

## I.4

Mettere in pratica il *Formative Assessment*: uno studio di caso sui fattori connessi all'implementazione del *Formative Assessment* negli insegnanti di scuola secondaria di secondo grado.

### Implementing Formative Assessment: A Case Study on Factors Related to the Adoption of Formative Assessment by Upper Secondary School Teachers.

Alessandro Oro

Università di Bologna, Dipartimento di Scienze dell'Educazione "Giovanni Maria Bertin"  
alessandro.oro2@unibo.it

Nonostante le numerose evidenze circa l'efficacia delle strategie di *formative assessment* (FA), la messa in pratica di quest'ultime risulta difficoltosa per la maggioranza degli insegnanti. In letteratura è possibile distinguere due macro-categorie di fattori che influenzano (sia nelle intenzioni che nell'effettiva messa in atto di strategie di FA) gli insegnanti: fattori personali e fattori contestuali. Tramite uno studio di caso collettivo condotto con metodi misti (architettura parallela e strategia di ricerca complementare), la presente ricerca ha come obiettivo di studiare se e quanto (e con quale relazione fra essi) i valori personali degli insegnanti di scuola secondaria di secondo grado, il loro senso di auto-efficacia percepita, la loro *assessment literacy* influiscano sulle loro convinzioni rispetto alle funzioni della valutazione e degli errori in classe e alla loro disponibilità ad agire tramite strategie di FA. Inoltre, si intende studiare come tali relazioni si differenzino a seconda della cultura valutativa della scuola di appartenenza degli insegnanti. Sono state selezionate 4 scuole-caso (secondaria di secondo grado) nel territorio della regione Emilia-Romagna, per effettuare uno studio di caso collettivo condotto con *mixed methods* strutturato in 3 azioni. Una prima fase per rilevare la cultura valutativa dell'Istituto tramite interviste semi-strutturate ai Dirigenti Scolastici e analisi della documentazione strategica; una seconda azione di rilevazione dei costrutti relativi agli insegnanti tramite un questionario online; una terza azione di approfondimento tramite focus-group e restituzione dei dati. I primi risultati e lo stato dell'arte sono discussi nel contributo.

**Parole chiave:** convinzioni degli insegnanti; concezioni sulla valutazione; cultura valutativa; formative assessment; valori personali.

Despite extensive evidence supporting the effectiveness of formative assessment (FA) strategies, their implementation remains challenging for the majority of teachers. The literature identifies two broad categories of factors that influence teachers' intentions and actual use of FA strategies: personal factors and contextual factors. Through a collective case study employing mixed methods (parallel architecture with a complementary research strategy), this study aims to examine whether, to what extent, and in what interrelation teachers' personal values, perceived self-efficacy, and assessment literacy influence their beliefs about the functions of assessment, the role of errors in the classroom, and their willingness to adopt FA strategies. Additionally, the study investigates how these relationships vary according to the assessment culture of the teachers' respective schools. Four secondary schools in the Emilia-Romagna region were selected as case studies, with three main actions undertaken to achieve the research objectives. The first phase involved assessing the institutional assessment culture through semi-structured interviews with school principals and an analysis of strategic documents. The second phase focused on gathering data on teachers' constructs via an online questionnaire. The third phase provided in-depth exploration through focus groups and feedback on the collected data. Preliminary findings and the current state of the research are discussed in this contribution.

**Keywords:** teacher beliefs; conceptions of assessment; assessment culture; formative assessment; personal values.

## 1. Quadro teorico di riferimento

Il dibattito sulla funzione formativa della valutazione, ovvero l'idea che la valutazione possa essere utilizzata per migliorare la qualità del processo didattico oltre che per la misura dei suoi effetti, si è diffuso nella ricerca e nella pratica educativa a partire dagli anni '60 (Cizek; 2010; Wiliam; 2019). La prima distinzione funzionale tra i termini "sommativa" e "formativa" riferiti alla valutazione si ritrova in Michael Scriven (1967), per il quale la valutazione sommativa forniva informazioni per giudicare l'efficacia complessiva di un programma didattico, mentre i risultati della valutazione formativa erano utili per facilitare l'implementazione di tale programma. Negli anni immediatamente successivi, nel 1969, Benjamin Bloom descrisse la strategia didattica del *mastery learning* (apprendimento per padronanza), in un periodo durante il quale era credenza comune che soltanto gli studenti più abili potessero raggiungere alti livelli di standard educativo. Rifiutando questa convinzione, Bloom ritenne che la maggior parte degli allievi può raggiungere un alto livello di apprendimento, purché vengano create le condizioni adatte ai bisogni di ciascuno di loro. In questo approccio, il tempo necessario allo studente sarebbe stato rilevato attraverso valutazioni periodiche del suo progresso. In un testo del 1969, sempre Bloom scrive che "[...] in contrasto alla "valutazione sommativa" [...] si utilizza il termine "valutazione formativa" per fornire feedback e correttivi ad ogni passaggio del processo di insegnamento-apprendimento. Con valutazione formativa intendiamo una valutazione tramite brevi test utilizzati da insegnanti e studenti come aiuti nel processo di apprendimento" (Bloom, 1969, pp. 47-48). Dal lavoro pionieristico di Bloom, l'idea che la valutazione possa migliorare l'apprendimento e al tempo stesso misurare i suoi risultati è diventata centrale nelle prassi didattiche e nella ricerca educativa nel mondo.

Questo concetto generale di funzione formativa della valutazione ha negli ultimi anni guidato differenti studi sull'efficacia delle pratiche di formative assessment (FA) nell'apprendimento. Il primo studio, divenuto poi il più citato nel campo di ricerca sul FA, è un articolo di Black e Wiliam (1998), "*Inside the Black Box*", nel quale viene riportato un *effect size* dell'efficacia delle strategie di FA tra lo 0,4 e lo 0,7, estratto da una meta-analisi di 250 studi riguardanti la valutazione in classe. A partire da questo studio, si è aperto un intero filone di ricerca a proposito dell'efficacia delle strategie di FA. Intorno agli anni '10 del nuovo millennio si apre una fase di approfondita riflessione sulle effettive evidenze raccolte da Black e Wiliam, ed iniziano ad emergere alcuni articoli critici (Bennett, 2011; Kingston and Nash; 2011; Briggs et al., 2012; McMillan et al., 2013), volti a rianalizzare queste evidenze, proponendo nuove modalità di raccolta di queste. Nella loro meta-analisi, Kingston e Nash (2011), hanno rilevato un *effect size* medio di 0,25, suggerendo un'efficacia complessivamente minore rispetto a quella riportata da Black e Wiliam (1998). Briggs e colleghi (2012), rispondendo allo studio di Kingston e Nash, ne evidenziano alcune lacune metodologiche che li hanno guidati, pur sottolineando la necessità di un maggiore rigore metodologico e l'esigenza di più studi di tipo sperimentale. Allo stesso tempo, Bennett (2011), nel suo articolo "*Formative Assessment: A Critical Review*", ritiene che l'ampiamente citato *effect size* di 0,4-0,7 sia problematico tout-court in quanto basato su evidenze metodologicamente deboli o impossibili da rintracciare e poter visionare. È importante sottolineare che questi articoli "di critica" non nascono con l'intento di negare la possibilità che le strategie di FA possano migliorare il processo di insegnamento-apprendimento, sono invece proposti per indirizzare al meglio le scelte di ricerca in questo campo ed arrivare a delle evidenze quanto più accurate possibili. Più recentemente, altri studi e meta-analisi hanno evidenziato l'efficacia delle strategie di FA nelle capacità di lettura (Lee et al., 2020; Xuan et al., 2022), nell'apprendimento della matematica (Andersson e Palm, 2017; Lee et al; 2020). Altri studi sull'efficacia del feedback, elemento chiave del FA, hanno evidenziato associazioni significativamente positive tra feedback e apprendimento (Hattie, 2009; Van der Kleij et al., 2015; Wisniewski et al., 2020).

Nonostante le numerose evidenze dell'efficacia delle strategie di FA sul processo di insegnamento-apprendimento, la realizzazione di quest'ultime risulta difficile e complesso per la maggioranza degli insegnanti (Baird et al., 2016; Furtak et al., 2016; Yan & Brown, 2021). Questo aspetto è particolarmente rilevante e significativo nel contesto italiano, poiché, nonostante la ricerca pedagogica sperimen-

tale si sia occupata di definire il costrutto di valutazione formativa ed operazionalizzarlo, le pratiche di FA raramente sono entrate a far parte delle programmazioni didattiche e dei percorsi di formazione per insegnanti delle scuole secondarie (Benvenuto, 2023; Corsini, 2023; Vannini, 2023). Appare necessario soffermarsi su quali sono i fattori coinvolti nella messa in pratica di strategie di FA. Nella letteratura internazionale, sono presenti diverse rassegne sistematiche di studi che si sono occupate di analizzare i fattori coinvolti nell'implementazione del FA da parte degli insegnanti. Fulmer et al. (2015) hanno proposto un modello concettuale dei fattori che influenzano le pratiche valutative degli insegnanti suddivisa in fattori contestuali a livello micro- (valori, concezioni e conoscenze sulla valutazione) e livello meso- (legati all'istituto stesso) e macro- (influenze culturali di ampio grado). In Heitink et al. (2016) sono state individuate quattro macro-fattori (in questo caso intesi come prerequisiti) necessari per la messa in pratica di strategie di FA: il contesto (cultura della valutazione, supporto professionale), l'insegnante (conoscenze, abilità, atteggiamenti, convinzioni), lo studente. Schildkamp e colleghi. (2020), fornendo un modello di prerequisiti fondamentali degli insegnanti per la messa in atto di strategie di FA nella forma di Data-Based Decision Making e Assessment for learning, hanno individuato tre macro-aree: la conoscenza e le abilità (a proposito della valutazione, delle conoscenze pedagogiche specifiche, nel fornire i feedback, nel definire gli obiettivi) i fattori psicologici (atteggiamenti, pressione sociale, percezione di controllo) e i fattori sociali (collaborazione con i colleghi e gli studenti, cultura della scuola di appartenenza). Infine, Yan et al. (2021), analizzando 52 ricerche, hanno distinto due macro-categorie di fattori che influenzano (sia nelle intenzioni che nell'effettiva messa in atto di strategie di FA) gli insegnanti ovvero i fattori personali e i fattori contestuali. Tra i primi, hanno una particolare significatività le intenzioni e le convinzioni degli insegnanti sulla valutazione formativa, il loro senso di auto-efficacia (nell'effettuare e gestire la valutazione) e la presenza/assenza di formazione specifica sul FA. Tra i secondi, quelli con maggiore impatto sono il supporto scolastico interno (a livello di istituto e colleghi), le condizioni di lavoro, le politiche esterne (nazionali e sovranazionali) e le norme culturali della scuola. È possibile notare quindi un punto comune tra le varie rassegne sui fattori antecedenti alla messa in atto di strategie di FA ovvero quello di identificare almeno 2 macro-variabili: i fattori disposizionali, propri di ogni singolo insegnante, e i fattori contestuali che possono essere intesi a più livelli o sistemi, ad esempio riferendosi al meso-sistema (l'istituto) o al macro-sistema (il contesto scolastico nazionale).

Ai fini della presente ricerca, sono stati presi in considerazione i fattori che più significativamente influiscono sulle convinzioni degli insegnanti e sulla loro disponibilità ad agire tramite strategie di FA. Esistono varie terminologie per indicare gli aspetti attraverso i quali i docenti interpretano l'insegnamento, l'apprendimento, la valutazione, il ruolo dell'insegnante e della scuola: nel presente studio è stato utilizzato il termine "convinzioni" (*teacher beliefs*) (Fives & Buhel, 2012). Studiare le convinzioni, le pratiche e gli atteggiamenti degli insegnanti (nonché le relazioni che intercorrono tra loro), è particolarmente importante per orientare l'agire educativo verso la qualità e l'equità dell'apprendimento, tendendo verso una didattica dimostrata come efficace dalla ricerca educativa empirica. In particolare, le convinzioni in merito alla valutazione sono strettamente collegate con le pratiche valutative (Brown, 2004) e possono essere considerate degli aspetti "latenti" che, in termini di pensiero e atteggiamento, si pongono "dietro l'azione" (Vannini, 2012). Interrogarsi sulle modalità più efficaci per promuovere processi di innovazione e miglioramento delle pratiche di insegnamento implica dunque la necessità di considerare e agire sulle concezioni degli insegnanti (Ciani e Rosa, 2020).

Un altro fattore preso in considerazione nel presente studio, questa volta di tipo contestuale, è la "cultura" condivisa nel meso-sistema e nel macro-sistema rispetto sulla valutazione. Con "cultura della valutazione", concetto principalmente sviluppato in ambito francofono (Michel, 2004; Rey, 2008), si intende l'insieme degli scopi, dei fini, delle funzioni che vengono assegnati alla valutazione all'interno di un contesto scolastico. Capperucci, Corsini e Ugolini (2018) problematizzano la questione facendo notare che in Italia cultura del testing e cultura della valutazione tendono a coincidere. Ad esempio, l'INVALSI nel 2008 identifica "lo sviluppo di una cultura della valutazione (*teach to test*)" con "l'introduzione nelle scuole di prove di verifica strutturate (quesiti a scelta multipla) e semi-strutturate (quesiti

con stimolo chiuso e risposta aperta)” (INVALSI, 2008, p. 99). Questo tipo di cultura della valutazione inserita nelle politiche nazionali riguardanti la scuola, e quindi di meso-sistema, si allontana da una visione inclusiva e centrata sullo studente come invece accade nelle pratiche di FA (Corsini, 2018). Attualmente, sono ancora scarse le evidenze in merito alla cultura valutativa delle singole scuole, intesa come insieme di convinzioni e pratiche delle modalità valutative, delle loro funzioni e dei loro scopi. Inoltre, si rileva come la cultura della valutazione diffusa tra gli insegnanti spesso è centrata su procedure valutative che derivano esclusivamente dalle competenze disciplinari degli insegnanti, dalle loro esperienze passate, mancando l’uso di procedure accurate di misurazione: è una cultura che riverbera quella precedente alla scolarizzazione di massa (Vertecchi, Agrusti e Losito, 2010). Diversi fattori di tipo personale sono stati considerati nel presente studio. Il primo è il senso di auto-efficacia percepito dall’insegnante, essendo il predittore più stabile della qualità dell’insegnamento in quanto favorisce la formazione di un clima supportivo, la gestione efficace della classe e l’attivazione cognitiva degli studenti (Kunsting et al, 2015). Inoltre, l’auto-efficacia degli insegnanti è anche il principale fattore personale che predice la messa in atto di strategie di valutazione formativa (Yan et al., 2021, 2022). L’auto-efficacia dell’insegnante è solitamente definita come la percezione della capacità di gestire e mettere in pratica un processo di insegnamento-apprendimento efficace, e questo vale anche in riferimento all’implementazione di strategie di FA (Karaman e Shain, 2017). Insegnanti con un alto livello percepito di auto-efficacia nelle proprie capacità di mettere in atto strategie di FA saranno più disposti a farlo effettivamente (De Luca et al., 2019). Un altro fattore personale rilevante per la presente ricerca è costituito dall’*assessment literacy*, traducibile come “alfabetizzazione sulla valutazione”, ovvero quell’insieme di conoscenze pratiche e teoriche, nonché di consapevolezza, riguardo ai vari aspetti della valutazione e sulle sue funzioni. Il termine *assessment literacy*, coniato da Stiggins (1991), definisce la comprensione dei concetti fondamentali della valutazione e le sue procedure, di modo da poterle consapevolmente durante le decisioni educative. In letteratura, esistono riferimenti a specifici comportamenti che evidenziano il livello di conoscenza e competenza degli insegnanti rispetto alla valutazione, come ad esempio mettere in pratica valutazioni di qualità e coerenti con gli obiettivi di apprendimento, interpretare le performance degli studenti alla luce del tipo di prova valutativa messa in atto, somministrare accuratamente le valutazioni, comunicare i risultati in modo accurato ed evitare giudizi di valore sullo studente (Brookhart, 2011; Popham, 2008; Stiggins, 1991). Nel presente studio, l’insieme di conoscenze e competenze che formano l’alfabetizzazione alla valutazione verrà misurato considerando le caratteristiche del contesto italiano (formazione degli insegnanti, pratiche rilevate da rilevazioni internazionali, Indicazioni Nazionali per il curriculum). Ulteriore fattore personale coinvolto, spesso poco considerato nelle ricerche, riguarda l’influenza dei valori personali nelle pratiche valutative (Barni et al., 2019, 2018). Nella teorizzazione di Schwartz (1992, 1994, 2001, 2012), i valori personali sono credenze o convinzioni riguardanti obiettivi desiderabili o modalità di comportamento, che trascendono le situazioni specifiche e guidano le persone nella selezione e nella valutazione degli eventi. Il sistema proposto da Schwartz (2001) è composto da dieci valori (Universalismo, Benevolenza, Conformismo, Tradizione, Sicurezza, Potere, Successo, Edonismo, Stimolazione, Autodirezione) che formano un sistema organizzato caratterizzato da similarità e incompatibilità in grado di influenzare le scelte del soggetto; questi valori sono organizzati intorno a due diverse dimensioni bipolari a cui ogni soggetto tende preferenzialmente ad appartenere: “Autotrascendenza” (Universalismo e Benevolenza) vs “Autoaffermazione” (Edonismo, Successo, Potere); “Apertura al Cambiamento” (Autodirezione, Stimolazione, Edonismo) vs “Conservatorismo” (Conformismo, Tradizione, Sicurezza). Per ciascuna dimensione bipolare, l’appartenenza a una categoria porta a considerare in maniera negativa i valori opposti.

## 2. Obiettivi e interrogativi/ipotesi della ricerca

Basandosi sui costrutti presentati del paragrafo precedente, il presente progetto di ricerca ha come obiettivo generale esplorare i fattori che influiscono sulle convinzioni degli insegnanti di scuola secondaria

di secondo grado in merito alla valutazione e all'errore, e quindi sulla disponibilità ad agire secondo strategie di *formative assessment* al fine di creare un clima di classe positivo. Nello specifico, possono essere identificati due principali obiettivi:

- 1) Studiare se e quanto (e con quale relazione fra essi) i valori personali degli insegnanti di scuola secondaria di secondo grado, il loro senso di auto-efficacia percepita, la loro *assessment literacy* influiscano sulle loro convinzioni rispetto alle funzioni della valutazione e degli errori in classe. Inoltre si intende studiare quale tipo di relazione intercorre tra questo tipo di influenza e le pratiche valutative dichiarate dagli insegnanti (Main study).
- 2) Studiare come tali relazioni si differenzino a seconda della cultura valutativa (in termini di apertura/chiusura nei confronti di strategie di FA) della scuola di appartenenza degli insegnanti.

In particolare, il presente studio intende rispondere alle seguenti ipotesi correlazionali:

- H1: esiste una relazione tra i valori personali degli insegnanti e le loro concezioni rispetto alla valutazione degli apprendimenti e rispetto all'errore in classe;
- H2: esiste una relazione tra l'*assessment literacy* degli insegnanti e le loro concezioni rispetto alla valutazione degli apprendimenti e rispetto all'errore in classe.
- H3: esiste una relazione tra le concezioni dell'insegnante rispetto alla valutazione degli apprendimenti e rispetto all'errore in classe e le loro dichiarazioni di pratiche di strategie di Formative Assessment.
- H4: la relazione tra concezioni e pratiche dichiarate è mediata dal livello di Auto-Efficacia percepita dall'Insegnante.
- H5: la cultura valutativa dell'istituto influenza le concezioni dell'insegnante.
- H6: la cultura valutativa dell'istituto influenza le pratiche dichiarate dell'insegnante.

Attraverso queste ipotesi, si vuole testare il modello proposto nella figura 1, tramite analisi della correlazione, di mediazione e moderazione e di mediazione moderata.

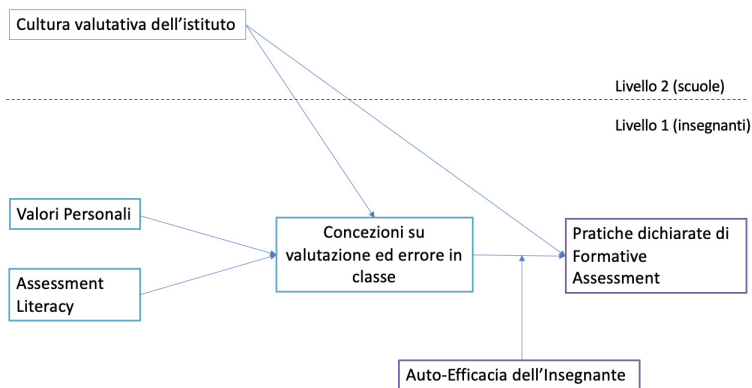


Figura 1 – Modello teorico da testare

### 3. Scelte metodologiche e procedurali

Il presente progetto di ricerca ha seguito un approccio metodologico "*mixed methods*", ovvero quel tipo di ricerca che combina insieme elementi degli approcci quantitativi e qualitativi, con l'obiettivo di raggiungere una maggiore "profondità" nella comprensione del fenomeno studiato (Creswell e Plano Clark,

2017). Nello specifico, questo approccio è stato applicato ad un disegno di ricerca di studio di caso, configurazione indicata da Creswell e Plano Clark (2017, p. 182) come “*mixed methods case study design*”, nel quale la raccolta di dati quantitativi e qualitativi sono integrati per fornire evidenze *in depth* rispetto ai diversi casi presi in esame. Come descritto da Stake (1995) e Yin (2009), un “caso” può essere di varia natura, dal singolo individuo fino ad una specifica organizzazione e dunque gli studi di caso seguono queste distinzioni. Nella presente ricerca, è stato scelto lo studio di caso collettivo (Stake, 1995). Questo approccio prevede lo studio di più casi contemporaneamente o in sequenza al fine di aumentare la rappresentatività e permettere una comparazione tra diverse rilevazioni. Sono stati selezionati come casi quindi gli istituti secondari di secondo grado in quanto organizzazioni e unità di analisi, mentre come fonti di raccolta dei dati saranno considerati gli insegnanti e i Dirigenti Scolastici.

La scelta di optare per questa metodologia di ricerca è collegata alle ipotesi operative precedentemente delineate. Il focus del presente studio riguarda i “contesti” scuola nella forma delle loro culture condivise (in termini di pratiche, visioni e obiettivi dell’educazione) sulla valutazione e sugli errori e come queste influenzano le convinzioni e le pratiche degli insegnanti. Da qui la scelta di considerare un intero istituto come *caso*, inteso come insieme di insegnanti *in-service*, studenti e Dirigente Scolastico. Nello specifico, le scuole-caso selezionate sono state scelte tramite un’interlocuzione con l’Ufficio Scolastico Territoriale della Città Metropolitana di Bologna, in particolare l’ufficio Didattica e Cultura. Grazie alla disponibilità delle responsabili dell’Ufficio menzionato, e dopo una prima fase di ricognizione con un numero maggiore di scuole, sono stati scelti quattro Istituti di Istruzione Superiore nell’area Bolognese che si sono detti disponibili a partecipare alla ricerca. Un elemento importante che ha guidato la scelta è stata la posizione dimostrata dalla scuola in merito alla valutazione degli apprendimenti: in due scuole sono stati attivati o erano attivi dei progetti specifici sulla valutazione degli apprendimenti, mentre le altre due scuole dimostravano un approccio più “tradizionale alla valutazione”. Il disegno di ricerca ha previsto tre distinte fasi, attuate parallelamente e in modo analogo in ogni scuola caso descritte nei seguenti sotto-paragrafi.

### 3.1 Fase 1: conoscenza e rilevazione della cultura valutativa della scuola-caso

In questa prima fase, di stampo principalmente qualitativo, l’obiettivo è stato quello di far emergere gli elementi caratterizzanti della cultura valutativa di ogni singolo caso, ovvero l’insieme di dichiarazioni e azioni che racconta come la scuola si pone nei confronti del tema della valutazione degli apprendimenti. Per fare questo, si è optato per un processo che ha previsto l’intersezione tra un’analisi documentale dei principali documenti strategici di ogni scuola, e le dichiarazioni dei Dirigenti Scolastici e delle Funzioni Strumentali (nello specifico Area 4 – Sistema autovalutazione d’istituto e valutazione apprendimento e comportamento). I documenti strategici selezionati sono stati il Piano Triennale dell’Offerta Formativa e il Rapporto di Auto Valutazione. Il PTOF rappresenta il “documento fondamentale costitutivo dell’identità culturale e progettuale delle istituzioni scolastiche ed esplicita al progettazione curricolare, extracurricolare, educativa e organizzativa che le singole scuole adottano nell’ambito della loro autonomia” (DPR 275/1999, art. 3 comma 1), mentre il RAV (DPR 80/2013) è un documento di auto-analisi delle scuole che permette loro di identificare punti di forza e debolezza, tenendo conto del contesto nel quale sono inserite. I documenti sono stati analizzati tramite una griglia semi-strutturata con 3 indicatori fondamentali (presenza e grado di approfondimento cui viene presentata la valutazione degli apprendimenti, la sua funzione formativa, se e come questa sia collegata alla programmazione didattica), a loro volta utilizzato una gradazione per vedere quanto le sezioni dei documenti prese in esame si avvicinasero o meno ad elementi di valutazione formativa così come descritti nella letteratura. Parallelamente, ai Dirigenti Scolastici e alle Funzioni Strumentali sono state somministrate delle interviste semi-strutturate della durata di 45 minuti volte ad analizzare due macro-temi: i valori che compongono la cultura della specifica scuola in senso generale (per esaminare il posizionamento della scuola su un

continuum di tradizionalismo-apertura al cambiamento) e la cultura valutativa di istituto (prassi e indirizzi condivisi in senso top-down, dalla Direzione agli insegnanti).

### 3.2 Fase 2: rilevazione dei dati relativi agli insegnanti

Al fine di rilevare i costrutti precedentemente presentati specificamente riferiti agli insegnanti, e per fornire una risposta alle ipotesi correlazionali presentate, è stato scelto come strumento di raccolta dei dati un questionario, strutturato in versione online su piattaforma OneDrive con licenza dell'Università di Bologna. Per poter partecipare, gli insegnanti erano obbligati a dare il proprio consenso al trattamento dei dati, in ottemperanza con le normative sulla privacy GDPR.

Nella tabella 1 è possibile visionare la struttura e le dimensioni indagate dal questionario.

Sezione	Variabili	Tipologia di quesito	Riferimenti
Informazioni socioanagrafiche	Età, Anzianità di servizio, Genere, scuola di appartenenza	Scelta multipla categoriale	Creato ad hoc
Autoefficacia dell'Insegnante	Livello di autoefficacia percepita dall'insegnante	Scala Likert a 7 passi	Scala SAED, validazione italiana di Biasci et al. (2014) della <i>Teachers Self-Efficacy Scale</i> (Taschannen-Moran e Wolfhook Hoy, 1999)
Convinzioni sulla valutazione	Convinzioni sul ruolo della valutazione	Scala Likert a 4 passi	Scala CONVAL (Ciani e Rosa, 2021)
Convinzioni sulla valutazione formativa	Convinzioni sul ruolo formativo della valutazione	Scala Likert a 4 passi	Scala VALFOR (Ciani e Vannini, 2017)
Convinzioni sull'errore in classe	Convinzioni sul ruolo dell'errore nell'apprendimento	Scala Likert a 5 passi	<i>Error Orientation Questionnaire</i> , versione adeguata per i contesti educativi usata da Corazza, Matteucci e Santagata (2015)
Situazione reale di valutazione formativa	Percezione di utilità di strategie di valutazione formativa, frequenza di applicazione di strategie in classe e a scuola	Scala Likert a 5 passi	Creato ad hoc, viene presentata una situazione reale di formative assessment in un'interazione insegnante-alunno seguita da asserzioni a cui rispondere su scala Likert. Gli item si riferiscono alle variabili indicate.
Frequenza del formative assessment	Frequenza con la quale vengono messe in atto strategie di formative assessment	Scala Likert a 4 passi	Tradotta ( <i>back-to-back translation</i> ) da scala usata in Young e Jackman (2014)
Conoscenza del formative assessment	Literacy sul formative assessment	Scala Likert a 5 passi	Creato ad hoc basandosi sulla scala <i>Assessment Practices Inventory</i> utilizzata in Zhang e Burry-Stock (2003)
Valori Personali	Valori Personali degli insegnanti (continuum "apertura al cambiamento - tradizionalismo")	Scala Likert a 6 passi	Portrait Value Questionnaire (Schwartz, 2001), versione validata in italiano da Capanna et al. (2005)

Tabella 1 - Struttura del questionario rivolto agli insegnanti

### 3.3 Fase 3: Approfondimento dei risultati e restituzioni alle scuole

Come fase conclusiva, è stata presa in considerazione l'idea di effettuare dei focus-group, uno per ogni scuola-caso, con insegnanti disponibili a partecipare. I focus-group seguono un'articolazione della durata di 60 minuti con massimo 10 partecipanti. Questa fase, che al momento della scrittura del presente contributo deve ancora essere effettuata, ha come obiettivo quello di andare ad indagare altre dimensioni non rilevate nel questionario, come ad esempio il percorso euristico compiuto dalle e dagli insegnanti nel momento dell'operare una valutazione degli apprendimenti od anche l'influenza percepita del contesto rispetto alla loro possibilità di mettere in pratica strategie di *formative assessment*.

Altro componente della fase conclusiva della ricerca è rappresentato da momenti di restituzione ai Collegi dei Docenti delle scuole-caso selezionate nei quali vengono presentati i risultati del questionario e delle analisi qualitative (selezionati per scuola). A partire dai risultati, vengono successivamente proposte delle attività di dibattito e brainstorming per raccogliere ulteriori impressioni sul significato che hanno i risultati per gli insegnanti, ad esempio se trovano coerenza con quanto riportato oppure se hanno ulteriori visioni da aggiungere.

## 4. Stato dell'arte della ricerca

Al momento della scrittura del presente contributo, sono state effettuate le interviste semi-strutturate ai Dirigenti Scolastici degli Istituti coinvolti e sono in corso le analisi documentali dei documenti strategici. Al tempo stesso, è stato effettuato un try-out del questionario rivolto agli insegnanti presso un Istituto di Istruzione Superiore non facente parte dei casi presi in considerazione dalla ricerca, alla quale hanno partecipato 70 insegnanti. La somministrazione di un questionario pilota è una fase necessaria per controllare la validità (interna e di facciata) dello strumento che si intende utilizzare (Benvenuto, 2015). Grazie ai dati raccolti, sono state effettuate due tipi di analisi statistica. In primo luogo, è stata effettuata un'Analisi delle Componenti Principali, per verificare se gli item delle scale si raggruppavano in modo coerente rispetto alle dimensioni teoriche previste e per identificare e rimuovere quegli item che non contribuivano in modo significativo alle componenti. Successivamente, è stata effettuata un'Analisi di Affidabilità tramite il calcolo del coefficiente alfa di Cronbach per ogni scala utilizzata. Nella Tabella 2 è possibile vedere i risultati del calcolo del coefficiente di affidabilità per ogni scala, risultato da adeguato a ottimo.

Dimensione e scala	Coefficiente
Auto-efficacia dell'Insegnante (SAED)	0,89
Concezioni sulla valutazione (CONVAL)	0,79
Concezioni sulla valutazione formativa (VALFOR)	0,65
Convinzioni sull'errore (EOQ)	0,78
Frequenza di strategie di <i>Formative Assessment</i>	0,80
<i>Formative Assessment Literacy</i>	0,81
Valori Personali (PVQ)	0,88

Tabella 2 – Coefficienti delle scale utilizzate nel questionario

Allo stato attuale, è anche in fase di ultimazione la rilevazione ufficiale dei dati presso gli insegnanti degli istituti coinvolti nella ricerca.

## 5. Punti di forza e di criticità della ricerca

Il presente studio ha come principale punto di forza quello di esplorare le dinamiche antecedenti alle pratiche di valutazione formativa degli insegnanti di scuola secondaria di secondo grado, una tipologia di studi poco presente nel contesto educativo italiano. Infatti, nonostante la riflessione sulle pratiche valutative degli insegnanti (Gattullo, 1978; Vannini, 2012) e quella docimologica in generale (Visalberghi, 1955; Vertecchi, 1976; Benvenuto, 2003) siano molto approfondite e sviluppate, rimane comunque difficoltoso il passaggio dalla riflessione teorica all'applicazione pratica nel contesto scolastico. La presente ricerca propone un disegno che permette l'ingresso nel contesto e il suo studio da più punti di vista (i documenti prodotti dalla scuola, la visione del Dirigente, la visione degli insegnanti) per capire al meglio in che misura e quali sono i fattori che influenzano maggiormente il lavoro dei docenti nell'ambito valutativo. Questo aspetto è rilevante anche dal punto di vista dell'impatto che la specifica ricerca può avere nel contesto indagato, coinvolgendo nella riflessione i partecipanti stessi alla ricerca per esplorare la possibilità che un lavoro di questo tipo possa effettivamente influenzare il cambiamento delle pratiche negli insegnanti.

Dall'altro lato, la presente ricerca ha come punti di debolezza la volontà di considerare un grande numero di costrutti nello stesso momento, rischiando di rimanere in superficie e non approfondire mai veramente le variabili considerate in questa analisi esplorativa. Inoltre, come ricordato dagli stessi metodologi Creswell e Plano Clark (2017) la scelta dello specifico disegno di ricerca proposto dovrebbe essere svolto in presenza di un'expertise elevata nel ricercatore per poter affrontare tutti i passaggi con coerenza e validità e questo fattore è distante dalla reale esperienza del ricercatore che ha proposto il presente studio. Inoltre, trattandosi di uno studio di caso collettivo, i risultati emersi potranno essere applicabili soltanto per le scuole che hanno preso parte alla ricerca, mancando così la possibilità di un'estensione dei risultati anche ad altri contesti educativi simili.

## Riferimenti bibliografici

- Andersson, C., & Palm, T. (2017). The impact of formative assessment on student achievement: A study of the effects of changes to classroom practice after a comprehensive professional development programme. *Learning and Instruction*, 49, 92–102. <https://doi.org/10.1016/j.learninstruc.2016.12.006>
- Barni, D., Danioni, F., & Benevene, P. (2019). Teachers' Self-Efficacy: The Role of Personal Values and Motivations for Teaching. *Frontiers in Psychology*, 10, 1645. <https://doi.org/10.3389/fpsyg.2019.01645>
- Barni, D., Russo, C., & Danioni, F. (2018). Teachers' Values as Predictors of Classroom Management Styles: A Relative Weight Analysis. *Frontiers in Psychology*, 9, 1970. <https://doi.org/10.3389/fpsyg.2018.01970>
- Benvenuto, G. (2015). *Stili e metodi della ricerca educativa*. Carocci.
- Benvenuto, G. (2003). *Mettere i voti a scuola. Introduzione alla docimologia*. Carocci
- Benvenuto, G. (2023). Lezioni dal passato. Cosa (non) ha funzionato nella formazione degli insegnanti? Costruire il futuro: Le emergenze dalla ricerca per la formazione. Lessons from the past. What (didn't) work in teacher education? Building the future: Emergencies from education research. *Lifelong, Lifewide Learning*, 19(42), 14–24. <https://doi.org/10.19241/lll.v19i42.747>
- Black, P., & Wiliam, D. (1998a). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan*, 80(2), 139–148. <https://doi.org/10.1177/003172171009200119>
- Black, P., & Wiliam, D. (1998b). Assessment and classroom learning. *Assessment in Education*, 5(1), 7–74. <https://doi.org/10.1080/0969595980050102>
- Black, P., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2004). Working inside the black box: Assessment for learning in the classroom. *Phi Delta Kappan*, 86(1), 8–21. <https://doi.org/10.1177/003172170408600105>
- Bloom, B.S., (1969). Some theoretical issues relating to educational evaluation. In *Educational evaluation: New roles, new means*. The 63rd yearbook of the National Society for the Study of Education, part 2 (Vol. 69), ed. R.W. Tyler, 26–50. Chicago, IL: University of Chicago Press.
- Biasci, V., Domenici, G., Patrizi, N., & Capobianco, R. (2014). Teacher Self-Efficacy Scale (Scala sull'Auto-

- Efficacia del Docente – SAED): Adattamento e validazione in Italia. *Journal of Educational, Cultural and Psychological Studies*, 1(10), 485–509. <https://doi.org/10.7358/ecps-2014-010-bias>
- Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability*, 21(1), 5–31. <https://doi.org/10.1007/s11092-008-9068-5>
- Briggs, D. C., Ruiz-Primo, M. A., Furtak, E., Shepard, L., & Yin, Y. (2012). Meta-Analytic Methodology and Inferences About the Efficacy of Formative Assessment. *Educational Measurement: Issues and Practice*, 31(4), 13–17. <https://doi.org/10.1111/j.1745-3992.2012.00251.x>
- Broadfoot, P. M., Daugherty, R., Gardner, J., Gipps, C. V., Harlen, W., James, M., et al. (1999). *Assessment for learning: Beyond the black box*. Cambridge, UK: University of Cambridge School of Education. <https://doi.org/10.13140/2.1.2840.1444>
- Capanna, C., Vecchione, M., & Schwartz, S. H. (2005). La misura dei valori. Un contributo alla validazione del Portrait Values Questionnaire su un campione italiano. *Giunti Organizzazioni Speciali*, 246, 29–41. <https://doi.org/10.1421/86906>
- Capperucci D., Corsini C., Ugolini F. C. (2018). Una nuova cultura della valutazione. Valutare alunni e insegnanti, in S. Ulivieri, L. Binanti, S. Colazzo, M. Piccinno (a cura di), *Scuola Democrazia Educazione. Formare ad una nuova società della conoscenza e della solidarietà*. Lecce, Pensa Multimedia, 1013-1022.
- Ciani, A., & Vannini, I. (2017). Equità e didattica: Validazione di scale sulle convinzioni di insegnamento democratico. *Cadmo Giornale Italiano di Pedagogia Sperimentale*, 2, 5-23. <http://digital.casalini.it/10.3280/CAD2017-002003>
- Ciani, A., Pasolini, E., & Vannini, I. (2021). Il formative assessment nelle convinzioni e nelle pratiche degli insegnanti. Analisi secondarie da una indagine sui docenti di scuola media di due regioni italiane. *Journal of Educational, Cultural and Psychological Studies*, 24(2), 45-65. <https://doi.org/10.7358/ecps-2021-024-cian>
- Ciani, A., & Rosa, A. (2021). CONVAL: validazione di una scala sulle concezioni degli insegnanti relative alla valutazione. Validation of the CONVAL scale on teachers' conceptions of assessment. *Lifelong, Lifewide Learning*, 17(39), 94–113. <https://doi.org/10.19241/lll.v17i39.648>
- Cizek, G. J. (2010). An introduction to formative assessment: history, characteristics, and challenges. In Andrade HL, Cizek GJ (Eds.), *Handbook of Formative Assessment* (pp. 3-17). Routledge.
- Creswell, J. W., & Clark, V. L. P. (2017). *Designing and Conducting Mixed Methods Research*. Sage Publications.
- Coggi, C. & Ricchiardi, P. (2005). *Progettare la ricerca empirica in educazione*. Carocci.
- Corsini C., Scierri I.D.M., Scionti A. (2018). La validità delle prove INVALSI di comprensione della lettura. In A.M. Notti (Ed.), *La funzione educativa della valutazione* (pp. 46-61) Pensa.
- Corsini, C. (2023). *La valutazione che educa*. FrancoAngeli.
- Decreto del Presidente della Repubblica, 8 marzo 1999, n. 275. *Regolamento recante norme in materia di autonomia delle istituzioni scolastiche*, ai sensi dell'art. 21 della legge 15 marzo 1997, n. 59.
- Decreto del Presidente della Repubblica, 28 marzo 2013, n. 80. *Regolamento sul sistema nazionale di valutazione in materia di istruzione e formazione*.
- Ferretti, F., Paraskevi M., C., & Vannini, I. (2018). *Formative Assessment for mathematics teaching and learning. Teacher Professional Development Research by Videoanalysis Methodologies*. FrancoAngeli.
- Fives, H., & Buehl, M. M. (2012). Spring cleaning for the “messy” construct of teachers’ beliefs: What are they? Which have been examined? What can they tell us? In K. R. Harris, S. Graham, T. Urdan, S. Graham, J. M. Royer, & M. Zeidner (Eds.), *APA educational psychology handbook, Vol 2: Individual differences and cultural and contextual factors*. (pp. 471-499). American Psychological Association. <https://doi.org/10.1037/13274-019>
- Fulmer, G. W., Lee, I. C. H., & Tan, K. H. K. (2015). Multi-level model of contextual factors and teachers’ assessment practices: An integrative review of research. *Assessment in Education: Principles, Policy & Practice*, 22(4), 475-494. <https://doi.org/10.1080/0969594X.2015.1017445>
- Furtak, E. M., Kiemer, K., Circi, R. K., Swanson, R., de León, V., Morrison, D., & Heredia, S. C. (2016). Teachers’ formative assessment abilities and their relationship to student learning: Findings from a four-year intervention study. *Instructional Science*, 44(3), 267–291. <https://doi.org/10.1007/s11251-016-9371-3>
- Furtak, E. M., & Ruiz-Primo, M. A. (2008). Making students’ thinking explicit in writing and discussion: An analysis of formative assessment prompts. *Science Education*, 92(5), 799–824. <https://doi.org/10.1002/sc.20270>
- Furtak, E. M., Ruiz Primo, M. A., & Bakeman, R. (2017). Exploring the Utility of Sequential Analysis in Studying Informal Formative Assessment Practices. *Educational Measurement: Issues and Practice*, 36(1), 28–38. <https://doi.org/10.1111/emip.12143>
- Gattullo, M. (1978). *Voti, test, schede: ricerche sulla valutazione scolastica*. La Nuova Italia.

- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81–112. <https://doi.org/10.3102/003465430298487>
- Heitink, M. C., Van der Kleij, F. M., Veldkamp, B. P., Schildkamp, K., & Kippers, W. B. (2016). A systematic review of prerequisites for implementing assessment for learning in classroom practice. *Educational Research Review*, 17, 50–62. <https://doi.org/10.1016/j.edurev.2015.12.002>
- Higgins, M., Grant, F., & Thompson, P. (2010). Formative Assessment: Balancing Educational Effectiveness and Resource Efficiency. *Journal for Education in the Built Environment*, 5(2), 4–24. <https://doi.org/10.11120/jebe.2010.05020004>
- Karaman, P. (2017). Adaptation of Teachers Conceptions and Practices of Formative Assessment Scale into Turkish Culture and a Structural Equation Modeling. *International Electronic Journal of Elementary Education*, 10(2), 185–194. <https://doi.org/10.26822/iejee.2017236114>
- Kingston, N., & Nash, B. (2011). Formative Assessment: A Meta-Analysis and a Call for Research. *Educational Measurement: Issues and Practice*, 30(4), 28–37. <https://doi.org/10.1111/j.1745-3992.2011.00220.x>
- Künsting, J., Neuber, V., & Lipowsky, F. (2015). Teacher self-efficacy as a long-term predictor of instructional quality in the classroom. *European Journal of Psychology of Education*, 31(3), 299–322. <https://doi.org/10.1007/s10212-015-0272-7>
- Lee, H., Chung, H. Q., Zhang, Y., Abedi, J., & Warschauer, M. (2020). The Effectiveness and Features of Formative Assessment in US K-12 Education: A Systematic Review. *Applied Measurement in Education*, 33(2), 124–140. <https://doi.org/10.1080/08957347.2020.1732383>
- Leighton, J. P., Guo, Q., & Tang, W. (2022). Measuring preservice teachers' attitudes towards mistakes in learning environments. *Learning Environments Research*, 25(1), 287–304. <https://doi.org/10.1007/s10984-021-09362-1>
- Lui, A. M., & Andrade, H. L. (2022). The Next Black Box of Formative Assessment: A Model of the Internal Mechanisms of Feedback Processing. *Frontiers in Education*, 7, 751548. <https://doi.org/10.3389/educ.2022.751548>
- Matteucci, M. C., Corazza, M., & Santagata, R. (2015). Learning from errors, or not: An analysis of teachers' beliefs about errors and error-handling strategies through questionnaire and video. In *Progress in Education* (Vol. 37, pp. 33–54). Novascience.
- McMillan, J. H., Venable, J. C., & Varier, D. (2013). *Studies of the Effect of Formative Assessment on Student Achievement: So Much More is Needed*. <https://doi.org/10.7275/TMW-7792>
- Michel A. (2005), “France. Vers une culture de l'évaluation”, *Revue internationale d'éducation de Sèvres*, 40, 39-41. <https://doi.org/10.4000/ries.1241>
- Popham, W. J. (2008). *Transformative assessment*. Association for Supervision and Curriculum Development.
- Schellekens, L. H., Bok, H. G. J., de Jong, L. H., van der Schaaf, M. F., Kremer, W. D. J., & van der Vleuten, C. P. M. (2021). A scoping review on the notions of Assessment as Learning (AaL), Assessment for Learning (AfL), and Assessment of Learning (AoL). *Studies in Educational Evaluation*, 71, 101094. <https://doi.org/10.1016/j.stueduc.2021.101094>
- Schwartz, S. H., Melech, G., Lehmann, A., Burgess, S., Harris, M., & Owens, V. (2001). Extending the Cross-Cultural Validity of the Theory of Basic Human Values with a Different Method of Measurement. *Journal of Cross-Cultural Psychology*, 32(5), 519–542. <https://doi.org/10.1177/0022022101032005001>
- Schwartz, S. H. (1992). Universals in the Content and Structure of Values: Theoretical Advances and Empirical Tests in 20 Countries. In M. P. Zanna (Ed.), *Advances in Experimental Social Psychology* (Vol. 25, pp. 1–65). Academic Press. [https://doi.org/10.1016/S0065-2601\(08\)60281-6](https://doi.org/10.1016/S0065-2601(08)60281-6)
- Schwartz, S. H. (1994). Are There Universal Aspects in the Structure and Contents of Human Values? *Journal of Social Issues*, 50(4), 19–45. <https://doi.org/10.1111/j.1540-4560.1994.tb01196.x>
- Schwartz, S. H. (2012). An Overview of the Schwartz Theory of Basic Values. *Online Readings in Psychology and Culture*, 2(1). <https://doi.org/10.9707/2307-0919.1116>
- Schwartz, S. H. (2014). Rethinking the Concept and Measurement of Societal Culture in Light of Empirical Findings. *Journal of Cross-Cultural Psychology*, 45(1), 5–13. <https://doi.org/10.1177/0022022113490830>
- Schwartz, S. H., Melech, G., Lehmann, A., Burgess, S., Harris, M., & Owens, V. (2001). Extending the Cross-Cultural Validity of the Theory of Basic Human Values with a Different Method of Measurement. *Journal of Cross-Cultural Psychology*, 32(5), 519–542. <https://doi.org/10.1177/0022022101032005001>
- Scriven, M. (1967). The Methodology of Evaluation. In R. W. Tyler, R. M. Gagnè, & M. Scriven (Eds), *Perspectives of Curriculum Evaluation* (pp. 39–83). Rand McNally.
- Seifried, J., & Wuttke, E. (2010). Student errors: How teachers diagnose and respond to them. *Empirical Research in Vocational Education and Training*, 2(2), 147–162. <https://doi.org/10.1007/BF03546493>
- Stake, R. E. (1995). *The art of case study research*. SAGE.

- Stiggins, R. J. (1991). Assessment Literacy. *Phi Delta Kappan*, 72, 534-539. . <https://eric.ed.gov/?id=EJ514726>
- Tulis, M. (2013). Error management behavior in classrooms: Teachers' responses to student mistakes. *Teaching and Teacher Education*, 33, 56-68. <https://doi.org/10.1016/j.tate.2013.02.003>
- Van der Kleij, F. M., & Lipnevich, A. A. (2021). Student perceptions of assessment feedback: A critical scoping review and call for research. *Educational Assessment, Evaluation and Accountability*, 33(2), 345-373. <https://doi.org/10.1007/s11092-020-09331-x>
- Vannini, I. (2022). *Una valutazione che analizza e ricostruisce. Rileggere oggi "Didattica e docimologia" in dialogo con Mario Gattullo*. Armando.
- Vannini, I. (2023). La scuola di tutti e di tutte. Equità e qualità del sistema e professionalità dell'insegnante: Un'analisi incompleta. *Lifelong, Lifewide Learning*, 19(42), 25-34. <https://dx.doi.org/10.19241/lll.v19i42.748>
- Vertecchi, B. (1976). *Valutazione formativa*. Loescher.
- Vertecchi, B., Agrusti, G., Losito, B. (2010), *Origini e sviluppi della ricerca valutativa*. FrancoAngeli.
- Visalberghi, A. (1955). *Misurazione e valutazione nel processo educativo*. Edizioni di Comunità
- Wiliam, D. (2010). An Integrative Summary of the Research Literature and Implications for a New Theory of Formative Assessment. In H. L. Andrade, & G. J. Cizek (Eds.), *Handbook of Formative Assessment* (pp. 18-40). Taylor & Francis
- Wiliam, D., & Thompson, M. (2017). Integrating Assessment with Learning. What Will It Take to Make It Work? In C. A. Dwyer (Ed.), *The Future of Assessment* (pp. 53-82). Routledge. <https://doi.org/10.4324/9781315086545>
- Wiliam, D. (2019). Conclusion. Why Formative Assessment is Always Both Domain-General and Domain Specific and What Matters is the Balance Between the Two. In H. L. Andrade, R. E. Bennett, & G. J. Cizek (Eds.), *Handbook of Formative Assessment in the Disciplines* (1st ed., pp. 243-264). Routledge.
- Wisniewski, B., Zierer, K., & Hattie, J. (2020). The Power of Feedback Revisited: A Meta-Analysis of Educational Feedback Research. *Frontiers in Psychology*, 10, 3087. doi:10.3389/fpsyg.2019.03087
- Xu, Y., & Brown, G. T. L. (2016). Teacher assessment literacy in practice: A reconceptualization. *Teaching and Teacher Education*, 58, 149-162. <https://doi.org/10.1016/j.tate.2016.05.010>
- Xuan, Q., Cheung, A., & Sun, D. (2022). The effectiveness of formative assessment for enhancing reading achievement in K-12 classrooms: A meta-analysis. *Frontiers in Psychology*, 13, 990196. <https://doi.org/10.3389/fpsyg.2022.990196>
- Yan, Z., Li, Z., Panadero, E., Yang, M., Yang, L., & Lao, H. (2021). A systematic review on factors influencing teachers' intentions and implementations regarding formative assessment. *Assessment in Education: Principles, Policy & Practice*, 28(3), 228-260. <https://doi.org/10.1080/0969594X.2021.1884042>
- Yan, Z., & Pastore, S. (2022). Are teachers literate in formative assessment? The development and validation of the Teacher Formative Assessment Literacy Scale. *Studies in Educational Evaluation*, 74, 101183. <https://doi.org/10.1016/j.stueduc.2022.101183>
- Yin, R. K. (2009). *Case study research: Design and Methods*. SAGE.
- Young, J. E. J., & Jackman, M. G.-A. (2014). Formative assessment in the Grenadian lower secondary school: Teachers' perceptions, attitudes and practices. *Assessment in Education: Principles, Policy & Practice*, 21(4), 398-411. <https://doi.org/10.1080/0969594X.2014.919248>
- Zhang, Z., & Burry-Stock, J. A. (2003). Classroom Assessment Practices and Teachers' Self-Perceived Assessment Skills. *Applied Measurement in Education*, 16(4), 323-342. [https://doi.org/10.1207/S15324818AME1604\\_4](https://doi.org/10.1207/S15324818AME1604_4)