delle domande dei test e dei contenuti rientri in questa categoria. Non riguardano solo il ciclo primario e secondario, ma anche l'istruzione universitaria. Al fine di fornire agli studenti un **feedback** mirato e di organizzare il processo didattico in modo responsabile ed efficace (allineamento curricolare) è necessario fare una distinzione nell'ambito della categoria *Applicare*. L'approccio **RTTI** distingue esplicitamente due livelli di applicazione attraverso il *Training* (T₁) e il *Transfer* (T₂). T₁ e T₂ mostrano somiglianze con le due abilità sottostanti la categoria principale dell'applicazione di Bloom: da un lato, eseguire una procedura in un compito routinario o in un contesto già noto (*Applicazione-Esecuzione*), dall'altro, eseguire una procedura in un nuovo compito o in un contesto inedito (*Applicazione-Implementazione*). Nel processo di apprendimento degli studenti c'è un'importante differenza tra i due livelli di applicazione. In pratica, ci sono studenti che sono perfettamente capaci di risolvere domande come è stato loro insegnato (T₁) con risposte semplici e incontrovertibili, ma con domande in cui c'è un'applicazione orientata del transfer (T₂) devono rispondere in base al contesto che cambia. Con i soggetti che presentano danni nella corteccia prefrontale, per le domande orientate al transfer, si osserva come altre aree del cervello siano attive rispetto a quelle implicate nelle decisioni orientate alla correttezza (Sousa, 2009). Oltre alle domande T₁ orientate all'esercizio è anche importante distinguere le domande T₂ orientate al transfer perché «imparare a generalizzare (transfer) è il cuore di una buona istruzione. Se i nostri studenti possono solo riprodurre ciò che hanno appreso da noi, non potranno fare molto, perché quelle esatte circostanze probabilmente non si verificheranno di nuovo. L'apprendimento è utile soltanto se lo studente può applicare ciò che ha appreso fuori dal contesto dell'apprendimento» (William, Leahy, 2018, p. 43). Per una domanda della categoria *Training* (T₁), ci si esercita, la fonte è riconoscibile e la domanda presenta un grado di difficoltà simile a quella presente nell'eserciziario. Gli studenti riconoscono le variabili nella nuova fonte, e gli aspetti a cui prestare attenzione sono riconoscibili e abbinati alla pratica durante le lezioni. Come menzionato prima, *Fare esemplificazioni* e *Classificare* rientrano nella categoria T₁. *Fare esemplificazioni* si presenta quando uno studente deve selezionare o produrre un esempio specifico o un aspetto di un concetto generale o di un principio, ad esempio, “[<<]fare degli esempi di vari stili artistici nella pittura”[>>] (Anderson *et al.*, 2001).

Classificare è un processo complementare all'eseplificazione e si ha quando uno studente riconosce che qualcosa appartiene a una determinata categoria (*ibid.*[Anderson *et al.*?], 2001).

Le domande orientate al Training (T₁) sono finalizzate a valutare l'applicazione di una data conoscenza disciplinare in una situazione nota, comparabile a quella in cui è stata acquisita. Le domande della categoria T₁ stabiliscono fino a che punto lo studente ha padroneggiato le procedure apprese passo dopo passo, i metodi o le "ricette", simili alle situazioni in cui ha appreso. È importante che le domande del T₁ non richiedano la scoperta di nuove informazioni. Per tale ragione, queste domande non dovrebbero valutare le eccezioni a una regola appresa o non dovrebbero essere di "falsa pista".

Per le domande relative all'applicazione del transfer (T₂), la nuova fonte è diversa rispetto a quella su cui gli studenti si sono esercitati, e questa deve pertanto essere prima resa utilizzabile. Inoltre, una domanda della categoria T₂ non può essere affrontata in modo sequenziale passo dopo passo, ma richiede per essere risolta che i diversi passaggi siano messi assieme e combinati in modo pratico. Come già anticipato, *Interpretare*, *Comparare* e *Spiegare* nella versione rivista della tassonomia di Bloom rientrano nella categoria T₂.

Si ha l'*Interpretazione* quando uno studente è in grado di convertire le informazioni da una forma di rappresentazione a un'altra. Per aumentare la probabilità che l'interpretazione sia valutata, l'informazione inclusa nel compito di valutazione deve essere nuova, inedita rispetto al percorso didattico (*ibid.*[Anderson *et al.*?], 2000). La *Comparazione* include il ricercare somiglianze, cercare corrispondenze tra elementi e modelli in un oggetto, evento o idea e quelli in un altro oggetto, evento o idea. Nel comparare, quando l'informazione è fornita, uno studente ricerca le corrispondenze con la conoscenza più familiare (*ibid.*[Anderson *et al.*?], 2000). La *Spiegazione*, infine, si ha quando uno studente è capace di costruire e usare un modello causa-effetto di un sistema, o cercare spiegazioni per un sintomo o provare a risolvere problemi per fare una diagnosi (*ibid.*[Anderson *et al.*?], 2001).

Le domande orientate al transfer (T₂) valutano l'applicazione di una disciplina a una nuova situazione che richiede un transfer. Queste domande valutano la raccolta, l'applicazione e il trasferimento dei materiali appresi da parte dello studente a nuovi contesti. Sono finalizzate cioè a stabilire l'abilità di uno studente a selezionare le variabili o i dati corretti, deter-

TABELLA 7.1

RTTI e Tassonomia rivista di Bloom a confronto

| | |
|---|---|
| R | Ricordare – Riconoscere; Ricordare – Richiamare |
| T | Comprendere – Esempificare; Comprendere – Classificare Applicare – Eseguire |
| T | Comprendere – Interpretare; Comprendere – Comparare; Comprendere – Spiegare; |
| T | Applicare – Implementare |
| I | Comprendere – Sintetizzare; Comprendere – Fare inferenze; Analizzare – Differenziare; Analizzare – Organizzare; Analizzare – Fare attribuzioni Valutare – Controllare; Valutare – Criticare |
| I | Creare – Generare; Creare – Pianificare; Creare – Produrre |

minare quale procedura sia più opportuna, e combinare parte dei piani o delle procedure per risolvere le domande di tipo T₂.

Le tre categorie principali della versione rivista della tassonomia di Bloom, *Analizzare*, *Valutare* e *Creare*, sono state tutte inserite nella categoria *Insight-Innovation* del modello RTTI. Nella versione di Bloom tra l'altro queste categorie sono spesso agganciate alla conoscenza metacognitiva ([Anderson *et al.* o Bloom?]p. 239). Le tre abilità principali raramente sono insegnate a scuola, o all'università, in quanto di solito si enfatizzano altre abilità. Diversi argomenti scolastici ricalcano le tre celle (*Ricordare*, *Comprendere*, *Applicare*) della Tabella della Tassonomia (*ibid.*[Anderson *et al.*?], 2001). Per efficienza, sono state combinate assieme nella categoria *Insight-Innovation* del modello RTTI.

Uno studente può rispondere a una domanda *Insight-Innovation* se è in grado di costruire il contesto o il metodo in modo indipendente e lavorare in maniera innovativa, fuori e dentro i confini curriculari. Lo studente deve pertanto essere capace di collocare e definire la fonte da sé e deve costruire da sé la strategia di risoluzione, che non sia prestrutturata e già utilizzata per esercitarsi, così da giungere alla soluzione corretta. Come già ribadito, *Sintetizzare* e *Fare inferenze* rientrano nella categoria *Insight-Innovation* del modello RTTI. *Sintetizzare* comprende la costruzione di una rappresentazione dell'informazione (*ibid.*[Anderson *et al.*?], 2001). *Fare inferenze* invece comprende un modello all'interno di una serie di esempi o di istanze; un[o] studente è in grado di astrarre un concetto e far confronti tra istanze nel contesto dell'intero insieme (*ibid.*[Anderson *et al.*?], 2001).

Le domande della categoria Insight-Innovation (I) richiedono che lo studente analizzi e valuti in modo autonomo e sistematico da diverse prospettive. È lo studente stesso a dover creare e costruire il contesto e le procedure, dentro e fuori il curriculum, che sono necessari per arrivare alla soluzione.

Di seguito Nella TAB. 7.1 riportiamo una rassegna di come le sei categorie principali e le diciannove abilità della tassonomia rivista di Bloom siano correlate al modello RTTI.

7.11

I processi di apprendimento non gerarchici

«In una tassonomia, le categorie giacciono lungo un continuum» (Anderson *et al.*, 2001, p. 4) e descrivono un sistema di classificazione gerarchico (Block, 1975). L'apprendimento non è sempre sequenziale, motivo per cui RTTI non è definito come una tassonomia, ma è chiamato meta-[elim]linguaggio sull'apprendimento. R non è sempre la fase preliminare alla fase di I: a volte gli studenti che rispondono correttamente alle domande di tipo I potrebbero non fare altrettanto per le domande R e/o T1 e/o T2. Alcuni studenti pensano che le cose complicate siano semplici e che cose semplici siano difficili (William, Leahy, 2018). Un insegnante deve scoprire la logica di apprendimento di uno studente. Questo è possibile solo se riesce a mappare tutti i livelli cognitivi, da R a I, al fine di raccogliere informazioni sul processo di apprendimento e sui talenti degli studenti (e delle classi) così da comprendere i livelli di partenza di ognuno. Pertanto, i test che sono utilizzati per la valutazione a sostegno dell'apprendimento devono includere domande da R a I.

7.12

La raccolta dei dati per la valutazione per l'apprendimento

Definire fino a che punto gli obiettivi formativi siano stati realizzati, (Tyler, citato da Nevo, 1995, p. 10) comincia con la valutazione per l'apprendimento e la raccolta dei dati. «Utilizziamo, in generale, il termine di valutazione per riferirci a tutte quelle attività realizzate dagli insegnanti – e dai loro studenti per valutare se stessi – in grado di fornire informazioni

che possano essere utilizzate come **feedback** per modificare le attività di insegnamento e apprendimento» (Black, Wiliam, 1998, p. 2). Questo dovrebbe implicare lo sviluppo di test da parte dei docenti, il che richiede la conoscenza relativa a come costruire i test e i relativi *items*. La costruzione dei test non sempre ottiene la dovuta attenzione. La letteratura disponibile sulla costruzione dei test è prevalentemente in chiave psicometrica e non è sufficientemente adatta all'azione quotidiana dei docenti. Un insegnante spesso, ad esempio, non ha l'opportunità di fare dei pre-test, e molto frequentemente non ha il tempo e la competenza per realizzare e interpretare calcoli complessi. Tuttavia, se ci si aspetta che gli insegnanti ottengano più informazioni dai test invece che dai voti, allora saranno in grado di controllare in modo condizionato se la qualità di base di un test sia o meno sufficiente. Saranno anche in grado di correggere errori di costruzione, di realizzare analisi e interpretazioni di base nel campo psicometrico, e di decidere quali **step**[passi] sia necessario intraprendere per passare dall'analisi alla diagnosi e infine alla terapia.

Nel Manuale **RTTI** (Drost, Verra, 2018a) sono offerti una serie di strumenti pratici per stimolare e **supportare**[sostenere] gli insegnanti nello sviluppo e controllo dei loro test, incluse le indicazioni per l'uso del materiale nel modo più opportuno, a seconda della situazione di riferimento. Una lista di Controllo e Progettazione aiuta in modo funzionale gli insegnanti, prima di somministrare un test, a vagliare la qualità di base delle domande inserite nel test. Per sviluppare una competenza di base rispetto al metodo **RTTI** in termini di meta-[elim]liguaggio e psicometria, è stato ideato uno strumento e-learning, assieme a un[o] strumento web-based, per supportare la valutazione per l'apprendimento (**RTTI-online**).

7.13

Per un uso della valutazione a sostegno dell'apprendimento

Di solito, si distingue tra valutazione dell'apprendimento e valutazione per l'apprendimento, tuttavia a essere importante non è tanto l'etichetta quanto ciò che si fa con le informazioni raccolte per promuovere l'apprendimento degli studenti. Le informazioni raccolte attraverso i test di carattere sommativo possono essere utilizzate per **supportare**[sostenere] l'apprendimento (Laveault, Allal, 2016). Se la valutazione è a supporto dell'apprendimento, obiettivi di apprendimento e criteri di successo devono essere riportati in termini generali così da promuovere il transfer (Arter, McTigh,

TABELLA 7.2

Esempi di costruzioni di ulteriori domande **RTTI**

| | |
|----------------|--|
| R | Dai una definizione di; Elabora un piano step-by-step; Dai la formula di; Enumera; Nomina; Indica; Come/dove/quando/chi (fatto)?; Chiama/traduci "conoscenza di base" |
| T ₁ | Calcola; Cerca; Raggruppa; Descrivi a parole tue; Fai un esempio; Completa; Cerca di utilizzare le caratteristiche apprese; Coniuga; domande "Cerca-trova" |
| T ₂ | Argomenta cosa va meglio; Dimostra con; Fai un'analisi dell'errore; Spiega; Spiega con; Mostra come; Fornisci una relazione/connessione causa-effetto; Fai una combinazione; Confronta; Osserva; Apprezza; Coordina |
| I | Costruisci un'ipotesi; Sviluppa/progetta/sposta; Analizza; Fornisci un nuovo argomento per; Fai una sintesi rilevante; Valuta criticamente; Ricava (somiglianze/differenze); Fai previsioni; Monitora e concludi; Progetta un percorso di ricerca; Valuta; Allena; Prova |

2001 in Wiliam, Leahy, 2018). I test sono utilizzati sempre meno e non solo per misurare in termini di voti la performance di uno studente. Al contrario, si sta incredibilmente diffondendo la tendenza a ricavare le informazioni per lo sviluppo dello studente. I risultati dei test dovrebbero fornire informazioni rispetto a[lla] profondità dell'apprendimento, diversi livelli di qualità, autonomia di apprendimento, guida o supporto, transfer di apprendimento, qualità per l'apprendimento e/o situazioni inedite (Laveault, Allal, 2016). Per quanto riguarda l'uso dei test a supporto dell'apprendimento, gli insegnanti devono trovare la causa per un risultato (tardivo). Per questa valutazione a **supporto[sostegno]** dell'apprendimento, il test dovrebbe contenere domande dalla categoria R alla categoria I. Ad esempio, se uno studente non risponde a una domanda della sezione T₂, la ragione potrebbe risiedere in una conoscenza carente a livello di R, in una pratica insufficiente con le domande della sezione T₁ o, se lo studente va bene nella sezione I, nell'ambito linguistico della domanda della sezione T₂. Va da sé che una logica **RTTI** deve essere scelta in modo che si adatti agli obiettivi di apprendimento, al tipo e al grado di scuola. Oltre a ciò, gli insegnanti possono stimolare gli studenti in base ai loro punti di forza perché **RTTI** prevede che si spieghi il voto attribuito agli studenti. Per esempio, per uno studente che ha ottenuto un buon punteggio in I ma non in R, cercare la logica in ciò che deve essere appreso sarà molto più utile che dividere la materia in pezzi da ripetere. In pratica, però, non tutti i test contengono domande da R a I. Per esempio, espressioni per le domande, come quelle

riportate ~~nella seguente tabella nella~~ **TAB. 7.2**, possono essere usate dagli insegnanti per progettare ulteriori domande per test a uso formativo se ci sono lacune in determinato livello cognitivo.

7.14

Interpretare i risultati per valutare, fare diagnosi, supportare[sostenere] l'apprendimento

Se gli insegnanti dispongono di un test *formative*, ciò non significa che debbano anche valutare per supportare l'apprendimento. La base per l'interpretazione dei risultati dei test richiede non solo la conoscenza di norme, standard e livelli di apprendimento desiderati, ma anche la conoscenza dell'interpretazione diagnostica e formativa dei risultati ottenuti. Inoltre è necessario applicare tale conoscenza nella restituzione di un **feedback** e nella definizione di azioni sia per i singoli studenti sia per gruppi di studenti (uso diagnostico). Questo include anche pensare a come articolare le azioni di recupero e di miglioramento.[:] **S[s]**i tratta di una fase davvero difficile per molti insegnanti. Gli insegnanti considerano la valutazione a supporto dell'apprendimento come non particolarmente complicata,[:] allo stesso tempo, però, è un carico di lavoro ulteriore rispetto alle loro attività didattiche e molto raramente ne è parte integrante (Dannow *et al.*, 2012). Supportare l'apprendimento con un **feedback** concreto, *feedforward* e con azioni di recupero è fondamentale per gli insegnanti. **Gli insegnanti[Questi ultimi]** non sono formalmente formati come ricercatori, tuttavia ci si aspetta che affrontino questi aspetti. La formazione, in tal senso, è davvero importante per via della centralità attribuita al rendimento scolastico e alla strategia che influisce sul processo primario dell'apprendimento degli studenti. Può avvenire, al meglio, come formazione *on the job*, passando dal tradizionale focus sul transfer di conoscenza e abilità allo sviluppo di una comprensione *della e nella* pratica professionale (Dall'Alba, Sandberg, 2006). **RTTI-online** contiene importanti elementi per aiutare gli insegnanti nella loro valutazione a supporto dell'apprendimento.

7.15

Applicazione **ICT** di tipo *user-friendly*

L'applicazione web di **RTTI-online** fornisce il supporto per la raccolta sistematica di informazioni, l'analisi, la diagnosi e le azioni, il **feedback** su

misura e il *feedforward* per l'insegnante, gli studenti e il management della scuola. L'applicazione online di **RTTI** può essere utilizzata per ogni forma di valutazione, in ogni momento, per un test realizzato dal docente, a livello nazionale o per una valutazione integrata all'azione di insegnamento:

- aiuta gli insegnanti a bilanciare adeguatamente la valutazione rispetto alle differenti categorie R, T₁, T₂ e I;
- analizza in modo automatico i risultati e ne monitora e controlla la qualità;
- suggerisce un **feedback** su misura e un *feedforward* per insegnanti e studenti;
- fornisce una valutazione a supporto dell'apprendimento.

Lo strumento di riflessione automatica QA (Analisi della Qualità) per l'insegnante (*formative assessment* del test) consiste nei seguenti quattro step.

Step 1. *Monitorare l'allineamento curricolare*

Quando si evidenzia una deviazione consistente tra la relazione effettiva e quella ottimale con il modello **RTTI** si effettuano aggiustamenti progressivi. Questo solleva la domanda sul posizionamento di quel test nel processo dell'apprendimento continuo. Un processo di apprendimento continuo indica come gli studenti, per un dato corso durante la loro carriera scolastica, progrediscono dalla posizione di partenza fino al livello dell'esame. È un programma continuo in cui si è pensato deliberatamente a cinque aspetti che assicurassero un allineamento sostanziale: quali sono i nostri obiettivi di apprendimento, quale materiale è offerto, quando, in che modo e perché (validità di contenuto). Oltre a questo allineamento sostanziale, il percorso di apprendimento continuo richiede anche un aggiustamento rispetto al livello cognitivo in cui la materia deve essere trattata in rapporto con **RTTI** (validità di costruito). Ciò impedisce che lo scarto cognitivo per gli studenti sia troppo grande (cosa che li demotiverebbe) o che il risultato medio di una classe diminuisca improvvisamente. Per tale ragione si discutono, per ogni periodo dell'anno, gli obiettivi di apprendimento/gli argomenti che, nell'ottica del modello **RTTI**, saranno insegnanti e testati.

Mappatura a ritroso Il primo passo nell'Analisi della Qualità di **RTTI** online è di controllare l'allineamento del test con il processo di apprendimento continuo. Quando si evidenzia una deviazione consistente tra il modello di **RTTI** nei passaggi e il modello ottimale, i passaggi sono progressivamente aggiustati. Gli insegnanti hanno bisogno di tempo e delle opportunità per questi aggiustamenti: non possono cambiare radicalmente la sequenza

del corso corrispondente al rapporto **RTTI** richiesto in una sola volta. Gli studenti devono essere in grado di abituarsi al nuovo modo di definire il problema e il curriculum e devono avere le chance per acquisire le strategie necessarie. Rendendo gli obiettivi di apprendimento e i corrispondenti test trasparenti e concreti prima di ciascun ciclo di lezione, il focus per tutti gli insegnanti è posto sugli obiettivi di apprendimento da raggiungere e le lezioni diventano più efficaci ed efficienti (Davies *et al.*, 2016). È importante per la motivazione degli studenti che sia loro chiaro quali siano gli obiettivi di apprendimento del ciclo di lezioni (Laveault, Allal, 2016) e che gli insegnanti allineino le valutazioni e i compiti a casa con quegli obiettivi di apprendimento. In pratica, accade che i compiti a casa coprano il 90% delle domande relative alla categoria T₁, mentre l'obiettivo di apprendimento per T₁ è solo del 40%. E questa potrebbe essere una delle possibili cause per spiegare la bassa motivazione degli studenti a svolgere i compiti a casa.

Step 2. Monitorare la qualità delle domande del test

Il secondo step nell'analisi della qualità è il monitoraggio della qualità delle domande del test. Per analizzare la qualità delle domande del test, si esamina il valore determinante di una domanda spesso attraverso i valori di Rit³ e Rir⁴. Per molti insegnanti, è un po' difficile calcolare e interpretare questi valori. Il p-value, la percentuale di studenti che hanno risposto correttamente a una domanda, è spesso utilizzato per giudicare la qualità delle domande del test. Quando un insegnante lavora con **RTTI** può soltanto usare (sebbene sia poco affidabile) i p-value come valida alternativa per monitorare la qualità delle domande di un test. Ci si aspetta che i p-value delle domande di T₂ e di I, per esempio, in una classe prima di una scuola secondaria siano più bassi rispetto a quelli delle domande per R e T₂. In fin di conti, di solito ci sono pochissimi studenti in una classe che risponderanno alle domande di T₂ e I in maniera appropriata, a differenza di quanto accade per le domande di R e T₁. Similmente, in una classe quinta di una scuola secondaria ci si aspetta che siano molti di più gli studenti che rispondano in modo corretto alle domande di T₂ e I rispetto al grado primario; pertanto ci si aspetta che i p-values[elim] per queste domande siano quantomeno più alti. Per garantire la qualità delle domande del test, **RTTI-online** segnala automaticamente quando una domanda codificata in **RTTI** eccede rispetto al valore atteso per quel tipo di percorso didattico. Tutte le

3. Rit value indica la performance di un *item* rispetto all'intero test.

4. Rir value indica la performance di un *item* rispetto al test meno il punteggio dell'*item*.

domande che vanno fuori dal valore atteso sono mostrate all'insegnante per una successiva analisi:

- ci potrebbe essere un errore nella costruzione, in qual caso la domanda necessita di essere controllata per essere rimossa o migliorata prima di inserirla in un altro test;
- la causa potrebbe anche essere una codifica errata di **RTTI**: la domanda potrebbe riferirsi a un livello cognitivo differente, di conseguenza la codifica va aggiustata per una corretta diagnosi;
- la causa potrebbe essere didattica: l'obiettivo di apprendimento non si aggancia con la disposizione didattica offerta. L'insegnante può ricorrere a interventi didattici a livello di classe per portare gli studenti al livello desiderato. I colleghi di classi che ottengono performance di livello migliore possono anche essere coinvolti in modo da consentire lo scambio e la condivisione di competenza;

Gli insegnanti controllano e assicurano in modo autonomo la qualità dei loro test e possono persino decidere di rimuovere domande e/o di migliorarle per un uso futuro. L'insegnante può anche redigere programmi di recupero per l'organizzazione dell'azione didattica e dell'apprendimento.

Step 3. Interventi didattici

Nel passaggio 3 dell'analisi della qualità si propongo[no] e si rendono visibili gli interventi didattici funzionali per il livello cognitivo più basso della classe. Per prima cosa, si possono trasferire a tutta la classe raccomandazioni per strategie di apprendimento efficaci, per esempio, attraverso l'esempio o l'imitazione o attraverso l'uso esplicito delle strategie di apprendimento fornite quando si discutono i compiti a casa. In seconda battuta, l'applicazione fornisce raccomandazioni concrete per l'uso di ruoli alternativi dell'insegnante che può, inoltre, scegliere direttamente fra più di 100 metodi didattici da utilizzare in varie fasi della lezione. Questo passaggio **supporta[assiste]** l'insegnante nell'analizzare gli errori che possono frequentemente manifestarsi al fine di progettare nuove attività che aiutino gli studenti a progredire nella loro comprensione dei concetti chiave scientifici studiati (Laveault, Allal, 2016, p. 9).

Step 4. Interventi didattici

Nel passaggio 4, il modello di **RTTI** di ciascuno studente è mostrato, ciò fornisce **insight[una comprensione?]** rispetto al processo di apprendimento che c'è dietro i loro voti. Questo modello può essere confrontato dall'insegnante con la media dello studente nel suo modello; si tratta di una ricerca idiografica. L'insegnante fornisce una descrizione (narrativa) e la ra-

gione possibile del modello di **RTTI**. «Il feedback è più efficace quando è restituito nella giusta modalità (ad esempio, dettagliato, narrativo, non attraverso voti)» (Andrade, 2013, p. 25). Le strategie di apprendimento più promettenti che incontrano il modello **RTTI** sono rese evidenti così che l'insegnante non solo riceve un'analisi e una diagnosi, ma anche suggerimenti per azioni direttamente mirate all'uso immediato da parte degli studenti. L'insegnante può anche vedere cosa lo studente ha fatto da sé nel suo account di **RTTI**-online e offrire azione di **supporto[ausilio?]** se necessario. Infine, gli insegnanti e i loro colleghi possono discutere dei casi inseriti per i quali non hanno ancora iniziato a implementare azioni e supportarle attraverso la competenza professionale che può essere disseminata all'intero gruppo docenti per ulteriori elaborazioni durante la discussione con gli studenti.

*Il ciclo della qualità di **RTTI*** Il ciclo della qualità di **RTTI** supporta la struttura di una valutazione per l'apprendimento nel contesto scuola in accordo con la raccomandazione di D. William e S. Leahy (2018) di creare opportunità dentro la scuola per promuovere dialogo, per lavorare su un curriculum coerente in cui istruzione, apprendimento e pratiche di testing siano connessi in un processo continuo.

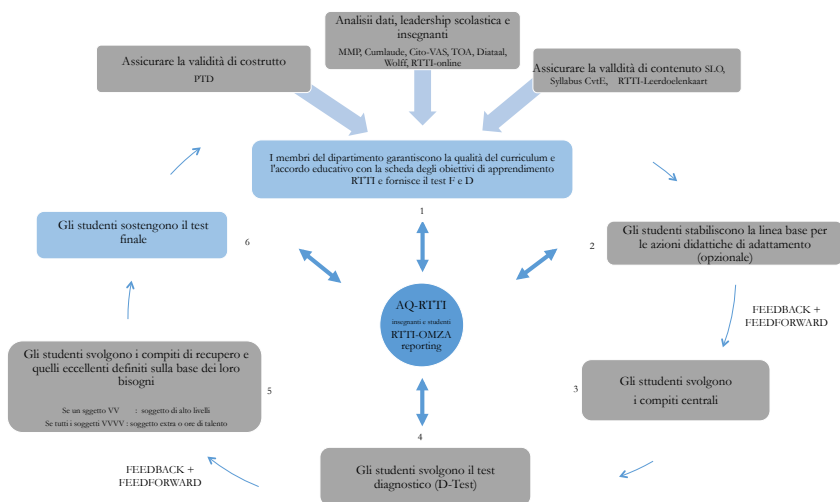
Il ciclo della qualità di **RTTI** si sviluppa sulla base del numero di cicli di insegnamento in cui è organizzato l'anno scolastico in una scuola; numerose scuole lavorano su 3-4 cicli per anno. Analizzando un test **RTTI** con i relativi voti (costruito secondo la relazione ottimale **RTTI**) in ciascun ciclo e utilizzando i 4 punti del controllo qualità, si può ottenere una panoramica valida e affidabile del processo di apprendimento di ciascuno studente per tutte le discipline. Il ciclo della qualità di **RTTI** è facilitato da diversi eventi specifici e attività. Quelle di seguito presentate seguono la numerazione riportata nella figura del ciclo di qualità di **RTTI** (FIG. 7.1).

1. *Compiti degli insegnanti* Gli insegnanti forniscono test di tipo *high[c.vo]* (V-test) e di tipo diagnostico (D-test) che saranno ripresi nel successivo ciclo di insegnamento e produrranno dei metodi di studio **RTTI** e ne controllano la qualità. D. William e S. Leahy (2018) raccomandano di realizzare subito il test finale, in modo da avere chiare le conoscenze e le abilità e sapere cosa si intende che gli studenti siano in grado di fare alla fine del modulo o della serie di lezioni. È preferibile considerare i test simultaneamente se è previsto un test o una valutazione congiunta.

2. *Misurazione zero* Durante la prima lezione del ciclo gli studenti stabiliscono, fondamentalmente, una linea di base per la misurazione. Gli insegnanti possono usare il test diagnostico (D) oppure usare una serie di compiti a casa o

FIGURA 7.1

Ciclo della qualità di RTTI [fornire jpeg 300dpi]



di esercitazioni relative all'obiettivo di apprendimento. Sulla scorta del risultato, gli studenti svolgeranno i compiti fondamentali per quegli aspetti che non padroneggeranno ancora, e svolgeranno compiti eccellenti per quegli aspetti che invece già padroneggiano. Con un'azione didattica di adattamento, gli studenti che hanno davvero padroneggiato a pieno l'argomento possono seguire il corso al livello superiore o seguire un programma accelerato. L'insegnante può anche ricorrere alla linea di base della misurazione per accomodare gli studenti che sembrano acquisire la disciplina per la prima volta attraverso le esercitazioni principali e svolgere in seguito un test diagnostico.

3. *Compiti fondamentali* Si tratta di una selezione limitata di compiti nell'ambito dell'intero curriculum, composto dal programma minimo con le componenti cruciali e rilevanti necessarie per raggiungere l'obiettivo di apprendimento. Le scuole olandesi **L.O.O.T.[LOOT][sciogliere]**⁵ hanno spesso sviluppato alcuni programmi chiave perché hanno ridotto il curriculum agli aspetti fondanti per il percorso formativo dei loro migliori atleti. Per gli studenti che seguono il programma regolare si tratta di un modo efficace per la-

5. Si tratta di scuole di secondo grado dove atleti particolarmente dotati hanno l'opportunità di combinare il loro migliore sviluppo sportivo con l'istruzione.

TABELLA 7.3
 Scheda degli obiettivi di apprendimento **RTTI**

| | Periodo: | | | Peso (Finale) test: |
|--|---|--|---|---|
| | R | T ₁ | T ₂ | |
| 1. Lettura di testi di economia 3 [°] ginnasio | | | | |
| RTTI -ratio | 5% | 55% | 35% | 5% |
| 1.1g. Puoi sintetizza- re dei testi | Quadro concettuale; a.o. Aspetti principali Aspetti collaterali Sintetizza su indicazioni Sintetizza in modo infor- mativo Sintetizza in modo schema- tico | Puoi distinguere gli aspetti principali e secondari in un testo e riconosci diversi tipi di sintesi. | Puoi connettere gli aspetti principali e secondari in un testo in modo logico | Puoi fare una sintesi ade- guata di un testo per gli altri che si adatta all'obiet- tivo comunicativo che deve essere raggiunto |
| Testi dietro il link | | Gli aspetti chiave sono parte dei testi realmente necessari per comprendere la storia. Gli aspetti collate- rali... ecc. | https://www.youtube.com/watch?v=(gKpe-com/watch?y=4F81-tWCfjw | https://www.youtube.com/watch?v=8647-bGT43w |
| Supporto | Overhearing[c.vo]* | Signalwords[c.vo] | 25-words- summary questions[c.vo] | 5W- OV-SLIM |
| Compiti centrali | B5-1a** | B5-1be, 2abc, 3, 8 | B5-4, 5c, 6, 7b | B5-5ab, 7a |
| Compiti di recupero | B5-R | B5-T1 | B5-T2 | B5-I |
| Compiti eccellenti | D5-1abc, 2° | D5-1de, 2bc, 4, 9 | D5-5, 3, 6c, 7, 8b | D5-6ab, 8a, 9 |
| Testitems[c.vo] data- base | 4, 7 | 1-3, 8-12, 15-17, 22 | 5, 6, 13, 19-21 | 14, 18 |

sciare più tempo alla personalizzazione (recupero), per approfondire una materia, o una materia diversa, o per essere capace di accelerare l'apprendimento.

4. *Chiave diagnostica (chiave D)* Approssimativamente due settimane prima del test sui progressi ottenuti (V-test) si svolge un test diagnostico (D), in linea con gli obiettivi di apprendimento e **RTTI**, così che lo studente, oltre allo specifico controllo sulla materia, possa ottenere informazioni sulla comprensione dei differenti livelli cognitivi. In questo modo i riscontri ottenuti dal test V potranno essere usati per la discussione. Tutto ciò può anche rappresentare parte del compito a casa che è stato valutato dall'insegnante o dallo studente. L'insegnante può anche considerare questo test come uno "strumento" se è necessario un ulteriore voto. La valutazione di un test-D non è sempre fatta dall'insegnante.[.] **L[I]**o stesso studente o un compagno può farlo. Quando si usa la valutazione per l'apprendimento di un test diagnostico gli studenti possono intraprendere azioni specifiche per prepararsi al test reale di tipo V.

– 5. *Compiti correttivi di livello eccellente* Lo studente (dopo la valutazione per l'apprendimento del test diagnostico nel formato online **RTTI**) conosce anche **i[le] suoi[sue]** possibili **gap[carenze]** e aree di forza. Nel caso di lacune, lo studente svolgerà dei compiti di recupero, come indicato nella scheda degli obiettivi di apprendimento di **RTTI** o nella guida. Per livelli cognitivi più forti, lo studente può cominciare con compiti di eccellenza. Se si dominano già tutti i livelli cognitivi lo studente dovrebbe terminare il corso al livello più alto e cominciare a lavorare su compiti corrispondenti al livello più alto della scheda degli obiettivi di apprendimento (TAB. 7.3).

6. *Test di progresso (V-test)* Infine si somministra il test sul progresso (V). Dopo il controllo dei test, gli studenti possono controllare i punteggi nel programma **RTTI[-]**online e, con l'aiuto del processo di analisi della qualità, ottenere riscontri rispetto ai loro risultati e al possibile approccio. «La cosa più importante è che aiutiamo gli studenti ad imparare senza di noi» (William, Leahy, 2018, p. 182). Sulla base di quella analisi, gli stessi studenti indicano quali azioni useranno per ottenere migliori risultati nell'immediato futuro. Queste azioni degli studenti sono combinate con le azioni possibili dell'insegnante rispetto all'analisi della qualità. In seguito prenderà avvio il ciclo di qualità successivo.

Una valutazione per l'apprendimento supportata da **RTTI**-online rende possibile per gli insegnanti di portare avanti la ricerca valutativa, migliorare l'istruzione, generare rinnovamento per le loro pratiche didattiche attraverso l'apprendimento al lavoro e allo stesso tempo fornire un **feedback** mirato e orientato allo sviluppo dei loro studenti.

7.16

RTTI e lo sviluppo della competenza nella valutazione per l'apprendimento

L'approccio **RTTI** dimostra come ciascuna delle sfide di base per gli insegnanti indicate nella sezione introduttiva di questo capitolo si possano affrontare attraverso un processo che potrebbe essere descritto come “una forma di schaffolded learning” **supportato[sostenuto]** dall'elaborazione e strutturazione di un pacchetto multimediale per il testing e l'insegnamento. Questo si evidenzia nel fornire informazioni di **background**, strutture, suggerimenti e diverse esemplificazioni pratiche nelle seguenti aree:

- costruzione dei test;
- interpretazione diagnostica dei risultati dei test;
- supporto nel fornire il **feedback** e nel pianificare l'azione didattica;
- implementazione di controlli sui nuovi processi di insegnamento e apprendimento.

Inoltre, una formazione di base su questi aspetti della valutazione a supporto dell'apprendimento è fornita nella forma di e-learning e il Manuale di **RTTI** per la competenza su **RTTI** e **OPSA**, allineamento curricolare, concetti di base del testing nel campo dell'istruzione con liste di controllo per la valutazione per l'apprendimento e interventi da realizzare in classe;

L'applicazione web con **RTTI-online** che **supporta[aiuta]** periodicamente gli insegnanti per un miglioramento continuo nella loro valutazione a sostegno dell'apprendimento con importanti elementi rispetto al monitoraggio dell'allineamento curricolare, la qualità delle domande dei test e gli interventi didattici e formativi.

Riferimenti bibliografici

- ANDERSON L. W., KRATHWOHL D. R. (2001), *A Taxonomy for Learning, Teaching and Assessing. A Revision of Bloom's Taxonomy of Educational Objectives*, Addison Wesley Longman, New York, NY.
- ANDRADE H. L. (2013), *Classroom assessment in the context of learning theory and research*, in J. McMillan, [**SAGE?**] *Handbook of Research on Classroom Assessment*, SAGE, New York, NY, pp. 17-34.
- BANGERT-DROWNS R. L., ~~KULIK C. C., KULIK J. A., MORGAN M. T.~~ [*et al.*] (1991), *The Instructional Effect of Feedback in Test-Like Events*, in “Review of Educational Research”, 61, pp. 213-23.

- BLACK P., WILLIAM D. (1998), *Assessment and Classroom Learning*, in “Assessment in Education: Principles, Policy & Practice”, 5, pp. 7-74.
- BLOCK A. DE (1975). *Taxonomie van leerdoelen*, Scriptoria, Antwerpen.
- BLOOM B. S. (ed.) (1956), *Taxonomy of Educational Objectives: The Classification of Educational Goals*, David McKay, New York, **NY**.
- BURNS M. K., SYMINGTON T. (2002), *A Meta-Analysis of Prereferral Intervention Teams: Student and Systemic Outcomes*, in “Journal of School Psychology”, 40, pp. 437-47.
- DATNOW A., PARK V., KENNEDY-LEWIS B. (2012), *High School Teachers’ Use of Data to Inform Instruction*, in “Journal of Education for Students Placed at Risk”, 17 (4), pp. 247-65.
- DAVIES A., HERBST S., SHERMAN A. (2016), *Assessment for Learning: A Framework for Educator’s Professional Growth and Evaluation Cycles*, in D. Laveault, L. Allall (eds.), *Assessment for Learning: Meeting the Challenge of Implementation*, Springer Nature.
- DROST M., VERRA P. (2108[2018]a), *Handboek rtti*, Uitgeverijplus, Bodegraven.
- DROST, M., VERRA P.[ID.] (2108[2018]b), *Slimmer Leren*, Uitgeverijplus, Bodegraven.
- ERAUT M. (1994), *Developing Professional Knowhow and Competence*, Falmer Press, London.
- FABER J. M., VISSCHER A. J. (2014), *Digitale leerlingvolgsystemen: een review van de effecten op leerprestaties*, University of Twente, Enschede.
- FOLMER E., KOOPMANS-VAN NOOREL A., KUIPER W. (2017), *Curriculum spiegel 2017*, **SLO** – nationaal expertisecentrum leerplanontwikkeling, Enschede.
- FUCHS L. S., FUCHS D. (1986), *Effects of Systemic Formative Evaluation: A Meta-Analysis*, in “Exceptional Children”, 53 (3), pp. 199-208.
- HATTIE J. (2009), *Visible Learning: A Synthesis of Over 800 Meta-analyses Relating to Achievement*, Routledge, London.
- HATTIE J., TIMPERLEY H. (2007), *The Power of Feedback*, in “Review of Educational Research”, 77, pp. 81-112.
- HENDRIKS M. A., SCHEERENS J., SLEEGERS P. J. C. (2014), *Effects of Evaluation and Assessment on Student Achievement: A Review and Meta-analysis*, University of Twente, Enschede.
- KIM S. E. (2005), *Effects of Implementing Performance Assessments on Student Learning: Meta-Analysis Using hlm*. (Unpublished Doctoral dissertation), The Pennsylvania State University, University Park, **PA**.
- KLUGER A. N., DENISI A. (1996), *The Effects of Feedback Interventions on Performance: A Historical Review, a Meta-Analysis, and a Preliminary Feedback Intervention Theory*, in “Psychological Bulletin”, 119, pp. 254-84.
- LAVEAULT D., ALLALL L. (eds.) (2016), *Assessment for Learning: Meeting the Challenge of Implementation*, Springer Nature, Switzerland.
- LYONS N. (ed.) (1998), *With Portfolio in Hand: Validating the New Teacher Professionalism*, Teachers College Press, New York and London.

- MINTZBERG H. (1979), *The Structuring of Organizations*, Prentice Hall, Englewood Cliffs
- NEVO D. (1995), *School-based Evaluation: A Dialogue for School Improvement*, Pergamon, Oxford.
- ONDERWIJSINSPECTIE (2015), *Informatiegebruik voor kwaliteitsverbetering*. url: rijksoverheid.nl/documenten/rapporten/2015/01/08/informatiegebruik-voor-kwaliteitsverbetering.
- ONSTENK J. (2004), *De school en de lerende organisatie; Denkend aan scholen zie ik lerende organisaties traag door oneindig papierland gaan*, meso magazine jrg. 24 (138).
- OPRINS E., ANDRIESSEN J. (2002), *Leerprocessen bij kennismanagement*. Uit: Kennismanagement - leerprocessen organiseren in werkomgevingen (Dankbaar, M (red), Oprins, E. (red), Andriessen, J., Hoek, F. van, and Tonnenam, J.), **CINOP** Centrum voor innovatie van opleidingen.
- SANDERS P. (2013), *Toetsen op school*, Cito, Arnhem.
- SCHEERENS J. (2016), *Educational Effectiveness and Ineffectiveness. A Critical Review of the Knowledge Base*, Springer, Dordrecht, Heidelberg, New York, London.
- SCHEERENS J.**[ID.] (2004), *The Evaluation Culture*, in “Studies in Educational Evaluation”, 30 (2), pp. 105-24.
- SCHILDKAMP K., KARBAUTZKI L., VANHOOF J. (2014), *Exploring Data Use Practices around Europe: Identifying Enablers and Barriers*, in “Studies in Educational Evaluation”, 42, pp. 15-24.
- SHUTE V. J. (2008), *Focus on Formative Feedback*, in “Review of Educational Research”, 78, pp. 153-89.
- SOUSA D. A. (2009), *How the Gifted Brain Learns*, **SAGE**, Thousand Oaks, **CA**.
- VISSCHER A. J., EHREN M. (2011), *De eenvoud en complexiteit van opbrengstgericht werken*, Universiteit Twente, Vakgroep Onderwijsorganisatie en -management, Enschede.
- WILIAM D. (2011), *What is Assessment for Learning?*, in “Studies in Educational Evaluation”, 37, pp. 3-14.
- WILIAM D., LEAHY S. (2018), *Formatieve assessment; integreren in de praktijk* (bewerkt door: P. Pijl, R. Adrian **and[e?]** I. de Bont), Bazalt Educatieve uitgaven, Rotterdam.