

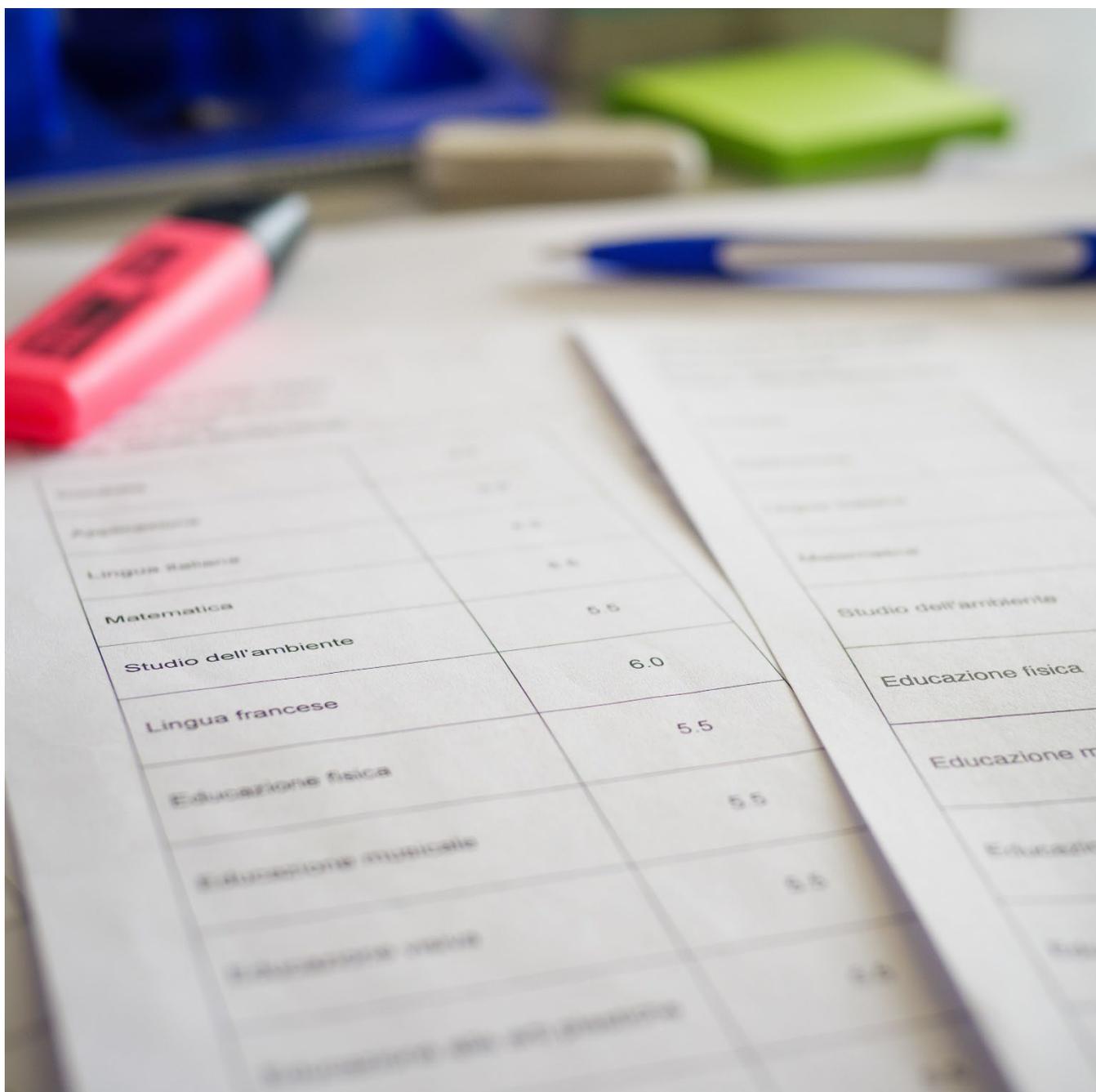
**SUPSI**

Quaderni di ricerca

## Che cosa c'è dietro a un voto?

Un'analisi multilivello applicata alle valutazioni di fine anno d'italiano e matematica nella scuola elementare

Giovanna Zanolla, Franck Petrucci, Emanuele Meier





# Che cosa c'è dietro a un voto?

Un'analisi multilivello applicata alle valutazioni di fine anno d'italiano e matematica nella scuola elementare

di Giovanna Zanolla, Franck Petrucci, Emanuele Meier

Proposta di citazione:

Zanolla, G., Petrucci, F. e Meier, E. (2023). *Che cosa c'è dietro a un voto? Un'analisi multilivello applicata alle valutazioni di fine anno d'italiano e matematica nella scuola elementare*. Centro competenze innovazione e ricerca sui sistemi educativi.

Locarno, 2024

CIRSE – Centro competenze innovazione e ricerca sui sistemi educativi

Piazza San Francesco 19, 6600 Locarno

[dfa.cirse@supsi.ch](mailto:dfa.cirse@supsi.ch)

ISBN (print) 978-88-85585-87-4

ISBN (online) 978-88-85585-89-8

Responsabilità del progetto: Giovanna Zanolla

Impaginazione e revisione: Elena Camerlo

Immagine di copertina: credits @Luca Ramelli

Il volume è distribuito con Licenza Creative Commons

*Attribuzione – Condividi allo stesso modo 4.0 Internazionale (CC BY-SA 4.0)*

*Quando in questo documento è usato il maschile, unicamente a scopo di semplificazione, la forma è da intendersi riferita in maniera inclusiva a tutte le persone, quale che sia la personale definizione di genere.*

# Ringraziamenti

Ai colleghi del CIRSE Alberto Crescentini e Marina Pettignano che, con Emanuele Meier, hanno messo a disposizione i dati relativi alle prove standardizzate cantonali di matematica e italiano nella scuola elementare da loro organizzate.

Allo staff del supporto GAS-GAGI (Gestione Amministrativa delle Scuole – Gestione Allievi e Gestione Istituti) del Centro di risorse didattiche e digitali del Dipartimento dell'educazione, della cultura e dello sport del Canton Ticino, che ci ha fornito i dati indispensabili per lo svolgimento di questo lavoro con la consueta disponibilità.

A Elena Camerlo, che ha curato l'impaginazione e la formattazione di questo rapporto.



# Abstract

La valutazione è una pratica educativa che influisce sulla vita di tantissimi giovani di tutti gli ordini scolastici. Nonostante assegnare un voto sia un atto di responsabilità dal quale può dipendere l'avvenire di un giovane, il dare un voto è spesso soggetto a un'elevata variabilità e soggettività, con il rischio di compromettere la credibilità della valutazione e l'equità dell'istruzione.

Dei modelli multilivello finalizzati a identificare le variabili individuali e attinenti alla classe e al docente che incidono sul voto d'italiano e quello di matematica all'ultimo anno della scuola elementare nell'anno scolastico 2020/21 e sul voto in matematica, sempre in quinta elementare, nell'anno scolastico 2014/15 hanno permesso di evidenziare che i voti dipendono in gran parte dall'abilità dell'allievo/a nella disciplina considerata (è stata riscontrata una certa coerenza tra i voti e i risultati nei test standardizzati). Tuttavia, ci sono fattori extra cognitivi come il genere dell'allievo/a, l'origine sociale, la madrelingua, l'età e l'abilità media della classe nella disciplina che, presi singolarmente non sembrano esercitare un grande effetto sui voti, ma complessivamente incidono in maniera consistente su di essi. In particolare, si è evidenziato come i fattori extra cognitivi arrivino a modificare il voto di fine anno fino a un punto, il che non è poco in un intervallo che normalmente nella scuola elementare è compreso tra 4.0 e 6.0.

Questo studio richiama l'attenzione sul tema dei cosiddetti effetti terziari, ovvero della possibile non neutralità degli insegnanti rispetto alle caratteristiche ascritte (origine sociale, sesso e trascorso migratorio) degli allievi e del ruolo della scuola nella riproduzione delle disuguaglianze di partenza. Nelle valutazioni, che incorporano appunto elementi cognitivi e non cognitivi e condizionano i destini scolastici e professionali dei giovani, tale assenza di neutralità si esplicita infatti molto bene.



# Sommario

Introduzione .....	1
1 Quadro teorico.....	3
1.1 La società meritocratica, gli “effetti terziari” e altri effetti collaterali .....	3
1.2 Di che cosa sono fatti i voti? .....	5
1.2.1 Bias generali della valutazione .....	6
1.2.2 Bias scolastici della valutazione .....	7
1.2.3 Bias sociali della valutazione .....	7
1.2.4 La contestualità delle pratiche valutative .....	9
1.3 Voti e punteggi nei test standardizzati .....	10
2 Obiettivi e metodi.....	13
3 Risultati.....	15
3.1 Le caratteristiche proprie dell’allievo/a e attinenti alla classe e al docente che incidono sul voto di fine anno .....	15
3.1.1 Informazioni generali sui modelli multilivello: la popolazione coinvolta .....	15
3.1.2 Risultati dei modelli vuoti .....	15
3.1.3 Risultati dei modelli definitivi.....	15
3.2 Punteggi nelle prove standardizzate e voti: quale relazione? .....	17
4 Conclusioni.....	21
5 Appendice .....	23
Riferimenti bibliografici .....	29



## Introduzione

La valutazione è una pratica educativa che influisce sulla vita di tantissimi giovani di tutti gli ordini scolastici. I voti ottemperano infatti a diverse funzioni, tra cui: dare agli studenti (e alle loro famiglie) un riscontro sul loro apprendimento e sul livello di conoscenza di una materia; motivarli nello studio e accrescere il loro impegno e la fiducia in loro stessi; condizionare le scelte educative permettendo l'ingresso ai curricoli caratterizzati da esigenze estese oppure limitando l'accesso a determinate formazioni post-obbligatorie; fornire ai potenziali datori di lavoro un indicatore delle capacità e delle qualità di un candidato (Argentin e Triventi, 2015). Nonostante sia un atto di responsabilità dal quale può dipendere l'avvenire di un giovane, il dare un voto è spesso soggetto a un'elevata variabilità e soggettività, con il rischio di compromettere la credibilità della valutazione e l'equità dell'istruzione (Benvenuto, 2003).

Questo progetto giunge dopo anni di analisi sui voti nelle scuole elementari (SE), medie (SM) e medie superiori (SMS) ticinesi, volte a confrontare la distribuzione dei voti (e le medie degli stessi) attribuiti dai docenti delle varie scuole, classi e materie a bambine e bambini, ragazze e ragazzi di provenienze sociali e geografiche diverse e ad analizzarne la tendenza da un ordine scolastico a quello successivo.

Dalle analisi dei dati GAS-GAGI (Gestione allievi e istituti) del Dipartimento dell'educazione, della cultura e dello sport, relative all'arco temporale compreso tra l'anno scolastico 2012/13 e 2021/22 per le scuole elementari e tra l'anno 2008/09 e 2021/22 per gli altri due ordini scolastici menzionati, è emerso un quadro complessivamente eterogeneo (le medie dei voti variano a dipendenza dei docenti e delle scuole e a cambiare è anche lo stile metrico degli insegnanti, ad esempio c'è chi non utilizza le code estreme dell'intervallo di punteggio e chi non scende mai sotto una data soglia) con alcune costanti. Tra queste ultime si annoverano:

- il fatto che le allieve e gli allievi di estrazione sociale avvantaggiata e di nazionalità svizzera riportino voti medi più elevati in quasi tutte le materie di tutti gli ordini scolastici analizzati;
- il fatto che le bambine/ragazze riportino voti medi più elevati in quasi tutte le materie. Fanno eccezione studio dell'ambiente e matematica alla SE, matematica e educazione fisica alla SM, matematica, fisica e educazione fisica al liceo, educazione fisica e scienze naturali alla Scuola cantonale di commercio;
- il peggioramento dei voti nel passaggio dalla SE alla SM e soprattutto dalla SM alle scuole medie superiori<sup>1</sup>. La transizione dalla SM al liceo comporta, infatti, un peggioramento del voto in italiano e storia per il 90% dei giovani e per l'80% dei casi per la matematica. Quanto alla transizione dalla SM alla Scuola cantonale di commercio, nei diversi anni si osservano percentuali di peggioramenti del voto pari all'80% in geografia e scienze naturali, intorno al 60-70% in matematica, tra il 50 e 60% in tedesco e francese e prossime al 50% in educazione fisica.

A partire dal 2022 sono stati avviati due progetti spin-off, ovvero derivanti da quello che incorporava le analisi sopra citate, uno dal taglio più quantitativo e uno più qualitativo.

Lo spin-off qualitativo, attualmente in corso, è volto ad approfondire, attraverso una serie di interviste, le concezioni che i docenti della SE e della SM hanno dei voti, i criteri con cui li attribuiscono, gli strumenti utilizzati per la verifica degli apprendimenti, il modo di comunicare i risultati delle prove alle allieve, agli allievi e alle rispettive famiglie e l'interpretazione dei voti da parte di questi.

Lo spin-off quantitativo, oggetto del presente rapporto, si pone gli obiettivi di:

---

<sup>1</sup> Va precisato che i voti non sono gli stessi in tutti gli ordini scolastici. Alla SE essi variano tra 3 e 6, alla SM tra 2 e 6 e alle SMS tra 1 e 6. I mezzi punti sono concessi per tutti i voti alla SE e alle SMS e solo per i voti da 4 a 6 alla SM (v. Regolamento delle scuole comunali del 3 luglio 1996, Regolamento della scuola media del 30 maggio 2018 e Regolamento delle scuole medie superiori del 15 giugno 2016).

- identificare le variabili individuali e attinenti alla classe e al docente incluse nella banca dati GAS-GAGI che incidono sul voto;
- capire quanto i voti di un certo ordine scolastico in determinate materie siano predittivi dei voti che si conseguono nello stesso livello scolastico l'anno seguente e dei voti riportati nella stessa materia nel livello scolastico successivo e di confrontare la capacità predittiva del voto con quella dei punteggi nelle prove standardizzate.

Le analisi presentate in questo rapporto si riferiscono alla sola SE e alle sole discipline della matematica e dell'italiano in quanto, per i motivi che vedremo nel paragrafo 1.3, si è reso necessario ai fini dell'elaborazione dei modelli statistici multilivello tenere conto dei punteggi nei test standardizzati nella disciplina di cui si studia il voto di fine anno e gli unici test di cui disponiamo al momento riguardano la SE e appunto la matematica e l'italiano.

L'intento di questo studio è anche di richiamare l'attenzione su un tema in auge nella sociologia degli anni Settanta del secolo scorso e che ora sotto il nome di "effetti terziari" sta conoscendo una fase di revival: quello della possibile non neutralità degli insegnanti rispetto alle caratteristiche ascritte (origine sociale, genere e background migratorio) delle allieve e degli allievi e del ruolo della scuola nella riproduzione delle disuguaglianze di partenza. Nelle valutazioni, che incorporano elementi cognitivi e non cognitivi e condizionano i destini scolastici e professionali dei giovani, tale assenza di neutralità si esplicita infatti molto bene (Argentin e Pavolini, 2020). Se è vero che nessuna economia avanzata si è finora dimostrata in grado di annullare lo svantaggio di partenza delle figlie e dei figli degli strati sociali meno favoriti e il massimo che il sistema educativo riesce a fare è di controbilanciare le potenti forze esterne alla scuola responsabili del divario educativo (Major e Machin, 2018), è bene che gli insegnanti siano consapevoli del loro ruolo nella potenziale riproduzione delle disparità e della loro possibilità di attenuare la disuguaglianza.

# 1 Quadro teorico

## 1.1 La società meritocratica, gli “effetti terziari” e altri effetti collaterali

Secondo la tesi dell'*increased merit selection* (Jonsson, 1992) le società contemporanee tendono verso un tipo di organizzazione sempre più meritocratico. Il merito è particolarmente valorizzato, perché visto come un criterio legittimo di differenziazione sociale. L'idea di garantire opportunità lavorative, posizioni di responsabilità e mobilità sociale esclusivamente ai *meritevoli*, indipendentemente dalle loro condizioni di provenienza, e di penalizzare i *non meritevoli*, esercita infatti, come sostiene Boarelli (2019), una comprensibile attrazione. In una società in cui le posizioni sociali sono correlate al livello d'istruzione degli individui, è la scuola l'istituzione cardine che attribuisce e segnala i meriti individuali (Tenret, 2014) e lo fa naturalmente attraverso i voti e i diplomi.

Il termine *meritocrazia* è stato utilizzato per la prima volta da un sociologo inglese, Alan Fox, il quale in un articolo del 1956 aveva messo in evidenza come la selezione basata sull'intelligenza e il talento fosse alla base di un'inaccettabile divisione sociale. Due anni dopo uscì il famoso racconto fantasciologico di Michael Young (1958) ambientato nel 2033 in un'Inghilterra in cui, dopo anni di battaglie, si è insediato un regime meritocratico. Sconfitto il nepotismo della vecchia società preindustriale, fondato su privilegi ascritti, nel nuovo sistema le cariche si assegnano solo in base al merito e alla competenza. La perfetta meritocrazia si traduce però ben presto in un incubo: ancor prima della nascita, basandosi su previsioni di tipo genetico, viene stimato il quoziente intellettivo di ciascun individuo e sulla base di questo si stabilisce che tipo di lavoro potranno svolgere le persone e in quale percorso scolastico andranno incanalate. Nata dunque con una connotazione negativa in quanto volta a giustificare le disuguaglianze sociali, la parola *meritocrazia* con il passare del tempo ha acquisito un'accezione positiva. In particolare, la trattazione neoliberista a cavallo tra il ventesimo e il ventunesimo secolo della formula originale di Michael Young (1958), che eguagliava il merito alla somma di talento e impegno, ha messo da parte il primo addendo, ereditato per via genetica, ambientale e familiare, e ha ridotto il merito a una sola questione di sforzo individuale, attribuendo così al singolo individuo la responsabilità dei propri successi e fallimenti. Solo in tempi recenti si è ripreso a concepire il fenomeno della meritocrazia in chiave critica e cioè come un'ideologia volta a legittimare le crescenti disparità sociali (Markovits, 2019; Piketty, 2019) e con effetti disgreganti sul tessuto sociale, dal momento in cui chi giunge a determinati traguardi accademici e professionali, tende a pensare che il proprio successo sia legato esclusivamente ai propri sforzi e ad attribuire il fallimento altrui allo scarso impegno e alla carente forza di volontà (Sandel, 2020). La meritocrazia invoca infatti una società 'attiva' e 'competitiva', non una società solidale (Benadusi e Giancola, 2021).

In un periodo in cui la meritocrazia è quanto di più prossimo a un'ideologia universale che oggi abbiamo (Woolridge, 2021) e in cui al tempo stesso si sente molto il bisogno di realizzare l'uguaglianza delle opportunità educative, diventa importante capire che cosa siano i voti, ovvero i mattoni della “fabbrica del merito”. Ciascun voto, infatti, unito agli altri, consente (o impedisce) di ottenere un certo titolo di studio, di proseguire in certi percorsi formativi e di raggiungere successivamente una determinata posizione sociale. In antitesi con la visione della meritocrazia sopra descritta, la teoria della riproduzione sociale evidenzia come il sistema scolastico spesso non riesca a individuare il potenziale di coloro che non sono portatori del linguaggio, della cultura e dei valori delle classi privilegiate e la scuola contribuisca così a legittimare le disparità di partenza (Bowles e Gintis, 2011). Se gran parte degli studi sulle disuguaglianze educative si è finora concentrata sull'associazione tra alcune variabili ascritte (in primis l'origine sociale) e i risultati scolastici degli studenti e sul divario dei diversi gruppi sociali nella probabilità di effettuare, a parità di risultati, una certa transizione scolastica, il ruolo della scuola e in particolare degli insegnanti nella riproduzione delle disuguaglianze è stato poco esplorato e spesso relegato dai sociologi in una sorta di

scatola nera. Si tratta, dunque, di prendere in considerazione anche i cosiddetti “effetti terziari”<sup>2</sup>, ovvero il ruolo dei docenti, il cui agire sarebbe orientato dalle caratteristiche sociodemografiche degli allievi, nel rafforzare le disuguaglianze di partenza. Ciò implica la considerazione dei processi, di cui gli insegnanti spesso sono inconsapevoli essendo essi “passivi portatori di interessi delle classi dominanti per la posizione sociale e il ruolo attribuito loro nei sistemi scolastici» (Argentin, 2021, p. 155), che tuttavia contribuirebbero a ostacolare la promozione dell'equità all'interno delle scuole. Si tratta di meccanismi, magari non così palesi come un tempo, “ma non per questo meno influenti sui destini scolastici (e quindi in seguito sociali) degli individui diversi per background sociale o altre caratteristiche ascritte” (Argentin, 2021, p. 135).

Come si producono dunque gli effetti terziari? Attraverso le valutazioni degli insegnanti e la loro influenza sulle decisioni scolastiche dei loro allievi. I voti e i consigli orientativi, infatti, sono un messaggio fondamentale per i genitori, in particolare per quelli con capitale culturale limitato e una più debole conoscenza del sistema educativo (Argentin e Pavolini, 2020). Uno studio di Helbig e Morar (2018) mostra come gli insegnanti valutino le abilità e il comportamento degli allievi di estrazione sociale elevata più favorevolmente in quanto li reputano più capaci e ambiziosi degli allievi di origine sociale più svantaggiata. Secondo lo stesso studio a rendere ulteriormente iniqua dal punto di vista sociale la valutazione è il fatto che i docenti considerino, a torto o a ragione, i genitori una risorsa cruciale per le carriere scolastiche dei loro figli e questo elemento li porta ad attribuire voti più alti a chi pensano benefici di maggiore sostegno a casa. La sotto/sopravalutazione degli studenti di basso/elevato background socioculturale da parte degli insegnanti rinforzerebbe la disuguaglianza doppiamente: in maniera diretta, attraverso l'impatto sulla riuscita degli allievi (incoraggiamenti e scoraggiamenti), e indirettamente, attraverso i messaggi non privi di *bias* forniti ai genitori (Argentin e Pavolini, 2020).

La riproduzione delle disuguaglianze di partenza non è l'unico effetto collaterale della “fabbrica del merito”. Secondo Duru-Bellat (2009) ce ne sono almeno altri due: quello di privare gli studenti di una risorsa fondamentale dell'apprendimento, la cooperazione, e il fatto che la valutazione sia destinata a trasformarsi in un marchio che accompagnerà lo studente lungo tutto la sua esperienza scolastica.

Per quanto riguarda la cooperazione, Butera (2006) ritiene che un sistema basato sul merito, di cui i voti sono una delle espressioni concrete, porti gli individui a un confronto sociale permanente, ovvero a valutare le proprie capacità utilizzando gli altri come termine di paragone, e generi competizione impedendo di trarre beneficio dalla cooperazione. La competizione ostacola l'apprendimento in quanto gli sforzi degli individui sono indirizzati sul fare meglio degli altri piuttosto che sull'imparare. Gli epidemiologi Wilkinson e Pickett (2019) sottolineano come, in maniera analoga, nella popolazione in età lavorativa l'ansia sociale, legata al timore di essere valutati per la propria posizione economica, lavorativa e sociale possa indurre effetti psicologici quali bassa stima di sé, insicurezza e depressione, ma anche l'“aumento dell'io” e il narcisismo. Non solo: quanto maggiore è l'ansia, tanto maggiore la riluttanza a interagire con altre persone e a partecipare alla vita associativa; più ampie le distanze sociali, più grande la diffidenza verso gli altri. Si comprende dunque perché una delle cause di spaccatura del tessuto sociale di questi anni sia proprio l'istruzione (Collier, 2018).

Quando i voti diventano una sorta di etichetta che sancisce il valore di una persona, essi smettono di avere una qualsiasi funzione formativa e i comportamenti che ne derivano possono essere o di disimpegno verso

---

<sup>2</sup> Essi si aggiungono agli effetti “primari” e “secondari” descritti da Boudon (1973) al fine di comprendere meglio i processi con cui si riproducono le disuguaglianze educative. Gli effetti primari si riferiscono a disuguaglianze nei risultati scolastici tra studenti di diversa classe sociale e sono legati ad esempio al fatto che i figli dei genitori più istruiti già dai primi anni d'età possiedono un vocabolario più ampio, fruiscono di maggiori consumi culturali nel tempo libero, beneficiano di maggiore e migliore supporto nello studio oltre che di un ambiente domestico più adeguato per quest'ultimo, delle più elevate aspettative dei genitori riguardo all'istruzione e del fatto che questi tendono ad esercitare maggiori pressioni su dirigenti e insegnanti affinché venga loro erogato un servizio migliore (Argentin, 2021; Calarco, 2018). Gli effetti secondari sono invece le differenze tra le classi sociali nelle probabilità di effettuare una data transizione scolastica per studenti con eguali risultati scolastici. In pratica, a parità di risultati scolastici, è probabile che i figli delle famiglie più acculturate optino per percorsi di studio più impegnativi e remunerativi (Argentin, 2021). Secondo alcuni autori c'è più variabilità tra un Paese e l'altro in fatto di effetti secondari che di effetti primari (Jackson e Jonsson, 2013). Allo stesso tempo si evidenzia una mancanza di attenzione agli effetti primari e terziari che sembra essere legata alla complessità dei meccanismi coinvolti.

le attività di studio o, al contrario, di perfezionismo con tendenza a voler strafare (Kendall e di Pietro, 1995). Curran e Hill (2019) hanno rilevato che il perfezionismo nelle sue tre forme (auto-orientato, ovvero riferito agli standard elevati che gli individui fissano per se stessi; orientato verso l'altro, cioè caratterizzato dalla tendenza a definire standard elevati per gli altri e a criticare le prestazioni altrui; socialmente prescritto, che porta a credere che le altre persone abbiano riguardo alle proprie prestazioni alte aspettative, che si teme di non riuscire a soddisfare vivendo così una costante preoccupazione per una propria mancanza di perfezione) è aumentato negli ultimi anni con pesanti ripercussioni sul benessere dei giovani. Tale tendenza secondo Curran e Hill è legata all'enfasi dell'individualismo a scapito del collettivismo, all'ideologia meritocratica che, come si è scritto, porta chi non raggiunge determinati traguardi scolastici e professionali ad essere etichettato come "non meritevole" e colpevole del proprio insuccesso e alle pressioni dei genitori, che oggi spesso si sentono corresponsabili dei successi e dei fallimenti dei figli e adottano uno stile ansioso e controllante. È importante sottolineare che lo stile genitoriale è a sua volta il riflesso del clima economico: quanto più la ricchezza è distribuita in maniera diseguale, più i genitori spingeranno i figli a studiare e a impegnarsi duramente per conseguire le credenziali scolastiche più prestigiose e avremo fenomeni quali quello delle "mamme tigri" e dei "genitori elicottero", quando invece nelle società più egualitarie sono esaltati valori come l'autonomia, la creatività o il senso civico (Doepke e Zilibotti, 2019).

Vi è poi un dibattito anche sugli effetti dei docenti con standard valutativi elevati, cioè quelli che tendono a fissare l'asticella in alto e ad assegnare voti elevati solo agli allievi davvero eccellenti, e degli insegnanti più indulgenti, che danno buoni voti anche agli allievi mediamente competenti. I primi nel lungo termine possono stimolare gli studenti già forti a sforzarsi di più e a migliorarsi ulteriormente, ma al tempo stesso possono scoraggiare inesorabilmente gli studenti più deboli (Betts e Grogger, 2003). Valutazioni molto generose, che caratterizzano il fenomeno dell'"inflazione dei voti", possono al contrario influenzare la percezione di autoefficacia degli individui e avere un effetto benefico sulla loro riuscita scolastica (Bandiera et al., 2015), ma anche portarli a credere di avere delle competenze che in realtà non possiedono e condurli in imprese per le quali non hanno sufficienti conoscenze e abilità, oltre a dissuaderli dal lavorare di più e dal pretendere di più da se stessi, nel momento in cui essi credono di aver già ottenuto ottimi risultati (Finefer-Rosenbluh e Levinson, 2015). Alcuni studi hanno evidenziato come soggetti con caratteristiche sociodemografiche diverse reagiscano in maniera eterogenea a standard valutativi elevati, ad esempio i maschi sarebbero più spronati a impegnarsi rispetto alle ragazze in quanto più sensibili agli incentivi a breve termine e meno dotati delle compagne di motivazione intrinseca (Fallan e Opstad, 2012; Vecchione et al., 2014). Gli studenti con background migratorio sarebbero secondo certi autori caratterizzati da una maggiore motivazione intrinseca e, analogamente alle ragazze, sarebbero dunque meno inclini ad aumentare il proprio impegno in presenza di una maggiore severità nella valutazione (Gillen-O'Neel et al., 2011); secondo altri sarebbero invece dotati di maggiore motivazione estrinseca rispetto agli autoctoni (Triventi, 2020).

## 1.2 Di che cosa sono fatti i voti?

Per quanto l'assegnazione di un voto sia un momento importante e delicato per l'impatto che esso ha, chi lo decide è un essere umano che, pur cercando di essere più obiettivo possibile, sarà inevitabilmente condizionato dalla propria esperienza personale, dalle proprie credenze, dal contesto in cui opera e relative dinamiche psicologiche, dai propri pregiudizi e da altri elementi di distorsione che ricadono in quelli che siamo soliti chiamare *bias valutativi*.

Numerosi autori hanno sottolineato il fatto che i voti non siano una misura attendibile delle prestazioni di un individuo e che i docenti, nel momento in cui li assegnano nelle pagelle, adottino criteri diversi (Randall ed Engelhard, 2009 e 2010). In particolare, c'è convergenza sul fatto che nell'attribuzione delle valutazioni gli insegnanti utilizzino una combinazione di elementi cognitivi e non cognitivi, che nella letteratura anglosassone viene chiamata "hodgepodge" (Brookhart, 1991), ovvero un miscuglio che incorpora l'abilità, l'impegno, il comportamento, le disposizioni e la progressione nel corso dell'anno scolastico. In particolare, il peso attribuito al comportamento varia tra gli insegnanti (Guskey, 2009) e questa eterogeneità nelle

pratiche valutative è ascrivibile secondo alcuni autori alla mancanza di una formazione in materia di valutazione (Stiggins, 2002). Guskey e Link (2019) hanno rilevato che gli insegnanti del Secondario II attribuiscono maggiore rilevanza alle verifiche e ai compiti per casa rispetto ai docenti della scuola media ed elementare. In quest'ultima i maestri sembrano dare, invece, maggiore peso alle valutazioni formative e all'osservazione degli allievi in classe, cosa che dipende anche dalle minori competenze linguistiche dei bambini della scuola primaria rispetto ai ragazzi della scuola secondaria. Da questi ultimi i docenti si aspettano l'assunzione della responsabilità del proprio apprendimento. Sempre in questo studio si stima che una quota del voto compresa tra il 10% e il 20% dipenda da fattori non cognitivi come la partecipazione in classe, le abitudini di lavoro, l'ordine, l'impegno e la puntualità nelle consegne. Il peso dei fattori non cognitivi sul voto finale non varierebbe in misura significativa secondo i diversi ordini scolastici né secondo l'anzianità dei docenti.

Il sociologo francese Merle (2018) ha individuato una serie di *bias* generali, scolastici e sociali ricorrenti nelle pratiche valutative, alcuni dei quali saranno richiamati nei prossimi paragrafi.

### 1.2.1 Bias generali della valutazione

La letteratura mette innanzitutto in evidenza come i docenti tendano ad adeguare il livello del loro insegnamento e le valutazioni al rendimento dei loro allievi in modo da mantenere di anno in anno approssimativamente la stessa distribuzione gaussiana dei voti (De Landsheere e De Landsheere, 1992). Questi ultimi sono dunque comparati, ovvero non dipendono solo dalla prestazione dello studente in termini assoluti, ma dalla sua prestazione in relazione a quella dei suoi compagni di classe (Leclercq, Nicaise e Demeuse, 2004). Ne deriva che chi frequenta una classe con un livello medio di competenza piuttosto alto, avrà una valutazione più penalizzante di chi è inserito in una classe debole, perché i docenti si trovano a dover riprodurre una distribuzione di voti che sia innanzitutto credibile all'esterno (Merle, 2018).

La tendenza alla distribuzione gaussiana dei voti attribuiti, denominata anche "effetto Posthumus" (Merle, 2018) è legata alla credenza che, come accade per molti fenomeni naturali, vi siano frequenze esigue nei valori estremi, ovvero per i risultati altamente positivi o decisamente negativi, e più elevate nei valori centrali, corrispondenti alle valutazioni mediane. Gli insegnanti, implicitamente, tenderebbero credere che solo una percentuale limitata di studenti possa ottenere valutazioni ottimali, considerandoli di fatto come soggetti appartenenti a una popolazione con caratteristiche distribuite casualmente e la scuola e l'insegnamento come limitatamente incidenti sulle potenzialità di base dei discenti (Domenici, 2001)<sup>3</sup>.

Collegata all'effetto Posthumus, c'è quella che Merle (2018) definisce "costante macabra", che impone una percentuale minima di allievi sotto la media della classe controbilanciata da una percentuale di allievi sopra la media, anche se gli allievi sotto la media sono in realtà pochi perché ci si trova in presenza di una classe buona. La costante macabra è un pregiudizio che prova l'esistenza di norme implicite nell'attribuzione dei voti che si impongono ai docenti e contribuiscono alla costruzione dell'insuccesso scolastico designando una percentuale di allievi minima con difficoltà.

La valutazione può risentire anche del fatto che la correzione di un compito segua una verifica molto buona o, al contrario, una piuttosto scarsa con cui viene fatto un confronto diretto in cui si enfatizzano le differenze (effetto contrasto) e, più in generale, dell'ordine in cui le verifiche vengono corrette. Spesso accade che il primo compito corretto sia quello che viene valutato con maggiore severità, perché l'unico termine di paragone è il compito ideale o quello che svolgerebbe il docente. I compiti successivi riceverebbero una valutazione più indulgente per evitare di abbassare eccessivamente il voto medio, preoccupazione che si attenuerebbe per l'ultimo terzo di compiti che sarebbe oggetto di una valutazione più severa (Merle, 2018).

---

<sup>3</sup> Da uno studio di Addimando (2019) è emerso come il 12% dei docenti della scuola dell'obbligo ticinese abbia una concezione innatista dell'intelligenza, ovvero la consideri prevalentemente come un bagaglio genetico e biologico dato, modificabile solo entro certi limiti; un 39% ha invece secondo lo stesso studio una concezione costruttivista e ritiene pertanto che l'intelligenza sia plastica e modificabile grazie all'intervento del docente e della famiglia, mentre il restante 49% non è categorizzabile in uno di questi due gruppi.

C'è poi l'effetto alone, che consiste nella tendenza ad attribuire a una prova un giudizio complessivamente positivo o negativo, poiché una sua caratteristica specifica, considerata positiva o negativa, ha influenzato il giudizio a tal punto da estendere quel particolare tratto ad altri aspetti inficiando così il giudizio complessivo.

### 1.2.2 Bias scolastici della valutazione

Secondo Sarrazin e Trouilloud (2011) le aspettative che i docenti nutrono nei confronti dei discenti agiscono come una sorta di filtro interpretativo e portano a delle distorsioni della realtà nel momento in cui il docente percepisce, interpreta e valuta le azioni di un allievo. In generale i docenti si rapportano più spesso agli allievi che reputano migliori, rivolgono loro maggiore attenzione, li incoraggiano, le loro aspettative in termini di riuscita scolastica sono superiori e ciò favorisce la profezia che si auto-adempie (Bressoux e Pansu, 2003).

Anche gli allievi che si trovano a ripetere la classe o comunque in ritardo scolastico sono oggetto di sistematica sottovalutazione da parte dei docenti, ovvero tenderanno a riportare voti inferiori a chi non è incorso in bocciature o non ha un'età superiore a quella regolare, anche a parità di punteggio nel test standardizzato nella stessa disciplina. È come se per l'istituzione scolastica la bocciatura non fosse legata a una difficoltà contingente ma costituisse un impedimento permanente (Merle, 2018).

Anche frequentare una scuola o una filiera che gode di una buona reputazione comporta per gli allievi iscritti una valutazione più favorevole (Merle, 2018). In Ticino non è da escludere che il sistema con corsi attitudinali e base, che consentono di adattare l'insegnamento della matematica e del tedesco alle caratteristiche degli alunni della SM, incida sui voti: la frequenza di un alunno di un determinato curriculum fornisce informazioni che possono potenzialmente alterare il giudizio del docente, anche per una materia diversa dalla matematica o dal tedesco (Crotta et al., 2020).

### 1.2.3 Bias sociali della valutazione

#### a) Il genere dell'allievo

Le distorsioni della valutazione a vantaggio di uno dei due sessi emergono spesso dal confronto dei voti con i punteggi in test standardizzati (e anonimizzati ai correttori) (Iacus e Porro, 2011). Da questo confronto risulta frequentemente che sono le ragazze a riportare le valutazioni più favorevoli (Argentin e Triventi, 2015; Kiss, 2013; Lievore e Triventi, 2022; Terrier, 2015; Voyer e Voyer, 2014; Zanolla, 2023), anche se ci sono studi che, al contrario, non rilevano differenze tra le valutazioni dei due sessi (Hinrich et al., 2010) o le rilevano a vantaggio dei ragazzi (Tiedemann, 2000). Altri autori sottolineano che in lettura, matematica e scienze i ragazzi che hanno prestazioni analoghe alle ragazze ottengono un voto inferiore, ma che questo trattamento apparentemente meno favorevole scompare quando si considerano le competenze socio emotive (Cornwell et al., 2013; Zanolla, 2018). In effetti, come si ribadirà più avanti, ci sono evidenze che le pratiche di valutazione degli insegnanti riflettano anche l'impegno degli studenti, la motivazione e il loro comportamento (Bowers, 2011) e le ragazze, che sono spesso ritenute socialmente più competenti, più coscienti (Perander et al., 2020), più dotate di autodisciplina (Duckworth e Seligman, 2006) e più impegnate nelle attività di classe (Van Houtte, 2020) verrebbero ricompensate con voti più alti (Calarco, 2018; Cornwell et al., 2013). A essere particolarmente premianti sarebbero la loro coscientosità e il loro interesse per le materie, mentre la minore frequenza di comportamenti problematici avrebbe minore peso (Triventi et al., 2023). Altre ricerche hanno messo in evidenza lo stereotipo di genere secondo cui i ragazzi sono più logici delle ragazze e queste ultime meno portate per la matematica (Tomasetto, 2019; Giberti, 2019). Come scrive l'economista Rinaldi (2023), gli stereotipi di genere riguardo alle discipline STEM (le discipline scientifiche, tecnologiche, ingegneristiche e matematiche che svolgono un ruolo cruciale nella società attuale) nascono in famiglia e si consolidano nelle scuole: *A forza di ripetere alle bambine che non sono portate per le materie scientifiche, finiscono per crederci. E purtroppo è vero anche il contrario: come conseguenza i maschi tendono a sovrastimare la loro competenza. Insomma, gli uomini pensano di essere più bravi* (Rinaldi, 2023, pp. 137 -138). Se l'interiorizzazione di certi stereotipi da un lato è causa della scarsa fiducia delle ragazze nella matematica oltre che del loro fallimento nel raggiungere il loro pieno potenziale (Carlana, 2019), dall'altro può creare una sorta di discriminazione positiva a loro favore poiché

gli insegnanti assegnerebbero loro voti più alti proprio per incoraggiarle in una disciplina in cui le considerano più deboli (Terrier, 2020). In conclusione, la letteratura accademica sui *bias* della valutazione legati al sesso è ricca, ma non c'è sempre consenso unanime.

#### b) L'origine sociale dell'allievo

Spesso anche la mancata corrispondenza tra la provenienza sociale dell'insegnante e quella dell'allievo è fonte di distorsioni e pregiudizi e le ricerche sociologiche in numerosi casi hanno sottolineato che i giovani appartenenti alla classe media ricevono più supporto dai docenti e sono dunque più attrezzati per superare le sfide della scuola (Calarco, 2018; Romito, 2016). Gli insegnanti, secondo Argentin (2021), tendono a consigliare agli studenti di origine svantaggiata percorsi meno ambiziosi, ad assecondare i genitori più esigenti e critici nei loro confronti e ad avere una scarsa consapevolezza dei modi con cui gli stereotipi su origini sociali, genere e immigrazione plasmano le dinamiche nelle classi scolastiche, condizionando anche il loro agire. Alle pratiche discriminatorie messe in atto consapevolmente o inconsapevolmente dal personale scolastico si aggiunge un conflitto di tipo simbolico intorno a che cosa nella scuola viene definito "appropriato", "pertinente" e "di valore" e cosa al contrario viene definito "inappropriato" e "mediocre" (Romito, 2016). Le scuole, per quanto possano essere aperte a persone di tutte le provenienze sociali, non sono dunque contesti neutrali (Bourdieu, 1998), ma in genere i docenti si aspettano che studenti e genitori si comportino in maniera conforme alla classe media (Lareau, 2000) e chi mostra un elevato livello di capitale culturale e un insieme di comportamenti che corrispondono ai valori e alle aspettative della scuola riceve voti più alti (Bourdieu e Passeron, 1990), perché ritenuto accademicamente più dotato (Calarco 2014; Jæger e Breen 2016). Al contrario, i docenti tendono a sottostimare le capacità degli studenti degli strati più svantaggiati e questi ultimi finiscono per interiorizzare le percezioni degli insegnanti e per comportarsi in maniera aderente a esse (Jussim e Harber, 2005). I bambini senza background migratorio e con genitori con un titolo di studio elevato in generale hanno maggiore accesso a libri, riviste e giornali (Todd e Wolpin 2007; Yeung, Linver e Brooks-Gunn, 2002) e i loro genitori sono più propensi a leggere loro ad alta voce, coinvolgerli in attività e insegnare loro comportamenti e valori che vengono premiati dagli insegnanti (Calarco 2014; Lareau, 2011). Inoltre, se le famiglie di classe media educano i loro figli all'influenza, ovvero a chiedere direttamente supporto ai docenti nel momento in cui si trovano in una situazione di difficoltà, i genitori degli strati sociali più svantaggiati prediligono la strategia della deferenza, ossia educano i loro figli ad arrangiarsi il più possibile da soli e a non arrecare disturbo al docente (Calarco, 2018; Romito, 2016). Le prime danno priorità ai buoni voti (Lareau et al., 2011; Nelson 2010), i secondi al buon carattere dei figli (Edwards, 2004). Come sottolinea Romito (2016), il capitale culturale ha due dimensioni interconnesse: la prima riguarda la familiarità con il sistema educativo e la capacità di reperire informazioni e muoversi al suo interno (*embeddedness*); la seconda fa riferimento al senso di *entitlement*, ovvero di sicurezza o di fiducia in sé stessi con cui famiglie e studenti si interfacciano con il personale scolastico. Tutte queste differenze possono spiegare parte del divario nei voti (Boudon 1974; Fryer Jr e Levitt 2004; Jonsson e Rudolphi 2011; Mayer 1997; Yeung, Linver e Brooks-Gunn 2002), tuttavia non è chiaro quanto del divario socioeconomico nelle valutazioni sia dovuto a effettive differenze di abilità e quanto dipenda da fattori non correlati all'abilità, come differenze sistematiche nel capitale culturale o altre forme di pregiudizio degli insegnanti (Hjorth-Trolle et al., 2022).

#### c) Comportamento in classe dell'allievo e attrattività fisica

Secondo alcuni studi i voti attribuiti dai docenti sono più fortemente correlati con i fattori non cognitivi di quanto non lo siano i punteggi nei test standardizzati (Devries et al., 2018; Farrington et al., 2012; Lechner et al., 2017). Al di là dell'abilità degli allievi e delle loro caratteristiche sociodemografiche, i comportamenti prosociali costituiscono un forte predittore dei voti (Gerbino et al., 2018). Tra gli aspetti non cognitivi che possono influenzare la valutazione dei docenti figurano i problemi di comportamento, la capacità attentiva, l'organizzazione del materiale scolastico e la capacità di tener traccia delle consegne (Angelo 2014; Enzi, 2014). Tra i tratti della personalità più premianti si annoverano l'apertura mentale (Gerbino et al., 2018; Laidra et al., 2007; Zuffianò, 2013), come pure la già menzionata coscienziosità (Conrad e Patry, 2012) e ovviamente la motivazione e l'impegno (Chamorro-Premuzic e Furnham, 2014). Al contrario nevroticismo ed estroversione penalizzerebbero la valutazione (Ackerman e Heggestad, 1997).

Per quanto riguarda l'aspetto fisico, una sintesi di 28 ricerche realizzata da Ritts, Patersson e Tubs (1992) ha mostrato che l'attrattività fisica esercita un effetto positivo sulla percezione dell'intelligenza e del potenziale scolastico e sui voti degli allievi. È risaputo che nelle nostre società occidentali ci sia una preferenza per la magrezza e il sovrappeso sia visto negativamente (Swami et al., 2010). Le persone in sovrappeso sono ritenute responsabili della propria condizione, meno dotate di autodisciplina, pigre e poco coscienti (Puhl e Heuer, 2009) e spesso certi giudizi vengono espressi senza troppe remore (Latner et al., 2008; Wang et al., 2004). Da alcuni studi è emerso che anche i docenti possono nutrire pregiudizi nei confronti degli allievi in sovrappeso (Wilson et al., 2015) e questi ultimi riportano, a parità di caratteristiche con gli allievi normopeso, voti medi significativamente inferiori (Florin et al., 2011). Secondo Dian e Triventi (2021) questa penalizzazione appare più marcata tra i maschi, sui quali ricadono dunque due stereotipi, quello dell'essere meno diligenti delle ragazze e quello dell'essere meno dotati di autodisciplina legato all'essere in sovrappeso, con conseguente doppio svantaggio.

#### d) Le caratteristiche del docente

I modi in cui gli insegnanti valutano i loro allievi sono importanti, perché i voti sono sia un indicatore dell'abilità degli allievi, sia qualcosa che può influenzare la loro abilità, in quanto i giudizi degli insegnanti condizionano il modo in cui gli studenti percepiscono sé stessi (effetto Pigmalione). I docenti si comportano in maniera diversa con i diversi allievi anche a dipendenza delle aspettative che nutrono nei loro confronti (Rubie-Davies, 2018) e gli allievi tenderanno a conformarsi alle loro attese. Come sopra scritto, la letteratura evidenzia che ci sono stereotipi circa la minore competenza delle ragazze nelle materie STEM (Ridgeway e Smith-Lovin, 1999) e aspettative inferiori nei confronti degli allievi di origine sociale svantaggiata o appartenenti alle minoranze etniche (Tobisch e Dresel, 2017) e da questi i docenti possono mettere in atto comportamenti discriminatori o al contrario antidiscriminatori proprio per compensarli (Behaghel et al, 2015).

Alcuni insegnanti sono poi più inclini di altri a distorcere le informazioni relative alle competenze scolastiche degli allievi e, nonostante i risultati delle ricerche non siano completamente in linea tra loro, sembra che coloro che hanno una personalità dogmatica presentino una tendenza più pronunciata a formulare stereotipi sociali per semplificare le categorie sociali (Merle, 2018).

Bisogna inoltre ricordare che i docenti operano in classi e scuole con determinate caratteristiche che influenzano la relazione insegnante – studente, come per esempio il numero di studenti per classe, la composizione di genere, la composizione sociale, la leadership del direttore scolastico.

#### 1.2.4 La contestualità delle pratiche valutative

Il sociologo italiano Giancola (2019) sottolinea come l'attribuzione del voto scolastico non sia un'operazione che avviene in un momento isolato dalla quotidiana esperienza scolastica, ma al contrario sia immersa nella realtà della classe scolastica, nella quale coesistono le finalità degli insegnanti, le strategie degli studenti, le regole e le pratiche inscritte negli spazi, nei tempi, negli strumenti della scuola.

*Gli attori e gli oggetti che tessono la tela delle relazioni entro la quale si costruisce socialmente il voto sono tanti e a vario livello c'è l'insegnante, c'è lo studente - con le sue infinite peculiarità individuali, le sue inclinazioni, il suo carattere e la sua capacità performativa – c'è la classe – tanto nel suo insieme quanto nei suoi sottogruppi stereotipici tipo "i bravi", i "distratti" ecc. - ci sono le griglie di valutazione e gli apparati di registrazione della performance (veri e propri attori "non umani"), nonché le relazioni tra i docenti e i rispettivi giudizi "incrociati" sullo studente Tizio o Caio. Volendo, in questo elenco (per forza di cose assolutamente incompleto) mancano i Dirigenti scolastici, il ruolo delle famiglie e quello degli operatori della scuola. Insomma, l'attribuzione del voto (...) è un'operazione densa di significati e di relazioni che configura un livello di complessità tale che, per effetto del principio di autorità del docente (...) viene annullato e naturalizzato (ovviamente non senza conflitti, dinamiche di opposizione o di messa in discussione, ecc.) entro il quadro dell'asimmetria di potere tra studente e docente (Giancola, 2019, p. 136).*

Giancola rileva come gli insegnanti utilizzino i voti anche come strumento motivante, ovvero per spronare gli studenti a continuare in un determinato modo, nell'impegno, nel rispetto, nella collaborazione, ecc. Per esempio, capita che un docente dia una sufficienza allo studente la cui prestazione non è ancora del tutto

sufficiente per mostrargli che impegnandosi può arrivare a un buon voto. L'impegno, in particolare, secondo Giancola è un artificio retorico che sottolinea il permanere di un atteggiamento che premia l'elemento volontaristico del processo di apprendimento e miglioramento e che richiama da vicino *l'equità al ribasso*, ossia il fatto che la scuola non ricompensi tanto le prestazioni, quanto piuttosto un atteggiamento morale. Secondo il sociologo, i voti sono poi oggetti instabili: sono aggiustati in itinere in un processo che cumula correzioni e distorsioni, assestamenti e adattamenti. Per ultimo, dopo il processo di aggiustamento per singolo docente o sottogruppi (diadi, triadi, ecc. di docenti), i voti sono ulteriormente negoziati a livello di classe, dove confluiscono elementi comportamentali, di atteggiamento, di congruenza performativa tra le varie discipline, casualità assortite e scelte idiosincriche (su singoli studenti o gruppi di essi). Dalle osservazioni dei consigli di classe si evince una tendenza alla convergenza (anche se non mancano conflitti e attriti) che secondo Giancola può riorientare anche significativamente il voto (Giancola, 2019).

### 1.3 Voti e punteggi nei test standardizzati<sup>4</sup>

Per tutti gli elementi evidenziati nel paragrafo precedente, è dunque chiaro che i voti, oltre ad attestare il possesso di determinate abilità cognitive, a fornire un riscontro sull'apprendimento di una particolare materia da parte degli studenti e a costituire dei criteri formali e informali per stabilire chi può proseguire verso percorsi formativi più o meno ambiziosi, certificano anche l'avvenuta acquisizione di abilità non cognitive (Pattison et al., 2013).

Il carattere multidimensionale dei voti e in particolare la loro componente non cognitiva sono stati ampiamente sottolineati nelle ricerche degli anni Ottanta e Novanta del secolo scorso, in certi casi anche per metterne in dubbio la validità (Brookhart, 2015; Cross & Frary, 1999), in contrapposizione all'oggettività dei test standardizzati (Geiser & Santelices, 2007).

La natura multidimensionale dei voti li rende in ogni caso, secondo alcuni autori, predittori migliori del successo scolastico di un individuo nei cicli scolastici successivi rispetto ai punteggi nei test standardizzati (Cliffordson, 2008; Geiser & Santelices, 2007) e favorirebbe le ragazze, tipicamente più aderenti alle aspettative degli insegnanti (Dee, 2005; Dee, 2007; Lekholm & Cliffordson, 2008, 2009). Secondo alcuni studi anglosassoni la relazione tra i voti e l'origine sociale degli individui è più debole di quella che sussiste tra quest'ultima e i punteggi nei test standardizzati (Atkinson e Geiser, 2009) e i voti costituirebbero pertanto uno strumento di valutazione più equo. Ciò che può sembrarci alquanto controintuitivo è probabilmente legato alle finalità del test standardizzato: nei casi in cui il suo superamento sia il requisito necessario per accedere a delle formazioni prestigiose, cosa che avviene nel mondo anglosassone ma non in Ticino, è lecito attendersi che i genitori dei ceti superiori siano disposti a investire ingenti risorse per la preparazione al test dei figli. Per quanto riguarda l'appartenenza nazionale, invece, pare che i voti penalizzino più del test gli allievi stranieri a causa delle minori aspettative degli insegnanti nei loro confronti (Burgess & Greaves, 2009).

Nei test standardizzati l'elemento della relazione tra insegnante e allievo scompare, perché ogni studente è sottoposto nelle stesse condizioni alle stesse prove, che sono poi corrette secondo gli stessi criteri da esaminatori esterni e imparziali in modo da garantire che i risultati siano fra loro comparabili. Tuttavia i test variano, oltre che secondo gli obiettivi che sono volti a misurare, anche secondo l'accuratezza delle stime, che a sua volta dipende da fattori quali i contenuti e le abilità che vengono inserite, il peso che si decide di attribuire a esse, la formulazione verbale degli item, la motivazione e lo stato psicofisico dei giovani cui è sottoposto il test, la qualità dell'ambiente in cui la prova viene eseguita, oltre al tempo che la scuola investe nel preparare i ragazzi in vista della prova (Koretz, 2008). In alcuni studi è stato inoltre riscontrato che maschi e femmine differiscono nel loro atteggiamento verso la competizione, in quanto le femmine è più

---

<sup>4</sup> Gran parte di questo paragrafo è tratto da: Crotta, F., Salvisberg, M., & Zanolli, G. (2020). Note scolastiche e competenze fondamentali in matematica degli allievi dell'11° anno HarmoS in Ticino. *Schweizerische Zeitschrift für Bildungswissenschaften*, 42(2), 323-342.

probabile che si intimidiscano in una situazione competitiva e occorre pertanto cautela nell'imputare l'eventuale scarto tra voti e punteggio nel test totalmente alla discriminazione degli insegnanti (Zanga e De Gioannis, 2023).

Il confronto tra voti e punteggi nei test standardizzati può tuttavia fornire delle indicazioni sulla severità del giudizio dei docenti e sulla correttezza del posizionamento degli studenti secondo la distribuzione dei voti (Bonesrønning, 1999; Lindahl, 2007) e costituire un'utile occasione di riflessione per i docenti sui criteri di valutazione adoperati e i loro eventuali pregiudizi (Wiberg, 2019; Zwick e Green, 2007) o, come nel caso italiano, contribuire a fare luce su eventuali disparità regionali negli standard di assegnazione dei voti (Argentin e Triventi, 2015).

Lo scarto tra i due tipi di valutazione può anche essere utilizzato per l'individuazione dei giovani con un maggior rischio di dropout, in quanto sembra che tali soggetti riportino una valutazione nei test migliore di quella assegnata loro dagli insegnanti (Marcenaro-Gutierrez e Vignoles, 2015). Inoltre, considerare congiuntamente i voti e i punteggi nei test consentirebbe di predire con più accuratezza le prestazioni all'università (Atkinson e Geiser, 2009). Il confronto può infine essere utile per fornire indicazioni sulla regolamentazione del sistema educativo, ad esempio Salvisberg e Zampieri (2015) hanno evidenziato come la correlazione tra i voti scolastici in matematica e i risultati in PISA nella stessa disciplina sia inferiore alle correlazioni tra i voti in italiano e scienze e i risultati in PISA nelle rispettive materie. Una spiegazione potrebbe risiedere nel regolamento della Scuola media (Consiglio di Stato della Repubblica e Cantone Ticino, 30 maggio 2018): gli allievi, pur di non precludersi l'opportunità di iscriversi a una scuola media superiore, potrebbero decidere di restare nel corso A in matematica anche con un'insufficienza, piuttosto che frequentare il corso B, dove otterrebbero verosimilmente un voto più alto.

Vista l'utilità qui esplicitata di considerare congiuntamente i voti e i risultati nei test standardizzati svolti nelle stesse materie, nei modelli multilivello volti a identificare le variabili individuali e attinenti alla classe e al docente che incidono sul voto di fine anno, si è deciso di focalizzarci sulla sola scuola elementare, e in particolare sulle classi quinte, disponendo per queste classi dei risultati di un test standardizzato in matematica riferito all'anno scolastico 2014/15 (Crescentini, 2017b) e di due test in italiano e in matematica organizzati nell'anno scolastico 2020/21 dal team delle prove standardizzate del CIRSE (Meier et al., 2022).



## 2 Obiettivi e metodi

Gli obiettivi principali<sup>5</sup> di questo lavoro sono:

- 1- individuare le caratteristiche proprie dell'allievo/a e attinenti sia alla classe sia al docente che incidono sul voto di fine anno;
- 2- capire quanto i voti di fine anno, in determinate materie di un certo ordine scolastico, siano predittivi dei voti di fine anno nell'anno scolastico seguente, sia esso nello stesso ordine scolastico o nell'ordine scolastico successivo.

Per quanto riguarda il primo obiettivo, la ricerca in campo educativo, come si è visto nel paragrafo 1.2, ha dimostrato che il profitto scolastico di un allievo in un dato momento non dipende solo dalle sue caratteristiche personali e dall'insieme di abilità e conoscenze accumulato negli anni precedenti, ma anche dalle caratteristiche e dal livello medio delle competenze possedute dagli allievi della sua classe e della sua scuola e dalle caratteristiche del docente che attribuisce il voto. In altre parole, esiste un effetto di contesto, o di composizione del gruppo di cui un alunno fa parte, che condiziona in qualche modo il suo stesso apprendimento. Volendo separare l'effetto delle caratteristiche degli allievi da quello del contesto, la regressione multilivello, consentendo di analizzare gli effetti del contesto sui risultati scolastici degli studenti, si rivela un modello più adeguato della regressione ordinaria allo studio delle relazioni tra variabili in campo educativo (Raudenbusch e Bryk, 2002). Essa tiene conto della struttura gerarchica dei dati che i fenomeni in campo educativo tipicamente presentano: gli allievi, infatti, non sono normalmente isolati gli uni dagli altri ma sono raggruppati in classi, che a loro volta sono riunite in scuole, le scuole in circondari, ecc., ed è probabile che gli alunni appartenenti a uno stesso gruppo (ad esempio la stessa classe o scuola) siano tra loro più simili rispetto agli alunni di gruppi diversi, e ciò anche a causa dell'esposizione a fattori ambientali comuni una volta che i gruppi si sono costituiti. Tecnicamente si procede elaborando un primo modello "vuoto", che non contiene variabili esplicative ma si limita a scomporre la varianza complessiva delle medie dei voti nella componente tra le classi e in quella tra gli studenti della stessa classe. Successivamente, si può procedere valutando l'esito sul voto delle caratteristiche degli studenti (genere, origine sociale, nazionalità, ecc.) e, nei limiti delle possibilità offerte dalla banca dati GAS-GAGI, delle caratteristiche della classe (intesa sia come composizione degli studenti che ne fanno parte sia come docente o collegio dei docenti) ed eventualmente della scuola.

Una tra le variabili relative all'individuo, che concorrono a determinare il voto di fine anno di una determinata materia, è l'abilità dell'allievo/a in quella determinata materia, che può essere indicato attraverso il punteggio riportato in un test standardizzato. Se vogliamo individuare i "bias" della valutazione, cioè le variabili individuali e riferite al contesto che giocano un ruolo significativo nella determinazione dei voti, a parità di livello di abilità degli studenti, dobbiamo inserire il punteggio riportato in un test standardizzato nella stessa materia del voto su cui si sta indagando e svolto nello stesso anno scolastico. Ciò ci porta a focalizzare l'analisi su:

- il voto in matematica in quinta elementare negli anni scolastici 2014/15 e 2020/21;
- il voto in italiano nelle quinte elementari nell'anno scolastico 2020/21.

Per tali classi, materie e anni scolastici disponiamo infatti per l'intera coorte di allievi e allieve dei punteggi nei test standardizzati (Crescentini, 2017b; Meier et al., 2022).

Il secondo obiettivo di questo lavoro implica il confronto della capacità predittiva del voto con quella dei punteggi nella prova standardizzata nella stessa disciplina. Saranno dunque analizzate:

---

<sup>5</sup> Negli ultimi anni sono state eseguite delle analisi volte a confrontare la distribuzione i voti e le medie degli stessi attribuite dai vari docenti, nelle diverse scuole, classi e assegnate ai ragazzi di genere, origine sociale e nazionalità diversi e ad analizzarne la tendenza da un ordine scolastico a quello successivo. Tali calcoli sono aggiornati all'anno scolastico 2021/22 e forniti corredati da un commento sintetico in un documento a parte.

- a) le relazioni tra il punteggio nella prova standardizzata in italiano realizzata in terza elementare e il voto in italiano di terza elementare e il voto riportato in quarta elementare (all'interno quindi del livello primario);
- b) le relazioni tra il punteggio nella prova standardizzata in italiano realizzata in quinta elementare e il voto in italiano ottenuto in quinta elementare e quello nella stessa materia in prima media (a cavallo tra livello primario e secondario I).

## 3 Risultati

### 3.1 Le caratteristiche proprie dell'allievo/a e attinenti alla classe e al docente che incidono sul voto di fine anno

#### 3.1.1 Informazioni generali sui modelli multilivello: la popolazione coinvolta

Con riferimento all'obiettivo 1 del capitolo 2, si è proceduto nell'elaborazione di tre modelli multilivello, la cui popolazione è indicata nella figura 1.

Figura 1 Popolazione relativa ai modelli multilivello elaborati

Modello relativo al voto in <b>matematica</b> in V SE a.s. 2014/15	Modello relativo al voto in <b>matematica</b> in V SE a.s. 2020/21	Modello relativo al voto in <b>italiano</b> in V SE a.s. 2020/21
2.369 allieve/i di 173 classi	2.238 allieve/i di 181 classi	2.161 allieve/i di 181 classi
In tutti e tre i modelli sono stati esclusi le bambine e i bambini delle pluriclassi (volendo inserire nel modello la variabile "punteggio di abilità medio nel test della classe" veniva a mancare il risultato di una parte dei compagni, quelli iscritti in altri anni diversi dal quinto) e quelli frequentanti classi con due docenti di genere diverso (volendo inserire nel modello la variabile "genere del docente", nelle classi in cui i due docenti sono di genere diverso non è possibile attribuire il genere di chi vi insegna).		

#### 3.1.2 Risultati dei modelli vuoti

Dai modelli vuoti<sup>6</sup> risulta che la percentuale di varianza legata alla classe frequentata è pari a 7.0% nel modello volto a stimare il voto in matematica nel 2014/15, a 4.9% nel modello volto a stimare quello riferita all'anno 2020/21 e a 4.3% nel modello a stimare il voto in italiano nel 2020/21 (v. Appendice, modello 1 nelle figure A, B e C). Tali percentuali sono sufficientemente elevate per procedere nell'elaborazione di modelli multilivello in tutti e tre i casi.

#### 3.1.3 Risultati dei modelli definitivi

Introducendo nella stima dei modelli successivi al modello base una o più variabili esplicative, di primo (relative all'allievo) e secondo (relativo alla classe frequentata) livello, le componenti di varianza "tra" ed "entro" si riducono e questo consente di valutare, per differenza rispetto alla varianza iniziale, il contributo di ciascun predittore, o gruppo di predittori, alla spiegazione della variabilità dei dati.

Nei modelli finali (v. Appendice, modello 3 nelle figure A, B e C) sono state inserite le seguenti variabili, che, salvo nel caso dei punteggi riportati nei test standardizzati, sono incluse nella banca dati GAS-GAGI:

- Livello 1 (**allievo/a**): punteggio riportato nel test standardizzato<sup>7</sup>, professione dei genitori<sup>8</sup>, età, madrelingua, genere, voto di condotta;

<sup>6</sup> I modelli vuoti sono modelli in cui non ci sono predittori né a livello 1 (individuo) né a livello 2 (classe) e in cui si assume che il valore della variabile dipendente (nel nostro caso il voto di fine anno) dipenda solo dalla media dell'intera popolazione, dalla componente di errore associata al gruppo di appartenenza (la classe) e da una componente di errore individuale. La variabilità del voto viene così scomposta nella somma di due componenti: la varianza "tra" le classi e la varianza "entro" le classi. È così possibile calcolare il "coefficiente di correlazione intraclasse" ( $\rho$ ), che permette di valutare il grado di omogeneità tra gli allievi e le allieve iscritti/e alla stessa classe. Il coefficiente  $\rho$  assume valori compresi tra 0 e 1 ed esprime la proporzione tra la varianza di secondo livello (legata alla classe frequentata) e la somma delle varianze dei due livelli (individuale e della classe), o varianza totale. Quanto più  $\rho$  si avvicina ad 1, tanto più forte è l'effetto del raggruppamento e quindi maggiore l'utilità di un'analisi multilivello.

<sup>7</sup> Per il test in matematica relativo al 2014/15, v. Crescentini (2017b), per il test in matematica e quello in italiano relativi al 2020/21, v. Meier et al., (2022).

<sup>8</sup> L'origine socioeconomica delle allieve e degli allievi è stata calcolata utilizzando l'indice ISEI (International Socio-Economic Index of Occupational Status) basato sulla professione dei genitori. Nella procedura di calcolo si classificano le professioni di entrambi i

- Livello 2 (**classe**): punteggio medio della classe nel test standardizzato, tipo di contratto del docente titolare (un docente a tempo pieno o due a tempo parziale), genere del docente<sup>9</sup>.

Ciò che emerge è che:

- il risultato nel test standardizzato ha un effetto positivo "grande" (il coefficiente  $f^2$  di Cohen è  $> 0.35$ ) in tutti e tre i modelli;
- avere almeno uno dei genitori con professione altamente qualificata esercita un effetto positivo sul voto in matematica e su quello in italiano nel 2020/21, mentre per quanto riguarda il voto in matematica del 2014/15 a fare la differenza, in senso negativo, è soprattutto avere entrambi i genitori con professioni scarsamente qualificate;
- l'età superiore a quella dei compagni e delle compagne di classe in tutti e tre i modelli ha un effetto negativo sul voto di fine anno;
- la madrelingua non italiana non esercita alcun effetto significativo sul voto in matematica, ma abbassa quello in italiano;
- il genere femminile dell'allievo ha un effetto negativo sul voto finale in matematica nel 2020/21 e un effetto positivo sul voto finale in italiano nello stesso anno, mentre non esercita alcun effetto sul voto in matematica nel 2014/15;
- il voto di condotta ha un effetto positivo sui voti finali di tutti e tre i modelli;
- maggiore è il punteggio medio della classe nel test standardizzato, indicatore della sua abilità nella materia, minore sarà il voto di fine anno dell'individuo;
- se il genere del docente non sembra esercitare alcun effetto significativo sul voto di fine anno in nessuno dei tre modelli, avere due docenti titolari a metà tempo ha un effetto negativo sul voto in matematica nel 2014/15 e su quella in italiano nel 2020/21, mentre non esercita un effetto significativo sul voto in matematica nel 2020/21.

In pratica:

- si osserva una certa coerenza tra l'abilità dell'allievo/a risultante dal test e la valutazione nella stessa disciplina del test attribuita dal docente;
- l'effetto positivo dell'origine sociale avvantaggiata sul voto in matematica e italiano del 2020/21 e quello negativo dell'origine svantaggiata sul voto in matematica nel 2014/15 si evidenziano a parità di livello di abilità nel test standardizzato delle materie in questione: le motivazioni possono risiedere nella maggiore o minore prossimità culturale del soggetto con l'istituzione scolastica o nella relazione scuola-famiglia;
- c'è probabilmente un effetto di etichettamento (*labelling*) che grava sugli allievi con uno o più anni in più per progressi rallentamenti, bocciature o trasferimenti dall'estero;

---

genitori secondo le categorie della classificazione standard svizzera delle professioni CH-ISCO-19 che incorpora i primi quattro livelli dell'International Standard Classification of Occupation (ISCO-08) cui ha aggiunto un quinto livello per tener conto delle particolarità del mercato del lavoro svizzero. Si applica quindi il criterio di dominanza, che consiste nel selezionare il punteggio di status più elevato tra quelli del padre e della madre. A ciascun individuo viene quindi assegnato uno tra i seguenti livelli di origine socioeconomica (OECD, 2018): professioni altamente qualificate (gruppi professionali ISCO dall'1 al 3); professioni mediamente qualificate (gruppi professionali dal 4 all'8); professioni scarsamente qualificate (gruppo professionale 9). Nella classificazione ISCO-19 il codice 1 si riferisce ai dirigenti, 2 alle professioni intellettuali e scientifiche, 3 alle professioni tecniche intermedie; 4 agli impiegati d'ufficio; 5 alle professioni nelle attività commerciali e nei servizi; 6 al personale specializzato addetto all'agricoltura, alle foreste e alla pesca; 7 agli artigiani e operai specializzati; 8 ai conduttori di impianti e macchinari e addetti al montaggio; 9 alle professioni non qualificate.

<sup>9</sup> Purtroppo, non è stato possibile ottenere il dato relativo all'anzianità di servizio al netto dei congedi e, risultando l'effetto della variabile al lordo dei congedi non significativo sul voto, si è deciso di scartare questa variabile dal modello.

- l'effetti negativo della madrelingua non italiana sul voto in italiano si palesa a parità di abilità dimostrata nel test e potrebbe pertanto dipendere da un pregiudizio del docente nei confronti degli allievi e delle allieve di madrelingua straniera;
- i modelli del 2020/21 sembrano confermare il pregiudizio che vuole le ragazze meno portate per la matematica e più predisposte per l'italiano;
- un comportamento non rispettoso delle regole non abbassa solo il voto di condotta, ma anche i voti delle materie curriculari;
- la valutazione, come si è scritto nel paragrafo 1.2.1, è comparativa ed essere confrontati con un gruppo forte è svantaggioso;
- il risultato nel test standardizzato è l'unica variabile che esercita un effetto "grande": tutte le altre variabili qui menzionate, anche quando esercitano un effetto significativo sui voti di fine anno, hanno un effetto "piccolo" (il coefficiente  $f^2$  di Cohen è  $< 0.15$ ). La somma degli "effetti piccoli" dà però luogo a qualcosa d'importante e arriva a modificare il voto di fine anno fino a un punto, il che in un intervallo che normalmente alla SE è compreso tra 4.0 e 6.0, è di assoluto rilievo. Due modelli analoghi realizzati all'ultimo anno della SE nel 2020/21 nel Canton Ginevra relativamente ai voti di fine anno in matematica e francese hanno evidenziato lo stesso risultato (Petrucci e Zanolla, 2023), corroborando così i risultati di queste analisi.

### 3.2 Punteggi nelle prove standardizzate e voti: quale relazione?

Il progetto della valutazione delle competenze disciplinari degli allievi ticinesi delle scuole elementari e scuole medie è attivo sul territorio cantonale dal 2013, anno in cui è stata realizzata la prima rilevazione delle competenze disciplinari in matematica in quarta elementare. Nel corso degli anni sono state realizzate diverse valutazioni delle competenze nelle discipline di matematica, italiano e geografia (figura 2).

Figura 2 Prove standardizzate cantonali realizzate nelle scuole elementari ticinesi e rispettive materie

	2013	2014	2015	2021	2022	2023	2024
<b>Matematica</b>	4 <sup>a</sup> Elementare		5 <sup>a</sup> Elementare	5 <sup>a</sup> Elementare			4 <sup>a</sup> Media
<b>Italiano</b>		3 <sup>a</sup> Elementare		5 <sup>a</sup> Elementare			4 <sup>a</sup> Media
<b>Geografia</b>						4 <sup>a</sup> Media	

Queste valutazioni hanno interessato gli allievi ticinesi che frequentavano i vari anni del ciclo di scolarità obbligatoria. In questo contesto i risultati alle prove standardizzate hanno mostrato una relazione (in termini di correlazione), consistente negli anni e trasversalmente alle discipline scolastiche considerate, con i voti degli allievi ottenute nello stesso anno (Crescentini, et al., 2014; Crescentini, 2017a; Crescentini, 2017b; Meier, et al., 2022). Ciò è in linea con diverse ricerche, che mostrano come i punteggi nei test standardizzati siano dei buoni predittori dei voti (Kuncel, 2007) nonostante le finalità valutative differenti (Crahay e Mottier Lopez, 2009; Salvisberg e Zampieri, 2015). Tuttavia, il potere predittivo dei punteggi alle prove standardizzate sui voti degli allievi negli anni successivi alla realizzazione delle prove stesse non è ancora stato oggetto di analisi approfondite.

Al fine di investigare la relazione tra i punteggi alle prove standardizzate e i voti, in questo paragrafo verranno analizzati i dati di due coorti di allievi ticinesi che hanno svolto le prove standardizzate in italiano

in terza (Crescentini, 2017a) e quinta (Meier, et al., 2022) elementare<sup>10</sup>. Questo approccio di analisi permette di valutare l'effetto predittivo dei risultati alle prove standardizzate nel corso del tempo (un anno dopo lo svolgimento delle prove) e di comprendere se l'effetto predittivo possa dipendere dal contesto in cui vengono realizzate le valutazioni degli allievi. Verranno dunque analizzate 1) le relazioni tra i punteggi alle prove standardizzate realizzate in terza elementare e i voti ottenuti in terza e quarta elementare (all'interno del livello primario); 2) le relazioni tra i punteggi alle prove realizzate in quinta elementare e i voti ottenuti in quinta elementare e in prima media (a cavallo tra livello primario e secondario I).

Le relazioni tra i punteggi alle prove standardizzate e i voti per le due coorti di allievi considerate sono state analizzate tramite delle analisi di correlazione che permettono di apprezzare l'associazione (covariazione) di due o più variabili numeriche (Figura E in appendice). Le analisi realizzate indicano che i voti in italiano ottenuti dagli allievi in terza elementare correlano fortemente ( $r = .87$ ) con i voti ottenuti l'anno successivo in quarta elementare. In modo analogo i punteggi alla prova standardizzata di italiano effettuata in terza elementare (nei mesi di maggio e giugno) correlano in modo elevato ( $r = .66$ ) con i voti in italiano in terza elementare e con i voti ottenuti in quarta elementare nella stessa materia ( $r = .65$ ). Si osservano inoltre delle correlazioni elevate tra i voti in italiano di quinta elementare e di prima media ( $r = .67$ ), tuttavia inferiori alle correlazioni osservate tra i voti in italiano in terza e quarta elementare (Figura F in appendice). Si osserva altresì come i punteggi alle prove standardizzate in italiano effettuate in quinta elementare si correlino fortemente con i voti in italiano in quinta elementare ( $r = .72$ ) ma in misura minore con i voti in prima media ( $r = .63$ ).

Le analisi di correlazione precedenti indicano, in sintesi, come i punteggi alle prove standardizzate correlino fortemente con i voti ottenuti nel corso dello stesso anno scolastico e con i voti ottenuti nell'anno scolastico successivo indifferentemente dall'anno scolastico in cui sono state realizzate le prove standardizzate. Tuttavia, la relazione tra i punteggi alle prove standardizzate ed i voti ottenuti nell'anno successivo è inferiore quando gli allievi passano dalla quinta elementare alla prima media comparativamente a quella che si osserva al passaggio tra terza e quarta elementare. In altre parole, come era lecito aspettarsi, la relazione tra i punteggi alle prove standardizzate e il voto in italiano ottenuto l'anno scolastico successivo è più debole quando gli allievi intraprendono il passaggio tra livello primario e livello secondario I con il relativo cambiamento di classe e di docente responsabile dell'insegnamento della materia.

Al fine di comparare l'effetto predittivo dei punteggi alla prova standardizzata sui voti in italiano ottenuti l'anno scolastico successivo, è stata dunque realizzata un'analisi di regressione lineare multipla. Questa tipologia di analisi permette di analizzare l'effetto dei punteggi alle prove standardizzate sui voti ottenuti l'anno successivo indipendentemente dal biennio considerato (terza-quarta elementare e quinta elementare–prima media) e introducendo al contempo delle variabili di controllo (voti ottenuti lo stesso anno in cui sono state somministrate le prove, genere, lingua madre e livello socioeconomico dell'allievo). L'obiettivo è quello di valutare il potere predittivo delle prove standardizzate quando gli allievi affrontano il passaggio dalle scuole elementari alle scuole medie (dal livello primario al secondario I) comparativamente al passaggio tra classi di scuola elementari successive (all'interno del livello primario). Le variabili di controllo sono state incluse al fine di valutare l'effetto predittivo delle prove standardizzate indipendentemente dall'effetto di queste variabili.

Il modello di regressione utilizzato (Figura G in appendice) predice i voti degli allievi un anno scolastico dopo la realizzazione delle prove standardizzate (anno  $n+1$ ) utilizzando come predittori i punteggi ottenuti alle prove standardizzate, i voti nell'anno di realizzazione delle prove (anno  $n$ ), il genere, la lingua madre,

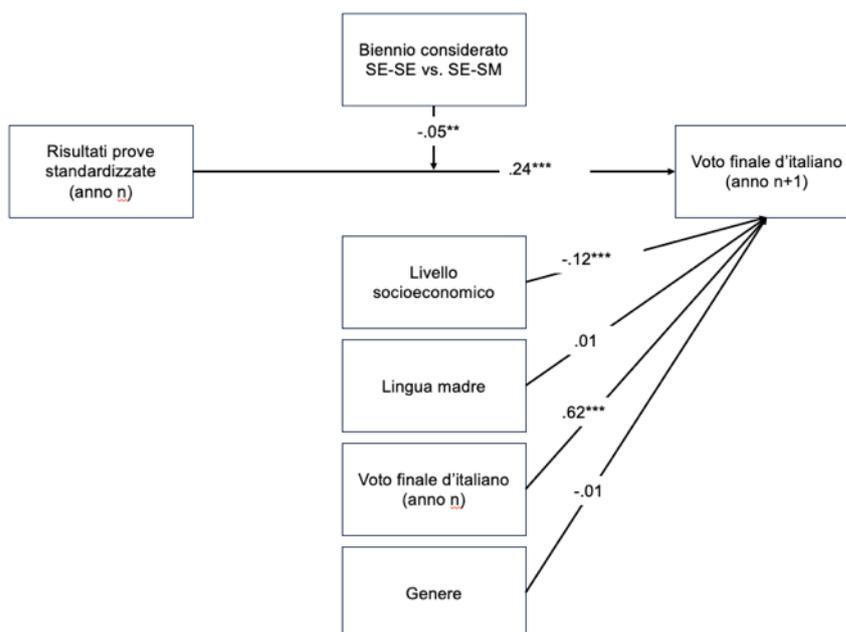
---

<sup>10</sup> La prova d'italiano svolta in terza elementare nel corso dell'anno scolastico 2013/14 comprendeva complessivamente 107 item dei quali 28 attinenti al lessico, 19 alla punteggiatura, 34 all'ortografia morfologica e 26 all'ortografia fonologica" (Crescentini, 2017a, p. 9). Hanno partecipato alla prova tutti gli allievi ( $N = 2982$  bambini di cui 50.8% maschi e 49.2% femmine) delle classi di terza elementare del Cantone. La prova d'italiano svolta in quinta elementare nell'anno scolastico 2020-2021 era composta da 228 domande di cui: 20 attinenti alla metacognizione, 122 all'ortografia, 26 alla punteggiatura e 60 alle inferenze, al lessico, alla lettura e agli aspetti linguistici. Hanno partecipato alla prova tutti gli allievi di quinta elementare del Cantone ( $N = 2903$ , di cui 49.8% maschi e 50.2% femmine). I voti degli allievi per gli anni scolastici considerati, così come le loro caratteristiche sociodemografiche sono state ottenute e messe in relazione ai risultati delle due prove standardizzate tramite i registri amministrativi GAGI.

il livello socioeconomico e una variabile intergruppo ("biennio considerato") che indica se i dati si riferiscono al biennio terza-quarta elementare (SE-SE) o al biennio quinta elementare-prima media (SE-SM). È stato inoltre aggiunto un termine d'interazione al fine di valutare se l'effetto predittivo delle prove standardizzate sui voti all'anno  $n+1$  dipenda dal biennio considerato. Il modello teorico dell'analisi di regressione lineare multipla è rappresentato nella figura 3 dove sono pure indicati i coefficienti standardizzati ottenuti che indicano la forza e la direzione dell'associazione tra le varie variabili del modello.

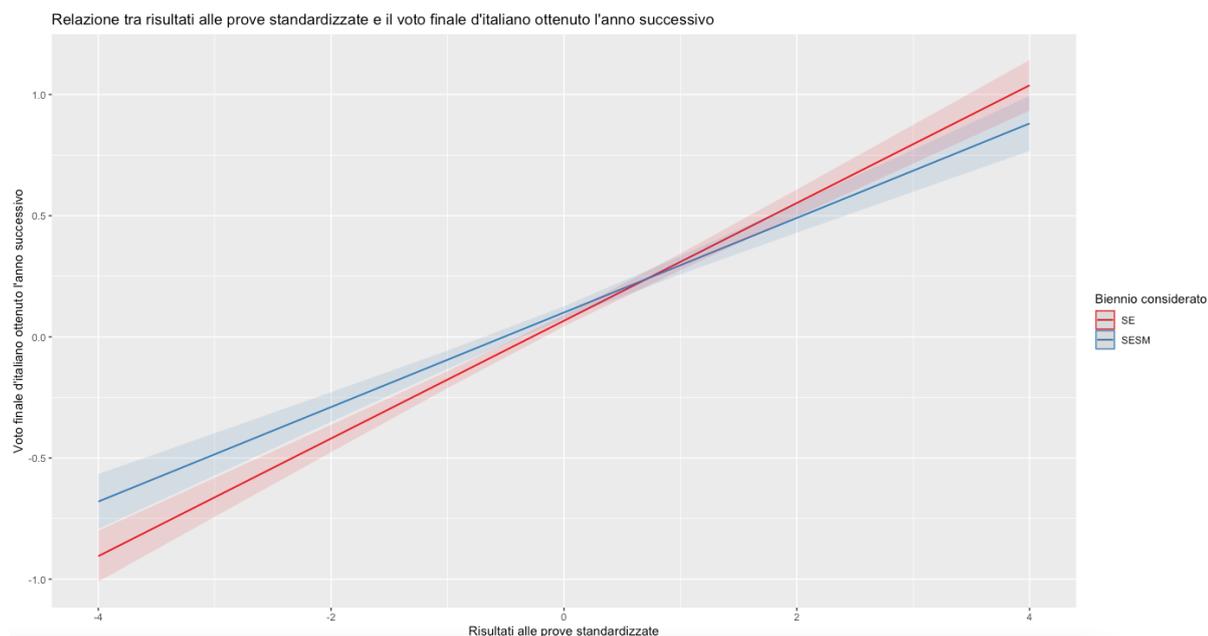
I risultati dell'analisi di regressione multipla (Figura 3) mostrano come più alto è il voto il voto finale d'italiano (ottenuto in terza elementare o in quinta elementare), più alto è il voto d'italiano ottenuto l'anno successivo (in quarta elementare o in prima media). In modo analogo, più alti sono i punteggi ottenuti nelle prove standardizzate (realizzate in terza elementare o in quinta elementare) e più alto sarà il voto d'italiano ottenuto l'anno successivo la somministrazione delle prove (in quarta elementare o in prima media). Va tuttavia sottolineato come la relazione tra i voti di italiano ottenuti in due anni successivi sia più forte (.62) rispetto a quella osservata tra i punteggi ottenuti nelle prove standardizzate ed il voto ottenuto l'anno successivo (.24).

Figura 3 Modello teorico della regressione lineare multipla e coefficienti standardizzati ottenuti.



Inoltre, le analisi permettono di evidenziare come la relazione tra i punteggi ottenuti nelle prove standardizzate e il voto d'italiano ottenuto l'anno successivo, sia più elevata se si prende in considerazione il biennio terza - quarta elementare comparativamente al biennio quinta elementare - prima media. In altre parole, la relazione tra i punteggi ottenuti nelle prove standardizzate in terza elementare e il voto d'italiano in quarta elementare (biennio SE-SE) è più forte della relazione tra i punteggi ottenuti alle prove standardizzate in quinta elementare ed il voto d'italiano in prima media (biennio SE-SM). Nella figura 4 è possibile osservare come la relazione (rappresentata dalla pendenza della retta di regressione) tra i punteggi nelle prove standardizzate e il voto scolastico ottenuto l'anno successivo sia più forte se si considera il biennio terza-quarta elementare in rosso (SE-SE) comparativamente al biennio quinta elementare-prima media in blu (SE-SM).

Figura 4 Rappresentazione grafica dell'interazione tra i risultati alle prove standardizzate e il biennio considerato sui voti ottenuti l'anno successivo (anno n+1).



I risultati di queste analisi indicano nel loro insieme che il miglior predittore dei risultati scolastici nell'anno n+1 è rappresentato dai voti ottenuti l'anno precedente. Ciò nonostante, anche i risultati alle prove standardizzate in italiano mostrano una relazione positiva con i voti ottenuti l'anno successivo. Inoltre, questa relazione risulta essere più forte quando i risultati alle prove standardizzate vengono messi in relazione con i voti ottenuti in classi dello stesso livello scolastico. Nello specifico, i risultati delle analisi realizzate indicano che le prove standardizzate predicano in modo leggermente più marcato i voti ottenuti l'anno successivo quando questo si situa all'interno di uno stesso livello educativo. Va sottolineato che le prove standardizzate, che nella loro costruzione si focalizzano su alcune competenze come indicato nel piano di studio, rilevano in modo parziale le competenze disciplinari degli allievi, e che i voti degli allievi, oltre alle valutazioni riguardanti le competenze disciplinari, si compongono anche di valutazioni maggiormente soggettive ma rilevanti per una valutazione globale dell'allievo (Crescentini, 2017a). In tal senso le due tipologie di valutazione possono ritenersi complementari al fine di permettere una valutazione strutturata dell'allievo nella quale di soggettività e oggettività si possono integrare (Crahay e Mottier Lopez, 2009).

## 4 Conclusioni

Il principale risultato delle analisi qui presentate è che i voti dipendono in gran parte dall'abilità dell'allievo nella disciplina considerata (è stata riscontrata una certa coerenza tra voti e i risultati nei test standardizzati), ma ci sono fattori extra cognitivi (il genere dell'allievo, l'origine sociale, la madrelingua, l'età e l'abilità media della classe nella disciplina) che presi singolarmente non sembrano esercitare un grande effetto sui voti, ma complessivamente incidono in maniera consistente su di essi. In particolare, si è evidenziato come la somma degli "effetti piccoli" arrivi a modificare il voto di fine anno fino a un punto, il che non è poco in un intervallo che normalmente alla SE è compreso tra 4.0 e 6.0.

Questa ricerca ha inoltre permesso di evidenziare come i punteggi nelle prove standardizzate abbiano un potere predittivo del voto scolastico anche negli anni successivi al loro svolgimento. Tuttavia, sono da sottolineare gli effetti contestuali, che indicano come la relazione tra i risultati nel test standardizzato e i voti ottenuti l'anno successivo allo svolgimento del test dipenda, tra le altre cose, dal livello scolastico considerato. Nello specifico, il potere predittivo dei punteggi delle prove standardizzate sui voti ottenuti l'anno successivo è più debole quando vengono messi in relazione i risultati alle prove standardizzate ottenuti a livello di scuola elementare (quinta) e i voti scolastici ottenuti a livello di scuola media l'anno successivo (prima) comparativamente a quanto osservato all'interno di uno stesso livello scolastico (terza-quarta elementare).

Trattandosi di uno studio quantitativo basato su dati secondari, è evidente che sono necessari ulteriori approfondimenti sulle modalità con cui vengono attribuiti i voti e sulle concezioni dei docenti della valutazione: difatti attualmente è in corso da parte del CIRSE una ricerca volta a indagare su questi aspetti. Appurato che allievi e allieve con una determinata origine sociale e una certa madrelingua, a parità di livello di abilità nella materia considerata, riportano voti significativamente più alti di altri, occorre capire che cosa ci sia alla base di questo differenziale di successo e addentrarsi dunque nel tema degli effetti terziari, ovvero del contributo della scuola alla riproduzione della disuguaglianza. Tale riflessione sociologica è particolarmente importante visto che si tratta di bambini e bambine della scuola elementare, le cui carriere educative e professionali possono essere plasmate dalle valutazioni riportate ai primi stadi del percorso scolastico. Inoltre, gli stessi effetti dei voti possono variare secondo le caratteristiche individuali, dal momento che, ad esempio, gruppi sociali diversi reagiscono in maniera diversa a successi e fallimenti e che le decisioni sugli investimenti educativi includono anche una valutazione dei loro costi, benefici e possibilità di successo (Breen e Goldthorpe, 1997)<sup>11</sup>.

Se è utopistico che la scuola diventi il grande equalizzatore delle condizioni di partenza degli individui auspicato da Horace Mann, è indubbio che la mobilità sociale passi da un'istruzione di qualità per tutti e, come sostengono Argentin e Pavolini (2020), per realizzare una società più equa è più facile intervenire sui meccanismi con cui le scuole e i docenti operano che cambiare l'allocazione delle risorse tra le famiglie degli allievi. La ricerca sociologica, indagando sulle logiche che sottendono l'attribuzione dei voti, può stimolare un processo riflessivo tra i docenti e contribuire a scardinare, anziché rafforzare, i meccanismi responsabili della riproduzione della disuguaglianza.

---

<sup>11</sup> Più specificamente, secondo Breen e Goldthorpe (1997) l'obiettivo di tutti i gruppi sociali è ridurre il rischio di retrocessione sociale ma, partendo da classi sociali diverse, hanno aspirazioni professionali diverse.







Figura C: modelli multilivello relativi al voto finale in italiano, a.s. 2020/21

	Modello 1 - Modello vuoto senza predittori. Solo pendenza aleatoria			Modello 2 - Modello 1 + Effetti fissi del livello 1 (allievi)			Modello 3 - Modello 2 + Effetti fissi del livello 2 (classi)		
	Coefficiente	S.E.	p	Coefficiente	S.E.	p	Coefficiente	S.E.	p
<b>Effetti fissi</b>									
Costante	5.130	0.016	***	5.038	0.020	***	5.051	0.022	***
Punteggio nel test standardizzato*				0.028	0.001	***	0.029	0.001	***
Professione dei genitori ( <i>rif. professioni mediamente qualificate</i> )									
Altamente qualificate				0.074	0.017	***	0.072	0.016	***
Scarsamente qualificate				-0.113	0.037	**	-0.118	0.037	***
Età ( <i>rif. nato/a nel 2010 o dopo</i> )									
<i>Nato/a prima del 2010</i>				-0.088	0.020	***	-0.088	0.020	***
Madrelingua ( <i>rif. italiana</i> )									
Altra lingua				-0.119	0.021	***	-0.118	0.020	***
Sesso ( <i>rif. maschio</i> )									
femmina				0.086	0.016	***	0.084	0.016	***
Nota di condotta*				0.164	0.019	***	0.167	0.019	***
Risultato medio della classe nel test standardizzato							-0.017	0.002	***
Tipo contratto docente ( <i>rif. docente titolare a tempo pieno</i> )							-0.066	0.034	*
Due docenti titolari a tempo parziale									
Sesso del docente ( <i>rif. femmina</i> )							-0.020	0.031	ns
Maschio									
<b>Effetto aleatorio</b>									
Varianza di livello 2 (classe)	0.015	0.005	***	0.026	0.004	***	0.018	0.003	***
4.29%									
Varianza di livello 1 (allievo/a)	0.337	0.011	***	0.119	0.004	***	0.119	0.004	***
<b>Misure di fit del modello</b>									
Devianza (-2 log V)	3'864.23			1'807.25			1'778.21		
AIC	3'868.23			1'811.25			1'782.21		
BIC	3'879.59			1'822.60			1'793.56		
ns non significatif. * p ≤ 0.05; ** p ≤ 0.01; *** p ≤ 0.001									
ICC = coefficient de corrélation intraclass									
* variabile centrata sulla media della popolazione ( <i>Grand-mean centering</i> )									

Figura D: Dimensione degli effetti dei predittori

Variabile*	f <sup>2</sup> di Cohen (effect size)			Interpretazione
	2014-15	2020-21		
	Matematica	Matematica	Italiano	
Punteggio nel test standardizzato	0.810	0.677	1.409	Grande
Professione dei genitori	0.012	0.007	0.017	Piccolo
Età	0.009	0.007	0.009	Piccolo
Nota di condotta	0.060	0.043	0.040	Piccolo
Sesso	–	0.009	0.010	Piccolo
Madrelingua	–	–	0.015	Piccolo
Punteggio medio della classe nel test standardizza	0.108	0.088	0.061	Piccolo
Tipo contratto docente	0.013	–	0.003	Piccolo

\* Solo per le variabili con effetto statisticamente significativo

Secondo i criteri proposti da Cohen (1988),  $f^2 \geq 0,02$ ,  $f^2 \geq 0,15$  et  $f^2 \geq 0,35$  costituiscono degli effetti rispettivamente dei piccoli, medi e grandi

Figura E: Media (M), deviazione standard (DS) e correlazioni tra i risultati alle prove standardizzate in italiano realizzate in terza elementare e i voti di terza e quarta elementare.

Variabile	M	DS	Punteggi Prova Standardizzata di italiano (Terza Elementare)	Voto in italiano (Terza Elementare)
1. Punteggi Prova Standardizzata di italiano	55.47	13.78		
2. Voto in italiano (Terza Elementare)	4.91	0.56	.66	
3. Voto in italiano (Quarta Elementare)	4.92	0.56	.65	.87

Nota. N = 2762

Figura F: Media (M), deviazione standard (DS) e correlazioni tra i risultati alle prove standardizzate in italiano realizzate in quinta elementare e i voti di quinta elementare e prima media.

Variabile	M	DS	Punteggi Prova Standardizzata di italiano (Quinta elementare)	Voto in italiano (Quinta Elementare)
1. Punteggi Prova Standardizzata di italiano	49.19	13.91		
2. Voto in italiano (Quinta Elementare)	5.03	0.58	.72	
3. Voto in italiano (Prima Media)	4.84	0.61	.63	.67

N = 2919.

Figura G: Analisi di regressione lineare multipla. Predizione del voto finale d'italiano nel corso dell'anno scolastico successivo alla realizzazione delle prove standardizzate d'italiano.

<i>Predittori</i>	<b>Voto anno n+1</b>		
	<i>Coefficiente standardizzato</i>	<i>Intervallo di fiducia al 95%</i>	<i>p</i>
(Intercetta)	0.18	0.12 – 0.24	<b>&lt;0.001</b>
Punteggi Prova Standardizzata di Italiano	0.24	0.22 – 0.27	<b>&lt;0.001</b>
Voto in italiano (anno n)	0.62	0.60 – 0.64	<b>&lt;0.001</b>
Biennio considerato	0.03	-0.00 – 0.07	0.087
Genere	-0.01	-0.04 – 0.01	0.325
Lingua madre	0.01	-0.02 – 0.05	0.469
Livello socioeconomico	-0.12	-0.14 – -0.09	<b>&lt;0.001</b>
Interazione: Punteggi Prova Standardizzata * Biennio considerato	-0.05	-0.08 – -0.02	<b>0.003</b>
N	4387		
R <sup>2</sup> / R <sup>2</sup> aggiustato	0.707 / 0.707		

Il livello socioeconomico è stato codificato come segue: 1= "Professioni altamente qualificate"; 2= "Professioni mediamente qualificate"; 3=" Professioni scarsamente qualificate"<sup>12</sup>. Per il genere: 0 = "Maschio"; 1="Femmina". Per la lingua madre: 1 = "Italiano"; 0 = "Altre lingue".

<sup>12</sup> L'origine socioeconomica delle allieve e degli allievi è stata calcolata utilizzando l'indice ISEI (International Socio-Economic Index of Occupational Status) basato sulla professione dei genitori e procedendo come già descritto nella nota 8.



## Riferimenti bibliografici

- Ackerman, P. L., e Heggestad, E. D. (1997). Intelligence, personality, and interests: Evidence for overlapping traits. *Psychological Bulletin*, 121(2), 219–245.
- Addimando, L., (2019). *Le pratiche didattiche nella Scuola dell'obbligo*. Locarno: Centro innovazione e ricerca sui sistemi educativi.
- Angelo, C. S. (2014). Is there a bias towards girls in non-anonymous evaluation? Online: <https://2014.economicsofeducation.com/user/pdfsesiones/166.pdf> (luglio 2023).
- Argentin, G. (2021). *Nostra scuola quotidiana. il cambiamento necessario*. Il Mulino.
- Argentin, G., Pavolini, E. (2020). How Schools Directly Contribute to the Reproduction of Social Inequalities. Evidence of Tertiary Effects, Taken from Italian Research. *Politiche Sociali*, 1, 149-176.
- Argentin, G., e Triventi, M. (2015). The north-south divide in school grading standards: new evidence from national assessments of the Italian student population. *Italian Journal of Sociology of Education*, 7(2), 157-185. Armand Colin.
- Atkinson, R. C., & Geiser, S. (2009). Reflections on a century of college admissions tests. *Educational Researcher*, 38(9), 665-676.
- Bandiera, O., Larcinese, V., & Rasul, I. (2015). Blissful ignorance? A natural experiment on the effect of feedback on students' performance. *Labour Economics*, 34, 13-25.
- Behaghel, L., Crépon, B., e Le Barbanchon, T. (2015). Unintended effects of anonymous resumes. *American Economic Journal: Applied Economics*, 7(3), 1-27.
- Benadusi, L., e Giancola, O. (2021). *Equità e merito nella scuola: teorie, indagini empiriche, politiche*. Franco Angeli.
- Betts, J. R., e Grogger, J. (2003). The impact of grading standards on student achievement, educational attainment, and entry-level earnings. *Economics of Education Review*, 22(4), 343-352.
- Boarelli, M. (2019). *Contro l'ideologia del merito*. Gius. Laterza e Figli Spa.
- Bonesrønning, H. (1999). The variation in teachers' grading practices: Causes and consequences. *Economics of Education Review*, 18(1), 89-106.
- Boudon, R. (1973). *L'inégalité des chances. La mobilité sociale dans les sociétés industrielles*. Paris:
- Bourdieu, P. (1998). *The state nobility: Elite schools in the field of power*. Stanford University Press.
- Bourdieu, P., & Passeron, J. C. (1990). *Reproduction in education, society and culture* (Vol. 4). Sage.
- Bowers, A. J. (2011). What's in a Grade? The Multidimensional Nature of What Teacher-assigned Grades Assess in High School. *Educational Research and Evaluation*, 17(3), 141-159.
- Bowles, S., e Gintis, H. (2011). *Schooling in capitalist America: Educational reform and the contradictions of economic life*. Haymarket Books.
- Breen, R. and Goldthorpe, J. H. (1997). Explaining educational differentials towards a formal rational action theory. *Rationality and society*, 9(3), 275-305.
- Bressoux, P., e Pansu, P. (2003). *Quand les enseignants jugent leurs élèves*. Presses universitaires de France.
- Brookhart, S. M. (1991). Grading practices and validity. *Educational Measurement: Issues and Practice*, 10(1), 35–36.
- Brookhart, S. M. (2015). Graded achievement, tested achievement, and validity. *Educational Assessment*, 20(4), 268-296.

Burgess, S., & Greaves, E. (2009). *Test scores, subjective assessment and stereotyping of ethnic minorities*. Centre for Market and Public Organisation, University of Bristol. <http://www.bristol.ac.uk/media-library/sites/cmpo/migrated/documents/wp221.pdf>.

Butera, F. (2006). La meritocrazia a scuola: Un serio ostacolo all'apprendimento. *Psicologia sociale*, 1(3), 431-448.

Calarco, J. M. (2014). Coached for the classroom: Parents' cultural transmission and children's reproduction of educational inequalities. *American Sociological Review*, 79(5), 1015-1037.

Calarco, J. M. (2018). *Negotiating opportunities: How the middle class secures advantages in school*. Oxford University Press.

Calarco, J. M. (2018). *Negotiating opportunities: How the middle class secures advantages in school*. Oxford University Press.

Cantone, T. I. (2015). *Piano di studio della scuola dell'obbligo ticinese*. Lugano: Società d'arti grafiche Veladini.

Carlana, J. M. (2019). Implicit Stereotypes: Evidence from Teachers' Gender Bias. *The Quarterly Journal of Economics*, 134(3), 1163-1224.

Chamorro-Premuzic, T., e Furnham, A. (2014). *Personality and intellectual competence*. Psychology Press.

Cliffordson, C. (2008). Differential prediction of study success across academic programs in the Swedish context: The validity of grades and tests as selection instruments for higher education. *Educational Assessment*, 13(1), 56-75.

Collier, P. (2018). *The future of capitalism: Facing the new anxieties*. New York: Harper.

Conrad, N., e Patry, M. W. (2012). Conscientiousness and academic performance: A mediational analysis. *International Journal for the Scholarship of Teaching and Learning*, 6(1), n1.

Consiglio di Stato della Repubblica e Cantone Ticino. (30 maggio 2018). Regolamento della scuola media. <https://m3.ti.ch/CAN/RLeggi/public/index.php/raccolta-leggi/legge/num/646>

Cornwell, C., Mustard, D. B., e Van Parys, J. (2013). Noncognitive Skills and the Gender Disparities in Test Scores and Teacher Assessments. *Journal of Human resources*, 48(1), 236-264.

Crahay, M., & Mottier Lopez, L. (2009). *Évaluations en tension. Entre la régulation des apprentissages et le pilotage des systèmes*. Bruxelles: De Boeck.

Crescentini, A. (2017a). *Prove standardizzate ticinesi. Italiano nella classe III Scuola Elementare*. Locarno: Centro innovazione e ricerca sui sistemi educativi.

Crescentini, A. (2017b). *Prove standardizzate ticinesi. Matematica nella classe V Scuola Elementare*. Locarno: Centro innovazione e ricerca sui sistemi educativi.

Crescentini, A., Salvisberg, M. & Zanolla, G. (2014). *Prove standardizzate di Matematica per la SE*. Locarno: Centro Innovazione e Ricerca sui Sistemi Educativi (CIRSE).

Cross, L. H., & Frary, R. B. (1999). Hodgepodge grading: Endorsed by students and teachers alike. *Applied measurement in Education*, 12(1), 53-72.

Crotta, F., Salvisberg, M., & Zanolla, G. (2020). Note scolastiche e competenze fondamentali in matematica degli allievi dell'11 anno HarmoS in Ticino. *Schweizerische Zeitschrift für Bildungswissenschaften*, 42(2), 323-342.

Curran, T., e Hill, A. P. (2019). Perfectionism is increasing over time: A meta-analysis of birth cohort differences from 1989 to 2016. *Psychological bulletin*, 145(4), 410.

- DECS, D.S. (2015). *Piano di studi della scuola dell'obbligo ticinese*. Lugano: Società d'arti grafiche già Veladini e co SA.
- Dee, T. S. (2005). A teacher like me: Does race, ethnicity, or gender matter? *American Economic Review*, 95(2), 158-165.
- Dee, T. S. (2007). Teachers and the gender gaps in student achievement. *Journal of Human Resources*, 42(3), 528-554.
- De Landsheere, V., e De Landsheere, G. (1992). *L'éducation et la formation: science et pratique*. Presses universitaires de France, Paris, France.
- De Vries, J. M., Rathmann, K., e Gebhardt, M. (2018). How does social behavior relate to both grades and achievement scores? *Frontiers in Psychology*, 9, 857.
- Dian, M., e Triventi, M. (2021). The weight of school grades: evidence of biased teachers' evaluations against overweight students in Germany. *Plos one*, 16(2), e0245972.
- Doepke, M., & Zilibotti, F. (2019). *Love, money, and parenting: How economics explains the way we raise our kids*. Princeton University Press.
- Domenici, G. (2001). *Manuale della valutazione scolastica*, Roma-Bari, Laterza.
- Duckworth, A.L. e Seligman, M. E. (2006). Self-discipline Gives Girls the Edge: Gender in Self-discipline, Grades, and Achievement Test Scores. *Journal of educational psychology*, 98(1), 198.
- Duru-Bellat, M. (2009). *Le mérite contre la justice*. Lectures, Les livres.
- Enzi, B. (2015). *Gender differentials in test scores and teacher assessments: Evidence from Germany*. Working Paper. Online: [http://www.edge-page.net/jamb2014/papers/Enzi% 20-% 20Draft. Pdf](http://www.edge-page.net/jamb2014/papers/Enzi%20-%20Draft.Pdf) (luglio 2023).
- Fallan, L., e Opstad, L. (2012). Attitudes towards Study Effort Response to Higher Grading Standards: Do Gender and Personality Distinctions Matter? *Journal of education and learning*, 1(2), 179-187.
- Farrington, C. A., Roderick, M., Allensworth, E., Nagaoka, J., Keyes, T. S., Johnson, D. W., e Beechum, N. O. (2012). Teaching Adolescents to Become Learners: The Role of Noncognitive Factors in Shaping School Performance--A Critical Literature Review. Consortium on Chicago School Research. Online: <https://files.eric.ed.gov/fulltext/ED542543.pdf> (luglio 2023).
- Finefter-Rosenbluh, I., & Levinson, M. (2015). What is wrong with grade inflation (if anything)? *Philosophical Inquiry in Education*, 23(1), 3-21.
- Florin, T. A., Shults, J., e Stettler, N. (2011). Perception of overweight is associated with poor academic performance in US adolescents. *Journal of School Health*, 81(11), 663-670.
- Fox, A. (1956). Class and equality. *Socialist Commentary*, 1113.
- Fryer Jr, R. G., e Levitt, S. D. (2004). Understanding the black-white test score gap in the first two years of school. *Review of economics and statistics*, 86(2), 447-464.
- Geiser, S., & Santelices, M. V. (2007). *Validity of high-school grades in predicting student success beyond the freshman year: High-school record vs. standardized tests as indicators of four-year college outcomes*. <https://cloudfront.escholarship.org/dist/prd/content/qt7306z0zf/qt7306z0zf.pdf>.
- Gerbino, M., Zuffianò, A., Eisenberg, N., Castellani, V., Luengo Kanacri, B. P., Pastorelli, C., et al. (2018). Adolescents' prosocial behavior predicts good grades beyond intelligence and personality traits. *Journal of Personality*, 86, 247-260.
- Giberti, C. (2019). Differenze di Genere in Matematica: dagli Studi Internazionali alla Situazione Italiana. *Didattica della matematica*, (5), 44-69.

Gillen-O'Neel, C., Ruble, D. N., e Fuligni, A. J. (2011). Ethnic stigma, academic anxiety, and intrinsic motivation in middle childhood. *Child development*, 82(5), 1470-1485.

Guskey, T. R. (2009). Bound by Tradition: Teachers' Views of Crucial Grading and Reporting Issues. Online: <https://files.eric.ed.gov/fulltext/ED509342.pdf> (luglio, 2023).

Guskey, T. R., e Link, L. J. (2019). Exploring the factors teachers consider in determining students' grades. *Assessment in Education: Principles, Policy e Practice*, 26(3), 303-320.

Helbig, M. e Morar, T. (2018). Why Teachers Assign Socially Unequal Marks.: A Case for Establishing Tertiary Origin Effects in the Model of Primary and Secondary Effects of Educational Research. *International Journal of Learning, Teaching and Educational Research*, 17(7), 1-25.

Hinnerich, B. T., Höglin, E., e Johannesson, M. (2011). Are boys discriminated in Swedish high schools? *Economics of Education review*, 30(4), 682-690.

Hjorth-Trolle, A., Rosenqvist, E., e Hed, A. (2022). Grading Practices and the Social Gradient in GPA: Quasi-Experimental Evidence from Sweden. *European Sociological Review*, 38(3), 455-471.

Iacus, S. M., e Porro, G. (2011). Teachers' evaluations and students' achievement: a 'deviation from the reference' analysis. *Education Economics*, 19(2), 139-159.

Jæger, M. M., e Breen, R. (2016). A dynamic model of cultural reproduction. *American Journal of Sociology*, 121(4), 1079-1115.

Jonsson J. O. (1992). *Toward the merit selective society?* Swedish institute for social research, University of Stockholm.

Jonsson, J. O., e Rudolphi, F. (2011). Weak performance—strong determination: school achievement and educational choice among children of immigrants in Sweden. *European Sociological Review*, 27(4), 487-508.

Jussim, L., e Harber, K. D. (2005). Teacher expectations and self-fulfilling prophecies: Knowns and unknowns, resolved and unresolved controversies. *Personality and social psychology review*, 9(2), 131-155.

Kendall, P., e Di Pietro, M. (1995). *Terapia scolastica dell'ansia. Guida per psicologi e insegnanti* (Vol. 7). Edizioni Erickson.

Kiss, D. (2013). Are immigrants and girls graded worse? Results of a matching approach. *Education Economics*, 21(5), 447-463.

Kiter Edwards, M. L. (2004). We're Decent People: Constructing and Managing Family Identity in Rural Working-Class Communities. *Journal of Marriage and Family*, 66(2), 515-529.

Koretz, D. M. (2008). *Measuring up*. Harvard University Press.

Kuncel, N. R., e Hezlett, S. A. (2007). Standardized Tests Predict Graduate Students' Success. *Science*, 315(5815), 1080–1081. <https://doi.org/10.1126/science.1136618>

Jackson, M. and J.O. Jonsson, (2013), *Why does inequality of educational opportunity vary across countries? Primary and secondary effects in comparative context*, in Jackson, M. (ed.), *Determined to succeed? Performance versus choice in educational attainment*. California, Stanford University Press.

Laidra, K., Pullmann, H., e Allik, J. (2007). Personality and intelligence as predictors of academic achievement: A cross-sectional study from elementary to secondary school. *Personality and Individual Differences*, 42(3), 441–451.

Lareau, A. (2000). *Home advantage: Social class and parental intervention in elementary education*. Rowman & Littlefield Publishers.

Lareau, A. (2011). *Unequal childhoods: Class, race, and family life*. Univ of California Press.

Lareau, A., Weininger, E. B., Conley, D., e Velez, M. (2011). Unequal childhoods in context: Results from a quantitative analysis. *Unequal childhoods: Class, race, and family life*, 333-41.

Latner, J. D., O'Brien, K. S., Durso, L. E., Brinkman, L. A., e MacDonald, T. (2008). Weighing obesity stigma: the relative strength of different forms of bias. *International Journal of Obesity*, 32(7), 1145-1152.

Lechner, C., Danner, D., e Rammstedt, B. (2017). How is personality related to intelligence and achievement? A replication and extension of Borghans et al. and Salkever. *Personality and Individual Differences*, 111, 86–91.

Lekholm, A. K., & Cliffordson, C. (2008). Discrepancies between school grades and test scores at individual and school level: effects of gender and family background. *Educational Research and Evaluation*, 14(2), 181-199.

Lekholm, A. K., & Cliffordson, C. (2009). Effects of student characteristics on grades in compulsory school. *Educational Research and Evaluation*, 15(1), 1-23.

Leclercq, D., Nicaise, J., e Demeuse, M. (2004). Docimologie critique: des difficultés de noter des copies et d'attribuer des notes aux élèves. In M. Demeuse (a cura di), *Introduction aux théories et aux méthodes de la mesure en sciences psychologiques et en sciences de l'éducation* (pp. 273- 292). Liège: Editions de l'Université de Liège.

Lievore, I., e Triventi, M. (2022). Do teacher and classroom characteristics affect the way in which girls and boys are graded? A multilevel analysis of student–teacher matched data. *British Journal of Sociology of Education*, 44(1), 97-122.

Lindahl, E. (2007). *Comparing teachers' assessments and national test results: Evidence from Sweden* (No. 2007: 24). Working Paper, IFAU-Institute for Labour Market Policy Evaluation, Svezia.

Major, L. E., e Machin, S. (2018). *Social mobility: And its enemies*. Penguin UK.

Marcenaro-Gutierrez, O., & Vignoles, A. (2015). A comparison of teacher and test-based assessment for Spanish primary and secondary students. *Educational Research*, 57(1), 1-21.

Markovits, D. (2019). *The meritocracy trap*. Penguin UK.

Mayer, S. E. (1997). *What money can't buy: Family income and children's life chances*. Harvard University Press.

Meier, E., Crescentini, A., e Pettignano, M. (2022). *Prove standardizzate cantonali: Matematica e Italiano in V SE. Dipartimento Formazione e Apprendimento*. <https://epic-almeida-fbd15f.netlify.app/intro.html>

Merle, P. (2018). *Les pratiques d'évaluation scolaire: historique, difficultés, perspectives*. PUF.

Mottier Lopez, L. & Crahay, M. (Ed.) (2009). *Évaluations en tension. Entre la régulation des apprentissages et le pilotage des systèmes*. Bruxelles: De Boek.

Nelson, M. K. (2010). *Parenting out of control: Anxious parents in uncertain times*. NYU Press.

OECD (2019). *PISA 2018: Insights and interpretation*. OECD Publishing.

Pattison, E., Grodsky, E., & Muller, C. (2013). Is the sky falling? Grade inflation and the signaling power of grades. *Educational Researcher*, 42(5), 259-265.

Perander, K., Londen, M., e Holm, G. (2020). Anxious Girls and Laid-back Boys: Teachers' and Study Counsellors' Gendered Perceptions of Students. *Cambridge Journal of Education*, 50(2), 185-199.

Petrucci, F. e Zanolla, G. (2023). *Au-delà de la Réalité Objective des Performances Scolaires: les Biais de la Notation à l'École Primaire dans les Cantons du Tessin et de Genève*. In: Congrès SSRE/SSFE 2023 - Une époque en mutation ? L'éducation entre continuité et changement, 28-30.06.2023, Zürich.

Piketty, T. (2019). *Capital et idéologie*. Éditions du Seuil.

- Puhl, R.M. e Heuer, C.A. (2009). The stigma of obesity a review and update. *Obesity*, 17, 941-964.
- Randall, J., e Engelhard, G. (2009). Examining teacher grades using Rasch measurement theory. *Journal of Educational Measurement*, 46, 1–18.
- Randall, J., e Engelhard, G. (2010). Examining the grading practices of teachers. *Teaching and Teacher Education*, 26, 1372–1380.
- Raudenbush, S. W., e Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (Vol. 1). Sage.
- Ridgeway, C.L., Smith-Lovin, L. (1999): The Gender System and Interaction, in *Annual Review of Sociology*, 25:191–216.
- Rinaldi, A. (2023). *Le signore non parlano di soldi*. Rizzoli.
- Ritts, V., Patterson, M. L., e Tubbs, M. E. (1992). Expectations, impressions, and judgments of physically attractive students: A review. *Review of educational research*, 62(4), 413-426.
- Romito, M. (2016). *Una scuola di classe. Orientamento e disuguaglianza nelle transizioni scolastiche* (pp. 1-281). Guerini e Associati.
- Rubie-Davies, C.M. (2018): *Teacher Expectations in Education*, New York, NY: Routledge.
- Salvisberg, M., e Zampieri, S.(2014). *Valutazioni sotto esame. Piste esplorative per un confronto tra PISA e note scolastiche 2009 e 2012*. Locarno: Centro Innovazione e Ricerca sui Sistemi Educativi
- Sandel, M. J. (2020). *The tyranny of merit: What's become of the common good?* Penguin UK.
- Salvisberg, M., & Zampieri, S. (2015). *Valutazioni sotto esame. Piste esplorative per un confronto tra PISA e note scolastiche 2009 e 2012*. Centro Innovazione e Ricerca sui Sistemi educativi.
- Sarrazin, P. e Trouilloud, D. (2011). *Les attentes des enseignants : une menace pour l'évaluation ? L'évaluation, une menace?* Presses Universitaires de France.
- Stiggins, R. J. (2002). Assessment crisis: The absence of assessment FOR learning. *Phi Delta Kappan*, 83(10), 758–765.
- Swami, V., Frederick, D. A., Aavik, T., Alcalay, L., Allik, J., Anderson, D., ... e Zivcic-Becirevic, I. (2010). The attractive female body weight and female body dissatisfaction in 26 countries across 10 world regions: Results of the International Body Project I. *Personality and social psychology bulletin*, 36(3), 309-325.
- Tenret, E. (2014). La meritocrazia a scuola, attraverso la scuola. Il caso italiano in prospettiva comparata. *Quaderni di sociologia*, (64), 45-71.
- Terrier, C. (2015). *Giving a little help to girls? Evidence on grade discrimination and its effect on students' achievement*. CEP Discussion Paper N. 1341. Centre for Economic Performance, London School of Economics and Political Science. Online: <https://eprints.lse.ac.uk/61696/1/dp1341.pdf>
- Terrier, C. (2020). Boys Lag behind: How teachers' Gender Biases Affect Student Achievement. *Economics of Education Review*, 77, 101981
- Tiedemann, J. (2000). Gender-related beliefs of teachers in elementary school mathematics. *Educational studies in Mathematics*, 41(2), 191-207.
- Tobisch, A., e Dresel, M. (2017). Negatively or positively biased? Dependencies of teachers' judgments and expectations based on students' ethnic and social backgrounds. *Social Psychology of Education*, 20, 731-752.
- Todd, P. E., e Wolpin, K. I. (2007). The production of cognitive achievement in children: Home, school, and racial test score gaps. *Journal of Human capital*, 1(1), 91-136.

Tomasetto, C. (2019). Gender Stereotypes, Anxiety, and Math Outcomes in Adults and Children. In Mammarella, I. C., e Caviola, S. and Dowker, A. (Eds.). *Mathematics Anxiety: What Is Known, and What is Still Missing*. Routledge.

Triventi, M. (2020). Are children of immigrants graded less generously by their teachers than natives, and why? Evidence from student population data in Italy. *International Migration Review*, 54(3), 765-795.

Triventi, M., Lievore, I., e Dian, M. (2023). *Why are girls better graded in school? The role of students' personality traits, school-related attitudes, and behavior*. Online: <https://osf.io/erydh/download> (luglio 2023).

Van Houtte, M. (2020). Understanding the Gender Gap in School (Dis) engagement from Three Gender Dimensions: The Individual, the Interactional and the Institutional. *Educational Studies*, 1-19.

Vecchione, M., Alessandri, G., e Marsicano, G. (2014). Academic motivation predicts educational attainment: Does gender make a difference? *Learning and Individual Differences*, 32, 124-131.

Voyer, D., e Voyer, S. D. (2014). Gender differences in scholastic achievement: a metaanalysis. *Psychological Bulletin*, 140(4), 1174-1204.

Wang, S. S., Brownell, K. D., e Wadden, T. A. (2004). The influence of the stigma of obesity on overweight individuals. *International journal of obesity*, 28(10), 1333-1337.

Wiberg, M. (2019). The relationship between TIMSS mathematics achievements, grades, and national test scores. *Education Inquiry*, 10(4), 328-343.

Wilkinson, R., e Pickett, K. (2019). *The inner level: How more equal societies reduce stress, restore sanity and improve everyone's well-being*. Penguin UK.

Wilson, S. M., Smith, A. W., e Wildman, B. G. (2015). Teachers' perceptions of youth with obesity in the classroom. *Advances in School Mental Health Promotion*, 8(4), 231-243.

Wooldridge, A. (2021). *The aristocracy of talent: How meritocracy made the modern world*. Penguin UK.

Yeung, W. J., Linver, M. R., e Brooks-Gunn, J. (2002). How money matters for young children's development: Parental investment and family processes. *Child development*, 73(6), 1861-1879.

Young, M. (1958). *The Rise of the Meritocracy*. London: Pelican Books.

Zanga, G., e De Gioannis, E. (2023). Discrimination in grading: A scoping review of studies on teachers' discrimination in school. *Studies in Educational Evaluation*, 78, 101284.

Zanolla, G. (2014). Why do children differ in mathematical competencies? The experience of a standardized test in the primary school of Canton Ticino, Switzerland. *International Journal of Teaching and Education*, 2 (3), 155-167.

Zanolla, G. (2018). The gender gap in math. Evidences of a study in the primary school in the Swiss canton of Ticino. *International Journal of Teaching and Education*, 6 (1), 103-125.

Zanolla, G. (2023). Equità. In Castelli, L. e Plata, A. (a cura di), *Scuola a tutto campo*, SUPSI.

Zuffianò, A., Alessandri, G., Gerbino, M., Luengo Kanacri, B. P., Di Giunta, L., Milioni, M., et al. (2013). Academic achievement: The unique contribution of self-efficacy beliefs in self-regulating learning beyond intelligence, personality traits, and self-esteem. *Learning and Individual Differences*, 23(1), 158-162.

Zwick, R., & Greif Green, J. (2007). New perspectives on the correlation of SAT scores, high school grades, and socioeconomic factors. *Journal of Educational Measurement*, 44(1), pp. 23-45.





Repubblica e Cantone Ticino  
Dipartimento dell'educazione, della cultura  
e dello sport

Scuola universitaria professionale  
della Svizzera italiana

Che cosa c'è dietro a un voto?

Un'analisi multilivello applicata alle valutazioni  
di fine anno d'italiano e matematica nella  
scuola elementare

Quaderni di ricerca – n. 47

Scuola universitaria professionale della Svizzera italiana  
Dipartimento formazione e apprendimento / Alta scuola  
pedagogica  
Centro competenze innovazione e ricerca sui sistemi  
educativi

Piazza San Francesco 19  
6600 Locarno

ISBN (print) 978-88-85585-87-4  
ISBN (online) 978-88-85585-89-  
8

