

L'introduzione della Prova Nazionale nell'Esame di Stato al termine del I ciclo d'istruzione: i primi effetti sugli insegnanti di italiano

Massimo Marcuccio

Università degli Studi di Parma, Dipartimento di Scienze della Formazione e del Territorio

massimo.marcuccio@unipr.it

THE NATIONAL TEST IN THE FINAL EXAMINATION OF THE LOWER SECONDARY SCHOOL: INITIAL EFFECTS ON TEACHERS OF ITALIAN

ABSTRACT

The article presents the results of a study aimed at describing the initial effects of the National Test (in the Final Examination of lower secondary schools) on teachers of Italian. The study lies within the research line on the use of evaluation and employs an adaptation of the theory of influence by Mark and Henry (2004) as a theoretical framework. The opinions and statements of practices of 41 teachers were collected through eight focus groups conducted in eight lower secondary schools of a city in northern Italy. The analysis of transcripts was performed with a quanti-qualitative approach which identified eleven areas of theme concerning the importance and procedures of the National Test, the teachers' activities related to it and the students' reactions. The analysis of the thematic contexts shows that the National Test has caused some changes in teachers' opinions and practices, but not all the possible ones. This is due to teachers' resistance to the pedagogical assumptions of the National Test. The conclusions discuss the heuristic relevance of the theoretical framework and some proposals for the development of research with regard to the impact of evaluations.

Keywords: Evaluation use, Final Examination (lower secondary school), National Test, School system evaluation, Teachers' beliefs.

1. IL CONTESTO DELLA RICERCA

1.1. *La valutazione del sistema scolastico italiano*

Il Servizio Nazionale di Valutazione del sistema educativo di istruzione e formazione viene istituito in Italia con il Dlgs. 286/2004 e affidato alla gestione dell'INVALSI, l'Istituto Nazionale di Valutazione del Sistema Educativo di Istruzione e Formazione. A seguito del mandato ricevuto, e dopo le esperienze di tre Progetti Pilota (PP1, PP2, PP3), l'INVALSI avvia dall'anno scolastico 2004/2005 una Rilevazione Nazionale del Sistema Istruzione che prevede sia la valutazione degli apprendimenti degli studenti, mediante la somministrazione di prove strutturate (italiano, matematica e scienze), sia la valutazione della qualità complessiva dell'offerta formativa degli istituti scolastici per mezzo di un questionario di sistema. Sono coinvolte in questa prima rilevazione la scuola primaria (classi II e IV), la secondaria di I grado (classi I) e la scuola secondaria di II grado (classi I e III)¹.

A partire dall'a.s. 2007/2008 la rilevazione si estende alle classi III della secondaria di I grado con l'inserimento di una Prova Nazionale (PN) nell'Esame di Stato a conclusione del I ciclo di istruzione (Legge 176/2007 e Direttiva 16/2008). La prova, predisposta dall'INVALSI, intende verificare i livelli di apprendimento degli studenti in italiano e matematica.

Nel giugno 2008 e 2009 i criteri per definire l'incidenza e il peso della PN sulla valutazione finale dell'Esame di Stato sono a discrezione delle sottocommissioni d'esame. Dal 2010, invece, il MIUR sancisce che l'esito della PN concorra all'esito dell'Esame di Stato al pari degli esiti delle prove scritte e orali e del giudizio di idoneità all'ammissione; il voto finale viene elaborato mediante il calcolo della media dei voti in decimi ottenuti nelle singole prove e nel giudizio di idoneità (DPR 122/2009).

In questo contributo focalizzeremo la nostra attenzione sugli effetti delle PN di comprensione del testo somministrate prima di quest'ultima innovazione.

1.2. *La PN di italiano nell'Esame di Stato al termine del I ciclo*

La PN di italiano è stata costruita a partire da un Quadro di riferimento teorico elaborato in base alle *Indicazioni Nazionali* del 2003, alle *Indicazioni*

¹ Per il dettaglio dei cambiamenti intercorsi nel tempo circa le classi coinvolte, le discipline indagate e i diversi strumenti utilizzati (prove e questionari) si rimanda alle pagine del sito dell'INVALSI (<http://www.invalsi.it>) riservate al Servizio Nazionale di Valutazione.

per il curriculum del 2007 e ad alcune ricerche internazionali (PIRLS, 2006 e PISA, 2006).

La struttura della prova, dal 2008 all'anno successivo, ha subito alcune trasformazioni (Tabella 1).

La PN 2008 era articolata in due parti: la prima di comprensione di un testo narrativo, la seconda di riflessione sulla lingua con domande di grammatica. I quesiti somministrati erano a scelta multipla, con quattro alternative di risposta, e a risposta aperta univoca (INVALSI, 2008). Anche la PN 2009 era strutturata in due parti (comprensione e riflessione sulla lingua) ma i testi per la comprensione erano due (narrativo ed espositivo) e i quesiti somministrati erano solo a scelta multipla, sempre con quattro alternative di risposta (INVALSI, 2009). Gli ambiti misurati dalle PN 2008 e 2009 sono riportati nella Tabella 2.

In tutti e due gli anni, la medesima PN è stata somministrata agli studenti di cittadinanza sia italiana sia straniera. Sul sito dell'INVALSI erano disponibili esempi di prove che le scuole potevano utilizzare a propria discrezione. Spettava alle sottocommissioni d'esame predisporre prove specifiche per gli alunni con disabilità intellettiva: una raccolta di suggerimenti ed esempi elaborati dalle scuole erano a disposizione sul sito dell'INVALSI. Per gli alunni con diagnosi di dislessia o di altri DSA è stata messa a disposizione una prova in versione informatizzata così come una prova in formato *Braille* e in formati ingranditi per gli alunni con disabilità visiva.

Il tempo a disposizione per lo svolgimento della PN di italiano è stato di un'ora in entrambi gli anni. Nel 2008 è stato somministrato prima il fascicolo di matematica e poi quello di italiano; nel 2009 è avvenuto il contrario ma sempre con una pausa di 15 minuti tra un fascicolo e l'altro. Gli alunni con una diagnosi di dislessia o di altri DSA avevano a disposizione gli strumenti compensativi e un tempo maggiore così come gli alunni con disabilità visiva.

La somministrazione e la sorveglianza erano affidate a insegnanti della scuola ma non della disciplina oggetto di valutazione. Ai somministratori era consentito esclusivamente fornire indicazioni sulle modalità di risposta e sul tempo rimanente; nulla poteva essere detto circa i contenuti dei quesiti.

La sottocommissione d'esame, nel 2008, utilizzando una griglia predisposta dall'INVALSI, aveva il compito di correggere la PN registrando su una scheda predisposta per il riconoscimento ottico le risposte degli studenti o il punteggio grezzo attribuito nel caso delle risposte univoche aperte. Per la PN 2009 l'INVALSI introduce un duplice criterio per identificare la soglia della sufficienza (60/100) e fornisce alle commissioni d'esame un prospetto per l'attribuzione dei punteggi degli studenti che si collocano al sotto o al di sopra di tale soglia. Le schede di correzione compilate andavano inviate all'INVALSI.

Tabella 1. – Numero di quesiti per tipo di esercizio e anno della PN.

TIPO DI ESERCIZIO	ANNO DELLA PN	
	2008	2009
Comprensione del testo (narrativo)	15	17
Comprensione del testo (espositivo)	–	13
Riflessione sulla lingua	10	10
TOTALE QUESITI	25	40

Tabella 2. – Numero di quesiti della sezione «comprensione del testo» della PN per ambito, anno e tipo di testo.

AMBITO		Competenza lessicale	Comprensione locale del testo	Organizzazione logica entro e oltre la frase	Aspetti formali e retorici	Comprensione globale del testo	TOTALE QUESITI
ANNO	E TIPO DI TESTO						
2008	Testo narrativo	1	8*	2	1	3*	15
	Testo narrativo	2	7	4	1	3	17
2009	Testo espositivo	3	5	3	–	2	13
	TOTALE QUESITI	5	12	7	1	5	30

* La distinzione tra ambito di comprensione «locale» e «globale» del testo è stata effettuata da noi in base all'analisi dei quesiti poiché la PN 2008 conteneva item che misuravano complessivamente la «comprensione locale e globale del testo» (INVALSI, 2008).

Era inoltre previsto che l'INVALSI restituisse gli esiti in due fasi: dapprima, con la diffusione di un rapporto con i dati campionari a livello nazionale e regionale alla fine dell'estate; in un secondo momento mettendo a disposizione online i dati delle singole scuole nei primi mesi dell'anno scolastico successivo.

1.3. La situazione problematica

L'obiettivo della *valutazione esterna* dell'INVALSI mediante la PN è, da un lato, rilevare i livelli di apprendimento degli studenti, dall'altro, favorire «il progressivo allineamento a standard di carattere nazionale» (Dir. 16/2008). È insito, quindi, nella PN un intento trasformativo del sistema scolastico che, tuttavia, presenta caratteristiche non funzionali al raggiungimento di tale scopo.

Infatti, nella seconda indagine IARD sulle condizioni di vita e di lavoro degli insegnanti italiani (Cavalli, 2000), solo il 22,5% degli insegnanti della secondaria di I grado era d'accordo con l'«imposizione di strumenti di valutazione esterna» per garantire l'omogeneità nella qualità dell'istruzione. Nella medesima ricerca soltanto il 37% degli insegnanti della media inferiore statale dichiara di utilizzare regolarmente test di profitto costruiti dallo stesso docente o da colleghi o tratti dai libri di testo, mentre una percentuale ancora più bassa (3,2%) utilizza in modo regolare test di profitto forniti da istituti specializzati. Nella terza indagine IARD (Cavalli & Argentin, 2010), l'80% degli insegnanti afferma di utilizzare in modo regolare o saltuario test costruiti direttamente dal docente e dai colleghi o suggeriti dai libri di testo. L'uso delle prove strutturate fornite da agenzie specializzate (tra cui l'INVALSI) viene effettuato regolarmente solo dal 5,2%, saltuariamente dal 41,6%. Tenendo presente che l'espressione «test costruiti in modo autonomo o da colleghi o prelevati da libri di testo» fa riferimento in larga parte a strumenti tradizionali (esercizi e/o problemi), la conclusione è che i docenti italiani abbiano una preferenza per «metodologie più consolidate, preferite sia alle prove destrutturate quali relazioni individuali e di gruppo, sia a quelle fortemente strutturate, come i test standardizzati» (Barone, 2010, p. 169). I dati dell'ultima rilevazione sono ancor più rilevanti se si tiene conto dell'avvio, a partire dall'a.s. 2001/2002, della sperimentazione del primo Progetto Pilota dell'INVALSI.

Inoltre, in relazione a reperimento, interpretazione, diffusione e uso dei dati restituiti dall'INVALSI, va tenuto presente che solo il 36,4% degli insegnanti della secondaria di I grado usa la rete per informarsi e consultare informazioni di carattere organizzativo (Cavalli & Argentin, 2010). A ciò va

aggiunto che in alcuni contesti i dirigenti di istituti secondari di II grado lamentano i «tempi insufficienti per analizzare con obiettività e criticità i risultati ottenuti», la «molteplicità degli impegni», il «limitato interesse da parte di alcuni insegnanti nei confronti dell'analisi dei dati», la «scarsa convinzione dei docenti in merito alla significatività e utilità dei dati» (Giovannini & Marcuccio, 2007, p. 71). E questa, presumibilmente, è una situazione che possiamo attenderci anche nei gradi scolastici inferiori.

Infine, sono state segnalate alcune criticità specifiche della PN. In particolare Domenici (2008 e 2009), in una cornice che riconosce alla PN l'importante funzione di supporto ai processi decisionali delle scuole, ha evidenziato problemi circa la validità e attendibilità dei risultati – legati al ruolo svolto dagli insegnanti nella somministrazione e attribuzione del punteggio e agli oggetti misurati dai quesiti – suggerendo «cautele metodologico-procedurali» e «interpretative».

L'introduzione della PN – realizzata peraltro in una situazione di valutazione certificativa quale l'Esame di Stato, la cui struttura era ormai consolidata da quasi un trentennio (DM 26/08/81) – è accaduta quindi in un contesto valutativo che non utilizza in modo diffuso prove strutturate di tipo standardizzato né sembra avere le risorse adeguate per avvalersi degli esiti della PN.

A partire da questa situazione problematica si è deciso di condurre un'indagine esplorativa allo scopo di rilevare quali conseguenze abbia provocato l'introduzione della PN sugli insegnanti che sono i principali soggetti, all'interno del contesto scolastico, coinvolti in questa rilevazione.

2. IL QUADRO TEORICO DI RIFERIMENTO

2.1. *Le valutazioni esterne nelle scuole*

La somministrazione nelle scuole di prove standardizzate costruite da soggetti esterni in conformità a standard di apprendimento definiti a livello centrale è un fenomeno diffuso da tempo in contesti educativi non italiani. Negli Stati Uniti, per esempio, a partire dagli anni '60 si è verificato un graduale incremento di questo fenomeno che ha raggiunto l'apice tra la fine degli anni '80 e i primi anni '90, periodo in cui i risultati di queste prove sono stati collegati a un numero sempre più elevato (*high*) di «poste in gioco» (*stake*) sia per la singola scuola (erogazione di finanziamenti, riconoscimento di ambiti di autonomia e privilegio) sia per gli studenti (promozione a una classe

successiva, ammissione a un ordine di scuola superiore) (Segall, 2003). Il fenomeno delle valutazioni *high-stake* ha toccato anche il continente europeo in particolar modo a partire dagli anni '90 a seguito della crescente rilevanza del movimento dell'*accountability*.

In letteratura non esiste una definizione univoca del concetto di valutazione *high-stake* (Chapman & Snyder, 2000; Abram *et al.*, 2003; Segall, 2003); alcune definizioni più restrittive individuano l'alto livello di posta in gioco per gli studenti nelle conseguenze di tipo formale quali la promozione o l'acquisizione di un titolo finale (Huebert & Hauser, 1999); altre estendono il significato di *high-stake* alle valutazioni che possono essere interpretate da diversi soggetti come un'espressione della qualità del sistema scolastico così da spingerli a chiedere un miglioramento dei risultati delle prove (Corbet & Wilson, 1991); altre ancora considerano *high-stake* le valutazioni che hanno influenze sociali su molteplici livelli delle realtà educative (Segall, 2003), non da ultimo il benessere degli insegnanti (Stake, 1998). In queste due ultime accezioni anche la valutazione effettuata mediante la PN 2008 e 2009 può essere considerata una valutazione *high-stake* per la ricaduta che ha avuto sia sugli studenti, sebbene dal punto di vista formale limitata, sia sugli insegnanti e sulle scuole.

Le ricerche empiriche sulle valutazioni «ad elevata rilevanza» ci offrono una serie di risultati contrastanti. Alcune evidenziano molteplici fenomeni critici di cui sono causa le prove strutturate (*teaching to test*, *narrowing curriculum*, *narrowing pedagogy*); altre, invece, dimostrano che i possibili esiti negativi possono essere prevenuti o attenuati soprattutto attraverso particolari interventi sugli insegnanti (Yeh, 2005). Tuttavia, in questo dibattito, vanno tenuti attentamente in considerazione i diversi sistemi di valutazione all'interno dei quali vengono somministrate le prove standardizzate nonché i diversi formati delle prove. Tra queste ricerche alcune riguardano in modo specifico i punti di vista degli insegnanti di lettere (Zancanella, 1992; Boardman & Woodruff, 2004; Crocco & Costigan, 2007; Diamond, 2007; Assaf, 2008).

Il dibattito sulle valutazioni esterne *high-stake* si è sviluppato anche in Europa, parallelamente al diffondersi del fenomeno. Le ricerche empiriche, il cui numero è ormai abbastanza consistente (ad es., Scott, 2007; Lafontaine & Dierendonck, 2008; Nidegger, 2008), fanno emergere preoccupazioni degli insegnanti simili a quelle rilevate oltreoceano.

In Italia, la riflessione degli esperti di docimologia sulle valutazioni esterne in ambito scolastico da parte dell'INVALSI ha accompagnato il nascere e l'evoluzione dei Progetti Pilota e dei primi interventi del Servizio Nazionale di Valutazione (Domenici, 2003 e 2004; Lucisano, 2003 e 2005; Vertecchi, 2004; Losito, 2006) e si è arricchita anche di contributi di esperti disciplinari

(ad es., in ambito matematico, Fandiño Pinella, 2005 e Domingo, 2006; in ambito linguistico, Bertocchi, 2010 e Manzotti & Zampese, 2010). Sulle PN sono stati realizzati alcuni studi tesi a evidenziarne i presupposti di tipo pedagogico-normativo (ad es., Bertagna, 2004; Sandrone, 2009); altri, invece, che ne hanno rimarcato la rilevanza per i processi decisionali in ambito scolastico mettendone in luce anche potenziali aspetti critici connessi alle scelte metodologico-procedurali (Domenici, 2008 e 2009). Poche, invece, sono le ricerche empiriche sull'impatto delle restituzioni dei dati delle prove INVALSI (Giovannini, 2007). Non ci risultano essere state effettuate specifiche ricerche di tipo empirico sulla PN.

2.2. *Le teorie sull'uso delle valutazioni*

L'indagine qui presentata si colloca nel filone di indagini sull'uso delle valutazioni, in continuità con precedenti nostre ricerche (Marcuccio & Truffelli, 2007), all'interno di una cornice teorica più generale che considera le valutazioni esterne come attività integrate alle valutazioni interne in una prospettiva formativa (Nevo, 2002; Bonami, 2005; Looney, 2011).

Le indagini sulle conseguenze delle valutazioni, che rientrano nel filone di ricerca denominato *evaluation utilization* o *evaluation use*, mirano a indagare in che modo le valutazioni e gli esiti influenzino i soggetti e le organizzazioni coinvolte nelle valutazioni stesse (Alkin, 2005). Gli studi avviati intorno agli anni '60 grazie ai contributi di autori quali Carol Weiss, Michael Patton, Marvin Alkin, hanno contribuito a identificare tre macroaree di fattori associati alla probabilità di un ampio uso delle valutazioni: (a) le *caratteristiche degli utilizzatori* (aspettative, precedenti esperienze valutative, atteggiamenti verso la valutazione, percezione dei rischi provenienti dalla valutazione); (b) il *contesto* della valutazione; (c) la *valutazione in sé stessa* (procedure, strumenti, modelli di valutazione, interazioni tra valutatore e utilizzatore, rilevanza e specificità delle informazioni, modalità delle restituzioni).

Dalla letteratura sull'uso degli esiti della ricerca scientifica sono stati assunti, nel dibattito valutativo, i primi due tipi di uso: l'*uso strumentale*, ossia l'uso degli esiti della valutazione per assumere decisioni o per risolvere problemi; l'*uso concettuale*, ossia l'uso – spesso non previsto inizialmente – degli esiti di una valutazione per modificare il modo di pensare dei destinatari. La riflessione successiva ha portato a sviluppare altri tipi di uso: l'*uso per la legittimazione* (*legitimative utilization*), per supportare una decisione politica già assunta; l'*uso simbolico*, per dare rilevanza simbolica al soggetto che ha deciso di realizzare la valutazione stessa; l'*uso del processo* (introdotto di recente da Jennifer Green, Jean King e Michael Patton), ossia le conseguenze che ha

il processo della valutazione e l'interazione tra valutatori e stakeholders sui soggetti e sull'organizzazione in cui è realizzata. Da un approfondimento di queste concettualizzazioni è stato proposto di utilizzare la distinzione tra *uso dei risultati* e *uso del processo* tenendo presente che entrambi possono essere utilizzati sia in modo strumentale sia in modo concettuale (Alkin, 2005).

Una tappa dello sviluppo degli studi sull'*evaluation use* è segnata dall'introduzione del concetto di «influenza» («capacità o potere delle persone o delle cose di produrre effetti sugli altri attraverso mezzi intangibili e indiretti») proposto da Karen Kirkhart (2000, p. 7) per sostituire quello di «uso», ritenuto inadatto a cogliere i diversi tipi di impatto della valutazione. A partire da questo concetto la ricercatrice ha elaborato un modello di influenza delle valutazioni in cui si intersecano le fonti dell'influenza (processi ed esiti della valutazione), l'intenzionalità del cambiamento provocato dalla valutazione (previsto [*intended*] e non previsto [*unintended*]) e il momento in cui accadono i cambiamenti (immediati, finali, a lungo termine); lo scopo è quello di riuscire a cogliere effetti «multidirezionali, incrementali, non intenzionali e strumentali».

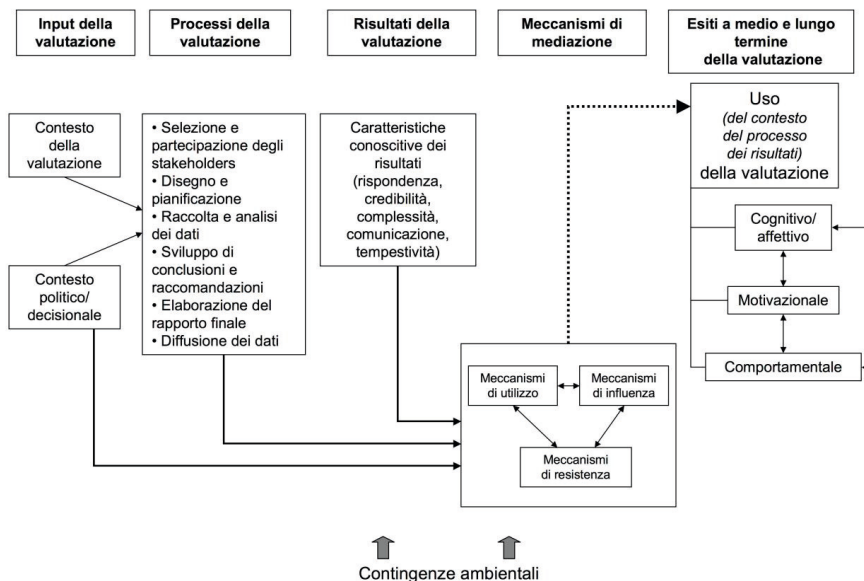
Il modello di Kirkhart è stato successivamente sviluppato da Melvin Mark e Gary Henry (Henry & Mark, 2003; Mark & Henry, 2004), sulla base dei risultati di ricerche empiriche e teoriche realizzate in diversi ambiti delle scienze sociali. Gli autori hanno elaborato una *teoria dell'influenza della valutazione* che, a partire dal modello di Cousin (2003), individua le *dimensioni* dell'intervento di valutazione che possono provocare un'influenza (il contesto politico-decisionale, le procedure valutative e le caratteristiche degli esiti), i *meccanismi generali* (processi) che favoriscono l'influenza, gli *esiti misurabili* dell'influenza a livello cognitivo/affettivo, motivazionale e comportamentale. Gli esiti, che hanno come obiettivo finale quello del benessere sociale, possono a loro volta trasformarsi in processi/meccanismi capaci di favorire l'influenza.

La proposta di introdurre il concetto di influenza ha aperto un vivo dibattito tra i ricercatori. Alkin (2005) rivendica l'importanza di mantenere una distinzione tra i concetti di «uso» e «influenza». Il primo si colloca in modo esclusivo nel contesto del programma sottoposto a valutazione ed entro un arco di tempo limitato; il concetto di influenza, invece, fa riferimento ad altri contesti e ad altri intervalli temporali. Inoltre l'influenza e l'uso si differenziano per il grado di intenzionalità con cui utilizzatori e valutatori usano la valutazione. Per questo il concetto di influenza può essere riconducibile al concetto di «impatto» oppure a quello di uso concettuale poiché più ci si allontana dal luogo e dal momento della valutazione più i risultati e i processi possono essere considerati come conoscenza. Patton (2008) riconosce l'importanza del framework della Kirkhart e della sua funzione nell'elaborazione di nuovi modelli di *ricerca* relativi agli effetti della valutazione, tra cui

quelli di Melvin Mark e Gary Henry. Tuttavia egli ritiene, facendo propri gli argomenti di Alkin, che esso sia «meno utile» per guidare la *pratica* valutativa (Patton, 2008, p. 110). Johnson *et al.* (2009), inoltre, affermano che il modello di Mark e Henry non fornisce una chiara definizione operativa del concetto di influenza così come dei concetti di percorso, processo e meccanismo di influenza. A queste osservazioni critiche aggiungiamo che il modello di Mark e Henry non sembra riuscire a spiegare il fenomeno, già codificato in letteratura, del «non uso» delle valutazioni (Patton, 2008). Infatti, anche in presenza di fattori che possono facilitare l'uso della valutazione, l'insorgere di dinamiche di «resistenza» (Taut & Brauns, 2003) può attenuare, se non addirittura bloccare, i processi di utilizzo della valutazione.

Nonostante queste osservazioni critiche riteniamo utile assumere il modello di Mark e Henry (2004) come riferimento teorico di questa ricerca per i seguenti motivi: il modello è stato proposto come punto di partenza per avviare future ricerche empiriche e quindi il suo ambito di riferimento è principalmente quello della ricerca e non tanto quello dell'uso pratico; il modello contiene una nuova variabile (quella dei meccanismi di influenza) che ne incrementa il valore esplicativo rispetto ai modelli precedenti; inoltre esistono già ricerche finalizzate a raccogliere evidenze empiriche per validare la portata euristica del modello (Christie, 2007). Tuttavia, per ovviare alle critiche sollevate e per rispondere alle esigenze della presente ricerca, abbiamo ritenuto opportuno procedere a un adattamento del modello apportando alcuni aggiustamenti. Il primo è l'inserimento nel modello della variabile *meccanismi di mediazione* che racchiude: (a) i *meccanismi di utilizzo*, ossia i meccanismi progettati intenzionalmente per promuovere un utilizzo dei vari aspetti della valutazione stessa; (b) i *meccanismi di influenza*, ossia i meccanismi non previsti dal piano di valutazione, attraverso i quali possono essere provocati dei cambiamenti, *anche* al di fuori del contesto della valutazione; (c) i *meccanismi di resistenza* messi in atto per impedire un cambiamento potenzialmente promosso dalla valutazione (Taut & Brauns, 2003). I meccanismi di mediazione, interagendo tra di loro, influiscono sugli esiti a medio e lungo termine delle valutazioni. In secondo luogo, abbiamo inserito il termine «uso» quale termine di sintesi che racchiude gli effetti provocati dai vari aspetti valutazione (contesto politico/istituzionale, processi ed esiti). Infine, abbiamo previsto anche la possibilità del «non uso» tra le possibili relazioni tra i meccanismi di mediazione e gli usi della valutazione (ossia i cambiamenti provocati dalla valutazione). Nella Figura 1 quest'ultimo aspetto è rappresentato dalla freccia tratteggiata.

Nel nostro adattamento il concetto di *influenza* assume un significato simile a quello della Kirkhart (2000, p. 7): «[...] capacità o potere delle persone o delle cose di produrre effetti, intenzionali o no, sugli altri attraverso mezzi intangibili (*intangibile*) e indiretti (*indirect*)».



Adattato da: Mark & Henry, 2004.

Figura 1. – Il modello di utilizzo della valutazione.

Va precisato che i meccanismi di influenza sono meccanismi *non previsti dal disegno iniziale della valutazione* e sono stati inseriti per differenziarli dai processi di utilizzo previsti sin dall'inizio da parte della valutazione. Inoltre il contesto di riferimento d'uso della valutazione è quello in cui è avvenuta la valutazione, e il riferimento temporale del fenomeno descritto dal modello è quello finale e a medio termine.

3. IL DISEGNO DELLA RICERCA

Alla luce della cornice teorica assunta abbiamo formulato i seguenti interrogativi di ricerca: L'introduzione della PN ha favorito l'uso della valutazione da parte degli insegnanti di italiano? Se sì, in quali ambiti del contesto scolastico è avvenuto tale uso? Quali sono stati gli usi effettuati? Si sono verificati dei non usi? Sono stati attivati dei meccanismi di mediazione? Se sì, quali?

I dati emersi dalle ricerche empiriche in letteratura ci portano a ipotizzare che gli usi della PN, effettuati dagli insegnanti di italiano, siano esigui e stret-

tamente funzionali all'assolvimento burocratico della PN a causa di meccanismi di resistenza e che, tra i cambiamenti, siano più frequenti quelli provocati dal processo valutativo e dai presupposti di contesto politico/istituzionale.

L'indagine ha coinvolto un campione non rappresentativo di docenti di italiano che, al momento dell'indagine (a.s. 2008/2009)² o l'anno precedente, insegnavano nelle classi terze degli istituti secondari di I grado di un comune del Nord Italia. Si tratta dei due anni scolastici in cui alle commissioni d'esame era ancora riconosciuta la competenza di determinare in modo autonomo il peso della PN sull'esito del voto finale dell'Esame di Stato. La scelta di delimitare la ricerca ai soli insegnanti di italiano scaturisce dal convergere in questa indagine di un interesse di ricerca personale relativo alla comprensione del testo scritto.

L'obiettivo di far emergere i diversi tipi di usi (cognitivi, motivazionali e comportamentali) e le relative argomentazioni ha portato a privilegiare il focus group quale tecnica di raccolta dei dati. Ne sono stati realizzati 9, ciascuno per ogni istituto secondario di I grado statale del territorio comunale che ha aderito all'indagine, a cui hanno partecipato complessivamente, su base volontaria, 46 insegnanti. Anche se non probabilistico, il campione può essere ritenuto adeguato alla natura esplorativa dell'indagine. Due istituti scolastici hanno preferito non partecipare all'indagine.

Gli incontri, della durata di circa due ore, si sono svolti nei locali delle scuole nei mesi di ottobre-novembre del 2009 e sono stati condotti dall'autore affiancato da un osservatore³ secondo una metodologia ormai consolidata (Krueger, 1994 e 1997; Morgan, 1997; Greenbaum, 1998; Zammuner, 2003) con particolare attenzione all'uso specifico di questa tecnica nei contesti educativi (Vaughn *et al.*, 1996). La conduzione della discussione è stata effettuata adottando uno stile comunicativo non direttivo (Lumbelli, 1981) che si è avvalso della tecnica del rispecchiamento (Lumbelli, 1993) al fine di attenuare le distorsioni nell'emergere delle dichiarazioni.

Durante l'incontro ai partecipanti sono stati distribuite copie delle PN e un estratto dei Rapporti dell'INVALSI per agevolare il riferimento ai vari aspetti delle rilevazioni. I temi guida utilizzati per il rilancio della discussione, che verteva sia sulla PN 2008 che sulla PN 2009, sono stati i seguenti: reazioni e valore attribuito alla PN, uso delle PN e dei risultati, impatto sulla didattica personale e sull'organizzazione, bisogni formativi degli insegnanti

² Colgo l'occasione per rinnovare il mio sentito ringraziamento, per la loro disponibilità, accoglienza e collaborazione, a tutti i dirigenti scolastici e agli insegnanti che hanno partecipato alla ricerca.

³ Il mio grazie va anche a Giorgia Agolini, Beatrice Bosini, Stella Chiussi e Giovanna Sineri, che hanno svolto con professionalità il ruolo di osservatrici.

circa la PN, valutazione complessiva della PN. È stata chiesta l'autorizzazione all'audioregistrazione, che in una scuola è stata negata. I dati utilizzati per l'indagine, quindi, sono quelli relativi alle trascrizioni di 8 focus group per un totale di 41 insegnanti.

Il corpus delle trascrizioni è stato analizzato con due procedure distinte: la prima informatizzata, di tipo lessico-testuale, utilizzando il software T-LAB versione 7.0.2 (Lancia, 2004) ed elaborando alcuni suoi output; la seconda mediante un'analisi tematica del contenuto dell'intero corpus (Boyatzis, 1998; Ryan & Bernard, 2003). La scelta di utilizzare T-LAB è in primo luogo legata agli obiettivi della ricerca: far emergere dapprima una mappa degli ambiti dei possibili usi della PN attraverso un approccio induttivo che limitasse l'intervento del ricercatore. Il software, infatti, mediante la possibilità di lavorare sui lemmi del corpus con una «codifica semiautomatica» (Tuzzi, 2003) consente di limitare inizialmente l'interpretazione del ricercatore. Il secondo motivo è legato a un interesse di ricerca trasversale a questa indagine per gli aspetti metodologici delle ricerche con metodi misti (Giovannini *et al.*, 2009) che ha portato a escludere per esempio un approccio di tipo *grounded theory*. La presentazione dei risultati di questo particolare aspetto della ricerca, tuttavia, esula dagli obiettivi del presente contributo che si limita alla restituzione degli esiti della prima procedura di analisi.

Il corpus è stato preliminarmente suddiviso in base ai turni conversazionali (testi) identificati da due variabili: focus e persona (conduttore/docente). Dopo l'importazione, il corpus risultava così caratterizzato: testi importati: 2239 (conduttore: 668; docente: 1571); contesti elementari: 3264; occorrenze (*token*): 94332; forme (*type*): 8071; lemmi: 4191; *hapax*: 3761. La soglia minima iniziale di occorrenza per effettuare l'analisi era 5; dopo le procedure di disambiguazione e lemmatizzazione abbiamo elevato la soglia a 8. Sono state inserite nell'analisi sia le parole chiave individuate dal software sia altre di particolare interesse per la ricerca.

Sull'insieme del subcorpus «docente» è stata effettuata, in primo luogo, un'analisi tematica dei contesti elementari (*cluster analysis*) con il vincolo di 4 co-occorrenze all'interno di ogni singolo contesto. Il numero di contesti classificati è stato di 960 su un totale di 1571. La soluzione dell'analisi ha individuato 11 cluster che sono stati scelti senza procedere a ulteriori partizioni. Hanno guidato la costruzione dei cluster le seguenti parole chiave: *risposta, differente, modifica didattica, abituato, prova INVALSI, straniero, antologia, prova di esercitazione, valutazione, testo, difficoltà*.

In un secondo momento è stata effettuata un'analisi tematica del contenuto, di tipo qualitativo (Boyatzis, 1998; Ryan & Bernard, 2003; Massey, 2011), sui primi 20 contesti elementari significativi individuati dal software per ciascun cluster.

Nei primi giorni di settembre 2010, mediante l'invio postale di un opuscolo cartaceo, i dati di questa fase dello studio sono stati restituiti a tutti i partecipanti alla ricerca.

4. I RISULTATI

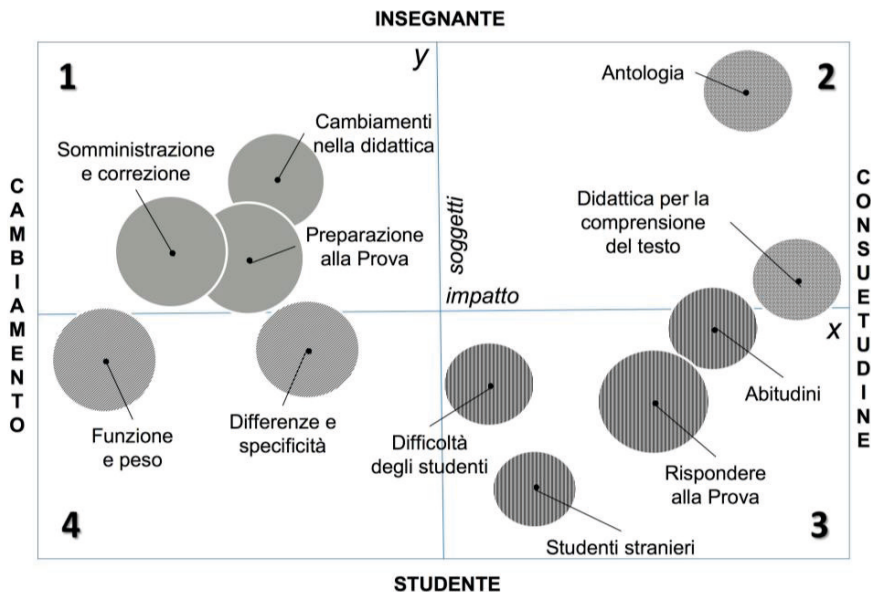
4.1. *L'analisi tematica dei contesti elementari*

L'analisi tematica dei contesti elementari del corpus ha consentito di individuare 11 cluster di parole chiave (rinominati in base alle parole chiave che li componevano) e collocarli su un piano bidimensionale lungo due assi ortogonali, da noi denominati «Impatto» (l'orizzontale) e «Soggetti» (il verticale), che delimitano quattro macroaree che abbiamo così etichettato:

- Macroarea 1 – *Cambiamenti per gli insegnanti*. Delimita gli ambiti tematici relativi alle attività degli insegnanti che sono cambiate in relazione alla PN: *somministrazione e correzione, preparazione alla PN e cambiamenti nella didattica*.
- Macroarea 2 – *Consuetudini degli insegnanti*. Raggruppa due ambiti tematici, *didattica per la comprensione del testo e antologia*, che descrivono l'attuale didattica della comprensione del testo.
- Macroarea 3 – *Consuetudini degli studenti*. Raccoglie quattro ambiti tematici: i primi due, *rispondere e abitudini*, descrivono abitudini e modalità di risposta degli studenti rispetto alla PN e alle prove strutturate in generale; gli altri, *studenti stranieri e difficoltà degli studenti*, descrivono le difficoltà che gli studenti, stranieri e non, trovano nell'affrontare la PN.
- Macroarea 4 – *Cambiamenti per gli studenti: funzione e peso*. Comprende due ambiti tematici che toccano gli aspetti organizzativo-istituzionali più generali della prova percepiti, però, come rilevanti per le loro conseguenze sugli studenti: la *funzione* e il *peso* della PN e le *differenze e specificità* della PN (Figura 2).

4.2. *L'analisi dei contesti elementari significativi*

Dall'analisi tematica del contenuto dei primi 20 contesti elementari significativi individuati dal software per ciascuno degli 11 cluster è stato possibile ricostruire la seguente rete concettuale relativa a entrambe le PN qui prese in esame (2008 e 2009).



L'estensione delle aree dei cerchi indica la diversa numerosità dei lemmi che compongono i cluster.

Figura 2. – Rappresentazione spaziale delle macroaree e degli ambiti tematici emersi durante i focus group.

La somministrazione e la correzione. La somministrazione e la correzione sono gli ambiti di cambiamento «obbligatori» per gli insegnanti poiché la normativa richiedeva che fossero loro, in prima persona, a compiere queste operazioni. Per la somministrazione delle prove nelle classi, in molte situazioni sono state seguite le indicazioni fornite dall'INVALSI; in alcuni casi, invece, sono stati incaricati per la distribuzione delle prove e l'assistenza gli insegnanti della stessa materia.

Tuttavia la presenza degli insegnanti della stessa disciplina è considerata «rassicurante» per gli studenti e in grado di garantire egualmente le «condizioni di oggettività» se gestita senza «intrusioni» e «suggerimenti» sollecitati anche dalla «paura di fare brutta figura».

Gli insegnanti non comprendono, però, la scelta di somministrare la PN in sede d'esame. Un'alternativa che viene proposta è quella di somministrare la PN verso la fine dell'anno scolastico rimandando a una successiva decisione se utilizzarne gli esiti in sede di Esame di Stato. È evidente che in tal modo si renderebbe meno pesante l'esame per gli studenti ma non si darebbe l'«idea di omogeneità» e di «importanza» alla prova.

La mancanza di informazioni sulle procedure e il ritardo dell'invio delle griglie di correzione hanno reso critica la correzione della PN 2008. Non così per la PN 2009 sebbene anch'essa sia stata particolarmente difficile per tre motivi: i tempi stretti per la restituzione dei dati, che hanno impedito il confronto tra colleghi e l'analisi dei risultati; il sistema di registrazione delle risposte (una scheda da riempire con i «pallini»); infine il «complicato» e «cervellotico» calcolo dei punteggi, che ha portato via «ore e ore» rendendo «pazzi» e «disperati» gli insegnanti. Il doppio sistema di attribuzione del punteggio della PN 2009 è stato gestito in modo differente dalle scuole ma con la tendenza a privilegiare la modalità che salvaguardasse l'andamento annuale della classe.

Durante la correzione sono emerse talvolta discordanze sulle chiavi di soluzione fornite dall'INVALSI. Inoltre, dati i diversi ambiti di valutazione e le caratteristiche dei tipi di testo e delle opzioni di risposta, ci si sarebbe aspettata una differenziazione in fase di attribuzione del punteggio.

Persorgono, inoltre, perplessità circa il tentativo di «rendere oggettiva la valutazione di una materia che è estremamente difficile». Le risposte chiuse sono certo più «comode», più «oggettive» per la correzione, ma sono inadeguate a rilevare aspetti più importanti. Questo ha portato a ridimensionare il peso della PN nella valutazione complessiva dell'esame.

I cambiamenti nella didattica e la preparazione alla prova. Il cambiamento più rilevante ha toccato il *programma di grammatica*. La tradizionale scansione triennale in analisi grammaticale, analisi logica e analisi del periodo è stata rivista alla luce delle scelte della PN che si differenziano, tuttavia, da una tendenza presente in alcuni ambiti della scuola che considera la grammatica «obsoleta» o funzionale allo studio delle lingue straniere o al proseguimento in alcuni tipi di scuola. Precedentemente, pur dedicando tempo all'insegnamento della grammatica, si «dava per scontato che non rientrasse nella valutazione d'esame». Da qui la necessità, per chi già non lo faceva, (a) di portare «avanti parallelamente» l'analisi logica e l'analisi del periodo; (b) di «mantenere ripassato» durante l'ultimo anno «tutto il programma di grammatica»; (c) di «intrecciare» la comprensione del testo con l'analisi grammaticale; (d) di effettuare prove intermedie di preparazione. Questo cambiamento tuttavia si scontra con il modo di ragionare per «comparti stagni» degli studenti.

La seconda novità indotta dalla PN riguarda la somministrazione di *prove* di «simulazione» in preparazione all'esame sia per abituare gli studenti a «osservare cosa la domanda chiede» e a «far vedere come deve fare», così da diminuirne i livelli di «ansia» e di «panico»; sia per consentire di «capire quali sono le difficoltà», e «intervenire» per «correggere il tiro». Questo tipo di attività, a differenza dei colleghi di matematica, è avvenuta in molti casi senza un vero confronto e accordo con gli altri colleghi di italiano. Le prove utilizzate

per queste esercitazioni sono state reperite dalla rete internet, andando «a caccia sul sito dell'INVALSI». Esse, tuttavia, contengono testi «poco logici», poco «chiari» e «di una complessità incredibile». Un minore spazio hanno avuto le pubblicazioni *ad hoc* e le appendici presenti in alcune antologie perché difficilmente reperibili e non «complete» come la PN.

L'introduzione della PN, invece, non ha provocato cambiamenti nel tradizionale e continuo lavoro, realizzato per tutto il triennio, di «lettura, analisi ed esercitazione» dei testi scritti; questo perché «non è una prova che fa cambiare» l'insegnamento ma la particolare «situazione della classe», ciò che vi «sperimenti» e «vedi che funziona». Per altri, invece, è «servita» a superare alcuni «limiti» nell'impostazione didattica. Questo aspetto viene sviluppato qui di seguito.

La didattica della comprensione del testo. Pur con differenze, esistono strategie condivise tra i docenti per favorire la comprensione del testo. In primo luogo la richiesta di leggere «prima di tutto il testo» per averne un'«idea» complessiva per poi passare a un preliminare lavoro sulle «competenze lessicali». La scelta dei successivi livelli di analisi è dettata da un approccio prevalentemente strutturale: individuazione e descrizione di personaggi, luoghi, sequenze, tempo narrativo e trama. Tuttavia sono presenti anche diverse concezioni del testo e del processo di comprensione che ritengono inutile «parcellizzare così tanto» la comprensione sia perché «c'è una ricchezza in un testo che non è riducibile all'insieme di tutto questo» sia per il «contributo personale» alla comprensione del soggetto che legge.

La didattica della comprensione del testo è impostata tenendo conto di studenti per i quali la lettura non è più percepita come un'attività rilevante. La loro interazione con il testo è inadeguata: «fretta», mancanza di una «strategia di approccio al testo», preferenza per un «linguaggio immediato e povero», incapacità di «inferire il significato della parola dal contesto». Ecco perché risulta per loro difficile discriminare, nelle alternative di risposta della PN, termini «talmente vicini» così come cogliere aspetti più complessi come il genere letterario. Da qui i dubbi sollevati sulla validità della PN: con una o poche domande «misuro la competenza lessicale di un ragazzino?».

L'antologia. L'antologia è lo strumento principe della didattica della comprensione del testo. L'impostazione delle antologie varia «estremamente» da autore ad autore e non sempre fornisce un valido e «accurato» aiuto. Alcune antologie si prestano meglio di altre a lavorare sulle «strategie di lettura», ma risulta egualmente difficile trovarne una valida perché talvolta vi sono «ambiguità interpretative» e scelte lessicali che sono fonte di «preoccupazione» all'atto dell'adozione.

L'impostazione tradizionale delle antologie segue uno sviluppo triennale di progressivo approfondimento degli aspetti del testo che sono trattati in relazione a brani raggruppati per argomento o tipologia testuale. La comprensione dei testi è supportata da alcuni esercizi impostati con domande quasi sempre uguali, al punto da risultare «ripetitive», che privilegiano la «comprensione globale del testo» rispetto a quella «analitica» e la «produzione scritta, libera, aperta» diversamente dalla PN. Le domande a risposta chiusa non condividono «per niente» l'impostazione della PN, anzi risultano piuttosto «banali», «semplici», «molto diverse» e, quindi, «inutili»; in alcuni casi «fanno quasi un po' ridere». Inoltre gli ambiti e i processi misurati dai quesiti non sempre sono dichiarati in modo «dettagliato», «peculiare», «specifico» e «puntiforme» rendendo difficile il raffronto con la PN. Circa lo sviluppo delle competenze lessicali, poi, molte antologie suggeriscono il significato dei termini sconosciuti in nota invece di costruire esercizi a tal scopo. I redattori della PN avrebbero dovuto tener in maggiore considerazione tutti questi aspetti.

Le abitudini degli studenti e il rispondere. Per comprendere il comportamento degli studenti durante la PN è necessario considerare che la scuola italiana privilegia la produzione scritta e ricorre alle prove strutturate solo per «correggere poco».

Gli studenti, a scuola, sono abituati generalmente a rispondere a domande aperte che «scavano più in profondità», «più chiare», più «evidenti»; le domande a scelta multipla presenti nelle antologie hanno alternative di risposta molto più «differenziate», con la risposta giusta che «si capisce bene»; non così nella PN dove le domande sono di livello «molto medio» e le alternative sono «ambigue», «cervellotiche», «non sempre puntuali», con «sottili sfumature» da «sembrare giuste tutte». Questo può spingere gli studenti a considerare le domande della PN «banali» e «troppo facili» e, in alcuni casi, a supporre che vi sia «qualcosa di nascosto». Così, soprattutto gli studenti più «bravi» si sono «persi in quisquiglie» e sono caduti in errore.

Inoltre, gli studenti sono abituati a insegnanti che «accettano» dagli studenti anche risposte diverse alla stessa domanda quando c'è un ragionamento che giustifica la loro «interpretazione» e «rielaborazione personale» del testo, mentre nella PN la risposta giusta è «solo una».

Infine gli studenti sono soliti operare con tempi più dilatati. Durante la PN essi «continuano a sfogliare, a guardare quante ne ho fatte, che ore sono». E l'«ottimizzazione dei tempi» è stato uno degli obiettivi principali delle prove di esercitazione.

Quelli elencati sono solo alcuni dei fattori che spingono gli studenti a non prestare attenzione a «cosa chiede la domanda» e ad affidarsi alla «me-

moria» cadendo spesso in errore. D'altro canto, vi è anche chi, preso dall'incertezza, segna subito una risposta, il più delle volte quella giusta, ma in seguito la modifica.

Le caratteristiche della PN non sembrano aiutare in questo gli studenti. Vi sono tuttavia aspetti della PN che riescono a suscitare il loro coinvolgimento. Per esempio, il brano narrativo della PN 2009 è stato affrontato dagli studenti «con più voglia di mettersi alla prova». Inoltre il fatto di «sapere che sarebbe stata valutata» ha di certo provocato una particolare attenzione verso la PN. Infine, la presenza, nelle domande, del «preciso riferimento alla riga» del brano che «devono andare a controllare» per rispondere è una «guida» che, sebbene possa favorire un comportamento «meccanico», aiuta gli studenti «superficiali», che rispondono «a caso» senza fermarsi a «osservare cosa chiede la domanda».

Gli studenti «stranieri» e le difficoltà degli studenti. La presenza di studenti stranieri di prima e seconda generazione è molto forte in quasi tutte le scuole. A questi si affiancano gli studenti «meno inseriti» e «in difficoltà», i «ragazzini del sud», quelli con certificazioni di disabilità e con problemi di dislessia, quelli con «complessità a livello personale e sociale» e con «svantaggi socio-culturali». Spesso tutti concentrati nella stessa classe.

Questo porta a chiedersi: «Chi sceglie i testi» per la PN conosce «i ragazzi che abbiamo davanti?». Infatti, le situazioni che si creano con questi studenti durante la PN sono abbastanza particolari: una ragazzina straniera «molto brava» è scoppiata «in lacrime alla riconsegna» perché «non capiva il termine corpolento». Altri sono andati «a intuito»; uno studente dislessico «ha fatto zero su tutto». È «terribile» vedere un «ragazzino in difficoltà che non capisce, che si dispera e non può chiedere aiuto». Tuttavia, laddove le condizioni lo hanno permesso, un lavoro di «recupero» durante l'anno ha consentito agli studenti stranieri e a quelli in difficoltà di raggiungere risultati accettabili nella PN.

Anche la situazione degli studenti cosiddetti normodotati si presenta abbastanza problematica. Nelle classi vi sono studenti che, oltre a difficoltà nella «produzione» e a un vocabolario «sempre più povero», rivelano carenze anche nella «padronanza del lessico», nella capacità di compiere «inferenze» e nell'usare i «connettivi». Tuttavia la PN può valorizzare parzialmente anche questi studenti. Coloro che hanno problemi di lessico, per esempio, posti di fronte a esercizi di grammatica «ci saltano fuori molto meglio». Per questi, un tipo di lavoro come l'analisi, anche «se in maniera meccanica», può risultare un'agevolazione. Inoltre va da sé che chi «ha meno competenze» riesca meglio in una prova in cui sono presenti «già» le possibili scelte, così come riesce meglio chi «senza bisogno di studio» è «sveglio» e sfrutta il suo «livello di logica»

rispetto allo studente «bravo» che «studia molto» e «ha una preparazione forte» ma «non sempre la logica» lo supporta. Per questo, in fase di valutazione della PN, sarebbe opportuno avere un «confronto» tra colleghi per evitare di «premiare qualcuno che ha lavorato in un certo modo».

Per riuscire a favorire pienamente i ragazzi con difficoltà si può ipotizzare la predisposizione di una prova strutturalmente uguale ma con «testi semplificati». Permangono in ogni caso difficoltà difficilmente superabili per gli studenti con DSA che non hanno più una certificazione di disabilità.

Funzione e peso. Qual è la funzione della PN: valutare i livelli degli studenti, le scuole o gli insegnanti? La PN si rivela sostanzialmente adeguata a verificare i livelli di un «sapere minimo uguale per tutti», anche se non mancano dubbi sulla capacità di certi quesiti di misurare effettivamente determinati obiettivi: come mai «la stessa abilità poteva avere risultati altissimi e risultati bassissimi» nella prova dello stesso studente? Ma se quello è lo scopo, perché inserire la PN nell'Esame di Stato e utilizzarne il voto per la valutazione finale? Nel caso in cui la finalità fosse anche quella di valutare gli insegnanti, sarebbe inevitabile il manifestarsi di un comportamento di «aiuto» nei confronti degli alunni durante lo svolgimento della PN.

Il *peso da attribuire alla PN* è stato definito in tutti i casi dal collegio docenti delle singole scuole. In alcuni contesti, tuttavia, si è deciso di inserire la valutazione della PN nella valutazione complessiva solo in caso di positività consentendo così di valorizzare studenti che avevano ricevuto valutazioni basse nelle altre prove, ma nello stesso tempo di non penalizzare studenti con buone valutazioni durante tutto il triennio.

Insegnanti e genitori sono preoccupati circa la possibilità che il peso della PN sul voto finale dell'esame sia vincolata dal Ministero e venga aumentata in modo «eccessivo» rispetto a un «percorso di tre anni di valutazioni». In tal modo si «toglierebbe valore» al «tema», che costituisce un «punto forte» della valutazione in italiano. Inoltre si correrebbe il rischio di trattare le differenti situazioni che sono presenti nelle classi con la stessa valutazione andando in senso contrario alle indicazioni sempre più frequenti di una personalizzazione dell'insegnamento.

Differenze e specificità. La prima «differenza» che emerge è tra gli esiti della PN e quelli ottenuti dagli stessi studenti durante l'anno scolastico. In alcuni casi studenti «modesti» o che «annaspavano» hanno ottenuto nella PN dei buoni risultati. Si registrano anche casi opposti. Diverse le ipotesi interpretative: il ruolo giocato dalla «superficialità» degli studenti meno bravi e dalla «scrupolosità» dei bravi; lo stile cognitivo più «schematico» di alcuni; il fatto

che in classe si misurano certe abilità con «poche domande» diversamente da quanto avviene con la PN. Questo fenomeno ha «creato dubbi» anche sui «metodi» di insegnamento soprattutto in presenza di giudizi «decisamente negativi» dati a chi ha ottenuto nella PN esiti positivi «non con risposte fortunate».

Il secondo ambito di «differenza» riguarda il valore della *produzione scritta*, un «punto forte» della disciplina, e della sua valutazione rispetto a prove complessivamente «molto rigide» che «non sempre piacciono all'insegnante». La produzione scritta è più importante perché consente la libertà e la «creatività» del soggetto; perché è priva dei limiti insiti nelle prove strutturate; perché, infine, spinge gli studenti «a pensare e a sviluppare un pensiero» senza trovarselo «già dato». Certo, «è una bella cosa che uno capisca quello che c'è scritto» in un testo, tuttavia «speriamo che qualcuno riesca anche a fare la quinta risposta, insomma, riesca una volta tanto a dire qualcosa di suo».

Tuttavia esiste anche una componente di «soggettività» nella valutazione della produzione scritta che porta a ritenere «utile» anche l'introduzione delle prove strutturate senza però che vi sia l'«aspirazione» a far diventare «matematicamente oggettiva» la valutazione del «tema».

Inoltre non va dimenticato che è possibile una «integrazione» tra la valutazione della PN e quella dell'«italiano scritto» in quanto valutano abilità diverse della stessa competenza linguistica generale anche se l'attività della produzione può essere utilizzata in prove di comprensione laddove siano presenti domande a risposta aperta.

5. DISCUSSIONE

L'obiettivo di questa indagine era costruire una prima mappa dei cambiamenti provocati sugli insegnanti di italiano dall'introduzione della PN nell'Esame di Stato al termine del I ciclo di istruzione all'interno di una cornice teorica rappresentata dal modello di uso delle valutazioni adattato dalla teoria dell'influenza di Mark e Henry (2004). I risultati dell'analisi dei dati di tipo quanti-qualitativo, oggetto del presente contributo, hanno consentito di rilevare gli usi e i non usi della PN e alcuni meccanismi di mediazione.

Dall'analisi quantitativa (*cluster analysis*) delle trascrizioni è stata messa in luce l'esistenza di due ambiti di discorso ben distinti relativi all'introduzione della PN: uno del «cambiamento» e uno del «non cambiamento». Nel primo ambito emergono i cambiamenti comportamentali previsti dalla norma, ossia la somministrazione e la correzione delle prove (che vanno interpretati come partecipazione ai processi della valutazione), e quelli non «normati»,

vale a dire le attività di «preparazione alle prove» e i cambiamenti nella «didattica». Nel secondo, rientrano i «non cambiamenti» sia degli insegnanti (didattica della comprensione del testo) sia degli studenti (modalità di reazione alle prove strutturate).

Questo primo dato conferma la nostra ipotesi poiché tra i cambiamenti sollecitati dalla PN (usi della valutazione) rileviamo non solo quelli che coincidono con la partecipazione degli insegnanti, normata per legge, ai processi della PN (somministrazione e correzione), ma anche cambiamenti che toccano sia l'adempimento burocratico della PN (preparazione alla prova) sia il curriculum della disciplina (cambiamenti nella didattica).

L'analisi qualitativa dei contesti tematici significativi ci consente di approfondire e integrare questo primo risultato. Prendiamo dapprima in esame gli *usi riconducibili* con certezza *al processo* della valutazione. A questo proposito va sottolineato l'emergere di manifestazioni cognitive/affettive e motivazionali associate alla partecipazione degli insegnanti alle attività della PN. Durante il momento della *somministrazione* è emersa l'incomprensione per il rigore di certe regole di comportamento (presenza di insegnanti di classi diverse, limiti alle interazioni verbali con gli studenti), dall'altro la consapevolezza del potenziale rischio di comportamenti non etici in caso di una PN legata alla valutazione degli insegnanti. Anche la *correzione* ha posto gli insegnanti di fronte ad alcune novità rispetto alla tradizionale pratica didattica: la registrazione delle risposte su un modulo per la lettura ottica; il calcolo del punteggio secondo un duplice criterio; i tempi stretti per la correzione. Questo, da un lato, ha provocato una forte irritazione negli insegnanti (uso affettivo), dall'altro, ha fatto emergere le differenze tra i vari aspetti della PN e i punti di vista degli insegnanti (uso cognitivo): (a) tra l'«approccio qualitativo» dei docenti e l'«approccio quantitativo» della PN; (b) tra le chiavi di risposta e le modalità per l'attribuzione del punteggio dell'INVALSI e quelle ritenute più opportune dai docenti; (c) tra la prospettiva «di sistema» della PN e quella «locale» degli insegnanti. Questo insieme di concezioni emergenti ha favorito sia il ridimensionamento del peso riconosciuto alla PN in sede di attribuzione del voto finale dell'esame sia l'incomprensione di somministrare la PN in sede d'esame finale sebbene si riconosca il valore dell'«omogeneità» e dell'«importanza» che deve avere l'esame. Va segnalato, inoltre, che i tempi riservati alla correzione hanno impedito il confronto tra gli insegnanti e l'analisi approfondita dei risultati privando così gli insegnanti di un'occasione per attivare meccanismi di influenza. In sintesi la partecipazione degli insegnanti alla somministrazione e alla correzione della PN ha avuto come conseguenza principale (uso del processo) quella di alimentare (non è possibile qui determinare se ex-novo oppure no) un atteggiamento non favorevole nei confronti della PN che non si è tradotto, certo, in un

suo boicottaggio ma in un'attenuazione delle sue implicazioni sull'esame (uso comportamentale).

Risulta più complesso riuscire a valutare se la ristrutturazione dell'attività di insegnamento della grammatica e la somministrazione di prove per far esercitare gli studenti alla PN siano imputabili a un uso del processo oppure a un uso dei presupposti politico/organizzativi della prova; oppure a entrambi. In ogni caso, resta il fatto del verificarsi di questi due fenomeni. La rivisitazione della didattica della grammatica ha implicato, in alcuni casi, una vera e propria ristrutturazione del curriculum e l'adozione di precise strategie didattiche (integrazione tra analisi logica e del periodo e tra queste e la comprensione del testo; continuità dei riferimenti durante tutte le attività; realizzazione di verifiche intermedie). Sembra che vi sia stato un rafforzamento, negli insegnanti, della convinzione dell'importanza della grammatica in contrapposizione a movimenti di opinione diversa (uso cognitivo e comportamentale). Questo dato non conferma la nostra ipotesi che prevedeva solo usi strettamente funzionali all'adempimento burocratico della PN. In questo caso, invece, rileviamo uno stimolo a mantenere o a rivitalizzare il curriculum esistente previsto dalle indicazioni nazionali.

Per quanto riguarda la somministrazione di prove di simulazione in preparazione all'esame, gli insegnanti hanno attivato deboli procedure di ricerca di prove sia in rete sia nella pubblicistica. In questo caso viene confermata l'ipotesi di un uso funzionale alla PN ma, a differenza di quanto rilevato nelle ricerche, non possiamo parlare di un vero e proprio *teaching to test* poiché questo tipo di attività viene relegata nel secondo quadrimestre a ridosso della data dell'esame e viene limitata a poche esercitazioni che, a differenza di quelle realizzate dai colleghi di matematica, sono state progettate e realizzate in molti casi senza un vero confronto e accordo con gli altri colleghi di lettere. E anche in questo caso è mancata una delle possibili situazioni per attivare i meccanismi di influenza.

Circa l'*uso dei dati*, rileviamo che questo si è verificato in occasione del confronto, in sede d'esame, degli esiti della PN con le valutazioni date dagli insegnanti durante l'anno scolastico. Questa comparazione, nel caso di marcate discrepanze considerate non imputabili al caso (ottimi risultati nella PN a fronte di valutazioni negative dell'insegnante), ha fatto sorgere «dubbi» sull'efficacia delle strategie didattiche adottate in relazione ai diversi stili cognitivi. A questi dubbi, tuttavia, non hanno fatto seguito modifiche nei comportamenti didattici.

Il *non uso* più rilevante tocca l'ambito della didattica della comprensione del testo. Esiste una differenza abbastanza marcata tra la didattica della comprensione del testo attuata dagli insegnanti e la PN. Per rendersi conto di questo è sufficiente confrontare il quadro teorico di riferimento della PN con

il tipo di esercizi presenti sulle antologie utilizzate dai docenti. Perché, allora, non sono stati attivati dagli insegnanti processi di cambiamento? Di certo, in questo caso, abbiamo la messa in atto di un comportamento di resistenza da parte degli insegnanti, suscitato sia dalla partecipazione ai processi della valutazione sia dai presupposti culturali della prova che confermano quanto emerso dalle indagini IARD. Tuttavia la nostra ricerca aggiunge un aspetto ulteriore ossia una diversità culturale tra il promotore della PN e il mondo della scuola che tocca molteplici aspetti: le scelte didattiche (comprensione del testo vs. produzione scritta; sincronia vs. diacronia; standardizzazione vs. personalizzazione); l'approccio ai testi (comprensione vs. interpretazione); la valutazione (prove non strutturate vs. prove strutturate; eterovalutazione vs. autovalutazione; approccio quantitativo vs. qualitativo). Questa rete di concezioni ha svolto, di fatto, una funzione di «resistenza» alla valutazione (Taut & Brauns, 2003) favorendone un suo «non uso». Questo dato conferma quanto sostenuto già dalla metà degli anni '80 (Clark & Peterson, 1986; Pajares, 1992), e riaffermato di recente (Scott, 2007), vale a dire l'importanza delle concezioni degli insegnanti nei processi valutativi.

Un ulteriore uso mancato riguarda i dati restituiti dall'INVALSI. Gli unici dati relativi alla PN 2008 di cui avevano preso visione alcuni gli insegnanti erano quelli distribuiti dai dirigenti e riferiti agli esiti delle singole classi; non abbiamo registrato nessuna iniziativa specifica per rintracciare ulteriori dati. E questo conferma quanto ipotizzato. Tuttavia, va tenuto presente che per quanto riguarda i risultati della PN 2009 essi non erano stati ancora resi disponibili dall'INVALSI al momento della nostra indagine.

In sintesi gli esiti della ricerca dimostrano una situazione più articolata rispetto a quella delineata in fase di formulazione delle ipotesi. L'introduzione della PN ha provocato negli insegnanti di italiano cambiamenti nella didattica non solo per adempiere ai doveri istituzionali (somministrazione di prove di simulazione), come ipotizzato, ma anche a seguito del riconoscimento del valore pedagogico di una delle scelte della PN (didattica della grammatica) che ha promosso non tanto un «restringimento del curriculum», come in altri contesti internazionali, quanto una sua «rivitalizzazione», diversamente da quanto ipotizzato. In entrambi i casi, tuttavia, il cambiamento è legato all'uso del processo e/o dei presupposti di contesto della valutazione così come previsto. I dati, infatti, hanno avuto un ruolo marginale nel provocare cambiamenti di alcun tipo anche se gli esiti delle PN rilevati immediatamente in sede di correzione hanno favorito l'avvio di un interessante processo riflessivo (uso cognitivo) diversamente da quanto ipotizzato. Possiamo supporre che sul loro uso parziale abbia influito un'esigenza di «utilità immediata» che riuscirebbe a giustificare il non interesse per i dati che riguardano livelli organizzativi più complessi rispetto alla classe.

6. CONSIDERAZIONI CONCLUSIVE

Nella relazione tra insegnanti e prove standardizzate l'impatto sui docenti è uno degli aspetti di una realtà complessa (van den Berg, 2002; Segall, 2003) che il nostro studio esplorativo ha potuto indagare solo in parte poiché condotto su un campione non rappresentativo, autoselezionatosi e con un tipo di analisi che ha consentito una rilevazione degli ambiti macro del problema lasciando sullo sfondo aspetti individuabili mediante altri approcci. Inoltre, l'aver svolto l'indagine a ridosso delle prime somministrazioni della PN ha fatto sì che alcuni aspetti, ad esempio il timore di una ricaduta sugli insegnanti degli esiti della PN, a causa del loro carattere di intrusività siano emersi solo in modo marginale nelle trascrizioni e abbiano trovato spazio, peraltro ridotto, solo nelle note del conduttore prese a «microfono spento». In tal modo, l'uso del contesto politico/istituzionale della valutazione è stato solo marginalmente indagato.

Tuttavia riteniamo che gli esiti a cui siamo giunti possano dare un contributo di un qualche interesse. Sul piano teorico fanno emergere, anche se in modo parziale, la rilevanza della variabile «meccanismi di mediazione» e della sua struttura complessa (meccanismi di utilizzo, di influenza e di resistenza). Dal punto di *vista pratico*, il conflitto culturale emerso tra insegnanti e PN circa il significato della comprensione del testo e la cultura valutativa suggerisce di sviluppare con attenzione, nelle future formazioni di dirigenti e insegnanti, anche questi aspetti.

Gli esiti, inoltre, lasciano intravedere possibili linee di sviluppo per future ricerche. In primo luogo sarà interessante indagare gli effetti sugli insegnanti delle PN 2010 e 2011 il cui esito ha concorso alla valutazione finale degli studenti in egual misura agli esiti delle prove scritte e orali e al giudizio di idoneità (DPR 122/2009). Inoltre, ulteriori indagini sull'uso dei dati contribuiranno a chiarire possibili percorsi di integrazione tra i dati restituiti dall'INVALSI e i sistemi di autovalutazione di istituto così come auspicano anche Domenici e Lucisano (2011).

Infine, la rilevanza delle concezioni degli insegnanti dovrà essere indagata all'interno di disegni di ricerca che abbraccino anche una prospettiva organizzativa per comprendere le strette interrelazioni tra i punti di vista individuali e i limiti che questi trovano nei vincoli organizzativi (Argyris, 1992; Owen & Lambert, 1995).

RIFERIMENTI BIBLIOGRAFICI

- Abrams, L. M., Pedulla, J. J., & Madaus, G. F. (2003). Views from the classroom: Teachers' opinions of Statewide Testing Programs. *Theory into Practice*, 42(1), 18-29.
- Alkin, M. C. (2005). Utilization of evaluation. In S. Mathison (Ed.), *Encyclopedia of evaluation* (pp. 434-436). Thousand Oaks, CA: Sage.
- Argyris, C. (1992). *On organizational learning*. Cambridge, MA: Blackwell.
- Assaf, L. C. (2008). Professional identity of a reading teacher: Responding to high-stakes testing pressures. *Teachers and Teaching: Theory and Practice*, 14(3), 239-252.
- Barone, C. (2010). La valutazione: verso una spirale al ribasso? In A. Cavalli & G. Argentin (a cura di), *Gli insegnanti italiani: come cambia il modo di fare scuola. Terza indagine dell'Istituto IARD sulle condizioni di vita e di lavoro nella scuola italiana* (pp. 159-180). Bologna: il Mulino.
- Bertagna, G. (2004). *Valutare tutti, valutare ciascuno. Una prospettiva pedagogica*. Brescia: La Scuola.
- Bertocchi, D. (2010). Indagini internazionali e nazionali sulla competenza di lettura. *Italiano LinguaDue*, 2(2), 231-246.
- Boardman, A., & Woodruff, A. (2004). Teacher change and «high-stakes» assessment: What happens to professional development? *Teaching and Teacher Education*, 20(6), 545-557.
- Bonami, M. (2005). Évaluation interne et évaluation éternelle: concurrence ou complémentarité? *Les Cahiers de Recherche en Éducation et Formation*, 38, <http://www.girsef.ucl.ac.be>.
- Boyatzis, R. E. (1998). *Transforming qualitative information: Thematic analysis and code development*. Thousand Oaks, CA: Sage.
- Cavalli, A. (a cura di). (2000). *Gli insegnanti nella scuola che cambia. Seconda indagine IARD sulle condizioni di vita e di lavoro nella scuola italiana*. Bologna: il Mulino.
- Cavalli, A., & Argentin, G. (a cura di). (2010). *Gli insegnanti italiani: come cambia il modo di fare scuola. Terza indagine dell'Istituto IARD sulle condizioni di vita e di lavoro nella scuola italiana*. Bologna: il Mulino.
- Chapman, D. W., & Snyder, C. W., Jr. (2000). Can high stakes national testing improve instruction: Reexamining conventional wisdom. *International Journal of Educational Development*, 20, 457-474.
- Christie, C. A. (2007). Reported influence of evaluation data on decision makers' actions: An empirical examination. *American Journal of Evaluation*, 28(1), 8-25.
- Clark, C. M., & Peterson, P. L. (1986). Teachers' thought processes. In M. C. Wittrock (Ed.), *Handbook of research on teaching* (3rd ed., pp. 255-296). New York: Macmillan.

- Corbet, H. D., & Wilson, B. (1991). *Testing, rebellion and reform*. Norwood, NJ: Ablex.
- Cousin, J. B. (2003). Utilization effects of participatory evaluation. In T. Kellaghan, D. Stufflebeam, & L. Wingate (Eds.), *International handbook of educational evaluation* (pp. 245-266). Boston: Kluwer.
- Crocco, M. S., & Costigan, A. T. (2007). The narrowing of curriculum and pedagogy in the age of accountability: Urban educators speak out. *Urban Education*, 42(6), 512-535.
- Diamond, J. B. (2007). Where the rubber meets the road: Rethinking the connection between high-stakes testing policy and classroom practice. *Sociology of Education*, 80, 285-313.
- Domenici, G. (2003). Allargare l'accesso al sapere. *Valore Scuola*, 11, 14-19.
- Domenici, G. (2004). Contro la valutazione di bassa qualità. *Tuttoscuola*, 443, 50-52.
- Domenici, G. (2008). Arriva la quarta prova scritta per gli esami di Stato. *Tuttoscuola*, 483, 36-37.
- Domenici, G. (2009). Consolidare e migliorare una buona scelta. *Tuttoscuola*, 492, 30-32.
- Domenici, G., & Lucisano, P. (2011). Valutazione, conoscenza, processi decisionali. Dibattito. *Journal of Educational, Cultural and Psychological Studies*, 3, 147-167.
- Domingo, P. (2006). Le prove PISA ed INVALSI: possibili conseguenze sulla pratica didattica. *Notiziario della Unione Matematica Italiana*, 33(11/b), 63-69.
- Fandiño Pinella, M. I. (2005). La valutazione in matematica e le prove INVALSI. *La Matematica e la sua Didattica*, 3, 359-371.
- Giovannini, M. L. (a cura di). (2007). *L'impatto delle rilevazioni nelle scuole. I punti di vista della pubblica amministrazione e delle istituzioni scolastiche bolognesi*. Roma: Carocci.
- Giovannini, M. L., & Marcuccio, M. (2007). Un'indagine tra i dirigenti scolastici. In M. L. Giovannini (a cura di), *L'impatto delle rilevazioni nelle scuole I punti di vista della pubblica amministrazione e delle istituzioni scolastiche bolognesi* (pp. 57-94). Roma: Carocci.
- Giovannini, M. L., Marcuccio, M., & Truffelli, E. (2009). L'applicazione dei «metodi misti» nell'analisi di percorsi integrati tra istruzione e formazione professionale. In G. Domenici & R. Semeraro (a cura di), *Le nuove sfide della ricerca didattica tra saperi, comunità sociali e culture* (pp. 21-34). Roma: Monolite.
- Greenbaum, T. L. (1998). *The handbook for focus group research* (2nd ed.). Thousand Oaks, CA: Sage.
- Henry, G. T., & Mark, M. M. (2003). Beyond use: Understanding evaluation's influence on attitudes and actions. *American Journal of Evaluation*, 24(3), 293-314.

- Heubert, J. P., & Hauser, R. M. (Eds.). (1999). *High-stakes: Testing for tracking, promotion and graduation*. Washington, DC: National Academy Press.
- INVALSI (2008). *La prova nazionale al termine del primo ciclo. Aspetti operativi e prime valutazioni sugli apprendimenti degli studenti*, <http://www.INVALSI.it/EsamiDiStato/pagine/evidenza.php>.
- INVALSI (2009). *Esame di Stato I ciclo. A.S. 2008/2009. Prova Nazionale 2009. Prime analisi*, <http://www.INVALSI.it/esamidistato0809/>.
- Johnson, K., Greenesid, L. O., Toal, S. A., King, J. A., Lawrenz, F., & Volkov, B. (2009). Research on evaluation use. A review of the empirical literature from 1986 to 2005. *American Journal of Evaluation*, 30(3), 377-410.
- Kirkhart, K. E. (2000). Reconceptualizing evaluation use: An integrated theory of influence. *New Directions for Evaluation*, 888, 5-23.
- Krueger, R. A. (1994). *Focus group. A pratical guide for applied research*. Newbury Park, CA: Sage.
- Krueger, R. A. (1997). *Analyzing and reporting focus group results*. Thousand Oaks, CA: Sage.
- Lafontaine, D., & Dierendonck, C. (2008). *Comment les évaluations externes des acquis des élèves sont-elles perçues par les enseignants en Communauté française de Belgique et en Suisse romande?*, Actes du 20^e Colloque de l'ADMEE-Europe, Université de Genève, <http://plone.unige.ch/sites/admee08/tables-rondes/m-tr3/m-tr3-1>.
- Lancia, F. (2004). *Strumenti per l'analisi dei testi. Introduzione all'uso di T-LAB*. Milano: Franco Angeli.
- Looney, J. W. (2011). *Integrating formative and summative assessment: Progress toward a seamless system?* OECD Education Working Papers, 58. Paris: OECD Publishing, http://www.oecd-ilibrary.org/education/integrating-formative-and-summative-assessment_5kghx3kbl734-en.
- Losito, B. (2006). Quando la scuola si interroga. Per un sistema di valutazione «amico» delle scuole. *Cadmo*, 1, 43-54.
- Lucisano, P. (2003). Validità e affidabilità delle pratiche valutative: a proposito del Progetto Pilota 2. *Cadmo*, 11(2), 37-56.
- Lucisano, P. (2005). E se qualcuno valutasse i valutatori? Dal pasticcio pilota al pasticcio generale. *Valore Scuola*, 11, 6-11.
- Lumbelli, L. (1981). *Il colloquio non direttivo*. Milano: Franco Angeli.
- Lumbelli, L. (1993). Tecniche di intervista, processi cognitivi e ricerca educativa. *Età Evolutiva*, 46, 38-95.
- Manzotti, E., & Zampese, L. (2010). Le prove d'italiano dell'INVALSI. Un quadro di riferimento. In E. Lugarini, *Valutare le competenze linguistiche* (pp. 47-73). Milano: Franco Angeli.
- Marcuccio, M., & Truffelli, E. (2007). Uno studio di caso. In M. L. Giovannini (a cura di), *L'impatto delle rilevazioni nelle scuole. I punti di vista della pubblica*

- amministrazione e delle istituzioni scolastiche bolognesi* (pp. 111-135). Roma: Carocci.
- Mark, M. M., & Henry, G. T. (2004). The mechanisms and outcomes of evaluation influence. *Evaluation*, 10(1), 35-57.
- Massey, O. T. (2011). A proposed model for the analysis and interpretation of focus groups in evaluation research. *Evaluation and Program Planning*, 34, 21-28.
- Morgan, D. L. (1997). *Focus group as qualitative research* (2nd ed.). London: Sage.
- Nevo, D. (2002). Dialogue evaluation: Combining internal and external evaluation. *School-Based Evaluation: An International Perspective*, 8, 3-16.
- Nidegger, C. (2008). *Epreuves externes et enseignants: des positions inconciliables?*, Actes du 20^e Colloque de l'ADMEE-Europe, Université de Genève, <http://plone.unige.ch/sites/admee08/tables-rondes/m-tr3/m-tr3-2>.
- Owen, J. M., & Lambert, F. C. (1995). Roles for evaluation in learning organizations. *Evaluation*, 1(2), 237-250.
- Pajares, M. F. (1992). Teachers' beliefs and educational research: Cleaning up a messy construct. *Review of Educational Research*, 62(3), 307-332.
- Patton, M. Q. (2008). *Utilization-focused evaluation*. Thousand Oaks, CA: Sage.
- Ryan, G. W., & Bernard, H. R. (2003). Techniques to identify themes. *Field Methods*, 15, 85-109.
- Sandrone, G. (2009). Forse una proposta per il collasso educativo della scuola? A proposito delle ipotesi INVALSI sulla valutazione esterna. *Orientamenti Pedagogici*, 56(3), 419-437.
- Scott, C. (2007). Stakeholder perceptions of test impact. *Assessment in Education*, 14(1), 27-49.
- Segall, A. (2003). Teachers' perceptions of the impact of state-mandated standardized testing: The Michigan Educational Assessment Program (MEAP) as a case study of consequences. *Theory and Research in Social Education*, 31(3), 287-325.
- Stake, R. E. (1998). Some comments on assessment in U.S. education. *Education Policy Analysis Archives*, 6(14), <http://olam.ed.asu.edu/epaa/v6n14.html>.
- Taut, S., & Brauns, D. (2003). Resistance to evaluation. A psychological perspective. *Evaluation*, 9(3), 247-264.
- Tuzzi, A. (2003). *L'analisi del contenuto*. Roma: Carocci.
- van den Berg, R. (2002). Teachers' meanings regarding educational practice. *Review of Educational Research*, 72(4), 577-625.
- Vaughn, S., & Schumm, J. S., & Sinagub, J. (1996). *Focus group interviews in education and psychology*. Thousand Oaks, CA: Sage.
- Vertecchi, B. (2004). Due linee per la valutazione del sistema scolastico. *Tuttoscuola*, 446, 36-38.

- Yeh, S. S. (2005). Limiting the unintended consequences of high-stakes testing. *Education Policy Analysis Archives*, 13(43), 1-21, <http://epaa.asu.edu/epaa/v13n43/>.
- Zammuner, V. L. (2003). *I focus group*. Bologna: il Mulino.
- Zancanella, D. (1992). The influence of state-mandated testing on teachers of literature. *Educational Evaluation and Policy Analysis*, 14(3), 283-295.

RIASSUNTO

L'articolo presenta i risultati di un'indagine che intende descrivere i primi effetti sugli insegnanti di italiano dell'introduzione della Prova Nazionale (PN) nell'Esame di Stato al termine del I ciclo d'istruzione. Lo studio si colloca nel filone di ricerche sull'uso delle valutazioni e assume come riferimento teorico un adattamento della teoria dell'influenza di Mark e Henry (2004). Le opinioni e le dichiarazioni di pratiche di 41 docenti sono state raccolte mediante otto focus group realizzati in altrettanti istituti secondari di I grado di un comune del nord Italia. L'analisi delle trascrizioni è stata effettuata con un approccio di tipo quanti-qualitativo che ha consentito di individuare undici ambiti del discorso che riguardano la rilevanza e le procedure della PN, le attività degli insegnanti ad essa correlate e le reazioni degli studenti. L'analisi dei nuclei tematici mette in evidenza che la PN ha provocato alcuni cambiamenti comportamentali e cognitivi negli insegnanti ma non tutti quelli possibili a causa di resistenze da parte dei docenti verso i presupposti pedagogici della PN. Nelle conclusioni è discussa la portata euristica del modello assunto e sono avanzate alcune proposte per lo sviluppo delle ricerche sull'impatto delle valutazioni.

Parole chiave: Concezioni degli insegnanti, Esame di Stato (I ciclo), Prova Nazionale, Uso della valutazione, Valutazione del sistema scolastico.