

PROGRAMMA EDUCATION

FGA WORKING PAPER

N. **9** (12/2008)

**La misurazione
del valore aggiunto nella scuola**

Roberto Ricci
Ricercatore INVALSI

© Fondazione Giovanni Agnelli, 2008

Le opinioni espresse in questo testo sono responsabilità dell'autore e non necessariamente riflettono quelle della Fondazione Giovanni Agnelli.

The opinions expressed in this paper are the sole responsibility of the author and do not necessarily reflect those of the Fondazione Giovanni Agnelli.

www.fondazione-agnelli.it

segreteria@fga.it

1. Introduzione¹

Recentemente anche in Italia si sta sviluppando un dibattito sulla misurazione degli esiti scolastici e sul loro utilizzo. La questione riveste un'importanza strategica non solo per il mondo della scuola in senso stretto, ma per l'intera società di cui il sistema scolastico è parte importante e vitale.

Al di fuori dei confini nazionali i problemi di come giungere a misurazioni soddisfacenti degli esiti della scuola hanno trovato ampio spazio di discussione già a partire dagli anni Ottanta del secolo scorso, dando origine a confronti, talvolta anche accesi, sia sotto il profilo più generale di opportunità e congruenza con le finalità proprie dei sistemi scolastici sia sotto l'aspetto più propriamente statistico-misuratorio. Mai come in questo ambito, i due piani sembrano intersecarsi, influenzandosi reciprocamente e rendendo, quindi, necessaria una consapevolezza generale di tutti i soggetti interessati alla conoscenza dei risultati che si producono all'interno delle scuole.

Nello sviluppo del presente lavoro saranno delineate le principali caratteristiche dei metodi di misurazione mediante il cosiddetto *valore aggiunto*, puntando agli aspetti essenziali che rendono possibile comprendere potenzialità e limiti del metodo. Senza indulgere sulle complesse metodologie statistiche che consentono di giungere ad una misurazione appropriata del valore aggiunto in ambito scolastico, sarà posta particolare attenzione alle assunzioni di fondo e metodologiche alla base di questo processo, cercando di chiarire ed evidenziare quali debbano essere le modalità di applicazione per evitare l'insorgere di fenomeni distorsivi e non desiderabili che, a diversi livelli, possono produrre difficoltà e tensioni interne ed esterne al sistema scolastico superiori ai benefici che si desidera ottenere. E' noto infatti in letteratura che la mancanza di trasparenza e l'inefficacia nella comunicazione verso tutti i portatori d'interesse di ogni fase del processo possono ingenerare comportamenti opportunistici di alcuni soggetti coinvolti nella misurazione e valutazione, vanificando in tutto o in parte i risultati ottenuti.

2. Il quadro di riferimento teorico per la misurazione dell'effetto scuola

La misurazione delle prestazioni di una scuola relativamente agli apprendimenti degli studenti richiede l'attenta valutazione di diversi aspetti che non possono essere ignorati se non si vuole correre il rischio di fornire informazioni ed indicazioni fuorvianti che,

¹ Le opinioni espresse sono da attribuirsi all'autore e non impegnano la responsabilità dell'Istituto di appartenenza dell'autore stesso (Istituto Nazionale per la Valutazione del Sistema Educativo di Istruzione e di Formazione - INVALSI)

certamente, produrrebbero effetti ancora più negativi dell'assenza di qualsiasi azione di misurazione dell'efficacia dei processi formativi.

Come sovente accade, ripercorrere le fasi storiche attraversate dai paesi che nel corso degli ultimi decenni hanno realizzato sistemi di misurazione degli apprendimenti consente di comprendere limiti e potenzialità delle metodologie scelte.

A partire dalla seconda metà degli anni Ottanta del secolo scorso nel mondo anglosassone, in particolare in Inghilterra, Galles e in alcuni stati degli Stati Uniti, sono stati avviati dei processi di riforma del sistema scolastico che hanno visto come elemento caratterizzante, tra gli altri, la pubblicizzazione dei risultati sugli apprendimenti conseguiti da ciascuna istituzione scolastica in un determinato intervallo di tempo.

In una prima fase la misurazione degli apprendimenti è stata effettuata confrontando i risultati medi ottenuti da ciascuna scuola in tutte o in alcune classi (*grade*) in prove standardizzate somministrate agli studenti. A seconda delle scelte effettuate, al termine di questa fase di misurazione ciascuna scuola viene posizionata in una graduatoria in ragione del risultato medio conseguito. Se si guarda all'esperienza dell'Inghilterra, la pubblicazione delle suddette graduatorie (*league tables*) ha posto e pone tuttora non pochi problemi, sia di carattere generale che di natura squisitamente metodologica. Limitandosi al secondo aspetto, si può comprendere abbastanza facilmente che se si utilizzano solo dei dati cosiddetti "grezzi", ovvero non depurati dall'effetto di fattori che poco o nulla dipendono dalla scuola oggetto di rilevazione, le valutazioni sulla prestazioni di un'unità scolastica forniscono principalmente informazioni circa le origini e le caratteristiche degli studenti piuttosto che sulla qualità della scuola. E' noto, infatti, che i risultati conseguiti da un allievo in una prova sono fortemente influenzati dalle sue caratteristiche, sia personali che socio-culturali. Pertanto, stilare una graduatoria dei risultati medi delle scuole sulla base dei semplici valori medi grezzi significherebbe fornire informazioni scarse e poco rilevanti sull'efficacia dell'azione formativa ed educativa esercitata dalle singole istituzioni scolastiche, mancando quindi l'obiettivo principale che solitamente ci si pone quando si decide di misurare gli esiti prodotti da un sistema scolastico sui livelli di apprendimento.

Le suddette considerazioni richiedono quindi una diversa definizione dell'*effetto scuola* in modo che sia possibile una sua operazionalizzazione più appropriata ed informativa. A tal riguardo, diverse sono le proposte che si trovano in letteratura, ma quella formulata da Grisay (1999) riprendendo Bosk e Witziers (1995) sembra essere la maggiormente interessante. Secondo Grisay l'effetto scuola può essere misurato in quattro modi differenti: 1) con la differenza tra il punteggio medio grezzo di tale scuola e quello medio generale di un certo territorio, 2) mediante la misurazione dei progressi medi che gli allievi realizzano in un determinato arco di tempo (guadagni cognitivi), 3) per mezzo della differenza media tra i punteggi osservati ed i punteggi attesi in relazione alle caratteristiche degli allievi (condizione socio-economico-culturale, attitudini, ecc.), 4) mediante il guadagno cognitivo medio netto rispetto a tutti i fattori di contesto che non sono controllati dalla scuola.

La misurazione dell'effetto scuola sulla base della prima e, anche se in misura minore, della seconda definizione determina le distorsioni richiamate in precedenza. Come noto, esse producono una classificazione delle scuole fortemente influenzata da elementi che sfuggono alla sfera di azione delle istituzioni scolastiche. La terza operazionalizzazione del concetto di effetto scuola rappresenta l'accezione più diffusa di valore aggiunto all'interno del filone di ricerca sull'efficacia in campo scolastico. La quarta definizione si riferisce a una concezione di valore aggiunto ancora più articolata ed approfondita poiché si basa sui guadagni cognitivi netti degli allievi con riferimento a prove standardizzate somministrate ad intervalli di tempo predefiniti.

Sia che si intenda il valore aggiunto secondo la definizione più diffusa che in termini di guadagni cognitivi netti, risulta cruciale poter disporre di misure appropriate e statisticamente corrette sugli apprendimenti. Specie all'interno dei nostri confini nazionali, quest'ultimo aspetto non è stato sempre tenuto nella debita considerazione, basando quindi le analisi successive su misure deboli degli esiti delle prove cognitive proposte agli allievi. Negli ultimi anni la ricerca internazionale ritiene necessario costruire delle scale metriche, dette di Rasch, che siano in grado di misurare simultaneamente i risultati degli studenti e il livello di difficoltà di ciascuna domanda (Bond e Fox, 2007). Ciò è determinante poiché, specie se si ragiona in termini di guadagni cognitivi medi netti, è fondamentale disporre di risultati che siano confrontabili secondo una metrica appropriata e coerente con le finalità di misurazione adottate. Inoltre, l'utilizzo di queste metodologie ormai consolidate facilita l'individuazione della corrispondenza di ciascuna domanda con il livello di abilità o competenza che essa intende descrivere e misurare. Solo se si riserva particolare attenzione a questi aspetti preliminari, ma fondamentali, è possibile costruire un sistema di misurazioni che dia un sufficiente grado di affidabilità e, quindi, di reale utilizzabilità per costruire un sistema di valutazione delle scuole di una regione, di un paese o di qualsiasi ambito territoriale oggetto di studio.

3. La misurazione del valore aggiunto

La misurazione del valore aggiunto si realizza mediante procedure tecnico-statistiche piuttosto complesse la cui descrizione dettagliata esula dallo scopo di presentazione generale e non strettamente tecnica del presente lavoro. Tuttavia, ciò che maggiormente interessa in questa sede è la descrizione delle metodologie alle quali si fa riferimento in modo che risulti chiaro il loro significato ed il loro valore interpretativo, sia sotto il profilo delle potenzialità sia sotto quello dei limiti.

La misurazione del valore aggiunto può avvenire facendo riferimento a diverse aggregazioni di studenti, principalmente la classe o la scuola. Se dal punto di vista tecnico, almeno per gli aspetti generali, quest'ultima distinzione non riveste particolare rilevanza, sul piano teorico più generale la questione è tutt'altro che secondaria e, soprattutto, per

nulla pacifica nella letteratura internazionale. Alcuni eminenti ricercatori ritengono che la misurazione più appropriata del valore aggiunto debba essere effettuata a livello di classe poiché tale grandezza è determinata principalmente dall'effetto degli insegnanti, specie per la scuola di base, mentre, invece, è messa in discussione l'efficacia informativa del valore aggiunto a livello di scuola (Bressoux, 2004). Non mancano esempi nei quali si cerca di avvicinare le due posizioni mediante la costruzione di modelli statistici a tre livelli (studenti, classi, scuole) in grado di misurare sia un effetto classe che un effetto scuola (Ponisciack e Bryk, 2005).

Tipicamente la misura del valore aggiunto, sia esso a livello di scuola o di classe, si basa sull'utilizzo di modelli multilivello (*multilevel analysis*) che permettono di considerare la struttura annidata dei dati scolastici. Infatti le unità statistiche, gli studenti, sono raggruppati in classi, in scuole e così via. Ignorare questa natura dei dati vorrebbe dire perdere una parte considerevole dell'informazione e, soprattutto, significherebbe non poter distinguere tra variabili che spiegano il loro effetto a livelli diversi, ovvero a livello di studente, di classe, di scuola, ecc. (Raudenbusch e Bryk, 2002).

Per semplicità di esposizione si consideri dapprima un modello non particolarmente complesso in cui si desidera misurare il valore aggiunto a livello di scuola secondo la terza o la quarta definizione di effetto scuola proposta da Grisay (1999). In questo modo si introduce quindi un modello statistico descritto da variabili di due livelli: il livello studente ed il livello scuola. Mediante la predisposizione di questa tecnica di analisi si cerca di spiegare il rendimento in una determinata prova cognitiva (tipicamente una prova di uscita) in funzione di alcune variabili che operano a livello individuale e a livello di scuola. Nell'impostazione più semplice della determinazione del valore aggiunto, utile per comprenderne le caratteristiche fondamentali, si cerca di studiare l'effetto prodotto, in termini associativi e non causali, da alcune variabili ritenute cruciali. Se il modello di analisi è ben formulato, è quindi possibile introdurre gli elementi fondamentali che concorrono a spiegare i risultati individuali e di scuola che si osservano. Tutto ciò che non può essere spiegato in funzione delle variabili predette, ovvero i cosiddetti residui dei modelli di regressione multilivello, può essere considerato in un certo qual modo come rappresentativo del *quid* aggiunto da ciascuna scuola.

Se dal punto di vista meramente descrittivo, la misurazione del valore aggiunto sembra essere un'operazione abbastanza semplice, sotto il profilo tecnico-statistico essa pone diversi problemi la cui soluzione non è neutrale rispetto ai risultati ottenuti.

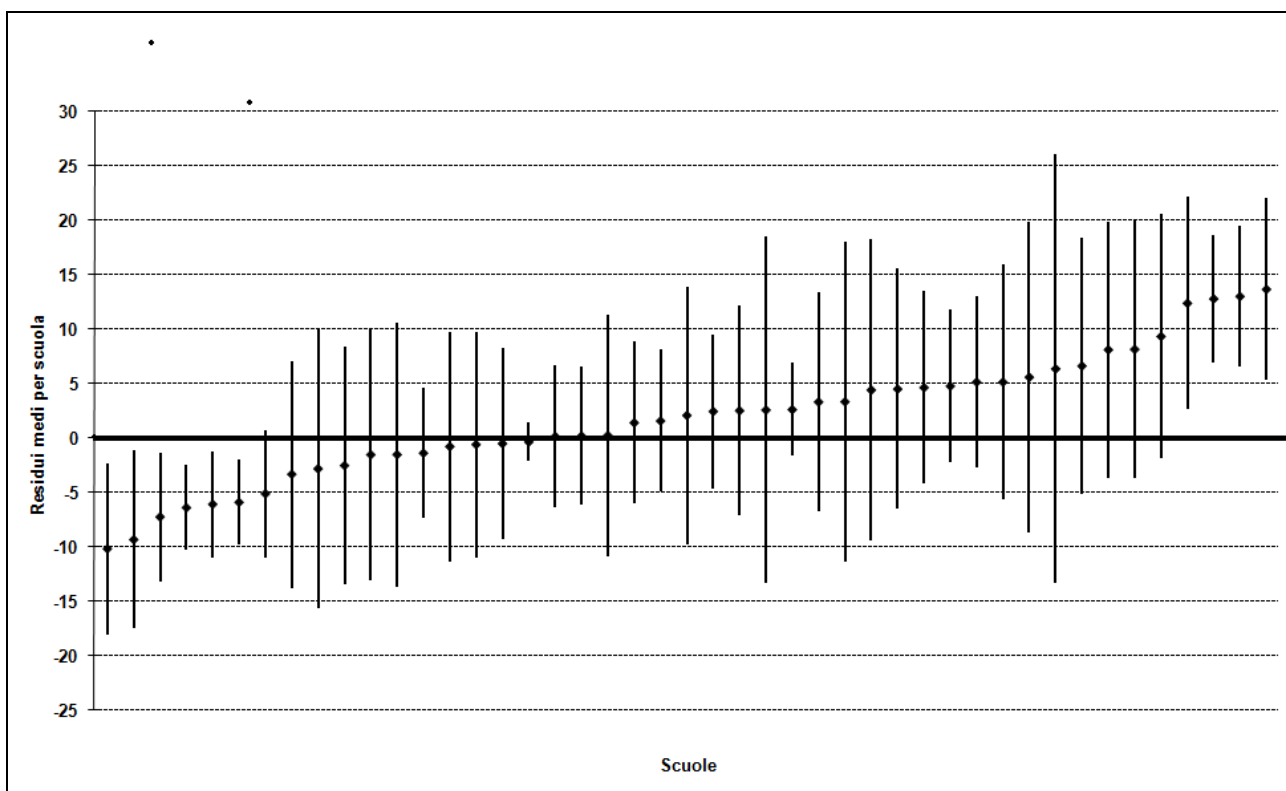


Figura 1. La misura del valore aggiunto

La figura 1 contiene la simulazione di un grafico tipico che si utilizza per rappresentare la misura del valore aggiunto di un gruppo di scuole. Al centro di ciascuna barra verticale si trova il valore medio dei residui di scuola al netto di tutte le variabili di contesto e, eventualmente, di preparazione iniziale. In altre parole il valore medio dei residui indica l'apporto specifico medio in termini di rendimento di ciascuna scuola al risultato dei suoi allievi. La barra verticale rappresenta l'intervallo di confidenza del suddetto valore, ovvero l'intervallo che con una probabilità solitamente del 95% comprende al suo interno il valore medio che si vuole stimare. Se i segmenti verticali intersecano la linea orizzontale corrispondente allo zero, ciò sta ad indicare che il valore medio osservato non è significativamente diverso da zero. Nel caso specifico questo significa che il valore aggiunto della scuola in esame è praticamente nullo rispetto alla media generale attesa.

La figura 1, come detto poc'anzi, rappresenta un caso tipico in cui solo una piccola minoranza di scuole si discosta significativamente dalla linea orizzontale: le prime sei da sinistra in senso negativo e le ultime quattro in senso positivo. Tuttavia ciò non deve erroneamente indurre a pensare che sia limitata l'informazione che si può trarre da un'analisi come quella rappresentata in figura 1. Essa permette di distinguere quali siano le scuole che si differenziano significativamente dalle altre. Inoltre, proprio alla luce di questa informazione è possibile immaginare un interessante avvicinamento tra ricerca qualitativa e quantitativa, sovente considerate come antitetiche, quando invece si possono proficuamente posizionare in un'utile ed interessante relazione di complementarità. Una

volta che la misurazione del valore aggiunto ha consentito di individuare le scuole con livelli medi di prestazioni degli allievi significativamente inferiori o superiori alla media generale, può essere molto interessante e produttivo dare corso ad un'indagine qualitativa approfondita che renda possibile l'individuazione degli elementi di debolezza delle prime e dei punti di forza delle seconde. Ciò potrebbe avere una ricaduta molto positiva sull'intero sistema scolastico, sia per ridurre le conseguenze di comportamenti scarsamente efficaci, sia per diffondere pratiche che producono effetti positivi sugli apprendimenti degli allievi.

Non da ultimo, la definizione del valore aggiunto riveste un'importanza notevole se si vuole mettere a confronto i risultati conseguiti da ciascuna scuola. La comparazione degli istituti in funzione del valore aggiunto significa confrontare le scuole sulla base di ciò che esse mediamente riescono ad aggiungere ai risultati medi dei loro allievi al netto dell'effetto di variabili non controllate dalla scuola stessa.

Se la finalità comparativa tra istituzioni scolastiche è ciò che interessa, allora riveste un'importanza cruciale il fatto che essa venga effettuata secondo modalità che garantiscano, allo stato attuale delle conoscenze scientifiche, risultati affidabili in grado di misurare effettivamente ciò che interessa e non elementi di contesto che scarsamente attengono alla sfera di influenza di ciascuna scuola.

La figura 2 mostra un caso emblematico che è opportuno tenere ben presente quando si auspica che la misurazione dei livelli di apprendimento rispetto ad una prova standardizzata si traduca anche, più o meno esplicitamente, nella formulazione di una classificazione delle singole istituzioni scolastiche (Goldstein, 2001). La figura 2 riporta i risultati medi di scuola nel test di matematica per la scuola primaria, livello KS2, somministrato in alcune scuole inglesi nel 1998. Il modello A della figura 2 rappresenta l'ordinamento delle scuole rispetto ai punteggi medi grezzi ottenuti nella prova suddetta². In particolare sono state contrassegnate due scuole, una con un cerchietto all'ultimo posto della graduatoria ed un'altra con un triangolo posizionata nei posti più elevati.

² Sull'asse verticale di tutti e tre i modelli si leggono gli scarti dalla media di riferimento e le barre verticali corrispondenti a ciascuna scuola rappresentano, come indicato in relazione alla figura 1, gli intervalli di confidenza.

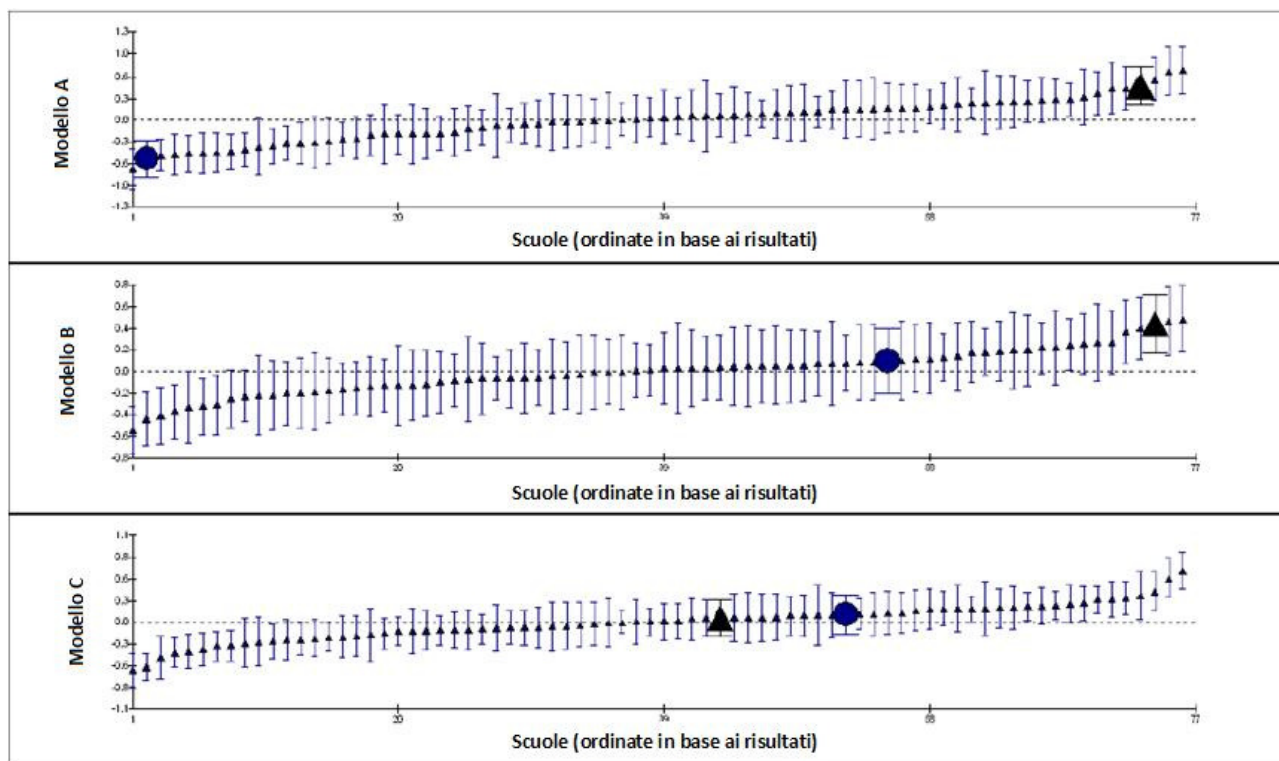


Figura 2. Il confronto dei risultati di scuola³

Il modello B rappresenta l'ordinamento dei risultati medi di scuola mediante l'aggiustamento in base ad alcune variabili di contesto (*covariate*) di carattere socio-economico⁴. Come si può notare facilmente, non solo la posizione della scuola contrassegnata con il cerchietto è sensibilmente migliorata, ma, poiché l'intervallo di confidenza associato alla sua media interseca la linea orizzontale corrispondente alla media generale, si può affermare che tale scuola non differisce in termini di risultati medi in modo significativo dalla maggior parte delle scuole esaminate, mentre secondo il modello A essa conseguiva risultati significativamente peggiori della media.

Il modello C rappresenta la graduazione (*ranking*) secondo una valutazione di valore aggiunto in base alla terza definizione fornita da Grisay (1999). Infatti, i risultati sono stati aggiustati non solo in base a covariate di natura socio-demografica, ma anche ad una misura del livello di apprendimento al termine della classe (*grade*) precedente, ovvero i risultati conseguiti nel test di matematica per il livello KS1. Anche in questo caso si producono dei cambiamenti piuttosto rilevanti e che devono essere considerati con particolare attenzione ed interesse. In primo luogo si nota che la posizione relativa delle due scuole prese ad esempio si capovolge e la scuola che in base ad una valutazione effettuata solo in funzione dei punteggi medi grezzi si collocava addirittura in ultima

³ Figura adattata da GOLDSTEIN H. (2001), *Using Pupil Performance Data for Judging Schools and Teachers: scope and limitations*, pag. 399.

⁴ La covariata principale introdotta è il diritto al "free school meals" che, solitamente, in Inghilterra si associa a condizioni socio-economiche svantaggiate.

posizione, ora si trova nella parte medio-alta della graduatoria davanti all'istituzione scolastica che secondo il modello A si trova ai vertici. Ma c'è ancora di più. In base al modello C entrambe le scuole conseguono dei risultati che non sono significativamente diversi dalla media generale.

L'esempio tratto ed adattato da Goldstein (2001) suggerisce diverse riflessioni, specie in una prospettiva come in quella attuale in cui anche in Italia sempre più si parla di sistemi di *accountability* nel quadro dell'autonomia delle scuole (Martini, 2007 e 2007). La misurazione del valore aggiunto ha rimarcato l'importanza di riservare particolare attenzione alle modalità secondo le quali le scuole vengono comparate rispetto ai risultati conseguiti. Chiaramente quest'ultimo aspetto non attiene solo la misurazione del valore aggiunto, ma poiché esso si basa sull'attenta individuazione degli elementi misuratori la conseguenza è stata che il tema ha assunto una rilevanza ancora maggiore.

Sempre in un'ottica comparativa dei risultati conseguiti dalle scuole alcuni autori rimarcano la necessità di prendere in considerazione l'effetto compositivo (*peer effect*) a livello di classe, ma anche di scuola. Esso consiste, in linea del tutto generale, nell'insieme degli effetti di interazione sugli apprendimenti degli studenti e sulle modalità di insegnamento che la composizione del gruppo genera in misura più o meno marcata. Dal punto di vista misuratorio l'effetto di composizione determina la violazione delle due assunzioni fondamentali alla base dei modelli per la misurazione della efficacia scolastica: 1) l'esistenza di un rapporto unidirezionale fra i processi ed i risultati, 2) l'indipendenza dei processi dalle variabili che descrivono le caratteristiche individuali e di gruppo degli studenti.

Per le suddette ragioni alcuni autori ritengono opportuno che, anche in termini di valore aggiunto, la comparazione tra scuole, classi e allievi debba tenere in adeguato conto l'effetto di composizione, confrontando quindi solo realtà caratterizzate da un certo grado di omogeneità. Ma anche quest'ultimo aspetto non è privo di conseguenze sotto il profilo metodologico, almeno allo stato attuale delle conoscenze statistiche. Se da un lato la comparazione tra realtà omogenee fornisce maggiori garanzie di effettuare confronti appropriati, dall'altro lato ciò si traduce giocoforza in una riduzione della variabilità dei dati. E' pertanto chiaro che, in base alla definizione precedentemente fornita di valore aggiunto, l'attenuazione del peso del *peer effect* determina anche una sottostima del reale effetto di scuola, ovvero del valore aggiunto. Come sovente accade, compito del ricercatore è quello di trovare un soddisfacente punto di mediazione tra aspetti contrastanti e, soprattutto, quello di rendere esplicite le scelte effettuate.

4. Potenzialità e limiti nell'uso del valore aggiunto

L'esposizione generale delle modalità secondo le quali si giunge alla determinazione del valore aggiunto dovrebbero aiutare a comprenderne le potenzialità ed i limiti.

Le analisi di valore aggiunto sono realmente informative solo a condizione che siano valide le misurazioni sulle quali esse si basano. Quest'ultimo è un aspetto molto importante, specie in una realtà come quella italiana che si è avvicinata solo di recente all'idea di introdurre queste metodologie per misurare l'efficacia delle scuole e che non può vantare una lunga tradizione consolidata e radicata nella predisposizione ed utilizzo di prove standardizzate.

Un altro aspetto metodologico molto rilevante che non riguarda in via esclusiva la misurazione del valore aggiunto, ma che ne permea il significato ed il valore informativo, è dato dalla natura associazionale e non causale dei modelli statistici utilizzati. In termini più espliciti, ciò significa che non è possibile individuare una relazione di causa ed effetto tra le variabili esplicative prescelte (*predittori*) e gli esiti, ma è necessario interpretare i risultati ottenuti come la forza con la quale questi predittori si associano agli esiti medesimi.

La comparazione dei risultati delle scuole mediante la misurazione del valore aggiunto permette di individuare, pur con tutte le cautele suddette, le scuole o le classi che conseguono risultati particolarmente positivi o preoccupanti, suggerendo quindi piste di approfondimenti sia quantitativi che qualitativi per comprendere meglio ciò che determina tali situazioni particolari.

Diversi sono i casi nei quali le misurazioni del valore aggiunto sono state introdotte in maniera stabile e su ampia scala. Certamente il caso più noto è quello dello stato del Tennessee (USA) e che prevede in alcuni casi anche la misurazione del valore aggiunto a livello classe, anche se i risultati non vengono resi pubblici (Sanders e Horn, 1995). Inoltre non mancano anche esperienze in Texas (USA) (Webster, 2005) ed in altri paesi, soprattutto anglosassoni, oltre al già citato caso inglese.

Le critiche recentemente avanzate al sistema del Tennessee sono molto interessanti poiché consentono, da un lato, di cogliere i punti sui quali a livello internazionale si sta dibattendo e, dall'altro, quali possono essere i rimedi per far fronte alle difficoltà emerse (Amrein-Beardsley, 2008). Si è infatti riscontrato che il sistema di incentivi e sanzioni conseguenti all'applicazione della misurazione del valore aggiunto hanno in alcuni casi generato comportamenti opportunistici nei dirigenti che in alcune situazioni sono intervenuti sulla composizione delle classi per ottenere delle valutazioni più convenienti per la scuola, a danno degli allievi più deboli per diverse ragioni, sia individuali che a causa dell'appartenenza a gruppi socialmente svantaggiati.

Il rischio dell'insorgere di comportamenti opportunistici può riguardare anche i docenti e non solo gli studenti che in un qualche modo possono essere indotti a far sì che i loro allievi conseguano dei risultati più elevati nelle prove cognitive finali (Jacob e Levitt, 2002).

La soluzione tecnica dei suddetti problemi passa attraverso metodi e tecniche la cui complessità esula ampiamente dagli scopi del presente lavoro, specie quando si devono introdurre delle metodologie per far fronte al grosso problema dei dati mancanti quando il

sistema della misurazione del valore aggiunto avviene in un'ottica longitudinale. Tuttavia, ciò che può essere importante per il caso italiano è conoscere il contesto operativo e metodologico di riferimento, almeno nelle sue linee generali in modo che prima che si intraprenda qualsiasi strada si possa tenere ben conto dei problemi in gioco e quali sono le questioni maggiormente controverse e delicate là dove tali sistemi sono stati applicati con risultati molto interessanti e suggestivi, ma non senza problemi, anche seri.

5. Alcune riflessioni conclusive

Gli elementi essenziali alla base della misurazione del valore aggiunto esposti nei paragrafi precedenti permettono di cogliere gli aspetti essenziali di questa metodologia ed i vantaggi che essa può determinare, senza però ignorare limiti e punti di attenzione.

Certamente uno dei motivi principali per i quali è stata adottata la misurazione dei risultati in termini di valore aggiunto è la possibilità di misurare in modo appropriato il progresso degli allievi. A questo riguardo è importante osservare che il valore aggiunto, specie se applicato a livello individuale, consente di impostare le rilevazioni e le misurazioni non solo secondo una prospettiva sincronica (*cross-sectional*), ma anche diacronica, ovvero longitudinale. In quest'ultima ipotesi, disponendo di prove cognitive valutate su scale comparabili⁵, è possibile tracciare curve di crescita di ogni singolo allievo giungendo quindi ad una valutazione molto informativa, non solo sotto il profilo sommativo, ma anche quello formativo (Gori, 2003).

La valutazione delle scuole e degli studenti mediante il valore aggiunto ha riscontrato, in genere, un certo favore di tutti gli attori che operano nei sistemi scolastici, poiché permette di tenere in adeguata considerazione i livelli di partenza degli allievi, le condizioni socio-economico-culturali in cui ogni istituzione scolastica si trova ad operare. Quest'ultimo aspetto è certamente molto importante ed utile, senza però sottovalutare i problemi e le difficoltà di vario genere che un sistema di misurazione di valore aggiunto porta con sé.

Se si pensa ad una strutturazione longitudinale di valutazione di valore aggiunto, non bisogna sottovalutare alcuni aspetti logistici che esercitano un impatto molto forte, sovente sottovalutato, sulla tenuta concettuale dei metodi statistico-misuratori adottati. In questo particolare caso è necessario prevedere un sistema di codifica dei dati che sia in grado di seguire l'allievo per un periodo di tempo molto lungo nel quale possono intervenire molte variazioni. Se ciò non avviene, si determina una crescita incontrollata di dati mancanti che vengono di fatto a ridurre notevolmente, se non addirittura in alcuni casi a vanificare, le capacità di stima dei modelli statistici prescelti (Amrein-Beardsley, 2008).

⁵ La costruzione di scale *ancorate* mediante i principi della Rasch Analysis consente di effettuare delle comparazioni appropriate sulla base di scale metriche omogenee (Bond e Fox, 2007).

Qualsiasi sia l'ottica nella quale si pensa di costruire un sistema di valore aggiunto, *cross-sectional* o longitudinale, a livello di studente, di classe (insegnante) o di scuola è necessario prevedere l'insorgere di comportamenti opportunistici di alcuni dei soggetti interessati a vario titolo al processo di misurazione, specie se ad esso è collegato un sistema di incentivi e penalità. I riflessi di questo fenomeno si manifestano su vari piani: metodologici e politici. Dal primo punto di vista esistono metodi che consentono di studiare e di individuare in buona parte sia i tentavi degli studenti di copiare le risposte sia dei docenti di favorire risultati migliori (Jacob e Levitt, 2002). Il secondo aspetto, invece, riguarda tutte le azioni che il decisore politico può intraprendere per evitare che questi fenomeni distorsivi si manifestino.

A conclusione di queste brevi considerazioni sulla misurazione del valore aggiunto è necessario rimarcare l'importanza strategica e fondamentale che riveste la cura del processo di informazione che deve accompagnare la costruzione di un impianto per molti aspetti complesso e delicato. Se in generale è importante per chi è valutato conoscere i metodi ed i fini per i quali ciò avviene, nel caso della misurazione dell'efficacia delle scuole e, di conseguenza, di tutto il sistema scolastico, ciò è a maggior ragione importante. Non bisogna dimenticare che in Italia si è cominciato a dibattere e ad operare su questi temi solo negli ultimi anni con forti resistenze e diffidenze. Solo cercando di rendere chiaro e trasparente ciò che in generale un sistema di valutazione e in particolare la misurazione del valore aggiunto possono dire e cosa non possono dire è pensabile che tutti i portatori d'interesse verso il sistema scolastico si possano avvicinare a queste metodologie senza preclusioni aprioristiche o aspettative sovradimensionate ed improprie.

RIFERIMENTI BIBLIOGRAFICI

AMREIN-BEARDSLEY, A. (2008), *Methodological Concerns About the Education Value-Added Assessment System*. Educational Researcher, 37, 2, 65-75.

BOND, T.G., FOX C.M. (2007), *Applying The Rasch Model, Fundamental Measurement in the human Sciences*, Second Edition, Lawrence Erlbaum Associates, Publishers.

BOSKER, R., WITZIERS, R. (1995). *School Effects: Problems solutions and a meta analysis*. Paper presented at the Eighth Annual International Congress for School Effectiveness and Improvement, CHN, Leeuwarden, The Netherlands, January.

BRESSOUX P. (2004), *Fomalisation et modélisation dans lee sciences sociales: un etude del la construction du jugement des enseignants*. Revue Française de Pédagogie, 148, 61-74.

GOLDSTEIN H. (2001), *Using Pupil Performance Data for Judging Schools and Teachers: scope and limitations*. British Educational Research Journal, 27, 4, 391-405.

GORI E. (2003), *Quali prospettive dalla ricerca sulla qualità e l'efficacia della scuola per la costruzione di sistemi di Accountability dell'istruzione*. Non profit, Maggioli Editore, 9, 2, 335-420.

GRISAY A. (1999), *Comment mesurer l'effet des systèmes scolaires sur les inégalités entre élèves ?* In La justice du système éducatif, edited by Meuret, Denis. De Boeck, Bruxelles.

JACOB, B.A., LEVITT S. (2002), *An Investigation of the Prevalence and Predictors of Teacher Cheating*, In Law and Economics Workshop (University of California, Berkeley), Anno 2002, Paper 4.

MARTINI A. (2007), *Accountability*, Voci della scuola VI, Tecnodid, Napoli

MARTINI A. (2007), *Autonomia "chiama" valutazione*, Rivista dell'Istruzione, Maggioli Editore, 23, 2, 35-41.

PONISCIAK, S.M., BRYK A.S. (2005), *Value Added Analysis of the Chicago Public Schools: An Application of Hierarchical Models*, In Value Added Models in Education: Theory and Applications edited by R.W. Lissitz, JAM Press, Maple Grove, Minnesota, USA.

RAUDENBUSH S.W., BRYK A.S., (2002), *Hierarchical Linear Models. Applications and Data Analysis Methods*, Second Edition, Sage Publications, Newbury Park, CA, USA

SANDERS, W.L., HORN, S.P. (1995), *The Tennessee value-added assessment system (TVAAS): mixed methodology in educational assessment*, in: A.J. Shinkfield and D.L. Stufflebeam (ed.), "Teacher evaluation: guide to effective practice", Kluwer Academic Publishers, Boston, pp. 337-376.

WEBSTER, W.J. (2005), *The Dallas School-Level Accountability Model: The Marriage of Status and Value-Added Approaches*, In *Value Added Models in Education: Theory and Applications* edited by R.W. Lissitz, JAM Press, Maple Grove, Minnesota, USA.