

Il debito può essere formativo? Una ricerca per indagare l'efficacia didattica degli esami a settembre

Can the school failure be formative? A study to investigate the educational effectiveness of the summer exams

Davide Parmigiani

Università di Genova - davide.parmigiani@unige.it

Cecilia Marsic

Università di Genova - cecilia.marsic@gmail.com

Elisabetta Nicchia

Università di Genova - elisabettanicchia@gmail.com

Cecilia Russo

Università di Genova - russocecilia@outlook.it

ABSTRACT

The students of the upper secondary schools can have poor grades in one or more subjects at the end of the school year. If so, they have to attend summer courses and face exams either at the end of August or beginning of September. The department of Education at University of Genoa carried out a 2-year research aimed at investigating the educational relevance of those paths. The purposes of this study were: calculating the number of students and the frequency of subjects with poor grades; examining the critical aspects related to these exams (economical, organisational, emotional, etc.); focusing the summative and formative factors supported or not by these assessment tests. During the first year, we conducted a probing study aimed at creating a reliable research instrument whilst, during the second year, the questionnaire was administered to 900 students, 150 teachers and 150 parents from 5 upper secondary schools. The findings indicate several difficulties connected to these exams. In particular, we noticed limits concerning the economical, organisational and formative aspects.

Gli studenti delle scuole secondarie di II grado possono giungere al termine dell'anno scolastico con debiti formativi in una o più discipline. In questo caso, essi devono frequentare dei corsi di recupero e affrontare delle prove

* Attribuzione delle parti. Davide Parmigiani: organizzazione e strutturazione della ricerca, stesura dei paragrafi 3, 4, 5 (tranne 5.2) e 6; Cecilia Marsic, Elisabetta Nicchia e Cecilia Russo: somministrazione degli strumenti, analisi dei dati qualitativi, stesura dei paragrafi 1, 2 e 5.2.

aggiuntive, calendarizzate a fine estate. Il dipartimento di Scienze della Formazione dell'Università di Genova ha condotto una ricerca biennale finalizzata ad indagare la significatività didattica e valutativa di tali percorsi. Gli obiettivi dello studio intendevano far emergere il numero di studenti e la frequenza delle discipline oggetto di sospensione del giudizio. Inoltre, sono stati analizzati gli aspetti critici relativi sia ai corsi di recupero sia agli esami a settembre. Infine, sono stati investigati gli aspetti sommativi e formativi che sono supportati o meno dalle prove valutative estive. Nel corso del primo anno della ricerca, è stato svolto uno studio pilota che ha permesso di elaborare uno strumento di ricerca affidabile mentre, nel secondo anno, lo strumento è stato somministrato a 900 studenti, 150 insegnanti e 150 genitori di cinque scuole secondarie. I risultati sottolineano diverse criticità legate a queste prove valutative. In particolare, vengono segnalati limiti che riguardano gli aspetti economici, organizzativi e formativi.

KEYWORDS

Supplementary Study Credits, Summative Assessment, Formative Assessment, Upper Secondary School.

Debiti Formativi, Valutazione Sommativa, Valutazione Formativa, Scuola Secondaria di II Grado.

1. Introduzione

Nella storia della scuola italiana, si sono alternate diverse modalità valutative e svariate prove di verifica. Inizialmente, gli esami conclusivi di un periodo scolastico o gli esami per riparare le insufficienze, erano presenti in tutti gli ordini di scuola. Progressivamente, molti esami sono stati soppressi oppure trasformati in prove di recupero.

In ogni caso, alcune domande erano sempre presenti: sono efficaci? Servono allo studente per colmare le lacune disciplinari? O, viceversa, si rivelano superflui, tendenzialmente sterili e improduttivi da un punto di vista disciplinare?

Ancora oggi, esse assalgono gli studenti, gli insegnanti e i genitori che devono affrontare quelli che venivano chiamati esami di riparazione. Il termine "riparazione" è particolarmente evocativo: lo studente ha evidenziato carenze e deve recuperare il tempo perduto. Dal 1995 la terminologia è mutata per sottolineare che anche la scuola deve farsi carico di un percorso di recupero attraverso l'organizzazione di attività di studio aggiuntive e prove che aiutino lo studente a sanare un debito formativo.

La presente ricerca ha provato a rispondere alle domande illustrate in precedenza, coinvolgendo gli studenti, gli insegnanti e i genitori di cinque scuole secondarie di II grado della provincia della Spezia. Lo studio è stato biennale ed ha indagato le idee, le percezioni e gli aspetti che caratterizzano i percorsi che, da un'insufficienza, portano lo studente a sperimentare corsi di recupero e prove valutative, generalmente calendarizzate a fine estate. Sulla base di queste motivazioni, l'obiettivo generale della ricerca può essere sintetizzato nel modo seguente: qual è la significatività didattica e valutativa dei corsi di recupero e delle prove finali?

2. Cornice storico-normativa

L'espressione "esami di riparazione" venne introdotta dalla Riforma Gentile (1923). Al termine di ogni anno, chiunque non riuscisse a raggiungere una votazione pari a 6/10, in qualunque ordine e grado, era tenuto a sostenere degli esami di riparazione, per poter accedere alla classe successiva. Queste prove di verifica furono regolamentate da due Regi Decreti (653 del 4/05/25 e 2049 del 21/11/29) che dettavano le tempistiche e le modalità del recupero.

Con la L. 119 del 5/04/69, il Ministro Sullo modificò la struttura degli esami di maturità e abolì la sessione autunnale.

Il Ministro Malfatti abolì gli esami di riparazione nelle scuole elementari e medie (L. 517 del 4/08/77, artt. 3 e 6) e introdusse una nuova forma di valutazione. Gli insegnanti dovevano compilare, ogni tre mesi, una scheda di osservazione dell'alunno, da discutere con i genitori. Infine, si elaborava un giudizio di idoneità per il passaggio alla classe o all'ordine di scuola successivi. Iniziava a delinearsi il contorno di una valutazione formativa, che non tenesse conto solo dei risultati raggiunti ma anche dei processi di apprendimento.

Nelle scuole superiori, il processo di avvicinamento ad una valutazione formativa fu più complesso. Gli esami di riparazione furono sospesi per un breve periodo: il Ministro D'Onofrio li abolì con il DL 253 del 28/06/95. Al loro posto vennero introdotti gli IDEI, interventi didattici ed educativi integrativi, adottati ed organizzati dal Collegio dei docenti e dai Consigli di classe. Ogni istituto aveva il compito di stabilire «le modalità temporali ed organizzative, anche con opportuni adattamenti del calendario scolastico» (art.193 ter). La messa in atto di azioni di sostegno durante la fase di apprendimento, e non al suo termine, fu uno degli aspetti maggiormente innovativi. Risalgono a questo periodo i primi corsi di recupero organizzati dalle scuole che dovevano essere frequentati dagli studenti che non avevano conseguito la sufficienza in alcune discipline. Il Consiglio di classe aveva anche il compito di attuare verifiche periodiche affinché si potesse monitorare l'efficacia di questi interventi (art.193 bis).

La valenza formativa degli esami di riparazione emerse nel 2007, quando questi furono riattivati dal Ministro Fioroni, sebbene con nuova terminologia e attraverso nuove modalità. Sulla base della L. 1 dell'11/01/07, fu introdotta l'espressione "debito formativo" (DM 42 del 22/05/07 e 80 del 3/10/07) e venne decretato l'obbligo di recupero del debito entro la fine dello stesso anno scolastico per poter essere ammessi alla classe successiva. Il Collegio docenti ed i singoli Consigli di classe avevano il compito di predisporre interventi didattici di recupero e di sostegno, che divennero parte integrativa del POF. Vennero introdotte delle verifiche in itinere, al fine di colmare le carenze rilevate durante gli scrutini intermedi, che potevano svolgersi secondo diverse modalità. Il mancato raggiungimento della sufficienza portava al rinvio della formulazione del giudizio finale. Durante il periodo estivo, ogni scuola doveva organizzare dei corsi «finalizzati al recupero dei debiti formativi» (DM 80, art. 5), al termine dei quali era prevista una verifica. Il percorso doveva concludersi «entro il 31 agosto dell'anno scolastico di riferimento e comunque non oltre la data di inizio delle lezioni dell'anno scolastico successivo» (art. 6). Tecnicamente, quindi, non si parla più di esami di riparazione bensì di prove di verifica da attuarsi sia in itinere che al termine del percorso di recupero. Queste modalità sono tuttora in vigore e costituiscono il contesto in cui si inserisce la nostra ricerca.

3. I fondamenti teorici

La valutazione sommativa può essere definita come l'indicazione dei livelli di apprendimento raggiunti dall'alunno in un determinato momento di un percorso didattico. In pratica, la valutazione sommativa sottolinea l'aspetto relativo alla misurazione intesa «come accertamento del profitto, [in modo da] definire le prestazioni degli allievi e apprezzarle secondo una gamma di valori» (Becchi, 1995, p. 5). In genere, l'aspetto sommativo della valutazione viene focalizzato ed evidenziato nel corso di esami o prove di vario tipo: strutturate, semistrutturate o destrutturate (Vertecchi, 2003).

La valutazione formativa, invece, può essere definita come «un'informazione comunicata all'alunno finalizzata al cambiamento delle sue strategie di studio con l'obiettivo di migliorare il suo apprendimento» (Shute, 2008, p. 154). Essa accompagna tutte le fasi del processo valutativo e ha lo scopo di offrire un feedback allo studente e al docente sullo sviluppo dell'apprendimento e sull'evoluzione del processo didattico (Scriven, 1967; Parmigiani, 2016). La valutazione formativa viene effettuata durante lo sviluppo e la realizzazione di un intervento didattico, quindi viene realizzata in itinere con funzioni di regolazione e monitoraggio (Weurlander et al., 2012; Wiliam, 2011; Cizek e Andrade, 2010; Heritage, 2010; McMillan, 2007; Bell e Cowie, 2001; Torrance e Pryor, 1998; Wiliam e Black, 1996; Sadler, 1989). In pratica, l'obiettivo profondo della valutazione formativa risulta quello di avviare percorsi indirizzati a fornire agli alunni feedback circa le loro azioni al fine di un riorientamento migliorativo (Bondioli e Ferrari, 2008, pp. 20-21).

Da un lato, la caratteristica essenziale della valutazione sommativa è rappresentata da un giudizio sull'apprendimento dello studente al fine di evidenziarne gli esiti. Invece, è «formativa qualsiasi valutazione che aiuta l'alunno a imparare; egli partecipa alla regolazione degli apprendimenti e dello sviluppo nella direzione di un progetto educativo» (Perrenoud, 1991 citato da Dionisi e Tadiello, 2007, p. 34).

Queste due prospettive appaiono distanti e separate ma sono, di fatto, contigue: da un lato, è necessario conoscere gli sviluppi dell'apprendimento attraverso una certificazione degli esiti; dall'altro, è opportuno accompagnare il percorso didattico di momenti formativi al fine di sollecitare meccanismi di adattamento e autoregolazione (Giannandrea, 2009).

Nel contesto della nostra ricerca, intendiamo analizzare se la valutazione sommativa e quella formativa sono effettivamente bilanciate oppure se viene evidenziata una loro separazione. In pratica, vogliamo investigare se, nel corso dei corsi di recupero e nelle prove finali ad agosto/settembre, vengono maggiormente evidenziati gli aspetti sommativi, quelli formativi oppure se emergono entrambi.

La base teorica su cui fondiamo la nostra azione di ricerca è espressa da Wininger (2005), il quale afferma che la valutazione sommativa e formativa, insieme, possono diventare strategie didattiche da offrire all'alunno per aiutarlo a sviluppare un processo di autovalutazione e di auto-orientamento, teso a sviluppare al meglio le proprie potenzialità. In particolare, egli presenta una sorta di unione – un *marriage* – fra la valutazione sommativa e quella formativa e la definisce *formative summative assessment*. Essa consiste nello stretto collegamento fra gli aspetti quantitativi, tipici della valutazione sommativa, e quelli qualitativi, specifici di quella formativa. Nel caso della nostra ricerca, intendiamo vedere se i corsi di recupero e le prove finali estive riescono ad armonizzare gli aspetti sommativi e formativi della valutazione, come suggerito da Black e Wiliam (1998) con la teoria del *balanced assessment*.

4. Il disegno della ricerca

4.1. Una ricerca biennale: lo studio pilota

L'obiettivo generale della ricerca è finalizzato a scandagliare la significatività didattica e valutativa degli esami a settembre. Per perseguire tale obiettivo, è stata organizzata un'indagine biennale che ha compreso uno studio pilota (condotto nel corso dell'as 17/18) e la ricerca allargata ad un campione significativo nel corso dell'as 18/19.

Lo studio pilota è stato condotto presso il Liceo Mazzini di La Spezia ed è stato rivolto all'elaborazione e alla taratura di uno strumento per poterlo sottoporre, successivamente, alle altre scuole secondarie della provincia. Tale studio esplorativo era necessario per esaminare i punti critici relativi agli esami a settembre, in modo da far emergere le categorie principali da inserire nello strumento.

La domanda di ricerca riferita allo studio pilota era la seguente: quali sono le principali problematiche, legate agli esami a settembre, percepite da studenti, famiglie e insegnanti?

Fra febbraio e maggio 2017, sono stati coinvolti 42 partecipanti: 23 studenti (STU) di diverse classi che hanno sostenuto gli esami a settembre; 15 insegnanti (INS) di diverse discipline; 4 genitori (GEN) di studenti che hanno sperimentato gli esami a settembre. Tutti gli INS e 9 STU sono stati sottoposti ad un'intervista semistrutturata. Infine, sono stati organizzati 5 focus group: 1 con i GEN e 4 con i restanti 14 STU.

L'analisi qualitativa dei dati ha consentito di far emergere le categorie concettuali relative alle problematiche legate ai corsi di recupero e agli esami a settembre. In questo modo, è stato possibile realizzare un pilota di un questionario strutturato, formato da item aperti e chiusi che è stato sottoposto a tutti i partecipanti. L'analisi quantitativa dei dati ha consentito di testare l'affidabilità dello strumento.

4.2. Le domande e il disegno della ricerca

Sulla base dei risultati dello studio pilota, è stata avviata la seconda fase della ricerca centrata sulle seguenti domande:

1. Qual è la percentuale di STU con debito formativo (DF) nelle scuole secondarie di II grado della provincia di La Spezia?
2. Quali sono le discipline in cui si riscontrano il maggior numero di STU con DF?
3. In che modo incidono gli aspetti economici, temporali, relazionali e organizzativi sui corsi di recupero e sulle prove di verifica a settembre?
4. Quali aspetti sommativi sono sviluppati dalle prove di verifica a settembre?
5. Quali aspetti formativi sono sviluppati dalle prove di verifica a settembre?

Le prime due domande sono state approssimate attraverso la raccolta e l'analisi dei dati forniti dalle scuole e dall'USR della Liguria. Le altre domande sono state affrontate attraverso una *survey*, costruita sulla base di un disegno di ricerca fondato sui *mixed methods*. In particolare, abbiamo adottato un *concurrent mixed method design*, una struttura di ricerca di tipo quanti-qualitativo, ideata seguendo le indicazioni di Tashakkori e Teddlie (1998, 2003, 2009). Il disegno viene definito

concurrent in quanto i dati quantitativi e qualitativi vengono raccolti simultaneamente nel corso dell'indagine. Questo tipo di *mixed method design* viene definito *multistrand* in quanto entrambe le tipologie di dati sono raccolte ed analizzate per rispondere alle medesime domande di ricerca. Le inferenze finali si basano sull'analisi concomitante e simultanea dei dati.

Inoltre, non ci sono priorità o subalternità di una tipologia di dati rispetto all'altra, in quanto l'obiettivo principale è la triangolazione dei dati al fine di una profonda comprensione del fenomeno. Tale approccio, definito *concurrent triangulation design*, prevede che i dati quantitativi e qualitativi siano utilizzati per «confirm, cross-validate, or corroborate findings within a single study» (Creswell et al., 2003, p. 229).

4.3. Il contesto e i partecipanti

La seconda fase della ricerca, condotta nell'as 2018/19, è stata allargata a 5 scuole secondarie di II grado della provincia della Spezia: oltre al Liceo Mazzini, sono stati coinvolti l'IIS Parentucelli-Arzelà, l'ITET Fossati-Da Passano, l'IIS Cardarelli e l'IPSEOA Casini.

I partecipanti alla ricerca sono stati: 910 STU, 142 INS e 151 GEN. Per quanto riguarda gli STU, nell'ambito dei diversi istituti e degli indirizzi, sono state sorteggiate le classi a cui sottoporre lo strumento di ricerca. In questo modo abbiamo potuto comporre un campione stratificato proporzionale sulla base di tre strati: genere, tipologia di scuola e debiti formativi.

Strato	Modalità	Frequenze osservate	Frequenze attese
Genere	F	56,71	57,93
	M	43,29	42,07
Tipologia di scuola	LIC	49,78	45,73
	TEC	32,19	33,88
	PROF	18,02	20,42
Debito formativo	SI	31,87*	22,34
	NO	68,13*	77,66

Tab. 1 – Struttura del campione degli STU (percentuali)

Nella tabella 1, vengono illustrate le caratteristiche del campione STU. Le frequenze osservate si riferiscono alle percentuali rilevate nel campione mentre le frequenze attese sono le percentuali riferite alla popolazione scolastica degli istituti secondari della provincia della Spezia. Gli strati "genere" e "tipologia di scuola" non presentano differenze statisticamente significative. Viene rilevata, invece, una differenza fra le percentuali relative allo strato "debito formativo", in quanto le classi sorteggiate comprendevano le classi dalla seconda alla quinta mentre gli STU potenzialmente interessati fanno parte delle classi dalla prima alla quarta. Siccome il numero totale degli STU diminuisce con il progredire delle classi, la percentuale di STU col debito formativo nel campione risulta più elevata. Per tali motivi, possiamo definire il campione STU significativo e rappresentativo della popolazione.

Strato	Modalità	Frequenze osservate	Frequenze attese
Genere	F	78,17	69,78
	M	21,83	30,22
Età	fino a 34	5,6	2,8
	fra 35 e 44	18,3	16,02
	fra 45 e 54	28,9	31,31
	oltre 54	47,2	49,87
Aree disciplinari	Ling/Lett	18,31	36,84*
	Mat/Sci	23,94*	45,51
	Ling/Str	18,31	17,65
Anzianità	meno di 20	48,59	
	più di 20	51,41	
Debiti formativi negli ultimi 5 anni	Nessuno	33,81	
	da 1 a 19	36,62	
	da 20 a 49	21,13	
	da 50 a 180	8,45	

Tab. 2 – Struttura del campione degli INS (percentuali)

Nella tabella 2, vengono illustrate le caratteristiche del campione INS. In questo caso, sono stati esaminati 5 strati. Il genere e l'età non presentano differenze statisticamente significative rispetto alla popolazione degli INS della provincia della Spezia. Si rileva, invece, una differenza importante fra le aree disciplinari. Pochi INS di area matematico/scientifica hanno partecipato alla ricerca. Per gli strati relativi all'anzianità e alla quantità di debiti assegnati dagli INS negli ultimi 5 anni, non vi sono dati disponibili per verificarne la rispondenza fra campione e popolazione ma ci serviranno per le analisi. In definitiva, il campione insegnanti non è pienamente rappresentativo della popolazione ma può fornire spunti di riflessione rilevanti.

Strato	Modalità	Frequenze osservate
Genere	F	86,09
	M	13,91
Età	fino a 34	0,66
	fra 35 e 44	15,23
	fra 45 e 54	73,51
	oltre 54	10,60
Tipologia di scuola	LIC	79,47
	TEC	17,22
	PROF	3,31*
Debito formativo	SI	33,1
	NO	66,9

Tab. 3 – Struttura del campione dei GEN (percentuali)

Nella tabella 3, vengono illustrate le caratteristiche del campione GEN. In questo caso, non vi sono dati a cui fare riferimento. L'elemento più critico è rappresentato dalla scarsa partecipazione dei GEN degli studenti degli istituti professionali.

4.4. Lo strumento

Lo strumento di ricerca era un questionario composto sia da item chiusi che aperti e suddiviso in 5 aree principali:

1. L'area "dati personali" era rivolta a raccogliere i dati personali utili per descrivere il rispondente:
 - a. STU: età, genere, tipologia di scuola, classe frequentata
 - b. INS: età, genere, anzianità, tipologia di scuola, area disciplinare
 - c. GEN: età, genere, tipologia di scuola
2. L'area "debiti formativi" era focalizzata sulle eventuali esperienze di corsi di recupero e prove a settembre:
 - a. STU: avere avuto o no DF, numero di debiti nelle diverse classi, discipline oggetto di sospensione del giudizio, eventuali bocciature al termine delle prove
 - b. INS: avere o no assegnato DF, numero di debiti assegnati negli ultimi 5 anni, eventuali bocciature al termine delle prove
 - c. GEN: avere o no figli con DF, numero di debiti nelle diverse classi, discipline oggetto di sospensione del giudizio, eventuali bocciature al termine delle prove
3. L'area "aspetti rilevanti" era centrata sui fattori che incidono sui corsi di recupero e le prove a settembre; in particolare, erano evidenziate le seguenti sottoree:
 - a. Aspetti economici, legati ai costi (5 item)
 - b. Aspetti temporali, legati al periodo di studio (7 item)
 - c. Aspetti emozionali, legati alle relazioni (3 item)
 - d. Aspetti organizzativi, legati alle questioni logistiche relative ai corsi e alle prove (4 item)
4. L'area "valutazione" era mirata a rilevare se i corsi di recupero e le prove di verifica sostengono e/o sviluppano gli aspetti
 - a. Sommativi (6 item)
 - b. Formativi (12 item, suddivisi in 3 scale: 5 item per la scala metacognitiva; 3 item per la scala auto-riflessiva; 4 item per la scala emotivo-motivazionale)
5. L'area "idee" era indirizzata a suscitare idee ed opinioni sui corsi di recupero e sulle prove di verifica.

Le prime due aree presentavano diversi strumenti di misura. La terza area era costituita da un numero complessivo di 19 item con una scala Likert a 4 livelli (da Assolutamente SI a Assolutamente NO) con la possibilità di indicare "Non ho intenzione di rispondere" per evitare una risposta forzata con conseguente bias cognitivo. Sono state costruite tre versioni in parallelo del questionario, ciascuna rivolta ad un gruppo di partecipanti, per comparare i dati fra STU, INS e GEN.

Le procedure di somministrazione erano differenziate per le tre categorie di partecipanti. Dopo il sorteggio delle classi, un ricercatore si recava nelle classi e guidava gli STU nella compilazione dello strumento attraverso l'utilizzo del proprio dispositivo digitale (cellulare, tablet). Un ricercatore è rimasto nella sala professori in modo da guidare gli INS nella compilazione del questionario in forma cartacea o digitale. I genitori sono stati avvisati dalle segreterie e sono stati forniti del link per compilare online lo strumento.

5. Analisi dei dati

5.1. I dati quantitativi

La prima domanda di ricerca era finalizzata ad individuare la percentuale di STU con DF. Sono stati analizzati gli STU suddivisi per genere, per tipologia di scuola e per classi di frequenza.

Variabili	Modalità	Campione	Popolazione (as 2017/18)
Genere	F	28,11	19,13
	M	38,61* ¹	26,63* ²
Tipologia di scuola	LIC	24,28* ³	20,79* ⁴
	TEC	37,88* ³	25,16* ⁴
	PROF	42,07* ³	21,71

¹ $\chi^2=7,790$ $p<,005$ residui standardizzati corretti +2,8
² $\chi^2=31,298$ $p<,000$ rsc +5,6
³ $\chi^2=24,755$ $p<,000$ rsc -4,9 (LIC), +2,7 (TEC), +3,1 (PRO)
⁴ $\chi^2=8,494$ $p<,014$ rsc -2,3 (LIC), +2,9 (TEC)

Tab. 4 – Studenti con debito formativo suddivisi per genere e tipologia di scuola (percentuali)

Gli STU di genere maschile presentano una percentuale di DF più elevata, statisticamente significativa, sia nel campione che nella popolazione di riferimento. Fra le tipologie di scuola, gli STU dei licei mostrano un livello più basso di DF sia nel campione sia nella popolazione, mentre fra gli STU dei tecnici il valore è più elevato. Gli STU con DF dei professionali del campione hanno ricevuto maggiori DF rispetto a quelli della popolazione. Per concludere l'analisi della prima domanda di ricerca, nel campione sono stati evidenziati gli STU bocciati dopo le prove di verifica a settembre: rappresentano il 3,1% degli STU con DF. Non vi sono dati relativi alla popolazione.

Le classi in cui sono presenti maggiori STU con DF sono la 1a (22,2%), la 2a (24,4%) e la 3a classe (23,4%). Non vi sono differenze statisticamente significative fra queste classi. Mentre in 4a classe, gli STU con DF scendono statisticamente al 18,5% ($\chi^2=11,106$ $p<,011$ rsc -3,1). È interessante notare che, in classe 2a, gli STU con maggiore percentuale di DF, sono quelli dei tecnici: 30,1% ($\chi^2=7,931$ $p<,019$ rsc +2,6).

La seconda domanda di ricerca intendeva far emergere le discipline che riscontrano il maggior numero di STU con DF.

Area	% area	Discipline	% disciplina all'interno dell'area
Linguistico/Letteraria	20,64	Italiano	33,6
		Storia	21,6
		Latino	19,5
		Filosofia	6,2
Matematico/Scientifica	57,81	Matematica	49,9
		Fisica	10,1
		Economia aziendale	6,6
		Chimica	6,5
Lingue Straniere	21,55	Informatica	5,1
		Inglese	65,9
		Francese	19,7
		Tedesco	12,8

Tab. 5 – Discipline oggetto di debito formativo

Come si può notare, le discipline di ambito matematico/scientifico rappresentano la maggior parte di DF con, in particolare, il numero particolarmente elevato di debiti in Matematica. All'interno delle altre due aree, Inglese e Italiano sono le discipline con maggior numero di debiti. L'elenco è parziale in quanto non sono state indicate le discipline con numeri molto bassi di debiti.

È interessante notare che, all'interno delle differenze di genere, gli STU maschi hanno un numero di debiti statisticamente maggiore in Italiano, Storia, Latino, Chimica e Informatica, invece, in Matematica, Fisica, Economia aziendale, Inglese e Francese non vengono rilevate differenze di genere.

La terza, la quarta e la quinta domanda di ricerca miravano ad evidenziare l'incidenza di alcuni aspetti critici relativi ai corsi di recupero e le prove di verifica a settembre.

L'affidabilità dello strumento è stata misurata mediante l' α di Cronbach che ha presentato i seguenti risultati: aspetti economici (,622); aspetti temporali (,789); aspetti emozionali (,679); aspetti organizzativi (,620); aspetti sommativi (,734); aspetti formativi (,905).

Aspetti rilevanti	item	STU	INS	GEN
Economici	(ec1) necessità ripetizioni	75,5	52,3*	77,5
	(ec2) costo ripetizioni	66,4	63,3	87,7*
	(ec3) acquisto materiali didattici	21,6	2,3*	13,7
	(ec4) acquisto materiali digitali	11,6	5,4	8,6
	(ec5) costo generale esami	42,8	32,6	74,5*
Temporali	(t1) estate periodo troppo breve	64,3*	45,8*	52,8
	(t2) meglio a giugno	49,2*	22,1*	40,1
	(t3) estate non adatta studio approfondito	81,9*	55,6*	68,8*
	(t4) no tempo interessi personali	79,3*	33,1*	60,2*
	(t5) no tempo per amici	70,4*	18,2*	50,8*
	(t6) no tempo per famiglia	43,2*	9,4*	27,1*
	(t7) famiglia costretta riorganizzare vacanze	60,8	51,8*	70,9*
Emozionali	(em1) esami incidono positivamente sul clima familiare	18,7	16,8	27,3
	(em2) esami incidono positivamente su serenità personale studente	20,7	16,5	27,7
	(em3) esami incidono positivamente su rapporto insegnanti/studenti	30,9	37,2	36,7
Organizzativi	(o1) prove esami adatte	71,6	80,3*	65,1
	(o2) modalità corsi recupero adatte	59,4	59,4	43,2*
	(o3) esami sono difficili	46,3	5,1*	38,1
	(o4) esami svolgono ruolo positivo e sono utili	51,4	57,1	60,2

Tab. 6 – Dati area “Aspetti rilevanti” between i partecipanti (% risposte positive)

La tabella 6 riporta i dati relativi alle risposte positive nelle diverse aree, relative alla terza domanda di ricerca, suddivise fra i partecipanti. I dati sono stati aggregati nel livello “positivo” (somma delle risposte Assolutamente SI e più SI che NO) e nel livello “negativo” (somma delle risposte Assolutamente NO e più NO che SI). I dati sono stati esaminati come non parametrici e le differenze statisticamente significative sono state misurate con il test di Kruskal-Wallis per campioni indipendenti.

Oltre alle differenze “between” i partecipanti, è interessante sottolineare le differenze “within” i partecipanti. In particolare, si notano differenze riguardanti il genere e fra gli STU con e senza DF. Le differenze statisticamente significative sono state misurate con il test U di Mann-Whitney e di Kruskal-Wallis per campioni indipendenti e il χ^2 . Nel primo caso, le studentesse denotano maggiore consapevolezza per gli aspetti economici (ec1, ec2), maggiori difficoltà legate al periodo di studio (t6 e t7), maggiori limiti per quelli emozionali (em1, em2, em3). Nel secondo caso, invece, gli STU con DF sottolineano le difficoltà organizzative dei corsi di recupero (o2) e la scarsa utilità/efficacia delle prove di verifica finali (o4).

Queste ultime considerazioni valgono anche per gli INS che assegnano molti DF e quelli con più di 20 anni di anzianità. Queste due categorie segnalano punteggi più bassi per o2 e o4.

La quarta domanda di ricerca focalizza gli aspetti sommativi della valutazione. Ad una prima analisi, le risposte delle tre tipologie dei partecipanti, si assomigliano molto. Per i primi 3 item (vedi Tab. 7), relativi alla possibilità di imparare gli argomenti, raggiungere gli obiettivi e sviluppare le competenze previste dal DF, le risposte si polarizzano intorno al valore del 50%. STU, INS e GEN si suddividono in due gruppi di dimensioni analoghe, senza differenze significative.

Aspetti sommativi item	STU	INS	GEN
(s1) durante l'estate posso imparare tutti gli argomenti previsti	53,9	47,9	51,5
(s2) durante l'estate posso raggiungere tutti gli obiettivi previsti	51,6	48,9	51,9
(s3) durante l'estate posso sviluppare tutte le competenze previste	54,9	47,1	54,6
(s4) le valutazioni assegnate rispecchiano il percorso che ho effettuato in estate	56,2	64,5	63,2
(s5) le valutazioni devono basarsi sui livelli richiesti (assoluto)	81,3	88,6	86,5
(s6) le valutazioni devono considerare i miei miglioramenti (individuale)	89,4	93,6	92,5

Tab. 7 – Dati area “Valutazione”, aspetti sommativi between i partecipanti (% risposte positive)

Un'analisi più approfondita dei dati within i partecipanti, ci consente riflessioni più articolate. Nella fig. 1, vengono mostrati i dati differenziati fra gli STU con debito formativo (DF) e quelli senza debito formativo (NON DF). Si nota come gli STU DF percepiscano in misura minore l'efficacia sommativa dei corsi di recupero e delle prove a settembre, in particolare per gli item s2 e s4.

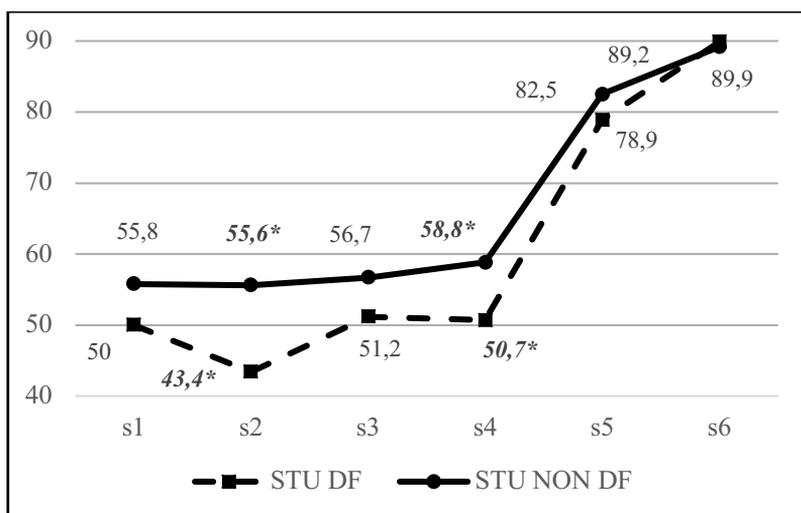


Fig. 1 – Dati area “Valutazione”, aspetti sommativi within gli studenti (% risposte positive)

L’analisi dei dati relativi agli aspetti formativi (quinta domanda di ricerca) riserva risvolti maggiormente interessanti. Nella tab. 8, vengono elencati i dati relativi agli aspetti formativi suddivisi nelle tre scale indicate. Gli INS assegnano punteggi minori nella scala metacognitiva rispetto agli STU, mentre tutti i partecipanti attribuiscono punteggi molto bassi alla scala emotivo-motivazionale.

Aspetti formativi		STU	INS	GEN
scale	item			
m1	autoregolazione	66,1	58,7	59,8
m2	gestione discipline	60,1	58,4	49,2*
m3	elaborare informazioni e materiali	63,1*	48,9*	60,1
m4	flessibilità organizzazioni info e materiali	57,2*	47,1*	50,8
m5	nuove strategie di studio	62,2*	55,5	53,5
ar1	cause insuccesso	67,2	74,8	69,2
ar2	concentrazione	68,7	69,1	63,7
ar3	autonomia	63,3*	40,3	46,2
em1	autostima	36,2	40,7	40,9
em2	ansia	34,7	35,1	41,2
em3	atteggiamento positivo	31,6	35,6	36,7
em4	motivazione allo studio	50,5	43,1	48,8

Tab. 8 – Dati area “Valutazione”, aspetti formativi between i partecipanti (% risposte positive)

Le differenze più significative si evidenziano nell'analisi within i partecipanti, in particolare STU e INS. Nella fig.2 si notano le differenze fra STU DF e STU NON DF e gli INS che assegnano o meno i DF. Si rilevano differenze nelle scale metacognitiva ed autoriflessiva, con punteggi inferiori per gli STU DF e per gli INS DF. Inoltre, si notano punteggi molto bassi nella scala emotivo-motivazionale per tutti.

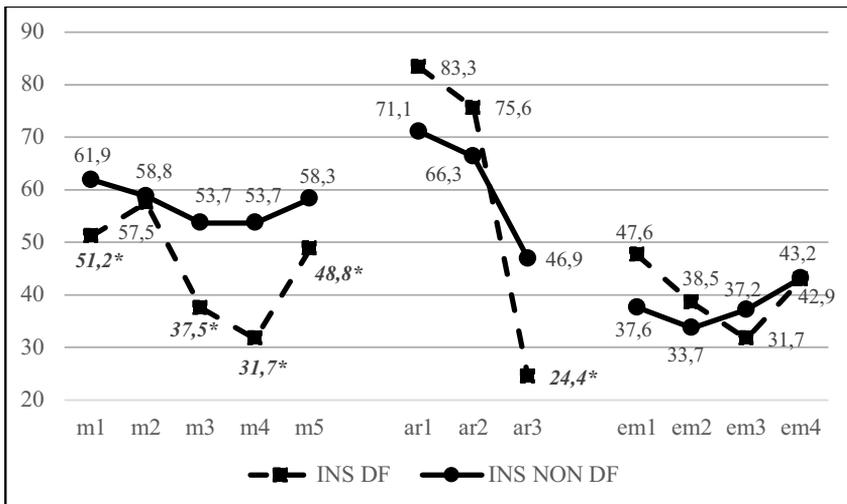
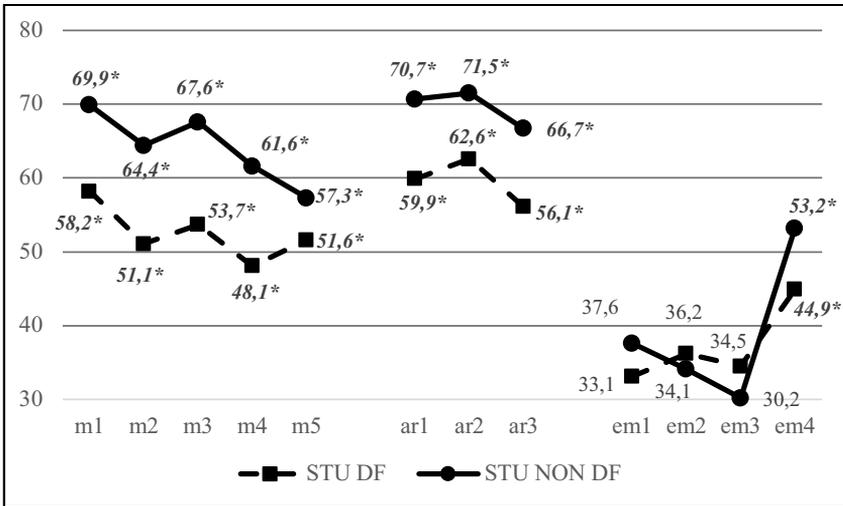


Fig. 2 – Dati area “Valutazione”, aspetti formativi within gli studenti e gli insegnanti (% risposte positive)

Un'ultima analisi interessante è quella relativa alla Regressione lineare. Inserendo, come variabile dipendente, gli studenti con e senza debito formativo, risultano come predittori ($t = 11,179$) gli item m5 (strategie di studio, $p < ,002$), m2 (gestione delle discipline, $p < ,002$) ed em4 (atteggiamento positivo, $p < ,007$).

5.2. I dati qualitativi

Per l'analisi dei dati qualitativi, è stato adottato il metodo di codifica a posteriori. Le risposte alle due domande aperte, incluse nella quinta area "idee", dopo essere state trascritte, sono state analizzate in modalità *triple-blind*. Tre ricercatori indipendenti hanno condotto un'analisi preliminare del corpus delle interviste trascritte. In seguito, le diverse osservazioni sono state confrontate, al fine di giungere ad un'interpretazione condivisa. In questo modo, è stata limitata l'influenza delle precomprensioni dei ricercatori (Trincherò, 2004).

I passaggi dell'analisi qualitativa sono stati i seguenti:

1. Innanzitutto, è stato generato un *code system* dei diversi elementi rinvenuti nel testo, assegnando un'etichetta che categorizza in forma semplificata le diverse frasi (*code*); per stabilire i codici è stato utilizzato un approccio induttivo, partendo dai dati raccolti, e non un approccio deduttivo, basato su una fondazione teoretica preesistente;
2. Dopo la codifica del testo, i codici sono stati raggruppati in *cluster*, secondo relazioni di affinità tematica e di somiglianza;
3. Per ogni codice inserito nei cluster, sono state calcolate le occorrenze;
4. Infine, i dati sono stati inseriti in una tabella in cui vengono elencati i cluster, i codici compresi nel cluster e le occorrenze di ciascun codice.

Dopo una prima analisi, è emersa un'evidente differenza tra le risposte date dagli STU DF e quelle date dagli STU NON DF. Analogamente, è risultata una chiara differenziazione fra gli insegnanti che assegnano o meno debiti formativi. Per questo motivo, è stato deciso di ripetere l'analisi di tutti i dati qualitativi raccolti, distinguendo e analizzando le risposte in base a questo criterio. Infine, è stato misurato il peso del codice all'interno del corpus per ciascun partecipante, in modo da evidenziare i codici ritenuti maggiormente significativi.

CLUSTER	CODE	PESO DEL CODE NEL CORPUS					
		STU DF	STU NON DF	INS DF	INS NON DF	GEN DF	GEN NON DF
ASPETTI POSITIVI							
SONO UTILI	Se danno l'opportunità di migliorare nella disciplina	25,01	22,45	22,49	26,32	23,32	25,42
	Per affrontare l'anno successivo	6,38	5,10	7,50	10,53	13,33	13,55
	Se adeguatamente realizzati	5,32	3,40	10,00	5,26	10,00	11,86
VALENZA FORMATIVA	Dimostrare le proprie capacità	0	2,04	0	0	0	0
	Migliorare le proprie competenze	1,06	1,02	0	0	0,00	0
SUPPORTO ALLA RIFLESSIONE	Riflettere sui propri errori	4,26	7,14	2,50	26,32	3,33	1,69
	Essere consapevoli delle strategie di studio	1,06	4,42	7,50	5,26	0	0
totali		43,09	45,58	49,99	73,68	49,99	52,54
ASPETTI NEGATIVI							
SONO INUTILI	Andrebbero aboliti	18,62	13,61	25,00	15,79	0	0
	Non si migliora nella disciplina	10,64	6,12	12,51	5,26	16,67	15,25
	Male organizzati e mal strutturati	4,26	3,06	12,50	5,26	16,67	16,94
DIFFICOLTÀ NEL RECUPERO	Periodo estivo troppo breve per recuperare	8,51	7,48	0	0	0	0
	Periodo estivo non incentiva lo studio	3,19	9,18	0	0	0	0
	Occorre un supporto maggiore da parte della scuola e dei docenti durante l'anno	2,13	3,40	0	0	0	0

DIFFICOLTÀ EMOTIVO- MOTIVAZIO- NALI	Provocano ansia e stress	7,98	10,20	0	0	10,01	8,47
	Aumentano l'avversione per la materia	1,60	1,36	0	0	6,67	6,77
totali		56,91	54,42	50,01	26,32	50,01	47,46

Tab. 9 – Analisi dei cluster e dei codici

Nella tab. 9 vengono presentati i dati relativi alla distribuzione dei codici e dei cluster relativi agli aspetti positivi e negativi legati ai corsi di recupero e alle prove di verifica. È importante notare che non sono stati riportati tutti i codici ma solo quelli con un livello di occorrenze numericamente rilevante. Si possono sottolineare alcuni aspetti.

L'utilità degli esami è subordinata ad alcuni fattori, in particolare, "se danno l'opportunità di migliorare nella disciplina" (per tutti i partecipanti) e "se adeguatamente realizzati" (per INS e GEN). La valenza formativa è citata solo dagli STU mentre il supporto alla riflessione viene sottolineata dagli INS NON DF.

L'inutilità degli esami è sottolineata da tutti i partecipanti anche se con sfumature differenti. I GEN rimarcano le difficoltà nell'organizzazione e strutturazione. Le difficoltà nel recupero vengono identificate solo dagli STU (in particolare, riferite all'inadeguatezza del periodo estivo) mentre le difficoltà emotivo-motivazionali sono segnalate da STU e GEN, con particolare riferimento all'ansia e stress correlati.

Altri cluster, invece, sono finalizzati ad evidenziare suggerimenti e soluzioni alternative per l'organizzazione e la conduzione dei corsi di recupero e delle prove di verifica finali. Il primo cluster è riferito in particolare ai corsi di recupero. INS e STU segnalano la necessità di rendere intensivi i corsi, possibilmente tenuti dagli insegnanti di classe. Si propongono anche corsi con un numero limitato di alunni e percorsi di tutoraggio, anche a distanza (quest'ultimo indicato dai GEN).

Le tempistiche di svolgimento delle prove rappresentano un altro cluster. Gli STU richiedono che le prove si tengano tra giugno e luglio, principalmente per due motivi: ritengono di essere più ben disposti verso lo studio, visto come una sorta di continuazione dell'anno scolastico, e fanno notare la possibilità di trascorrere serenamente i seguenti mesi estivi. Anche GEN e INS sono a favore di questa proposta.

Infine, è interessante considerare in parallelo gli ultimi due cluster, che riguardano la modalità di strutturazione degli esami e l'oggetto di valutazione. Gli STU chiedono di diminuire il livello di difficoltà e il carico di studio previsti, seguiti dall'esigenza di criteri di valutazione palesi. Interessante è, poi, la proposta di personalizzare gli esami per ciascun studente: con questa, si pone l'attenzione al percorso individuale di ognuno e si sottende la necessità di una valutazione più formativa che sommativa. In linea con questo pensiero, gli STU affermano che l'oggetto di valutazione dovrebbe riguardare il percorso di miglioramento svolto nella disciplina, piuttosto che il recupero di tutto il programma. Diverse occorrenze si riferiscono ad una valutazione che riguardi unicamente gli argomenti non pienamente appresi dall'alunno durante l'anno o, comunque, solo gli aspetti nodali di una disciplina.

6. Discussione e conclusioni

Dall'analisi dei dati emergono alcune osservazioni importanti. La risposta alla prima domanda di ricerca indica che, ad una percentuale variabile fra un quinto e un quarto degli STU, vengono assegnati DF. Le questioni più rilevanti, però, sono incentrate sul genere, la tipologia di scuola e le classi. Ai maschi vengono assegnati molti più debiti, così come i debiti si concentrano negli istituti tecnici e nelle prime classi, in particolare la seconda. Queste prime osservazioni ci suggeriscono di investire le risorse disponibili, soprattutto, in questi ambiti.

La seconda domanda di ricerca conferma un trend nazionale che vede le discipline matematico-scientifiche oggetto dei DF, seguite dalle lingue (in particolare, l'inglese) e dal binomio italiano/storia. A differenza di alcune credenze, è importante sottolineare che non vi sono differenze di genere fra F e M nei debiti relativi alle discipline scientifiche.

La tabella 6 sintetizza i risultati della terza domanda. Si può osservare che, per gli aspetti economici, i GEN percepiscono i DF come molto costosi mentre, per gli INS, la percezione dei costi è più bassa. Gli aspetti temporali indicano come, per gli STU, l'estate non sia un periodo adatto per lo studio e, per tale motivo, i DF limitino fortemente il tempo per i propri interessi. Gli aspetti emozionali indicano che, per tutti i partecipanti, i DF hanno un impatto non positivo sullo studente da un punto di vista emotivo. Infine, gli aspetti organizzativi sottolineano che, per gli INS, le prove degli esami sono adatte mentre, per i GEN, i corsi di recupero non sono ben organizzati.

Le osservazioni più interessanti, però, si riscontrano nelle analisi within i partecipanti. Le femmine affrontano lo studio per i DF con maggiore preoccupazione e si nota come gli STU DF dichiarino che i corsi di recupero sono poco adatti e le prove finali poco utili. Parallelamente, gli INS con maggiore anzianità e quelli che assegnano molti DF notano difficoltà nei corsi di recupero e per le prove finali. Questo aspetto sottolinea una questione rilevante: gli INS DF, che assegnano molti debiti, devono utilizzare strumenti valutativi che non ritengono adeguati.

L'analisi degli aspetti sommativi e formativi fornisce le risposte alla quarta e quinta domanda. Gli STU non credono di poter raggiungere tutti gli obiettivi previsti dal percorso di recupero ma, soprattutto, si notano criticità negli aspetti formativi. Gli STU DF e gli INS DF evidenziano criticità nella scala metacognitiva, in quella auto-riflessiva e, soprattutto per gli INS, nell'autonomia.

Infine, l'analisi della regressione ci fornisce la riflessione più importante in quanto sottolinea come i percorsi di recupero debbano focalizzarsi sull'evoluzione delle strategie di studio, la gestione delle informazioni relative alle discipline e sulla possibilità di incentivare un atteggiamento positivo nei confronti delle discipline oggetto di DF.

I dati qualitativi confermano e sottolineano le difficoltà dei STU DF e degli INS DF che denotano criticità e perplessità nell'utilizzo di percorsi valutativi che non sortiscono gli effetti desiderati. Sono particolarmente interessanti le proposte e le soluzioni alternative relative all'organizzazione e alla conduzione dei corsi di recupero e delle prove di verifica finali.

In conclusione, questa ricerca ci suggerisce sia alcuni aspetti critici sia suggerimenti e indicazioni per il futuro.

Gli aspetti maggiormente problematici si concentrano nei percorsi di recupero, in particolare, nella loro organizzazione e nella modalità di erogazione, ritenute carenti dai partecipanti che hanno esperienza dei DF. Le prospettive indicano una rimodulazione dei percorsi, suggerendo percorsi di tutoraggio, anche online, e focalizzando gli obiettivi del recupero fra quelli basilari e fondanti.

In parallelo, la tempistica delle prove finali indica una preferenza per un anticipo a giugno/luglio o, addirittura, nel corso dell'anno. In questo senso, si privilegierebbe l'aspetto di monitoraggio piuttosto che quello di prova finale.

L'aspetto più complesso da affrontare è l'inadeguata incidenza degli aspetti formativi. La scala emotivo-motivazionale è decisamente la più delicata ma, insieme a quella metacognitiva, rappresentano il maggior limite dei debiti formativi. Gli insegnanti segnalano più volte questa difficoltà tanto che propongono attività di formazione per gli insegnanti di discipline che sono spesso oggetto di DF in modo da essere in grado di supportare lo sviluppo di aspetti quali l'autonomia, la capacità di elaborare le informazioni, sviluppare nuove strategie di studio e un atteggiamento positivo verso la disciplina.

I passi futuri, quindi, sono orientati in tre direzioni: la formazione degli insegnanti, l'organizzazione di percorsi più flessibili e mirati e, infine, la chiara indicazione di utilizzare un insuccesso (il debito formativo) per incentivare il cambiamento nell'approccio e nelle strategie di studio.

Riferimenti bibliografici

- Becchi, E. (1995). Valutare sistemi, programmi e profitto educativi: operatività complesse. *Quaderni di RES*, 9, pp. 5-8.
- Bell, B., Cowie, B. (2001). *Formative assessment and science education*. Dordrecht, The Netherlands: Kluwer.
- Black, P., Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education*, 5(1), pp. 7-74.
- Bondioli, A., Ferrari, M. (2008). *Verso un modello di valutazione formativa. Ragioni, strumenti e percorsi*. Bergamo: Junior.
- Cizek, G.J., Andrade, H.L. (Eds.) (2010). *Handbook of formative assessment*. New York, NY: Routledge.
- Creswell, J. W., Plano Clarke, V. L., Gutmann, M. L., & Hanson, W. E. (2003). Advanced mixed methods research designs. In A. Tashakkori, & C. Teddlie (Eds.), *Handbook of mixed methods in social and behavioral research* (pp. 209-241). Thousand Oaks, CA: Sage.
- Dionisi, G., Tadiello, R. (2007). *Valutazione formativa, pedagogia del contratto e differenziazione didattica*. Milano: FrancoAngeli.
- Giannandrea, L. (2009). *Valutazione come formazione. Percorsi e riflessioni sulla valutazione scolastica*. Macerata: EUM.
- Heritage, M. (2010). *Formative assessment. Making it happen in the classroom*. Thousand Oaks, CA: Corwin.
- McMillan, J.H. (Ed.) (2007). *Formative classroom assessment: Theory into practice*. New York: Teachers College Press.
- Parmigiani, D. (2016). *L'aula scolastica 2. Come imparano gli insegnanti*. Milano: FrancoAngeli.
- Perrenoud, P. (1991). Towards a pragmatic approach to formative evaluation. In P. Weston (Ed.), *Assessment of pupil achievement*. Amsterdam, The Netherlands: Swets & Zeitlinger (pp. 79-101).
- Sadler, R.D. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, 18, pp. 119-144.
- Scriven, M. (1967). *The methodology of evaluation*. Washington, DC: American Educational Research Association.
- Shute, V.J. (2008). Focus on Formative Feedback. *Review of Educational Research*, 78(1), pp. 153-189.
- Tashakkori, A., & Teddlie, C. (1998). *Mixed methodology. Combining qualitative and quantitative approaches*. Thousand Oaks (CA): Sage.
- Tashakkori, A., & Teddlie, C. (2003). *Handbook of mixed methods in social and behavioural research*. Thousand Oaks (CA): Sage.

- Tashakkori, A., & Teddlie, C. (2009). *Foundations of mixed methods research*. Thousand Oaks (CA): Sage.
- Torrance, H., Pryor, J. (1998). *Investigating formative assessment. Teaching, learning and assessment in the classroom*. Maidenhead, Philadelphia: Open University Press.
- Trincherò, R. (2004). *I metodi della ricerca educativa*. Roma-Bari: Laterza.
- Vertecchi, B. (2003). *Manuale della valutazione. Analisi degli apprendimenti e dei contesti*. Milano: FrancoAngeli.
- Weurlander, M., Söderberg, M., Scheja, M., Hult, H., Wernerson, A. (2012). Exploring formative assessment as a tool for learning: students' experiences of different methods of formative assessment. *Assessment & Evaluation in Higher Education*, 37(6), pp. 747-760.
- William, D. (2011). *Embedded Formative Assessment*. Bloomington, IN: Solution Tree Press.
- William, D., Black, P. (1996). Meaning and consequences: A basis for distinguishing formative and summative functions of assessment? *British Educational Research Journal*, 22(5), pp. 537-548.
- Winingar, S.R. (2005). Using Your Tests to Teach: Formative Summative Assessment. *Teaching of Psychology*, 32(3), pp. 164-166.