

DIPARTIMENTO DI PSICOLOGIA
DEI PROCESSI DI SVILUPPO
E SOCIALIZZAZIONE

FACOLTÀ DI MEDICINA
E PSICOLOGIA



SAPIENZA
UNIVERSITÀ DI ROMA

**Dottorato di Ricerca
in Pedagogia Sperimentale**

TESI DI DOTTORATO

**Tra pratiche e percezioni sulla valutazione dei sistemi
scolastici: confronto tra l'Italia e altri paesi europei**

Dottoranda
Eleonora Mattarelli

Tutors
Prof. Giorgio Asquini
Prof. Emilio Lastrucci

Ciclo XXXV

Anno Accademico 2022 – 2023

DIPARTIMENTO DI PSICOLOGIA
DEI PROCESSI DI SVILUPPO
E SOCIALIZZAZIONE

FACOLTÀ DI MEDICINA
E PSICOLOGIA



SAPIENZA
UNIVERSITÀ DI ROMA

**Dottorato di Ricerca
in Pedagogia Sperimentale**

Tesi di Dottorato

Dottoranda
Eleonora Mattarelli

Tutors
Prof. Giorgio Asquini
Prof. Emilio Lastrucci

XXXV Ciclo

Tra pratiche e percezioni sulla valutazione dei
sistemi scolastici: confronto tra l'Italia e altri
paesi europei

Nuova Cultura Roma

Anno Accademico
2022 – 2023

Composizione grafica a cura dell'Autore

Indice

Premessa	XI
Introduzione	XIII
Parte prima – Il quadro teorico	1
Capitolo primo – L’epistemologia della valutazione educativa	2
1.1. Il campo epistemico della valutazione	2
1.2. Misurazione e valutazione, il contributo di Visalberghi	6
1.3. I modelli della valutazione	8
1.4. La controversia anglofona	19
1.5. L’OCSE e le recenti riflessioni sulla valutazione educativa	21
Capitolo secondo – Indicatori per l’educazione e qualità dei sistemi scolastici	28
2.1. La qualità dei sistemi scolastici	28
2.2. La proposta dell’OCSE per la valutazione della qualità dell’istruzione	34
2.3. Il lento sviluppo degli indicatori per l’educazione	35
2.4. Gli indicatori dell’educazione dell’OCSE	41
2.5. <i>Education at a Glance 2022</i>	47
Capitolo terzo – Le indagini internazionali sugli apprendimenti	53
3.1. Verso le indagini comparative internazionali	53
3.2. La nascita dell’IEA	54
3.3. Le indagini TIMSS e PIRLS dell’IEA	58
3.4. L’OCSE e l’idea di una indagine	62
3.5. L’indagine PISA dell’OCSE	65
Capitolo quarto – La valutazione delle scuole, valutazione esterna ed autovalutazione	69
4.1. Qualcosa si muove... Per la valutazione scolastica	69
4.2. Tra valutazione esterna e autovalutazione delle scuole	72
4.3. La valutazione esterna	76
4.4. L’autovalutazione	79
4.5. Un ponte tra passato e presente	82
Parte seconda – La ricerca	85

Capitolo primo – Il disegno di ricerca dottorale	86
1.1. Lo studio di caso, lo stile e il metodo.....	86
1.2. Tra lo studio di caso e la comparazione nella ricerca dottorale	91
1.3. La definizione e la comparazione dei casi, i sistemi scolastici e di valutazione	98
1.4. Il problema della ricerca e i casi oggetto di analisi	103
1.5. Le domande di ricerca e gli obiettivi.....	106
1.6. Scelte metodologiche: le procedure, i partecipanti e gli strumenti di ricerca.....	109
1.6.1. Le procedure.....	110
1.6.2. I partecipanti.....	111
1.6.3. Gli strumenti di ricerca	114
1.6.4. La timeline del progetto di ricerca.....	120
 Capitolo secondo – L’analisi documentale qualitativa.....	 122
2.1. L’analisi documentale come metodo di ricerca.....	122
2.2. In quali contesti avviene la valutazione di sistema?.....	127
2.2.1. Il sistema educativo finlandese	128
2.2.2. Il sistema educativo spagnolo	130
2.2.3. Il sistema educativo italiano	132
2.3. Perché valutare la scuola?	134
2.3.1. La legislazione sulla valutazione del sistema scolastico finlandese	135
2.3.2. La legislazione sulla valutazione del sistema scolastico spagnolo	136
2.3.3. La legislazione sulla valutazione del sistema scolastico italiano	138
2.4. Quale ente è preposto alla valutazione del sistema scolastico?.....	140
2.4.1. KARVI-FINEEC, l’istituto di valutazione scolastica finlandese	141
2.4.2. INEE, l’istituto di valutazione scolastica spagnolo	143
2.4.3. INVALSI, l’istituto di valutazione scolastica italiano	144
2.5. Cosa valutano gli istituti nazionali preposti?	146
2.5.1. Il piano di valutazione per il triennio 2020-2023 di KARVI-FINEEC.....	147
2.5.2. La valutazione degli apprendimenti di INEE	150
2.5.3. La valutazione degli apprendimenti di INVALSI	155
2.6. Come avviene la valutazione delle scuole?.....	161
2.6.1. La valutazione esterna e l’autovalutazione delle istituzioni scolastiche in Finlandia	162

2.6.2. La valutazione e l'autovalutazione delle istituzioni scolastiche in Spagna.....	172
2.6.3. La valutazione esterna e l'autovalutazione delle istituzioni scolastiche in Italia.....	180
Capitolo terzo – Le interviste.....	190
3.1. Le interviste qualitative.....	190
3.2. La trascrizione e l'analisi delle interviste.....	194
3.3. Nota operativa: la traccia delle interviste e il consenso informato.....	198
3.4. Alcune riflessioni dalle interviste agli NPM dell'indagine OCSE-PISA – valutazione standardizzata degli apprendimenti.....	206
3.5. Alcune riflessioni dalle interviste agli NPM dell'indagine OCSE-PISA – valutazione della scuola e per la scuola.....	218
3.6. Alcune riflessioni dalle interviste agli NPM dell'indagine OCSE-PISA – valutazione e policy processes.....	224
Capitolo quarto – Il questionario.....	231
4.1. Il questionario ai dirigenti scolastici, verso la triangolazione italiana.....	231
4.2. La struttura del questionario.....	233
4.3. I risultati del questionario – le variabili di sfondo.....	241
4.4. I risultati del questionario – l'autovalutazione delle istituzioni scolastiche.....	246
Conclusioni.....	258
<i>Appendice</i>	282
<i>Bibliografia</i>	299
<i>Sitografia</i>	344
<i>Ringraziamenti</i>	345

To my grandparents.

*Not everything that counts can be counted
and not everything that can be counted counts.*

Premessa

Il primo ricordo del mio interesse per il sistema educativo finlandese risale alla scuola media, oggi rinominata scuola secondaria di primo grado. L'allora professoressa di lettere, per altro con un cognome molto simile al mio, stava spiegando di come Don Abbondio aveva comunicato a Renzo che il matrimonio non s'aveva da fare, quando chiedemmo di interrompere la lezione in quella noiosa ultima ora di lunedì mattina. Eravamo stanchi e di certo la lettura dei Promessi Sposi non ci aiutava nella concentrazione. Fu così che prendemmo a sfogliare e leggere insieme alla professoressa un quotidiano, uno di quelli che venivano portati gratuitamente nella nostra scuola. Ancora non lo sapevo, ma da quel giorno quello sarebbe divenuto un rito consueto in quel primo e ultimo anno insieme alla professoressa M., prossima al pensionamento. Forse c'era qualche articolo sulla Finlandia o qualche indiscrezione sull'indagine PISA del 2006... Ricordo solo che alzai la mano e, presa dall'entusiasmo, raccontai ai compagni presenti che mia mamma veniva dalla Finlandia e aveva frequentato le scuole lì. Non erano delle scuole come tutti le conoscevamo, lì c'era l'ora di cucito, di cucina, di lavorazione del ferro e del legno per tutti. La pausa si faceva ogni 50 minuti di lezione e, cosa ancora più importante per me, tutti potevano uscire fuori in giardino a prendere una boccata di aria con il sole, con la pioggia o con la neve. La risposta fu "tesoro bello, ma te lo immagini se noi del secondo piano dovessimo scendere a ogni ricreazione in giardino?! Riesci ad immaginare la confusione per le scale tra chi sale e chi scende?!". Quelle parole mi lasciarono un po' basita e spensero subito il mio entusiasmo.

Mi portai dentro quella risposta, raccontai spesso questo aneddoto anche durante gli anni del liceo, il classico. Quello che proprio in questo momento Massimo Gramellini descrive come la cyclette: se ci finisci sopra è faticoso e sembra che non porti da nessuna parte, quando termina la pedalata, invece, scopri che ti ha fornito i muscoli per andare dappertutto. Per una serie di coincidenze che fanno parte della vita, il mio percorso scolastico e universitario si è sempre esplicitato all'interno

delle aule italiane, nonostante la mia doppia cittadinanza. Mi sono sempre chiesta cosa ne sarebbe stato di me, se invece di frequentare quelle, avessi frequentato una scuola finlandese. Più didattica laboratoriale, più pause tra un'ora e un'altra, meno nozionismo, meno ore trascorse a studiare nelle "notti prima degli esami", soprattutto quelle universitarie. La faccenda della cyclette ha toccato anche Enrico Galiano, portando l'attenzione sulla questione dell'equità educativa: c'è chi può vivere la scuola come una pedalata e chi, invece, ancora oggi il tempo e i mezzi per stare in sella alla bicicletta non li ha, quindi scende presto, svolta l'angolo ed entra nel mondo del lavoro. Di fatto non è una questione di sudate e di muscoli, ma di polmoni e di aria per arrivare ovunque.

Da queste riflessioni passate, divenute nuovamente attuali, scaturisce il bisogno della ricerca, ovvero analizzare la valutazione realizzata nei sistemi scolastici per comprendere, attraverso i processi attuati, quali sono le pratiche messe a punto e le percezioni avvertite e, conseguentemente, come vengono rilevati i fattori che influiscono sulla qualità di un sistema educativo. Spesso la valutazione e la qualità, infatti, sono coniugate insieme per il miglioramento, anche se procedono su due sentieri montani differenti. Nello specifico, gli studenti italiani, finlandesi e spagnoli hanno la possibilità di rimanere in sella alla bicicletta per arrivare ovunque lo desiderino? Come è fatto il casco con il quale vengono equipaggiati? Quali sono le pratiche che i rispettivi governi nazionali mettono in atto per assicurarsi che la catena sia pulita e che le pedivelle girino normalmente senza riportare problemi nel movimento centrale? L'augurio e la speranza è che la ricerca condotta durante il percorso di dottorato, a più di quindici anni da quel lunedì mattina in compagnia del Manzoni e del sistema scolastico finlandese, trattando un tema così delicato, complesso, ma al contempo urgente e non più ignorabile, possa contribuire ad aprire nuovi interrogativi e nuove riflessioni sull'importanza di assicurare alle nuove generazioni un'educazione di qualità in cui le politiche educative dal livello macro al livello micro utilizzino la valutazione come strumento di monitoraggio costante e gli esiti di quest'ultima come base per prendere decisioni ragionate.

Introduzione

Il lavoro di ricerca condotto durante il percorso di dottorato in Psicologia Sociale, dello Sviluppo e della Ricerca Educativa si inserisce nel dibattito sulla valutazione dei sistemi educativi con l'obiettivo di esplorare e descrivere le pratiche e le percezioni legate ai processi valutativi. La discussione su queste tematiche già da qualche anno anima il contesto nazionale e internazionale con riflessioni che ruotano attorno al costrutto di qualità dell'educazione declinato e affiancato dalla sua efficienza ed efficacia (OECD, 2013; Dell'Anna, 2021). La definizione del buono o del cattivo funzionamento di un sistema educativo, infatti, è da rintracciarsi in diversi fattori che concorrono a definirne le peculiarità. Muovendo da questa consapevolezza, si vogliono analizzare le caratteristiche di questi aspetti scattando una fotografia dello stato dell'arte della valutazione dei sistemi scolastici in tre realtà europee: l'Italia, la Finlandia e la Spagna (OECD, 2004).

Riprendendo lo studio realizzato per il lavoro di tesi magistrale in Pedagogia Sperimentale, la consultazione dei documenti e l'ascolto delle voci dei diversi testimoni privilegiati tenta di disporre le delicate e complesse tessere del puzzle della valutazione.

Una valutazione autentica, valida e affidabile, quella che porta al miglioramento delle pratiche educative a tutti i livelli e degli apprendimenti degli studenti, è fondamentale per un sistema di istruzione di qualità. È anche determinante ed essenziale per il riconoscimento del lavoro degli operatori educativi e per la certificazione dell'apprendimento degli studenti. Promuovere la valutazione è, dunque, nell'interesse di studenti e famiglie, dei professionisti dell'educazione e dei sistemi educativi. Per questo motivo, sono sempre più numerosi i paesi che intraprendono ambiziosi programmi di riforme scolastiche che includono pratiche valutative del sistema educativo. Questo può consolidare la cultura della valutazione, rafforzandone il ruolo nei framework e nei programmi che guidano le riforme. (OECD, 2013, p. 3, traduzione a cura dell'autrice)

La struttura della tesi si divide in due parti. La prima parte, articolata in quattro capitoli, propone una ricostruzione del framework teorico che, partendo dalla valutazione come costruito inteso in senso lato, si focalizza sulla valutazione dei sistemi educativi in senso specifico e sulle forme che questa può assumere.

Il primo capitolo si apre contestualizzando il campo epistemico della valutazione e la questione linguistica che intorno ad esso si avvicenda. Dopo aver delineato la valutazione così come è stata teorizzata da Dewey (1939, 1939), prosegue con il contributo offerto da Visalberghi (1955, 1958) e con la ricostruzione storico-epistemologica dei modelli di valutazione (Stame, 2001; Castoldi, 2013; Trincherò, 2014a), per concludersi riprendendo la controversia linguistica anglofona (OECD, 2013; Corsini, 2015; Corsini, 2018c).

Il secondo capitolo inizia a delineare una delle forme che la valutazione dei sistemi educativi può assumere, la valutazione della qualità attraverso l'elaborazione di indicatori (Vertecchi, 2003). Numerosi, infatti, sono gli autori e gli organismi che hanno tentato di sistematizzare il concetto di qualità riferito all'istruzione, tra questi vi è l'OCSE (Lastrucci, 2017b; Kyriakides et alii, 2020; Dell'Anna, 2021). La proposta empirica ed operativa dell'OCSE ha avuto inizio nel secolo scorso con il lento cammino degli indicatori che, dopo la battuta di arresto degli anni Settanta e Ottanta, sono tornati nel dibattito internazionale a seguito della conferenza tenutasi a Washington D.C. nel 1987 per discutere del report *A Nation at Risk: the imperative for educational reform* (Gardner, 1983; Bottani, 2009; Asquini, 2018) che ha dato il via, a partire dal 1992, alla pubblicazione annuale di una raccolta intitolata *Education at a Glance* (OECD, 2022).

Il terzo capitolo prosegue nella delineazione delle diverse forme della valutazione dei sistemi educativi soffermandosi sulle indagini internazionali e sugli *outcomes* dell'istruzione. L'IEA, *International Association for the Evaluation of Educational Achievement*, e l'OCSE, Organizzazione per la Cooperazione e lo Sviluppo Economico, ormai da qualche decennio si occupano di misurare in modo standardizzato gli apprendimenti degli studenti. Mentre l'IEA basa le proprie indagini sui contenuti dei curricula scolastici dei paesi partecipanti e ha come target

gli studenti della scuola primaria e della scuola secondaria inferiore (Mullis et alii, 2016; Martin et alii, 2017), l'OCSE nell'indagine PISA, uno dei più grandi studi sperimentali mai realizzati, rileva le competenze dei quindicenni di tutto il mondo in lettura, matematica e scienze con cadenza triennale (OECD, 2019abc).

Il quarto capitolo conclude la delimitazione delle diverse forme della valutazione dei sistemi educativi, arrivando a quella che riguarda le istituzioni scolastiche (OECD, 2004). L'urgenza di valutare sia esternamente che internamente le scuole si è mostrata dopo l'esperienza inglese dell'OFSTED (Matthews & Smith, 1995; Poliandri & Romiti, 2012). Se la valutazione esterna prevede la realizzazione di visite da parte di valutatori, l'autovalutazione delle istituzioni scolastiche si pone come momento di riflessione corale tra gli agenti interni (OECD, 2013). Entrambe frutto dei processi di decentralizzazione e *accountability*, si mescolano e sintetizzano all'interno dei piani di miglioramento nei quali vengono indicate le azioni da intraprendere per il progresso e lo sviluppo (Reding, 2002; Robasto, 2017; Marzano, 2019).

La seconda parte, relativa all'indagine empirica, anche essa articolata in quattro capitoli, ripercorre le scelte metodologico-procedurali fatte durante il lavoro di ricerca e ne illustra i risultati.

Nel primo capitolo è descritto il disegno della ricerca dottorale. È introdotto lo studio di caso (Stake, 1995; Merriam, 1998; Yin 2009) e il modello di comparazione utilizzato (Jullien, 1817; Bray & Thomas, 1955; Brickman, 1960), sono enucleati i bisogni conoscitivi della ricerca ed è argomentata la scelta dei casi "rivelatori" (Sena, 2021). Inoltre, sono formulate le domande di ricerca e gli obiettivi e sono illustrate le scelte metodologiche, ovvero le procedure seguite, i partecipanti selezionati e gli strumenti adoperati. La ricerca segue due livelli di analisi (Zanazzi, 2014): il livello internazionale si sofferma sulla documentazione prodotta dagli organismi sovranazionali, il livello nazionale corrobora i dati raccolti attraverso il primo analizzando caso per caso (Scheerens, 2019).

Nel secondo capitolo è descritta l'analisi documentale qualitativa condotta (Bowen, 2009) e i risultati ad essa collegati. Il modello di Cardone e colleghi (2010) ha guidato la restituzione delle evidenze emerse. Per ciascun paese oggetto di indagine, Italia, Finlandia e Spagna,

sono contestualizzati il sistema educativo di riferimento e gli atti legislativi e sono descritti gli istituti nazionali di valutazione, Invalsi, KARVI ed INEE, e il contenuto delle valutazioni da loro condotte, in particolar modo quelle che riguardano gli apprendimenti su scala nazionale. Infine, sono presentate le pratiche di valutazione esterna e autovalutazione delle istituzioni scolastiche.

Nel terzo capitolo sono affrontate le riflessioni raccolte dalle interviste realizzate telematicamente con i National Project Manager dell'indagine PISA dell'OCSE, riportandone il processo di messa a punto che ha previsto una traccia uguale per tutti e alcune domande suppletive di approfondimento. Le suggestioni, analizzate attraverso l'analisi tematica riflessiva, sono sintetizzate attraverso i tre temi principali emersi dall'analisi e aderenti alle domande di ricerca: valutazione standardizzata degli apprendimenti, valutazione della scuola e per la scuola e valutazione e policy processes (Zammuner, 1996; Braun & Clarke, 2006; Di Fraia, 2014; Terry et al, 2017; Pagani, 2020).

Nel quarto e ultimo capitolo, invece, sono affrontati i risultati emersi dal questionario dedicato ai dirigenti scolastici italiani. La descrizione della costruzione ad hoc dello strumento è seguita dalla modalità duplice di somministrazione: online-manuale attraverso Google Moduli e telefonica (Denzin & Lincoln, 1994; Vitalini, 2010; Lombi, 2015; Benvenuto, 2015). Le pratiche e le percezioni legate alla valutazione delle istituzioni scolastiche sono analizzate in due momenti: nel primo si descrive il gruppo di riferimento con le caratteristiche socio-ambientali di sfondo, nel secondo, invece, è approfondita la tematica dell'autovalutazione come pratica per il miglioramento.

Il paragrafo finale, quello delle conclusioni, riprendendo le fila del discorso, problematizza i risultati emersi a partire dai bisogni conoscitivi scaturiti all'inizio del lavoro, racconta dell'incontro avvenuto con un dirigente scolastico finlandese e apre la possibilità ad alcuni sviluppi futuri per la ricerca.

Parte prima

IL QUADRO TEORICO

Capitolo primo

L'epistemologia della valutazione educativa

1.1. Il campo epistemico della valutazione

Nel 2008, mentre nella scuola primaria si passava dai giudizi ai voti, il ministro Tremonti commentava quel cambiamento nelle pagine del Corriere della Sera.

Il '68 ha portato via i voti sostituendoli con i giudizi. I numeri sono una cosa. I giudizi sono una cosa diversa. I numeri sono una cosa precisa, i giudizi sono spesso confusi. Ci sarà del resto una ragione perchè tutti i fenomeni significativi sono misurati con i numeri. Un terremoto è misurato con i numeri della scala Mercalli o Richter. Il moto marino è misurato in base alla scala numerica della "forza", la pendenza di una parte di montagna in base ai "gradi", la temperatura del corpo umano ancora in base ai "gradi". La mente umana è semplice e risponde a stimoli semplici. I numeri sono insieme precisi e semplici. Il messaggio che trasmettono è un messaggio diretto. Se gli stessi fenomeni - terremoto, moto marino, pendenza, temperatura corporea - fossero espressi non con numeri ma attraverso frasi complesse con finalità descrittive, il messaggio resterebbe impreciso. [...] Ogni valutazione deve mettere capo a una classifica. Questa è la logica della valutazione. Se non c'è una classifica, non c'è neanche una reale valutazione. Nella scuola inglese ad esempio gli studenti sono addirittura classificati in un ordine rigido. In ogni classe esiste un primo classificato, un secondo classificato e così via. Mi sembra francamente un'esagerazione. Ma non mi sembra affatto un'esagerazione tornare a dare i voti come una volta: 10, 9, 8, e così via, perchè la verità è semplice: dare un giudizio senza una classifica significa non dare affatto un giudizio reale. (Il passato e il buon senso, in Corriere della Sera, 22 agosto 2008)

Il termine valutazione è presente ormai da molti anni nel lessico educativo. Se ne parla tra i banchi di scuola, sui giornali, in televisione, ma non sempre in modo consapevole e ancor più in chiave pedagogica. Non tutte le valutazioni, infatti, possono essere definite educative e la questione linguistica che ruota intorno a questo termine è tutto fuorchè in via di risoluzione.

John Dewey è tra i primi studiosi del ventesimo secolo ad occuparsi di valutazione e di utilizzo dei vocaboli ad essa collegati. Secondo il pedagogista statunitense, affinché una valutazione possa essere definita educativa, deve possedere tre caratteristiche imprescindibili (Dewey 1938, 1939):

1. deve esprimere un *giudizio di valore*;
2. deve basarsi sulla *distanza rilevata* tra una *situazione auspicata* e una *situazione effettivamente riscontrata*;
3. deve avere come obiettivo la *presa di decisioni* finalizzata alla *riduzione di tale distanza*.

Per Dewey (1938, 1939) e per i suoi successori il processo valutativo ha inizio da una situazione che presenta una o più problematiche e per la quale si vuole trovare una soluzione. La partecipazione cosciente e consapevole degli attori coinvolti nel processo, infatti, dà loro modo di esprimere un giudizio di valore (Corsini, 2015).

Le scienze sociali hanno bisogno di riflettere sulla valutazione poiché discipline come la matematica, la fisica e l'astrologia, essendo esatte, per l'appunto, non esprimono un giudizio di valore ma guardano ai numeri. Il nocciolo della questione, quando si discute della teoria della valutazione, non verte sulle scienze esatte che godono di una rilevanza scientifica in quanto tali, bensì sulla possibilità che anche le *faccende umane* e le scienze sociali possano essere oggetto di *proposizioni genuine* sulla loro direzione (Dewey, 1939).

A proposito della questione linguistica, invece, c'è un'interessante riflessione che deriva dal lessico utilizzato nel processo valutativo: entrambi i termini, "valutazione" e "valutare", nel linguaggio quotidiano sono utilizzati sia per intendere l'atto del *pregiare*, quindi del ritenere di valore qualcosa o qualcuno, sia l'atto dell'*apprezzare*, quindi del dare un valore, un prezzo a qualcosa o a qualcuno (Dewey, 1939). La radice di questi termini deriva dal latino "valuto" e "valitus" ovvero essere forte, sano e robusto e da "valeo, valere" ovvero avere valore¹. La questione linguistica intorno ai due termini, tuttavia, non si esaurisce

¹ Per approfondimenti circa l'etimo della parola "valutazione": <https://www.etimo.it/?term=valutare> (ultima consultazione 21 febbraio 2023)

con l'etimo: quando si attribuisce un valore a qualcosa o a qualcuno lo si fa all'interno di una sfera emotiva che ha riferimenti strettamente personali e soggettivi, quando, invece, si attribuisce un prezzo a qualcosa o a qualcuno, questo avviene sulla base di intrinseche proprietà oggettive dell'oggetto sottoposto a valutazione.

Qual è dunque l'accezione più indicata nelle valutazioni educative?

Stando al significato originale della parola "valutazione" questa rimanda all'idea di stima e di considerazione. Questo è il motivo per il quale, tornando alla prima considerazione di Dewey (1938, 1939), se si intende la valutazione come atto attraverso il quale il valutatore esprime un giudizio e, dunque, attribuisce un valore all'oggetto valutato, non si può prescindere dal considerarlo come un atto in cui emerge chiaramente la dimensione soggettiva. La soggettività per lungo tempo ha rappresentato un punto di forza e un punto di debolezza per la docimologia (Vertecchi, 2003; Calenda & Milito, 2020): se da una parte può dar vita a distorsioni che pregiudicano il processo valutativo, dall'altra gli può conferire autenticità (Corsini, 2015, 2018ab). Infatti, è la percezione soggettiva che abbiamo di una situazione problematica che funge da molla per l'azione e per il miglioramento della stessa.

La cultura della valutazione che si è affermata si percepisce invece come oggettiva, fondata sul numero e incentrata sulla classificazione di meriti e colpe. Spesso si parla di "fotografia del presente", una metafora abusata e del tutto ingiustificata: una fotografia dipende da inquadratura, focale, tempi di scatto, sensibilità iso. Non esiste una fotografia oggettiva. (Corsini, 2018a, p. 204)

La seconda considerazione deweyana (Dewey, 1938, 1939) riporta l'attenzione sulla misura, vale a dire sulla lontananza che c'è tra il punto A, la situazione osservata, e il punto B, la situazione auspicata. A suo parere nel processo valutativo, quando si misura la distanza tra la situazione attuale e quella osservata, c'è sempre una discrepanza.

Qualche osservazione del risultato ottenuto in paragone e in contrasto con quello atteso, tale che il paragone getti luce sulla effettiva adeguatezza delle cose adoperate come mezzi. (Dewey, 1939, p. 37)

In tal senso la misurazione è un momento particolarmente importante del processo valutativo secondo il pedagogista statunitense, infatti, più gli attori coinvolti sono coscienti della distanza che intercorre tra il punto A e il punto B, più prenderanno decisioni che avranno come obiettivo quello di colmare tale distanza, collegando l'astrattezza della prima alla concretezza dello scopo ultimo da raggiungere (Lucisano, 1989). Si tratta di una misurazione che può essere espressa tanto in formato numerico e quantitativo tanto in formato qualitativo. Sebbene secondo alcuni sia importante dare una forma quantitativa agli esiti di tale misurazione (Gattullo, 1989), secondo altri a fare la differenza non è la formulazione, bensì il fine della misurazione stessa, molto spesso di natura qualitativa.

È assurdo insistere sulla misurazione numerica quando il fine a cui la proposizione quantitativa si connette come mezzo a conseguenza è qualitativo, come è assurdo accontentarsi di una misurazione qualitativa (che in tal caso sarebbe un lavoro congetturale) nel caso di altre finalità. È in ultima analisi la natura del problema trattato che decide che genere di comparazione-misurazione sia richiesta al fine di ottenere una determinata soluzione. Alcuni si lamentano della riduzione a termini numerici, altri tutto vorrebbero ridurre a numero [...], gli uni e gli altri perdono di vista il significato logico della misurazione, che è determinato dal riferimento strumentale delle proposizioni quantitative ad una conseguenza intenzionale oggettiva. Gli uni e gli altri considerano le proposizioni come definitive e compiute, quando in realtà sono intermedie e strumentali. (Dewey, 1938, p. 260)

Sebbene sembra essersi diffuso il pensiero comune secondo il quale "numerico" equivale a "degnò di fiducia" e di "affidabilità", non sempre questa equazione è confermata (Corsini, 2019ab). La terza e ultima considerazione deweyana (Dewey, 1938, 1939), effettivamente, chiarisce il motivo alla base dei processi valutativi: se una valutazione non è percepita come utile, se il fine ultimo non è evidente nella mente di chi valuta e di chi viene valutato, allora si rischia di guardare alla valutazione come a un adempimento burocratico calato dall'alto e non autenticamente vissuto dai partecipanti in essa coinvolti (Ellerani, 2006;

Swaffield, 2011; Lucisano & Corsini 2015; Lucisano, 2015; Corsini, 2018ab; Trinchero, 2018).

1.2. Misurazione e valutazione, il contributo di Visalberghi

Qualche anno dopo le pubblicazioni di Dewey (1938, 1939) il pedagogista e filosofo Aldo Visalberghi ne divenne promotore nel panorama nazionale: dapprima conobbe le opere deweyane durante un'esperienza da borsista negli Stati Uniti, poi si fece promulgatore del maestro statunitense nel contesto italiano. L'Erasmus americano, così lo potremmo definire ai giorni nostri, fu per Visalberghi un'occasione importante per avvicinarsi alla visione deweyana della ricerca pedagogica (Vertecchi, 2007): gran parte del Vecchio Continente di allora era avulso delle spinte innovative che provenivano dagli Stati Uniti tant'è che il pedagogista italiano non solo tradusse alcune opere dell'autore americano che confluirono in adattamenti italiani (Visalberghi, 1951), ma a partire dalle concettualizzazioni di Dewey elaborò una serie di considerazioni in merito alla valutazione che racchiuse in *Misurazione e valutazione nel processo valutativo* pubblicato nel 1955, in *Esperienza e valutazione* pubblicato nel 1958 e in altre opere (Visalberghi, 1955, 1958, 1977b, 1978, 1988).

Il suo interesse per la valutazione non derivava dalla rincorsa di una razionalità fine a sé stessa, che trovava nelle misurazioni un punto d'approdo, consentendo di osservare dall'esterno i fenomeni educativi, individuando i punti di forza e più spesso di debolezza dell'attività delle scuole. Per Visalberghi la valutazione era una forma alta di conoscenza, che supposeva l'impegno solidale di quanti, a vario titolo, condividevano la responsabilità dell'azione educativa. Alla valutazione dovevano offrire il loro contributo i politici (cui spetta assumere decisioni circa l'indirizzo del sistema), gli amministratori (dal cui impegno derivano condizioni più o meno positive per lo sviluppo dell'attività delle scuole), le forze sociali (il cui ruolo è di far emergere gli orientamenti della domanda di istruzione), e soprattutto gli insegnanti (i più vicini agli allievi, i più pronti a rilevare le loro esigenze e a fornire ad esse prime risposte). (Vertecchi, 2007, p. 2)

Negli anni in cui scriveva Visalberghi, anni particolarmente importanti per la scuola italiana e per i cambiamenti che in essa stavano avvenendo, si pensi ad esempio all'avvento della scuola media unica (Lamma Valentini et alii, 1964), il dibattito docimologico era particolarmente caldo (Melchiori, 2002; Benvenuto, 2008; Castoldi, 2012). Molti studiosi insieme a Visalberghi si interrogavano sul dualismo oggettività-soggettività nella valutazione. Richiamando Dewey (1939), la discussione verteva sul possibile utilizzo di test standardizzati per misurare e valutare gli esiti negli apprendimenti degli studenti eliminando la componente soggettiva, in particolar modo dal momento della correzione. Visalberghi (1955), inserendosi in questa tavola rotonda, invitava a riflettere sui due momenti che caratterizzano i processi valutativi, ovvero la misurazione e la valutazione. Mentre la misurazione permette di raccogliere dati quantitativi circa l'oggetto di indagine, la valutazione accompagna tutto il processo, dalla costruzione della prova all'elaborazione di un giudizio di valore che sia quanto più equilibrato possibile (Giacomantonio, 2008; Losito, 2010).

La parola «misura» ha due significati principali, che non sono scollegati come taluni aspetti della vita moderna potrebbero far temere. Non c'è nessuna ragione di fondo per cui la «misura» intesa come operazione di conteggio o confronto non debba accompagnarsi con la «misura» intesa come abito di equilibrio e discrezione. Si potrebbero fare, è vero, sottili analisi circa l'origine classica dei due significati ed il loro uso rinascimentale, ma non crediamo che i risultati sarebbero in contrasto con la semplice osservazione di buon senso che l'abito stesso del misurare, implicando l'attitudine a vedere un più ed un meno dove il giudizio affrettato scorge qualità assolute è esso stesso un abito di riflessività, di moderazione e di prudenza. Come si spiega allora che la tendenza contemporanea a «misurare» tutto si accompagni spesso con atteggiamenti completamente opposti? [...] Mi limiterò a congetture che forse vale per la «misura» ciò che vale per la scienza in generale: la cosa più grave è di volersi fermare a metà strada, dimenticando o non volendo ricordare che essa sola, come la famosa lancia d'Achille, può guarire i mali che produce. Fermarsi a metà strada significa, di solito, tentare uno sfruttamento affrettato di risultati provvisori e parziali, tentarlo in base a vecchi abiti che non sono quelli che la scienza stessa coraggiosamente perseguita, fa maturare, ma appartengono invece all'oscuro passato. [...] Ma il test come tale non valuta più,

misura. Esso rappresenta un momento intermedio, di pura misurazione, inserito fra serie complesse di atti valutativi. Ciò che nelle prove tradizionali è fuso assieme, e talvolta confuso, nella pratica del testing viene accuratamente distinto. Nonché sopprimere la valutazione, la misurazione nasce dalla valutazione e nella valutazione confluisce. Essa mantiene però una provvisoria, ma nettissima autonomia, che le permette di raggiungere, come vedremo, una precisione ed un'attendibilità altrimenti impossibili. (Visalberghi, 1955, pp. 11-18)

Il modo di intendere il rapporto tra misurazione e valutazione ha, secondo il pedagogista italiano, implicazioni sul piano socio-culturale e politico (Corsini, 2018b). L'esortazione da parte sua, infatti, non era quella di schierarsi per la misurazione o per la valutazione, per la corrente dei "quantitativi" o per quella dei "qualitativi", bensì di distinguere all'interno dei processi valutativi fasi differenti: le fasi strettamente quantitative-misurative di somministrazione degli strumenti e di raccolta dei dati sono seguite dalle fasi più qualitative in cui ciò che è stato quantificato viene analizzato e interpretato da un punto di vista qualitativo - i cosiddetti *atti valutativi* che richiedono la formulazione di un giudizio.

Le fasi di misurazione e di valutazione sono fasi di uno stesso processo che si conclude con la formulazione di giudizi o con l'attribuzione di voti. Chi intende i due termini come antitetici o contrapposti confonde le fasi del processo di valutazione, rilevazione e analisi, con i possibili momenti nei quali esso si sviluppa. (Benvenuto & Giacomantonio, 2008, p. 4)

Le riflessioni di Visalberghi (1955) suggeriscono dunque un rapporto dialettico e interconnesso tra misurazione e valutazione in ambito educativo di cui non si può non tenere conto (Corsini, 2018c).

1.3. I modelli della valutazione

Barbier (1989) dopo Dewey (1938, 1939) e Visalberghi (1955) torna a interrogarsi sulla valutazione e sull'accezione che questa acquista quando viene "praticata" nei contesti educativi. Riprendendo quanto già

messo in luce dal suo predecessore (Dewey, 1939), rimarca sia la dimensione soggettiva dei processi valutativi sia la dimensione rappresentativa. Ogniqualvolta ci si trova a valutare si aziona un processo di mediazione tra chi valuta e chi/cosa viene valutato. In altre parole ogni processo valutativo è doppio.

Un duplice processo di rappresentazione, il cui punto di partenza consiste in una rappresentazione fattuale di un oggetto e il punto di arrivo in una rappresentazione codificata di questo stesso oggetto. (Barbier, 1989, p. 76)

Nello specifico, il duplice processo di rappresentazione si articola in due momenti, così come già messo in luce da Visalberghi (1955). Nella rappresentazione fattuale il valutatore raccoglie dati e informazioni per farsi un'idea di ciò che andrà a valutare, nella rappresentazione codificata interpreta, valuta ed esprime un giudizio elaborato con uno specificato codice comunicativo sulla realtà oggetto di indagine (Barbier, 1989).

In termini culturali la definizione proposta sottolinea il ruolo ineludibile affidato al soggetto nel processo valutativo; poiché nella sensibilità attuale lo stesso processo conoscitivo non può essere pensato come puro rispecchiamento di una realtà esterna al soggetto, tanto più un processo valutativo, che implica l'attribuzione di un giudizio di valore, non può prescindere dal soggetto che valuta. (Castoldi, 2013, p. 1)

Dello stesso avviso sono Beeby (1977) e Galliani (2009). Il primo riprende i due momenti della valutazione educativa (raccolta sistematica dei dati e interpretazione di essi) già espressi da Visalberghi (1955) e l'idea deweyana di sviluppo di un giudizio di valore avente come fine l'azione migliorativa (Dewey, 1938, 1939; Beeby, 1977). Il secondo definisce la valutazione come una "disciplina finalizzata a emettere giudizi sulle azioni formative e di insegnamento" (Galliani, 2009) che sono state progettate intenzionalmente per avere delle ricadute sui destinatari in termini di apprendimenti con risultati non solo formativi, ma afferenti anche alla sfera economica e sociale (Galliani, 2009, 2015).

Nel secolo scorso, così come si sono annoverati molti studiosi intorno alle riflessioni sui concetti di misurazione e valutazione in educazione, si è sviluppato un florido dibattito anche intorno alla ricerca valutativa e ai modelli ad essa riconducibili.

Dagli anni Novanta, infatti, i processi valutativi non solo sono stati oggetto di riflessione circa il significato, le forme e le funzioni che assolvono, ma hanno seguito anche gli sviluppi degli approcci e dei paradigmi teorici cui fanno riferimento (Stame, 2001; Galliani & Notti, 2014; Tammaro et alii, 2017).

La dicotomia soggettività-oggettività è soltanto uno degli esempi che spiega la così coinvolgente discussione sulla varietà degli approcci teorici cui si può far appello in campo valutativo (Castoldi, 2013). Guba e Lincoln (1982, 1989), ad esempio, distinguono due paradigmi di riferimento: un paradigma docimologico-positivista e un paradigma sistemico-fenomenologico. Le principali differenze sono:

- *Natura della realtà*: nel paradigma docimologico-positivista la realtà è unica e tangibile e scomponibile; nel paradigma sistemico-fenomenologico la realtà è multipla, complessa e studiabile solo olisticamente;

- *Relazione osservatore-oggetto osservato*: nel paradigma docimologico-positivista l'osservatore è ben distinto dall'oggetto osservato; nel paradigma sistemico-fenomenologico l'osservatore ha un'influenza non trascurabile sull'oggetto osservato e viceversa;

- *Natura degli enunciati di verità*: nel paradigma docimologico-positivista l'indagine è finalizzata a raggiungere obiettivi generalizzabili e non contestualizzabili; nel paradigma sistemico-fenomenologico l'indagine è finalizzata a descrivere casi specifici con riferimento a particolari contesti;

- *Spiegazione dell'azione*: nel paradigma docimologico-positivista per ogni azione/effetto/risultato reale c'è una causa tangibile; nel paradigma sistemico-fenomenologico ogni azione/effetto/risultato è il frutto dell'interazione tra molteplici fattori;

- *Il ruolo dei valori nell'indagine*: nel paradigma docimologico-positivista si cerca di ridurre al minimo le influenze soggettive; nel paradigma sistemico-fenomenologico l'indagine è influenzata

esplicitamente dai valori espressi dal ricercatore/valutatore e dalle scelte da esso operate (Lincoln & Guba, 1982; Castoldi, 2013).

Cosa valutare nel momento in cui si valuta? Perché valutare? Come valutare? Quando valutare? Le risposte a tutte queste domande delineano le strade che possono essere intraprese dai processi valutativi e gli approcci e i paradigmi che si possono chiamare in causa. Ad essere valutati possono essere, infatti, i risultati negli apprendimenti, le politiche educative e i sistemi educativi nel loro complesso. Il valutatore può decidere di effettuare una valutazione all'inizio, una sommativa alla fine o una formativa in itinere (Weeden et alii, 2009; Notti, 2010, 2014; Falcinelli, 2014).

Se si prende a riferimento la definizione di valutazione data da Bezzi (2000) e Palumbo (2001) quest'ultima può essere considerata come un'*attività cognitiva* che ha l'obiettivo di fornire un giudizio su una o più azioni intenzionali che producono effetti esterni di qualsiasi natura e che si basano su un'attività di ricerca che segue procedure rigorose e strumenti progettati per realizzarla (Tammaro et alii, 2017). Palumbo (2001), inoltre, sottolinea come nei processi valutativi non solo l'oggetto della valutazione, ma anche le finalità - elaborare un giudizio di valore basato su criteri specifici - e le procedure - comparare secondo specifiche modalità di raccolta e interpretazione dei dati - hanno ricadute sulla riuscita del processo stesso.

Sembra a questo punto evidente nell'esposizione fin qui illustrata che il problema della valutazione non è solo riconducibile alla questione linguistica, ma anche all'approccio/paradigma di riferimento. Al variare del focus della valutazione stessa, infatti, varia anche il modello cui fare riferimento. Pensando di voler dare enfasi all'impatto che i processi valutativi generano nei contesti in cui avvengono, ad esempio, è possibile richiamare quella corrente presente in letteratura che ha come criterio principe le necessità sociali specifiche di un determinato contesto, come riportato più avanti (Wahba & Bridwell, 1976).

Trincherò (2014b) individua otto modelli di valutazione:

- *Goal-based*: il processo valutativo ha lo scopo di verificare che gli obiettivi prefissati siano stati raggiunti e che la distanza tra il punto A e il punto B sia stata colmata (Dewey, 1938, 1939); spesso attraverso questo

approccio si vuole verificare l'efficienza e l'efficacia di un intervento educativo (Tyler, 1942);

- *Goal-free*: il processo valutativo ha lo scopo di descrivere ciò che è stato effettivamente raggiunto, a prescindere dagli obiettivi prefissati; l'attenzione è concentrata sulla situazione osservata e sugli effettivi effetti di un intervento educativo (Scriven, 1972);

- *Responsive*: il processo valutativo ha lo scopo di analizzare la responsività degli interventi educativi rispetto alle problematiche e ai bisogni posti dagli utenti; è una valutazione che emerge dal contesto, di tipo bottom-up, e dal significato che le danno i soggetti coinvolti in esso (Stake, 1975, 1980, 1983);

- *Orientato al giudizio*: il processo valutativo, attuato da esperti esterni, ha lo scopo di formulare giudizi autorevoli (Patton, 1997);

- *Antagonista*: il processo valutativo ha lo scopo di evidenziare i pro e i contro, i punti di forza e le criticità dell'intervento valutativo (Wolf, 1979);

- *Empowerment*: il processo valutativo, nel momento in cui si esplica, ha lo scopo di stimolare nei partecipanti all'intervento educativo la consapevolezza dei propri punti di forza e di debolezza (Fetterman, 2001);

- *Decision-based*: il processo valutativo è sorretto dalla presa di decisione che avviene prima che esso abbia luogo (Stufflebeam, 1973);

- *Casi di successo*: a partire dai casi di successo sono messe in luce, attraverso il processo valutativo, le potenzialità dell'intervento educativo (Brinkerhoff, 2003; Trincherò, 2014b).

Castoldi (2013), invece, individua cinque dimensioni chiave di analisi del processo valutativo:

- *Scopi*: quando lo scopo è quello del controllo, la valutazione assume la funzione di accertamento e rendicontazione, soprattutto quando sono presenti rapporti gerarchici tra gli attori coinvolti essa è volta a verificare che siano stati raggiunti gli obiettivi prefissati o a sanzionare se questo non è avvenuto; quando lo scopo è quello dello sviluppo, la valutazione assume le caratteristiche di un'azione retroattiva, ovvero orienta e ricalibra le azioni in funzione degli obiettivi

prefissati proponendo sia linee di miglioramento che responsabilizzazione degli attori coinvolti;

- *Approcci*: quando l'approccio è top-down-applicazionista le azioni si originano in alto nella struttura gerarchica e a cascata procedono verso il basso e verso la periferia, a riprodursi è una proposta *preconfezionata* poichè i soggetti deresponsabilizzati dipendono dal vertice; quando l'approccio è bottom-up-costruttivista le azioni si originano dal basso nelle reti dei soggetti, questi ultimi infatti sono direttamente coinvolti e presentano istanze reali in tutta autonomia e responsabilità;

- *Soggetti*: i responsabili del processo valutativo possono essere interni o esterni; quando sono interni coloro che sono oggetto di valutazione sono contemporaneamente responsabili di essa; quando sono esterni a decidere sulla valutazione sono coloro che non sono direttamente coinvolti nell'azione da valutare (Castoldi, 2013).

Da un lato una valutazione interna può rafforzare la valutazione esterna in molteplici direzioni:

- affinare la sensibilità della valutazione esterna in rapporto alle peculiarità e ai tratti contestuali della singola esperienza;
- arricchire e calibrare in rapporto ai processi reali e alle condizioni di contesto la lettura e l'interpretazione dei risultati;
- favorire l'impiego dei risultati della valutazione in funzione della regolazione e della progettazione della propria attività progettuale;
- coinvolgere in prima persona gli attori nella verifica del proprio operato.

Dall'altro la valutazione esterna consente a quella interna di:

- disporre di parametri di riferimento comuni e di indici di comparazione utili a posizionarsi;
- stimolare processi autovalutativi utili ad esplorare in profondità e comprendere il profilo emergente dall'osservazione esterna;
- legittimare il valore e il grado di confidenza dei processi autovalutativi attraverso il confronto con altri dati ed opinioni;
- potenziare l'impatto e il grado di utilizzabilità dei processi di autovalutazione in rapporto agli scopi e alle conseguenze connesse ad una valutazione esterna. (Castoldi, 2013, p. 3)

La bibliografia più recente in merito alla polarizzazione interno-esterno suggerisce, infatti, di integrare i due poli (Cerini, 2016).

- *Oggetti*: nell'ambito scolastico possono essere distinti in quattro focus:

- risultati negli apprendimenti degli studenti;
- azioni didattiche degli insegnanti;
- organizzazione del servizio scolastico (a livello di scuola);
- sistema scolastico nel suo complesso.

- *Funzioni*: nell'ambito scolastico possono essere distinte in cinque tipologie:

- valutazione predittiva/orientativa: avviene prima dell'inizio del percorso formativo e ha lo scopo di prevedere ed orientare quest'ultimo in base alle caratteristiche dei soggetti coinvolti;

- valutazione diagnostica: avviene all'inizio del percorso formativo e ha lo scopo di raccogliere informazioni sui soggetti coinvolti;

- valutazione formativa: affianca tutto il percorso formativo e fornisce continuamente feedback ai soggetti coinvolti sull'andamento dello stesso;

- valutazione sommativa: avviene alla fine del percorso formativo e ha lo scopo di restituire un feedback conclusivo ai soggetti coinvolti;

- valutazione certificativa: è una conseguenza del percorso formativo svolto e ha lo scopo di rilasciare un attestato e riconoscere il conseguimento di determinati risultati (Domenici, 2001; Palumbo, 2001; Coggi & Notti, 2002; Pellerey, 2006; Castoldi, 2009, 2012).

Un'altra tassonomia dei modelli di valutazione è suggerita da Stame (2001). Stame (2001) raggruppa le numerose proposte di modelli di valutazione presenti in letteratura in tre famiglie principali che ruotano attorno ai paradigmi:

- *Positivistica - sperimentale*: è un approccio nato negli Stati Uniti con l'obiettivo di trovare una soluzione tangibile a problemi urgenti di ordine sociale (es. disoccupazione, disagio e criminalità). È indicato particolarmente per programmi politici in cui con un andamento ciclico si passa dalla presa di decisione, all'implementazione, alla valutazione e a una nuova presa di decisione. I programmi, infatti, sono caratterizzati da obiettivi da perseguire, risultati attesi e strumenti attraverso cui

raggiungerli. La valutazione ha quindi lo scopo di verificare, tramite indicatori costruiti ad hoc per la misurazione, il raggiungimento di quanto prefissato e se è presente una discrepanza tra quanto prefissato e la reale situazione. Grazie al programma di intervento si è verificato un cambiamento? Questo cambiamento è dovuto davvero al programma di intervento o alle concause ad esso associate?

Lo scopo della valutazione è verificare l'efficacia del programma nel conseguire un obiettivo che dovrebbe risolvere un problema, e quindi la generalizzabilità di un tale intervento in altre situazioni simili. Per tali motivi, si ritiene più adatto il metodo sperimentale (o quasi-sperimentale, Campbell 1969), nel quale si mette a confronto un gruppo sperimentale sottoposto al programma con un gruppo equivalente che non riceve il programma (gruppo di controllo): in questo modo, le differenze osservate tra i due gruppi dovrebbero indicare le responsabilità del programma nel realizzare il risultato desiderato. (Stame, 2001, p. 21-46)

Le tecniche di misurazione dei risultati raggiunti sono perlopiù quantitative, si passa dal questionario ai modelli econometrici, ovvero quei modelli che permettono di fare sia un confronto pre e post-intervento sia un confronto tra il gruppo sperimentale che riceve il trattamento e il gruppo di controllo che non lo riceve. I limiti di questo approccio sono molteplici: innanzitutto non è possibile spiegare perché avviene un cambiamento, quando ciò avviene, ma è solo possibile rilevare l'avvento del cambiamento. Risulta difficile trovare dati sulla situazione pre-intervento e controllare che i risultati raggiunti non siano il frutto di concause che hanno agito durante il trattamento stesso. Inoltre, risulta difficile anche scorgere effetti inattesi che non rientrano nel disegno iniziale poiché si è predisposti a vedere solo ciò che è stato inizialmente stabilito. I risultati, infine, si possono conoscere solo molto tempo dopo la fine del programma in contrasto con la brama di conoscenza dei politici che sulla base dei risultati devono prendere decisioni.

Non si fa una valutazione, costruttiva (*formative*), per decidere cosa sta funzionando o meno ed eventualmente modificare il programma insieme agli

amministratori, ma si attende di vedere i risultati del programma per permettere al decisore di stabilire se esso merita di essere proseguito o interrotto (valutazione riepilogativa, o *summative*). [...] A questo approccio corrisponde un'idea di utilizzazione strumentale della valutazione: la valutazione dovrebbe fornire input di informazione che i politici trasformino in decisioni. (Stame, 2001, p. 21-46)

Il fatto che questo approccio sia in grado di fornire risposte chiare, anche se non sempre immediate, lo elegge come quello più adatto agli occhi dei decisori politici sia nei programmi di investimenti pubblici che nei programmi sociali (Hyman, 1967; Suchman, 1967; Weiss, 1972; Chen, 1990; Stame, 1998; Campbell & Russo, 1999; Rossi et alii, 2018).

- *Pragmatista - della qualità*: è un approccio nato in opposizione al positivismo che non riconosce il ruolo dei valori portati all'interno dei processi valutativi dagli attori interni o dai valutatori. Per il pragmatismo, diversamente dal positivismo, i valori sono centrali nella ricerca valutativa che esiste perchè esiste l'atto del giudicare.

Secondo Scriven, fondatore dell'approccio (egli lo chiama in vario modo: prospettivista, transdisciplinare), il valutatore non deve farsi influenzare dagli obiettivi del programma, e la sua deve essere una "goal free evaluation"; mettendo il programma e i suoi obiettivi sopra tutto - come nell'approccio "goal oriented" dei positivisti - ci si fornirebbe invece un alibi per una pretesa neutralità valoriale del valutatore. Quest'ultimo deve dare un giudizio sul programma in base alle proprie competenze e ai valori, che sono oggettivi e propri alla situazione ed al programma. (Stame, 2001, p. 21-46)

L'oggetto della valutazione è dunque la qualità, che essendo di difficile delimitazione, viene scomposta in vari aspetti che coincidono con altrettanti indicatori e standard di raggiungimento. Un'attività può avere una qualità intrinseca es. un corso di francese di qualità grazie alle competenze dell'insegnante che lo tiene, ma può avere anche una qualità estrinseca nel momento in cui incontra le necessità degli utenti es. anche se l'insegnante di francese è bravo, il corso non ha valore perchè l'obiettivo degli utenti è imparare l'inglese o un'altra lingua (Stame, 2001). In questo approccio il valutatore deve in un primo momento

stabilire gli standard di qualità, poi deve misurare la performance dandole un punteggio, infine deve creare una graduatoria di merito dei programmi ed elaborare per ciascuno un giudizio di valore.

Qui dunque l'elemento di confronto non è più ciò che un singolo programma può voler raggiungere in una data situazione (obiettivi), ma ciò a cui tutti i programmi simili dovrebbero mirare su una certa scala di merito se vogliono essere considerati di buona qualità. Si tratta di una differenza di fondo, che viene spesso sottovalutata considerando che anche gli obiettivi potrebbero essere formulati come standard, ma senza tener conto del diverso ruolo assunto dal valutatore nel giudicare un programma rispetto ad altri simili. (Stame, 2001, p. 21-46)

Nel misurare la performance rispetto agli standard e agli indicatori, la valutazione ha un ampio campo di azione attraverso i "giudizi degli esperti" che formulano opinioni rispetto allo scostamento dagli standard. Questo approccio viene utilizzato soprattutto nei sistemi di *performance review* e nel *New Public Management* poichè è in grado di restituire un feedback sulla produttività di un servizio o di confrontare servizi simili gestiti da altre strutture. In questo caso la valutazione viene utilizzata strumentalmente per prendere decisioni amministrative. Il limite di questo approccio sta nella certificazione. Negli sviluppi più recenti, infatti, l'approccio della qualità si è accostato sempre di più alla certificazione e il valutatore è diventato colui che certifica il raggiungimento di standard e assicura che vengano seguite determinate procedure (Scriven, 1982; Gore 1993; Pollitt 1995).

- *Costruttivista - del processo sociale*: è un approccio che, più che riporre la propria attenzione al disegno iniziale del programma, si focalizza sul contributo dei vari attori e su come evolve e si modifica il programma quando viene attuato. Viene applicato soprattutto su programmi pilota poichè la pietra della valutazione, vale a dire il successo, non è possibile definirla a priori, ma è un concetto che si costruisce in un momento successivo all'avvio del programma.

La valutazione deve tener conto del fatto che ogni volta che si attua un programma esso muta a contatto con il contesto: i problemi sentiti dai vari

stakeholders saranno diversi, e le conclusioni raggiunte per un programma non potranno essere generalizzate ad altre situazioni in cui vengono attuati programmi simili. Ciò non significa tuttavia che non si possano trarre delle lezioni da applicare anche altrove. (Stame, 2001, p. 21-46)

Ogni programma deve essere analizzato nel contesto sociale e istituzionale in cui viene attuato, infatti per avere una visione quanto più ampia possibile ci si può servire di interviste in profondità e analisi di casi o di strumenti quantitativi se adeguati a cogliere aspetti ritenuti centrali.

La cosa importante è che le parti siano coinvolte in una valutazione partecipata, che il processo sia seguito nelle sue varie fasi da valutatori che interagiscono con gli stakeholders, e che in presenza di interessi e posizioni in conflitto i valutatori possano assumere ruoli che vengono chiamati in letteratura con nomi che vanno da arbitri, a negoziatori, a facilitatori. Dall'altra parte, il valutatore dovrà fare appello alla propria esperienza per confrontare il caso analizzato con altri, più o meno simili. Dal confronto nasceranno giudizi e lezioni che potranno poi essere reinterpretate in altri contesti. (Stame, 2001, p. 21-46)

L'interpretazione è centrale in questo approccio e la valutazione è utilizzata nel suo uso conoscitivo (Cronbach et alii, 1980; Stake, 1980).

I modelli e le tassonomie di valutazione descritti da Trincherò (2014b), Castoldi (2013) e Stame (2001) benchè non esaustivi, risultano interessanti da un punto di vista epistemologico, anche se più o meno distinti. Nella pratica educativa quotidiana, dal momento che la valutazione parte dalla conoscenza e poichè per conoscere abbiamo bisogno di avvalerci di più fonti informative (Dewey, 1939), risultano spesso mescolate in un metodo combinatorio-creativo (Eco, 2004; Notti, 2014; Tammaro et alii, 2017). Ogni approccio, infatti, permette di conoscere meglio un aspetto della sfaccettata realtà.

1.4. La controversia anglofona

La controversia tra *assessment* and *evaluation* ha caratterizzato il dibattito anglofono degli ultimi decenni.

Se per *assessment* si intende l'*accertamento educativo*, con *evaluation* si vuole indicare il più generale *processo valutativo* (Corsini, 2018c). Non tutti, tuttavia, concordano con questa accezione. Argentin (2016) richiama il termine *assessment* per esprimere un giudizio sui risultati raggiunti o sulle procedure seguite per raggiungerli nel piano di erogazione di un servizio e il termine *evaluation* per la valutazione degli effetti cioè per l'analisi delle conseguenze dell'applicazione di una politica pubblica. Qualsiasi sia il significato che si vuole attribuire ai due termini anglofoni, in una *cultura valutativa inclusiva* emerge la funzione formativa e di sviluppo della valutazione in cui è chiara, nello specifico dell'ambito scolastico-educativo, sia l'integrazione tra valutazione e insegnamento sia l'integrazione tra valutazione e apprendimento (Scriven, 1967; Black & Wiliam, 1998; Greenstein, 2010; Corsini, 2018c).

Le evoluzioni anglosassoni degli ultimi decenni circa la valutazione educativa sono state caratterizzate dalla dicotomia tra *assessment for learning* e *assessment as learning* e dall'introduzione dell'*accountability*. A partire dalla valutazione formativa e dalle teorie dell'apprendimento l'*assessment for learning* si rifà alla valutazione per l'apprendimento, mentre l'*assessment as learning* alla valutazione come apprendimento (Corsini, 2015, 2018abc). Sull'*accountability* la questione è più spinosa e nasce insieme all'autonomia scolastica e universitaria. Il termine indica, infatti, la valutazione delle procedure e quindi la rendicontazione di un intervento, di una politica o di un servizio (Argentin, 2016), ma in un'ottica di controllo delle prestazioni. La cultura della valutazione che fa propria l'*accountability* dà spesso alla valutazione una funzione sommativa.

In modo da rendere conto dell'efficienza degli investimenti in istruzione e in vista di un miglioramento che dovrebbe scaturire da meccanismi competitivi. [...] L'*accountability* chiede riferimenti tali da assicurare l'opinione pubblica sul carattere rigoroso del controllo. A scuole e università vengono pertanto

imposti tempi, modalità e strumenti valutativi concepiti per dare luogo a misure oggettive. (Corsini, 2018c, p. 89)

Nella cultura dell'*accountability* le misure, intermedie e strumentali in Dewey (1939), diventano il fine ultimo della valutazione, la distanza dal punto A al punto B diventa una distanza da valori medi che viene percepita come un deficit e l'obiettivo non è più quello di produrre un giudizio di valore, ma di assegnare premi e punizioni (Corsini, 2015; Lucisano & Corsini, 2015; Corsini, 2018c).

Si tratta di un meccanismo in cui a farla da padrone, come direbbe la cultura popolare, sono il bastone e la carota. Un esempio concreto è il modello anglofono del *pay for performance* che implica l'utilizzo dell'*assessment* e dell'*evaluation* per dispensare riconoscimenti o sanzioni soprattutto a insegnanti e dirigenti scolastici (Ballou, 2001). È un approccio alla valutazione nato negli Stati Uniti dove negli anni 2000 è stata emanata la legge "*No child left behind act – test based accountability*" in cui, per certificare i risultati scolastici raggiunti dai singoli istituti, era richiesto di applicare la logica del *pay for performance*. Da questa legge sono nati alcuni programmi federali, come *Teacher incentive fund* o *Race to the top*, che sono stati criticati qualche anno dopo dagli stessi studiosi che avevano costruito gli strumenti per quel tipo di *accountability* (McGiunn, 2012; Simpson et alii, 2014; Malen et alii, 2015).

Nonostante l'*assessment* e l'*evaluation* abbiano acquisito funzioni molto radicali in alcuni momenti della storia anglofona, non sono mancati modelli di valutazione meno polarizzati. Si pensi, ad esempio, al *formative assessment* (Scriven, 1967), all'*authentic assessment* (Wiggins, 1998), al già citato *assessment for learning* (William, 2000; Stiggins, 2002; Wiliam, 2011), all'*inclusive assessment* (Watkins, 2007), all'*assessment as learning* (Earl, 2013) o all'*impact evaluation* (Kreber et alii, 2001). Quest'ultima è la valutazione di ciò che accade a livello sociale quando viene attuata una politica pubblica, anche educativa. È realizzata grazie all'aiuto di tecniche controfattuali di analisi che stimano gli effetti causali prodotti da un intervento. Produce una quantità di dati cospicua e non è trascurabile in ambito scolastico dal momento che è un'utile cartina al tornasole per gli stakeholders nella comprensione

dell'efficienza e dell'efficacia delle politiche e degli interventi educativi. Favorendo l'autoriflessione, fornisce indicazioni sulla funzionalità e sulla fattibilità dei programmi di intervento così da supportare i contesti educativi tramite *evidence based policy* e *evidence informed policy* (Leiber et al, 2015; Argentin, 2016).

1.5. L'OCSE e le recenti riflessioni sulla valutazione educativa

Sulle procedure, sugli scopi, sugli oggetti e sui soggetti della valutazione negli ultimi venti anni si è espressa anche l'OCSE, Organizzazione per la Cooperazione e lo Sviluppo Economico, un'organizzazione di portata mondiale di stampo economico che, sostenendo la discussione e il confronto tra i paesi membri e i paesi partner, ha l'obiettivo di fornire delle chiavi di lettura e di risoluzione dei problemi che accomunano la collettività². In ambito pedagogico ed educativo l'OCSE a tutt'oggi analizza l'importanza dell'istruzione e le ricadute sociali ed economiche di quest'ultima (Hanushek, 2011; Hanushek & Woessmann, 2015; OECD, 2019abc; OECD, 2020abc).

Nel 2009, mentre si realizza la quarta edizione dell'indagine PISA, approfondita nel terzo capitolo della prima parte, viene istituito il programma di monitoraggio intitolato *OECD Review on Evaluation and Assessment Frameworks for Improving School Outcomes* (OECD, 2013). La review, conclusa in quattro anni, si prefiggeva di analizzare, sulla scorta del crescente interesse mostrato dagli stakeholders, le politiche di *assessment* ed *evaluation* attuate nei 25 paesi parte del disegno di ricerca. L'interesse dei team di ricerca nazionali e internazionali non era rivolto solamente alle modalità con le quali vengono misurati i livelli di apprendimento raggiunti dagli studenti, ma anche alla performance di scuole, insegnanti, dirigenti scolastici e sistemi educativi nel loro complesso.

² Per approfondimenti circa la storia dell'OCSE: https://it.wikipedia.org/wiki/Organizzazione_per_la_cooperazione_e_lo_sviluppo_economico (ultima consultazione 21 febbraio 2023)

In questa sede l'OCSE distingue tre possibili tipologie di valutazione dei sistemi scolastici (OECD, 2013; Dell'Anna, 2021):

- *Assessment*: per la valutazione delle competenze degli studenti;
- *Appraisal*: per la valutazione di insegnanti, dirigenti scolastici e professionisti coinvolti nell'ambito educativo;
- *Evaluation*: per la valutazione dell'efficacia e dell'efficienza delle istituzioni scolastiche.

La figura 1 esemplifica graficamente il quadro concettuale della misurazione e della valutazione, gli aspetti coinvolti e il modo in cui questi ultimi si interconnettono tra di loro.

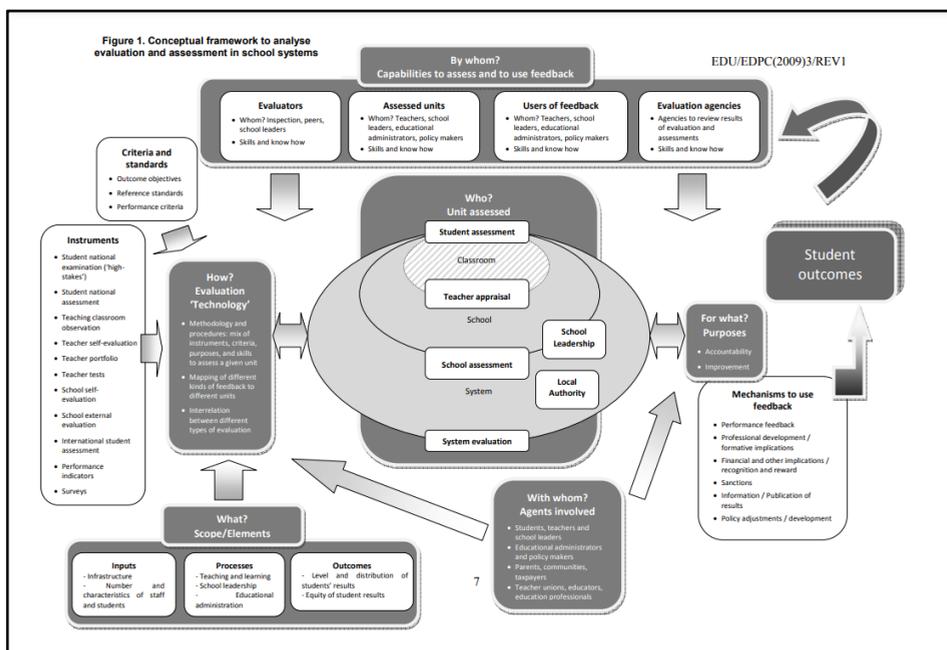


Figura 1. Framework per l'analisi dell'*assessment* e dell'*evaluation* nei sistemi scolastici. Fonte: OECD, 2009b.

Il quadro concettuale riportato raffigura sei aspetti interconnessi tra di loro:

- **Unità valutata:** chi? I soggetti della valutazione possono essere gli studenti, un insegnante, un dirigente scolastico, una scuola o un ente

locale educativo, una parte di un sistema educativo o il sistema educativo nel suo complesso;

- Capacità di valutare e utilizzare il feedback: da chi? Per valutare bisogna essere preparati a valutare, essere preparati ad essere valutati e saper diffondere e utilizzare i risultati di una valutazione: i valutatori devono essere scelti e formati, i valutati devono essere consapevoli di esserlo, i risultati devono essere utilizzati efficacemente in ottica migliorativa, le agenzie valutative devono contemporaneamente rivedere i risultati della valutazione e informare gli stakeholders;

- Scopi: cosa? Gli oggetti della valutazione possono essere i risultati come i livelli raggiunti dagli studenti negli apprendimenti o l'equità, i processi come la qualità dell'insegnamento e dell'apprendimento o l'efficacia della dirigenza scolastica, inoltre anche input come la qualità dell'infrastruttura scolastica o la proporzione tra studenti e insegnanti possono essere oggetti della valutazione;

- Metodologia: come? La valutazione può essere organizzata in modi differenti, ovvero può far riferimento a un mix di strumenti, criteri e standard, competenze, obiettivi generici e specifici;

- Finalità: per cosa? La valutazione generalmente ha l'obiettivo di migliorare la situazione preesistente tramite *l'accountability* e *l'improvement* e di sviluppare responsabilità in chi è coinvolto. Esempi di modalità di utilizzo dei risultati della valutazione includono feedback sulle prestazioni, premi e punizioni economiche, progetti di formazione professionale, pubblicazione dei risultati e revisione delle politiche;

- Agenti coinvolti: con chi? Nella valutazione possono essere più o meno coinvolte parti interessate: studenti, insegnanti e dirigenti scolastici, famiglie, sindacati, cittadini e decisori politici (OECD, 2009b).

Nel report pubblicato nel 2013 (OECD, 2013), a partire dal framework appena delineato dell'*assessment* e dell'*evaluation*, vengono descritti comparativamente i processi e le politiche adottate, ma anche le ricadute di queste ultime nei paesi aderenti al programma coordinato dall'OCSE.

Per quel che riguarda i trend delle politiche valutative appare chiaro che negli ultimi decenni è aumentata l'importanza rivestita da tali pratiche.

È evidente che l'*assessment* e l'*evaluation* sono sempre più considerate leve che guidano il cambiamento, il miglioramento, la responsabilità, la pianificazione educativa e lo sviluppo delle politiche all'interno dei sistemi scolastici. I paesi stanno sviluppando quadri di riferimento più completi per l'*assessment* e l'*evaluation*, ponendo maggiore enfasi sulla misurazione e sugli indicatori educativi sviluppati, dando crescente importanza all'uso responsabile dei risultati e facendo affidamento sempre più sugli standard educativi. In secondo luogo, [...] l'*assessment* e l'*evaluation* hanno acquisito importanza in risposta a un maggiore livello di autonomia scolastica, a un ruolo più forte dell'istruzione nei meccanismi di mercato, all'emergere del *New Public Management*, all'imperativo sempre maggiore di un uso efficiente delle risorse pubbliche, alla necessità di puntare alla "qualità per tutti" e alla crescente importanza dell'istruzione in un mondo globale. Altri fattori contestuali che influenzano lo sviluppo di quadri per l'*assessment* e l'*evaluation* includono l'aumento delle aspettative sulla professionalità degli insegnanti, i genitori più istruiti, la promozione dell'uso del processo decisionale basato sull'evidenza, i progressi tecnologici, gli interessi commerciali emergenti in materia di istruzione e il ruolo dei media. (OECD, 2013, p. 35, traduzione a cura dell'autrice)

È chiaro, dunque, uno sforzo comune nel mantenere i sistemi educativi allineati con i rapidi cambiamenti della società e la ferma volontà di assicurare istruzione di qualità alle giovani generazioni. In tutti i paesi in modi più o meno evidenti si sta assistendo a trend politici comuni che vanno, come già riportato, dal maggiore decentramento alla maggiore autonomia e che, di conseguenza, richiedono forme di accertamento valutativo e *assessment* su larga scala e individualizzato. Per i processi valutativi nel loro complesso, invece, il bisogno espresso dai paesi è quello di maggiore *evaluation* dato dalla crescente autonomia acquisita dalle istituzioni scolastiche (OECD, 2013; Corsini 2018c).

Emerge una comune tendenza verso modalità di valutazione finalizzate a rinnovare il settore pubblico applicandovi logiche di mercato, proprie del mondo aziendalistico. Nell'analisi delle politiche e pratiche valutative, risultano rappresentativi i concetti di *quality assurance* e *quality management*, ma anche quelli di rapporto costi-efficacia (*cost-effectiveness*) e competizione. (Dell'Anna, 2021, p. 29)

Ciò che deriva dall'*assessment* e dall'*evaluation* diventa in questo modo uno strumento per comparare i "livelli di sviluppo economico" e l'"efficacia dei sistemi pubblici" e non più per comprendere solamente lo stato di salute di un sistema scolastico e i possibili interventi migliorativi (Dell'Anna, 2021). Non tutti i paesi hanno la medesima tradizione di ricerca valutativa e il medesimo approccio, tuttavia, ci sono alcune priorità su cui tutti sono concordi (OECD, 2013):

- Avere un approccio olistico per comprendere il pieno potenziale della valutazione;
- Allineare l'*assessment* e l'*evaluation* con gli obiettivi educativi per promuovere l'apprendimento degli studenti;
- Concentrarsi sul miglioramento delle pratiche didattiche in modo che qualsiasi valutazione abbia un valore educativo e un riscontro pratico;
- Evitare distorsioni ed effetti collaterali indesiderati nell'utilizzo di alcune forme di valutazione, ad esempio i test; il focus deve essere sempre incentrato sulle necessità degli studenti e degli insegnanti;
- Mettere gli studenti al centro dei processi di insegnamento-apprendimento poichè essi devono essere completamente coinvolti nel loro apprendimento e devono poterlo autovalutare; bisogna tenere conto dei risultati negli apprendimenti più ampi, come le competenze sociali e relazionali: l'*assessment* e l'*evaluation* relativi alla performance devono comprendere un ampio spettro di variabili e devono basarsi sia su dati quantitativi che qualitativi;
- Costruire un framework di riferimento per l'*assessment* e l'*evaluation* a tutti i livelli del sistema educativo (strumenti e linee guida), ricomprendendo anche insegnanti e dirigenti scolastici;
- Ascoltare le esigenze locali perchè è importante trovare un giusto equilibrio tra gli scopi centralizzati/nazionali e i bisogni del singolo contesto;
- Progettare e sperimentare con successo sulla base delle *best practices* e generare consenso tra gli stakeholders che sono molto più aperti al cambiamento se ne comprendono razionalmente il motivo (OECD, 2013).

È ampiamente riconosciuto che i modi in cui l'*assessment* e l'*evaluation* vengono utilizzati in un sistema scolastico hanno ricadute dirette sulla possibilità di migliorarne l'efficienza, l'efficacia e la qualità. Questo è il motivo per il quale hanno acquisito e stanno acquisendo crescente importanza nelle agende educative nazionali. Ci sono, infatti, diversi fattori che negli ultimi anni hanno fatto sì che l'*improvement* e l'*accountability* venissero percepiti come necessità urgenti. Le sfide economiche e sociali sono diventate sempre più imponenti e questo fa sì che all'istruzione sia richiesta sempre più efficacia, efficienza, equità e qualità accompagnata da spinte all'autonomia e alla decentralizzazione e al miglioramento dell'utilizzo delle Tecnologie per l'Informazione e la Comunicazione (TIC). Ciò consente lo sviluppo di valutazioni sia su larga scala che individualizzate, facilitando la condivisione e la gestione dei dati, ma anche un maggiore senso di fiducia nei confronti della valutazione, favorendo l'*evidence-based decision making* (OECD, 2013).

Al netto, dunque, del significato che si vuole dare alle parole *misurazione* e *valutazione* o *assessment* ed *evaluation* e al netto degli effetti collaterali negativi a cui possono dare vita, queste ultime dovrebbero essere utilizzate come strumenti per monitorare il sistema scolastico e il livello raggiunto dagli studenti negli apprendimenti e per raccogliere informazioni da fornire agli stakeholders. Gli insegnanti dovrebbero essere incoraggiati a utilizzare la pratica dell'autovalutazione in classe, a prestare attenzione al "valore aggiunto" e a monitorare regolarmente l'apprendimento degli studenti. Tuttavia, a livello di sistema per i paesi non è sempre facile attuare le pratiche legate all'*assessment* e all'*evaluation* poiché spesso si scontrano con una mancanza della cultura della valutazione che deriva da una non adeguata progettazione delle politiche e da una scarsa capacità di mettere in atto le procedure valutative che causa il verificarsi di effetti collaterali indesiderati, per l'appunto, e un uso non adatto dei risultati della valutazione. Se da una parte i responsabili politici concedono sempre più autonomia gestionale e didattica alle scuole, dall'altra vogliono essere sempre più informati sull'andamento del sistema scolastico, studenti, insegnanti e dirigenti scolastici compresi. Ciononostante, questi ultimi non sempre sono in grado di utilizzare i dati per prendere decisioni informate. Questo è il

motivo per il quale *assessment* ed *evaluation* sono fondamentali per definire strategie per il miglioramento delle pratiche nei sistemi scolastici, ma debbono essere utilizzate considerando la valutazione come mezzo e non come fine dell'educazione (Dewey, 1951; Corsini, 2012; OECD, 2013; Zanazzi & Corsini, 2015; Castellana & Corsini, 2018; Capperucci, 2018; Corsini, 2018; Corsini, 2019; Corsini, 2020b; Corsini & Gueli, 2022).

Capitolo secondo

Indicatori per l'educazione e qualità dei sistemi scolastici

2.1. La qualità dei sistemi scolastici

La valutazione si compone di due fasi: la prima prevede la raccolta empirica dei dati, la seconda, invece, prevede l'analisi delle informazioni raccolte e la formulazione di un giudizio di valore (Dewey, 1939; Giovannini, 2014; Corsini, 2018ab; Dell'Anna, 2021).

Come messo in evidenza nel capitolo precedente, negli anni si sono sviluppate diverse teorie sulla ricerca valutativa che hanno portato allo sviluppo di diversi modelli e di diverse tassonomie di valutazione con ricadute differenti in base all'utilizzo (Stame, 2001; Castoldi, 2013; Trincherò, 2014b; Fondazione Agnelli, 2014). Infatti, ogni momento del processo valutativo, dall'implementazione alla diffusione dei risultati, non può non essere contestualizzato.

Data la fluidità del concetto di qualità, la descrizione e l'analisi dei sistemi di valutazione non possono prescindere da considerazioni e riflessioni sulla dimensione valoriale, di tipo teorico, metodologico e culturale, dalla fase di raccolta dati fino a quella della disseminazione delle evidenze valutative. La definizione di modelli per l'analisi e lo studio dei fenomeni risponde, in primo luogo, all'esigenza di ridurre la complessità dei sistemi al minimo. (Dell'Anna, 2021, p. 23)

Lastrucci (2017a) propone una gamma di variabili input che entrano a far parte dei modelli di valutazione e che possono essere riassunti in variabili di processo, ad esempio il livello degli apprendimenti degli studenti e l'esperienza degli insegnanti, e indicatori utilizzabili per rappresentare sinteticamente fenomeni complessi, ad esempio la motivazione. Nella valutazione dei sistemi scolastici e della loro qualità è doveroso prendere in considerazione il sistema di indicatori costruiti grazie alla combinazione di indici differenti (Lastrucci, 2017a; OECD, 2019ab).

I modelli di valutazione si distinguono tra loro per tre aspetti (Fondazione Agnelli, 2014):

- Il *perchè* della valutazione (obiettivi e finalità);
- Il *cosa* della valutazione (oggetti);
- Il *come* della valutazione (processi seguiti, strumenti adoperati, analisi effettuate).

In relazione all'ultimo aspetto, quello relativo ai processi, è necessario inoltre prestare attenzione a chi valuta, ossia alle istituzioni, enti, professionisti e soggetti impegnati nei processi di raccolta, analisi, interpretazione e disseminazione. (Dell'Anna, 2021, p. 24)

Nel tenere in considerazione il *perchè*, il *cosa* e il *come* della valutazione, tuttavia, bisogna contemporaneamente riflettere sul concetto di qualità riferito all'istruzione. Pionieri nell'utilizzo di questo termine nei processi di valutazione sono stati gli ambiti economici e aziendalistici da cui il costrutto è stato successivamente mutuato per l'educazione e nello specifico per la docimologia, le teorie del curricolo e la pedagogia sperimentale (Lastrucci, 2017b).

Nonostante il concetto di qualità dell'istruzione sia oggi largamente utilizzato anche nel parlato quotidiano, quest'ultimo è legato a molteplici accezioni e molteplici modalità attuative. Il dibattito internazionale degli ultimi decenni, infatti, ha sottolineato che ha anche delle forti accezioni politiche: la qualità della scuola e la qualità dei processi educativi sono legati innegabilmente alla qualità del sistema di istruzione e alla valutazione per il miglioramento di quest'ultimo (Rondini, 1995).

Tribus (1994), inoltre, evidenzia che la scuola ha a che fare con una pluralità di utenti, dunque, il concetto di qualità dell'istruzione non può non prendere in considerazione tutti i loro punti di vista, nonchè tanto i processi educativi quanto gli *outcomes* (Harvey & Green, 1993).

Negli anni Cinquanta in Giappone e negli Stati Uniti si è sviluppata intorno al concetto di "qualità dell'educazione" la teoria del *Total Quality Management*, un framework di gestione aziendale secondo il quale l'azienda non produce solo *outcomes* ma anche qualità e secondo il quale

tutti i dipendenti possono concentrarsi sul miglioramento della qualità dell'azienda stessa contribuendo alla costruzione della *customer satisfaction*³. Alcuni aspetti di questo modello di gestione aziendalistico sono stati ripresi nella valutazione dell'ambito educativo seguendone alcuni assunti: la qualità del prodotto è collegata alla qualità del processo produttivo; il miglioramento è un processo che cresce su se stesso; il coinvolgimento di tutti gli attori nei processi di implementazione e miglioramento è importante; le risorse umane che popolano il mondo dell'istruzione e della formazione debbono essere valorizzate; l'attenzione sia all'individuo sia all'organizzazione nel suo complesso è fondamentale (Sutcliffe & Pollock, 1992; Tribus, 1994; Hubbard, 1994; Hansen, 1994; Rusconi, 1995; Shank et alii, 1996; Bondioli, 2000).

La singola istituzione educativa, come ogni organizzazione, è un organismo complesso (Romei, 1995) la cui qualità è legata a una serie di aspetti difformi, a livelli di contesto dissimili, a prospettive di analisi differenziate, a ottiche a volte in contrasto tra loro. Non è facile dare una definizione di qualità di una istituzione educativa: si tratta comunque di declinare un'idea globale di "buona scuola" in una serie di determinazioni particolari, legate a differenti livelli di contesto e a prospettive difformi. Esistono aspetti che potremmo definire di "qualità intrinseca", relativi cioè a determinazioni particolari e obiettivamente verificabili su cui concordano esperti del settore, insegnanti, dirigenti; aspetti di qualità percepita da diversi soggetti che hanno a che fare con l'istituzione educativa (genitori, allievi, insegnanti, ordini di scuola successivi, mercato del lavoro); aspetti di qualità organizzativa e gestionale verificabili obiettivamente, spesso predefiniti da esperti (amministratori, dirigenti, insegnanti, ricercatori) ma apprezzabili da molte istanze che sono in rapporto con la scuola. Bisogna fare in ogni caso i conti con le aspettative dei diversi attori sociali in gioco, ma anche con le risorse disponibili e con le precondizioni della situazione presa in esame, con le garanzie che consentono a tutti di svolgere al meglio il proprio compito. I diversi aspetti della qualità - qui se ne sono elencati solo alcuni - differiscono per le prospettive di analisi che di volta in volta vengono adottate,

³ Per approfondimenti circa la Qualità Totale: https://it.wikipedia.org/wiki/Qualit%C3%A0_totale (ultima consultazione 21 febbraio 2023)

per i soggetti chiamati ad esprimere un giudizio, ad attribuire valore a delle determinazioni particolari che, nel loro insieme, concorrono a definire il profilo qualitativo della singola istituzione. Il concetto di qualità di una istituzione educativa è allora legato a quella di valore, oltre che a quella di contesto. Se compiere una operazione valutativa a livello educativo significa raccogliere sistematicamente dei dati, interpretarli secondo un giudizio di valore che ha immediati risvolti operativi (Beeby, 1977; Wolf, 1987), dire che una scuola è di qualità significa emettere un giudizio di valore sulla base di presupposti legati ad un gruppo particolare e ad un livello di contesto altrettanto determinato, pensando però di incidere sulla situazione di partenza. Altro è analizzare la qualità degli arredi e delle attrezzature nella singola classe, altro riflettere sulle modalità in cui si articola il processo educativo, altro indagare sulla percezione del servizio che hanno gli utenti esterni (genitori, alunni) o interni (insegnanti, personale ausiliario e di segreteria), altro mettere in evidenza gli aspetti gestionali del servizio nel suo complesso, altro riflettere sui rapporti tra la singola istituzione e l'Amministrazione da cui dipende, altro ancora sulla normativa in vigore e sulle implicazioni che le scelte assunte a livello di macrosistema hanno nel quotidiano. (Ferrari, 2000, pp. 71-72)

La valutazione della qualità è, dunque, un'operazione che assume significati propri in base ai punti di vista degli stakeholders coinvolti, ai differenti livelli di contesto e alle teorie a cui si fa riferimento. Lastrucci (2017b) sintetizza le principali teorie della qualità in:

- *Quality assurance*: capacità di definire e raggiungere standard formativi;
- *Contract conformance*: capacità delle singole istituzioni scolastiche e del sistema di "garantire la qualità del prodotto formativo" (Lastrucci, 2017b, p. 174);
- *Customer driver*: capacità di raggiungere obiettivi formativi costruiti sulla base delle necessità dell'utenza;
- *School-tailored management*: capacità di impiegare le logiche imprenditoriali per i fini educativi delle istituzioni scolastiche.

Mentre la prima teoria si rifà a una visione aziendalistica della scuola produttrice di competenze, la seconda e la terza guardano all'istruzione come a un bene pubblico e dunque alla qualità dell'istruzione come un

diritto sancito dalla Costituzione, la quarta, infine, concilia i primi tre modelli di qualità.

Scheerens (2011a), al contrario, per la gestione dei sistemi scolastici propone un modello a livelli.

Di descrivere il loro funzionamento attraverso un modello a livelli gerarchicamente connessi da legami deboli: il sistema nel suo complesso, la scuola, la classe o il gruppo di apprendimento e gli alunni. Ciascun livello possiede una certa libertà e autonomia, e può essere schematizzato tenendo in considerazione le condizioni preesistenti, come l'ambiente scolastico (livello scuola), le caratteristiche degli insegnanti (livello classe) e quelle degli studenti (livello alunni), i processi in atto, come – rispettivamente - la leadership, l'insegnamento e l'apprendimento, e i risultati specifici che ciascun livello produce. I livelli gerarchicamente superiori influiscono sul funzionamento di quelli inferiori ed esercitano forme di controllo. A seconda delle caratteristiche del sistema, in particolare del livello di decentralizzazione, si generano differenti forme di controllo e vi possono essere modalità di dialogo inverso, come ad esempio un feedback che dal basso produce cambiamenti nelle politiche di sistema. (Dell'Anna, 2021, p. 25)

Al variare del modello di qualità preso come riferimento, variano anche i modelli di misurazione e valutazione.

Tra i modelli di qualità propri dell'istruzione e non mutuati da altri settori, non si può non citare quello dell'*Educational Effectiveness Research* (EER) (Reynolds et alii, 2014), il risultato di riflessioni sui modelli e sulle sintesi empiriche (Kyriakides et alii, 2018; Kyriakides et alii, 2020). Il modello dell'EER, anche detto modello di Creemers, consente di riunire in uno schema formato da quattro livelli prospettive differenti prendendo in considerazione tutte le variabili che possono avere ricadute sugli apprendimenti degli studenti. Secondo gli autori, gli esiti degli studenti, ad esempio, possono essere influenzati da: caratteristiche individuali, qualità dei processi di insegnamento e aspetti organizzativi e strutturali del sistema educativo nel suo complesso.

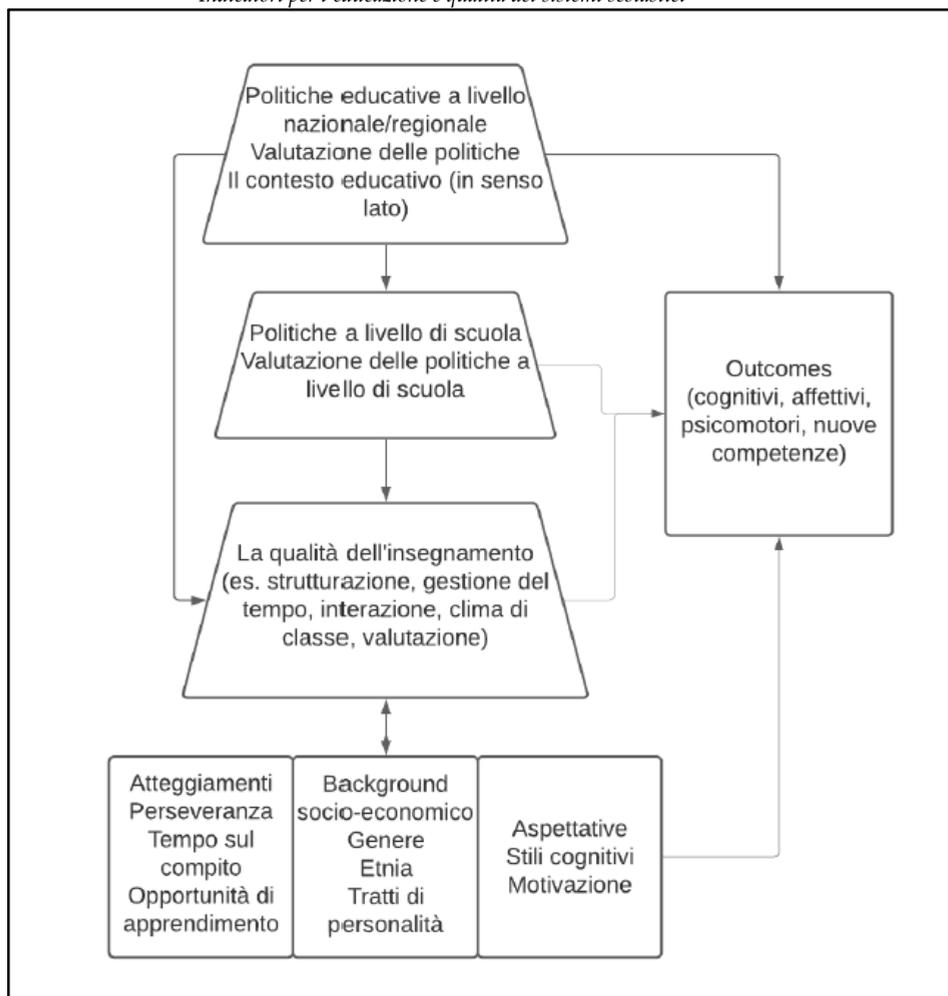


Figura 2. Modello dell' *Educational Effectiveness Research*. Fonte: Dell' Anna, 2021

Si tratta di un modello dinamico a fattori multidimensionali per la cui rilevazione è richiesta da una parte la misurazione della frequenza, dall'altra la descrizione più qualitativa del funzionamento del sistema di istruzione e formazione (Kyriakides et al., 2018; Kyriakides et al., 2020). Gli ultimi anni sono stati caratterizzati dallo sforzo di sviluppare nel settore dell'educazione un modello multilivello, condiviso dagli addetti ai lavori, sull'efficacia e sulla qualità dei sistemi di istruzione (Creemers & Kyriakides, 2010). Tuttavia, l'obiettivo è stato raggiunto a metà, infatti se da una parte la condivisione dell'EER è aumentata, dall'altra non

esiste ancora un modello di riferimento per la valutazione della qualità condiviso all'unisono (Scheerens, 2011b; Kyriakides et alii, 2018; Kyriakides et alii, 2020).

2.2. La proposta dell'OCSE per la valutazione della qualità dell'istruzione

Cosa si intende, dunque, per qualità dell'istruzione? Nella comunità scientifica la mancanza di una comune comprensione e un comune accordo su questo concetto sfocia in una mutevolezza tale da lasciare agli studiosi molteplici opzioni per l'analisi del funzionamento dei sistemi educativi (Dell'Anna, 2021).

In questa "liquidità" di riferimenti l'OCSE da circa tre decenni, mediante la pubblicazione annuale della raccolta degli indicatori per l'educazione, propone il suo punto di vista sulla questione. Nel caso specifico si tratta di un modello di valutazione della qualità dell'istruzione che si focalizza sul livello di sistema e abbraccia al suo interno una combinazione di aspetti strutturali, organizzativi e di risultati ottenuti che faranno da guida per il lavoro di ricerca che verrà presentato nei capitoli successivi.

I governi negli ultimi anni si sono sempre più interessati alle comparazioni internazionali delle opportunità e dei risultati educativi poichè è cresciuta sempre di più la volontà di costruire politiche per migliorare le prospettive sociali ed economiche degli individui, incentivare l'efficacia e l'efficienza nella scuola e mobilitare risorse per soddisfare le richieste sempre più complesse che giungono dalla società. Con l'intento di dare supporto ai decisori politici nel costruire sistemi educativi di qualità, l'OCSE ha accompagnato e sta accompagnando questi sforzi tramite il lavoro svolto sugli indicatori per l'educazione, quantitativi e internazionalmente comparabili, pubblicati annualmente nella raccolta intitolata *Education at a Glance* (OECD, 2021b).

L'avvincente cammino degli indicatori dell'istruzione dell'OCSE inizia nei primi mesi del 1973, quando viene pubblicato il report *Una struttura per indicatori dell'educazione che serva da guida per la decisione dei governi* (*A Framework for Educational Indicators to Guide Government*

Decisions) su cui aveva lavorato il Gruppo di Lavoro su Statistiche e Indicatori dell'Educazione. Il documento composto da 46 indicatori suddivisi in sei gruppi era stato approvato dal Comitato dell'Educazione dell'OCSE e riportava una breve descrizione del funzionamento dei sistemi educativi dei paesi che allora erano membri dell'organizzazione, puntando l'attenzione sulle ricadute dell'istruzione sul benessere dell'individuo e della società. I sei gruppi erano: contributo dell'educazione alla trasmissione del sapere; contributo dell'educazione al conseguimento dell'equità nelle opportunità; contributo dell'educazione nel far fronte alle necessità economiche; contributo dell'educazione allo sviluppo individuale; contributo dell'educazione alla trasmissione e allo sviluppo dei valori; utilizzo efficace delle risorse per il conseguimento degli obiettivi precedenti (Bottani & Tuijman, 1990; CERI-OCSE, 1994).

2.3. Il lento sviluppo degli indicatori per l'educazione

All'inizio degli anni Settanta l'interesse della comunità scientifica nei confronti di una possibile messa a sistema degli indicatori per l'educazione era elevato. Già nei due decenni precedenti, infatti, l'entusiasta dibattito aveva occupato le riflessioni degli studiosi che, volendo migliorare la società industriale moderna alla luce dei boom demografici ed economici di quegli anni, erano intenzionati a produrre dati scientifici sulla base dei quali prendere decisioni mirate. Nonostante il terreno fertile, il problema paradigmatico era concreto: ci si era resi conto che, pur volendo, il paradigma positivista non era applicabile a discipline perlopiù socio-pedagogiche che in quel momento non avevano raggiunto ancora una maturità tale da permetterlo. Così, tradite le promesse e le aspettative della comunità scientifica e compresa l'impossibilità di applicare il paradigma positivista alle scienze umane, non solo la ricerca sociale e pedagogica e la sua validità vennero messe in discussione, ma si registrò anche il tramonto dell'entusiasmo nei confronti degli indicatori dell'istruzione (Bottani & Tuijman, 1990). L'OCSE, avvertito il clima di opposizione e di sfiducia, in un primo momento ridimensionò il progetto, poi, venuto meno l'appoggio dei

decisioni politiche e degli istituti di statistica nazionali, smise di condurre rilevazioni e analisi quantitative per l'educazione. Il motivo era duplice: a quello epistemologico-positivista si aggiungeva anche quello geopolitico. Gli anni Settanta sono stati caratterizzati dalla più importante crisi economica e petrolifera dopo il 1929, di conseguenza i decisori politici non erano propensi a disporre di dati che potevano richiedere interventi educativi dispendiosi in un momento in cui le finanze dello stato erano destinate ad altro (Varsori, 2007). Il mancato collegamento tra le importanti informazioni sul funzionamento dei sistemi educativi di cui erano portavoce gli indicatori, utilizzabili tra l'altro per mettere a punto politiche educative informate dai dati, e i programmi e gli interessi politici non si arrestò (CERI-OCSE, 1994).

L'entusiasmo degli anni Cinquanta e Sessanta nei confronti degli indicatori dell'educazione non tornò prima del 1987. In quei quindici anni l'educazione in senso lato e le scuole nello specifico, non potendosi avvalere di metodi scientifici, non erano annoverate tra quelle istituzioni in grado di costruire una società giusta e democratica, tanto è vero che la stessa ricerca pedagogica aveva ripiegato sulle micro analisi limitate nel tempo e nello spazio e svolte in contesti ristretti, nelle singole classi e con determinati insegnanti. La prospettiva olistica dei sistemi di istruzione era lontana dalle scrivanie degli studiosi che producevano solamente documenti che arricchivano le informazioni di studenti, insegnanti e dirigenti scolastici locali. Secondo molti ogni contesto educativo era da considerarsi a sè stante, dunque era infruttuoso produrre indicatori o comparazioni intra e internazionali.

Sono le scuole a fare la differenza, non il sistema educativo in quanto tale; i risultati degli studenti dipendono da insegnanti e genitori, non dalle linee di condotta di chi ha il potere decisionale. (CERI-OCSE, 1994, p. 29)

Solamente nel 1987 ha ripreso vita il progetto di svolgere macro analisi sui sistemi educativi per combattere l'elevata frammentazione ed eterogeneità del sapere. L'occasione è stata fornita da un incontro tenutosi a Washington D.C. tra il Dipartimento dell'Educazione statunitense e l'OCSE in cui si è discusso di nuovi modelli di ricerca

educativa, di più ampio respiro rispetto alle micro analisi condotte in aula, e della possibilità di elaborare statistiche tra loro confrontabili. L'idea era quella di mettere a punto una serie molto ristretta di indicatori che fossero validi in tutti i paesi. In quel meeting il pedagogista T. Neville Postlethwaite per la prima volta ha sottolineato un aspetto che sarebbe poi tornato più volte nella storia delle comparazioni internazionali: confrontando una serie di dati provenienti dagli indicatori, il beneficio ricavato non era legato tanto alle affinità che potevano emergere dalla comparazione, quanto alle differenze che avrebbero spinto a nuove riflessioni e nuove domande di ricerca. Il dibattito in merito alla ricerca valutativa e alla costruzione di indicatori allora era tra l'individuazione di tratti simili o la legittimazione delle specificità di ciascun contesto. L'impegno preso a Washington D.C. è stato quello di considerare in particolare le differenze emergenti nel confronto tra indicatori in modo tale da collegare i dati macro-quantitativi con i riscontri qualitativi sul campo (Asquini, 2018).

Prima del 1987 la pubblicazione degli esiti delle indagini comparative internazionali condotte dall'IEA, *International Association for the Evaluation of Educational Achievement*, aveva riportato l'interesse sulle statistiche e sugli indicatori. Inoltre, nel 1983 la *United States National Commission on Excellence in Education* aveva redatto e pubblicato un preoccupante report sullo stato dell'educazione negli Stati Uniti intitolato *A Nation at Risk: the imperative for educational reform*⁴ (Gardner, 1983). Documento atipico, *Una Nazione a Rischio* ha rappresentato per molti versi un vero e proprio shock non solo per l'opinione pubblica americana, ma anche per il resto del mondo. Nell'indagine veniva messo chiaramente in luce quanto il sistema scolastico americano aveva un profondo bisogno di essere riformato. Le giovani generazioni non facevano mal sperare solamente nelle indagini IEA di quel periodo, ma anche in altri risultati allarmanti.

⁴ Per approfondimenti circa il documento *A Nation at Risk*: https://en.wikipedia.org/wiki/A_Nation_at_Risk (ultima consultazione 21 febbraio 2023)

- I punteggi medi bassi degli studenti americani nelle comparazioni internazionali sul profitto scolastico degli studenti. In 19 indagini internazionali alle quali gli Stati Uniti avevano a quell'epoca partecipato (dati anteriori al 1983), mai gli studenti americani si erano piazzati al secondo posto e nei confronti con gli altri paesi industrializzati erano finiti ultimi sette volte;
- La proporzione elevata di persone adulte praticamente "illetterate", ossia prive del bagaglio tecnico minimo necessario per capire o per redigere testi quotidiani semplici (circa 23 milioni di persone, dati del 1983);
- La proporzione dei diciassettenni che si trovano nelle stesse condizioni è del 17%. Questa proporzione sale al 40% tra i giovani delle minoranze etniche;
- La media dei punteggi conseguiti dagli studenti delle scuole secondarie nella maggioranza delle prove standardizzate è inferiore a quella di 26 anni prima, quando fu lanciato lo Sputnik;
- La discrepanza esistente tra i risultati conseguiti nei test dagli studenti ed i loro risultati scolastici;
- Il declino costante dei risultati conseguiti nei test alla fine della scuola secondaria superiore (le prove SAT per accedere all'università) che sono in calo costante tra il 1963 ed il 1980. Il punteggio medio nel test verbale è calato di quasi 50 punti mentre il punteggio medio nel test di matematica è diminuito di 40 punti;
- Il declino costante dei punteggi nei test di fisica ed inglese somministrati nelle università;
- Il declino drammatico della proporzione di studenti che conseguono punteggi elevati nei test SAT (per esempio gli studenti con un punteggio superiore a 650 punti);
- Il numero elevato di diciassettenni che non posseggono le competenze intellettuali complesse che ci si aspetterebbe da loro a quell'età. Quasi il 40% è incapace di effettuare inferenze da un testo scritto; solo un quinto riesce a scrivere un saggio in modo convincente; solo un terzo riesce a risolvere problemi di matematica che esigono una concatenazione di ragionamenti;
- La diminuzione continua dei punteggi conseguiti nei test di scienze tra i diciassettenni effettuati a livello nazionale nel 1969, 1973 e 1977. (Bottani, 2009, p. 4)

A seguito delle due Grandi Guerre e della crisi educativa che si era sviluppata nei primi anni Ottanta, l'America, sotto la guida del quarantesimo presidente degli Stati Uniti, Ronald Reagan, rischiava di non essere competitiva nè dal punto di vista dell'istruzione, nè dal

punto di vista lavorativo. Il rischio più grande era proprio quest'ultimo: non avere la forza lavoro necessaria per imporsi come potenza mondiale (Scotto Di Luzio, 2012; Lichtner, 2016).

La Commissione governativa americana addossava alla scuola la responsabilità del futuro del Paese e la scuola era additata come il capro espiatorio della crisi economica, morale e sociale della nazione. "A Nation at Risk" ha rappresentato una provocazione per la collettività scientifica che si occupa di scuola e d'istruzione, dapprima negli Stati Uniti e poi a livello planetario. È stato un gran tonfo che ha svegliato dal torpore i centri di ricerca accademici e quelli privati o para-pubblici, le organizzazioni internazionali, le associazioni scientifiche e professionali, mettendo in evidenza il divario profondo tra le aspettative del mondo politico da un lato e le finalità della comunità scientifica dall'altro. (Bottani, 2009, p. 5)

Molti ricercatori americani nel contrastare la tesi secondo la quale era colpa della scuola se gli Stati Uniti erano in crisi, tesi che per altro la scuola rigettava regolarmente (Berliner & Biddle, 1995), sostenevano la necessità più che impellente di innalzare la qualità dell'insegnamento e dell'apprendimento, ma anche la tesi secondo la quale il sistema scolastico rispecchia lo stato della società (Putnam 2004; Glass, 2008). Già dalle prime pagine del documento si leggeva che la società aveva perso di vista gli scopi fondamentali della scuola, le aspettative e gli sforzi metodici per conseguirli. Dalle parole dei 18 membri della commissione emergeva, inoltre, che solamente un quinto degli studenti era in grado di scrivere un saggio persuasivo e che un terzo di loro riusciva a risolvere un problema matematico che richiedeva più passaggi (Gardner, 1983). Tra chi accusava la scuola, chi era allarmato per i "non-risultati" raggiunti e chi sosteneva che *A Nation at Risk* era solamente una prova di indebolimento del sistema scolastico statunitense condotto dai neoliberali presenti nel ceto medio americano (Bracey, 2007), il quadro era chiaro: la situazione statunitense comparata con quella di altre nazioni industrializzate era grave tanto è vero che la commissione ha messo a punto trentotto raccomandazioni sugli standard e le aspettative, sui contenuti degli insegnamenti, sul tempo da dedicare alla scuola, sul supporto economico e sulla governance (CERI-OCSE, 1994).

Mentre all’inizio degli anni Ottanta in America i fattori analizzati fino ad ora hanno portato allo studio più approfondito delle variabili che influenzano l’apprendimento, nel resto del mondo, se inizialmente la crisi della qualità dei sistemi educativi è stata respinta, in un secondo momento ci si è resi conto che l’efficienza e l’efficacia dei sistemi scolastici erano in serio pericolo. I “complottoisti” statunitensi non erano riusciti a convincere del tutto gli altri paesi della teoria per la quale lo screditamento del sistema di istruzione americano era solo una mossa politica, infatti si era diffusa la volontà di giungere a una comune riflessione sul degrado della qualità della scuola pubblica (Weiss & Gruber, 1987).

La polemica divampò nella seconda parte degli anni Ottanta e sfociò nella sorprendente constatazione che mancavano prove documentate della validità di questi punti di vista contrapposti. Per evitare uno stucchevole dibattito dottrinario occorre prove ma queste non c’erano. Basti citare che i membri della Commissione americana sull’eccellenza scolastica, impegnati nella redazione del rapporto, non trovarono, all’epoca, nella banca dati federale sull’istruzione (*Common Core of Data*, CCD) i dati necessari per comprovare la tesi del degrado scolastico e costruire gli indicatori necessari per documentare le asserzioni formulate nel rapporto. [...] Una constatazione analoga avrebbe potuto essere fatta altrove e valeva anche per le banche dati internazionali sull’educazione come per esempio quella dell’UNESCO, oppure quella dell’OCSE. Nella massa di dati statistici sulla scuola raccolti e pubblicati negli annuari statistici non c’erano informazioni né sulla qualità dell’insegnamento, né sulle modalità d’insegnamento, né sulle competenze degli insegnanti né sui risultati degli studenti. Per altro pochissimi dati autorizzavano comparazioni attendibili tra sistemi scolastici diversi tra loro. Questa constatazione colse di sorpresa sia la comunità internazionale attiva nel campo dell’educazione sia i governi di molti Paesi costretti dalla crisi economica a controllare o a ridurre la spesa pubblica e quindi a gestire in modo oculato anche la spesa statale per il servizio scolastico. Per salvaguardare gli investimenti stanziati per l’istruzione occorre fornire prove convincenti della loro utilità, ossia occorre dati sui risultati, ma la maggioranza dei sistemi scolastici non ne possedeva e quando questi dati c’erano, spesso non era possibile collegarli tra loro perché i metodi di raccolta dei dati non usavano le stesse unità di analisi o le stesse definizioni. Le rivendicazioni di risorse a favore del sistema scolastico erano basate su assiomi

e postulati non dimostrati o su principi teorici, per non dire dogmi, non comprovati. (Bottani, 2009, p. 8)

Sul finire degli anni Ottanta questo contesto politico sfociò nella consapevolezza che le micro analisi dovevano essere affiancate da quelle macro, se si voleva dare validità e qualità ai sistemi educativi. Gli eventi storici avevano dimostrato che il successo dei sistemi scolastici ai livelli meso e macro è strettamente correlato con i modelli di monitoraggio e valutazione adoperati, le caratteristiche della direzione scolastica, le pratiche di consultazione degli attori coinvolti, la comunanza di interessi con la comunità, il tipo di scuola, la collaborazione scuola-famiglia, le pratiche di diversificazione scolastica, nonché l'assegnazione di risorse umane e finanziarie (CERI-OCSE, 1994). Senza la raccolta di dati al meso e al macro livello non si potevano avere linee guida comuni a tutti i paesi, infatti il successo di un sistema scolastico è il risultato dell'interazione di una serie di fattori in cui confluiscono sia il singolo studente e il singolo insegnante sia variabili di sfondo che accomunano l'intero sistema educativo.

2.4. Gli indicatori dell'educazione dell'OCSE

Gli anni Novanta si sono aperti con il consolidamento dell'idea secondo la quale l'educazione dovesse durare per l'intero arco di vita e dunque con l'urgenza di assegnare all'insegnamento e ai sistemi educativi nuove priorità (Field, 2000; Bottani, 2009). C'era una faglia molto netta tra le nuove richieste e le nuove urgenze della società e i sistemi scolastici costruiti e progettati in un'epoca del tutto diversa.

Quale senso dare a questi sistemi quando l'intera impalcatura del welfare state era messa in discussione? Si poteva contestare la legittimità del loro inserimento nelle responsabilità dello Stato o si doveva porre il problema della loro pertinenza come servizio pubblico, valutandone le prestazioni come si valutavano altri servizi nell'ambito delle politiche pubbliche? In altri termini, il settore scolastico era intoccabile oppure poteva essere passato al vaglio della critica fino ad immaginare l'impensabile, ossia la sua privatizzazione alla stessa

guisa di quanto si stava congetturando per il servizio postale, o quello ferroviario o quello sanitario? Con quali conseguenze? (Bottani, 2009, p. 8)

Il documento americano pubblicato qualche anno prima, *A Nation at Risk: the imperative for educational reform*⁵ (Gardner, 1983), aveva risvegliato l'attenzione dei decisori politici di tutto il mondo. A essere messo a nudo, infatti, non era stato solamente il sistema scolastico statunitense provato da una serie di malfunzionamenti e insuccessi, ma anche tutti gli altri. Ben presto ci si era resi conto che a livello globale i sistemi scolastici erano caratterizzati da questa faglia, da questo *miss-match* tra le richieste della società e l'offerta del mondo della scuola, che bisognava sanare. Per questo motivo tra la fine degli anni Ottanta e l'inizio degli anni Novanta, mentre i decisori politici si stavano interrogando su quali potessero essere le riforme migliori da varare (Bell, 1993), il progetto di riprendere a condurre ricerche a un livello macro trovava terreno fertile.

Da questa fase di confronti emerse, a livello internazionale, un ampio consenso su tre esigenze ritenute indispensabili per pilotare i sistemi scolastici:

1. l'adozione di metodi empirici di ricerca scientifica in grado di fornire prove attendibili sulla validità delle innovazioni e delle sperimentazioni;
2. la produzione di un insieme d'indicatori internazionali dei sistemi scolastici;
3. l'amplificazione e accelerazione delle valutazioni internazionali su vasta scala del profitto scolastico. (Bottani, 2009, p. 9)

Le tre esigenze fatte emergere dai decisori politici hanno trovato soddisfazione nei lavori che gli organismi sovranazionali hanno iniziato a partire dal decennio successivo. L'OCSE ha dato avvio al progetto sugli indicatori per l'educazione e ha iniziato a programmare l'indagine sui risultati negli apprendimenti dei quindicenni quali strumenti informativi (Asquini, 2018; OECD, 2019abc, 2020abc).

Nonostante le riluttanze di alcuni che non ritenevano affatto in crisi i sistemi scolastici, grazie al progetto INES, *International Indicators of*

⁵ Per approfondimenti circa il documento *A Nation at Risk*: https://en.wikipedia.org/wiki/A_Nation_at_Risk (ultima consultazione 21 febbraio 2023)

Education Systems, condotto da CERI, *Center for Educational Research and Innovation*, centro OCSE specializzato nella ricerca educativa e nelle innovazioni, infatti, alla fine degli anni Ottanta sono stati elaborati una serie di indicatori internazionali con l'obiettivo di farli diventare punto di riferimento per la qualità e per lo sviluppo dei sistemi scolastici (Bottani, 2009).

Nel 1992, dopo il lento inizio del 1973 e la conferenza a Washington D.C. del 1987, per la prima volta viene pubblicata la raccolta di indicatori intitolata *Education at a Glance* (OECD, 1992ab). Trentasei indicatori e tre aree di riferimento: contesti demografici, economici e sociali; costi, risorse e processi scolastici; esiti dell'istruzione. I costi e le risorse rappresentano l'ambito più approfondito, mentre il gruppo di indicatori sugli esiti dell'istruzione è a sua volta suddiviso in tre sottogruppi: il livello di istruzione raggiunto; i risultati di apprendimento; lo status post-scolastico (OECD, 1992ab; Asquini, 2018).

Così come era accaduto qualche anno prima alla Commissione americana sull'eccellenza, che non aveva trovato i dati per provare il degrado della scuola statunitense, così tra la fine degli anni Ottanta e l'inizio degli anni Novanta il mondo occidentale si era accorto di non avere modo di misurare lo stato di salute dei propri sistemi educativi. In questo panorama, gli indicatori del 1992 dovevano servire come base di partenza per prendere decisioni ragionate in merito a interventi educativi da implementare. Il rischio, tutt'altro che passato, era quello di mettere sul piatto della bilancia delle riforme educative ideologie politiche anziché dati raccolti scientificamente (Bottani, 1990).

La politica scolastica che non può avvalersi di una solida base di ricerca scientifica è sempre ricattata e ricattabile, impotente ad innovare, paralizzata di fronte a qualsiasi tentativo di cambiamento, condannata a salvaguardare i diritti acquisiti. Fu in questo contesto che fu presa la decisione di dotarsi sul piano internazionale di un insieme d'indicatori essenziali, tra loro comparabili, dello stato di salute dei sistemi scolastici. Moltissimi Paesi, Stati Uniti compresi, non sapevano, una ventina d'anni fa, quanto spendevano per la scuola e quale fosse il rendimento di tale investimento. La costituzione di un insieme d'indicatori della scuola non è un'operazione facile né dal punto di vista politico né da quello tecnico. Dal punto di vista politico si sapeva che dopo avere scelto e

costruito gli indicatori, gli attori all'interno dei vari sistemi scolastici avrebbero cercato di falsare, con varie motivazioni, i dati rilevati dagli indicatori.

Dal punto di vista tecnico, il problema era quello della comparabilità delle informazioni provenienti da sistemi scolastici tra loro molto diversi. Questa questione richiedeva l'adozione di una strategia per la raccolta di dati basata sulla trasparenza e la fiducia reciproca tra fornitori e produttori di statistiche e misure. (Bottani, 2009, p. 12)

Dopo la prima edizione di *Education at a Glance*, gli esperti si sono interrogati sui punti di forza e sulle criticità degli appena nati indicatori come strumento per monitorare i sistemi educativi (CERI-OCSE, 1994). Ogni indicatore, infatti, descrive l'andamento di un fenomeno e il comportamento di un sistema educativo, tuttavia, affinché ciò sia possibile, è necessaria un'interconnessione tra gli input, gli outputs e i processi educativi (Shavelson et alii, 1990a; Shavelson et alii, 1990b).

Lo sviluppo di una serie di indicatori internazionali dell'educazione non è semplicemente un esercizio tecnico programmato e controllato da esperti di statistica, ma è innanzitutto e soprattutto un impegno politico. Un indicatore non è una semplice espressione numerica o un dato statistico composito. Con esso si intende comunicare qualcosa circa il rendimento o il comportamento di un sistema educativo e può essere usato per informare o ispirare quanti vi risultano impegnati: amministratori, responsabili, insegnanti, studenti, genitori e il pubblico in genere. Di estrema importanza è che gli indicatori forniscono anche la base per creare nuove prospettive ed attese. Senza dati affidabili sui risultati del sistema educativo, una serie di indicatori non solo è incompleta, ma, ciò che è peggio, ha una limitata applicabilità. Lo scopo degli indicatori dell'educazione è quello di sollecitare un dibattito politico, di aiutare il processo decisionale e di ispirare l'azione politica. Ovviamente, un tale scopo non può essere raggiunto se si hanno a disposizione soltanto dati sulle risorse impiegate e sul flusso degli studenti. In una serie di indicatori debbono essere incluse anche le variabili dei risultati, perché queste offrono un modo di valutare in quale misura sistemi educativi diversi raggiungono i loro obiettivi. Tali misurazioni del rendimento scolastico non dovrebbero essere limitate ai dati sui risultati ottenuti dagli studenti, perché tutti i paesi si propongono in campo educativo obiettivi che vanno oltre quelli riguardanti l'apprendimento di ordine cognitivo. (Bottani & Tuijnman, 1994, pp. 33-34)

Tenendo conto del fatto che ogni processo educativo si situa in un determinato contesto e che in quest'ultimo gli vengono attribuiti dei significati piuttosto che altri, è emersa subito dopo il 1992 una mancanza di omogeneità nella definizione degli obiettivi dei processi educativi, nello specifico di quelli relativi all'insegnamento e all'apprendimento. Infatti, se la comparazione degli input e degli output sembrava verosimile, quella dei processi educativi era tutto fuorchè semplice (Shavelson et alii, 1990a; Shavelson et alii, 1990b).

La difficoltà di ottenere dati da autorità nazionali e di riferirli in una struttura comune che li renda confrontabili è enorme, ma non impossibile. Di nuovo, i due principali ostacoli sono un'insufficiente teoria e una conoscenza inadeguata circa l'approccio comparativo in pedagogia: confrontare significa raggruppare dati secondo una certa classificazione. Così, la conoscenza delle caratteristiche distintive dei sistemi e dei processi in esame è il cuore stesso del metodo comparativo. [...] L'approccio comparativo è un modo di descrivere la realtà o di rappresentare percezioni e osservazioni. [...] Qual è il valore cognitivo intrinseco del metodo comparativo in pedagogia e, conseguentemente, quali sono le implicazioni per gli strumenti da usare in modo che ne risulti una conoscenza comparata affidabile e valida? Il corpo di conoscenze e intuizioni che implicano una metodologia comparativa è altamente frammentato e sparso attraverso molte discipline. Esso ha bisogno di essere organizzato e sintetizzato, in modo tale che possa essere risolta la controversia circa i modelli in pedagogia comparata. (Bottani & Tuijnman, 1994, p. 36)

La comparabilità dei dati e la condivisione della nomenclatura è riportata anche da Bottani (2009).

La comparabilità delle informazioni sulla scuola di vari sistemi scolastici è stata un'altra rivendicazione politica espressa in primo luogo da alcuni paesi che volevano conoscere comparativamente quale fosse la qualità dei loro sistemi scolastici (per esempio il livello d'investimenti nella scuola correlato ai risultati ottenuti) e poi dalle maggiori organizzazioni internazionali intergovernative come l'UNESCO, la Banca Mondiale, l'OCSE che avevano bisogno di dati comparabili attendibili sia per informare i governi sullo stato

dell'istruzione sia per pianificare e finanziare progetti di riforma scolastica nei paesi in via di sviluppo. Orbene, per produrre dati comparabili si deve disporre di una nomenclatura comune dei programmi e dei tipi d'istruzione e formazione, basata su definizioni accurate e comprensibili da tutti. Per avere un'idea delle difficoltà da superare per produrre uno strumento del genere, basti qui accennare al fatto che non esiste ancora oggi una definizione condivisa da tutti i Paesi del concetto di scuola. Quindi quando si distribuiscono ai dirigenti scolastici questionari internazionali sull'organizzazione e il funzionamento delle scuole, si può essere certi che in ritorno non si avranno risposte tra loro comparabili perché i dirigenti delle scuole dei vari sistemi scolastici non condividono una stessa definizione di scuola. Uno dei meriti della produzione di un insieme d'indicatori comparabili a livello internazionale è stato quindi quello di avere stimolato la ripresa dei lavori sulla nomenclatura dei programmi d'insegnamento gestita dall'UNESCO. Senza una classificazione coerente, accettata da tutti, non si possono infatti produrre statistiche dell'istruzione comparabili tra Paesi e men che meno si possono realizzare insieme d'indicatori internazionali dei sistemi scolastici. L'uso di una tale classificazione è indispensabile per la raccolta di dati comparabili. L'adozione di una terminologia comune non impone ai Paesi di modificare la propria, ma li obbliga a completare il proprio sistema statistico qualora manchino i dati statistici relativi ad alcuni settori, e costringe inoltre a produrre tavole sinottiche che comparano la terminologia nazionale con quella internazionale. (Bottani, 2009, p. 14)

Il paese che si è dotato per primo di una serie di indicatori sul sistema educativo è la Francia. Dal 1991 in Francia prima della discussione della legge finanziaria in parlamento vengono elaborati 30 indicatori sul sistema scolastico, lo scopo è duplice: catalizzare l'attenzione sul dibattito riguardante la scuola e informare i parlamentari (Direction de l'évaluation et de la prospective, 1991). Meuret (2001), fautore di questa raccolta, sostiene che anche lo scarso successo degli indicatori francesi è da imputare alla settorialità del dibattito molto concentrato nelle mani dei professionisti e all'assenza nella raccolta degli indicatori delle preoccupazioni messe in luce da parte di chi la scuola la vive quotidianamente. Tuttavia, nonostante ciò, sulla scia dell'esperienza francese, anche altri paesi europei, tra cui l'Italia e la Germania, hanno

iniziato a pubblicare con maggiore o minore diffusione raccolte di indicatori nazionali (Bottani, 2009).

Senza dubbio il sistema di indicatori messo a punto dall'OCSE pone in risalto alcuni elementi del sistema educativo, lasciandone in ombra altri. Infatti, nonostante le iniziali discussioni tecniche ed epistemologiche e nonostante le problematiche legate al delicato rapporto tra la ricerca e la politica, alla comparabilità, alla raccolta e alla trasmissione dei dati e all'organizzazione sistematica di questi ultimi (Oakes, 1986; Bottani, 1990, 2009), dal 1992 ad oggi la serie *Education at a Glance* è stata pubblicata annualmente.

2.5. *Education at a Glance* 2022

Sono passati trenta anni dalla prima raccolta degli indicatori del 1992. Nell'edizione 2022, rilasciata nel mese di ottobre, i gruppi di indicatori sono quattro e non tre come nella prima serie dei trentasei indicatori. Come accade ormai da qualche anno a questa parte gli ambiti in cui si suddividono i 27 indicatori sono (OECD, 2022):

- gli outputs educativi e l'impatto dell'apprendimento;
- l'accesso all'educazione, la partecipazione e il progresso;
- le risorse finanziarie investite in educazione;
- gli ambienti di insegnamento e apprendimento e

l'organizzazione delle scuole.

La più recente raccolta di *Education at a Glance* non solo rappresenta l'edizione del trentennale, ma è anche la sintesi del funzionamento dei sistemi scolastici dopo due anni di pandemia.

Il 2022, rispetto al 2021, ha visto l'allentamento delle restrizioni sanitarie dovute al momento emergenziale in quasi tutti i paesi. L'educazione mondiale, tuttavia, ha risentito degli effetti negativi della persistente chiusura delle scuole o dell'assenza degli insegnanti e degli studenti per motivi di salute. Ciò ha interrotto i processi di insegnamento e apprendimento, ma anche quelli dovuti alla digitalizzazione attuati durante il primo periodo di didattica a distanza e in uso, in alcune realtà, ancora ora. Dal primo periodo ad oggi, infatti, si

è passati dalla gestione della crisi alla valutazione dell'impatto della pandemia.

Quasi tutti i paesi OCSE hanno implementato valutazioni standardizzate per quantificare le perdite negli apprendimenti a vari livelli di istruzione. La maggior parte dei paesi ha anche fornito un ulteriore sostegno agli studenti per alleviare gli effetti della pandemia. A livello primario e secondario, circa l'80% dei paesi, di cui sono disponibili i dati, ha implementato tali programmi di recupero. A livello pre-primario, questi programmi sono meno comuni, ma sono messi a disposizione in 19 dei 28 paesi con dati disponibili. Un ulteriore supporto psicologico e socio-emotivo per gli studenti della scuola primaria e secondaria è stato reso disponibile in 19 paesi su 29. (OECD, 2022, p. 21, traduzione a cura dell'autrice)

La pandemia ha richiesto ai paesi di trovare soluzioni alternative al tradizionale modo di fare scuola e, sebbene in un primo momento queste soluzioni siano state messe a punto come transitorie e di emergenza, a più di due anni dal primo caso Covid-19 registrato è innegabile aprire una riflessione in merito poichè molte di queste si sono rivelate preziose anche dopo i primi periodi di emergenza e hanno contribuito a facilitare in molti paesi il passaggio agli strumenti digitali nell'istruzione (Izzo & Ciurnelli, 2020; Lucisano, 2020; Batini et alii, 2020; Capperucci, 2020; Di Nunzio et alii, 2021; Castellana & Rossi, 2021; Mattarelli, 2021ab; OECD, 2022).

Le piattaforme digitali dovrebbero offrire risorse efficaci ed eque per l'apprendimento all'interno e all'esterno della scuola. Molti paesi riconoscono che, se ben realizzate, queste ultime facilitano l'accesso all'istruzione e riducono le disuguaglianze nei risultati degli apprendimenti. Per questi motivi, 17 paesi su 27 con dati disponibili prevedono di continuare a utilizzare prevalentemente gli strumenti digitali nell'istruzione secondaria inferiore, proseguendo al contempo con il fornire una migliore offerta di apprendimento a distanza o mista a tutti i livelli di istruzione, dalla scuola primaria alla scuola secondaria superiore. Di conseguenza, 13 paesi su 25 prevedono di aumentare la formazione digitale pre-servizio per gli insegnanti delle scuole secondarie inferiori a livello nazionale. Nella maggior parte dei paesi, la formazione digitale sarà fornita anche agli insegnanti già in servizio nell'istruzione

secondaria inferiore. Sebbene la maggior parte dei paesi intenda continuare a utilizzare gli strumenti digitali, le modifiche al quadro normativo o istituzionale, che disciplina l'educazione digitale, non sono state diffuse. Il 54% dei paesi con dati disponibili non ha apportato modifiche né al quadro normativo né a quello istituzionale sulla digitalizzazione e non ha intenzione di farlo. Infatti, anche se alcuni paesi hanno introdotto modifiche al proprio quadro normativo o istituzionale durante la pandemia, solo quattro di questi hanno in programma ulteriori revisioni: Italia, Lituania, Lussemburgo e Slovacchia. (OECD, 2022, p. 30, traduzione a cura dell'autrice)

Dall'istruzione pre-primaria all'istruzione terziaria la pandemia ha dimostrato il valore delle tecnologie nei processi di insegnamento e apprendimento. Molti paesi, per facilitare l'utilizzo di questi strumenti e supportare l'innovazione didattica, hanno riformato i propri quadri normativi e istituzionali e hanno acquistato nuove risorse e formato gli insegnanti al loro utilizzo.

Per beneficiare appieno della digitalizzazione, dobbiamo rafforzare la cultura dell'innovazione nell'istruzione. Ciò richiede i giusti quadri istituzionali e normativi, in particolare quelli che disciplinano l'educazione digitale. Richiede che gli appalti pubblici nel settore dell'istruzione diventino più sensibili alle opportunità digitali e creino maggiori incentivi per l'innovazione nel settore privato. Richiede, inoltre, agli insegnanti di acquisire le competenze necessarie non solo per utilizzare gli strumenti digitali in classe, ma anche per migliorare il proprio sviluppo professionale. (OECD, 2022, p. 9-10, traduzione a cura dell'autrice)

Oltre ai dati riguardanti la pandemia, l'edizione 2022 di *Education at a Glance* evidenzia anche altri aspetti propri dei sistemi educativi come la partecipazione all'educazione pre-primaria, la partecipazione alla professione dell'insegnante e l'accesso all'istruzione terziaria.

L'educazione pre-primaria ha acquisito sempre più importanza negli ultimi anni tanto è vero che nei paesi OCSE l'83% dei bambini tra i 3 e i 5 anni frequenta la scuola dell'infanzia, mentre il 4% frequenta già la scuola primaria. Quest'ultima, infatti, se di qualità, è un indicatore fondamentale di inclusione ed equità che dà agli studenti di qualsiasi

provenienza la possibilità di avere successo nella vita. Forte delle politiche comunitarie e dei benchmark educativi da raggiungere entro il 2020 (Allulli, 2010; Eurydice, 2019), la partecipazione all'educazione pre-primaria tra il 2005 e il 2020 è aumentata in media di 8 punti percentuali, sebbene i bambini al di sotto di 3 anni molto spesso vengano assistiti in casa o in istituzioni che non sono classificate ufficialmente come della prima infanzia: in media OCSE solamente il 27% dei bambini tra 0 e 3 anni sono inseriti in programmi educativi loro dedicati (OECD, 2022).

Negli ultimi anni è stato rilevato, inoltre, che gli insegnanti oltre al tempo dedicato all'insegnamento vero e proprio, trascorrono molte ore a preparare le lezioni e le valutazioni periodiche degli studenti. La situazione, ciononostante, è disomogenea: in alcuni paesi ci si aspetta che i docenti dedichino meno di un terzo del loro orario di lavoro all'insegnamento vero e proprio, in altri due terzi. Le linee guida internazionali stabiliscono che gli insegnanti devono dedicare in media 1000 ore all'anno all'insegnamento nella scuola dell'infanzia, 800 ore nella scuola primaria, 700 ore nella scuola secondaria. Non sempre questa ripartizione è rispettata e le oscillazioni tra i paesi sono ampie: per esempio, per quanto riguarda la scuola superiore, si passa da 1248 ore di insegnamento annue in Costa Rica a 483 ore in Polonia (OECD, 2019d, 2020e).

Gli ultimi venti anni hanno portato un profondo cambiamento nel mondo dell'istruzione e conseguentemente nel mondo del lavoro. La percentuale di giovani adulti che hanno un titolo di studio avanzato è aumentata considerevolmente nei paesi OCSE: se nel 2000 solamente il 27% dei giovani tra 24 e 34 anni aveva una laurea, nel 2021 il numero è raddoppiato fino ad arrivare al 48%. Ad oggi si è convinti che l'istruzione può avere profonde ricadute sulla vita delle persone e può prevenire i rischi economici e legati alla salute. Infatti, negli ultimi anni il tasso di disoccupazione è aumentato molto più significativamente per coloro che hanno il titolo di studio secondario di primo grado, rispetto a coloro che hanno un titolo terziario (Asquini, 2018; Proietti, 2020). La crescente richiesta di persone formate e con competenze avanzate da parte del mondo del lavoro, inoltre, ha ricadute non indifferenti sul futuro dell'istruzione e delle società nel loro complesso: tra gli adulti,

coloro che hanno un titolo di studio più elevato sono in grado di utilizzare più facilmente le tecnologie. C'è un numero crescente di studenti con background differenti che vuole accedere all'istruzione terziaria, questo si traduce in un adattamento di quest'ultima per andare incontro alle esigenze di tutti. Le università, infatti, devono essere preparate ad accogliere i giovani che vogliono formarsi e acquisire nuove competenze poichè i datori di lavoro, dal canto loro, si aspettano che tra qualche anno la laurea diventi il titolo di studio con il quale ci si presenta a chiedere un impiego. Nonostante i laureati siano avvantaggiati rispetto a coloro che hanno un titolo di studio inferiore e nonostante la spesa pubblica per studente sia più alta a livello terziario rispetto ad altri livelli di istruzione, la raccolta di indicatori dell'OCSE del 2022 sottolinea che molti studenti universitari non completano il proprio percorso di studio. Infatti, solamente il 39% in media OCSE degli studenti che frequentano un corso di laurea triennale non finiscono fuori corso. Dopo tre anni il tasso di completamento aumenta in tutti i paesi, ma in media non va oltre il 68%. Se si analizza la questione di genere, inoltre, gli uomini rispetto alle donne hanno meno probabilità di laurearsi nei tempi previsti legalmente dal corso di studi (OECD, 2022).

Dalla lettura degli indicatori per l'educazione emergono tre aspetti che sarebbe utile sostenere nei prossimi anni: i primi due riguardano i giovani, il terzo i datori di lavoro. Ai primi bisogna assicurare le condizioni per poter terminare il percorso di formazione terziaria e di avere la possibilità di trovare un lavoro senza doversi scontrare con un'offerta inadeguata, per i secondi bisogna fare in modo che incontrino possibili lavoratori con le competenze che necessitano per esercitare la professione. Per fare questo è bene che si inizi a pensare non solamente a un'istruzione terziaria adeguata, ma anche a sistemi educativi pronti ad accogliere l'eterogeneità, che siano efficaci ed efficienti e di qualità (OECD, 2022). Affinchè ciò si realizzi, occorre iniziare a mettere a punto strategie globali di valutazione dei sistemi educativi che tengano conto dell'istruzione nella prima infanzia, tanto quanto della formazione dei docenti o della possibilità di accedere e partecipare all'istruzione terziaria, senza tralasciare la valutazione del funzionamento dei sistemi di istruzione e formazione o i livelli degli apprendimenti raggiunti dagli

studenti che verranno trattati nel prossimo capitolo (OECD, 2019abcd, 2022).

Capitolo terzo

Le indagini internazionali sugli apprendimenti

3.1. Verso le indagini comparative internazionali

Parallelamente all'interesse per gli indicatori, negli anni Cinquanta del secolo scorso la comunità scientifica iniziò ad interessarsi anche alle indagini comparative su larga scala volte a misurare il livello degli apprendimenti degli studenti (Vertecchi et alii, 2010). L'intento era quello di misurare le competenze di una numerosa popolazione di studenti, ma anche di analizzare la correlazione tra i risultati raggiunti da questi ultimi e l'organizzazione del sistema scolastico nel suo complesso.

L'idea di valutare sistemi scolastici diversi con uno stesso strumento per cercare di capire se determinati modelli d'organizzazione e gestione della scuola o determinati curricula d'insegnamento dessero risultati migliori di altri fu discussa per la prima volta nel 1958 in seno alla riunione annuale dei direttori dei principali istituti di ricerca pedagogici dell'epoca, organizzata ad Amburgo dall'Istituto Internazionale dell'UNESCO per l'educazione. All'ordine del giorno fu inserito il tema della valutazione dei sistemi d'insegnamento per verificare innanzitutto se dal punto di vista pratico un'indagine del genere avesse senso, e per appurare poi se ci fosse un numero sufficiente di sistemi scolastici interessati ad effettuarla. (Bottani, 2009, p. 15)

I partecipanti ad Amburgo decisero che si poteva tentare di sperimentare un'indagine su larga scala così tra il 1959 e il 1962 si svolse il primo studio comparativo sulle competenze degli studenti. L'Italia non era tra i paesi partecipanti, tuttavia, con strumenti di misurazione molto diversi da quelli attualmente utilizzati dall'OCSE e dall'IEA vennero raggiunti 12.000 studenti di dodici paesi con un'età media di tredici anni. Il focus dei test cognitivi era sulle competenze in lettura, matematica, scienze, geografia e intelligenza non verbale, aspetto che non sarebbe stato mai più testato nelle indagini successive (Bottani, 2009).

Nonostante gli errori commessi, i direttori di ricerca poterono constatare che un'indagine comparata su vasta scala in diversi sistemi scolastici era fattibile. I primi risultati, per quanto grezzi e imprecisi, fornirono informazioni mai avute in precedenza sulle relazioni tra punteggi conseguiti dagli studenti nelle prove strutturate e svariati aspetti dei sistemi scolastici. (Bottani, 2009, p. 16)

Sebbene ci fossero da mettere a punto ancora alcuni aspetti tecnici e metodologici, l'indagine nata ad Amburgo e conclusa nel 1962 diede avvio alle indagini comparative internazionali

3.2. La nascita dell'IEA

L'IEA, *International Association for the Evaluation of Educational Achievement*, nasce contestualmente alla riunione annuale di Amburgo. Infatti, riconosciuta ufficialmente come un'entità legale nel 1967, già nel 1960 muove i suoi primi passi nel panorama delle indagini comparative internazionali. Al primo studio sulla fattibilità di un disegno di ricerca così ampio, in effetti, avevano partecipato i sociologi, gli psicometristi e gli psicologi che avrebbero fatto poi parte del primo team di ricerca IEA: nel misurare prima e valutare poi l'efficacia e l'efficienza dei sistemi scolastici, questi ultimi furono fin da subito concordi nell'affermare che bisognava mettere in relazione gli input con gli output (es. apprendimenti, partecipazione e attitudini) propri dei sistemi educativi.

I fondatori dell'IEA vedevano il mondo come un laboratorio educativo naturale, in cui diversi sistemi scolastici sperimentano modi diversi per ottenere risultati ottimali dall'educazione dei loro giovani. Hanno presupposto che se la ricerca potesse ottenere prove da un'ampia gamma di sistemi, la variabilità sarebbe sufficiente per rivelare relazioni importanti che altrimenti sfuggirebbero alla misurazione all'interno di un singolo sistema educativo. Hanno respinto con forza le affermazioni prive di dati sui meriti relativi ai vari sistemi educativi e hanno mirato a identificare i fattori che hanno influenze significative e coerenti sui risultati scolastici⁶. (traduzione a cura dell'autrice)

⁶ Per approfondimenti circa la storia dell'IEA: <https://www.iea.nl/about/org/history> (ultima consultazione 21 febbraio 2023)

Era chiaro che il mondo fosse un *laboratorio educativo* eterogeneo e che le strategie messe in atto da ciascun paese per garantire l'educazione delle giovani generazioni fossero diverse le une dalle altre. Andando oltre il dibattito degli anni precedenti sulla politica e sulla metodologia della ricerca educativo-valutativa, negli anni in cui si afferma l'IEA come agenzia per la misurazione delle competenze degli studenti, ci si rende conto che tramite la ricerca pedagogica si possono raccogliere informazioni sui sistemi educativi e che le differenze tra i singoli sistemi educativi non possono che emergere in quest'ottica comparativa. Lo scopo, infatti, è quello di intercettare le variabili di sistema, organizzative e politiche che influiscono sulla qualità, sull'efficienza e sull'efficacia dei sistemi educativi (Lehmann, 2011; OECD, 2013; Kyriakides et al, 2014; Sundberg, 2019).

Il loro interesse preminente sta nella grande quantità di variabili raccolte su notevoli campioni rappresentativi di allievi a vari livelli di età o scolarizzazione, e quindi sulla possibilità di studiarne le intercorrelazioni sia su scala internazionale, sia per singoli paesi ed anche per fasce o regioni degli stessi, in modo da poter formulare ipotesi empiricamente fondate circa l'influenza di vari fattori (economici, culturali, organizzativi, didattici) sul profitto scolastico e sugli atteggiamenti culturali. (Visalberghi, 1977a, p. 125)

Dopo lo Studio Pilota su Dodici Paesi del 1959-1962, la prima indagine sperimentale dell'IEA è FIMS, *First International Mathematics Study*, volta a misurare le conoscenze matematiche dei tredicenni. La ricerca ha luogo tra il 1962 e il 1965 ed è ancora oggi di particolare importanza poichè ha permesso di rivedere l'impianto tecnico e metodologico e di porre le basi per gli studi successivi. Grazie alla partecipazione di dodici paesi è stato possibile trovare un predittore statisticamente significativo che mette in relazione il rendimento scolastico con il successo. Lo studio FIMS, infatti, ha messo in luce che a partire dalle indicazioni contenute nei curricula è possibile affrontare in aula una determinata tematica con approcci didattici diversi. Questi ultimi, a loro volta, possono essere più o meno adeguati per i singoli studenti a seconda degli stili cognitivi. L'equità, dunque, accompagnata

dall'inclusione e intesa come la capacità di prendere in considerazione gli stili cognitivi di ciascuno, si è rivelata essere una delle chiavi per il successo dei sistemi educativi (McKnight, 1987; Brown, 1996; Bottani, 2009).

La pubblicazione dei risultati ottenuti da ricerche internazionali su vasta scala condotte dall'*International Association for the Evaluation of Educational Achievement* (IEA) e dall'*International Assessment of Educational Progress* (IEAP) sul profitto degli studenti in materie quali la matematica e le scienze, ha richiamato l'attenzione sia del pubblico che delle persone investite di poteri decisionali sui risultati della scuola. La valutazione del rendimento offerto dagli studenti e dalla scuola in genere, il controllo del funzionamento dei sistemi educativi, la conduzione della politica scolastica e il miglioramento dell'amministrazione di risorse sono diventati i temi centrali nel dibattito politico. Parallelamente a un'accresciuta attenzione ai risultati dell'educazione, si è avuto da parte dei politici uno spostamento di enfasi dai problemi connessi con la crescita quantitativa del sistema al miglioramento qualitativo della scuola. Conseguentemente, i risultati ottenuti dagli studenti sono divenuti un criterio chiave per giudicare la qualità e l'efficacia del sistema educativo. (Bottani & Tuijnman, 1994, p. 30)

In quegli anni si stava pian piano avviando una revisione sostanziale delle indagini comparative fino ad allora condotte.

Queste indagini si inseriscono nel filone di studi della pedagogia comparata, ma la arricchiscono di elementi propri della pedagogia sperimentale: se fino agli anni Cinquanta le comparazioni fra paesi si svolgevano esclusivamente su dati osservativi, forniti esclusivamente dai paesi stessi, IEA intende realizzare studi originali rilevando dati su alcuni gruppi di variabili per tutti i paesi che decidono di partecipare alle sue indagini. In tal modo si pongono necessariamente problemi squisitamente sperimentali, quali la definizione del modello di indagine, la costruzione di strumenti di rilevazione, la definizione di un campione di ricerca, le procedure di somministrazione dei materiali, le modalità di analisi dei dati. (Asquini, 2007, p. 58)

L'ipotesi iniziale avanzata dai ricercatori IEA ruotava attorno alla possibilità che il successo scolastico potesse dipendere da alcune

variabili ricorrenti nei sistemi educativi. Infatti, tra gli strumenti utilizzati tra gli anni Cinquanta e Sessanta c'erano le prove di profitto costruite con domande a risposta chiusa per misurare come i risultati degli studenti potessero essere influenzati non solo dalle loro naturali propensioni, ma anche da variabili di background che poco avevano a che fare con la scuola (Crosswhite, 1985; Martin, 1996; Ravecca, 2009; Maccario, 2012)

La caratteristica principale delle indagini IEA inizialmente era riconducibile al metodo attraverso il quale veniva costruito il framework di riferimento per ciascuno studio, compreso l'oggetto della misurazione. Rispetto alle indagini OCSE che sarebbero arrivate da lì a qualche anno, l'IEA si proiettò nel dibattito sulle indagini comparative sugli apprendimenti con la volontà di confrontare il profitto scolastico degli studenti: per fare ciò, c'era bisogno di misurare le conoscenze e dunque di costruire item intorno ai curricula che effettivamente erano stati affrontati nelle classi di tutti i paesi partecipanti (Bianchini, 2010). Questo comportava non solo uno studio particolarmente approfondito da parte dei team di ricerca chiamati a confrontare per ciascun paese il curriculum istituzionalmente definito con quello effettivamente svolto dal docente, percepito dallo studente e raggiunto, ma anche un cospicuo investimento di tempo per effettuare degli studi di caso su ogni singola realtà e, di conseguenza, per costruire le prove di profitto (Mullis & Martin, 2015, 2019).

Per un quarto di secolo, fin verso il 1990, l'IEA ha occupato una posizione di monopolio nell'organizzazione delle indagini comparate su vasta scala. La crisi della qualità della scuola scoppiata nel 1983 e l'insoddisfazione dei responsabili politici di molti sistemi scolastici sugli esiti della scolarizzazione nei loro Paesi furono fatali alla preminenza dell'IEA. Non si possono infatti moltiplicare le indagini, continuare a batter cassa, senza produrre rapidamente risultati e informazioni utilizzabili tempestivamente dai dirigenti dei sistemi scolastici per scegliere le riforme scolastiche da effettuare, per pilotarle o correggerne i difetti, in funzione dei risultati prodotti. Lo svolgimento di quelle numerose indagini rese però possibile un perfezionamento costante della metodologia della valutazione e l'accumulazione di un'importante esperienza sulle modalità d'impostazione e realizzazione delle indagini, di preparazione dei test e di

trattamento dei risultati. La diffusione dei risultati al di fuori della cerchia degli addetti ai lavori fu però alquanto limitata. Le conclusioni non arrivavano laddove avrebbero dovuto giungere, ossia ai responsabili dei sistemi scolastici, oppure vi arrivavano troppo tardi, perché i risultati venivano pubblicati con grande ritardo, quando avevano solo un interesse scientifico per gli specialisti dei sistemi scolastici o dell'educazione comparata ma non ne avevano più per i responsabili politici ai quali importava conoscere a breve termine gli effetti delle decisioni da loro prese. (Bottani, 2009, p. 16)

Sebbene abbia fatto da apripista per i ricercatori di tutto il mondo, nell'arco di qualche anno il "modello IEA" iniziò ad essere contestato e da un'iniziale pleora di indagini sul profitto degli studenti si passò alla sistematizzazione di alcuni di quegli studi che ancora oggi vengono ripetuti con regolarità (Walker 1976; Purves, 1987ab; Purves, 1992; Cumming, 1996; Asquini, 2007).

3.3. Le indagini TIMSS e PIRLS dell'IEA

Tra le più importanti indagini IEA vi sono TIMSS, *Trends in International Mathematics and Science Study*, e PIRLS, *Progress in International Reading Literacy Study*. All'inizio degli anni 2000, dopo l'avvento degli studi sulle competenze e sugli apprendimenti dell'OCSE, nel board dell'IEA si comprese che il "modello OCSE", che prevedeva la ripetizione di un'indagine all'interno di un programma di confronto longitudinale, era il modello più perseguibile in una società che stava mutando velocemente e che richiedeva competenze sempre più complesse. Rispetto all'OCSE, di cui si scriverà nei paragrafi successivi, l'IEA non disponeva di moltissimi fondi economici, ciononostante, la prima edizione dell'indagine PISA dell'OCSE e il terzo studio sulla matematica e le scienze dell'IEA delinearono un grande cambiamento nella valutazione su larga scala degli apprendimenti (Bottani, 2009; Asquini, 2007; Callegari & Gaudio, 2018).

Il primo studio IEA messo a sistema è stato *Trends in International Mathematics and Science Study*, TIMSS. Dopo il *First International Mathematics Study*, FIMS, degli anni 1959-1962, e il *Second International Mathematics Study*, SIMS, degli anni 1976-1989, rivolti agli studenti

tredicenni e a coloro che erano in procinto di intraprendere la carriera universitaria, il terzo studio sulla matematica fu svolto nuovamente negli anni 1994-1995 con il *Third International Mathematics Study* che di lì a poco sarebbe stato convertito in *Trends in International Mathematics and Science Study*, TIMSS (Garden, 1987; Härnqvist, 1987; Robitaille, 1990; Martin, 1996; Brown, 1996; Martin & Kelly, 1996; Beaton, 1996ab; Mullis, 1997; Mullis et al., 1998; Beaton & Robitaille, 1999; Battista, 1999; Hanna, 2000; Oldham, 2003; House, 2006; Mullis & Martin, 2007). La versione attuale dello studio delle tendenze in matematica e scienze dell'IEA, riproposta dal 1999 con ciclicità quadriennale, ha visto negli anni un numero sempre crescente di paesi partecipanti, arrivando all'ultima edizione disponibile, quella del 2019, con più di 60 nazioni. Gli studenti coinvolti in questo studio sono quelli del quarto anno di scolarità (9 anni circa) e quelli dell'ottavo anno di scolarità (13 anni circa). Il secondo target, gli studenti tredicenni, è stato mantenuto, nonostante l'indagine PISA dell'OCSE si occupi dei quindicenni, poichè già gli studi del 1995 e del 1999 si erano rivolti anche a loro e poichè la prima edizione di PISA aveva destato qualche perplessità nell'ambito della *mathematical literacy* (OECD, 2000; OECD, 2001). Lo studio IEA del 2019 è il settimo ciclo dell'indagine dopo 24 anni dalla prima edizione. Come accaduto negli anni precedenti sono stati somministrati due strumenti: un test cognitivo per indagare il grado di conoscenza matematica e scientifica degli studenti del quarto e dell'ottavo anno di scolarità e un set di questionari conoscitivi rivolti a vari destinatari per conoscere il background nel quale avviene l'apprendimento della matematica e delle scienze. Per avere un quadro completo sul contesto educativo e sociale sia formale che non formale, non solo vengono analizzati i curricula scolastici di ciascun paese, come già ricordato, ma attraverso il questionario scuola, il questionario insegnanti e il questionario studenti sono indagate variabili qualitative riferite al singolo e specifico contesto di riferimento. La prova cognitiva ha l'obiettivo di testare il profitto scolastico e si basa sul contenuto delle linee guida nazionali per la matematica e le scienze: da una parte gli esperti si concentrano su cosa è stato effettivamente affrontato in aula secondo le indicazioni ministeriali, ovvero cosa i paesi partecipanti all'indagine si aspettano, secondo le linee guida proprie di

ciascun contesto, che gli studenti di 8 e 13 anni sappiano fare, e dall'altra si concentrano sui domini cognitivi (cosa si è studiato), ovvero sulle performances auspicabili e sui risultati attesi (come si è studiato). Gli ultimi dati TIMSS disponibili, risalenti al 2019, descrivono il quadro relativo a 64 paesi e circa 600.000 studenti di tutto il mondo che hanno compilato la prova cognitiva e il questionario per la maggior parte telematicamente. Per descrivere il grado di conoscenza della matematica e delle scienze sono stati elaborati quattro benchmark internazionali a ciascuno dei quali corrisponde un livello diverso di competenza: al livello più basso gli studenti hanno una conoscenza di base della disciplina, mentre al livello più avanzato sanno applicare le competenze in una moltitudine di situazioni più o meno complesse e sanno argomentare circa il ragionamento che hanno seguito per giungere alla soluzione. In generale sia all'ottavo che al quarto grado di scolarità, sia in matematica che nelle scienze i paesi asiatici raggiungono punteggi elevati, spesso seguiti dalla Scandinavia. Il dato allarmante riguarda i livelli di competenza per il quarto grado: nel 2019 sia in matematica che in scienze circa il 92% degli studenti in 58 paesi su 64 partecipanti raggiungeva il livello minimo di competenza, mentre solamente il 6-7% degli studenti si trovava a un livello avanzato. Contrariamente alle aspettative, nei paesi che raggiungono i punteggi più elevati si trova anche il maggior numero di studenti che hanno un livello di competenza molto basso sia in matematica che nelle scienze. La situazione è molto simile per l'ottavo grado (Mullis & Martin, 2017; Fishbein et alii, 2018; Cotter et alii, 2019; Yin & Fishbein, 2019; Mullis et alii, 2020; Martin et alii, 2020; Vetterranta et alii, 2020ab; Cesareni et alii, 2020; Arias, 2020; Martin, 2020; Garcia et alii, 2022; Reynolds et alii, 2022; Tang et alii, 2022).

L'Italia ha partecipato a tutte queste indagini, con scarse ricadute sulla politica scolastica e sull'insegnamento. In Italia si è infatti parlato pochissimo dei risultati di queste ricerche anche perché sono state pubblicate poche analisi dei risultati italiani. Gli Stati Uniti che avevano agli inizi degli anni 90 fortemente sostenuto e finanziato l'indagine TIMSS dell'IEA per sapere a che punto erano gli studenti americani in matematica rispetto a quelli di altri Paesi, alla fine del secolo, si sono ritirati dall'indagine TIMSS-Advanced, come del resto la maggior parte degli altri Paesi, mentre l'Italia continua, si direbbe in

modo inerziale, a partecipare all'indagine e quindi a pagarne i costi, senza che si conoscano le ragioni di questa scelta. (Bottani, 2009, p. 17)

La seconda indagine dell'IEA iniziata nel secolo scorso e sistematizzata all'inizio degli anni 2000 è PIRLS, *Progress in International Reading Literacy Study*. Questo studio viene condotto ogni cinque anni dal 2001 ed è riconosciuto come l'indagine principale per la misurazione e la valutazione delle tendenze nel rendimento in lettura nel quarto anno di scolarità.

PIRLS fornisce dati comparativi a livello internazionale sul modo in cui i bambini leggono e offre informazioni rilevanti per le politiche per migliorare l'apprendimento e l'insegnamento. Lo studio è somministrato in una fase di transizione chiave nello sviluppo della lettura dei bambini: il passaggio dall'apprendimento della lettura alla lettura per imparare. La valutazione del rendimento nella lettura in questa fase cruciale fornisce agli educatori, agli insegnanti e ai responsabili politici informazioni chiave sull'efficacia del loro sistema educativo e aiuta a identificare le aree di miglioramento⁷. (traduzione a cura dell'autrice)

L'indagine PIRLS, in italiano ICONA, *Indagine sulla Comprensione dei bambini di Nove Anni*, ha come target i bambini della scuola primaria, nello specifico il quarto anno di scolarità. Il predecessore di questo studio è il *Reading Literacy Study*, RLS, condotto tra gli anni 1985 e 1992, che per quel tempo è stata la più grande indagine educativa mai svolta: 31 paesi partecipanti con più di 200.000 studenti, 10.000 insegnanti e 9.000 scuole. In quell'occasione per la prima volta a livello internazionale è stata elaborata la definizione di competenza in lettura, che sarebbe stata poi ripresa dalle indagini successive, ed è stato utilizzato il modello di Rasch per calcolare il punteggio nelle prove cognitive incrociando l'abilità del rispondente con il grado di facilità o difficoltà della domanda (Elley, 1992; Lundberg & Linnakyla, 1993; Binkley & Williams, 1996; Taube & Mejdning, 1996; Bechger et alii, 1998;

⁷ Per approfondimenti circa l'indagine PIRLS: <https://www.iea.nl/studies/iea/pirls> (ultima consultazione 21 febbraio 2023)

Elley, 1998; Gustafsson & Rosen, 2006). Oggi lo scopo dello studio è indagare le competenze in lettura in relazione agli obiettivi e agli standards dell'educazione alla lettura. I paesi partecipanti all'edizione del 2016 per la quale sono stati già rilasciati i dati, al contrario quella del 2021 è ancora in corso di pubblicazione, sono 50, mentre gli studenti coinvolti 319.000. Come per la matematica e le scienze, anche per la letteratura sono stati stabiliti quattro livelli di competenza, basso, intermedio, alto ed elevato. Similmente all'indagine TIMSS, anche in PIRLS 2016 nel 96% dei paesi gli studenti del quarto anno di scolarità hanno raggiunto un punteggio al test che denota una competenza bassa in lettura, nell'82% si trovano a un livello intermedio, nel 47% ad un livello alto e nel 10% al livello elevato. Questi ultimi sono in grado di distinguere, interpretare e valutare le informazioni che leggono. La Federazione Russa e Singapore sono i paesi con un punteggio medio più elevato, seguiti da Hong Kong, Irlanda e Finlandia. Per venti anni, dal 2001 al 2021, sono state analizzate le tendenze in lettura attraverso la somministrazione sia di prove cognitive, prima cartacee poi elettroniche, sia di questionari rivolti a genitori, dirigenti scolastici, insegnanti e studenti per esplorare il contesto scolastico e domestico in cui le competenze in lettura si sviluppano e per identificare le fragilità che gli studenti possono incontrare in termini di risorse materiali e opportunità.

La volontà che ha accompagnato l'IEA dalle prime indagini comparative alle versioni più recenti di TIMSS e PIRLS, infatti, è stata ed è tutt'ora quella di supportare i decisori politici dei paesi partecipanti nel prendere decisioni *evidence-based* (Mullis et alii, 2012; Raimondi et alii, 2013; Mullis et alii, 2016; Hooper et alii, 2016; Martin et alii, 2017; Warner-Griffin et alii, 2017; Howie, 2017; Leino, 2017; Garcia Crespo & Galian Nicolas, 2018; Cardarello & Pintus, 2019).

3.4. L'OCSE e l'idea di una nuova indagine

Verso la fine degli anni Novanta, dopo che la raccolta di indicatori per l'educazione era alla quinta pubblicazione annuale, i team di ricercatori dell'OCSE, Organizzazione per la Cooperazione e lo Sviluppo Economico, provenienti da tutto il mondo iniziarono a riflettere sulla

necessità e sulla fattibilità di un nuovo studio comparativo internazionale che si sarebbe andato ad aggiungere a quelli già condotti dall'IEA (OECD, 2000).

L'occasione venne fornita dalla raccolta di indicatori *Education at a Glance* che dal primo anno si era andata pian piano espandendo. Gli studiosi che stavano lavorando alla serie annuale di indicatori avevano delineato due ordini di problematiche tecnico-metodologiche: in primo luogo, fin dall'inizio, la costruzione degli indicatori si era rivelata essere un processo dispendioso temporalmente poichè i dati riferiti ai singoli contesti venivano inviati spesso in ritardo dagli enti di ricerca debitamente preposti di ciascun paese; in secondo luogo gli esiti degli apprendimenti scolastici su larga scala venivano monitorati solamente dall'IEA, che essendo un'associazione, non disponeva di fondi strutturali dunque conduceva ricerche a cadenza irregolare. Mancavano quindi i dati sui risultati negli apprendimenti e quelli invece sugli input e sugli output dei sistemi educativi erano indiretti e forniti da enti terzi. Questo non andava incontro alle esigenze dell'OCSE che, avendo l'obiettivo di acquisire l'indipendenza sperimentale e l'autonomia nella raccolta dei dati, voleva imporsi nel panorama internazionale con due novità: un'indagine sulle competenze degli adulti e un'indagine sulle competenze dei quindicenni (Castoldi, 2005).

I vari studi condotti a partire dagli anni Cinquanta dall'IEA avevano riguardato sempre la popolazione in età scolare, infatti, nessun ente di ricerca sovranazionale si era mai spinto nelle coorti di età più avanzate. L'indagine PIAAC, *Programme for the International Assessment of Adult Competencies*, condotta ancora oggi, ha l'obiettivo di analizzare le *skills* della popolazione adulta e fa parte del programma denominato *Survey of Adult Skills*. Si focalizza sulle competenze degli adulti e sulle abilità chiave quali l'alfabetizzazione, l'elaborazione delle informazioni, la matematica e la risoluzione dei problemi. Non solo raccoglie dati su come gli adulti utilizzano le proprie abilità nella vita quotidiana, a casa, al lavoro e nella comunità, ma fornisce anche informazioni molto importanti sull'impatto degli apprendimenti e sugli output dell'educazione (OECD, 2013bc; 2015; 2016; 2019ef; 2021c; 2022).

La seconda soluzione trovata dai ricercatori dell'OCSE, per raggiungere l'indipendenza sperimentale e l'autonomia nella raccolta dei dati, è stata progettare uno studio che fosse comparativo e longitudinale con lo scopo di misurare le competenze degli studenti che si apprestano a terminare la scuola dell'obbligo. I dati raccolti dall'IEA spesso non erano sufficienti poichè interessavano le coorti più giovani e poichè si focalizzavano sul profitto scolastico partendo dall'analisi delle linee guida ministeriali di ciascuna nazione partecipante. Più che misurare le conoscenze, quindi, si decise di misurare le competenze degli studenti dell'ultimo anno della scuola dell'obbligo (OECD, 2000, 2001).

L'insoddisfazione per tali esperienze precedenti era dovuta ad un insieme di ragioni, di cui si maturava una crescente consapevolezza: 1) la difficoltà di individuare conoscenze ed abilità omogeneamente presenti nei curricoli scolastici dei diversi Paesi, su cui poter effettuare comparazioni trasversali sufficientemente attendibili; 2) la quantità di risorse necessarie per condurre ricerche di ampia portata, con conseguente dilatazione dei tempi ed innalzamento dei costi; 3) la scarsa ricaduta dei dati raccolti, che tendevano a rimanere confinati nelle cerchie dei ricercatori e degli studiosi, con uno scarso impatto sulle politiche scolastiche e – ancor più – sul corpo docente delle scuole coinvolte; 4) l'episodicità delle varie iniziative, che non consentiva di disporre di serie storiche di dati omogenei in grado di effettuare comparazioni longitudinali sull'evoluzione dei diversi sistemi scolastici; 5) le forti resistenze, esplicite e latenti, presenti nell'establishment scolastico preoccupato degli effetti potenzialmente esplosivi di queste iniziative. In questo clima nasce il progetto PISA, un'indagine internazionale. (Castoldi, 2005, p. 26)

Nei documenti programmatici degli anni tra il 1997 e il 1999 si scriveva del *Project for International Student Assessment*, tuttavia, data la volontà degli stakeholders di rendere il nuovo progetto longitudinale e regolare nel tempo, il nome dell'indagine fu ben presto cambiato in *Programme for International Student Assessment* (OECD, 2000, 2001; Bottani, 2003; Bottani & Vrignaud, 2005; Bottani & Pegoraro, 2005; Viteritti & Giancola, 2015).

3.5. L'indagine PISA dell'OCSE

Il programma PISA, *Programme for International Student Assessment*, è un'indagine svolta nei paesi membri e partner dell'OCSE, Organizzazione per la Cooperazione e lo Sviluppo Economico, volta a misurare alcune competenze ritenute fondamentali nei quindicenni di tutto il mondo. Lo studio nacque dal progetto INES che, all'inizio degli anni Novanta, l'OCSE promosse in vista della costruzione di un sistema di indicatori quantitativi per la comparazione dei sistemi educativi di diversi paesi. Il progetto sugli *Indicators of Education Systems*, che sarebbe poi confluito nelle prime pubblicazioni del report *Education at a Glance* descritto nel capitolo precedente, riportò fin da subito una mancanza strutturale, già accennata precedentemente: molto spesso i dati a disposizione per l'elaborazione del sistema di indicatori erano di tipo indiretto, ovvero non prendevano in considerazione gli esiti negli apprendimenti degli studenti. Mancavano, in sostanza, dati sulla base dei quali discutere delle risorse impegnate nel funzionamento dei sistemi di istruzione o dei processi educativi attuati (Castoldi, 2005). Il campo dei risultati formativi comparabili internazionalmente era stato già esplorato qualche decennio prima dall'IEA, tuttavia vi erano alcuni aspetti non del tutto soddisfacenti, oltre alla periodicità non regolare, che non sostenevano l'indipendenza e l'autonomia sperimentale a cui puntava l'OCSE nell'indagare nel loro complesso i sistemi di istruzione, non solamente in termini di rendimento scolastico.

Con queste premesse sul finire degli anni Novanta si sviluppa l'indagine PISA, un'indagine comparativa internazionale che, rispetto a quelle condotte dall'IEA, focalizzate sui curricula scolastici, ha il compito di indagare le competenze trasversali in lettura, matematica e scienze con ciclicità triennale. I giovani sono pronti ad affrontare le sfide del futuro? Il target cui si rivolge l'indagine, i quindicenni di tutti i paesi partecipanti, è stato scelto per indagare quanto le giovani generazioni che si apprestano a concludere l'obbligo scolastico sono preparate per affrontare le sfide contemporanee. Come già ricordato poco più sopra, inizialmente denominato *Project for International Student Assessment*, l'indagine PISA è stata ben presto rinominata *Programme for International*

Student Assessment per dare continuità alla volontà di confrontare longitudinalmente aree ritenute rilevanti per essere cittadini attivi e partecipi (OECD, 1999, 2001).

Il programma PISA è frutto della collaborazione tra l'OCSE e altri enti riuniti all'interno del *PISA Governing Board* che ne definisce gli obiettivi, ne monitora i processi e orienta le analisi e la diffusione dei risultati. A sua volta gli istituti di ricerca e le università che compongono il *PISA Governing Board* sono in dialogo costante con i responsabili nazionali, i National Project Manager, che dirigono l'indagine in ciascun paese partecipante e che sono stati inclusi tra i testimoni privilegiati della presente ricerca dottorale, come si espliciterà nella seconda parte del lavoro. Negli anni il consenso nei confronti di questa indagine è aumentato in termini numerici: all'edizione del 2012 hanno partecipato sessantacinque paesi (OECD, 2013b), a quella del 2015 settantadue (OECD, 2016, 2017), mentre all'ultima edizione di cui sono disponibili i dati, il 2018, hanno partecipato settantanove paesi tra membri e partner dell'OCSE (OECD, 2019abc). Ogni ciclo triennale dedica particolare attenzione a uno degli ambiti principali misurati nell'indagine PISA: nel 2000 il dominio principale è stata la lettura, nel 2003 la matematica, nel 2006 le scienze, nel 2018 nuovamente la lettura (OECD, 2019abc). Gli altri due ambiti, definiti secondari, nel 2018 matematica e scienze, sono comunque rilevati, ma con un set di prove cognitive più contenuto rispetto a quelle dell'ambito primario, ma che permette comunque un confronto longitudinale. Il campionamento stratificato prevede l'individuazione di circa 150 scuole per ciascun paese dalle quali vengono estratti tra i 35 e i 40 studenti per un totale di circa 6.000 studenti partecipanti per ciascuna realtà nazionale (OECD, 2019e).

La *reading literacy*, competenza in lettura, è stata definita in PISA 2018 come segue.

La competenza in lettura è la comprensione, l'uso, la valutazione, la riflessione e l'interazione con i testi al fine di raggiungere i propri obiettivi, sviluppare le proprie conoscenze e potenzialità e partecipare all'interno della società. (OECD, 2019e, p. 28, traduzione a cura dell'autrice)

In ogni edizione dell'indagine vengono somministrate ai partecipanti sia prove cognitive che prevedono sia domande a risposta aperta che chiusa, per misurare le competenze nei vari ambiti principali e secondari, sia un questionario indirizzato a stakeholders differenti per conoscerne il background e permettere un'analisi di tipo contestuale.

Alcuni tratti distintivi qualificano il progetto, in risposta ad alcune difficoltà evidenziate nelle ricerche precedenti:

1) il focus su una fascia d'età – gli studenti quindicenni – posta a conclusione della scuola di base nella totalità dei Paesi coinvolti e quindi ritenuta cruciale per accertare la produttività del sistema scolastico complessivo;

2) la centratura non sui contenuti dei curricoli scolastici, bensì su un insieme di competenze ritenute necessarie per la vita adulta e per un inserimento attivo nella società;

3) la scansione in cicli periodici a cadenza triennale, in modo da consentire un monitoraggio nel tempo dell'evoluzione dei diversi sistemi scolastici, con una restituzione tempestiva dei risultati a distanza di circa un anno dalla rilevazione;

4) l'integrazione delle prove sui livelli di apprendimento con dati di processo, sia relativi al singolo studente, sia relativi ai contesti scolastici di appartenenza, allo scopo di esplorare le relazioni intercorrenti tra processi e risultati conseguiti;

5) la gestione coordinata del progetto da parte dei Governi dei Paesi partecipanti – sia in rapporto ai finanziamenti, sia in rapporto alla conduzione politica e tecnica dell'operazione – come condizione preliminare per la diffusione e l'impiego dei dati ottenuti.

L'insieme di questi tratti consente di cogliere il valore innovativo del progetto e le sue valenze sia sul piano culturale, sia sul piano politico. In termini culturali risulta particolarmente interessante la focalizzazione sul concetto di *literacy*, autentica parola chiave del progetto PISA, intesa come padronanza del soggetto di un determinato dominio culturale ad un livello adeguato da consentire una partecipazione attiva alla vita sociale. Ciò comporta uno spostamento di attenzione dalla cultura scolastica, tendenzialmente formale e decontestualizzata, alle situazioni reali di vita, come contesti nei quali verificare il possesso delle competenze indagate; dalle singole discipline curriculari, intese come settori chiusi e separati nei quali compartimentare gli apprendimenti, ad abilità generali, che attraversano il curriculum scolastico e richiamano un concetto di educazione comprensivo anche delle esperienze

extra-scolastiche. Tale opzione non è solo determinata da ragioni strumentali connesse al programma di accertamento, volte a distanziarsi dai contenuti dei curricoli scolastici necessariamente situati nella cultura dei diversi Paesi e di difficile comparazione, ma esprime anche un'idea di formazione di base dinamica e di più ampio respiro rispetto alla mera acquisizione dei traguardi enunciati nei programmi scolastici. Una formazione di base che acquista la sua rilevanza in rapporto alla sua capacità di attrezzare le nuove generazioni alle esigenze poste dall'attuale "società della conoscenza" e si misura con i contesti quotidiani che caratterizzano la vita reale e con le competenze presupposte da un inserimento attivo e consapevole nella vita sociale. Per un universo scolastico ancora incline a pensarsi "tra parentesi" rispetto alla vita reale, come portatore di un apprendimento formale e astratto – di tipo secondario –, l'impronta conferita al progetto PISA costituisce un forte impulso verso una visione più pragmatica e concreta dell'apprendimento. (Castoldi, 2005, p. 27-28)

Si tratta, dunque, di un'indagine innovativa che, nonostante i pareri contrari e nonostante la più o meno ampia diffusione, è in grado di andare incontro ai veloci cambiamenti della società e alle esigenze della globalizzazione, basti pensare, tra le altre cose, all'*electronic assessment*, e di misurare la relazione tra l'ambiente socio-familiare e l'acquisizione di competenze fondamentali per la vita di un cittadino (Prais, 2003; OECD, 2004; Castoldi, 2005; Bottani, 2009; Waldow, 2009; Corsini & Asquini, 2010; Asquini, 2011; Bieber & Martens, 2011; Corsini & Asquini, 2012; Corsini, 2013; Wiseman, 2013; Corsini & Losito, 2014; Asquini, 2014; Rutkowski & Rutkowski, 2016; Asquini, 2016; Goldstein, 2017; Corsini et al., 2017; Asquini, 2018; Lewis, 2020; Corsini, 2020a).

Capitolo quarto

La valutazione delle scuole, valutazione esterna ed autovalutazione

4.1. Qualcosa si muove... Per la valutazione scolastica

La prima edizione dell'indagine PISA dell'OCSE oltre a suscitare l'interesse degli studiosi sul piano tecnico e metodologico, suscitò anche l'interesse dell'opinione pubblica in particolar modo in alcuni paesi e per alcuni risultati raggiunti dagli studenti. Se da una parte il bisogno di dati comparabili sui sistemi educativi attraverso i quali interpretare i contesti nazionali aveva spinto gli stakeholders a intraprendere la strada di un'indagine internazionale sulle competenze dei quindicenni, dall'altra i risultati pubblicati nel 2001 avevano evidenziato alcune differenze nelle performances degli studenti in ambiti ritenuti fondamentali come la lettura, la matematica e le scienze soprattutto a discapito di alcuni paesi (OECD, 2000, 2001). Similmente a ciò che era avvenuto per il documento pubblicato negli Stati Uniti nel 1983 in cui si denunciava il preoccupante stato di salute del sistema scolastico americano (Gardner, 1983), di cui si è parlato nel capitolo due della prima parte, i primi risultati di PISA avevano destato preoccupazione in alcuni contesti scolastici, nonostante gli investimenti cospicui in termini economici. Allo stesso modo gli stessi risultati di PISA avevano lanciato segnali incoraggianti guardando soprattutto ai paesi *best performers*, come Finlandia, Corea, Canada, Giappone, Svezia, che avevano raggiunto risultati più che soddisfacenti, pur investendo ragionevolmente nell'educazione e puntando alla qualità e all'equità nella distribuzione delle risorse (OECD, 2004).

Questi esiti spostano ulteriormente il dibattito sui risultati PISA, poichè i paesi partecipanti cercano una migliore comprensione del motivo per cui alcune nazioni ottengono risultati negli apprendimenti più forti ed equi rispetto ad altri. A tal proposito, l'OCSE sta pubblicando una serie di rapporti tematici internazionali che analizzano in modo più approfondito l'impatto di alcuni

fattori individuali, scolastici e di sistema sulle prestazioni degli studenti. Allo stesso modo, la maggior parte dei paesi sta portando avanti ricerche e analisi nazionali per situare meglio i risultati di PISA nel proprio contesto educativo, sociale ed economico nazionale. Tuttavia, questa analisi dei risultati può spiegare le differenze nei modelli di prestazioni solo in misura limitata, dal momento che numerosi fattori – come il disegno sistemico dell'offerta educativa, le più ampie strategie di riforma e innovazione dell'istruzione e il contesto generale in cui operano i sistemi educativi – restano difficili da quantificare e non sono stati misurati da PISA. (OECD, 2004, p. 3, traduttrice a cura dell'autrice)

Su questa scia, il Ministero dell'Educazione tedesco, visti i livelli di competenza dei giovani quindicenni (OECD, 2001; Allmendinger & Leibfried, 2003; Ammermuller, 2004; Ammermuller, 2005; Ertl, 2006; Waldow, 2009; Neumann et alii, 2010; Martens & Niemann, 2013; Niemann et alii, 2017), progettò uno studio di portata internazionale con l'obiettivo di collegare i risultati di PISA con variabili che potevano avere alcuni effetti su questi ultimi, ad esempio lo sviluppo professionale degli insegnanti, le strategie di riforma e di innovazione dell'educazione, la valutazione e il monitoraggio del sistema scolastico, le governance e l'allocatione delle risorse, l'inclusione degli studenti stranieri, le linee guida nazionali per la definizione di standard (German Federal Ministry of Education and Research, 2003). Un gruppo di esperti proveniente da Canada, Inghilterra, Finlandia, Francia, Olanda, Svezia avviò uno studio sistematico, tematico e comparativo su alcune di queste variabili chiave ritenute responsabili dei buoni risultati raggiunti dai paesi oggetto di studio. Le evidenze furono sintetizzate nel rapporto intitolato *What Makes School Systems Perform? Seeing School Systems Through The Prism Of Pisa* pubblicato nel 2004 (OECD, 2004).

Sebbene sia lo studio multilaterale che questo rapporto possano essere visti solo come un primo passo verso l'integrazione dei risultati della ricerca educativa internazionale quantitativa e qualitativa, ciò stabilisce un'importante pietra miliare e allarga le basi per il dialogo politico e la collaborazione tra i paesi volta a definire e rendere operativi gli obiettivi educativi in modi innovativi, facendo un uso efficace della ricca fonte di prove fornite dai

confronti transnazionali. Insieme ai rapporti tematici internazionali di PISA, questo rapporto può supportare il continuo spostamento dell'attenzione politica dagli input educativi ai risultati dell'apprendimento e assistere i paesi nel portare a miglioramenti nella scolarizzazione e nella migliore preparazione per i giovani che entrano in una vita adulta in rapido cambiamento. (OECD, 2004, p. 4, traduzione a cura dell'autrice)

Oltre ad analizzare aspetti quali lo sviluppo professionale degli insegnanti, l'inclusione degli studenti non-nativi, le strategie per l'innovazione e le riforme del sistema scolastico, furono prese in considerazione anche le pratiche di monitoraggio del sistema educativo.

Gli standard di prestazione possono funzionare solo se vengono implementati e valutati in modo coerente. Le valutazioni del rendimento degli studenti sono ormai comuni in molti paesi OCSE e spesso i risultati sono ampiamente riportati e utilizzati nel dibattito pubblico e da coloro che si occupano del miglioramento della scuola. Inoltre, i paesi stanno monitorando sempre più le prestazioni rispetto alle norme internazionali, attraverso varie valutazioni transnazionali tra cui l'indagine PISA dell'OCSE, Organizzazione per la Cooperazione e lo Sviluppo Economico, e le indagini TIMSS e PIRLS dell'IEA, *International Association for the Evaluation of Educational Achievements*. Tuttavia, la ragione insita nelle valutazioni nazionali e la natura degli strumenti utilizzati variano notevolmente all'interno e tra i paesi. I metodi impiegati nei paesi OCSE includono diverse forme di valutazione e ispezione esterna e attività di garanzia della qualità e di autovalutazione proprie delle scuole. Le opinioni su come i risultati della valutazione possono e devono essere utilizzati sono differenti. Alcuni li vedono principalmente come strumenti per rivelare le migliori pratiche e identificare i problemi condivisi al fine di incoraggiare gli insegnanti e le scuole a migliorare e sviluppare ambienti di apprendimento più favorevoli e produttivi. Altri estendono il loro scopo fino alla contendibilità dei servizi pubblici o dei meccanismi di mercato nell'allocazione delle risorse, ad esempio, rendendo pubblicamente disponibili i risultati comparativi delle scuole per facilitare la scelta dei genitori o per ricevere fondi. Il tipo di benchmark riferito alle performances utilizzato e riportato ha quindi importanti implicazioni per le varie parti interessate e coinvolte, inclusi genitori, insegnanti, scuole e studenti. (OECD, 2004, p. 24, traduzione a cura dell'autrice)

All'inizio degli anni 2000 le pratiche messe a sistema nei paesi analizzati nel report dell'OCSE in collaborazione con il Ministero dell'Educazione tedesco erano differenti. In generale, l'importanza data ai benchmark nazionali e internazionali era considerevole sia in Canada, che in Inghilterra, Finlandia, Francia, Olanda e Svezia, tuttavia, c'erano alcune differenze negli scopi, negli strumenti utilizzati e nella diffusione dei dati della valutazione. I motivi alla base della valutazione di sistema erano differenti. La maggior parte dei paesi monitorava il proprio sistema educativo per mantenerne o migliorarne la qualità o per raccogliere dati empiricamente fondati sui quali costruire framework teorici. Inoltre, a seconda dell'organizzazione politica, la valutazione assumeva obiettivi diversi. Nei paesi federali era utilizzata per monitorare le province autonome e confrontarle tra di loro, in altri paesi invece per garantire a tutti pari opportunità indipendentemente dal background socio-economico di riferimento. Molta attenzione era dedicata anche ai livelli da raggiungere negli apprendimenti, al supporto da offrire agli insegnanti, all'*accountability* delle scuole e alla trasparenza dei risultati. Nel 2004 un numero cospicuo di paesi partecipava già almeno a due o tre indagini sugli apprendimenti tra TIMSS, PIRLS e PISA e conduceva valutazioni tematiche all'interno dei confini nazionali combinando la valutazione diagnostica e formativa. Spesso per motivi metodologici legati alla progettazione degli strumenti di valutazione, la diffusione dei risultati e degli esiti avveniva con modalità differenti, anche se la restituzione di un feedback individuale alle scuole come azione di supporto al miglioramento era una pratica utilizzata in tutti i paesi oggetto di studio (OECD, 2004).

4.2. Tra valutazione esterna ed autovalutazione delle scuole

Il primo paese ad implementare la valutazione degli istituti scolastici è stato l'Inghilterra tramite l'OFSTED, *Office for Standards in Education, Children's Services and Skills*, un dipartimento non-ministeriale con l'obiettivo di innalzare gli standard educativi e l'assistenza sociale delle giovani generazioni. Vista la sua presenza in tutto il paese, grazie ai distaccamenti regionali, all'OFSTED è dato il compito di ispezionare

ancora oggi regolarmente migliaia di istituti che forniscono istruzione, formazione e assistenza e di redigere report per il parlamento e per tutti gli stakeholders, genitori compresi⁸. Il motivo di tutto ciò è storicamente definito: l'*Education (Schools) Act* del 1992, nella sezione 9, stabiliva che entro la fine del 1994 bisognava ispezionare tutte le scuole secondarie finanziate dallo stato, alcune scuole dell'infanzia e alcune scuole primarie e scuole speciali al fine di garantire la qualità del sistema di istruzione inglese e un miglioramento continuo (Matthews & Smith, 1995). Oggi il raggiungimento degli standard educativi è monitorato dall'OFSTED attraverso studi regolari sia su larga scala che di caso con lo scopo di perseguire quattro obiettivi: miglioramento degli standard di apprendimento degli studenti, miglioramento della qualità dell'esperienza educativa vissuta dagli studenti, aumento dell'efficienza della gestione finanziaria delle scuole e sviluppo dell'etica scolastica e dell'autostima degli studenti (Matthews & Smith, 1995). Il sistema di ispezione scolastica inglese dal suo avvio è stato un modello educativo di valutazione delle scuole per gli altri paesi, ma ha suscitato anche alcune critiche (Ferguson et alii, 2000; Case et alii, 2000; Chapman, 2001; Chapman, 2002; Rosenthal, 2004; Plowright, 2007; Baxter & Clarke, 2013; Wilkins & Antonopoulou, 2020). Dopo l'esperienza inglese molti sono stati i paesi che si sono avvicinati all'idea di inserire all'interno della valutazione del sistema scolastico nel suo complesso anche quelle delle singole scuole (Poliandri & Romiti, 2012; Poliandri et alii, 2013; Serpieri, 2016).

Le ragioni storico politiche sono note. Basti ricordare che i cambiamenti avvenuti nei principali paesi industrializzati hanno posto ai sistemi educativi questioni cruciali alle quali far fronte: aumento della scolarizzazione, crescente eterogeneità della popolazione scolastica, conseguente aumento della spesa per l'istruzione, difficoltà di inserimento delle fasce deboli della popolazione, ma anche incremento dei servizi formativi offerti. In Italia - come nella maggior parte dei paesi dove si sono innescati questi processi - una prima risposta a

⁸ Per approfondimenti circa l'OFSTED: <https://www.gov.uk/government/organisations/ofsted/about> (ultima consultazione 21 febbraio 2023)

questi cambiamenti è stata l'introduzione dell'autonomia scolastica, che ha permesso una maggiore indipendenza nelle scelte delle singole istituzioni, ad esempio nell'attuazione del curriculum, ma anche negli aspetti di tipo organizzativo-gestionale. Nello stesso tempo, osservando le disparità esistenti sia a livello territoriale, sia fra le singole unità scolastiche, ci si è posti interrogativi sull'equità delle opportunità educative, sui costi del servizio, sulla qualità dell'offerta formativa, sull'aderenza alla legislazione in vigore e sulle innovazioni normative. (Cardone et alii, 2010, p. 698)

In merito alla valutazione delle scuole, alle visite esterne e ai processi di autovalutazione possono essere rilevate più tendenze. Come per la valutazione esterna, anche la funzione dell'autovalutazione, il significato attribuitogli e la compilazione dei relativi rapporti assumono significati propri a seconda del contesto di riferimento. Simons (2013) definisce la valutazione interna come "un processo di valutazione intenzionale delle pratiche scolastiche che fornisce approfondimenti sull'esperienza educativa degli studenti in misura maggiore rispetto a quelli misurati dai test". Le revisioni della letteratura, tuttavia, restituiscono un quadro poliedrico rispetto all'utilizzo che ne viene fatto a scuola (Nevo, 2001).

Tra i progetti che, a partire dalla fine del 1980, hanno utilizzato i concetti di "miglioramento", "efficacia", "autovalutazione", il progetto ISIP definisce la School Based Review (Van Velzen et al., 1985, cit. in Hopkins, 1988, 1990) come un'ispezione sistematica (descrizione e analisi) fatta da una scuola, un suo sottoinsieme o da individui (insegnante, dirigente scolastico) del suo attuale funzionamento, con la finalità di raccogliere informazioni per migliorarlo. L'autovalutazione, e la riflessione che ne deriva, può essere impiegata mediante i suoi strumenti «to the curriculum or organisation of the school» (Hopkins, 1988, p. 14), e nel progetto si insiste molto sulla differenza tra l'applicazione della revisione al miglioramento di un aspetto specifico, o diversamente, nella sua trasformazione in un atteggiamento generalizzato applicato abitualmente dalla scuola. (Perazzolo, 2019, p. 267-268)

L'autovalutazione scolastica implica differenti fasi tra progettazione, azione, monitoraggio e nuovo intervento che mettono gli attori coinvolti nella condizione di incontrarsi per discutere degli obiettivi della valutazione, raccogliere e analizzare i dati e implementare piani per il

miglioramento. Rispetto alla valutazione esterna, l'autovalutazione scolastica informa gli stakeholders sui punti di forza e sulle criticità e può avere sia funzioni formative nell'ottica dello sviluppo e del miglioramento, sia funzioni sommative quando vista come cartina al tornasole per l'*accountability* esterna. Gli studi condotti sull'*accountability orientation* (Pfau & Kay, 2002), infatti, hanno messo in luce che può essere rilevata una forma di valutazione consequenziale delle istituzioni scolastiche in cui l'autovalutazione precede le visite esterne e fornisce un framework all'interno del quale orientare le azioni di miglioramento.

Il modello cooperativo (Kyriakides e Campbell, 2004) e la valutazione collaborativa (Christie et al., 2004) integrano valutazioni interne ed esterne nelle prime tre fasi, Plan, Do e Check, in cui valutatori interni ed esterni collaborano per pianificare, progettare, condurre e riferire insieme sulla valutazione. Di conseguenza, gli interessi e i punti di vista dei valutatori esterni, così come i criteri definiti dalla scuola, vengono presi in considerazione contemporaneamente. Secondo Kyriakides e Campbell (2004) i criteri di misurazione emergono da entrambi i tipi di valutazione, tanto è vero che bisogna tentare di combinare i risultati per condurre una valutazione olistica e soddisfare le esigenze di entrambe le parti. Nella valutazione parallela, i valutatori interni ed esterni non partecipano alla valutazione reciproca. Secondo il modello di Kyriakides e Campbell (2004) e di Christie e colleghi (2004) sono la scuola e l'ente esterno a effettuare le proprie valutazioni condivisibili una volta completate. Questo è il caso di sistemi educativi come quello degli Stati Uniti, dove la valutazione esterna o l'*accountability* sono basate su test in cui le performances degli studenti, misurate con prove cognitive costruite esternamente, sono utilizzate per esprimere giudizi sulle prestazioni di scuole e insegnanti. (Nelson et al., 2015, p. 10, traduzione a cura dell'autrice)

A partire dall'avvio delle indagini internazionali sugli apprendimenti (OECD, 2000), dall'esperienza dell'OFSTED (Rosenthal, 2004) e dalla necessità di richiedere alle scuole di rendicontare il proprio operato per predisporre piani di miglioramento secondo lo schema "plan, do, check, act" (Deming, 1994; Nelson et al., 2015), si sono sviluppati differenti percorsi di valutazione delle istituzioni scolastiche che hanno subito modifiche ed evoluzioni nel tempo (Eurydice, 2015).

4.3. La valutazione esterna

Undici anni dopo l'interesse dell'OCSE e del Ministero dell'Educazione tedesco per la valutazione dei sistemi scolastici in ottica comparativa, grazie al lavoro dell'Unione Europea in collaborazione con Eurydice, la rete comunitaria di informazione sull'istruzione, è stato pubblicato un report intitolato *Assuring Quality in Education Policies and Approaches to School Evaluation in Europe* che restituisce un quadro dello stato dell'arte della valutazione delle scuole e della loro qualità in 32 paesi europei. Così come avvenuto nel 2004 (OECD, 2004; Eurydice, 2004), nel 2015 venivano comparati approcci, strutture e ruoli affidati tanto alla valutazione esterna delle istituzioni educative quanto all'autovalutazione di queste ultime prendendo in considerazione i processi valutativi, gli strumenti, la figura del valutatore esterno e l'utilizzo e la diffusione dei risultati delle valutazioni (Eurydice, 2015).

La valutazione esterna delle istituzioni scolastiche già utilizzata all'inizio degli anni 2000, nel 2015 veniva introdotta come progetto pilota in alcuni paesi europei, tra questi l'Italia, nazione in cui le visite esterne non erano ancora tra gli elementi principali di garanzia della qualità (Agnelli, 2014; Allulli, 2014; Falanga, 2014; Spallacci, 2015; Lucisano & Corsini, 2015). In 27 su 31 paesi partecipanti alla rilevazione di Eurydice (2015) in cui esisteva la valutazione esterna delle scuole, questa era condotta da organismi centrali, l'ispettorato, responsabili delle visite ispettive. Spesso gli obiettivi di queste ultime erano condivisi tra il livello centrale, quello regionale e quello locale, tuttavia vi erano anche paesi in cui organismi sub-regionali erano preposti alla definizione della valutazione esterna traducendo quest'ultima in diversi gradi di standardizzazione negli enti decentrati. Inoltre, in alcuni paesi, tra cui la Finlandia, le autorità locali e le singole scuole erano responsabili della valutazione delle istituzioni scolastiche in collaborazione con le autorità centrali e le linee guida da esse emanate. Era possibile ricoprire il ruolo di valutatore se in possesso di alcuni requisiti come l'abilitazione all'insegnamento o un certo numero di anni di esperienza nella scuola come docente o come dirigente. Le qualifiche pedagogiche, di ricerca o psicologiche potevano rappresentare un titolo

di preferenza per l'acquisizione del titolo di valutatore esterno (Eurydice, 2015). In Italia, ad esempio, l'aver acquisito determinate competenze al di fuori della scuola poteva rappresentare un titolo di preferenza per entrare nel team di valutatori esterni (Spallacci, 2015; Eurydice, 2015). Con le dovute differenze tra paese e paese, i criteri da seguire nella valutazione esterna per rilevare il buon funzionamento di una istituzione scolastica si concentravano intorno al rispetto delle normative, delle performances degli studenti e delle attività educative e gestionali. Nonostante in alcuni paesi i criteri di valutazione si concentrassero più su alcuni aspetti che altri e nonostante i diversi gradi di autonomia dei team di valutatori, l'iter valutativo era progettato in tre fasi ed era pressochè simile: analisi dei dati disponibili, visita ispettiva, rendicontazione (Cardone et alii, 2010). Gli strumenti a disposizione dei valutatori erano differenti e diversificavano le fonti informative della valutazione esterna favorendo un'ampia e articolata raccolta di dati e un'elaborazione finale *evidence-based* (Bechis & Tomatis, 2014).

Gli studiosi di Eurydice (2015) hanno analizzato anche i processi attuati durante la valutazione esterna e hanno rilevato due possibili approcci: il *risk-based approach* e il *profile-raising activities*. L'approccio basato sul rischio (Zimmerman, 1986; Klinke & Renn, 2002), utilizzato in quei casi in cui la scuola non raggiunge gli standard definiti, ha implicazioni sull'efficienza in termini economici, ovvero più fondi là dove sono necessari, e in termini di pertinenza degli indicatori presi in considerazione.

Rafforza il ruolo della valutazione esterna come processo mirato a identificare le debolezze del sistema e può potenzialmente avere l'effetto contrario di rendere invisibile la buona pratica. Tuttavia, l'approccio basato sul rischio e le sue implicazioni dovrebbero essere ulteriormente studiati e potrebbero essere un'area in cui i paesi cooperano tra pari. (Eurydice, 2015, p. 9, traduzione a cura dell'autrice)

Le attività di miglioramento, invece, si basano sulla diffusione delle *best practices*. In questo approccio i valutatori esterni non hanno solamente il compito di fare un bilancio di ciò che non funziona

all'interno dell'istituzione scolastica, ma hanno anche il compito di aumentare la visibilità di quelle scuole che raggiungono buoni risultati (Eurydice, 2015).

Le valutazioni esterne non hanno solo lo scopo di trovare difetti nelle prestazioni delle scuole, ma anche di aumentare la visibilità di quelle che stanno andando bene e stanno ottenendo buoni risultati. L'idea di utilizzare la valutazione esterna come strumento di identificazione e diffusione della visibilità delle buone pratiche consente di raccogliere e condividere prove su ciò che funziona e in quali circostanze con ritorni positivi sia a livello scolastico che di sistema. (Eurydice, 2015, p. 9, traduzione a cura dell'autrice)

Rispetto, infine, alla diffusione degli esiti delle visite ispettive, già nel 2015 erano presenti punti di vista diversi nei paesi oggetto di analisi che delineavano altrettanti modi diversi di concepire l'*accountability* delle scuole. Harris e Herrington (2006) individuano due tipologie di rendicontabilità: la *government-based accountability* e la *market-based accountability*. In quest'ultima l'operato della scuola è rendicontato per il pubblico esterno, o come dice la parola stessa per il mercato, mentre nella prima l'autorità centrale è responsabile della qualità dell'educazione e deve assicurare che le istituzioni scolastiche raggiungano i benchmark prestabiliti. Nell'*accountability* di mercato la responsabilità nella scelta della scuola è demandata ai genitori che, come in un mercato, dispongono della libertà di scelta del servizio da fruire e si basano sulla competitività delle scuole. Le scuole, dal canto loro, sono sanzionabili e incentivabili, possono ricevere più o meno fondi a seconda della performance raggiunta e devono offrire quante più attività agli studenti e alle loro famiglie. L'*accountability* di mercato si basa su due fondamenti: da una parte vi è l'accesso libero alle informazioni (es. quelle pubblicate sui siti istituzionali delle scuole), dall'altra la libertà di scelta degli studenti e delle famiglie. Nell'*accountability* demandata agli organi centrali, al contrario, l'approccio è di tipo top-down, le scuole ricevono le medesime linee guida e le informazioni circa la qualità delle istituzioni scolastiche devono essere accessibili in prima istanza a coloro che prendono decisioni con esse e per esse (Harris & Herrington, 2006).

Gli indicatori che inquadrano i sistemi scolastici nell'uno o nell'altro modello di *accountability* sono relativi alla pubblicazione e alla diffusione degli esiti delle visite ispettive e alla libertà delle famiglie e degli studenti di scegliere quale scuola frequentare (Eurostat, 2012).

Un rapporto che viene reso pubblico all'interno di un sistema che attribuisce pieni poteri a genitori e alunni di scegliere la propria scuola provoca dinamiche di mercato, dove la rendicontazione, e quindi il sistema di valutazione esterna che consente la produzione di tale report, diventa una leva che può influenzare la scelta dei genitori e, di conseguenza, fare pressione sulle istituzioni scolastiche per migliorare le prestazioni. Al contrario, una relazione che è non resa pubblica o che è diffusa con limitazioni, all'interno di un sistema che assegna gli studenti alle scuole sulla base di criteri predefiniti, come la vicinanza geografica, sposta la responsabilità di queste ultime all'interno del perimetro dei doveri dello Stato, che è in ultima analisi responsabile dell'educazione dei suoi cittadini e del miglioramento delle prestazioni delle istituzioni scolastiche. (Eurydice, 2015, p. 10, traduzione a cura dell'autrice)

La maggior parte dei sistemi educativi nel 2015 si disponeva all'interno di questo continuum: in alcuni di questi, anche quando i rapporti delle visite esterne erano pubblici, le famiglie non avevano la possibilità di scegliere l'istituzione scolastica, in altri, invece, la totale libertà delle famiglie era accompagnata dalla non diffusione pubblica dei rapporti delle visite esterne, spesso sostituita dalla nascita di reti informali di genitori che condividevano le loro esperienze sulle scuole o dalla diffusione dell'opinione pubblica (Olson Beal & Hendry, 2012; Eurydice, 2015; Burgess et alii, 2015; Fong, 2019; Zambelli et alii, 2021).

4.4. L'autovalutazione

Mentre le visite ispettive si andavano affermando in molti paesi europei, l'autonomia scolastica e la decentralizzazione ponevano l'autovalutazione degli istituti scolastici sempre più al centro dell'attenzione e la rendevano da consigliata ad obbligatoria in un numero sempre crescente di paesi (Eurydice, 2015). Nella maggior parte dei casi le politiche relative all'implementazione della valutazione

interna delle scuole lasciavano a queste ultime l'autonomia di decisione, con qualche differenza: in alcuni paesi gli istituti scolastici erano obbligati ad utilizzare lo stesso schema di valutazione degli ispettori esterni, in altri vi era un framework specifico per l'autovalutazione, in altri ancora il contenuto dell'autovalutazione era prescritto a livello centrale, dunque uguale per tutte le scuole. A variare erano anche gli stakeholders coinvolti nei processi valutativi: in alcuni casi il numero dei partecipanti era vasto e includeva anche i genitori e gli studenti, in altri includeva solo gli stakeholders interni alla scuola. Sull'utilizzo degli esiti della valutazione interna la situazione era eterogenea: spesso veniva lasciata libertà decisionale allo staff della scuola, nonostante le indicazioni sull'utilizzo per lo sviluppo redatte dagli organi centrali che includevano l'elaborazione di un piano per il miglioramento (Mathison, 1991; Barzanò, 2000; Eurydice, 2015; Agasisti et alii, 2015; Nelson et alii, 2015; Poliandri, 2016).

Nel 2004, solo un quarto dei paesi dava la possibilità alle scuole di utilizzare indicatori come i risultati dei test degli studenti per confrontare il proprio rendimento con quello di altre scuole che si trovavano in condizioni simili a quelle o alle medie nazionali. Attualmente, ciò si verifica in due terzi dei sistemi educativi, il che pone tali indicatori come il secondo strumento più frequentemente disponibile per la valutazione interna in tutta Europa. Questa tendenza ha coinciso con l'introduzione di meccanismi nazionali di test obbligatori in molti paesi negli ultimi anni, nonché con il fatto che un certo numero di paesi fornisce alle singole scuole i risultati dei test aggregati. (Eurydice, 2015, p. 11, traduzione a cura dell'autrice)

Nel 2015, invece, quasi tutti i paesi mettevano a disposizione delle istituzioni scolastiche almeno uno strumento per l'autovalutazione. Tra questi vi erano l'utilizzo del medesimo schema di valutazione adoperato dai valutatori esterni, la costruzione di indicatori per comparare le scuole tra di loro, la formazione del personale sull'autovalutazione, la redazione di manuali e linee guida specifiche, il supporto finanziario, lo scambio di esperienze in forum online e il supporto di specialisti esterni. In Spagna, ad esempio, le scuole disponevano di quasi tutte le misure di supporto elencate, comprese quelle finanziarie, mentre in altri paesi

erano presenti solamente linee guida elaborate centralmente (Mathison, 2011; Eurydice, 2015). In Italia, invece, veniva promosso dall'INVALSI, Istituto Nazionale per la Valutazione del Sistema Educativo di Istruzione e Formazione, e Indire, Istituto Nazionale Documentazione, Innovazione e Ricerca Educativa, il progetto VALeS⁹, di durata triennale, con l'obiettivo di accompagnare le 300 scuole campionate nei processi di autovalutazione, valutazione esterna e miglioramento (Palumbo & Pandolfini, 2016; Carro & Nencioni, 2016; Quadrelli et alii, 2018).

Il progetto Valutazione e Sviluppo Scuola (VALeS), promosso dal MIUR, è stato realizzato in collaborazione con l'INVALSI e l'Indire con la finalità di sperimentare un modello di valutazione delle istituzioni scolastiche orientato al miglioramento. A seguito della richiesta dell'Unione Europea di creare un Sistema Nazionale di Valutazione nel nostro Paese, le trecento scuole di primo e secondo ciclo del progetto VALeS, cui hanno aderito volontariamente nell'anno scolastico 2011/12, hanno rappresentato un cantiere per la sperimentazione di procedure, protocolli e strumenti funzionali alla realizzazione del Sistema Nazionale di Valutazione (SNV) così come sarebbe stato successivamente regolamentato nel d.p.r. 80 del 2013 per tutte le istituzioni scolastiche. [...] Il modello di valutazione proposto da VALeS si colloca all'interno di un approccio formativo che considera la valutazione come il punto di partenza per una continua azione di miglioramento, non circoscrivendo la sua funzione a quella esclusivamente volta alla rendicontazione. Il miglioramento è interpretato come quell'insieme di processi, sia organizzativi sia legati all'azione didattica, che è possibile implementare in modo partecipato con tutti i soggetti coinvolti e che si auspica possano avere un impatto positivo sui progressi e sui risultati degli studenti. Secondo questa prospettiva sia la valutazione interna sia quella esterna, e il quadro di riferimento a esse sotteso, sono considerate risorse a disposizione delle scuole per migliorare la propria efficacia come comunità di apprendimento. (Poliandri, 2016, p. 415-416)

Dal 2015 ad oggi l'attenzione dei ricercatori è stata focalizzata sullo studio della qualità e della validità dei processi di valutazione interna. Infatti, dall'inizio del nuovo millennio nei paesi occidentali,

⁹ Per approfondimenti circa il progetto VALeS: <https://www.invalsi.it/invalsi/ri/vales/index.php> (ultima consultazione 21 febbraio 2023)

indipendentemente dal numero di misure a supporto dei processi valutativi messi a disposizione nelle singole nazioni (Eurydice, 2015), è stata affidata sempre più importanza all'autovalutazione come momento di riflessione interna alla scuola. Il nuovo orientamento ha fatto sì che la conoscenza delle pratiche di autovalutazione sia divenuta fondamentale (Muzzioli et alii, 2016).

4.5. Un ponte tra passato e presente

Negli ultimi anni si è assistito a una crescente attenzione e a un crescente interesse nei confronti della valutazione degli istituti scolastici. Lo scarso utilizzo degli esiti della valutazione, il numero contenuto di visite ispettive soprattutto in alcuni contesti nazionali e il taglio delle spese per le valutazioni su larga scala hanno, tuttavia, richiesto agli stakeholders accurate riflessioni. Se da una parte l'affiancamento alla valutazione esterna dell'autovalutazione ha avviato processi di responsabilizzazione da parte delle scuole, dall'altra il *New Public Management* è diventato la strategia organizzativa per la programmazione sociale (Love, 1991; Mathison, 2011).

Negli anni più recenti l'*Education and Training 2020 Working Group Schools*, nell'ottica della creazione di un'Area Educativa Europea, ha contribuito al monitoraggio della misurazione e della valutazione dei sistemi educativi europei elaborando alcune considerazioni chiave per i decisori politici. Nello specifico il team di ricerca si è soffermato sulle modalità attraverso le quali coloro che prendono decisioni per e sulla scuola possono supportare i processi valutativi per il miglioramento del sistema educativo. Tra le considerazioni chiave vi sono quelle che comprendono l'autovalutazione e la valutazione esterna come due momenti posti lungo lo stesso continuum che si integrano e rafforzano a vicenda.

Gli ispettorati e altre agenzie nazionali e locali possono sostenere l'autovalutazione della scuola fornendo orientamenti e strumenti, sviluppando approcci basati sul dialogo, facendo della valutazione esterna, dell'autovalutazione e del miglioramento della scuola aspetti importanti

dell'ispezione, promuovendo la collaborazione nelle scuole e nelle reti e condividendo le buone pratiche. I responsabili politici possono sostenere altre forme di valutazione condotte da terzi, come ad esempio dirigenti scolastici e insegnanti di altre scuole, rappresentanti di locali autorità o ricercatori. (European Commission, 2020, p. 3, traduzione a cura dell'autrice)

In tal senso le indagini comparative internazionali sugli apprendimenti degli studenti sono uno strumento utile sia per i sistemi educativi nel loro complesso sia per le singole scuole poiché, se analizzati e letti accuratamente, forniscono informazioni sui livelli di competenze degli studenti e sulle possibili politiche da implementare sia nel contesto macro-nazionale che in quello micro-didattico. Tuttavia, risulta importante evitare di basare le indicazioni centrali indirizzate alle scuole solamente sulle aree risultate deficitarie nelle rilevazioni standardizzate e di tradurre le evidenze fornite in linee guida di cui la scuola nel singolo contesto può far uso (European Commission, 2020).

I piani di miglioramento sintetizzano in questa direzione le azioni che un'istituzione scolastica intende intraprendere per l'accrescimento (Reding, 2002).

Il Piano di Miglioramento (PdM) rappresenta il documento nel quale confluiscono le scelte progettuali, adottate dalla scuola sulla base delle esigenze contestuali, per pianificare percorsi formativi efficaci e di miglioramento mediante la chiara individuazione di obiettivi e di criteri non ambigui di valutazione. Il PdM è parte integrante del Piano Triennale dell'Offerta Formativa (PTOF), la cui predisposizione tiene conto di quanto stabilito nella sezione quinta del Rapporto di Autovalutazione (RAV), relativa all'individuazione delle priorità (declinate in traguardi e obiettivi di processo). Nel PdM si stabiliscono le azioni di miglioramento più adeguate da mettere in campo, sulla base delle criticità emerse e delle risorse a disposizione, concentrando la propria attenzione sull'ottimizzazione e integrazione funzionale delle risorse disponibili per il conseguimento dello scopo. (Marzano, 2019, p. 128)

La letteratura sul tema rileva, però, alcune criticità strutturali nell'elaborazione dei piani di miglioramento da parte delle scuole

(OECD, 2013; Robasto, 2017). Affinchè la valutazione delle scuole, sia esterna che interna che degli apprendimenti, abbia un impatto sul funzionamento di queste ultime, occorre che ci sia un utilizzo ragionato degli esiti della valutazione. Nel caso specifico dei piani di miglioramento, ma anche della valutazione interna degli istituti, la formazione da parte di chi utilizza questi strumenti e compila questi rapporti, gli insegnanti e i dirigenti scolastici, dunque, è essenziale (Schon, 1993; Damiano, 2007; Bonaiuti et al., 2012; Zanniello, 2016, 2017).

Parte seconda

LA RICERCA

Capitolo primo

Il disegno di ricerca dottorale

Nella seconda sezione della presente tesi verrà illustrata la parte empirica del lavoro che ha costituito il progetto di ricerca. Inizialmente verrà presentato il disegno della ricerca, verranno delineate le scelte metodologiche e procedurali operate e verranno esplicitate le domande di ricerca che hanno orientato il lavoro e gli strumenti che hanno permesso la raccolta dei dati. I capitoli successivi riporteranno i risultati emersi e le conseguenti implicazioni prima di giungere alle conclusioni finali.

1.1. Lo studio di caso, lo stile e il metodo

Il Merriam-Webster Dictionary, uno dei dizionari della lingua americana più diffusi, definisce lo studio di caso come un'analisi approfondita di singole unità, ad esempio persone o comunità, che collega i fattori di sviluppo con l'ambiente circostante¹⁰. Ciò che caratterizza le singole unità è quello che Ragin e Becker (1992) definiscono "casing", ovvero la scelta delle unità di analisi e dei confini che le determinano. Nel tempo molti studiosi si sono annoverati sulla definizione dello studio di caso, tuttavia, spesso sono finiti in una "palude classificatoria" (Gerring, 2004) che non ha fatto altro che rendere la questione più complessa. Negli ultimi trenta anni, infatti, le ricerche sullo studio di caso hanno prodotto più di venticinque definizioni differenti, ognuna delle quali con un'enfasi e una sfumatura diversa a seconda delle direzioni intraprese all'interno dell'indagine nella quale è stata messa a punto (VanWynsbergh & Khan, 2007). Nonostante il largo utilizzo che si fa dello studio di caso nell'esame dei fenomeni, spesso il mondo della ricerca "relega" questa metodologia alle

¹⁰ Per approfondimenti circa la definizione del Merriam-Webster Dictionary: <https://www.merriam-webster.com/dictionary/case%20study> (ultima consultazione 21 febbraio 2023)

fasi esplorative che precedono la messa a punto del disegno di ricerca e dello studio sul campo vero e proprio (George & Bennett, 2005; Flyvbjerg, 2011). Nel sottolineare il “limbo” dello studio di caso, molto utilizzato, ma poco considerato come metodologia di ricerca (Gerring, 2004), Flyvbjerg (2011) individua cinque *misunderstanding* che lo contraddistinguono:

- La conoscenza generica e teoretica è più apprezzabile di quella che deriva dallo studio di casi concreti;
- Non si può generalizzare sulla base di un caso specifico, dunque lo studio di caso non può contribuire all'avanzamento della conoscenza scientifica;
- Lo studio di caso è più indicato quando si devono formulare le ipotesi nelle prime fasi di un disegno di ricerca, le altre metodologie sono più adatte per la verifica di queste ultime e per la costruzione della teoria;
- C'è un bias nella verifica nello studio di caso che si traduce nella tendenza a confermare i preconcetti e la soggettività del ricercatore;
- Spesso è difficoltoso riassumere e sviluppare proposizioni e teorie generali sulla base di uno studio di caso.

Il primo malinteso si fonda sul fatto che lo studio di caso restituisce una conoscenza concreta e ancorata al contesto dal quale dipende. Nonostante le critiche (Campbell, 1975), è stato riconosciuto che questa conoscenza contestuale è importante tanto per i destinatari quanto per il ricercatore. Infatti, se quest'ultimo vuole sviluppare le proprie competenze di ricerca, è bene che si confronti con un'esperienza di indagine sul campo concreta e ancorata al contesto. Diversamente dalle altre discipline, però, nelle scienze sociali è molto complesso, se non impossibile, produrre una conoscenza che sia generica, indipendente dal contesto e riferibile a tutti gli eventi che rientrano in una determinata categoria. Il secondo malinteso si fonda sul fatto che i dati raccolti nello studio di caso non possono essere generalizzati, tuttavia, come ricorda Flyvbjerg (2011), anche Galileo, quando rigettò la teoria della gravità formulata da Aristotele, non si basò su una vasta gamma di osservazioni, bensì la falsificazione galileiana partì da un assunto teorico per arrivare in un secondo momento a un esperimento pratico. Il terzo

malinteso si fonda sulla verifica delle ipotesi correlata con lo studio di caso e sul fatto che bisogna identificare alcuni casi emblematici. Senza ombra di dubbio individuare un caso critico e/o paradigmatico non è semplice e richiede esperienza poichè si tratta di casi che dovrebbero confermare chiaramente o disconfermare del tutto le ipotesi di ricerca (Flyvbjerg, 2004). Il quarto malinteso si fonda sul fatto che nello studio di caso gli esiti della ricerca possono avere un'attendibilità dubbia poichè raccolti in modo tale da confermare alcuni preconcetti che il ricercatore aveva già in mente. L'"inconveniente paralizzante" dello studio di caso risiede nell'applicazione dei metodi scientifici, eppure, la tendenza a verificare in base ai propri schemi mentali non è propria solo dello studio di caso, ma anche della mente umana (Flyvbjerg, 2011). La critica mossa spesso nei confronti dello studio di caso è quella secondo la quale quest'ultimo sarebbe poco rigoroso da un punto di vista scientifico, ciononostante lo studio di caso gode di un suo rigore tanto quanto le metodologie più quantitative. Ragin e Becker (1992) e altri colleghi, a conferma di ciò, sottolineano che i ricercatori che hanno condotto studi di caso e approfondimenti spesso hanno riportato che i loro preconcetti erano errati e che i dati raccolti durante l'indagine li hanno costretti a rivedere le loro ipotesi. Il quinto e ultimo malinteso si fonda sul fatto che la sintesi e la generalizzazione a partire da uno studio di caso è spesso complessa poichè la componente narrativa che caratterizza questa metodologia di ricerca lo rende difficoltoso. Per un ricercatore esperto, tuttavia, uno studio di caso particolarmente articolato da sintetizzare non è arduo, infatti ciò indica che attraverso l'indagine è stata scoperta una problematica particolarmente ricca e da approfondire (Flyvbjerg, 2004, 2011).

Un altro paradosso sullo studio di caso è quello delineato da VanWynsbergh e Khan (2007). Essi sostengono che lo studio di caso non è un metodo, ovvero una tecnica o una procedura per raccogliere dati, poichè i ricercatori che lo utilizzano non raccolgono i dati in modo prescrittivo, al contrario fanno ricorso a vari metodi di ricerca per costruire o scoprire il caso. Inoltre, pur esistendo varie declinazioni dello studio di caso, si pensi a quello esplorativo, esplicativo, etnografico o longitudinale, nessuno di questi per essere definito tale richiede delle

procedure specifiche per la raccolta dei dati. Lo studio di caso, infatti, non è un disegno di ricerca all'interno del quale l'obiettivo dello studioso è quello di chiarire le caratteristiche di una classe di fenomeni simili (Garring, 2004), dal momento che non delinea un piano così prescrittivo. I disegni di ricerca, al contrario, guidano le ricerche dalla formulazione delle domande e delle ipotesi alle conclusioni, passando per la raccolta e l'interpretazione dei dati (Yin, 2003a). In questo senso, lo studio di caso, non delineando in modo prescrittivo i processi da seguire per la raccolta, l'analisi e l'interpretazione dei dati, non può essere definito un disegno di ricerca. Quest'ultimo, però, non può essere nemmeno una metodologia (spesso scambiata con il "metodo"), vale a dire un'interazione tra teoria, metodi e dati della ricerca (Sjoberg et alii, 1991).

Se Merriam (1998) definisce lo studio di caso come un metodo per indagare unità sociali complesse costituite da molteplici variabili di potenziale importanza per la comprensione del fenomeno oggetto di studio, Stake (1995) non lo spiega come una scelta metodologica che indica metodi e procedure di raccolta dei dati, ma come la decisione di ciò che si vuole studiare, e VanWynsbergh e Khan (2007) lo circoscrivono all'interno di un'euristica.

Lo studio di caso è un'euristica trans-paradigmatica e trans-disciplinare che implica l'attenta delimitazione dei fenomeni per i quali vengono raccolte le evidenze (evento, concetto, programma, processo, etc.). Per trans-paradigmatico intendiamo che il caso di studio è rilevante indipendentemente dal proprio paradigma di ricerca (ad esempio, postpositivismo, teoria critica, costruttivismo). Per trans-disciplinare intendiamo che lo studio di caso non abbia un orientamento disciplinare particolare; cioè può essere utilizzato nelle scienze sociali, nelle scienze applicate, negli affari, nelle belle arti e nella ricerca umanistica, per esempio. Consideriamo l'euristica al suo livello più generale come un approccio che focalizza la propria attenzione durante l'apprendimento, la costruzione, la scoperta o la risoluzione di problemi. Eckstein (2002), ad esempio, ha usato il termine euristico per descrivere un particolare tipo di studio di caso che impiega l'induzione analitica per scoprire l'essenza del caso. L'euristica nello studio di caso di Eckstein è l'induzione analitica, poiché l'induzione analitica è l'approccio consigliato per focalizzare la propria

attenzione sul caso. Riteniamo che ci sono diverse euristiche coinvolte nella ricerca degli studi di caso e proponiamo che queste euristiche vengano utilizzate continuamente per focalizzare la propria attenzione sulla localizzazione o sulla costruzione dell'unità di analisi (il fenomeno per il quale vengono raccolte le evidenze). (VanWynsbergh e Khan, 2007, p. 80)

La flessibilità dello studio di caso viene colta da Stenhouse (1985) e declinata in quattro stili:

- studio di caso etnografico: un caso specifico viene analizzato in profondità attraverso l'osservazione e le interviste; il ricercatore è coinvolto e partecipante e osserva con un occhio esterno le dinamiche di un contesto per comprenderne il funzionamento;
- studio di caso nella ricerca-azione: uno o più casi specifici vengono sviluppati per promuoverne il cambiamento mediante il feedback;
- studio di caso valutativo: uno o più casi specifici vengono analizzati in profondità per fornire agli stakeholders le informazioni necessarie per poter valutare programmi e politiche specifiche;
- studio di caso educativo: uno o più casi specifici vengono analizzati in profondità per comprendere le azioni al fine di arricchire la teoria educativa (Stenhouse, 1988; Benvenuto, 2015).

Alle già citate definizioni date da Merriam (1998) e Stake (1995) si aggiunge anche quella di Yin (2009) secondo cui lo studio di caso è l'indagine su un fenomeno che ha luogo in un contesto dal quale è difficile astrarlo e che si prefigge l'obiettivo di rispondere alle domande sul "come" e sul "perché". Yazan (2015) ha tentato una comparazione in termini di similitudini e divergenze tra le tre definizioni date dai tre autori e le successive implicazioni ed ha messo in evidenza come un ricercatore che intende avvalersi dello studio di caso all'interno del proprio disegno di ricerca trova dinnanzi a sé una molteplicità di approcci e un terreno conteso segnato da una varietà di prospettive. Ciò che può fare è combinare ecletticamente gli elementi delineati da Merriam (1988, 1998), Stake (1995, 2005) e Yin (2003a, 2009) in base alle necessità del proprio disegno di ricerca (Yazan, 2015).

1.2. Tra lo studio di caso e la comparazione nella ricerca dottorale

Il positivismo di Yin (2009) e il costruttivismo di Merriam (1998) e Stake (1995) rappresentano i capisaldi dai quali partire per la costruzione di un disegno di ricerca che si avvalga dello studio di caso che, a seconda dell'approccio richiamato, può differire per la definizione del caso, la raccolta, l'analisi e l'interpretazione dei dati.

Stake (1995) individua due caratteristiche dello studio di caso: l'intrinsichezza e la strumentalità. Lo studio di caso strumentale è più orientato a una preoccupazione esterna, mentre lo studio di caso intrinseco approfondisce una realtà specifica senza la preoccupazione di poter apprendere da quel caso per altri casi futuri, ma perchè il caso è interessante di per sé (Stake, 1995). Yin (2009), invece, nella sua disamina presenta tre dimensioni dello studio di caso che differiscono tra di loro per lo scopo ultimo dell'indagine. Lo studio di caso esplorativo esplora uno o più casi con lo scopo di definire i confini, le domande e le ipotesi di ricerca di uno studio successivo e si avvale di strumenti di misura quali-quantitativi in base alle necessità di osservazione. Lo studio di caso descrittivo descrive in modo dettagliato e approfondito un fenomeno nel contesto in cui ha luogo e anche in questo caso si avvale di strumenti di misura quali-quantitativi in base alle necessità di descrizione (Yin, 2009).

Certamente, nelle sue scelte sarà guidato più da un interesse di approfondimento e conoscenza dettagliata di una determinata realtà socio-educativa che da ipotesi di ricerca. [...] Lo studio di caso, come altri stili, può finalizzare le domande sia sul versante descrittivo (come), sia su quello esplicativo (perchè). (Benvenuto, 2015, p. 156-157)

Lo studio di caso esplicativo, invece, esplicita le ipotesi di ricerca analizzando le relazioni di causa ed effetto dei fattori che riguardano il singolo o i singoli casi alla stregua dei disegni di ricerca sperimentali (Yin, 2009).

Yin (2009) chiarisce ulteriormente che gli studi di caso possono essere condotti attraverso quattro disegni di ricerca (Benvenuto, 2015):

- *Holistic case study* (disegno di caso singolo): l'attenzione è rivolta ad un unico caso emblematico per un periodo di tempo prestabilito (Stake, 1995);
- *Embedded case study* (disegno di caso multiplo): l'attenzione è rivolta ad un unico caso emblematico all'interno del quale ci sono più unità di analisi (Scholz & Tietje, 2002);
- *Multiple case study* (disegno di casi comparativi): l'attenzione è rivolta a due o più casi che vengono comparati tra di loro, gli attori coinvolti sono eterogenei e non sono scelti tramite un disegno di campionamento (Stake, 2013);
- *Embedded multiple-case study* (disegno di casi multipli comparativi): l'attenzione è rivolta a due o più casi per i quali le condizioni vengono replicate, gli strumenti di raccolta dei dati dalle singole unità di analisi sono molteplici (Yin, 1994ab).

L'*embedded multiple -case study* di tipo esplorativo e descrittivo (Yin, 1994ab, 2009; Stake, 2013) è stato scelto come disegno di riferimento con l'obiettivo di analizzare più di un contesto contemporaneamente per metterne in luce similitudini e differenze attraverso l'utilizzo di molteplici tecniche di raccolta dei dati. Yin (2003) evidenzia, infatti, che è possibile far ricorso a studi di caso multipli e comparativi sia quando si prevedono risultati contrastanti, sia quando si prevedono risultati simili. Il lavoro di comparazione degli studi di caso fornisce alla letteratura di riferimento un'importante aggiunta proprio sulle divergenze e sulle convergenze che caratterizzano il fenomeno indagato (Vannoni, 2015). Nonostante il cospicuo investimento di tempo richiesto, è stato ritenuto opportuno assumere nel disegno di ricerca dottorale lo studio di caso multiplo e comparativo poiché ciò permette di analizzare i dati sia caso per caso che confrontandoli tra di loro (Cecconi, 2002; Baxter & Jack, 2008; Yin, 2009; Stake, 2013; Vannoni, 2015; Gustafsson, 2017).

Appare evidente dalle considerazioni fin qui riportate che lo studio di caso ha confini tutt'altro che delimitati. Le riflessioni condotte in ambito anglosassone fanno emergere come la "palude classificatoria" descritta da Gerring (2004) vada incontro al paradosso secondo il quale lo studio di caso è molto utilizzato in letteratura, ma poco considerato, soprattutto in ambito nazionale (Sena, 2021). Il non disporre di una definizione

condivisa dalla comunità scientifica lo classifica come un metodo di ricerca che non sempre gode di un suo valore scientifico e che può far ricorso perlopiù a strumenti di raccolta dei dati qualitativi, ad esempio le interviste o le osservazioni partecipanti, non quantitativi come i questionari. A seconda di come viene considerato, infatti, acquisisce un determinato status: che sia un metodo, una metodologia, un disegno di ricerca o una strategia di ricerca differente da quelle più strutturate, uno stile o un approccio, gode di molta flessibilità che gli permette di adattarsi bene al contesto e di raccogliere conoscenza concreta (Stake, 1995; Mettiam, 1998; Bassegy, 1999; Flyvbjerg, 2004; VanWynsberghe & Khan 2007; Brown, 2008; Yin, 2009; Cohen et al., 2011; Simons, 2014; Yazan, 2015; Tight, 2017; Sena, 2021; De Santis, 2022).

Khan (2007), infatti, per orientare il lavoro di chi decide di condurre uno studio di caso, individua sette *features* che fanno di una ricerca uno studio di caso:

1. Piccolo numero di partecipanti. Lo studio di caso si distingue per un'indagine approfondita e ricca di una o più unità di analisi che si traduce nella elaborazione di un report di ricerca dettagliato, spesso richiede un campione molto più ristretto rispetto a quello delle surveys basato su grandi numeri (Gomm et al., 2000).

2. Dettagli contestuali. Lo studio di caso punta a dare al lettore la sensazione di trovarsi immerso nello stesso contesto della ricerca fornendo particolari minuziosi e dettagliati attraverso "l'istanza in azione" (MacDonald & Walker, 1977).

3. Spontaneità del contesto. Lo studio di caso analizza spesso situazioni nelle quali non è possibile controllare comportamenti, organizzazione ed eventi (Yin, 2003b). Anderson e colleghi (2005) sostengono che la "spontaneità" del contesto è rintracciabile nell'impossibilità di ridurre setting complessi entro cui interagiscono diversi attori a un'unica relazione causale.

4. Limitatezza spaziale e temporale. Lo studio di caso fornisce un quadro dettagliato di uno spazio temporale e di un confine spaziale (Merriam, 1988). L'attenzione al tempo e allo spazio fa sì che il contesto venga preso in considerazione a seconda delle relazioni di interesse. L'aula, specifica Khan (2007), è un esempio di uno spazio delimitato

istituzionalmente all'interno del quale hanno luogo determinate dinamiche, curricula scolastici ed aspettative.

5. Ipotesi formulate e lezioni apprese. Le ipotesi di ricerca formulate in fase di progettazione possono essere confutate da ciò che il ricercatore scopre mentre conduce lo studio di caso e lavora nel contesto della ricerca raccogliendo e analizzando i dati (Lincoln & Guba, 2000; Eckstein, 2002). Khan (2002) sottolinea molto la caratteristica della serendipità nel ricercatore che conduce uno studio di caso poichè non è detto che egli riesca a testare le ipotesi formulate inizialmente, ma sicuramente ne genera diverse su altri aspetti implicati nello studio di caso.

6. Fonti di dati multiple. Lo studio di caso raccoglie informazioni da più fonti sviluppando "linee di indagine convergenti" (Khan, 2007) che facilitano la triangolazione e la restituzione di dati efficaci e precisi (Yin, 2003).

7. Estendibilità. Lo studio di caso può arricchire l'esperienza di un lettore dandogli la possibilità di comprendere più a fondo un fenomeno (Donmoyer, 1990). Nello studio di caso il ricercatore analizza le complesse dinamiche che hanno luogo riferendole a un determinato contesto nella speranza che il contesto stesso e le relazioni possano risuonare con il lettore (Khan, 2002, 2007; Yin, 2003).

Lo studio di caso prototipico messo a punto da Khan (2007) è utile ad orientare la ricerca dottorale qui presentata che ha uno scopo esplorativo-descrittivo (Yin, 2009). Come evidenziato nelle sette *features*, proprietà essenziali di uno studio di caso (Khan, 2007), quest'ultimo si avvale di fonti informative multiple e di dati raccolti attraverso strumenti differenti (caratteristica 6), coinvolge un numero ristretto di partecipanti (caratteristica 1) che si muovono nel loro contesto naturale (caratteristica 3) in un arco temporale e spaziale definito (caratteristica 4). Il ricercatore, dal canto suo, non è detto che verifichi le ipotesi inizialmente formulate, ma è quasi certo che scopra aspetti nuovi che non aveva preventivato (caratteristica 5): egli tenta, attraverso la redazione di un report di ricerca ricco e dettagliato, di far sentire il lettore parte viva del contesto (caratteristica 2) estendendo la sua esperienza su quel dato fenomeno (caratteristica 7).

L'indagine empirica presentata, oltre a far proprie le *features* dello studio di caso, si avvale anche di un disegno di ricerca che prevede la comparazione tra più casi, come esplicitato poco più sopra (Yin, 2009).

Gli studi di educazione comparata intra e internazionale affondano le radici lontano nel tempo. Già all'inizio dell'Ottocento Jullien (1817) in *Esquisse et vues préliminaires d'un ouvrage sur l'éducation comparée* indicava che, dopo le guerre che avevano macchiato il secolo precedente, stava emergendo come obiettivo comune la volontà di rinnovare e migliorare l'educazione pubblica. Sugeriva di formare una commissione esperta in ambito educativo e supportata dai principi sovrani e dalle associazioni di categoria già esistenti con l'obiettivo di raccogliere, sistematizzare e comparare tra loro i materiali per uno studio strutturato delle istituzioni e dei metodi di insegnamento-apprendimento presenti nei vari stati europei. Le analisi, che sarebbero state eseguite in meno di tre anni, avrebbero delineato il quadro dell'educazione europea puntando l'attenzione sui paesi progrediti e quelli regrediti, su quelli che rimanevano ancora in piedi, sulla popolazione bisognosa di aiuto, sulle cause dei malfunzionamenti, sulle forze opposte allo sviluppo del progresso religioso, morale e sociale e sul superamento di queste forze. Inoltre, sarebbe stato possibile aprire un focus su quelle aree nelle quali le azioni migliorative possono essere più facilmente "trasferite" da un paese a un altro tenendo sempre in considerazione gli adattamenti necessari per il contesto di riferimento (Jullien, 1817).

Se il lavoro pionieristico di Jullien (1817) rappresenta l'anno zero per la *Comparative Education* (Brickman, 1960), dalla metà del secolo scorso non sono mancati i contributi delle organizzazioni sovranazionali che hanno portato all'evoluzione sia dell'educazione comparata che della somministrazione di test standardizzati su larga scala (OECD, 2011; Hopfenbeck et alii, 2018; Mattarelli, 2022). I ricercatori comparativisti dell'IEA, ad esempio, sono soliti realizzare, prima di ogni indagine internazionale, uno studio di caso comparativo che prende in considerazione tutti i paesi partecipanti in vista della messa a punto del framework teorico di riferimento. Nello specifico dell'*Assessment Framework* dell'edizione 2016 dello studio PIRLS si legge che

Al fine di fornire una importante contestualizzazione per l'interpretazione dei risultati conseguiti in lettura, PIRLS raccoglie informazioni notevoli su quali opportunità educative vengono fornite agli studenti, nonché su quali fattori influenzano il modo in cui gli studenti fruiscono di queste opportunità. Questi dati sul background includono informazioni sulle politiche curriculari nazionali in lettura e sull'organizzazione del sistema educativo nell'ottica della facilitazione dell'apprendimento; su quale ambiente di apprendimento gli studenti dispongono a casa; su quale clima e su quali risorse scolastiche possono accedere; su come l'apprendimento si esplica effettivamente nelle aule. L'Enciclopedia PIRLS è stata pubblicata dopo ogni rilevazione a partire dal 2001. Ogni paese partecipante redige un capitolo descrittivo della struttura del proprio sistema educativo, dei curricula scolastici della scuola primaria e delle politiche generali adottate nell'insegnamento della lettura (es. formazione degli insegnanti, materiali didattici, valutazione). L'Enciclopedia PIRLS 2011 è un compendio prezioso di informazioni sulle pratiche di insegnamento della lettura in tutto il mondo e fornisce una risorsa indispensabile per le politiche educative e la ricerca nell'educazione comparata. (Mullis & Martin, 2015, p. 7, traduzione a cura dell'autrice)

La tradizione italiana in merito all'educazione comparata, pur essendo molto recente nel tempo, ha subito uno slancio nel dopoguerra favorito dai cambiamenti storico-sociali e politici e dai cambiamenti pedagogici.

I processi di decolonizzazione e l'internazionalizzazione dei problemi politici ed economici connessi sia alla competizione tra l'Est e l'Ovest del mondo (Malizia, 2008, p. 380), sia ai processi di industrializzazione sempre più rapidi e competitivi, hanno infatti influito sull'incremento degli studi comparativi, collegandoli in maniera sempre più vincolante alla possibilità di favorire il progresso economico, finanziario e politico, attraverso l'adozione di sistemi scolastici ed educativi che possano formare da un lato quadri dirigenti efficienti e dall'altro manodopera qualificata, oltre che cittadini responsabili in paesi democratici. L'entrata in campo delle scienze umane e l'aprirsi della pedagogia alla contaminazione, soprattutto metodologica, con le scienze dell'educazione (Visalberghi, 1986, pp. 20-23) hanno influito sulla volontà di rendere l'educazione comparata meno intuitiva ed empirica e più scientifica attraverso una maggiore attenzione al metodo. [...] A partire dagli anni Novanta il pluralizzarsi delle ricerche e il loro divenire complementari, il distinguersi dei

metodi in maniera sempre più evidente tra quantitativi e qualitativi – fenomeno che, anche se non ha ancora prodotto “costruzioni metodologiche compiute capaci di sostituire quelle tradizionali” (Malizia, 2008, p. 381), ha avuto il merito di allargare gli strumenti dell’indagine – e, infine, il rinnovarsi e moltiplicarsi dell’oggetto della ricerca hanno permesso all’educazione comparata di promuovere studi sia sul piano teorico che su quello pratico-fattuale, cioè sia nella dimensione “accademica” che in quella “applicata” (Paolone, 2016, pp. 43-77). [...] Dopo la fine della Seconda Guerra Mondiale alcuni dei principali esponenti della cultura pedagogica italiana hanno riservato grande attenzione agli studi comparativi: sono studiosi che hanno viaggiato molto per conoscere realtà diverse dalla loro e raccogliere sul campo dati utili alla comparazione, hanno tradotto le opere di autori stranieri e hanno collaborato spesso anche con gli importanti Organismi Internazionali che si erano andati costituendo in quegli anni. (Callegari, 2017, p. 94-95)

Nel 1955 prese vita una rivista bimestrale, *I Problemi della Pedagogia*, in cui scrivevano studiosi stranieri esperti di valutazione dei sistemi educativi con lo scopo di dare un contributo alla riflessione teorica, tenendo sempre conto del miglioramento dell’azione educativa (Zizioli, 2009). A partire da Volpicelli e Visalberghi fino ad arrivare a Laeng, vicepresidente del CESE e del CEDE, *Comparative Education Society in Europe e Centro Europeo dell’Educazione*, e rappresentante dell’Italia in organizzazioni internazionali come l’IEA e l’OCSE, la statistica anche nel nostro paese è stata vista come necessaria, ma non sufficiente per la comparazione.

La comparazione infatti non può limitarsi alla fase descrittiva, ma deve prevedere anche la fase esplicativa, cioè deve giungere a far comprendere l’unità nella differenza e la differenza nell’unità. Il comparatista deve insomma saper distinguere ciò che vi è di identico nelle differenti situazioni socio-culturali e pedagogico-educative studiate e ciò che vi è di differente in situazioni o istituzioni molto simili o affini. (Callegari, 2017, p. 98)

Il continuo rimando tra teoria e pratica è evidente per spiegare l’evolversi dei fatti educativi, tuttavia, nonostante lo sviluppo di indagini comparative internazionali quali-quantitative e il passaggio di testimone tra i primi studiosi comparatisti italiani, si pensi a Volpicelli,

Laeng e Visalberghi, il campo di studi della pedagogia comparata non ha suscitato particolare interesse nel panorama nazionale (Borri & Bovi, 2007; Palomba, 2011; Callegari, 2013; Benadusi, 2014; Pavan, 2014; Palomba, 2016; Barbieri et alii, 2016; Callegari & Gaudio, 2018; Callegari, 2019; Sicurello, 2020; Mincu, 2020).

1.3. La definizione e la comparazione dei casi, i sistemi scolastici e di valutazione

Saddler (1900) all'inizio del secolo scorso pubblicava un contributo in cui si interrogava su quanto lo studio dei sistemi educativi a livello internazionale potesse giovare praticamente alla conoscenza pedagogica. Dopo di lui Kandel (1933) ripercorreva l'organizzazione dei sistemi educativi e la sistematizzava in sei volumi (Kandel, 1936; Ferning, 1959; Blake, 1982; Rust et alii, 1999; Hans, 2012; Musa & Agbaire, 2013; Trehewey, 2014), mentre Cramer e Browne (1956) e Moehlman (1963) focalizzavano la loro attenzione sull'educazione comparata. L'interesse per questa prospettiva di ricerca si mantenne in ambito anglofono anche nei decenni successivi. Ai pionieri che avevano sottolineato l'importanza dei sistemi educativi negli studi di educazione comparata, seguirono i lavori di Ignas e Corsini (1981), Cameron e colleghi (1983), Postlethwaite (1988, 1995) e quelli pubblicati all'inizio del nuovo millennio da Steyn e Wolhuter (2000) e Marlow-Ferguson (2002).

Alcune di queste opere trascuravano comunque la chiarezza della definizione. [...] Il campo dell'educazione comparata è stato dominato da confronti locali che hanno dato particolare evidenza al paese o allo stato-nazione. Molti lavori [...] prevedevano quale unità di analisi i paesi. I loro autori possono essersi sentiti giustificati ad usare la parola "sistema" nella misura in cui si riferivano a sistemi educativi nazionali; ma pochi esploravano i confini concettuali di quei sistemi educativi nazionali o investigavano il punto fino a cui coesistevano altri sistemi dentro e oltre le frontiere nazionali. Molti degli autori presentavano i sistemi educativi nazionali come se le nazioni in questione avessero solo sistemi unici. (Bray et alii, 2009, p. 128)

Dall'inizio del secolo scorso fino agli anni 2000 l'educazione comparata e nello specifico la comparazione tra sistemi educativi ha visto una netta prevalenza di analisi uniformi anche in quei paesi, Stati Uniti d'America in primis, in cui il sistema di istruzione e formazione era tutt'altro che unificato. La tendenza a scorgere più le similitudini che le differenze all'interno dei sistemi educativi nazionali, infatti, ha generato un bias che ha fatto perdere una parte di conoscenza concettuale (Hayden et alii, 2003; Bray et alii, 2009).

Per i ricercatori che oggi si vogliono occupare di educazione comparata si apre, dunque, il problema prima della definizione e poi della comparazione delle unità di analisi.

Per quanto riguarda la definizione degli "oggetti" dello studio, Kandel (1933) sottolinea che circoscrivere ciò che si intende per "sistema" non è immediato.

Definire un sistema nazionale di educazione non è semplice, nonostante l'uso frequente di tale termine. [...] La difficoltà non è dovuta principalmente alla vasta gamma degli influssi, formali ed informali, che sono coinvolti nella formazione degli atteggiamenti e delle mentalità dei membri di una nazione, ma all'assenza di un criterio unico in base al quale l'esistenza di un sistema nazionale potrebbe essere esaminata. (Kandel, 1933, p. 83)

In questo studio così come nella definizione data da Allport (1955), psicologo statunitense, il "sistema educativo" è inteso come una serie di attori che interagiscono e si influenzano tra di loro.

Ogni aggregato, riconoscibilmente definito, di elementi dinamici che in qualche modo sono tra loro interconnessi e interdipendenti e che continua ad operare seguendo certe leggi in maniera tale da produrre un caratteristico effetto totale. Un sistema, in altre parole, è qualcosa che è coinvolto con un certo tipo di attività e persevera una certa integrità e unità; un sistema specifico può essere riconosciuto come distinto da altri che, comunque, mantengono una relazione dinamica con esso. (Allport, 1955, p. 469)

Tralasciando le ragioni storiche che hanno fatto sì che i governi nazionali dal XIX abbiano dato sempre più rilevanza all'educazione

come strumento di supporto e di potere, l'importanza della comparazione dei sistemi educativi risiede non solo in ragioni analitico-conoscitive e interpretative, ma anche nella volontà di "impedire la nozione di un paese, un sistema" (Bray et alii, 2009).

Identificare i paesi con i sistemi educativi, si può aggiungere, porta con sé il rischio di prospettive statiche: i confini nazionali infatti non cambiano frequentemente. L'analisi di sistemi che sono definiti dalla geografia al contrario si adatta più facilmente a mutamenti di confini e forme. In definitiva: la focalizzazione sui sistemi potrebbe talvolta ridurre i rischi di sovra-generalizzazione e iper-semplificazione ed aiutare a mostrare le dinamiche del cambiamento. (Bray et alii, 2009, p. 132)

La definizione dei sistemi educativi, unità di analisi, dunque non è rapida. Questi ultimi possono essere demarcati o in base a criteri spaziali e intra-nazionali, i confini nazionali, o in base a criteri funzionali e trans-nazionali, le strutture amministrative, tuttavia qualunque sia il criterio accolto, non si può non tenere in considerazione la loro interdipendenza.

Molti sistemi educativi e di formazione condividono storie comuni e gli odierni sistemi si influenzano reciprocamente sempre di più. Essi comparano l'un l'altro i processi e le performance; imparano lezioni dalle reciproche istituzioni e politiche; armonizzano i campi quali la preparazione e la mobilità degli studenti; e si sottomettono alle autorità degli enti sovra-nazionali quali la Commissione Europea e la Banca Mondiale. (Raffe et alii, 1999, p. 19)

Per quanto riguarda, invece, la comparazione, Bray e Thomas (1955) hanno messo a punto un modello-cubo per la classificazione degli studi comparativi in campo educativo. Nel lavoro dedicato ai livelli di comparazione i due autori riportano che la "*cross-fertilization*" tra i campi, che con diverse metodologie ed enfasi concettuali costituiscono gli studi educativi, è limitata. Come già evidenziato, spesso l'educazione comparata era più dedicata a ricerche trans-nazionali che a ricerche intra-nazionali e il rapporto tra la focalizzazione locale e quella internazionale era poco presente (Bray & Thomas, 1955).

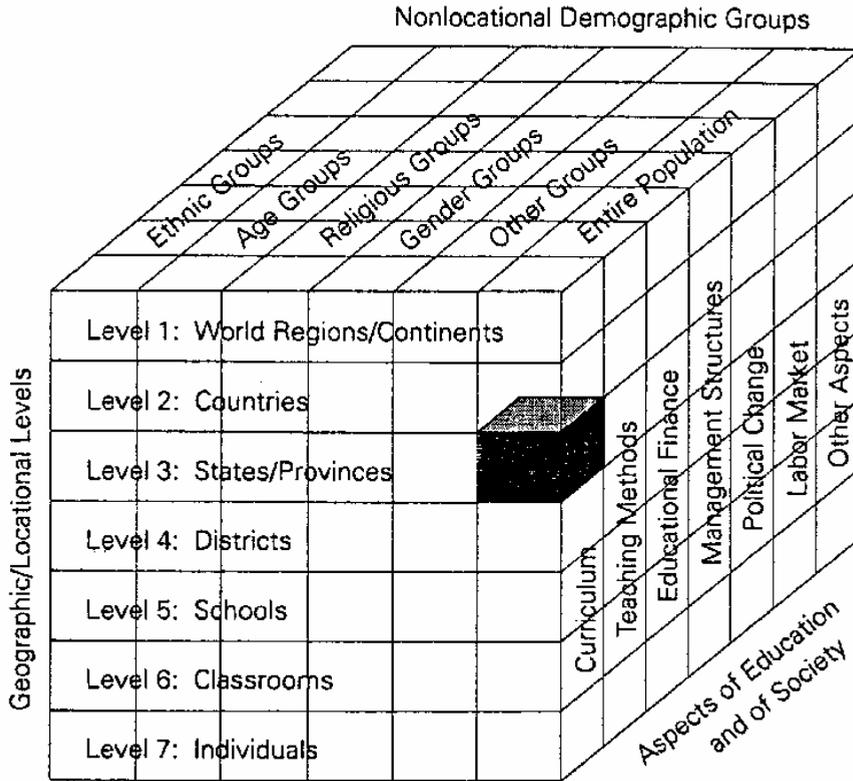


Figura 3. Cubo di Bray e Thomas. Fonte: Bray & Thomas, 1955, p. 475.

La figura 3 rappresenta gli aspetti che compongono ciascun lato del cubo. Nel lato anteriore sono contenuti i *livelli geografico/localazionali* di comparazione: regioni del mondo e continenti, paesi, stati e province, distretti, scuole, classi e individui. Nel lato superiore sono contenuti i *gruppi demografici non localazionali*: gruppi etnici, anagrafici, religiosi, di genere, altri gruppi e l'intera popolazione. Nel terzo lato, invece, sono contenuti gli *aspetti legati all'educazione e alla società*: i curricula, i metodi di insegnamento, gli aspetti economici, la struttura di gestione, i cambiamenti politici, i mercati del lavoro e altri aspetti. Gli studi comparativi, a seconda delle tematiche oggetto di indagine, coinvolgono una o più dimensioni dei rispettivi lati. La cella colorata di nero nella figura 3, ad esempio, indica che lo studio comparativo ha preso in esame i curricula della popolazione totale che abita in due o più stati (Espinoza,

1983; Bray et alii, 1995; Alexander et alii, 1999; Arnove, 2001; Garrido et alii, 2012; Sobe & Kowalczyk, 2013; Bray et alii, 2014; Mazon, 2014; Angiullar, 2018; Ballantine et alii, 2021).

A partire dal modello multi-livello messo a punto da Bray e Thomas (1955) per lo studio comparativo nel campo educativo, il lavoro presentato prende come unità di analisi i sistemi scolastici di più paesi dando rilevanza agli aspetti educativi che includono la valutazione e le strutture di gestione e che riguardano gli individui interessati. In aggiunta al modello-cubo (Bray & Thomas, 1955), l'indagine si è avvalsa anche della "analisi livellizzata" adottata in studi simili condotti precedentemente (Zanazzi, 2014). Zanazzi (2014), nel considerare le diverse fasi e i diversi aspetti implicati nella ricerca valutativa, individua due livelli di analisi che sono stati adattati alla ricerca dottorale. Nel primo livello, divenuto in questo caso il livello internazionale, gli enti sovranazionali mettono a punto linee guida da diffondere localmente. Nel secondo livello, divenuto in questo caso il livello nazionale, i governi nazionali recepiscono le indicazioni internazionali e le modellano in base ai propri contesti, costruendo sistemi di valutazione specifici.

In entrambi i casi, i sistemi di valutazione rispecchiano le culture nazionali e istituzionali, ma allo stesso tempo possono diventare un fattore di cambiamento in tali contesti. Pertanto, è molto importante guardare ai sistemi esistenti sapendo che si evolveranno, a volte più rapidamente di quanto ci aspettiamo. (Zanazzi, 2014, p. 18, traduzione a cura dell'autrice)

Similmente a quanto realizzato per le indagini standardizzate internazionali, si pensi agli studi comparativi dell'IEA tra i paesi partecipanti che precedono ogni *main study* e contribuiscono a definire il framework di misurazione e valutazione di riferimento (Mullis & Martin, 2015, 2019), per ogni realtà oggetto di indagine sono state esplorate e descritte le politiche di valutazione dei sistemi scolastici di cui si darà conto nei capitoli due, tre e quattro della seconda parte.

1.4. Il problema della ricerca e i casi oggetto di analisi

Scegliere uno o più casi oggetto di studio vuol dire indagare un fenomeno sociale e fornire non solo dati dettagliati sull'unità di analisi, ma anche guardare a quel fenomeno sociale da un'angolazione differente (Sena, 2021).

Seguendo i cinque passi individuati da Merriam (1998) nell'implementazione, a suo parere flessibile, di uno studio di caso, è stato fin qui definito il quadro teorico all'interno del quale si muove la ricerca dottorale attraverso la revisione della letteratura. Successivamente verrà illustrato il problema della ricerca e il campione selezionato intenzionalmente, *purposive sampling*, e verranno identificate le domande di ricerca (Merriam, 1998).

Come esplicitato nei paragrafi precedenti, l'indagine che ha preso forma durante il percorso di dottorato fa proprie le proprietà dello studio di caso prototipico illustrate da Khan (2007) e si avvale di un disegno di ricerca che prevede la comparazione tra più casi con scopi esplorativi-descrittivi (Yin, 2009). Per "sistema educativo" si intende un sistema di individui e di attori sociali interdipendenti che interagiscono tra di loro (Allport, 1955) e che non sono delimitati solo da fattori geografici, ovvero i confini politici, ma anche da fattori funzionali, ovvero le strutture amministrative (Bray et al., 2009). Lo studio comparativo, invece, segue il modello multi-livello di Bray e Thomas (1955) che include nei *livelli geografici/locazionali* i paesi, nei *gruppi demografici non locazionali* gli individui interessati e negli *aspetti legati all'educazione e alla società* le strutture di gestione e altri aspetti quali quelli legati alla valutazione della scuola. I casi "rivelatori" (Sena, 2021) sono rappresentati dai sistemi di valutazione scolastica di tre paesi europei: Italia, Finlandia e Spagna.

La valutazione del sistema scolastico è avvertita da qualche anno a questa parte non solo come il problema della ricerca qui presentata, ma anche come un problema di più ampio respiro poichè le scuole sono un elemento cruciale di ogni sistema educativo dal momento che al loro interno hanno luogo i processi di insegnamento-apprendimento (Zanchin, 2002). L'efficienza e l'efficacia di queste ultime, come illustrato

nei capitoli uno, due, tre e quattro della prima parte, è di fondamentale importanza, tanto è vero che per assicurare la qualità di un sistema educativo vi è bisogno di un sistema di monitoraggio e valutazione di quest'ultimo. Tuttavia, anche se la valutazione dell'efficienza e dell'efficacia dei sistemi pubblici, nello specifico dei sistemi educativi, è entrata da qualche anno a questa parte a far parte delle agende politiche nazionali e internazionali, ciò è requisito necessario, ma non un elemento sufficiente affinché il miglioramento del sistema educativo avvenga concretamente (Allulli, 2017). La volontà di osservare e monitorare il funzionamento delle scuole affonda le sue radici in differenti ragioni storiche e politiche. Il crescente incremento dell'offerta formativa, l'aumento della spesa per l'istruzione, la scolarizzazione di massa e l'eterogeneità degli studenti sono solamente alcuni dei fattori decisivi per l'introduzione dell'autonomia scolastica, non solo in Italia, e di conseguenza per l'introduzione di sistemi di valutazione scolastica per combattere le disparità e assicurare a tutti una scuola di qualità (Benadusi, 2009; Sykes et alii, 2012).

La maggior parte dei Paesi si è dotata di un sistema di valutazione dove le unità scolastiche sono poste al centro dell'analisi, intendendo la valutazione come mezzo per vedere ciò che accade nelle differenti realtà autonome. La valutazione non ha rappresentato quindi solo il contrappeso dell'autonomia, ma anche il suo volano, l'innescò del suo motore. (Cardone et alii, 2010, p. 699)

Il secondo problema legato alla valutazione dei sistemi scolastici e alla ricerca qui presentata è di ordine "logistico-organizzativo". Nonostante le spinte all'autonomia scolastica si siano manifestate già da qualche anno e nonostante l'accordo quasi unanime sulla necessità di mettere a punto un sistema di valutazione che monitori a sua volta il funzionamento del sistema scolastico, la diffusione della valutazione di sistema sconta ancora una divulgazione e una messa in pratica eterogenea nei paesi, che si alterna tra adempimenti burocratici e volani per il miglioramento (Giampietro et alii, 2016).

Prendendo le mosse dai due problemi delineati attraverso lo studio della letteratura, sono stati individuati i casi oggetto di analisi. Flyvbjerg

(2004, 2011), ricordando l'importanza della scelta del caso, sottolinea che è bene che il caso non sia "ordinario" ma un caso "emblematico" che porti all'ampliamento del dialogo e del dibattito su una tematica specifica. A tal proposito, oltre all'analisi approfondita del contesto italiano, sono state scelte come unità di analisi la Finlandia e la Spagna. Tralasciando per un momento l'Italia con la storia travagliata, le controversie che hanno caratterizzato la valutazione di sistema negli ultimi venti anni (Cardone et alii, 2010; Corsini, 2010) e le numerose esperienze di valutazione esterna ed autovalutazione che si sono susseguite e che verranno approfondite nel secondo capitolo della seconda parte (Poliandri & Romiti, 2012; Giampietro et alii, 2016), l'attenzione è ricaduta sulla Finlandia e sulla Spagna in qualità di importanti realtà educative europee (OECD, 2004). La prima perché è divenuta famosa dopo la prima edizione dell'indagine PISA dell'OCSE per gli ottimi risultati raggiunti dagli studenti quindicenni (OECD, 2001), anche se aveva già fatto parlare di sé nel secolo scorso. Nelle prime indagini realizzate dall'IEA negli anni Settanta, infatti, gli studenti finlandesi si erano distinti per i risultati soddisfacenti, fino ad arrivare all'inizio degli anni Novanta, quando, nell'indagine RLS sulla lettura, *Reading Literacy Study*, la più grande mai realizzata prima di PISA, gli stessi studenti sia di 9 che di 14 anni hanno raggiunto i livelli più elevati in lettura in tutti i domini indagati (Elley, 1992). La seconda, invece, perché sia in termini storici che socio-educativi è simile all'Italia e dunque agevolmente comparabile (OECD, 2019abc). Dal punto di vista del sistema valutativo delle istituzioni scolastiche, invece, i tre paesi oggetto di studio differiscono notevolmente poiché caratterizzati da modelli di valutazione molto diversi tra loro che incardinano credenze, ideologie e assunzioni differenti, come si vedrà nei capitoli due, tre e quattro della seconda parte.

Modelli centralizzati e gerarchici rispetto a modelli decentralizzati e partecipativi; modelli gestiti esternamente dalle autorità amministrative rispetto a modelli gestiti internamente dalle comunità scolastiche; modelli basati sul riconoscimento formale di premi e punizioni rispetto a modelli che si esauriscono nell'autovalutazione; modelli basati su indicatori di performance

standardizzati e comparabili rispetto a modelli basati su indicatori di performance “locali” non standardizzati e non comparabili; modelli alimentati da basi di conoscenza esclusivamente quantitativa rispetto a modelli alimentati da una ricca base di conoscenza qualitativa costruita tramite interviste e raccolta di opinioni. (Paletta, 2013, p. 1)

La valutazione, in questo caso, se da una parte si configura come uno strumento per monitorare lo stato di salute del sistema scolastico, dall'altra è ampiamente riconosciuto dalla letteratura che può dar vita tanto a interventi di miglioramento (Ehren & Visscher, 2006; Ehren et alii 2020) quanto a possibili effetti collaterali negativi (Volante, 2004; Firestone et alii, 2004), infatti non è stato individuato un modello di valutazione del sistema scolastico che funziona meglio rispetto a un altro.

L'ambiguità dei risultati di ricerca dipende non solo dalla mancanza di condivisione su quale significato attribuire al concetto di “valutazione” e quindi su come misurarlo, ma anche dalla scarsa comprensione che continuiamo ad avere in merito alle diverse condizioni di contesto in cui una certa configurazione dei sistemi di valutazione può dimostrare di essere più funzionale rispetto ad altre. (Paletta, 2013, p. 2)

Partendo da questo presupposto, dunque, l'interesse per i casi nasce dall'esistenza di sistemi scolastici e valutativi molto diversi tra di loro nei contesti italiani, finlandesi e spagnoli. La volontà è quella di superare i confini geograficamente intesi e di esplorare e descrivere questi sistemi educativi da un punto di vista più funzionale, alla luce delle pratiche messe in atto e dai risultati raggiunti (OECD, 2013; OECD, 2019abc).

1.5. Le domande di ricerca e gli obiettivi

L'individuazione dei problemi di ricerca, il primo di più ampio respiro legato alla valutazione dei sistemi scolastici e il secondo di tipo logistico-organizzativo orientato alla diffusione della valutazione di sistema, ha permesso, nelle prime fasi progettuali dello studio, di

delineare le domande di ricerca e dunque i casi oggetto di studio e gli obiettivi dell'indagine.

Dopo aver presentato la rassegna della letteratura e circoscritto il quadro teorico di riferimento dello studio, soddisfacendo i primi due punti delle linee guida proposte da Merriam (1998) per la realizzazione di uno studio di caso, sono stati descritti i problemi della ricerca e il campione selezionato intenzionalmente, *purposive sampling*, assolvendo ai secondi due punti delle linee guida individuate da Merriam (1998). Si procederà ora all'esplicitazione del quinto punto: le domande di ricerca interconnesse agli obiettivi del progetto di ricerca dottorale.

Lo studio ha lo scopo di indagare la valutazione del sistema scolastico da una parte nei contesti nazionali e internazionali di riferimento, Italia, Finlandia e Spagna, dall'altra all'interno di sistemi valoriali, sociali, politici e culturali che proprio nelle realtà nazionali la caratterizzano, tenendo conto degli stakeholders coinvolti.

La discussione intorno alla valutazione dei sistemi scolastici negli ultimi decenni ha animato le tavole rotonde degli addetti ai lavori. Facendo propri concetti e costrutti, come la qualità riferita alle istituzioni educative, dai confini non facilmente delineabili (Dell'Anna, 2021), ha dato vita a dibattiti accesi, a volte anche confusi (Corsini, 2015). Il motivo, l'oggetto e le modalità di valutazione sono ciò che differenziano un modello da un altro. Tuttavia, mentre sui perchè e sui cosa ci si sta avviando a un accordo condiviso dalla comunità scientifico-politica, grazie anche alla promulgazione di politiche comunitarie (Allulli, 2010; Colak & Ege, 2013; Arriazu & Solari, 2015; Rogge, 2019; Stumbriene et al, 2020), sul come si valuta sussiste ancora disaccordo (Fondazione Agnelli, 2014). Prendendo le mosse dal quadro teorico esposto, sono emerse le seguenti domande di ricerca:

1. Quali modelli di valutazione del sistema scolastico vengono attuati in Italia, Finlandia e Spagna?
2. Le istituzioni pubbliche come rispondono alla valutazione di sistema?
3. Quale supporto è fornito alle scuole nei processi di autovalutazione?

La prima domanda di ricerca ha l'obiettivo di esplorare e descrivere le pratiche di valutazione del sistema scolastico adoperate in Italia, Finlandia e Spagna. Come avviene la valutazione del sistema scolastico? Quali processi e quali strategie vengono messe in atto? La letteratura sul tema evidenzia che esistono diversi modelli di valutazione: si passa da quelli più centralizzati a quelli decentralizzati, da quelli che si fondano in particolar modo sull'autovalutazione a quelli che invece elargiscono premi o punizioni, da quelli più quantitativi a quelli più qualitativi, da quelli gestiti a livello centrale a quelli gestiti internamente dalle comunità scolastiche (Paletta, 2013). Non esiste un modello che funziona meglio di altri poichè ciascuno di essi va calato all'interno del contesto di riferimento e su portatori di interessi e valori differenti (Bondioli, 2000; Eurydice, 2015).

La seconda domanda di ricerca ha l'obiettivo di esplorare e descrivere la risposta delle istituzioni educative centrali alla valutazione del sistema scolastico. Gli esiti della valutazione vengono adoperati per prendere decisioni? Quale tipo di diffusione viene promossa? Quali percezioni provoca la valutazione di sistema negli stakeholders? I dati provenienti dalla valutazione di sistema, soprattutto se raccolti a livello internazionale, non sempre rispondono necessariamente alle questioni poste dai decisori politici e dagli stakeholders poichè spesso le questioni educative dipendono fortemente dai contesti sociali ed economici nei quali sono immerse. Infatti, quello che è rilevante per un sistema educativo, potrebbe non esserlo per un altro, tanto è vero che l'implementazione di una pratica in un paese sulla base dei risultati osservati in un altro paese non garantisce il successo di una determinata riforma (OECD, 2009a). Nonostante ciò, ai decisori politici è sempre più richiesto di utilizzare gli esiti del monitoraggio e della valutazione per prendere decisioni *evidence-based* sulle quali fondare gli sforzi per il miglioramento. Negli Stati Uniti, ad esempio, dopo la promulgazione dell'atto *No Child Left Behind* del 2001, venne richiesto che tutte le azioni educative messe a punto all'interno di quella iniziativa, prima di essere finanziate, venissero debitamente argomentate da prove, ricerche scientifiche e dati (Honig & Coburn, 2008).

La terza e ultima domanda di ricerca ha l'obiettivo di esplorare e descrivere la situazione dell'autovalutazione degli istituti scolastici in particolar modo in Italia, ma anche in Finlandia e Spagna. A livello centrale sono state messe a punto linee guida per l'autovalutazione? Con particolare riferimento all'Italia, come viene vissuta l'autovalutazione da coloro che sono al contempo destinatari e committenti? Nel rapporto pubblicato nel 2015 dalla rete Eurydice (Eurydice, 2015), l'autovalutazione scolastica era obbligatoria in 27 paesi dell'Unione Europea e consigliata nei restanti ed era ascrivibile all'interno di tre differenti prospettive: il miglioramento, lo sviluppo e la conoscenza e la rendicontazione (Chelimsky & Shadish, 1999). Nella prima prospettiva l'obiettivo dell'autovalutazione è dare all'istituzione scolastica gli strumenti per migliorare autonomamente, nella seconda prospettiva l'obiettivo è accrescere le conoscenze sul funzionamento dell'istituzione scolastica sia nell'organizzazione che nella qualità dei processi di insegnamento e apprendimento, mentre nella terza prospettiva l'obiettivo è fornire dati sull'efficienza e sull'efficacia per rispondere delle risorse investite nell'istruzione (Chelimsky & Shadish, 1999). Si tratta, in questo ultimo caso, di un'autovalutazione che precede la valutazione esterna poichè, avendo una funzione sommativa, si configura come "forma decentrata di rendicontazione" (Scheerens et alii, 2011; Giampietro et alii, 2016). A seconda della prospettiva adottata può essere fornito o meno supporto nei processi di autovalutazione, si pensi alla compilazione del Rapporto di Autovalutazione in Italia o al *Self Evaluation Form* elaborato in Inghilterra (Rosenthal, 2004; Plowright, 2007; Bubb et alii, 2007; McNamara & O'Hara, 2008; Giampietro et alii, 2016; Vincent-Lacrin, 2019).

1.6. Scelte metodologiche: le procedure, i partecipanti e gli strumenti di ricerca

Dopo aver individuato i problemi, i casi oggetto di studio e le domande di ricerca collegate agli obiettivi di quest'ultima, in questo paragrafo verranno illustrate le scelte metodologico-procedurali del disegno di ricerca, verranno presentati i partecipanti dello studio e

verranno descritti gli strumenti adoperati che saranno ripresi ulteriormente nei capitoli due, tre e quattro seguenti.

1.6.1. *Le procedure*

Parlett e Hamilton (1972, 1976) durante i loro lavori di ricerca presso l'Università di Edimburgo hanno sottolineato la necessità di uno studio interpretativo che attraversa tre fasi che a loro volta accompagnano il ricercatore che vuole analizzare i programmi educativi in ambienti complessi. Le tre fasi, osservazione, analisi revisionata, *renewed inquiry*, e argomentazione/spiegazione, sono interconnesse e si sovrappongono le une alle altre. Infatti, il passaggio da una fase all'altra, secondo i due autori, si verifica nel momento in cui i problemi dai quali ha origine la ricerca emergono e diventano progressivamente più chiari e definiti. "Il corso dello studio non può essere tracciato a priori" (Parlett & Hamilton, 1972, 1976), tanto è vero che il ricercatore, partendo da un database ampio, tende a ridurre man mano la portata dell'indagine per dare più enfasi alle problematiche emergenti (Parlett & Hamilton, 1976). Il *progressive focusing* dello studio, la focalizzazione progressiva introdotta dai due autori (1972, 1976), ha guidato le scelte metodologico-procedurali del disegno di ricerca (Parlett, 1976).

Lo studio di caso multiplo e comparativo (Yin, 2009; Gustafsson, 2017) che ha caratterizzato il disegno di ricerca, si è articolato su due livelli (Zanazzi, 2014; Scheerens, 2019). Come già puntualizzato, a livello internazionale gli organismi sovranazionali hanno prodotto e producono documenti e linee guida per la valutazione dei sistemi scolastici. L'analisi a questo primo livello si è concentrata sulla revisione della letteratura e sull'approfondimento dei report, delle politiche e della documentazione pubblicata, come esposto nella prima parte del presente lavoro. A livello nazionale, invece, nei prossimi capitoli verranno approfondite e corroborate le informazioni raccolte attraverso il primo, andando ad esaminare quali sono le pratiche di valutazione di sistema messe in atto in Italia, Finlandia e Spagna. La revisione della bibliografia ha restituito alcuni report prodotti a livello internazionale, si pensi per esempio al capitolo dedicato al "monitoraggio di sistema" nel

documento redatto dal Ministero dell'educazione tedesco in collaborazione con l'OCSE subito dopo le prime edizioni dell'indagine PISA (OECD, 2004) o al report curato da Eurydice e dall'Unione Europea sulle politiche e sugli approcci alla valutazione dei sistemi scolastici in Europa (Eurydice, 2015). Il primo risale all'inizio del nuovo millennio, il 2004, mentre il secondo al 2015 (OECD, 2004; Eurydice, 2015). Da allora alcune cose sono cambiate, ma a livello comunitario e comparativo non sono state più realizzate indagini su questa tematica. Questo è il motivo per il quale ci si è impegnati in un'analisi dei sistemi di valutazione scolastica, pur dovendo circoscrivere per motivi di tempo lo studio a una triade di realtà nazionali. Per l'analisi di quest'ultimo aspetto a livello nazionale è stato utilizzato lo schema proposto da Cardone e colleghi (2010).

I governi dei tre paesi oggetto di studio hanno costruito sistemi di valutazione della scuola che si basano su approcci diversi e si esplicano in processi differenti. Inoltre, le stesse scuole, supportate più o meno dai livelli centrali, realizzano la valutazione interna attraverso modalità differenti. I sistemi di valutazione, calati nei contesti nazionali e istituzionali specifici, riflettono la cultura della valutazione, ma possono configurarsi anche come volani per il cambiamento di quest'ultima (Zanazzi, 2014; Stame, 2018).

1.6.2. I partecipanti

Per rispondere alle domande di ricerca esposte nei paragrafi precedenti e agli obiettivi delineatisi dopo la focalizzazione dei problemi della ricerca, è stato individuato il gruppo dei partecipanti, sulla scorta della rilevanza dei casi oggetto di studio. Dal momento che l'intento della ricerca è esplorare e descrivere le pratiche legate alla valutazione dei sistemi scolastici in Italia, Finlandia e Spagna alla luce dei contesti istituzionali e valoriali all'interno dei quali si inserisce, ovvero del rilievo che è attribuito alla valutazione di sistema in ciascun paese oggetto di analisi, sono state realizzate alcune interviste semi-strutturate, telematiche a causa della pandemia, con alcuni testimoni privilegiati, i National Project Manager dell'indagine PISA dell'OCSE, che sono

coinvolti nella valutazione del sistema scolastico, nello specifico nella valutazione standardizzata del livello degli apprendimenti degli studenti. Inoltre, è stato somministrato un questionario-intervista a un gruppo di dirigenti scolastici italiani per indagare il loro punto di vista sulla valutazione del sistema scolastico, in particolare sull'autovalutazione di istituto nel contesto nazionale.

Ogni paese che partecipa allo studio sui quindicenni dell'OCSE ha un National Project Manager, vale a dire un rappresentante nazionale all'interno delle riunioni degli organi dell'indagine. Gli NPM, lavorando a stretto contatto con il *PISA Governing Board* (PBG), sono responsabili per la realizzazione dello studio e per il rispetto delle tempistiche e delle norme specificate internazionalmente. Inoltre, gli NPM, nel ruolo di ricercatori esperti, sono responsabili non solo dell'implementazione dell'indagine, ma anche dello sviluppo e della validazione degli strumenti di misura internazionali poiché, assolvendo al proprio compito in conformità con gli standard tecnici internazionali, devono garantire che i risultati di PISA siano di alta qualità. Salvo casi eccezionali, gli NPM ricoprono tale delicato ruolo per l'intero ciclo di durata dell'indagine. Essi generalmente hanno competenze elevate nella lingua inglese, tutte le riunioni organizzative si svolgono in inglese, esperienza pregressa nel progettare, organizzare e condurre survey su campioni molto ampi, capacità nel selezionare un team di ricerca composto da esperti con background differenti, familiarità con il controllo della qualità delle indagini e con le procedure di raccolta dei dati. Inoltre, alla luce del loro ruolo di rappresentanza, oltre ad avere conoscenza ed esperienza nel trattare con agenzie governative nazionali e regionali, sanno anche interloquire con il personale scolastico, i genitori e gli studenti, nonché utilizzare pacchetti statistici grazie alle conoscenze metodologico-procedurali relative al campionamento e alla gestione, all'analisi e all'interpretazione dei dati (OECD, 2020d). L'indagine PISA è un "progetto istituzionale" in cui gli attori nazionali giocano un ruolo cruciale nella sua promozione e realizzazione. Ciò non spiega solamente il rapido successo che il programma comparativo ha avuto dall'inizio degli anni 2000 ad oggi, ma anche il motivo per il quale si è espanso, come numero di paesi partecipanti, anche durante i

momenti di crisi economica che gli stati stanno attraversando (Grek, 2012).

La voce dei dirigenti scolastici italiani sull'autovalutazione di istituto non è nuova nell'ambito della ricerca valutativa. Molte sono le indagini che hanno ascoltato le opinioni di questi ultimi. Qualche anno fa, ad esempio, Pastori e Guadagni (2016) conducevano uno studio di caso qualitativo in Lombardia sull'esperienza di dirigenti scolastici, ai quali si univano anche insegnanti e studenti, a proposito della valutazione di sistema. Le due autrici indagavano attraverso interviste e focus group i processi organizzativi, comunicativi, informativi e formativi, nonché il coinvolgimento e le dinamiche relazionali, riflessive e didattiche che vedevano schierati in prima fila dirigenti scolastici, insegnanti e studenti. Le conclusioni dell'indagine mettevano in luce che occorreva una maggiore formazione sulle tematiche valutative, che c'era un'assenza consolidata di familiarità con le prove standardizzate e un'urgenza legata alla diffusione del concetto di valutazione di sistema come uno strumento utile alla ricerca sulla scuola, una riflessione rivolta alla scuola e un'occasione nella scuola per innescare processi riflessivi e autovalutativi, dando luogo non solo ad azioni di rendicontazione esterne, ma anche a processi di sviluppo e miglioramento interni (Pastori & Guadagni, 2016). Qualche anno più tardi Benvenuto e colleghi (2019) indagavano il punto di vista dei dirigenti scolastici sull'elaborazione del Rapporto di Autovalutazione, RAV. I risultati di quella ricerca se da una parte mettevano in luce che la compilazione del RAV era vista come un'occasione di crescita professionale che permetteva di attivare "competenze distribuite", dall'altra non tutti erano concordi sull'adeguatezza delle pratiche di autovalutazione. Emergevano difficoltà nella gestione dei processi di autovalutazione, ma anche la volontà di non gestire questi processi come adempimenti burocratici calati dall'alto (Benvenuto et alii, 2019). Nonostante la letteratura che analizza la figura dei dirigenti scolastici collegata all'autovalutazione delle scuole sia cospicua, sembra che la propensione di questi ultimi nel promuovere attività di sviluppo e valutazione nelle scuole sia la leva propulsiva che muove e sviluppa la cultura dell'autovalutazione negli istituti scolastici (Epifani et alii, 2016).

MacBeath (2008), infatti, sostiene che il ruolo del dirigente scolastico è fondamentale per l'interiorizzazione degli obiettivi dell'autovalutazione da parte del personale scolastico in vista delle azioni di miglioramento. La leadership del dirigente scolastico come motore per la sensibilizzazione alla cultura della valutazione è sottolineata anche nel D.P.R. n. 80 del 2013 che ha introdotto in Italia l'autovalutazione di istituto: in quest'ultimo documento i dirigenti scolastici hanno dovuto assumere un nuovo ruolo a metà tra le istanze di rendicontazione esterne e di sviluppo interno alla scuola. Tuttavia, essendo l'autovalutazione di istituto normalmente centralizzata, la motivazione dei dirigenti scolastici nei confronti di quest'ultima potrebbe venire meno e, dunque, l'autovalutazione delle scuole e l'elaborazione dei RAV potrebbe essere vista come un mero adempimento burocratico (Epifani et alii, 2016). Alla luce di ciò, prendendo in considerazione il "nuovo" ruolo dei dirigenti scolastici all'interno del Sistema Nazionale di Valutazione e tenendo conto del pericolo adempimentale insito nei processi di autovalutazione, per rispondere alla terza domanda di ricerca del presente studio, presentata nel paragrafo precedente (Quale supporto è fornito alle scuole nei processi di autovalutazione?), si è deciso di condurre un'analisi documentale per tutti e tre i paesi oggetto di indagine e di esplorare il punto di vista dei dirigenti scolastici italiani, in qualità di testimoni privilegiati, di occhi e orecchie interne alla scuola, nonché di committenti e destinatari dei processi di autovalutazione di istituto.

1.6.3. Gli strumenti di ricerca

Partendo da un triplice obiettivo di ricerca scaturito dall'individuazione dei problemi e delle domande di ricerca illustrate nel paragrafo precedente - analizzare i modelli e le pratiche di valutazione del sistema scolastico in Italia, Finlandia e Spagna, esplorare il rapporto tra gli esiti della valutazione e la loro diffusione e indagare i processi di autovalutazione delle istituzioni scolastiche nei tre paesi oggetto di studio, con uno sguardo particolare all'Italia - sono stati

predisposti strumenti di rilevazione specifici tenendo conto della multidimensionalità degli oggetti di indagine.

Lo studio di caso si caratterizza per l'assenza di protocolli specifici che delineano le modalità e gli strumenti attraverso i quali raccogliere i dati. Parlett e Hamilton (1972, 1976) introducendo il *progressive focusing* dell'indagine, la progettazione dello studio man mano che si realizza, poi ripresa da Stake (1995), rimarcano quanto lo sviluppo dello studio di caso non possa essere deciso del tutto a priori. Essendo una ricerca di tipo qualitativo, infatti, vi è l'interpretazione dei dati che emergono dalla realtà indagata: l'obiettivo è descrivere realtà complesse prestando attenzione alle informazioni esperienziali e alle analisi di tipo induttivo (Semeraro, 2011). Denzin e Lincoln (2005) rappresentano la ricerca qualitativa come un'attività situata che tiene conto dei contesti specifici e che consiste in un insieme di pratiche osservative che rendono evidenti uno o più fenomeni agli occhi del ricercatore. Di fondo è presente una multidimensionalità delle esperienze che sono altamente situate e che dipendono dal momento storico in cui vengono osservate. L'approccio della ricerca e delle tecniche di raccolta dei dati, infatti, è stato di tipo sociologico-etnografico, riconducibile in parte a quello tipico della *grounded theory*. All'*iterative study design*, la progettazione iterativa dello studio che prevede la raccolta e l'analisi dei dati a cicli in modo tale da poter essere utilizzati come base per lo step successivo, si sono aggiunti il *purposeful sampling* e il *constant comparison approach to data analysis* (Glaser & Strauss, 1967). Il primo ha previsto una selezione mirata delle fonti informative in grado di corroborare, confermare o disconfermare quanto evidenziato in precedenza. Il secondo, invece, ha previsto la comparazione dei "problemi di interesse" tra di loro per rilevarne similitudini e differenze (Glaser & Strauss, 1967; Hutchinson & Wilson, 1986; Strauss & Corbin, 1997; Cutcliffe, 2000; Walker & Myrick, 2006; Oktay, 2012; Allodola, 2014).

Per rispondere alla prima domanda di ricerca (Quali modelli di valutazione del sistema scolastico vengono attuati in Italia, Finlandia e Spagna?) è stata condotta un'analisi documentale delle linee guida internazionali, dei siti web istituzionali, dei report nazionali e della letteratura pubblicata, come riportato nei capitoli di apertura del

presente lavoro. Il primo strumento e tecnica di ricerca, lo studio bibliografico e reportistico, ha permesso di comparare i sistemi educativi tra di loro, prendendo a riferimento alcuni indicatori che verranno esplicitati nel secondo capitolo della seconda parte, e di analizzare le pratiche legate alla valutazione dei sistemi scolastici, una duplice modalità di valutazione a cui le scuole si sottopongono e vengono sottoposte (Giampietro et alii, 2016). I documenti, istituzionali, scientifici e di divulgazione, fanno parte del tessuto del nostro mondo e la loro quantità è cresciuta nel tempo, man mano che gli strumenti di scrittura e di diffusione cartacea e telematica si sono affinati (Love, 2013). Come sottolineato da Stake (1995), l'analisi documentale si configura come un importante mezzo informativo dello studio di caso poiché non richiede il coinvolgimento dei soggetti partecipanti alla ricerca ed è una fonte di dati non intrusiva. Inoltre, proprio perché i dati possono essere raccolti direttamente dai documenti, l'analisi documentale qualitativa, *Qualitative Document Analysis* (QDA), permette di ricostruire i significati dati all'oggetto di studio nel sistema valoriale del contesto di riferimento all'interno del quale quest'ultimo si colloca (Altheide et alii, 2008; Bowen, 2009; Sena, 2016, 2021; De Santis, 2022). Per questo motivo, quest'ultima è spesso utilizzata nelle scienze sociali per facilitare lo studio imparziale e coerente dei documenti scritti poiché il contenuto di questi ultimi può favorire l'approfondimento, la comprensione dei problemi educativi sia nella pratica della ricerca che nella teoria (Wach & Ward, 2013; Owen, 2014; Cardno, 2018).

Per rispondere alla seconda domanda di ricerca (Le istituzioni pubbliche come rispondono alla valutazione di sistema?) sono state svolte interviste semi-strutturate telematiche, a causa del momento emergenziale e pandemico che ha caratterizzato una buona parte del percorso di dottorato, rivolte ai National Project Manager dell'indagine PISA dell'OCSE. I dati raccolti attraverso il secondo strumento e tecnica di ricerca, l'incontro telematico mediato da uno schermo con i responsabili nazionali dell'indagine dell'OCSE sui quindicenni, ha permesso di corroborare e approfondire le informazioni acquisite attraverso lo studio documentale. Nel capitolo tre della seconda parte si entrerà nel merito della costruzione dello strumento, delle aree di

interesse indagate e dei dati raccolti. Si è deciso di ricorrere all'utilizzo dell'intervista poiché quest'ultima permette di acquisire informazioni dettagliate sugli aspetti chiave dell'indagine favorendo il dialogo e la costruzione partecipata del significato dell'oggetto di studio tra l'intervistato e l'intervistatore (deMarrais, 2003, 2004). Le interviste, dal momento che richiedono lo scambio di codici verbali tra i partecipanti alla discussione, possono essere di diverse tipologie (Zammuner, 1998). Quella utilizzata nel lavoro qui presentato è di tipo semi-strutturato.

Un terzo tipo di intervista è quella semi-strutturata in cui il ricercatore predispone una serie di domande su cui desidera ricevere risposta da parte dell'intervistato. Vengono mantenuti tuttavia anche criteri di flessibilità, nel senso che il ricercatore può utilizzare ulteriori domande rispetto a quelle previste, aumentando la quantità e qualità del materiale prodotto dall'intervistato. I dati raccolti con le interviste completamente strutturate possono essere analizzati tramite criteri di tipo quantitativo (individuando frequenze, correlazioni, regolarità, fatti oggettivi, ecc.), mentre i dati raccolti tramite le interviste libere o semi-strutturate possono essere analizzati secondo criteri di tipo qualitativo (ad esempio considerando modalità di spiegazione, di selezione dei fatti e di interpretazione degli stessi, usate dagli intervistati). Anche in questo secondo caso l'analisi dei dati si fonda su un contenuto di tipo linguistico, come nel caso dei materiali osservativi sopra segnalati. Le interviste vanno in genere registrate, previo assenso degli intervistati, e i dati raccolti, a partire dalla trascrizione delle risposte, possono essere analizzati tramite l'analisi del contenuto, da cui far emergere (come si vedrà nel prosieguo di questo saggio) i nuclei di significato presenti nel testo delle risposte degli intervistati. (Semeraro, 2011, p. 102)

Secondo deMarrais (2004) il rapporto faccia a faccia che si crea durante l'intervista permette di entrare in contatto con l'esperienza vissuta (De Santis, 2022).

Il mondo dell'esperienza vissuta non sempre corrisponde al mondo della descrizione oggettiva perché l'oggettività spesso implica il tentativo di spiegare un evento come separato dal suo contesto. Piuttosto che separare e poi oggettivare aspetti del mondo della vita, lo scopo è descrivere l'esperienza umana così come viene vissuta. In quest'ottica, il significato di una esperienza è

sempre situato nel contesto esperienziale attuale ed è coerentemente correlato al progetto in corso del mondo della vita. (deMarrais, 2004, p. 56)

Per rispondere alla terza domanda di ricerca (Quale supporto è fornito alle scuole nei processi di autovalutazione?) è stato somministrato un questionario-intervista ai dirigenti scolastici italiani, attori e protagonisti nel contesto nazionale dei processi di autovalutazione delle istituzioni scolastiche, ed è stata condotta un'analisi documentale per tutti e tre i paesi oggetto di studio. Il terzo strumento e tecnica di ricerca è stato somministrato sia telematicamente tramite Google Moduli che telefonicamente a chi si è reso disponibile: la doppia metodologia di somministrazione è stata individuata alla luce del sovraccarico di impegni che caratterizza il lavoro dei dirigenti scolastici, acuito ancora di più dalla situazione pandemica. Vista la ripresa delle attività didattiche non più a distanza e visto il perdurare delle condizioni emergenziali, per non gravare ancora di più su una situazione già delicata, è stata data la possibilità di rispondere al questionario sia in forma scritta che in forma orale attraverso un'intervista telefonica nel pieno rispetto delle persone coinvolte (Bassey, 1999). Il capitolo quattro della seconda parte descriverà le fasi di costruzione dello strumento, messo a punto ad hoc con stimoli chiusi e stimoli aperti, il tracciato degli item e le evidenze emerse. Se da una parte l'intervista telefonica permette di raccogliere dati in modo approfondito poiché si tratta sempre di un percorso condiviso di co-costruzione di significato, come descritto poco più sopra, dall'altra il questionario, modalità di partecipazione più scelta dal gruppo di dirigenti scolastici, permette di ampliare la quantità di dati raccolti grazie alla sua agile diffusione (Lombi, 2015). Coggi e Ricchiardi (2005) descrivono il questionario come una serie ordinata di domande.

Una lista organizzata di domande che vengono poste per iscritto, nelle stesse condizioni, a un gruppo abitualmente ampio di soggetti allo scopo di raccogliere informazioni, di conoscere opinioni, atteggiamenti, intenzioni e azioni compiute. (Coggi & Ricchiardi, 2005, p. 84)

Quest'ultimo, oltre a permettere la rilevazione delle opinioni dei partecipanti, si caratterizza anche per la semplicità di inserimento della risposta da parte del rispondente e per l'economicità e rapidità di somministrazione (Trincherò, 2002; Corbetta, 2003; Coggi & Ricchiardi, 2005; Lucisano & Salerni, 2009).

Ovviamente anche il questionario si presta a una varietà di forme, per offrire anch'esso la possibilità di indagare ed esplorare in modo più descrittivo e qualitativo determinate realtà, introducendo domande meno strutturate e possibilità di scelte aperte. Ma in questi casi si perderebbe lo specifico vantaggio della velocità di elaborazione e di analisi statistica, richiedendo quindi specifiche analisi più categoriali e/o descrittive delle risposte aperte. (Benvenuto, 2015, p. 247)

Si tratta dunque di uno strumento che, nonostante metta il soggetto nelle condizioni di non poter approfondire dettagliatamente alcune risposte date, presenta la possibilità di inserire domande poco strutturate che colgono le esperienze e le percezioni degli intervistati, coloro che rispondono alle domande scritte del questionario, su uno specifico oggetto di studio e non prevedono risposte predeterminate (Zammuner, 1996, 1998). Nonostante questo strumento sia stato aggiunto al sopraggiungere del diniego alla partecipazione alla ricerca da parte dei National Project Manager dell'indagine PISA dell'OCSE di alcuni paesi europei, di cui si darà conto nelle conclusioni, rifiuto causato a sua volta dal sopraggiungere della pandemia, dal posticipo di alcune indagini standardizzate sugli apprendimenti al 2022 e dal conseguente incremento del carico di lavoro per i responsabili nazionali dell'indagine PISA, ha permesso di indagare più in profondità il punto di vista di soggetti attivi nella scuola, i dirigenti scolastici, e coinvolti nei processi di autovalutazione delle istituzioni scolastiche. L'affondo su quest'ultima tematica, infatti, non era previsto nel disegno iniziale del progetto di ricerca dottorale ed è stato aggiunto quando i contatti con altre realtà europee sono venuti meno e gli spostamenti fisici sono stati limitati per motivi di sicurezza. Pur riconoscendo tutte le difficoltà legate al momento emergenziale che ha caratterizzato gli ultimi anni, lo

spin-off imprevisto creatosi sulle pratiche e sulle percezioni legate all'autovalutazione di istituto di un gruppo di dirigenti scolastici italiani si è rivelato prezioso e arricchente sia rispetto alle forme che la valutazione dei sistemi educativi può assumere (Vertecchi, 2003), sia rispetto alla triangolazione delle fonti, soprattutto in ambito nazionale (Denzin, 1970).

Per la descrizione della costruzione, della somministrazione, delle evidenze emerse e degli aspetti metodologico-procedurali che caratterizzano ciascun strumento e tecnica impiegata si rimanda alle rispettive note metodologiche presenti in apertura dei capitoli due, tre e quattro della seconda parte.

1.6.4. La timeline del progetto di ricerca

Di seguito si riporta la scansione temporale che ha accompagnato le varie tappe del progetto di ricerca. Il diagramma di Gantt riportato nella tabella 1 descrive le fasi che hanno caratterizzato il lavoro degli ultimi tre anni circa. Va precisato che, sfortunatamente, a poco più di quattro mesi dall'inizio del dottorato sono stati registrati i primi casi della patologia infettiva causata da SARS-CoV-2. Questo ha comportato l'interruzione dei seminari, dei convegni e degli incontri periodici di dottorato in presenza e l'inizio del primo *lockdown*. Contestualmente anche l'attività di ricerca in presenza ha subito una battuta di arresto che si è protratta anche dopo l'estate del 2020 per il perdurare di situazioni sanitarie potenzialmente pericolose e si è tramutata nel sistema a colori per il monitoraggio della pandemia, ovvero la classificazione delle fasce di rischio. Quest'ultima ha portato con sé le limitazioni agli incontri faccia a faccia e agli spostamenti che tutti tristemente conosciamo. Solamente durante l'ultimo anno accademico di dottorato, quello destinato all'analisi e all'interpretazione dei dati raccolti e alla stesura del lavoro finale, si è potuti tornare ad incontrarsi di persona rispettando le indicazioni operative per prevenire il diffondersi della malattia e del contagio.

	2020		2021		2022		
Fasi	Inverno-primavera	Estate-autunno	Inverno-primavera	Estate-autunno	Inverno-primavera	Estate-autunno	Proroga (3 mesi)
Analisi della letteratura	■	■	■	■	■	■	
Contatti per la definizione dei panel di lavoro		■					
Definizione e costruzione degli strumenti di ricerca		■	■	■			
Somministrazione degli strumenti agli stakeholders			■	■	■	■	
Analisi dei dati raccolti e interpretazione dei risultati				■	■	■	■
Stesura dell'elaborato finale					■	■	■

Tabella 1. Diagramma di Gantt riepilogativo sulla timeline della ricerca dottorale. Fonte: elaborazione propria.

Capitolo secondo

L'analisi documentale qualitativa

2.1. L'analisi documentale come metodo di ricerca

Da qualche anno a questa parte i documenti organizzativi e istituzionali sono un punto fermo nella ricerca qualitativa, infatti sempre più paper pubblicati menzionano l'analisi documentale come parte della metodologia, spesso utilizzata quando vengono condotti studi di caso qualitativi (Bowen, 2009). L'analisi documentale, infatti, è una procedura per la revisione sistematica dei documenti e dei report prodotti sia in formato cartaceo che digitale. Alla base vi è un esame e un'interpretazione dei dati in essi contenuti affinché questi ultimi generino significati condivisi, supportino lo sviluppo della consapevolezza e l'acquisizione di conoscenza (Corbin & Strauss, 2011). Atkinson e Coffey (1997) definiscono i documenti come "fatti sociali" prodotti, condivisi e utilizzati in modi socialmente organizzati. Si tratta di scritti di vario genere, dai documenti istituzionali, ai diari personali, ai quotidiani e alle riviste, che vengono analizzati in una procedura analitica che implica la selezione dei documenti oggetto di interesse, la valutazione di questi ultimi e la sintesi dei dati in essi contenuti. Di questo processo si darà conto nei paragrafi successivi di questo capitolo nel quale verranno riportate le risultanze dell'analisi documentale svolta. Sulla stregua della procedura individuata da Bowen (2009) verranno presentati, infatti, dati, estratti e citazioni organizzati in temi principali restituiti attraverso le domande che hanno guidato il processo di analisi (Losito, 1996; Labuschagne, 2003; Pagani, 2020).

Per l'analisi dei documenti presi in considerazione in questo studio (compresi quelli elettronici, i siti web istituzionali ad esempio) è stata condotta un'analisi dei contenuti in essi riportati: le informazioni sono state dapprima vagliate, poi, se pertinenti, organizzate in categorie rispondenti alle domande di ricerca illustrate nel capitolo precedente. Allo *skimming*, un esame superficiale dei documenti, è seguito un esame più approfondito con annessa interpretazione e valutazione del testo

scritto (Bowen, 2009). Nonostante le critiche mosse all'analisi dei contenuti, quest'ultima prende in considerazione i processi interpretativi poichè non è volta alla quantificazione, tendenza tipica dell'analisi del contenuto condotta dai mass media (Silverman, 2000). Nello specifico del lavoro qui presentato, sono stati dapprima letti i documenti in cui sono stati identificati dati o passaggi rilevanti ai fini degli obiettivi di ricerca e sono state distinte le informazioni pertinenti da quelle non pertinenti (Strauss & Corbin, 1998; Corbin & Strauss, 2008). Successivamente sono state individuate alcune tematiche emergenti all'interno dei documenti che sono diventate le categorie di riferimento per l'analisi di quest'ultimi e le domande guida per la restituzione delle risultanze (es. struttura del sistema educativo, modello di valutazione del sistema scolastico, legislazione nazionale etc.). A partire dal modello di categorizzazione si è poi proceduto con una rilettura più attenta, mirata e approfondita dei documenti, con l'individuazione delle tematiche pertinenti e con la presentazione dei risultati (Fereday & Muir-Cochrane, 2006).

Sebbene i documenti possano essere una ricca fonte di dati, i ricercatori dovrebbero esaminarli con occhio critico ed essere cauti nell'usarli nei loro studi. I documenti non devono essere trattati come registrazioni necessariamente precise, accurate o complete di eventi che si sono verificati. I ricercatori non dovrebbero semplicemente "riportare" parole e passaggi dai documenti disponibili nel loro rapporto di ricerca. Piuttosto, dovrebbero stabilire il significato del documento e il suo contributo alle questioni in esame. Il ricercatore, in qualità di analista, dovrebbe determinare la rilevanza dei documenti rispetto al problema e allo scopo della ricerca. Inoltre, il ricercatore dovrebbe verificare se il contenuto dei documenti sia adatto alla struttura concettuale dello studio. È inoltre necessario determinare l'autenticità, la credibilità, l'accuratezza e la rappresentatività dei documenti selezionati. [...] Nei casi in cui vengono utilizzati come fonte di dati gli studi precedenti, è importante andare oltre i risultati della ricerca. L'analista dovrebbe identificare il disegno di ricerca (trasversale o longitudinale), i metodi e gli strumenti, nonché il quadro teorico di ogni studio. "Quanti documenti dovrei analizzare?" Sento il ricercatore principiante fare questa domanda. Sebbene sia una domanda legittima, la preoccupazione non dovrebbe riguardare "quanti"; piuttosto, dovrebbe riguardare la qualità dei documenti e i dati che contengono,

alla luce dello scopo e del disegno dello studio. [...] L'analisi del documento, quindi, non consiste nel mettere insieme una serie di estratti da materiale stampato per trasmettere qualunque idea venga in mente al ricercatore. Si tratta di un processo di valutazione dei documenti che ha l'obiettivo di produrre conoscenza empirica e sviluppare la comprensione. In questo processo il ricercatore dovrebbe lottare affinché venga garantita l'obiettività e la sensibilità e mantenere l'equilibrio tra entrambi. (Bowen, 2009, p. 33-34, traduzione a cura dell'autrice)

I documenti possono essere utili per obiettivi di ricerca differenti, Bowen (2009) ne individua cinque. La prima funzione è quella di fornire dati sul contesto in cui ha luogo la ricerca e in cui i partecipanti si trovano ad operare. Queste informazioni possono aiutare il ricercatore a comprendere le ragioni storiche e i problemi attuali che possono incidere sul fenomeno oggetto di studio, nonché possono essere un valido supporto sia per individuare a monte le problematiche dalle quali scaturiscono i bisogni di ricerca, sia per contestualizzare i dati raccolti durante le interviste. La seconda funzione è quella di suggerire domande di ricerca che non erano emerse in una prima fase di problematizzazione e progettazione (Goldstein & Reibold, 2004). La terza funzione è quella di fornire dati "supplementari" per la ricerca, infatti le informazioni evinte dai documenti possono configurarsi come un'ulteriore e preziosa base di conoscenza (Hoepfl, 1997). La quarta funzione è quella di essere una memoria storica per tenere traccia dei cambiamenti e delle innovazioni. Se, ad esempio, sono accessibili varie bozze di un documento, il ricercatore le può confrontare per comprendere quali cambiamenti ci sono stati tra l'una e l'altra (Yin, 1994ab). La quinta funzione è quella di rendere possibile la triangolazione delle fonti. I documenti possono essere utilizzati per verificare i risultati e corroborare le informazioni raccolte da altre fonti. Qualora la triangolazione delle fonti e delle informazioni raccolte dovesse risultare contraddittoria e lontano dalla corroborazione, allora è bene per il ricercatore indagare ulteriormente. Al contrario, quando le informazioni provenienti da fonti differenti convergono, allora c'è maggiore fiducia e attendibilità dei risultati (Angrosino & Mays de Perez, 2000). Come ogni metodo di ricerca, anche l'analisi documentale

presenta punti di forza e di debolezza. Tra i primi rientrano l'efficienza poichè l'analisi dei documenti richiede meno tempo della raccolta di altri tipi di dati; la disponibilità poichè con l'avvento di internet molti documenti sono diventati di dominio pubblico dunque possono essere fruiti senza chiedere il consenso di qualcuno (Merriam, 1988); l'economicità poichè l'analisi dei documenti non richiede investimenti economici rispetto ad altri metodi di raccolta dei dati, questi ultimi nel caso specifico sono già stati raccolti e sistematizzati all'interno di report, di conseguenza ciò che rimane da fare è l'analisi del contenuto; mancanza di reattività poichè i documenti, essendo già redatti, non sono influenzati dal processo di ricerca come può accadere, ad esempio, quando si decide di ricorrere all'osservazione del fenomeno indagato; stabilità poichè non essendo reattivi, i documenti sono stabili nel tempo ovvero non sono alterati dalla presenza del ricercatore e possono essere revisionati più volte (Merriam, 1988); esattezza poichè i documenti contengono al loro interno informazioni, quali nomi e date precise e puntuali (Yin, 1994ab); copertura poichè i documenti possono fornire informazioni su un lasso di tempo prolungato o su un fenomeno vasto (Yin, 1994ab). I punti di debolezza, invece, possono essere rintracciati nella insufficienza di dettagli: nei documenti non è detto che siano sempre presenti le informazioni di cui si ha bisogno dal momento che sono stati prodotti per scopi diversi da quelli della ricerca. Di conseguenza, qualora si voglia rispondere a una domanda di ricerca attraverso l'analisi documentale, è bene corroborare sempre i dati raccolti con altre fonti di dati; nella difficoltà nella recuperabilità: non sempre i documenti sono disponibili, l'accesso potrebbe risultare bloccato e il recupero di documenti datati potrebbe risultare difficoltoso (Yin, 1994ab); nella selettività parziale: i documenti resi pubblici potrebbero essere stati selezionati a priori in base alle politiche e alle procedure aziendali/nazionali oppure potrebbero riflettere le linee di azione intraprese da chi ha in mano il potere in quel contesto e in quel momento, per esempio (Yin, 1994ab; Bowen, 2009).

Nel caso specifico della ricerca dottorale qui presentata l'analisi dei documenti prodotti a livello nazionale e internazionale è utilizzata da una parte per descrivere preliminarmente i sistemi educativi e valutativi

dei paesi coinvolti, Italia, Finlandia e Spagna, dall'altra per corroborare e comprendere meglio i dati raccolti attraverso la somministrazione delle interviste ai National Project Manager dell'indagine PISA dell'OCSE e la somministrazione del questionario ai dirigenti scolastici italiani. I documenti considerati durante l'analisi della bibliografia, infatti, hanno permesso di esplorare e descrivere i contesti nazionali di riferimento, in alcuni casi hanno sollevato ulteriori domande e interrogativi di ricerca che, attraverso il *progressive focusing* (Parlett & Hamilton, 1976), sono stati tenuti in considerazione nelle fasi successive dell'indagine, hanno permesso di ricostruire il cambiamento su una linea del tempo diacronica, infine hanno permesso di corroborare i dati raccolti attraverso le interviste e il questionario accompagnando sia prima, che dopo e durante la costruzione degli strumenti di ricerca e la raccolta dei dati, come già ricordato.

Nei paragrafi che seguono verranno riportati i risultati emersi dall'analisi documentale qualitativa seguendo sia gli approcci propri dello studio di caso, sia quelli propri della comparazione di realtà nazionali differenti (Bray & Thomas, 1955). In prima istanza verranno analizzate le strutture dei tre sistemi educativi nel loro complesso, successivamente verrà fatto un focus sugli atti legislativi riguardanti la valutazione scolastica e verranno delineati i compiti e le funzioni degli istituti di ricerca valutativa presenti nei tre paesi oggetto di studio. Per completare l'analisi verranno riportati gli aspetti indagati da questi ultimi, con particolare rimando alle indagini nazionali sugli apprendimenti, dove presenti, e verrà esplorata la valutazione delle scuole intesa come valutazione esterna ed autovalutazione. Le analisi si basano su un insieme di variabili identificate come significative rispetto alle domande e agli obiettivi della ricerca. Il modello utilizzato per la restituzione delle risultanze dell'analisi documentale è quello proposto da Cardone e colleghi (2010). In ogni paese oggetto di analisi ci si è posti alcune domande relative alla valutazione del sistema educativo. Chi/Quale ente o istituto è chiamato ad occuparsi della valutazione del sistema scolastico? Perché è necessaria la valutazione del sistema scolastico? Qual è lo scopo e l'oggetto della valutazione? Come avviene la valutazione delle scuole? Quest'ultima domanda, oltre ad avere

guidato il processo di restituzione dell'analisi documentale è anche uno degli interrogativi di ricerca descritti nel primo capitolo della seconda parte. Al modello di Cardone e colleghi (2010), si è aggiunta un'ulteriore questione: dove avviene la valutazione del sistema scolastico? Per rispondere a quest'ultima domanda e per contestualizzare meglio la valutazione in ciascun paese, si è deciso di procedere in primo luogo dalla descrizione del sistema educativo di riferimento, riportando, in secondo luogo, per ciascuna domanda del modello i dati riguardanti ciascuno studio di caso così da favorire la comparazione. Le evidenze emerse tratteggiano i contorni della valutazione del sistema scolastico così come è intesa dalle autorità centrali e dai documenti da esse diffusi. L'istantanea scattata attraverso l'analisi della documentazione prodotta in ciascun paese mostra un panorama variegato che verrà discusso più ampiamente nelle conclusioni del presente lavoro.

Per facilitare la consultazione, l'aggiornamento e la replicabilità delle comparazioni verranno riportati i rimandi ai documenti e ai siti web istituzionali consultati, alcuni dei quali in lingua inglese e finlandese.

2.2. In quali contesti avviene la valutazione di sistema?

La conoscenza della struttura di un sistema scolastico è importante perché, come ampiamente riconosciuto, da qualche decennio a questa parte l'educazione si è configurata come motore per la mobilità sociale. L'accesso all'istruzione, la permanenza all'interno di quest'ultima e la possibilità di personalizzazione del percorso educativo sono da molti ritenuti indicatori del benessere sociale e collettivo (Checci & Zollino, 2001). Bowen (2009) riconosce che i documenti possono essere utilizzati per obiettivi differenti, uno di questi ruota attorno alla possibilità di fornire informazioni sul contesto in cui ha luogo la ricerca. Per questo motivo si è pensato di avviare l'analisi documentale dallo studio della struttura dei sistemi educativi in Finlandia, Spagna e Italia, per meglio contestualizzare i dati raccolti non solo attraverso l'analisi documentale, ma anche attraverso le interviste e il questionario. Per tale analisi sono stati consultati i report comparativi pubblicati annualmente da Eurydice sulla struttura dei sistemi educativi europei, i siti web dei ministeri

dell'educazione di ciascun paese¹¹ e l'ultima edizione disponibile degli indicatori sintetizzati nella raccolta dell'OCSE intitolata *Education at a Glance* (Eurydice, 2022; OECD, 2022).

2.2.1. Il sistema educativo finlandese

L'articolazione del sistema educativo finlandese prevede un primo segmento dedicato all'educazione della prima infanzia e all'educazione pre-primaria, un secondo segmento dedicato all'educazione primaria e secondaria ed un terzo segmento dedicato all'educazione terziaria. Nella figura 4 che segue ne è riportata la struttura.

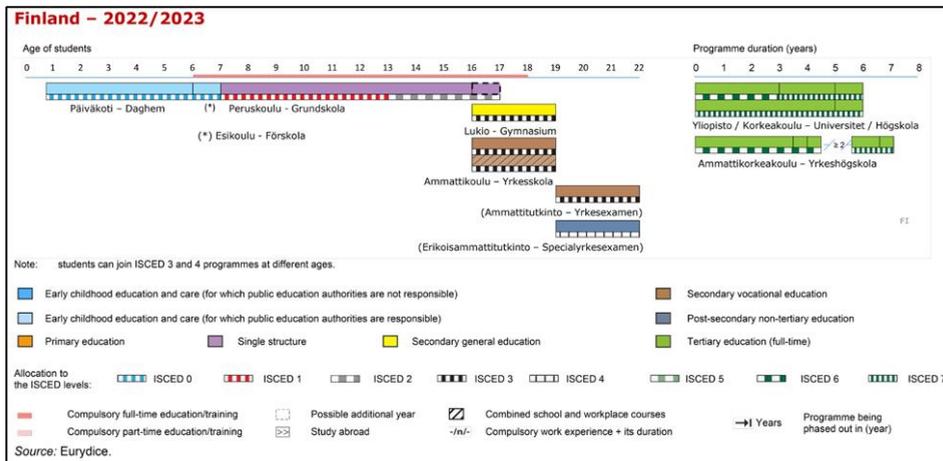


Figura 4. Struttura del sistema scolastico finlandese. Fonte: Eurydice, 2022.

L'educazione della prima infanzia, *varhaiskasvatus*, accoglie i bambini tra 0 e 6 anni ed è un diritto universalmente riconosciuto. Comporta un contributo economico minimo da parte delle famiglie ed è organizzata

¹¹ Per approfondimenti circa il Ministero dell'Educazione e della Cultura finlandese: <https://okm.fi/etusivu> (ultima consultazione 21 febbraio 2023).

Per approfondimenti circa il Ministero dell'Educazione e della Formazione professionale spagnolo: <https://www.educacionyfp.gob.es/portada.html> (ultima consultazione 21 febbraio 2023)

Per approfondimenti circa il Ministero dell'Istruzione e del Merito italiano: <https://www.miur.gov.it/> (ultima consultazione 21 febbraio 2023)

all'interno di strutture quali gli asili nido e le scuole dell'infanzia, *päiväkoti*, o in alternativa nel *family day care*, le strutture di educazione parentale. Tra il 2010 e il 2018 la percentuale di bambini frequentanti le strutture dedicate alla prima infanzia è aumentata di 16.2 punti percentuali, tanto è vero che nel 2019 i bambini inseriti all'interno dei servizi educativi per la prima infanzia erano l'83.9%, percentuale, tuttavia, non sufficiente per soddisfare il benchmark stabilito nella strategia Europa 2020 che prevedeva che almeno il 95% di loro frequentasse servizi dedicati (Allulli, 2010; Pasimeni, 2013)¹².

L'educazione pre-primaria, *esiopetus*, si articola in un anno di pre-istruzione obbligatoria prima di iniziare la scuola primaria all'età di 7 anni. È completamente gratuita ed è organizzata dalle autorità locali che hanno la responsabilità di garantirla a tutti i bambini. Una giornata tipo nella scuola pre-primaria è suddivisa in due momenti: nella prima parte i bambini svolgono attività in preparazione alla scuola primaria, nella seconda parte si dedicano ad attività più ludiche tipiche della scuola dell'infanzia.

L'educazione di base, *perusopetus*, racchiude al suo interno sia l'educazione primaria che l'educazione secondaria di primo grado. Da 7 a 16 anni gli studenti frequentano la scuola di base che si divide in nove annualità. È anche essa completamente gratuita per le famiglie a partire dai testi scolastici, fino ad arrivare ai pasti nella mensa, alla cancelleria e ai trasporti, qualora necessario, per raggiungere la scuola dalla propria abitazione.

L'educazione secondaria superiore, *toisen asteen koulutus*, invece, si divide in liceale, *lukiokoulutus*, e professionale, *ammattillinen koulutus*, e ha una durata di tre anni (da 16 a 19 anni). In particolar modo la scuola secondaria superiore a indirizzo professionale spesso accoglie anche studenti adulti che si inseriscono nuovamente nel contesto formativo per

¹² Per approfondimenti circa l'Education and Training Monitor 2020 (Finlandia): https://op.europa.eu/webpub/eac/education-and-training-monitor-2020/countries/finland_fi.html#five (ultima consultazione 21 febbraio 2023)

ottenere una riqualificazione professionale (Vertanen, 2002; Paaso, 2012; Niemi & Jahnujainen, 2018; Nokelainen et alii, 2018).

L'educazione terziaria, come l'educazione secondaria superiore, si divide in universitaria, *yliopisto*, e delle scienze applicate, *ammattikorkeakoulu*. La differenza tra le due risiede nel fatto che la prima è più orientata accademicamente mentre la seconda è più professionalizzante, si pensi ad esempio ai politecnici corrispettivi italiani (Vuorinen & Valkonen, 2005; Eurydice, 2022). Per chi conclude la scuola secondaria superiore la non iscrizione diretta a un corso di studi terziario è abbastanza comune. La tendenza a prendersi un "anno sabbatico" è dovuta alla selettività dei test di accesso alle università. Non è particolarmente raro, infatti, che uno studente debba tentare l'ingresso a un corso di laurea più volte prima di riuscirci, questo spiega come mai meno del 30% degli studenti transitano direttamente dall'istruzione secondaria superiore alle università o ai politecnici. Per arginare questo gap negli ultimi anni il governo finlandese ha varato una serie di misure che cercano di ridurre il numero di anni trascorsi tra il conseguimento del diploma di scuola secondaria superiore e l'ingresso all'istruzione terziaria. La prima di queste prevede l'attribuzione di un maggior peso ai diplomi di scuola secondaria superiore nella selezione degli studenti universitari. La seconda, invece, aumenta il numero di posti disponibili in ciascun corso di studi (OECD, 2022).

2.2.2. Il sistema educativo spagnolo

L'articolazione del sistema educativo spagnolo prevede un primo segmento dedicato all'educazione della prima infanzia e all'educazione primaria, un secondo segmento dedicato all'educazione secondaria ed un terzo segmento dedicato all'educazione terziaria. Nella figura 5 che segue ne è riportata la struttura.

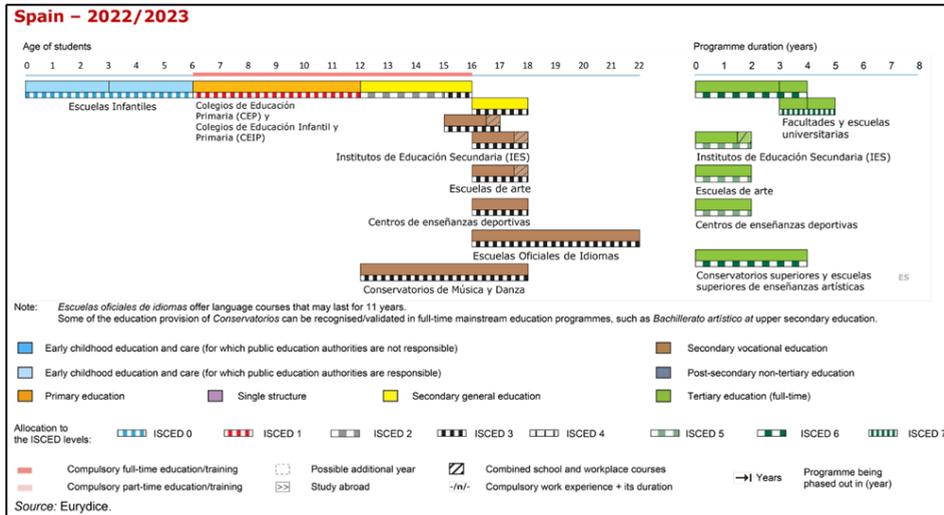


Figura 5. Struttura del sistema scolastico spagnolo. Fonte: Eurydice, 2022.

L'educazione dell'infanzia, *escuelas infantiles*, accoglie i bambini tra 0 e 6 anni. Pur non essendo obbligatoria, è gratuita poiché finanziata con i fondi pubblici. Similmente all'Italia, ci sono sia istituzioni che si concentrano solamente sull'educazione dell'infanzia, *escuelas infantiles*, sia istituzioni che racchiudono al loro interno la scuola dell'infanzia e la scuola primaria, *colegios de educacion infantil y primaria*. Nel 2019, nonostante persistessero differenze a livello regionale e nelle comunità autonome, il 98% dei bambini di 4 anni frequentava la scuola dell'infanzia, mentre il 50% dei bambini al di sotto dei 3 anni frequentava i nidi¹³.

L'educazione di base è obbligatoria e gratuita nelle istituzioni pubbliche e ha una durata di dieci anni. Si suddivide in educazione primaria, *colegios de educacion primaria*, tra 6 e 12 anni, ed educazione secondaria di primo grado obbligatoria, *institutos de educacion secundaria*, tra 12 e 16 anni, alla fine della quale lo studente riceve il diploma

¹³ Per approfondimenti circa l'Education and Training Monitor 2020 (Spagna): <https://op.europa.eu/webpub/eac/education-and-training-monitor-2020/countries/spain.html#five> (ultima consultazione 21 febbraio 2023)

conclusivo della scuola secondaria di primo grado, *Graduado en Educación Secundaria Obligatoria*. L'educazione secondaria superiore ha la durata complessiva di due anni, tra 16 e 18 anni. Si divide in liceale, *bachillerato*, e professionale, *formacion profesional*, erogata anche dai centri di formazione nazionali (Eurydice, 2022). Nel 2020 il 65% dei giovani diplomati, con un'età media di 18 anni, proveniva dagli istituti liceali, mentre il 35%, con un'età media di 24 anni, dalla formazione professionale. L'età media dei diplomati provenienti dagli istituti liceali non differisce molto da quella della media degli altri paesi (19 anni in media OCSE), mentre quella dei diplomati provenienti dalla formazione professionale si. Questo è spiegato dal fatto che in Spagna, così come in altri paesi (ad esempio Canada e Nuova Zelanda), mentre il diploma liceale si configura come conclusione della prima fase di scolarizzazione, la qualifica tecnico-professionale può anche essere un seconda opportunità, similmente alla Finlandia, per coloro che hanno abbandonato la scuola per muoversi verso il mondo del lavoro senza aver conseguito il titolo di studio secondario superiore, consentendogli di ottenerlo per migliorare le prospettive occupazionali (OECD, 2022).

L'educazione terziaria, come l'educazione secondaria superiore, si divide in universitaria-accademica, *facultades y escuelas universitarias*, e in istituti di formazione professionale di livello avanzato, *institutos de educacion secundaria* (Eurydice, 2022). La formazione professionale, sia di livello secondario che terziario, e l'utilizzo delle tecnologie a scuola sono negli ultimi tempi oggetto di dibattito poiché il sistema educativo spagnolo sta subendo un rinnovamento a seguito della nuova riforma dell'educazione (Alcaide, 2021; Casadella et alii, 2022; Roldan & Gimenez, 2022).

2.2.3. Il sistema educativo italiano

L'articolazione del sistema educativo italiano prevede un primo segmento dedicato all'educazione dell'infanzia, un secondo segmento dedicato all'educazione primaria e all'educazione secondaria di primo grado, un terzo segmento dedicato all'educazione secondaria di secondo

grado e un quarto segmento dedicato all'educazione terziaria. Nella figura 6 che segue ne è riportata la struttura.

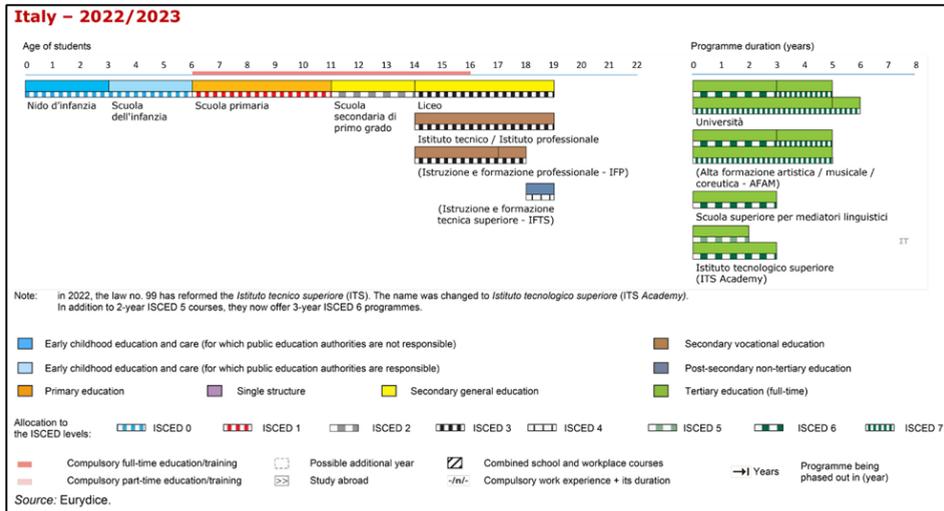


Figura 6. Struttura del sistema scolastico italiano. Fonte: Eurydice, 2022.

L'educazione dell'infanzia, il sistema integrato 0-6, non è obbligatoria e comprende al suo interno i servizi educativi per l'infanzia, gestiti da istituzioni pubbliche e private e rivolti ai bambini tra 0 e 36 mesi, e le scuole per l'infanzia rivolte ai bambini tra 3 e 6 anni. Nonostante il sistema integrato 0-6 sia di recente istituzione, risale infatti al D.lgs. n. 65 del 2017, l'integrazione tra i nidi di infanzia e la scuola dell'infanzia negli ultimi anni è diventata argomento sempre più discusso nel contesto nazionale poiché è stata sottolineata l'importanza di un'educazione integrale che sia in continuità tra i due segmenti che costituiscono il sistema integrato. Stornaiulo (2021), a tal proposito, la ha definita come una sfida tanto educativa quanto sociale (Bondioli & Savio, 2021; Fiorucci et alii, 2022). In Italia dal 2005 al 2020 c'è stato un decremento nel tasso di partecipazione dei bambini tra i 3 e i 5 anni all'educazione dell'infanzia, si è passati infatti dal 98% al 92%. Al contrario, nell'educazione della primissima infanzia, al di sotto dei 3 anni, si è registrato un incremento di un punto percentuale nello stesso arco temporale (OECD, 2022).

L'educazione primaria rientra nel primo ciclo di istruzione, ha una durata di cinque anni e accoglie i bambini tra 6 e 11 anni. Rappresenta il primo segmento dell'istruzione obbligatoria seguita dalla scuola secondaria di primo grado e dai primi due anni della scuola secondaria di secondo grado.

L'educazione secondaria di primo grado ha una durata triennale e accoglie i ragazzi tra 11 e 14 anni. L'educazione secondaria di secondo grado, invece, rientra nel secondo ciclo di istruzione ed è obbligatoria solamente nei primi due anni. Si divide in quattro indirizzi: licei, istituti tecnici, istituti professionali e formazione professionale. I primi tre hanno generalmente durata quinquennale e accolgono i ragazzi tra 14 e 19 anni, mentre la formazione professionale è organizzata a livello regionale e può avere una durata variabile tra i tre e i quattro anni (Eurydice, 2022).

L'educazione terziaria in Italia è offerta sia dalle università, con orientamento più accademico, sia dagli istituti di alta formazione, con orientamento più professionalizzante. A questi si aggiungono in particolare i percorsi di istruzione terziaria organizzati dalle istituzioni dell'Alta Formazione Artistica, Musicale e Coreutica (Eurydice, 2022).

2.3. Perché valutare la scuola?

La valutazione dei sistemi scolastici avviene per ragioni differenti che sono rintracciabili nella polarità che si snoda attorno allo sviluppo e al controllo (Castoldi, 2008).

Il primo polo mette in luce la capacità trasformativa dell'attività di valutazione che, richiedendo il coinvolgimento degli operatori, si configura come un vero e proprio apprendimento organizzativo, mentre il secondo polo si riferisce all'accertamento della produttività del servizio e alla verifica del raggiungimento degli esiti attesi. [...] I Paesi che hanno scopi valutativi legati a processi trasformativi sono quelli in cui le scuole affrontano processi di autovalutazione, sia questa obbligatoria o consigliata, centralizzata oppure scelta in modo autonomo nelle sue modalità dalla singola istituzione scolastica. Sono però anche quei paesi che creano team di supporto per il miglioramento e il monitoraggio dei progressi delle scuole, costituendo istituti a livello locale o

nazionale, stanziando fondi per l'attività di sostegno e fornendo le debite garanzie legislative. I sistemi più avanzati cercano di raggiungere entrambi gli scopi o i vari scopi presenti sul continuum sviluppo controllo. (Cardone et alii, 2010, p. 700-701)

Per meglio inquadrare le funzioni della valutazione, si è assistito a una risposta corale da parte dei paesi oggetto di indagine con atti legislativi chiarificatori degli obiettivi della valutazione, i cui estremi verranno puntualmente ripotati e illustrati nei sottoparagrafi successivi, e con la predisposizione di enti incaricati della sua realizzazione, aspetto approfondito nel paragrafo 2.4.

2.3.1. La legislazione sulla valutazione del sistema scolastico finlandese

La legge n. 1295 del 2013 e la legge n. 582 del 2015 e le successive modifiche definiscono il ruolo e le funzioni di KARVI-FINEEC, *Finnish Education Evaluation Centre* (Laki 1295/13, 2013; Laki 582/15, 2015). L'istituto è un'organizzazione esperta e indipendente che si occupa della valutazione esterna del sistema educativo, producendo informazioni basate sulle rilevazioni funzionali nel processo decisionale e nelle politiche educative per lo sviluppo dell'educazione. KARVI è un'unità dell'Agenzia Nazionale Finlandese per l'Educazione (OPH, *Opetushallitus*), come indicato nella legge n. 564 del 2016, operante sotto il Ministero dell'Educazione e della Cultura. Il direttore è nominato dal governo e rimane in carica per quattro anni, partecipa alle attività del Consiglio di Valutazione che monitora e sviluppa le attività di KARVI (Laki 1295/13, 2013; Laki 582/15, 2015).

Per quel che riguarda il piano triennale delle attività è compito del Consiglio di Valutazione elaborarne un disegno contenente gli elementi da valutare e il programma di valutazione che verrà approvato successivamente dal Ministero dell'Educazione e della Cultura. I principi da seguire nella predisposizione del piano sono definiti più dettagliatamente da decreti governativi. Il piano triennale delle attività valutative verrà analizzato nel paragrafo 2.5, con particolare rimando

alla valutazione degli apprendimenti prevista nella scuola dell'obbligo per il triennio 2020-2023 (KARVI-FINEEC, 2022).

Oltre alla realizzazione delle attività valutative descritte nel piano triennale delle attività, è compito di KARVI, così come stabilito dalle leggi sopra riportate, effettuare valutazioni relative all'istruzione di base, della prima infanzia e all'istruzione superiore. La valutazione dei risultati negli apprendimenti e del raggiungimento degli obiettivi prefissati per ciascun livello scolastico, infatti, avviene secondo l'art. 14 della legge n. 628 del 1998 sull'istruzione di base, secondo gli art. 11 e 12 della legge n. 714 del 2018 sull'istruzione della scuola secondaria superiore, secondo l'art. 15 della legge n. 531 del 2017 sull'istruzione erogata dagli istituti professionali e secondo l'art. 5 della legge 633 del 1998 sull'educazione artistica. Tra i compiti di KARVI, in ottica di valutazione del sistema educativo, rientra anche il sostegno agli enti erogatori di istruzione e formazione dalla prima infanzia all'istruzione terziaria nei processi di autovalutazione e nel monitoraggio della qualità (Laki 1295/13, 2013; Laki 582/15, 2015; Karjalainen et alii, 2016; Mikkola et alii, 2017; Repo et alii, 2018; Härkönen et alii, 2019; Venäläinen et alii, 2020; Vlasov et alii, 2020; Siippainen et alii, 2020; Ukkola & Metsämuuronen, 2021).

2.3.2. La legislazione sulla valutazione del sistema scolastico spagnolo

La Ley Organica n. 2 del 2006 e la Ley Organica n. 3 del 2020¹⁴ regolamentano la valutazione nazionale del sistema scolastico spagnolo. Le finalità di queste ultime sono molteplici e possono essere riassunte in cinque punti:

- Contribuire a migliorare la qualità e l'equità del sistema educativo;
- Orientare le politiche educative;
- Aumentare la trasparenza e l'efficacia del sistema educativo;

¹⁴ Per approfondimenti circa le leggi organiche: <https://boe.es/buscar/act.php?id=BOE-A-2006-7899&p=20220727&tn=0#tvi> (ultima consultazione 21 febbraio 2023)

- Fornire informazioni sul grado di raggiungimento degli obiettivi di miglioramento stabiliti dalle amministrazioni scolastiche;
- Fornire informazioni sul grado di raggiungimento degli obiettivi educativi fissati in ambito nazionale ed europeo, nonché sul rispetto degli impegni educativi assunti in relazione alla domanda della società spagnola e agli obiettivi fissati nel contesto dell'Unione Europea (Ley Organica 2/06, 2006; Ley Organica 3/20, 2020).

La valutazione, secondo le leggi organiche, è estesa ad ambiti differenti che comprendono i livelli degli apprendimenti degli studenti, i processi di insegnamento-apprendimento, le risorse disponibili, i contesti educativi, il personale docente e i dirigenti scolastici, il monitoraggio delle istituzioni scolastiche e le amministrazioni educative. La valutazione del sistema educativo è realizzata dall'INEE, *Instituto Nacional de Evaluacion Educativa*, organismo responsabile, e da tutti quegli enti incaricati territorialmente dall'INEE. Inoltre, è realizzata anche dai dirigenti scolastici e dai docenti che sono chiamati a collaborare con le amministrazioni centrali e periferiche delle comunità autonome nel momento in cui si svolgono monitoraggi e valutazioni (Ley Organica 2/06, 2006; Ley Organica 3/20, 2020).

In collaborazione con le comunità autonome, rientrano tra i compiti dell'INEE definiti dalle leggi organiche: la realizzazione di misurazioni e valutazioni che possano consentire di ottenere dati rappresentativi sia per gli studenti che per le scuole, per le comunità autonome e per la nazione nel complesso (le rilevazioni riguardano le competenze stabilite nel curriculum dell'istruzione primaria e secondaria e vengono effettuate applicando criteri stabiliti e omogenei); la messa a punto di metodologie e standard scientifici che garantiscono la qualità, la validità e l'affidabilità delle valutazioni; il supporto agli organi di valutazione delle comunità autonome nell'applicazione dei criteri e degli standard concordati e nella realizzazione delle misurazioni e delle valutazioni. Nell'ottica della attuazione di piani pluriennali di valutazione generale del sistema scolastico spagnolo, è responsabilità dell'INEE realizzare durante l'ultimo anno della scuola primaria e l'ultimo anno della scuola secondaria di primo grado una valutazione formativa delle competenze acquisite dagli studenti su base campionaria con il fine di informare le

istituzioni scolastiche, le famiglie e i decisori politici. L'INEE, inoltre, in collaborazione con le amministrazioni scolastiche locali, garantisce la partecipazione della Spagna alle indagini comparative internazionali condotte da OCSE, IEA e Commissione Europea. Infine, cooperando con gli organi competenti, l'INEE ha la responsabilità di predisporre il sistema statale degli indicatori dell'istruzione di base che contribuisce alla conoscenza approfondita del sistema educativo, anche da parte dell'opinione pubblica, e guida i processi decisionali in tutti i settori nei quali rientrano l'istruzione e la formazione. Gli indicatori forniscono informazioni che consentono di valutare il grado di equità raggiunto dal sistema educativo seguendone l'evoluzione in modo longitudinale. La messa a punto del sistema di indicatori avviene lungo un continuum in cui le scuole sono chiamate a fornire i loro dati alle amministrazioni centrali tramite i processi di autovalutazione di istituto, mentre le amministrazioni centrali, una volta redatto il report di indicatori, sono chiamate a restituire le statistiche elaborate in primo luogo alle amministrazioni delle comunità autonome ma, scendendo più in profondità, anche alle singole scuole (Ley Organica, 2/06, 2006, 3/20, 2020).

2.3.3. La legislazione sulla valutazione del sistema scolastico italiano

I primi due decenni degli anni 2000 hanno rappresentato un momento cruciale per l'introduzione della legislazione sulla valutazione del sistema scolastico italiano. Dopo la trasformazione del CEDE in Invalsi del 1999, ricordata nel paragrafo 2.4, e l'inizio delle attività di ricerca valutativa del nuovo istituto, nel 2004 è stato istituito ufficialmente il Sistema Nazionale di Valutazione con il decreto legislativo n. 286 del 19 novembre 2004 (D.Lgs 286/04, 2004; Legge 176/07, 2007) per sostenere l'attuazione dell'autonomia scolastica e favorire la crescita culturale, sociale ed economica del paese. Il fine era quello di monitorare la qualità, l'efficienza e l'efficacia del sistema educativo migliorandone l'offerta formativa attraverso la collaborazione di tre istituti: Invalsi, Istituto Nazionale per la Valutazione del Sistema

di Istruzione e Formazione, con il ruolo di coordinatore, Indire, Istituto Nazionale di Documentazione, Innovazione e Ricerca Educativa, e Contingente Ispettivo, ovvero dirigenti di seconda fascia con funzione tecnico-ispettiva (Legge 10/11, 2011; De Santis & Asquini, 2020). Tuttavia, è solamente nel 2013 che al Sistema Nazionale di Valutazione, SNV, è stato dato un vero e proprio ordinamento e regolamento attraverso il D.P.R n. 80 del 28 marzo 2013 (D.P.R 80/13, 2013). Quest'ultimo, fin dalla sua istituzione, ha l'obiettivo di implementare la qualità dell'offerta formativa e degli apprendimenti, valutando l'efficienza e l'efficacia del sistema educativo (D.P.R 80/13, 2013). Ai tre istituti che ne fanno parte sono affidati compiti specifici. L'Invalsi elabora i protocolli di valutazione e il programma delle visite alle istituzioni scolastiche ad opera dei nuclei di valutazione esterna, rende disponibili a queste ultime strumenti per la valutazione, definisce gli indicatori per la valutazione dei dirigenti scolastici, forma e seleziona gli esperti dei nuclei di valutazione esterna e assicura la partecipazione dell'Italia alle indagini internazionali. L'Indire supporta le istituzioni scolastiche per la messa a punto e l'implementazione dei piani di miglioramento della qualità dell'offerta formativa e dei risultati negli apprendimenti degli studenti e attiva progetti di ricerca finalizzati al miglioramento della didattica, al supporto dei processi di innovazione e alla formazione in servizio del personale della scuola. Il Contingente Ispettivo, invece, partecipa ai nuclei di valutazione tramite la dotazione organica disponibile dei dirigenti di seconda fascia con funzione tecnico-ispettiva (D.P.R 80/13, 2013). Lo stesso D.P.R 80/13 stabilisce anche che il procedimento della valutazione delle istituzioni scolastiche si sviluppa in quattro fasi sulla base dei protocolli di valutazione stabiliti comunitariamente e prendendo in considerazione il ruolo delle scuole nei processi di autovalutazione. La prima fase prevede l'autovalutazione delle istituzioni scolastiche.

1) Analisi e verifica del proprio servizio sulla base dei dati resi disponibili dal sistema informativo del Ministero, delle rilevazioni sugli apprendimenti e delle elaborazioni sul valore aggiunto restituite dall'Invalsi, oltre a ulteriori elementi significativi integrati dalla stessa scuola; 2) Elaborazione di un

rapporto di autovalutazione in formato elettronico, secondo un quadro di riferimento predisposto dall'Invalsi, e formulazione di un piano di miglioramento. (D.P.R 80/13, 2013, art. 6, comma 1)

La seconda fase prevede la valutazione esterna.

1) Individuazione da parte dell'Invalsi delle situazioni da sottoporre a verifica, sulla base di indicatori di efficienza ed efficacia previamente definiti dall'Invalsi medesimo; 2) Visite dei nuclei di cui al comma 2, secondo il programma e i protocolli di valutazione adottati dalla conferenza ai sensi dell'articolo 2, comma 5; 3) Ridefinizione da parte delle istituzioni scolastiche dei piani di miglioramento in base agli esiti dell'analisi effettuata dai nuclei. (D.P.R 80/13, 2013, art. 6, comma 1)

La terza fase prevede le azioni di miglioramento, ovvero la definizione e l'attuazione, con il supporto di Indire, di università o altri enti di ricerca, di interventi di miglioramento. La quarta e ultima fase, infine, prevede la rendicontazione sociale da parte delle istituzioni scolastiche attraverso l'elaborazione di indicatori, la pubblicazione e la diffusione dei risultati raggiunti sia in un'ottica di trasparenza che in un'ottica di condivisione con la comunità (D.P.R 80/13, 2013).

2.4. Quale ente è preposto alla valutazione del sistema scolastico?

Le motivazioni che spingono i paesi a dotarsi di sistemi per la valutazione delle istituzioni scolastiche sono spesso da rintracciarsi in ragioni storiche e politiche, l'autonomia scolastica e il decentramento ne sono un esempio (Benadusi, 2009). La distribuzione del potere decisionale e organizzativo tra i vari livelli coinvolti nella gestione dei sistemi educativi ha fatto emergere la necessità di "controllo sociale" da parte degli attori coinvolti (Cardone et alii, 2010). In alcuni paesi, Germania, Finlandia e Spagna compresi, ha prevalso l'idea alla base della valutazione come forma di rendicontazione alle amministrazioni nazionali e locali, in altri, Inghilterra, Danimarca e Austria ad esempio, la valutazione è considerata una forma di professionalizzazione di chi la scuola la vive quotidianamente (Cardone et alii, 2010). Negli ultimi anni,

oltre ad emanare leggi per regolamentare la valutazione dei sistemi educativi, i decisori politici di molti paesi hanno anche incaricato alcuni enti di ricerca per assolvere a quanto indicato negli atti legislativi (Cardone et alii, 2010).

Per la descrizione delle attività degli enti preposti alla valutazione dei sistemi scolastici, oltre a consultare la letteratura di settore di volta in volta segnalata, sono stati analizzati anche i siti web di ciascun istituto di valutazione¹⁵.

2.4.1. KARVI-FINEEC, l'istituto di valutazione scolastica finlandese

Il *Kansallinen Koulutuksen Arviointikeskus*, KARVI (FINEEC, *Finnish Education Evaluation Centre*, <https://karvi.fi/en/>), è l'istituto responsabile dal 2014 della valutazione del sistema scolastico finlandese a partire dalla prima infanzia fino ad arrivare all'età adulta. Le misurazioni e le valutazioni realizzate dall'istituto incrementano e supportano la conoscenza a livello locale, regionale e nazionale e rendono possibile la comparazione internazionale. La ricerca valutativa consiste, tra le altre cose, nel monitoraggio nazionale dei risultati negli apprendimenti e nelle valutazioni tematiche e di sistema. KARVI, inoltre, è deputato alla valutazione periodica e all'accreditamento del sistema terziario dell'istruzione. Nelle *missions* dell'istituto rientrano anche il sostegno agli stakeholders nella presa di decisioni riguardanti l'educazione e la gestione e la valutazione della qualità. L'ente è nato il 1 maggio del 2014 dall'unione tra *Finnish Higher Education Council*, il *Finnish Evaluation Council* e il *Finnish National Board of Education*.

Il KARVI opera come un'unità separata dell'Agenzia Nazionale Finlandese per l'Educazione (Opetushallitus, OPH), un'autorità indipendente responsabile

¹⁵ Per approfondimenti circa KARVI-FINEEC: <https://karvi.fi/> (ultima consultazione 21 febbraio 2023)

Per approfondimenti circa INEE: <https://www.educacionyfp.gob.es/inee/portada.html> (ultima consultazione 21 febbraio 2023)

Per approfondimenti circa INVALSI: <https://www.invalsi.it/invalsi/index.php> (ultima consultazione 21 febbraio 2023)

della valutazione dell'istruzione. Svolge valutazioni relative all'istruzione e agli istituti che erogano istruzione e formazione, dall'educazione della prima infanzia all'istruzione superiore e al mondo del lavoro. Conta un totale di circa 50 dipendenti con rapporti di lavoro sia a tempo determinato che a tempo indeterminato. Il personale è composto da esperti di valutazione e comunicazione, esperti di amministrazione e gestione e supervisori. KARVI opera come un'organizzazione decentralizzata, vale a dire la maggior parte del personale lavora principalmente a Helsinki, ma una parte del personale lavora a Jyväskylä. KARVI è coinvolto nel programma di sviluppo della cultura del coaching insieme all'Agenzia Nazionale Finlandese per l'Educazione e alla Commissione nazionale degli esami conclusivi della scuola secondaria superiore (Ylioppilastutkintolautakunta, YTL). (Peltola, 2021, p. 4, traduzione a cura dell'autrice)

Tutte le attività condotte da KARVI avvengono in ottemperanza dei principi di autonomia e indipendenza, fiducia, apertura e collaborazione. Se da una parte KARVI mantiene l'indipendenza e l'autonomia nel lavoro di ricerca, nella scelta delle metodologie e degli strumenti, dall'altra, attraverso la partecipazione degli stakeholders, incentiva la collaborazione per lo scopo comune. L'obiettivo delle valutazioni condotte da KARVI è quello di mettere a disposizione della comunità dati e informazioni cui far ricorso nei processi decisionali non solo a livello nazionale e internazionale, ma anche a livello locale e regionale.

Le attività, oltre a richiamare i livelli all'interno dei quali si articola il sistema educativo (locale, regionale e nazionale), includono, sia doveri statutari che attività di consulenza e supporto agli stakeholders nelle pratiche di autovalutazione (Mikkola et alii, 2017). La valutazione del sistema scolastico è articolata all'interno del piano di monitoraggio messo a punto per il triennio 2020-2023. La figura 7 sintetizza la *mission* di KARVI e delle attività di valutazione relative all'istruzione, alla didattica e ai risultati negli apprendimenti, di valutazioni tematiche ad ampio spettro, di sostegno agli enti erogatori di istruzione nell'autovalutazione e nel monitoraggio della qualità e di audit dell'istruzione terziaria.

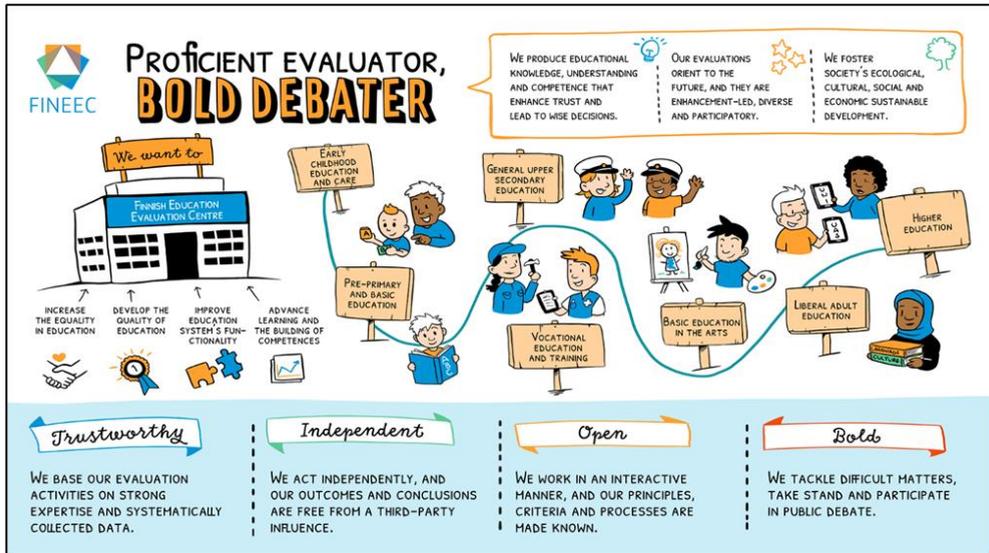


Figura 7. Strategia valutativa di KARVI. Fonte: <https://karvi.fi/en/fineec/>.

2.4.2. INEE, l'istituto di valutazione scolastica spagnolo

L'Instituto Nacional de Evaluación Educativa, INEE (<https://www.educacionyfp.gob.es/inee/portada.html>), è un organismo del Ministero dell'Educatione e della Formazione Professionale posto all'interno della Direzione Generale della Valutazione e della Cooperazione Territoriale. Realizza la valutazione del sistema educativo spagnolo in tre ambiti: valutazioni nazionali, valutazioni internazionali e indicatori educativi, *indicadores educativos* (*Sistema Estatal de Indicadores de la Educación*) (Ministerio de Educación y Formación Profesional, 2022). L'INEE, con la sua denominazione originale INCE, *Instituto Nacional de Calidad Educativa*, è stato fondato nel 1990, ma è dal 2012 che ha acquisito l'attuale denominazione (Ley Organica 1/90, 1990; Ley Organica 9/95, 1995).

L'INCE nasce nel 1990 dalla Legge Organica di Organizzazione Generale del Sistema Educativo (LOGSE), che dedica l'articolo 62, inserito nel titolo quarto, alla qualità dell'insegnamento. Nel giugno 1993 è stato approvato un regio decreto che ne regola la composizione, le funzioni e l'organizzazione (RD

928/1993). Quasi un anno dopo, il 17 maggio 1994, si è tenuta la riunione costitutiva del suo Consiglio Direttivo, nella quale sono stati discussi e approvati l'inquadramento e le linee generali che regolano l'organizzazione e il funzionamento dell'Istituto, nonché i suoi primi progetti e i suoi meccanismi di coordinamento. Inizia così il cammino di un'organizzazione che, con diversi cambi di nome e di orientamento, continua ad essere attiva e fiorente ancora oggi. (Ferrer, 2014, p. 3, traduzione a cura dell'autrice)

La mission e le attività che realizza rispondono alla Ley Organica n. 2 del 3 maggio 2006, Titolo VI art. 140-147 (*Evaluación del sistema educativo*) modificata dalla Ley Organica n. 3 del 29 dicembre 2020 (Ley Organica 2/06, 2006; Ley Organica 8/13, 2013; Ley Organica 3/2020, 2020).

2.4.3. INVALSI, l'istituto di valutazione scolastica italiano

L'Istituto Nazionale per la Valutazione del Sistema Educativo di Istruzione e di Formazione, INVALSI (<https://www.invalsi.it/invalsi/index.php>), è l'istituto preposto nel contesto italiano per la valutazione del sistema educativo. È un Ente Pubblico di Ricerca la cui sede si trova a Roma, in via Ippolito Nievo 35, sottoposto alla vigilanza del Ministero dell'Istruzione e del Merito.

L'Istituto, attraverso le proprie attività di studio e ricerca sul funzionamento dei sistemi formativi, delle politiche e delle prassi educative, la predisposizione e l'implementazione di strumenti di misurazione degli apprendimenti e delle competenze degli studenti e le attività di valutazione delle istituzioni scolastiche e formative da esso coordinate nell'ambito del SNV: promuove il miglioramento dei livelli di istruzione e della qualità del capitale umano, contribuendo allo sviluppo e alla crescita del Sistema d'Istruzione, motore di sviluppo dell'economia italiana e promotore di equità sociale, nel quadro degli obiettivi fissati in sede nazionale, europea e internazionale. L'Istituto assicura lo svolgimento delle rilevazioni annuali sugli apprendimenti e delle prove scritte standardizzate a carattere nazionale previste dalla normativa vigente in materia. L'Istituto, nel quadro degli obiettivi definiti dal Piano nazionale per la ricerca, di seguito denominato «PNR», esplica funzioni di rilevante interesse tecnico-scientifico, economico e sociale e, pertanto, nell'attuazione dei suoi compiti, promuove la collaborazione con gli altri enti di ricerca, le

amministrazioni pubbliche, le regioni, gli enti locali, le strutture universitarie e il mondo dell'impresa. (INVALSI, 2017, p. 3.)

L'Invalsi ha raccolto l'eredità del CEDE, *Centro Europeo dell'Educazione*, nato all'inizio degli anni Settanta e sostituito all'inizio degli anni 2000. Il Decreto del Presidente della Repubblica n. 419, art. 12, del 31 maggio 1974, infatti, sancisce ufficialmente la nascita di un nuovo centro europeo con sede a Frascati (RM), in villa Falconieri, sottoposto alla vigilanza del Ministero della Pubblica Istruzione. Al CEDE veniva attribuito il compito di raccogliere, elaborare e diffondere la documentazione pedagogico-didattica italiana e straniera, focalizzandosi particolarmente sugli ordinamenti scolastici di altri paesi, nello specifico quelli europei, e sull'attività in campo educativo delle organizzazioni internazionali. Erano nel mirino degli interessi di ricerca del CEDE la programmazione e i costi dei sistemi educativi, il rapporto tra formazione e occupazione, l'educazione permanente, i problemi dell'apprendimento e le strategie più opportune per valutarli, l'aggiornamento professionale del personale coinvolto nella scuola, le innovazioni educative e l'impiego delle tecnologie educative (D.P.R. 419/74, 1974; Carli et alii, 2015; De Santis & Asquini, 2020). Il consiglio direttivo di esperti, organo del CEDE nominato dal ministro e formato da undici membri, deliberava il programma delle attività, autorizzando la stipula di contratti e convenzioni con università e altri enti di ricerca (D.P.R. 419/74, 1974). Il CEDE è stato trasformato in Invalsi con il Decreto Legislativo n. 258 del 20 luglio 1999 (D.L. 258/99, 1999), a seguito dell'introduzione dell'autonomia scolastica. Inizialmente Istituto Nazionale per la Valutazione del Sistema di Istruzione, l'Invalsi, sottoposto alle direttive del ministro della pubblica istruzione e della ricerca che ne individua le priorità strategiche, ha il compito di coordinare il Sistema Nazionale di Valutazione.

In particolare, l'Istituto valuta l'efficienza e l'efficacia del sistema di istruzione nel suo complesso ed analiticamente, ove opportuno anche per singola istituzione scolastica, inquadrando la valutazione nazionale nel contesto internazionale; studia le cause dell'insuccesso e della dispersione scolastica con

riferimento al contesto sociale ed alle tipologie dell'offerta formativa; conduce attività di valutazione sulla soddisfazione dell'utenza; fornisce supporto e assistenza tecnica all'amministrazione per la realizzazione di autonome iniziative di valutazione e supporto alle singole istituzioni scolastiche anche mediante la predisposizione di archivi informatici liberamente consultabili; valuta gli effetti degli esiti applicativi delle iniziative legislative che riguardano la scuola; valuta gli esiti dei progetti e delle iniziative di innovazione promossi in ambito nazionale; assicura la partecipazione italiana a progetti di ricerca internazionale in campo valutativo e nei settori connessi dell'innovazione organizzativa e didattica. (D.L 258/99, 1999, art. 1, comma 3)

Le prime vicende della storia dell'istituto sono state costellate all'inizio del nuovo millennio di mutazioni e arricchimenti che hanno conferito una caratterizzazione più o meno stabile al nascente sistema di valutazione nazionale. Senza entrare nel merito delle circostanze sociali e politiche che hanno visto l'alternarsi di consigli di amministrazione e presidenti a partire dal dott. Vertecchi fino ad arrivare alla dott.ssa Ajello e all'attuale dott. Ricci, le rilevazioni standardizzate, gli obiettivi di queste ultime e la diffusione dei risultati sono da sempre al centro dell'attenzione (Corsini & Losito, 2014; Trincherò, 2014c).

2.5. Cosa valutano gli istituti nazionali preposti?

Quello che viene valutato e i dati che vengono rilevati dipendono in larga misura dagli scopi che sono stati prefissati. Là dove gli obiettivi dei processi valutativi sono maggiormente orientati al controllo, si fa più attenzione agli outcomes e ai benchmark. In un'ottica di produttività lo scarto tra i risultati attesi e i risultati ottenuti rappresenta ciò che è andato bene e ciò che è andato male. Contrariamente, là dove gli obiettivi dei processi valutativi sono orientati allo sviluppo, l'attenzione è concentrata sulla qualità globale dell'educazione erogata (Cardone et alii, 2010).

Oggi l'idea che la produttività - pur essendo condizione necessaria - non sia sufficiente per garantire la qualità del servizio è largamente condivisa. Diversi Paesi hanno quindi scelto di analizzare anche le componenti del sistema scuola,

ossia l'implementazione dei curricoli, le metodologie didattiche, le forme di gestione e organizzazione dell'unità scolastica, la formazione in servizio degli insegnanti, i rapporti con il territorio e le famiglie. (Cardone et alii, 2010, p. 702)

Gli elementi da osservare sono spesso racchiusi all'interno di un piano di valutazione biennale, triennale o quadriennale elaborato dagli istituti di valutazione preposti. Obiettivo dei seguenti sottoparagrafi è quello di analizzare nello specifico la valutazione degli apprendimenti prevista con cadenza regolare nei tre contesti nazionali oggetto di studio. Nel sottoparagrafo 2.5.2 si fa riferimento al più ampio piano di valutazione per il triennio 2020-2023 poiché in Finlandia non sono previste rilevazioni annuali, diagnostiche e censuarie su base nazionale circa i livelli negli apprendimenti degli studenti della scuola dell'obbligo come in Spagna e Italia, bensì delle analisi tematiche su alcune competenze ritenute fondamentali. Nell'esplicitare più dettagliatamente l'andamento delle rilevazioni nazionali svolte in Finlandia è stato ricostruito il loro sviluppo longitudinale nel tempo.

Per l'analisi documentale sono stati consultati in particolar modo i documenti programmatici riguardanti la valutazione standardizzata degli apprendimenti, di volta in volta indicati nei sottoparagrafi, di ciascun istituto nazionale di valutazione.

2.5.1. Il piano di valutazione per il triennio 2020-2023 di KARVI-FINEEC

Il piano di valutazione dell'istruzione per il triennio 2020-2023 (KARVI-FINEEC, 2022), nonché le linee guida e di sviluppo del sistema e dei metodi di valutazione, delinea quattro priorità:

- Sviluppo degli apprendimenti e delle competenze;
- Promozione dell'equità;
- Aumento dell'efficienza e dell'efficacia del sistema educativo;
- Sostegno dello sviluppo continuo.

Attraverso la prima priorità, sviluppo degli apprendimenti e delle competenze, si vogliono accrescere le conoscenze nel campo della ricerca valutativo-educativa rispondendo ai seguenti interrogativi (KARVI-FINEEC, 2022):

- Quanto sono stati raggiunti gli obiettivi fissati per gli apprendimenti e le competenze?
- Quali sono i fattori che promuovono o ostacolano l'apprendimento e il rafforzamento delle competenze?
- Come promuovere lo sviluppo dell'apprendimento e delle competenze rispetto ai bisogni futuri e in un'ottica di apprendimento permanente?

La valutazione dei risultati negli apprendimenti relativi all'istruzione di base, tramite il suo svolgimento campionario, raccoglie dati e informazioni utili per i decisori politici non solo per quel che riguarda la distribuzione dei livelli di apprendimento, intesa come conoscenze delle discipline insegnate a scuola, ma anche per quel che riguarda la realizzazione di un sistema educativo equo. La valutazione degli apprendimenti si basa sulle linee guida definite nazionalmente per i curricula scolastici ed è di tipo individuale solamente per gli studenti del nono anno di scolarità. Quest'ultima sempre più spesso è affiancata da valutazioni longitudinali che monitorano lo sviluppo delle competenze degli studenti per un periodo di tempo più lungo (KARVI-FINEEC, 2022).

Nel piano triennale delle attività 2020-2023 e nello specifico delle valutazioni dei risultati negli apprendimenti sono comprese valutazioni longitudinali e valutazioni delle tendenze alle quali possono essere aggiunte le valutazioni degli apprendimenti nella fase finale dell'istruzione dell'obbligo.

Le valutazioni longitudinali hanno l'obiettivo di misurare le competenze di base degli studenti analizzando i fattori che ne influenzano lo sviluppo negli snodi-annualità critici del sistema di istruzione. Al terzo, al sesto e al nono anno di scolarità sono misurate le competenze matematiche e linguistiche (nelle lingue madri, finlandese o svedese e svedese L2) - le ultime valutazioni nel primo anno di scolarità sono state effettuate nel 2018. Al nono anno di scolarità sono misurate le competenze linguistiche nella lingua straniera (inglese e finlandese per chi è madrelingua svedese) - le ultime valutazioni del sesto e del settimo anno di scolarità sono state effettuate nel 2018. I monitoraggi longitudinali avvengono con ciclicità triennale, si pensi ad esempio alla

valutazione delle competenze linguistiche in finlandese per chi è madrelingua svedese effettuate per il sesto anno di scolarità nel 2018 e per il nono anno di scolarità nel 2021 (KARVI-FINEEC, 2022).

Le valutazioni delle tendenze, così come intese da KARVI, sono monitoraggi ripetuti sulla stessa disciplina a distanza di qualche tempo e sempre nello stesso anno di scolarità. Infatti, hanno l'obiettivo di raccogliere dati sui possibili cambiamenti in termini di competenze in una determinata annualità. Le competenze matematiche e quelle nella lingua madre, ad esempio, rientrano tra quelle competenze misurate a distanza di qualche anno sempre nel nono anno di scolarità (KARVI-FINEEC, 2022).

Altri progetti di ricerca valutativa nel triennio 2020-2023 riguardano l'insegnamento della lingua madre, gli apprendimenti degli studenti che frequentano gli istituti tecnici e professionali, l'analisi della qualità dell'istruzione offerta dagli istituti tecnici e professionali, la valutazione degli istituti di istruzione superiore e delle innovazioni pedagogiche da loro implementate. Inoltre, nel biennio 2020-2021 appena concluso, sono state condotte ricerche valutative sul livello di alfabetizzazione degli immigrati e sui bisogni formativi dei bambini di cinque anni, compreso lo stato dell'arte dell'educazione artistica nella scuola della prima infanzia.

Attraverso la seconda priorità, promozione dell'equità, si vuole ampliare la conoscenza e la promozione dell'uguaglianza rispondendo ai seguenti interrogativi (KARVI-FINEEC, 2022):

- Qual è lo stato delle pari opportunità nell'educazione in Finlandia?
- Quali fattori promuovono o impediscono la realizzazione dell'uguaglianza educativa?

La risposta ai due interrogativi di ricerca proviene da una parte dalle valutazioni che avvengono nella scuola di base, dall'altra da progetti di ricerca costruiti ad hoc. Quelli in corso di svolgimento riguardano la prevenzione del bullismo e la costruzione di un ambiente di lavoro accogliente e la promozione della coesione sociale nella comunità scolastica (KARVI-FINEEC, 2022).

Attraverso la terza priorità, aumento dell'efficienza e dell'efficacia del sistema educativo, si vogliono approfondire alcune questioni rispondendo ai seguenti interrogativi (KARVI-FINEEC, 2022):

- Come funziona il sistema educativo nel suo complesso?
- Quali effetti hanno avuto le riforme educative?
- Qual è la capacità di risposta del sistema educativo ai cambiamenti degli ambienti e dei contesti attuali?

I progetti di ricerca che rientrano in questa priorità del piano triennale delle attività approfondiscono aspetti relativi ai curricula della scuola dell'infanzia e dell'istruzione di base, all'inclusione degli studenti stranieri, alle capacità del sistema educativo di rispondere a improvvisi cambiamenti strutturali, al possibile prolungamento della scuola dell'obbligo etc. (KARVI-FINEEC, 2022).

Attraverso la quarta priorità, sostegno dello sviluppo continuo, si vuole sostenere questo aspetto rispondendo ai seguenti interrogativi di ricerca (KARVI-FINEEC, 2022):

- Come funziona la gestione della qualità a supporto dello sviluppo continuo?
- In che modo è supportato il raggiungimento degli obiettivi prefissati?
- Come si realizza il senso di comunità e la partecipazione dei diversi attori allo sviluppo?

L'attuale piano delle attività 2020-2023 racchiude al suo interno i progetti di ricerca da realizzarsi entro il triennio. Ognuno di questi si concentra su una delle priorità stabilite, pur tenendo presente che l'attività di KARVI è regolata dalla legislazione specifica riportata nel paragrafo 2.3 del presente capitolo (KARVI-FINEEC, 2022).

2.5.2. La valutazione degli apprendimenti di INEE

La valutazione diagnostica degli apprendimenti degli studenti spagnoli è realizzata dall'INEE in collaborazione con le amministrazioni locali. Coinvolge annualmente i bambini e le bambine che frequentano il quarto anno della scuola primaria e i ragazzi e le ragazze del secondo anno della scuola secondaria di primo grado.

Le rilevazioni hanno un carattere censuario e hanno l'obiettivo di informare e formare le istituzioni scolastiche, i dirigenti scolastici, gli insegnanti, gli studenti, le famiglie e la comunità tutta. Inoltre, in ottemperanza a quanto stabilito per i piani di miglioramento (art. 121 della Ley Organica 2/06), le istituzioni scolastiche, a seguito degli esiti della valutazione, sono incoraggiate a sviluppare proposte di azione per far in modo che gli studenti possano raggiungere le competenze stabilite, per migliorare la qualità e l'equità dell'istruzione e per orientare i processi di insegnamento-apprendimento.

Le istituzioni scolastiche adotteranno le misure necessarie per compensare le carenze che possono esistere nella competenza nella comunicazione linguistica, nella lingua spagnola e, se del caso, nelle seconde lingue ufficiali, prendendo come riferimento le valutazioni effettuate in precedenza e includendo queste valutazioni e misure nel loro progetto educativo. Le amministrazioni scolastiche adotteranno le iniziative necessarie per facilitare le istituzioni scolastiche nell'applicazione delle misure. Il progetto educativo comprenderà al suo interno un piano di miglioramento, che sarà rivisto periodicamente, in cui, sulla base degli esiti delle valutazioni degli studenti e della scuola stessa, verranno proposte le strategie e le azioni necessarie per migliorare i risultati e le procedure educative, il rapporto con le famiglie e con il contesto circostante. (Ley Organica 2/2006, art. 121, comma 2bis-2ter, traduzione a cura dell'autrice)

La valutazione diagnostica degli apprendimenti, si legge nelle leggi organiche del 2006 e del 2020 (Ley Organica 2/06, 2006, 3/20, 2020), è volta a verificare le competenze linguistiche e matematiche. Le amministrazioni locali hanno il compito di supervisionare nel contesto di riferimento lo svolgimento della misurazione e della valutazione degli apprendimenti, fornendo alle istituzioni educative il supporto di cui necessitano. L'obiettivo non è quello di stilare classifiche, tanto è vero che spetta alle amministrazioni scolastiche locali disciplinare le modalità di diffusione alla comunità degli esiti delle valutazioni, nonché i piani di azione derivanti. Ai fini del miglioramento dell'organizzazione e dell'efficacia e dell'efficienza delle scuole, in aggiunta alla valutazione degli apprendimenti, possono essere predisposte valutazioni dei

dirigenti scolastici. Questa valutazione, unita a quella dei servizi e dell'offerta formativa, è effettuata dagli ispettori.

La valutazione diagnostica degli apprendimenti differisce da quella generale del sistema educativo non solo per il gruppo di studenti coinvolto, che da una parte è campionario dall'altra censuario, ma anche per il grado scolastico di riferimento. La figura 8 riporta le principali differenze. Come si può notare, se per la valutazione diagnostica degli apprendimenti sono responsabili le amministrazioni educative, per quella generale risponde l'INEE in collaborazione con le stesse amministrazioni educative. Inoltre, la prima è annuale mentre la seconda è pluriennale.

	Evaluación de Diagnóstico	Evaluación General del Sistema Educativo
¿Cuándo se realizan?	Mitad de etapa: 4.º EP y 2.º ESO	Final de etapa: 6.º EP y 4.º ESO
¿Qué se evalúa?	Competencias específicas, vinculadas al currículo articulado en áreas o materias	Competencias clave a partir de los descriptores operativos del Perfil de salida, no vinculadas a áreas o materias
¿Quién es el responsable de su realización?	Administraciones educativas	INEE en colaboración con las Administraciones educativas
Periodicidad	Anual	Plurianual
Carácter	Informativo, formativo y orientador para los centros, profesorado, alumnado, familias y conjunto de la comunidad educativa (informe del centro)	Informativo, formativo y orientador para los centros e informativo para las familias y conjunto de la comunidad educativa (informe del sistema o CC. AA.)
Marco normativo	Marco común de evaluación realizado por el INEE con los organismos correspondientes de las Administraciones educativas.	Estándares básicos metodológicos y científicos establecidos por el INEE en colaboración con las CC. AA. La Conferencia Sectorial de Educación velará para que se realicen con criterios de homogeneidad.
Alumnado participante	Censal. Participa todo el alumnado del sistema educativo español matriculado en los cursos a los que se dirige la prueba.	Muestral. Muestra representativa del Conjunto del Estado y de las CC. AA.

Figura 8. Differenze tra la valutazione diagnostica e la valutazione generale-formativa del sistema scolastico spagnolo. Fonte: <https://www.educacionyfp.gob.es/inee/evaluaciones-nacionales/evaluaciones-previstas-lomloe.html>.

La valutazione degli apprendimenti degli studenti è stata introdotta in Spagna con la Ley General de Educacion del 1970. Per la prima volta in quell'occasione il termine *valoracion* è stato utilizzato in riferimento

alla possibilità di misurare le competenze degli studenti da parte degli insegnanti. Ne è seguita l'istituzione di un centro di ricerca educativa chiamato CENIDE, riqualificato come INCIE dal 1974 e CIDE dal 1983. Successivamente la legge LOGSE, *Ley Organica de Ordenacion General del Systema Educativo*, del 1990 ha posto l'accento sulla valutazione come processo che può assicurare la qualità del sistema educativo. Si è trattato di una legge particolarmente impattante sulla struttura del sistema scolastico spagnolo poiché volta ad adeguarlo alle richieste da parte dell'Europa dopo la caduta del regime franchista (Pomante, 2016). Oltre ad innalzare l'obbligo scolastico al livello degli altri paesi europei e ad attenuare le differenze tra il percorso liceale e quello tecnico-professionale, prima ritenuto di secondo ordine, la legge LOGSE ha permesso l'intervento delle comunità autonome nella gestione e nell'organizzazione dei rispettivi sub-sistemi scolastici regionali. Contestualmente, è stato creato l'antesignano dell'attuale INEE, l'INCE, *Instituto Nacional de Calidad y Evaluación*, grazie al quale sono stati inaugurati progetti di ricerca sia locali che nazionali che hanno permesso alcune sperimentazioni circa la valutazione delle istituzioni educative, un esempio è il *plan EVA* (Garrido, 2000; Maestro Martin, 2006). È a partire dal 1995 che l'attuale INCE in collaborazione con le comunità autonome dà avvio alle valutazioni nella scuola primaria e in quella secondaria di primo grado. Gli ambiti oggetto di misurazione sono la lingua, la matematica, l'inglese, le scienze naturali e le scienze sociali con l'obiettivo di misurare la performance degli studenti tenendo conto dei processi sottesi e dell'influenza dell'ambiente circostante (Maestro Martin, 2005; Monarca & Vallev, 2014).

Oggi le valutazioni diagnostiche nazionali coordinate dall'INEE sono regolate dalla Ley Organica n. 3 del 29 dicembre 2020 (art. 21 e 29). Hanno luogo con carattere censuario nella quarta classe della scuola primaria e nella seconda classe della scuola secondaria di primo grado. Oggetto di analisi sono le competenze nella comunicazione linguistica e quelle matematiche. In questa sede non si entrerà nello specifico dei risultati raggiunti dagli studenti, tuttavia, è bene sottolineare che le amministrazioni educative supportano e monitorano le istituzioni scolastiche che da esse dipendono. Inoltre, queste ultime fanno in modo

che le scuole prendano in considerazione i risultati delle valutazioni per promuovere il cambiamento e il miglioramento. L'INEE, in quest'ottica, collabora con le amministrazioni educative locali per la realizzazione di quadri teorici comuni. Attualmente i team di esperti sono al lavoro per la realizzazione di nuovi quadri di riferimento e nuovi modelli di valutazione. A circa venti anni dall'istituzione dell'INEE, è utile ricordare la ricostruzione storica delle linee intraprese per la valutazione del sistema scolastico realizzata da Ferrer (2014): le politiche relative alla valutazione hanno adottato orientamenti diversi sia per quanto riguarda le finalità della valutazione, sia per quanto concerne il focus e gli strumenti utilizzati. L'autore individua nella storia della valutazione del sistema scolastico spagnolo tre tappe fondamentali: nella prima le prove sono state realizzate su un campione di studenti e a partire dai curricula esistenti per ciascun grado dell'istruzione, nella seconda le prove sono state realizzate con scopi diagnostici a partire dalle competenze di base acquisite, mentre nella terza le prove sono state individualizzate.

Il sistema generale di valutazione del sistema educativo costruito e sviluppato in questi due decenni presenta, quindi, continuità e rotture. Ci sono alcune linee di lavoro sostenute nel tempo, che hanno a che fare soprattutto con la costruzione di indicatori educativi e con la crescente partecipazione a studi internazionali. Ci sono anche elementi che sono cambiati, riferiti soprattutto all'oggetto della valutazione – contenuti curriculari o competenze di base -, alla sua finalità – sommativa o formativa - e alla collocazione delle prove – al termine o nelle fasi intermedie. Nel complesso, quello che si può dire è che questi vent'anni hanno visto la progettazione, l'organizzazione e lo sviluppo di un sistema generale di valutazione del sistema educativo, che dovrebbe dare ancora molti più frutti in futuro. (Ferrer, 2014, p. 19, traduzione a cura dell'autrice)

Sembrerebbe che dalle prime sperimentazioni degli anni Ottanta ai giorni nostri, il sistema di valutazione scolastica spagnolo abbia subito sviluppi e rotture nel tempo, confermandosi come un sistema ben radicato. Tuttavia, non tutti sono concordi. La *Ley Organica para la Mejora de la Calidad Educativa*, LOMCE, n. 8 del 9 dicembre 2013 ha introdotto alcuni cambiamenti nel sistema educativo spagnolo che coinvolgono

anche il significato attribuito alla valutazione del sistema scolastico. La nuova riforma, infatti, sostiene che l'obiettivo è quello di preparare al meglio gli studenti per il loro successivo inserimento nel mondo del lavoro. Nel fare ciò, però, secondo alcuni, si vuole lasciare il passo alla formazione di giovani più competitivi selezionati attraverso una serie di valutazioni che, se da una parte vogliono colmare alcuni vuoti lasciati dalle riforme precedenti, dall'altra portano le scuole a competere tra di loro per guadagnarsi un posto nella classifica del mercato educativo (Llorent Bedmar & Cobano-Delgado Palma, 2018). I più scettici, infatti, affermano che se la valutazione standardizzata viene applicata a contesti differenziati, non vengono tenute in considerazione le caratteristiche storiche e culturali proprie di quello specifico contesto che, essendo la Spagna uno stato federale, appaiono quanto più evidenti (Sanmartin, 2014; Saura & Luengo, 2015).

2.5.3. La valutazione degli apprendimenti di INVALSI

Il D.P.R. 80/13 oltre a implementare il Sistema Nazionale di Valutazione, regola anche le rilevazioni periodiche nazionali (D.P.R. 80/13, 2013).

Fermo restando quanto previsto dall'articolo 51, comma 2, del decreto-legge 9 febbraio 2012, n. 5, convertito, con modificazioni, dalla legge 4 aprile 2012, n. 35, le istituzioni scolastiche sono soggette a periodiche rilevazioni nazionali sugli apprendimenti e sulle competenze degli studenti, predisposte e organizzate dall'Invalsi anche in raccordo alle analoghe iniziative internazionali. Tali rilevazioni sono effettuate su base censuaria nelle classi seconda e quinta della scuola primaria, prima e terza della scuola secondaria di primo grado, seconda e ultima della scuola secondaria di secondo grado e comunque entro il limite, a decorrere dall'anno 2013, dell'assegnazione finanziaria disposta a valere sul Fondo di cui all'articolo 7 del decreto legislativo 5 giugno 1998, n. 204. (D.P.R. 80/13, 2013, art. 6, comma 3)

Nello Statuto dell'istituto è specificato che è compito di Invalsi, nell'ottica del supporto diffuso al sistema di istruzione e della riflessione didattica, costruire e utilizzare prove standardizzate (Invalsi, 2007). Le

rilevazioni periodiche nazionali, che si svolgono mediante la messa a punto di prove oggettive per la valutazione degli apprendimenti, prevedono, infatti, la misurazione di competenze ritenute fondamentali per l'apprendimento e per la vita degli studenti a partire dalle linee guida e dalle indicazioni nazionali, compresi i traguardi da raggiungere.

Di recente queste ultime hanno subito alcune modifiche tramite il Decreto Legislativo n. 62 del 13 aprile 2017 (D.L. 62/17, 2017) che ne ha modificato l'impatto e la relazione con gli esami conclusivi del primo e del secondo grado dell'istruzione superiore (Invalsi, 2018). Infatti, le prove standardizzate degli esami finali del primo e del secondo grado dell'istruzione superiore dall'anno scolastico 2017/2018 non fanno più parte dell'esame conclusivo, ovvero l'esito non incide più sul giudizio finale dello studente, ma rappresentano un requisito fondamentale per l'accesso a quest'ultimo: la prova viene somministrata nel mese di aprile e solo chi vi partecipa può accedere all'esame. L'esito è sintetizzato con un descrittore su una scala da 1 a 5 ed è riportato nella certificazione finale delle competenze dello studente. Nonostante le iniziali perplessità, tale soluzione permette di descrivere ciò che lo studente sa fare in relazione ai traguardi stabiliti nazionalmente, senza far pesare queste competenze sul voto finale dell'esame di stato (Invalsi, 2018).

Lo stesso decreto ha stabilito anche che a partire dall'anno scolastico 2017/2018 le prove standardizzate messe a punto da Invalsi si svolgono su piattaforma informativa e in modalità computerizzata (CBT, *Computer Based Tests*) nella scuola secondaria di primo e di secondo grado (terza classe della scuola secondaria di primo grado, seconda e quinta classe della scuola secondaria di secondo grado) (D.L. 62/17, 2017).

La modalità di svolgimento determina anche un cambiamento dell'impianto delle prove: non più lineari, cioè formate dalle stesse domande per tutti gli studenti, ma composte da un certo numero di quesiti differenti provenienti da un'unica banca di domande. Ogni prova formata in questo modo condivide con le altre la stessa difficoltà media e le stesse caratteristiche di contenuti e di tipologia di quesiti. Questo disegno consente inoltre di introdurre, entro certi limiti, degli elementi di differenziazione, pur mantenendo l'unitarietà della prova. In tal modo è possibile introdurre quesiti maggiormente appropriati per alcune filiere specifiche dell'istruzione e della formazione secondaria di

secondo grado, senza però perdere il vantaggio di risultati forniti su una scala unica, quindi direttamente comparabili. [...] Infine, ma non da ultimo, il D. Lgs. 62/2017 richiama esplicitamente il D.P.R. 80/2013, istitutivo del Sistema Nazionale di Valutazione, ribadendo il ruolo fondamentale delle prove nel più ampio contesto dell'intero processo di valutazione e autovalutazione delle scuole e del sistema educativo. (Invalsi, 2018a, p. 4)

Le prove misurano le competenze in italiano, matematica e inglese. Mentre per l'italiano e la matematica i Quadri di Riferimento sono ispirati alle linee guida e alle indicazioni nazionali, per l'inglese il punto di riferimento è il *Common European Framework of Reference for Languages*.

I Quadri di Riferimento messi a punto da Invalsi esemplificano i riferimenti concettuali, i collegamenti con le indicazioni di legge nazionali e la progettazione delle prove. Le indicazioni nazionali per i curricoli hanno subito negli ultimi anni profonde modifiche, in questo senso l'Invalsi ha cercato di sviluppare i Quadri di Riferimento cogliendo la continuità tra un ciclo di istruzione e l'altro (Invalsi, 2018a).

Il QdR è pensato, in primo luogo, per aiutare gli attori del sistema scolastico (insegnanti, dirigenti, famiglie) a interpretare i risultati ottenuti dalle singole scuole o dalle singole classi nelle prove delle Rilevazioni Nazionali. La comparazione dei risultati delle proprie classi o della propria istituzione scolastica con gli esiti complessivi delle prove, interpretati alla luce della conoscenza del contesto specifico in cui la propria scuola opera, può servire per individuare i punti di forza e di debolezza del percorso effettivamente realizzato in classe e delle scelte didattiche effettuate; può inoltre aiutare il coordinamento della progettazione didattica all'interno delle singole istituzioni scolastiche. Il QdR può essere adoperato dai responsabili ai diversi livelli (Ministero dell'Istruzione, Uffici Scolastici Regionali, Dirigenti scolastici) come strumento per interpretare i risultati del sistema nel suo complesso, per poter adottare opportune strategie di intervento, per esempio relativamente alla predisposizione di attività particolari di recupero o rafforzamento per gli studenti o di piani di formazione in servizio per i docenti. Il QdR, inoltre, può offrire agli studenti e alle famiglie informazioni utili per capire il significato della valutazione come momento cruciale del percorso scolastico e come momento di verifica del sistema. Il QdR si rivolge infine alle persone che propongono i quesiti e ai gruppi di lavoro che a diversi livelli li elaborano, ne

seguono i pretest sul campo e sulla base dei risultati di essi compongono le prove. Indica i vari aspetti dell'apprendimento da valutare e stabilisce un equilibrio tra i diversi ambiti. È quindi uno strumento fondamentale nella fase preparatoria delle prove. Tutte queste osservazioni portano a riflettere sull'importante effetto di ricaduta che il complesso delle prove delle Rilevazioni Nazionali ha sull'intero sistema scolastico e sulle sue scelte didattiche. È proprio in questo senso che un'attenta analisi dei risultati delle prove somministrate potrà contribuire a fornire una guida per il miglioramento dell'offerta del sistema di istruzione nel suo complesso e di ogni istituzione scolastica. (Invalsi, 2018a, p. 2-3)

Spesso e soprattutto in passato le prove delle rilevazioni nazionali sono state criticate per la loro impostazione e per i loro contenuti (Trincherò, 2014c). Secondo i Quadri di Riferimento Invalsi messi a punto, in italiano ad essere testate sono le competenze, ritenute fondamentali in ogni ciclo scolastico, di lettura e lessicali e la riflessione sulla lingua attraverso tipi e generi testuali diversi, similmente alle indagini internazionali sugli apprendimenti (OECD, 2019abc), per le classi seconda e quinta della scuola primaria, per il secondo anno della scuola secondaria di secondo grado e per le classi terminali della scuola secondaria di primo e di secondo grado.

La comprensione di un testo è frutto di un processo interattivo che risulta dall'integrazione ottimale del dato testuale con le conoscenze e le aspettative del lettore. Per comprendere, interpretare e valutare un testo il lettore deve essere in grado di individuare specifiche informazioni, ricostruire il senso globale e il significato di singole parti, cogliere l'intenzione comunicativa dell'autore, lo scopo del testo e il genere cui esso appartiene. Questo modo di intendere il processo di lettura è sotteso anche alle indagini internazionali sulla lettura promosse, per esempio, dall'OCSE (Organizzazione per la Cooperazione e lo Sviluppo Economico) e dall'IEA (*International Association for the Evaluation of Educational Achievement*). Ai fini dell'elaborazione della prova INVALSI di Italiano e della valutazione, si distinguono tre dimensioni costitutive della competenza complessiva di lettura: pragmatico-testuale, lessicale, grammaticale. Va tuttavia sottolineato che nel processo di comprensione del testo le competenze lessicale, pragmatico-testuale e grammaticale interagiscono

fra loro e non sono facilmente e nettamente distinguibili l'una dall'altra. (Invalsi, 2018b, p. 2)

La dimensione pragmatico-testuale è la capacità di ricostruire l'insieme di significati che il testo veicola a partire dal contesto in cui si inserisce. La dimensione lessicale è la capacità di ricostruire il significato di un vocabolo, riconoscendone le relazioni di significato con altri vocaboli. Infine, la dimensione grammaticale è la capacità di ricorrere alla grammatica esplicita ed implicita del testo per risolvere e favorire la comprensione (Invalsi, 2018b).

In matematica, con enfasi diverse per i differenti cicli, ad essere testate sono le competenze enucleate nelle indicazioni nazionali e raggruppate in dati e previsioni, numeri, relazioni e funzioni, spazio e figure che si articolano nelle tre dimensioni del conoscere, risolvere problemi e argomentare. I Quadri di Riferimento Invalsi per la matematica, in particolare, si collegano ai framework delle indagini IEA-TIMSS e OCSE-PISA. Tralasciando le differenze interne a questi due studi internazionali orientati il primo al curricolo e il secondo alle competenze imprescindibili per i quindicenni dei paesi partecipanti, tra la rilevazione nazionale Invalsi e quelle internazionali è possibile integrare le informazioni restituite: il lavoro di messa a punto dei Quadri di Riferimento italiani contempla sia la capacità di leggere la realtà circostante propria di PISA, sia gli aspetti relativi allo sviluppo dei curricoli propria di TIMSS (Mullis & Martin, 2017; Invalsi, 2018ab; OECD, 2019abc).

Naturalmente la continuità in termini di contenuti specifici e di conoscenze e abilità a essi legate si caratterizza in modo nettamente più marcato nel passaggio da un grado scolare al successivo, tanto è vero che diverse domande delle prove del Grado 5 possono essere utilizzate per le prove del Grado 8; allo stesso modo molte domande delle prove del Grado 8 possono essere utilizzate per valutare competenze degli studenti del Grado 10 e così via. In particolare, limitandoci a considerare le domande delle prove dei gradi 8 e 10, osserviamo che, per esempio, alcune di quelle relative ai concetti di percentuale, relazione funzionale fra grandezze, distribuzione statistica, area di una figura, rappresentazione polinomiale dei numeri, divisibilità, probabilità,

trasformazione geometrica, sono assai simili fra loro, tanto da poter essere proposte sia nelle prove del Grado 8 sia in quelle del Grado 10. (Invalsi, 2018a, p. 11)

L'utilizzo di prove standardizzate per la rilevazione degli apprendimenti se da una parte è un momento fondamentale del Sistema Nazionale di Valutazione italiano, dall'altra presenta dei limiti, si pensi per esempio alla loro discussa validità (Corsini et alii, 2017).

Ogni prova viene quindi costruita pensando non solo agli apprendimenti specifici di quel grado scolare, ma anche alle competenze generali e permanenti che lo studente, in quello specifico grado scolare, dovrebbe aver acquisito grazie a tutto il percorso precedente. Le prove possono quindi contenere domande che accertano il permanere di abilità o di conoscenze acquisite in gradi scolari precedenti, la cui padronanza è essenziale per l'esercizio di queste competenze. Quanto appena detto è fondamentale per chiarire limiti e compiti dell'azione del sistema di valutazione, in particolare attraverso l'uso di prove standardizzate. Questo strumento, che risulta particolarmente utile a dare informazioni su alcune criticità del sistema nel suo complesso, è poco adatto a misurare il conseguimento di alcuni dei Traguardi per il raggiungimento delle competenze elencati nelle Indicazioni. Per esempio mentre le prove standardizzate sono uno strumento utile per valutare i traguardi "Descrive, denomina e classifica figure in base a caratteristiche geometriche, ne determina misure" e "Riconosce e quantifica, in casi semplici, situazioni di incertezza", possono essere meno adeguate a valutare pienamente un traguardo come "Costruisce ragionamenti formulando ipotesi, sostenendo le proprie idee e confrontandosi con il punto di vista di altri" e sono sicuramente non adeguate a valutare il traguardo "Sviluppa un atteggiamento positivo rispetto alla matematica". In questo senso le prove standardizzate non devono né possono sostituirsi all'azione dell'insegnante per la valutazione degli studenti della propria classe: i risultati delle prove possono essere utilizzati come una fra le molte informazioni di cui l'insegnante deve entrare in possesso per valutare dinamicamente le prestazioni dei propri studenti e certificare le competenze raggiunte, ma non possono costituire l'unica sorgente di informazioni né quella più importante per la valutazione del singolo alunno, essenzialmente per tre motivi:

- le prove standardizzate, come già detto, non possono misurare né tantomeno valutare il conseguimento di traguardi caratterizzati da aspetti

metacognitivi o non cognitivi, come per esempio “Sviluppa un atteggiamento positivo rispetto alla matematica, attraverso esperienze significative, che gli hanno fatto intuire come gli strumenti matematici che ha imparato a utilizzare siano utili per operare nella realtà”; analogamente, almeno come sono attualmente strutturate, le prove standardizzate non consentono di valutare la competenza raggiunta dagli studenti nell’uso di tecnologie per costruire significati degli oggetti di studio;

- le prove standardizzate sono poco adatte a valutare pienamente il conseguimento di competenze nel: produrre congetture; sostenere argomentazioni e dimostrazioni complesse; porsi e risolvere problemi di una certa difficoltà e complessità, che richiedono diversi passi per essere affrontati e risolti; costruire e utilizzare modelli per situazioni complesse;

- le prove standardizzate tendono a porre domande per le quali la correzione delle risposte è indipendente da chi la effettua; in altri termini tendono, per quanto possibile, a eliminare ogni elemento di soggettività. Invece la valutazione che l’insegnante fa delle competenze conseguite da uno studente non può non essere soggettiva, dipendente anche da fattori sociali e affettivi, e tesa anche a promuovere l’acquisizione stessa delle competenze: tanto più è soggettiva, tanto più diventa adatta allo studente che viene valutato e quindi ricca e significativa per quello studente. (Invalsi, 2018a, p. 12-13)

In inglese ad essere testate sono le competenze riportate nel *Common European Framework of Reference for Languages: learning, teaching, assessment* (Council of Europe, 2001), un sistema di descrittori utili a comprendere a che livello si parla, si comprende e si scrive in una lingua straniera dal livello di base A1 al livello di padronanza C2. Al suo interno, grazie al lavoro del Consiglio Europeo, sono contenute anche indicazioni sui criteri di valutazione della padronanza linguistica (Council of Europe, 2001).

2.6. Come avviene la valutazione delle scuole?

Nei paragrafi precedenti sono state descritte le pratiche legate alla valutazione degli apprendimenti come parte di quella valutazione esterna intenta a monitorare la qualità dei sistemi educativi tramite la misurazione degli outcomes degli studenti. Nei sottoparagrafi che seguono verranno, invece, descritte la valutazione esterna intesa come la

possibilità di effettuare visite ispettive, là dove esistenti, e l'autovalutazione delle istituzioni scolastiche. I dati contenuti nel report Eurydice intitolato *Assuring Quality in Education: Policies and Approaches to School Evaluation in Europe* (Eurydice, 2015) verranno aggiornati e integrati in conformità con le pratiche attuali di valutazione delle istituzioni scolastiche. Le modalità attraverso le quali vengono valutate le scuole sono legate a ciò che si è scelto di osservare. Precedentemente ci si è soffermati sulla valutazione degli apprendimenti nazionali, là dove presente sistematicamente, tuttavia essendo la valutazione dei sistemi scolastici fondata su oggetti di osservazione eterogenei, si preferisce al giorno d'oggi integrare tecniche quantitative, si pensi ai test standardizzati, con tecniche più qualitative, si pensi ai colloqui tra il personale della scuola e i valutatori esterni o alle riflessioni interne agli attori che vivono in un determinato contesto scolastico (Cardone et alii, 2010).

2.6.1. La valutazione esterna e l'autovalutazione delle istituzioni scolastiche in Finlandia

La valutazione condotta dagli ispettori esterni in Finlandia non è regolamentata e sistematica. Ciò che è comunemente definito *quality assurance*, il sistema di assicurazione della qualità, nel caso specifico a livello scolastico, si basa prevalentemente sulle pratiche di autovalutazione messe in atto da coloro che si occupano di fornire istruzione o sui monitoraggi e sulle valutazioni esterne e tematiche effettuate da KARVI, di cui si è discusso nei paragrafi 2.4 e 2.5 del presente capitolo. Le valutazioni nazionali esterne condotte da questo istituto di ricerca, infatti, si focalizzano sulla rilevazione dello stato di salute dell'intero sistema educativo, non sulle singole scuole, e non hanno carattere ispettivo. I dati sono restituiti a queste ultime affinché ne possano tenere conto, ma non sotto forma di controllo. Le autorità locali e le istituzioni scolastiche, che sono responsabili per la gran parte dell'organizzazione e dell'erogazione dell'offerta formativa, ad eccezione delle linee guida emanate centralmente, hanno l'obbligo di sottoporre a valutazione la propria offerta educativa e di partecipare alle

valutazioni tematiche nazionali. Le pratiche e le procedure di valutazione locali sono stabilite dai municipi e dalle istituzioni scolastiche e possono prevedere valutazioni esterne delle singole scuole, quelle degli apprendimenti, per sostenerne lo sviluppo educativo e migliorarne le condizioni per l'apprendimento, e valutazioni interne (Eurydice, 2015).

L'autovalutazione in Finlandia è sotto la responsabilità di chi eroga i servizi educativi (es. municipi, singole istituzioni scolastiche, università), non è regolamentata da specifiche norme, tanto è vero che le pratiche di valutazione non sono strettamente definite ma lasciate alla libertà delle autorità locali, e fa riferimento al sistema di assicurazione della qualità (Eurydice, 2015). Chi eroga i servizi educativi, infatti, ha la possibilità di costruire un framework di riferimento ad hoc in base alle aree ritenute più rilevanti, alla frequenza e alle modalità e ai metodi di valutazione. Gli obiettivi dell'autovalutazione sono descritti nei curricula locali della scuola e nel piano annuale elaborato da quest'ultima. A livello centrale, invece, è stato messo a punto un piano guida per la valutazione e lo sviluppo della qualità. Nell'anno scolastico 2009/2010, infatti, il Ministero dell'Educazione e della Cultura ha reso disponibile a tutte le autorità locali uno strumento a supporto delle pratiche di autovalutazione per la garanzia della qualità sia a livello di singolo istituto scolastico che municipale (Opetusministeriö, 2010).

Lo scopo dei criteri nazionali di qualità per l'istruzione di base è sostenere l'offerta di un'istruzione di alta qualità e versatile e garantire i diritti educativi e culturali di base dei bambini e dei giovani, indipendentemente dal luogo di residenza, dalla lingua e dal background economico-familiare. L'uso di criteri di qualità aiuta le scuole, il personale, l'amministrazione e i decisori politici a prendere coscienza delle criticità e a migliorarle. Sono uno strumento di gestione utilizzato per valutare lo stato attuale e il possibile sviluppo dell'educazione. Nei complessi processi di valutazione e cambiamento a livello locale, i criteri di qualità dell'istruzione di base sono anche uno strumento per i professionisti dell'istruzione che si basa su un processo di ricerca, valutazione e validazione delle informazioni. Questi aiutano le scuole a valutare e sviluppare in modo flessibile gli interventi educativi locali. [...] Sono, inoltre, uno strumento che può essere utilizzato per produrre informazioni a livello locale

sulla qualità strutturale e funzionale delle istituzioni educative. Si tratta di un nuovo tipo di strumento di monitoraggio, che viene utilizzato per guidare l'organizzazione e i processi decisionali nell'istruzione di base. Tuttavia, non ha una natura legislativa. I criteri di qualità per l'educazione di base sono di natura raccomandativa, si configurano come strumento di gestione per la scuola e le autorità locali sia per un periodo più lungo che per un periodo più breve. I criteri di qualità determinano, quindi, sulla base di quali benchmark possono essere analizzate le attività e quali sono le aspettative in merito ad esse in modo che possano essere funzionali agli obiettivi prefissati, consentendo di valutare le attività educative implementate. [...] Per lo sviluppo e il miglioramento della scuola di base finlandese, non viene prestata attenzione solamente ai risultati negli apprendimenti, ma anche a quegli aspetti strutturali e logistici che consentono e promuovono il raggiungimento degli obiettivi di apprendimento prefissati. Lo scopo dei criteri di qualità è anche quello di sostenere l'insegnamento, l'organizzazione e la collaborazione tra le scuole e il territorio mediante modelli di azione che promuovono l'apprendimento e lo sviluppo degli studenti. L'attenzione è richiamata, infatti, sui fattori che implementano in modo significativo le condizioni di apprendimento degli studenti e dell'intera comunità scolastica. [...] In questo modo i criteri di qualità supportano anche la valutazione di sistema e la diffusione e l'utilizzo degli esiti della valutazione nei processi decisionali, soprattutto quelli locali. (Opetusministeriö, 2010, p. 9-10, traduzione a cura dell'autrice)

Nella messa a punto dei criteri di qualità le autorità centrali hanno tenuto conto delle differenti "culture operative" che chi eroga servizi può mettere in atto grazie all'autonomia decisionale e pedagogica. L'obiettivo non è quello di utilizzare questi ultimi in maniera acritica, bensì di fare in modo che, attraverso le domande guida che li compongono, ogni istituzione scolastica possa riflettere criticamente sul proprio operato.

Il modello adottato fa riferimento al Ciclo di Deming, *Ciclo PDCA - Plan - Do - Check - Act*, e all'interazione continua tra Pianificare - Fare - Verificare - Agire, quattro fasi adoperate per il monitoraggio e il miglioramento continuo dei prodotti e dei processi (Dahlgaard et alii, 1995). Il modello, utilizzato nello sviluppo della qualità, permette di monitorare le azioni in modo graduale e sistematico, separando la pianificazione, l'implementazione e la valutazione da altri stadi del

miglioramento. È possibile garantire la qualità nell'istruzione sviluppando in un processo continuo la qualità dell'insegnamento e dell'apprendimento. Al contrario, la qualità degli erogatori dei servizi educativi è possibile garantirla attraverso la messa a punto di una strategia negoziata con tutti gli stakeholders e attraverso la condivisione di valori, visioni e idee di interventi comuni. La qualità dell'istruzione di base può essere esaminata anche attraverso le prospettive, gli attori coinvolti e i processi. Per analisi delle prospettive, nel modello, si intende la conoscenza sottesa alle azioni intraprese dall'istituzione scolastica, un esempio sono le *best practices*. Per analisi degli attori coinvolti si intendono le loro percezioni e le mansioni svolte. Per analisi dei processi, infine, si intende l'approfondimento delle azioni intraprese dall'istituzione educativa in un'ottica di sviluppo migliorativo (Opetusministeriö, 2010).

In una scuola orientata allo sviluppo, l'accento è posto sull'autovalutazione e sull'implementazione delle operazioni a partire dai dati raccolti attraverso le valutazioni sistematiche. Nel processo di riforma e innovazione in prima istanza dovrebbero sempre essere presi in considerazione la cultura operativa e i punti di partenza propri dell'istituzione scolastica. L'autovalutazione delle scuole, infatti, risponde a quattro domande fondamentali: dove siamo adesso? Dove vorremmo essere in futuro? Come possiamo raggiungere al meglio il nostro obiettivo? Come valutiamo i cambiamenti raggiunti? (Opetusministeriö, 2010, p. 15, traduzione a cura dell'autrice)

Si tratta di linee guida ampiamente utilizzate dai fornitori di servizi educativi anche perchè prevedono il coinvolgimento di tutti gli stakeholders: il punto di vista delle autorità municipali, degli studenti, delle famiglie, degli insegnanti e dei dirigenti scolastici è fondamentale non solo per il coinvolgimento nei momenti decisionali, ma anche per la buona riuscita del sistema di assicurazione della qualità. Così come chi eroga servizi educativi ha la possibilità di costruire il framework di riferimento e i metodi propri delle pratiche valutative, così ha anche la possibilità di individuare il tipo di supporto o strumento fornito alle scuole, oltre ai criteri che guidano la qualità e il miglioramento elaborati dal Ministero dell'Educazione e della Cultura (Eurydice, 2015). Nello

specifico, lo strumento elaborato centralmente ha lo scopo di sostenere le scuole e le autorità pubbliche a identificare le carenze e sviluppare misure correttive per migliorare gli interventi educativi.

La legge n. 628 del 21 agosto 1998 (Laki 628/98, 1998) sancisce l'obbligo delle istituzioni che erogano servizi educativi di autovalutare il proprio operato ed esorta queste ultime a pianificare le proprie attività in base agli esiti delle valutazioni e dei monitoraggi. In quest'ottica i criteri di qualità dell'istruzione di base offrono ai decisori politici, compresi quelli municipali e interni alla scuola, una base a partire dalla quale prendere decisioni a breve e a lungo termine, in quanto, soprattutto nella rilevazione di criticità, sia la scuola sia il municipio ne possono tenere conto nella pianificazione annuale e nella ripartizione dei budget economici (Laki 628/98, 1998). Lo strumento, elaborato centralmente e messo a disposizione di tutte le scuole come supporto all'autovalutazione, comprende una parte dedicata alle autorità locali e ai municipi. I decisori politici delle autorità locali, infatti, dovrebbero essere in grado, secondo il documento, di rispondere alle seguenti domande:

- Come viene garantita la valutazione della qualità e come questa valutazione viene presa in considerazione nel processo decisionale?
- Quale tipologia di servizi, a partire dagli esiti delle valutazioni, si vogliono organizzare per la comunità? (Opetusministeriö, 2010)

La discrepanza tra il calendario finanziario dei municipi e il calendario scolastico, tuttavia richiede un'azione di allineamento sia da parte delle scuole che da parte delle autorità locali.

Durante il triennio di programmazione finanziaria l'amministrazione, in conformità con le aree definite prioritarie, può realizzare la valutazione della qualità dell'istruzione di base, in modo tale che in tutte le scuole siano valutati contemporaneamente gli stessi elementi sulla base dei criteri di qualità. Il consiglio di amministrazione responsabile dell'istruzione può decidere di integrare le attività di valutazione della scuola e il piano di valutazione nel piano annuale della scuola affinché gli esiti della valutazione vengano utilizzati sistematicamente dalle istituzioni scolastiche e a livello municipale nel processo decisionale e nella pianificazione degli interventi e delle operazioni finanziarie. (Opetusministeriö, 2010, p. 17, traduzione a cura dell'autrice)

Le autorità locali e gli erogatori di servizi educativi hanno la facoltà di decidere in che modo e in quale misura i criteri di qualità dell'istruzione di base possono essere utilizzati nell'implementazione dell'istruzione, di decidere come verrà valutato il loro raggiungimento, di decidere in che modo gli esiti della valutazione possono essere utilizzati in futuro in un'ottica migliorativa, di decidere come informare sui criteri di qualità e di decidere se il lavoro sulla qualità dell'istruzione di base è a livello regionale o locale. Le autorità locali e gli erogatori di servizi educativi, inoltre, hanno la responsabilità di organizzare il lavoro sui criteri di qualità dell'istruzione di base e di tenere in considerazione i risultati delle valutazioni (Opetusministeriö, 2010).

Nell'ottica dell'autovalutazione delle istituzioni scolastiche e dell'utilizzo degli esiti di quest'ultima nei processi decisionali a livello locale, la cooperazione tra i partner e il lavoro di rete per il raggiungimento della qualità nell'istruzione di base si configura come un elemento fondamentale. La collaborazione, infatti, permette di raggiungere un'intesa regionale sugli elementi e le modalità valutative, ma permette anche di ricevere feedback costruttivi da occhi esterni coinvolti allo stesso modo nei processi valutativi, nonché di mettere in primo piano gli obiettivi da raggiungere e le *best practices*.

Nel riportare i risultati della valutazione della scuola, in essi bisognerebbe trovare la risposta alle seguenti domande che guidano la pianificazione strategica delle autorità locali:

- Cosa ci è richiesto e come noi, professionisti dell'educazione, vogliamo agire?

- Quali dovrebbero essere le nostre attività, ad esempio l'insegnamento e l'osservazione dei regolamenti in materia di istruzione e ricerca, sulla base delle informazioni provenienti dalla valutazione?

- Come assicuriamo che tutti raggiungano i risultati di apprendimento richiesti?

La scuola ha bisogno dei criteri di qualità per l'istruzione e di informazioni basate sui punti di vista di decisori politici, studenti, famiglie, gruppi di lavoro legati alla scuola e personale docente. È essenziale fissare gli obiettivi in un formato comprensibile. Sulla base di questi ultimi è possibile valutare le attività

e gli interventi realizzati. Dal momento che i dati della valutazione sulle attività della scuola sono integrati nel processo decisionale e nel sistema di pianificazione operativa ed economica, le informazioni sull'autovalutazione dovrebbero essere raccolte sistematicamente nelle scuole dall'inizio dell'anno scolastico. (Opetusministeriö, 2010, p. 20, traduzione a cura dell'autrice)

Il processo di autovalutazione delle istituzioni scolastiche finlandesi, dunque, si esplica in cinque punti:

- Istituzione di un gruppo di lavoro che è a supporto del monitoraggio della qualità e dell'organizzazione delle attività;
- Definizione di descrittori e criteri a partire dalle linee guida ministeriali e a partire dai criteri adottati dal proprio comune o dalla propria regione;
- Processo di autovalutazione;
- Analisi dei dati e delle informazioni raccolte e messa a punto degli interventi di miglioramento;
- Diffusione degli esiti anche a livello locale per l'implementazione di un piano di azione in collaborazione con le autorità municipali.

La qualità è valutata annualmente sulla base dei criteri stabiliti dai gruppi di lavoro multidisciplinari municipali che comprendono dirigenti scolastici, insegnanti, rappresentanze amministrative, consulenti, psicologici scolastici e altri professionisti che operano nei servizi a sostegno della didattica e del benessere degli studenti. La cadenza annuale è progettata anche in base alla pianificazione del bilancio. Infatti, le istituzioni scolastiche si autovalutano e programmano interventi educativi sulla base dell'autonomia loro concessa (Opetusministeriö, 2010).

Nel quadro dei criteri, la qualità dell'istruzione di base si compone della qualità delle strutture e della qualità dei processi. La qualità delle strutture riflette la qualità delle strutture esterne, mentre la qualità dei processi riflette tutti i processi interni alla scuola, vissuti in particolar modo dagli studenti. La prima tipologia di qualità, delle strutture, si articola a sua volta in questioni regolate a livello nazionale che hanno un impatto a livello locale e organizzazione didattica e programmazione delle attività scolastiche. I descrittori specifici all'interno dei quali si

muove l'autovalutazione delle istituzioni scolastiche sono: gestione, personale, risorse finanziarie e valutazione. La seconda tipologia di qualità, dei processi, invece, pone al centro lo studente e il suo percorso scolastico ed educativo. I descrittori specifici della qualità dei processi sono: l'attuazione del curriculum, l'insegnamento, il sostegno all'apprendimento, alla crescita e al benessere, la partecipazione, il rapporto scuola-famiglia, l'ambiente di apprendimento, la sicurezza dell'ambiente di apprendimento (Opetusministeriö, 2010; Eurydice, 2015). Per ognuno di questi undici descrittori dal ministero è stata definita una scheda che comprende il framework di riferimento entro il quale si articola il descrittore e i criteri che lo compongono. Per ognuno di questi, inoltre, sono state definite delle domande guida per l'autovalutazione. A titolo esemplificativo si riporta un estratto della scheda del descrittore sulla valutazione.

Criteri di qualità:

- Vengono svolte attività di valutazione e sviluppo in modo sistematico e pianificato.
- Per l'organizzatore dell'insegnamento e della scuola si dispone di un sistema di valutazione adeguato alle esigenze.
- Le informazioni raccolte e prodotte tramite la valutazione sono utilizzate nell'implementazione dell'insegnamento e del funzionamento e nella pianificazione economica.
- Le competenze valutative degli erogatori di servizi educativi e delle scuole sono sviluppate in modo sistematico.
- Le autorità locali e la scuola hanno un corpus di principi valutativi concordati e condivisi.
- Ai processi valutativi e migliorativi partecipano anche studenti, famiglie, addetti alla cura degli studenti e altro personale, se necessario.

Domande guida per l'implementazione della qualità:

- Come garantire che gli erogatori di servizi educativi conducano con sistematicità attività di valutazione e che la valutazione produca sostegno allo sviluppo e alla conoscenza?
- Come vengono utilizzati i dati raccolti per l'implementazione della didattica e la pianificazione operativa?

- Come vengono sviluppate le competenze valutative dell'erogatore di servizi educativi e delle scuole?
- Come vengono restituiti i dati della valutazione agli erogatori di servizi educativi, al personale, agli studenti e alle famiglie?
- Come gli studenti e le loro famiglie partecipano alla valutazione e al suo sviluppo? (Opetusministeriö, 2010, p. 38, traduzione a cura dell'autrice)

In tal senso è compito del decisore locale decidere se e in quale misura i criteri stabiliti centralmente vengono adottati localmente. Inoltre, le scuole che partecipano alle valutazioni nazionali condotte da KARVI ricevono i dati di queste ultime che gli permettono di confrontare non solo se stesse con le medie, ma anche utilizzare quelle informazioni nelle pratiche di autovalutazione (Eurydice, 2015).

L'utilizzo dei dati delle pratiche valutative varia molto a seconda delle autorità locali e delle istituzioni scolastiche alla luce dell'autonomia che a queste ultime viene concessa. Chi eroga servizi educativi non è obbligato, infatti, a diffondere gli esiti del sistema di assicurazione di qualità o delle rilevazioni condotte da KARVI. Le linee guida emanate dal Ministero dell'Educazione e della Cultura, tuttavia, dedicano una parte alle raccomandazioni sull'uso ottimale dei risultati delle valutazioni come strumento di gestione e miglioramento della scuola, inoltre, promuovono la discussione aperta tra tutti gli stakeholders con una conseguente messa a punto di un piano di azioni migliorative, tenendo conto del territorio all'interno del quale si inserisce l'istituto scolastico (Eurydice, 2015).

L'autovalutazione della scuola significa la valutazione da parte della scuola dei propri punti di forza e delle proprie debolezze. Alcuni dati importanti vengono forniti dalle valutazioni esterne (es. valutazione locale, valutazione del benessere), dalla valutazione della qualità e dalle famiglie e dagli studenti sotto forma di sondaggi/questionari o altri momenti di confronto.

Spesso nel mese di aprile, la scuola organizza una giornata dedicata alla valutazione in cui viene applicato il metodo CAF, *Common Assessment Framework*. Il dirigente scolastico nomina un gruppo di

lavoro che comprende rappresentanti di insegnanti, studenti, famiglie e possibilmente altro personale scolastico.

Il gruppo si riunisce per una giornata lavorativa per rispondere a domande derivanti dai criteri di qualità nazionali per l'istruzione di base e dai criteri di qualità provinciali. Il referente per la qualità della scuola organizza la giornata, raccoglie i risultati e li consegna alla persona che valuta l'attività didattica. Di seguito è riportato un esempio di report redatto dal gruppo di lavoro e disponibile sul sito istituzionale della scuola.

Gli esiti specifici della valutazione della scuola, nel complesso e presi separatamente, formano obiettivi di sviluppo per l'autovalutazione della scuola. [...] L'obiettivo di sviluppo per il prossimo anno scolastico è mantenere le buone pratiche. Il punteggio più basso è stato assegnato all'area "partecipazione e influenza". Quest'ultimo sarebbe potuto essere più alto se ci fosse stata un'elaborazione di programmi scritti per il miglioramento delle problematiche riscontrate. [...] Ogni anno, il dirigente scolastico conduce un monitoraggio regolare dell'insegnamento al fine di ottenere ulteriori informazioni sull'attività didattica ed educativa della scuola attraverso il programma di seguito riportato e in base al numero di insegnanti nella scuola:

- 2-3 insegnanti, 2 sessioni di follow-up
- 4-9 insegnanti, 4 sessioni di follow-up
- 10-15 insegnanti, 6 sessioni di follow-up
- più di 15 insegnanti, 8 sessioni di follow-up

(<https://peda.net/orimattila/peruskoulut/pl/vkl2/arviointi/ki/ki>, traduzione a cura dell'autrice)

Il dirigente scolastico, inoltre, è chiamato ad organizzare momenti di discussione con i docenti della scuola durante l'anno scolastico per confrontarsi, ad esempio, sull'implementazione del curriculum. I docenti e il personale, a loro volta, sono chiamati a compilare un modulo che è discusso durante l'incontro. Gli incontri non sono pubblici, ma le idee emerse possono essere utilizzate per il miglioramento dell'istituzione scolastica. Inoltre, in ottica di monitoraggio, molto spesso nelle scuole si tiene una riunione settimanale del personale in cui, alla luce dei punti di forza e di debolezza evidenziati, si pianifica il lavoro

della settimana successiva. Le riunioni settimanali, infatti, supportano anche la valutazione degli interventi in base ai successi e alle sfide. Con cadenza annuale, generalmente nel mese di settembre, le istituzioni scolastiche sono chiamate, inoltre, ad elaborare un piano di azione annuale reso pubblico sul sito peda.net, un network di scuole in cui è possibile consultare i documenti pubblicati da queste ultime (<https://peda.net/>). Gli obiettivi valutativi contenuti nel piano di azione annuale, infatti, derivano dalla giornata di valutazione: nel mese di maggio-giugno la scuola prepara una relazione sulle attività svolte nell'anno scolastico precedente (Lehtonen, 2013).

2.6.2. La valutazione esterna e l'autovalutazione delle istituzioni scolastiche in Spagna

La valutazione esterna in Spagna è sotto la responsabilità del Ministero dell'Educazione e della Formazione Professionale, delle Comunità Autonome e dell'Ispettorato dell'Istruzione, ovvero funzionari pubblici nel ruolo di ispettori presenti in ogni comunità autonoma, Dipartimento Regionale e Divisione Territoriale (unità più piccola della comunità autonoma) (Eurydice, 2015). La Ley Organica n. 3 del 29 dicembre del 2020 che modifica le legge n. 2 del 3 maggio 2006 sull'istruzione (Ley Organica 3/20, 2020) stabilisce che, essendo di competenza del potere pubblico, è compito dello stato disporre e fare in modo che si realizzi la valutazione esterna delle istituzioni scolastiche in tutti gli ambiti territoriali. L'ispettorato, nel rispetto dei diritti fondamentali, in difesa delle discriminazioni e con imparzialità ed efficienza, supervisiona il funzionamento delle istituzioni educative e dell'offerta formativa sia pubblica che privata, avendone libero accesso, monitora le strategie di gestione di queste ultime, esaminando la documentazione pedagogica e amministrativa, supporta il miglioramento continuo, garantisce che le istituzioni educative rispettino le leggi emanate dallo stato e i regolamenti e collabora in qualità di autorità pubblica con i responsabili locali dei servizi educativi, dirigenti scolastici compresi. Inoltre, l'ispettorato, a livello centrale, ha il compito di realizzare sia rilevazioni che report periodici sulle

valutazioni ordinarie e straordinarie, vale a dire quelle segnalate dalle autorità educative in caso di mancata osservazione dei regolamenti, di verificare che le linee guida nazionali siano incluse nei curricula stabiliti localmente da ciascuna scuola, di garantire il rispetto e l'applicazione dell'equità e dell'uguaglianza per tutti gli studenti e tutte le studentesse e di verificare la concessione di borse di studio. Le comunità autonome possono disporre per l'ispettorato altre funzioni e organizzazioni specifiche, i valutatori esterni, infatti, possono partecipare, nel rispetto dell'esercizio dell'autonomia, alle riunioni collegiali scolastiche e ai consigli locali (Ley Organica 3/20, 2020). Localmente possono essere redatti piani di azione ispettivi annuali e pluriennali, possono essere definite aree di valutazione prioritarie o attività aggiuntive. Ogni comunità autonoma, quindi, stabilisce un proprio programma di valutazione che comprende le aree specifiche, la frequenza e gli indicatori da utilizzare (Eurydice, 2015).

In Andalusia, ad esempio, il Piano Generale di Attuazione della valutazione esterna per il periodo 2019-2023 comprende categorie generali di organizzazione e gestione della scuola e ha l'obiettivo di:

- Garantire l'esercizio dei diritti e l'adempimento dei doveri sulla base della normativa vigente;
- Supervisionare e valutare le istituzioni educative e l'offerta formativa in un'ottica migliorativa;
- Rilevare e promuovere la diffusione delle buone pratiche e delle innovazioni sviluppate e accreditate nelle istituzioni educative, nei servizi, nei programmi e nelle attività;
- Informare l'amministrazione centrale sulle risorse, sui processi e sui risultati del sistema educativo al fine di facilitare il processo decisionale;
- Promuovere la formazione, lo sviluppo professionale e il riconoscimento dell'ispettorato dell'istruzione per migliorare lo svolgimento delle sue funzioni;
- Attuare un approccio gestionale basato sui processi per rispondere adeguatamente alle nuove esigenze e realtà dei destinatari dell'azione dell'Ispettorato dell'Istruzione (Boja, 2019).

Per ciascuno di questi obiettivi sono state elaborate linee strategiche di lavoro, processi, azioni prioritarie da implementare e indicatori dei risultati attesi che i valutatori esterni devono considerare nel loro lavoro di valutazione e supervisione. Per il quarto obiettivo, diffusione dei risultati della valutazione a livello centrale, le linee strategiche di lavoro individuate rimandano alla restituzione dei dati al Ministero dell'educazione e della Formazione Professionale mediante l'elaborazione di relazioni che possano facilitare il processo decisionale da parte degli organi competenti. Di conseguenza i processi implicati sono: il coordinamento e la collaborazione con le amministrazioni scolastiche e altre istituzioni, la realizzazione di relazioni annuali sul funzionamento dei servizi regionali di ispezione e di una relazione annuale conclusiva dell'ispettorato. Le azioni prioritarie ricadono nella diffusione degli esiti della valutazione, mentre gli indicatori ruotano attorno alla realizzazione dell'azione secondo i processi prestabiliti e all'analisi, alla valutazione e alla realizzazione di proposte valutative e di diffusione dei risultati (Boja, 2019).

Nella Comunità Autonoma di Madrid, invece, il Piano di Azione Generale Pluriennale per la valutazione esterna ha una durata quinquennale (2021-2025) e ha l'obiettivo di:

- Aumentare l'efficienza nell'organizzazione e nel funzionamento delle istituzioni educative, dei programmi e dei servizi;
- Migliorare i risultati strategici delle istituzioni educative;
- Rafforzare la funzione di gestione e la sua leadership;
- Valutare la pratica didattica per il suo miglioramento.

Oltre alle azioni straordinarie, l'ispettorato della Comunità Autonoma di Madrid nel quinquennio considerato deve supervisionare il benessere all'interno delle scuole, supervisionare il rispetto delle diversità e l'inclusione, supervisionare l'organizzazione didattica e i risultati raggiunti negli apprendimenti e valutare la leadership scolastica. Come per la Comunità Autonoma dell'Andalusia, anche in questo caso sono stati elaborati scopi specifici, indicati i contesti e i caratteri di applicazione e gli indicatori di risultato per ciascun obiettivo. Per la valutazione della leadership scolastica, ad esempio, gli obiettivi da raggiungere sono:

- Supervisionare i documenti di pianificazione della scuola;
- Vigilare sul funzionamento degli organi collegiali e di coordinamento didattico;
- Supervisionare i risultati scolastici e i piani di miglioramento elaborati;
- Supervisionare l'organizzazione degli insegnamenti della scuola;
- Supervisionare l'efficienza nell'organizzazione delle risorse umane della scuola;
- Supervisionare la documentazione amministrativa e accademica della scuola;
- Supervisionare le attività complementari ed extrascolastiche;
- Preparare una relazione con proposte di miglioramento per le scuole e con raccomandazioni per l'amministrazione educativa (BOCM, 2021).

Le istituzioni interessate da questo obiettivo dell'ispettorato sono quelle pubbliche, infatti, l'azione viene realizzata su base censuaria durante il periodo di validità del Piano in tutte quelle scuole in cui il dirigente scolastico termina un periodo del suo mandato. I valutatori esterni coinvolti sono due, anche se solamente uno di loro è la figura di riferimento per la scuola. I criteri di valutazione, nella legislazione della Comunità Autonoma di Madrid, sono resi pubblici al dirigente scolastico e vertono sulla gestione e sull'amministrazione, sull'organizzazione didattico-pedagogica, sulle relazioni istituzionali e sui risultati (BOCM, 2021).

Come riportato nel paragrafo 2.5, l'oggetto delle valutazioni esterne dei valutatori sono anche le rilevazioni standardizzate sugli apprendimenti realizzate in ogni comunità autonoma per raccogliere dati e informazioni sulle istituzioni scolastiche e sugli studenti e, dunque, proporre piani di miglioramento (Eurydice, 2015).

L'ispettorato, infatti, durante le visite di valutazione, supervisiona dal punto di vista pedagogico ed organizzativo il funzionamento delle istituzioni scolastiche e, nel rispetto dell'autonomia gestionale, i progetti e i programmi offerti da queste ultime, supporta e collabora al miglioramento della didattica e partecipa alla valutazione del sistema educativo e degli elementi che la compongono (Ley Organica 3/20, 2020).

L'ispettorato è stato istituito con la legge n. 30 del 2 agosto 1984, modificata dalla legge n. 23 del 28 luglio 1988, tuttavia è con la Ley Organica n. 2 del 3 maggio 2006 modificata con la Ley Organica n. 3 del 29 dicembre 2020 (Ley Organica 3/20, 2020) che sono stati ridefiniti i criteri di partecipazione al corpo ispettivo. La formazione dei valutatori esterni è molto simile a quella richiesta agli insegnanti in servizio. Tuttavia, per entrare a far parte del corpo ispettivo, è necessario partecipare a un concorso pubblico di selezione nel quale viene elaborata una graduatoria sulla base di criteri di merito. Ogni comunità autonoma può stabilire i propri criteri, anche se una formazione professionale specifica e un tirocinio formativo sono parte integrante e obbligatoria della selezione dei valutatori esterni. Per questi ultimi i requisiti di ammissione, stabiliti nazionalmente, sono: almeno sei anni di servizio nell'insegnamento curricolare, padronanza della seconda lingua ufficiale (a seconda della comunità autonoma di riferimento) e altri requisiti stabiliti regionalmente dalle comunità autonome (es. titoli di studio specifici). È sempre compito delle comunità autonome, inoltre, programmare ed erogare la formazione in servizio e l'aggiornamento professionale dei valutatori esterni (Eurydice, 2015).

Nelle visite di valutazione i valutatori esterni, indipendentemente dalla comunità autonoma di riferimento, seguono una serie di procedure comuni descritte in ogni piano di valutazione regionale. In primo luogo essi consultano la documentazione pedagogica e amministrativa dell'istituzione scolastica, successivamente effettuano le visite con durata variabile per raccogliere informazioni sul funzionamento della scuola anche, là dove necessario, mediante osservazioni in aula, a seconda di quanto stabilito negli obiettivi della valutazione esterna, infine effettuano interviste a diversi stakeholders della comunità educativa, il dirigente scolastico, gli insegnanti, il personale, gli studenti e le loro famiglie. Le interviste rappresentano un momento di discussione su aspetti riguardanti l'istituzione scolastica, spesso vengono considerati i risultati negli apprendimenti raggiunti dagli studenti nelle prove standardizzate o altri dati e informazioni riguardanti la scuola. Molta importanza in questi momenti di confronto è data a eventuali strategie e misure per il miglioramento, tanto è vero

che anche durante il follow-up vengono effettuate interviste, in particolar modo con gli organi di gestione dell'istituzione, per monitorare i piani di miglioramento implementati (Eurydice, 2015).

L'obiettivo ultimo della valutazione esterna, infatti, è quello di supportare le scuole nel migliorare in quelle aree che hanno ricevuto una valutazione negativa attraverso misure di sostegno, guida e formazione erogate dall'ispettorato. Per questo motivo nei follow-up vengono realizzate una serie di visite programmate a cadenza regolare per valutare i progressi compiuti in tali aree. Non solo le comunità autonome, ma anche il Ministero dell'Educazione e della Formazione Professionale centrale richiede alle scuole di mettere a punto una serie di misure per migliorare l'offerta formativa da includere nel Piano di Miglioramento che ogni scuola redige tenendo conto, tra le altre cose, dei risultati della valutazione standardizzata degli apprendimenti, che devono essere resi pubblici agli organi di coordinamento della scuola, e dei feedback ricevuti durante la valutazione esterna. A questo proposito, non tutte le scuole ricevono le visite ispettive, ogni comunità autonoma, infatti, ha la facoltà di selezionare le istituzioni scolastiche da visitare sulla base dei propri criteri e del Piano annuale redatto da ogni ispettorato (Eurydice, 2015).

Ogni anno l'ispettorato e i valutatori esterni redigono un report finale annuale da inviare agli organi amministrativi di ciascuna comunità autonoma e al Dipartimento dell'Istruzione centrale (Ley Organica 3/20, 2020).

L'autovalutazione delle istituzioni educative in Spagna è definita da ciascuna comunità autonoma e dal Ministero dell'educazione e della Formazione Professionale centrale, con il supporto nella realizzazione dell'ispettorato dell'istruzione. A partire dall'autonomia pedagogica ed organizzativa concessa, ogni scuola valuta annualmente i propri punti di forza e di criticità per il miglioramento dei processi scolastici e dei risultati negli apprendimenti tenendo conto delle relazioni sugli esiti delle valutazioni esterne, ove realizzate, e dando particolare attenzione alla valutazione degli apprendimenti realizzata dagli organi preposti. Trovare pratiche di autovalutazione condivise non è semplice per il contesto spagnolo poichè, alla luce dell'organizzazione geo-politica

federale e dell'autonomia scolastica, ogni comunità autonoma può stabilire un proprio piano di azione per l'autovalutazione delle istituzioni educative e per il miglioramento della qualità dell'istruzione, può richiedere alle scuole di formulare un proprio protocollo di valutazione esterna o può elaborare una serie di indicatori sui quali richiedere a queste ultime di concentrarsi durante i processi autovalutativi. In generale le istituzioni educative alla fine di ogni anno scolastico redigono un report sulle attività svolte, sia curricolari che extracurricolari, i risultati raggiunti e la disponibilità di risorse economiche ed umane e attuano i piani di miglioramento della qualità proposti dalle comunità autonome. Gli esiti di queste due azioni si traducono per le scuole nella messa a punto di un Programma Generale Annuale che comprende la calendarizzazione delle attività per l'anno successivo e i piani di azione e miglioramento da realizzare. Il dirigente scolastico promuove l'autovalutazione di istituto all'interno della propria scuola e collabora con i valutatori esterni, mentre i consulenti scolastici interni o esterni alla scuola supportano l'autovalutazione, lo sviluppo e il monitoraggio dei piani di miglioramento. In ogni scuola è presente un coordinatore dei processi di autovalutazione di istituto, ovvero un insegnante responsabile della valutazione interna e della pianificazione delle azioni per il miglioramento. All'autovalutazione delle scuole partecipano, per le proprie competenze, anche l'ispettorato, che fornisce supporto nelle pratiche di autovalutazione e nell'implementazione dei piani di miglioramento, i rappresentanti degli studenti delle scuole secondarie superiori, gli enti coinvolti nel funzionamento della scuola (es. progetti di innovazione) e i team di autovalutazione interni alla scuola formati da insegnanti e dal dirigente scolastico, dove presente (Eurydice, 2015).

Nel processo partecipativo di autovalutazione, gli attori coinvolti possono disporre di alcuni strumenti di supporto: i report sui risultati raggiunti nelle indagini standardizzate sono il punto di partenza per le raccomandazioni per il miglioramento, ma a questi ultimi si aggiungono i corsi di aggiornamento professionale e formazione dei docenti presso i centri preposti che possono essere un indicatore per l'autovalutazione ma anche un elemento da implementare nei piani di miglioramento.

Alcune comunità autonome, inoltre, forniscono sostegno finanziario alle scuole per i processi di valutazione interna e per la gestione della qualità (DOGV, 2015). Altre, invece, negli anni hanno creato reti di scuole per lo sviluppo di progetti e di piani di miglioramento e per lo scambio delle *best practices* o, nel caso della Comunità Autonoma di Aragon, hanno incluso suggerimenti per la valutazione e l'autovalutazione delle istituzioni educative nelle indicazioni regionali per l'educazione secondaria (Centro del Profesorado de Sevilla, 2021; Gobierno de Aragon, 2022). Altre, ancora, hanno sviluppato un sistema di indicatori utili al monitoraggio e alla valutazione dell'attuazione dei piani di miglioramento da parte delle scuole del territorio. Nella Comunità Autonoma di Navarra sono stati messi a punto 30 indicatori per valutare i piani di miglioramento e la loro attuazione. Gli indicatori sono raggruppati in quattro aree, ognuna delle quali fa capo a un aspetto dell'implementazione e della programmazione interna alla scuola: disegno del piano di miglioramento, misure del piano di miglioramento, sviluppo e risultati, monitoraggio (follow-up), valutazione e proposte di miglioramento. Per ogni area è previsto un sotto-ambito che guida e orienta nell'utilizzo degli indicatori. Per ognuno di essi è esplicitato il framework di riferimento e la sua applicazione tramite una scala graduata di raggiungimento del risultato o di presenza/assenza della variabile, per aiutare a comprendere le differenti situazioni. La gradazione dell'indicatore è stabilita sulla base di due criteri: il numero di persone e la situazione descritta nell'indicatore e lo sviluppo e la sua realizzazione pratica (Servicio de inspeccion educativa, 2012). Un lavoro molto simile è stato realizzato anche dalla Comunità Autonoma di Castiglia-La Mancia. Anche in questo caso è stata messa a punto una serie di indicatori con l'obiettivo di sostenere le scuole nei processi di autovalutazione e miglioramento. Le aree interessate con le rispettive sottodimensioni sono: processi di insegnamento e apprendimento, organizzazione e funzionamento, rapporti con il contesto di riferimento, processi di valutazione, formazione e innovazione (Comunidad Autonoma de Castilla-La Mancha, 2003).

Il responsabile interno alla scuola per l'autovalutazione e il miglioramento, supportato dal team di esperti e aiutato dagli strumenti

messi a disposizione nelle differenti comunità autonome, redige annualmente un rapporto sulla base del quale viene messo a punto il piano di miglioramento. I valutatori esterni, a loro volta, possono utilizzare il rapporto come base di partenza per le visite valutative. Tutti gli attori coinvolti vengono informati degli esiti della valutazione, ma questi ultimi non sono utilizzati per stilare graduatorie di merito tra istituzioni scolastiche, infatti alcuni dati sensibili non possono essere pubblicati per motivi di privacy (Cabanes & Gimeno, 2004; Eurydice, 2015; Sanchez et alii, 2015; Juan & Alfageme-Gonzalez, 2021).

2.6.3. La valutazione esterna e l'autovalutazione delle istituzioni scolastiche in Italia

La valutazione esterna delle istituzioni scolastiche in Italia si sviluppa all'interno del Sistema Nazionale di Valutazione, regolato dal D.P.R. n. 80 del 2013 (D.P.R. 80/13, 2013; Eurydice, 2015). Le prime esperienze legate a quest'ultima risalgono alla fine degli anni Novanta e all'inizio degli anni 2000, periodo in cui alcune istituzioni scolastiche, soprattutto della provincia di Trento, sono state coinvolte in progetti per la messa a punto di un modello di valutazione esterna delle scuole (Poliandri et alii, 2013). Tuttavia, è negli anni scolastici 2012/2013 e 2014/2015 che l'Invalsi, coordinatore del Sistema Nazionale di Valutazione, implementa alcune sperimentazioni pilota per la sistematizzazione della valutazione esterna. I due progetti sono Valutazione e Sviluppo della Scuola, VALeS, e Valutazione e Miglioramento (Muzzioli et alii, 2012). Il secondo, Valutazione e Miglioramento, condotto grazie ai fondi strutturali europei, aveva l'obiettivo di supportare le scuole nelle attività di gestione e di realizzare osservazioni in un'ottica di valutazione sistemica e di implementazione di interventi di miglioramento. Ha coinvolto più di 400 scuole, tra primarie e secondarie, e ha visto la loro assegnazione casuale a due possibili percorsi valutativi: valutazione dei risultati e dei processi legati al contesto organizzativo o osservazione in classe e valutazione delle pratiche didattiche ed educative. Il primo, invece, VALeS, aveva l'obiettivo di testare e migliorare non solo gli strumenti, ma anche i percorsi e i protocolli per la valutazione esterna delle scuole.

In quest'ultimo progetto sono stati definiti due profili di valutatori esterni: esperti valutatori con esperienza scolastica (Profilo A) ed esperti valutatori con esperienza in altri settori (Profilo B) (Muzzioli et alii, 2012; Eurydice, 2015).

Negli ultimi anni, a seguito dei progetti pilota dei decenni precedenti, è stato messo a punto un protocollo per la valutazione esterna delle scuole stabilito dalla Conferenza per il coordinamento funzionale del Sistema Nazionale di Valutazione (Invalsi, 2016b). Composta dai presidenti di Invalsi e Indire e dal rappresentante del contingente ispettivo, la Conferenza è un organo di coordinamento del Sistema Nazionale di Valutazione che adotta i protocolli e la calendarizzazione delle visite esterne proposte da Invalsi, oltre a comunicare con il Ministero dell'Istruzione e del Merito per ciò che riguarda le priorità strategiche della valutazione del sistema educativo, compreso quello professionale.

Nel complesso sistema di valutazione delle scuole la Conferenza rappresenta dunque l'organismo preposto al coordinamento dei tre soggetti che lo compongono - l'Invalsi, l'Indire e il Contingente ispettivo - e, in particolare, all'adozione degli strumenti di lavoro necessari affinché la valutazione esterna delle scuole sia effettivamente orientata al miglioramento delle scuole e sia caratterizzata da procedure corrette, imparziali e uniformi su tutto il territorio nazionale. (Invalsi, 2016a, p. 3)

La valutazione esterna si configura come un momento di confronto e di riflessione sia tra chi vive nell'istituzione scolastica quotidianamente che tra chi, invece, si pone con un occhio esterno. Essa, infatti, ha lo scopo di fornire alla scuola informazioni per il miglioramento, inteso, per esempio, come miglioramento dei livelli negli apprendimenti, delle innovazioni o dell'offerta formativa (Invalsi, 2016ab).

La valutazione esterna condotta dai NEV ha il compito quindi di approfondire la lettura dei diversi aspetti dell'organizzazione scolastica e delle priorità sulle quali orientare il miglioramento, in un'ottica multi-prospettica, favorita anche dal confronto tra la rappresentazione che la scuola si è data

attraverso l'autovalutazione e quella fornita dai valutatori esterni. (Invalsi, 2016a, p. 5)

Ogni anno le scuole sono chiamate a compilare il Rapporto di Autovalutazione, RAV, di cui si parlerà più avanti, per valutare il proprio operato e le attività realizzate, tuttavia, i Nuclei Esterni di Valutazione, NEV, nelle loro visite prendono in considerazione non solo questo strumento messo a disposizione delle istituzioni scolastiche, ma anche tutti gli altri fattori che possono contribuire a far comprendere la situazione reale del contesto visitato. Le scuole da visitare sono individuate tramite un processo di campionamento casuale in cui è presa in considerazione l'area geografica di provenienza e il grado. In ogni istituzione scolastica si recano tre valutatori esterni, due individuati e contrattualizzati da Invalsi in qualità di esperti e un dirigente tecnico del Ministero dell'Istruzione e del Merito (Invalsi, 2016ab). Egli è il garante del percorso di valutazione, cura i rapporti con le scuole e ne assicura il corretto svolgimento. Uno dei due esperti individuati da Invalsi, docente o dirigente scolastico, si occupa dei processi pedagogici e didattici e del punto di vista degli insegnanti, l'altro, invece, esperto nella valutazione delle organizzazioni, assicura la corretta raccolta e interpretazione delle evidenze empiriche da un punto di vista metodologico-interpretativo. Ad entrambi gli esperti, il Profilo A e il Profilo B, è richiesta un'esperienza professionale di settore di almeno tre anni (Eurydice, 2015).

Le visite di valutazione esterna si svolgono in tre fasi. Nella prima fase i valutatori fanno conoscenza dell'istituzione scolastica navigando nel sito web, leggendo i documenti programmatici, il Rapporto di Autovalutazione e il Piano dell'Offerta Formativa, e consultando i dati e gli indicatori messi a disposizione dal Ministero dell'Istruzione e del Merito che offrono una fotografia della scuola dall'esterno. Nella seconda fase raccolgono dati e informazioni recandosi sul campo attraverso osservazioni, interviste e ulteriore lettura dei documenti.

La visita di valutazione esterna ha una durata di tre giorni e si articola di norma in questo modo: incontro iniziale con il dirigente scolastico, lo staff di

dirigenza e il nucleo interno di valutazione (NIV); raccolta delle evidenze attraverso interviste individuali e di gruppo ed esame della documentazione della scuola; visita e osservazione degli spazi della scuola; incontro conclusivo con il dirigente scolastico, lo staff di dirigenza e il nucleo interno di valutazione e breve comunicazione informale sugli esiti della visita. (Invalsi, 2016a, p. 10)

L'incontro iniziale è un momento di introduzione della visita esterna allo staff della scuola e di sintesi delle informazioni raccolte dalla consultazione dei documenti, mentre il secondo e il terzo giorno sono dedicati ai colloqui individuali e di gruppo con le componenti della scuola quali dirigente scolastico, direttore dei servizi generali e amministrativi, docenti facenti funzione strumentale, docenti disciplinari, docenti di sostegno, personale amministrativo non docente, studenti e famiglie. Inoltre, oltre alle interviste e ai focus group vengono svolte anche osservazioni degli spazi della scuola e dell'utilizzo che se ne fa. Prima della fine della visita e dell'invio del Rapporto di Valutazione Esterna, il dirigente tecnico, membro del nucleo esterno di valutazione, fa una prima restituzione informale di quanto emerso durante le tre giornate, mettendo in risalto i punti di forza e i punti di debolezza riscontrati (Invalsi, 2016ab). Il Rapporto di Valutazione Esterna, REV, redatto dopo la conclusione della visita, racchiude un giudizio motivato per ciascuna area oggetto di valutazione esplicitato attraverso una rubrica di valutazione. Successivamente all'invio alla scuola di quest'ultimo, il protocollo messo a punto da Invalsi per le visite di valutazione esterna prevede una nuova restituzione, in presenza e più dettagliata, da parte di uno dei componenti del nucleo esterno di valutazione. L'obiettivo dell'incontro è quello di riflettere criticamente sulle potenzialità e le sfide emerse prima della messa a punto dei piani di miglioramento, stabilendo un collegamento tra la fase "diagnostica" della valutazione con quella "formativa" e "operativa" (Invalsi, 2016ab).

Concretamente, la restituzione avviene nel contesto di un dialogo nel quale i risultati della valutazione sono commentati individuando possibili connessioni tra lo sviluppo di specifici processi e le ricadute sul funzionamento scolastico, oppure presentando casi nei quali criticità simili sono state affrontate con

successo, illustrando buone pratiche utilizzate in altri contesti nazionali o internazionali, fornendo indicazioni per approfondire tematiche specifiche e così via. (Invalsi, 2016a, p. 12)

La riflessione corale perseguita durante l'incontro conclusivo è un modo per analizzare la discrepanza tra le valutazioni date nel Rapporto di Autovalutazione da chi effettivamente vive nel contesto scolastico quotidianamente e le valutazioni formulate dagli occhi esterni del nucleo esterno di valutazione. I valutatori esterni, infatti, dopo essersi confrontati sintetizzano i loro giudizi in uno dei sette livelli previsti nella scala di valutazione (1 – situazione molto critica, 7 – situazione eccellente). Ogni valore numerico, inoltre, è motivato da una serie di evidenze ed è accompagnato da un giudizio più descrittivo (Invalsi, 2016b).

Per ogni area di processo (pratiche educative e gestionali) e per i risultati (voti scolastici, prove standardizzate, risultati nelle scuole successive e competenze di cittadinanza) il nucleo fornisce un giudizio descrittivo, in cui siano brevemente presentati alla scuola i principali elementi emersi e i punti di forza e di debolezza. Nei casi in cui il giudizio sia differente da quello che la scuola si era assegnata nel RAV, il nucleo esplicita chiaramente le motivazioni alla base di questa scelta. Complessivamente il NEV esprime un giudizio sulla coerenza tra il RAV compilato e le priorità e i traguardi indicati, anche in base alle risorse e al contesto della scuola. (Invalsi, 2016a, p. 11)

L'autovalutazione delle istituzioni scolastiche in Italia avviene, concretamente, attraverso la compilazione del Rapporto di Autovalutazione, RAV, che rappresenta il *trait d'union* tra la valutazione esterna e quella interna alla scuola, da parte del dirigente scolastico e del nucleo interno di valutazione, un team formato da docenti interni alla scuola incaricati di redigere il RAV. Quest'ultimo, infatti, è tra i primi documenti elaborati dalla scuola che i valutatori consultano prima della visita di valutazione esterna per comprendere meglio il contesto di riferimento e orientare le fasi successive. L'autovalutazione, inoltre, rappresenta la prima fase del processo di valutazione delineato nel SNV ed è un momento di riflessione corale interna per la condivisione di

strategie per il raggiungimento degli obiettivi delineati (D.P.R 80/13, 2013).

Nel nuovo SNV il miglioramento della scuola rappresenta la finalità ultima delle azioni di valutazione. Tutte le scuole italiane, siano esse statali o paritarie, sono tenute a riflettere sul proprio operato e a stendere un Rapporto di Autovalutazione (RAV), documento nel quale indicano i propri punti di forza e di debolezza e individuano le linee di azione per il miglioramento. Questo rapporto è un documento centrale per il successivo percorso valutativo. (Invalsi, 2016a, p. 8)

Il framework di riferimento dell'autovalutazione e della valutazione esterna, infatti, è il medesimo in entrambe le fasi del SNV. Ciò rende possibile in linea teorica la collaborazione tra il valutato e il valutatore e la conoscenza dell'oggetto della valutazione da parte del valutato (Invalsi, 2016a).

Il RAV così come le griglie per la valutazione esterna si articola in tre aree distinte, Contesto e risorse, Esiti e Processi, a cui si vanno ad aggiungere le Priorità, che indicano ciò che deve essere monitorato nei processi di valutazione sia interni che esterni e quali sono i traguardi da raggiungere. I Processi al loro interno si suddividono in pratiche educative e didattiche e pratiche gestionali ed educative. Per ciascuna area sono previsti indicatori, descrittori e fonti consultabili articolati come riportati nella tabella 2 che segue nella pagina successiva.

Contesto e risorse	Esiti	Processi	Priorità
Popolazione scolastica; Territorio e capitale sociale; Risorse economiche e materiali; Risorse professionali.	Risultati scolastici; Risultati nelle prove standardizzate; Competenze chiave europee; Risultati a distanza.	<u>Pratiche educative e didattiche:</u> -Curricolo, progettazione, valutazione; -Ambiente di apprendimento; -Inclusione e differenziazione; -Continuità e orientamento. <u>Pratiche gestionali e organizzative:</u> -Orientamento strategico e organizzazione della scuola; -Sviluppo e valorizzazione delle risorse umane; -Integrazione con il territorio e rapporti con le famiglie.	Traguardi e obiettivi di processo da raggiungere.

Tabella 2. Articolazione delle aree del RAV. Fonte: Invalsi, 2022a.

I contesti e le risorse indicano il luogo fisico all'interno del quale si esplicano i processi di insegnamento-apprendimento. Su di esso possono essere raccolti dati strutturali sulla popolazione scolastica, sul territorio e sulle risorse a disposizione, condizioni che non sempre possono essere modificate dagli interventi educativi. Gli esiti indicano il successo

formativo dello studente. Vengono presi in considerazione dati sul rendimento degli studenti, sui tassi di dispersione scolastica, di abbandono, sulle performances nelle prove standardizzate somministrate a livello nazionale e internazionale con l'obiettivo di assicurare a tutti il raggiungimento del livello minimo di competenze, sulle competenze trasversali di cittadinanza per essere cittadini attivi nella comunità e sul successo nei successivi studi e nel mondo del lavoro. I processi indicano ciò che avviene in aula nell'insegnamento e nell'apprendimento e ciò che avviene a livello sistemico nella scuola e nel rapporto che questa ha con le famiglie e il territorio. Vengono presi in considerazione la capacità della scuola di progettare un'offerta formativa aderente al contesto in cui si situa, di mettere a disposizione un ambiente innovativo e inclusivo e di orientare gli studenti nel passaggio da un livello dell'istruzione a un altro. Inoltre, vengono prese in considerazione anche le capacità della scuola di individuare e raggiungere le priorità attraverso un utilizzo attento delle risorse a disposizione, di promozione della qualità e della valorizzazione del personale e di coinvolgimento del territorio e delle famiglie nella persecuzione degli scopi prefissati (Invalsi, 2016ab, 2022ab).

Nel framework di riferimento comune, inoltre, sono individuati quattro principi guida per la riflessione e la valutazione delle pratiche educative tanto nella valutazione interna quanto in quella esterna: equità, partecipazione, qualità, differenziazione. Per equità si intende la necessità di assicurare a tutti gli studenti il raggiungimento dei livelli minimi di competenza. Per partecipazione si intende la necessità di fare in modo che tutti, indipendentemente dal background di provenienza, possano partecipare e usufruire delle attività della scuola. Per qualità si intende la necessità di fare in modo che i processi si caratterizzino per la presenza di elementi e fattori che ne innalzano la possibilità di riuscita. Per differenziazione si intende la capacità della scuola di individualizzare i processi e le attività sulla base dei bisogni e delle caratteristiche di ciascuno (Invalsi, 2016, 2022ab).

Di recente nel sito del Ministero dell'Istruzione e del Merito (<https://miurgov.static.istruzione.it/>) è stata dedicata una apposita sezione alla nuova versione del RAV elaborata per il triennio 2022-2025

(Invalsi, 2022ab, MIUR, 2022). In quest'ultima il nucleo interno di valutazione è chiamato a individuare vincoli e opportunità per il Contesto e le risorse e punti di forza e di debolezza per gli Esiti e i Processi. La stessa scala di valutazione adoperata dai valutatori esterni è utilizzata anche dal nucleo interno di valutazione che esprime i propri giudizi mediante uno dei sette livelli a disposizione (1 – situazione molto critica, 7 – situazione eccellente). Le situazioni in cui la scuola può trovarsi sono sintetizzate utilizzando esempi concreti che possono far comprendere più da vicino il livello della scala che meglio descrive la condizione della scuola.

L'espressione del giudizio dovrebbe derivare dalla riflessione scaturita dalla lettura degli indicatori e delle domande guida e dall'individuazione dei punti di forza e di debolezza. Per ciascuna rubrica di valutazione è previsto un campo aperto, in cui il NIV può motivare la scelta del livello di giudizio che si è assegnato nella scala di valutazione. La motivazione è richiesta nel caso la scuola si posizioni nelle situazioni 2, 4 e 6. (Invalsi, 2022a, p. 4)

Particolare rilevanza nel lavoro del nucleo interno di valutazione è data all'individuazione delle priorità, dei traguardi e degli obiettivi di processo.

Le priorità si riferiscono agli obiettivi generali che la scuola si prefigge di realizzare nel lungo periodo (tre anni scolastici) attraverso l'azione di miglioramento e devono necessariamente riguardare gli esiti degli studenti. Al fine di intraprendere un percorso di miglioramento efficace è opportuno individuare un numero limitato di priorità (si suggerisce di individuarne al massimo tre). Per ogni priorità individuata deve essere stabilito il corrispondente traguardo. Il traguardo articola in forma osservabile e/o misurabile i contenuti della corrispondente priorità e consiste in un risultato atteso nel lungo periodo (tre anni scolastici). L'insieme dei traguardi rappresentano le mete verso cui la scuola tende nella sua azione di miglioramento. Gli obiettivi di processo, invece, sono obiettivi operativi da raggiungere nel breve periodo (un anno scolastico) che riguardano una o più aree di Processo. (Invalsi, 2022a, p. 4)

Ogni priorità è collegata a un obiettivo di processo, ma ogni obiettivo può essere collegato a più priorità in modo da sviluppare per il contesto specifico di riferimento un piano di miglioramento e implementazione che si sviluppi su un arco temporale sia a breve che a più lungo termine.

Nel capitolo quattro della seconda parte verranno riportati i risultati emersi dalle analisi condotte a seguito della somministrazione di un questionario-intervista a un gruppo di dirigenti scolastici italiani sull'autovalutazione di istituto.

Capitolo terzo

Le interviste

3.1. Le interviste qualitative

Corbetta (1999) definisce l'intervista come una conversazione che l'intervistatore avvia con soggetti scelti sulla base di un piano di rilevazione con uno schema flessibile e non standardizzato per perseguire obiettivi conoscitivi. L'intervista qualitativa si differenzia da quella quantitativa, come il questionario, poichè concede all'intervistato e all'intervistatore maggiore libertà nell'esprimere i propri punti di vista. Pur essendo qualitativa, l'intervista può essere organizzata in tre modalità diverse: può essere strutturata, semi-strutturata o non strutturata. La differenza tra le tre tipologie risiede nella maggiore o minore apertura verso l'intervistato. Nell'intervista non strutturata l'intervistato è completamente libero di seguire un proprio filo logico nel riportare i propri pensieri, mentre nell'intervista strutturata egli è chiamato a seguire una traccia ben definita. Il ricercatore, invece, in base ai dati che vuole raccogliere, può decidere se porre a tutti gli intervistati le medesime domande, nella stessa sequenza e nella stessa formulazione, o se stabilire le tematiche da trattare, lasciando all'intervistato la facoltà di esporre le proprie argomentazioni (Corbetta, 1999).

Nella ricerca sociale l'intervista è uno degli strumenti per la raccolta di informazioni più utilizzati (Fideli & Marradi, 1996). Ai fini dell'affidabilità e della validità dei dati raccolti bisogna prestare attenzione soprattutto alla costruzione e alla conduzione di quest'ultima. Per quanto riguarda la costruzione della traccia dell'intervista è importante tenere a mente se quest'ultima è più o meno strutturata e standardizzata e il livello di direttività che verrà assunto durante il colloquio. Per andare incontro agli obiettivi della ricerca presentata, si è scelto di ricorrere a un'intervista semi-strutturata in cui la traccia ha previsto la formulazione di domande suddivise in sezioni aderenti ai bisogni conoscitivi. Nello specifico dello strumento costruito ad hoc che

verrà illustrato nei paragrafi successivi, la traccia prevedeva la definizione di domande stimolo-guida, ma non ne precisava la strutturazione interna: a ogni testimone privilegiato sono state proposte le stesse domande adattate al contesto nazionale e valutativo di riferimento a partire da ciò che è emerso dall'analisi documentale illustrata nel capitolo precedente, ma è stata mantenuta una flessibilità in base alle risposte fornite e all'esplicitazione di determinati elementi (Bichi, 2002).

Solo per fare un esempio: ciascuno risponde a suo modo a ciascuna domanda prefissata - e sono queste caratteristiche che riguardano la direttività - alla domanda "che cosa significa per lei indossare l'abito?" è possibile che un intervistato estenda la sua risposta in altri ambiti di interrogazione presenti nella traccia, come "come si sente con l'abito addosso?" oppure/anche "come vede la scelta di quei religiosi che non indossano l'abito?", costringendo così l'intervistatore a modificare l'ordine e/o il contenuto delle sue domande - caratteristiche che concernono la prima la standardizzazione, la seconda la strutturazione. (Bichi, 2002, p. 32)

Per quanto riguarda la conduzione, invece, è particolarmente importante il rapporto tra l'intervistatore e l'intervistato all'interno del quale avviene un'interazione sociale. A questo proposito, tra l'intervistato e l'intervistatore si instaura non solo un processo di negoziazione di significati, ma anche la creazione di una serie di aspettative reciproche in cui, per favorire la cooperazione tra le parti, il ricercatore deve fornire tutte le informazioni del caso all'intervistato attivando così la motivazione a perseguire l'obiettivo condiviso (Losito, 2015). Nel caso della conduzione di un'intervista con testimoni privilegiati e rappresentati istituzionali, come avvenuto nella ricerca dottorale presentata, l'attenzione richiesta nello svolgimento di quest'ultima è stata più che doverosa.

È difficile prevedere gli effetti di un qualsiasi tipo di interazione a meno che non si abbiano molte informazioni circa i valori e le aspettative dei partecipanti. Quando si dispone di tali conoscenze, è possibile usarle per interpretare e valutare le informazioni acquisite attraverso l'intervista. È possibile inoltre

utilizzare come un dato la stessa situazione d'intervista in quanto la comprensione della dinamica che ne sta alla base può allargare la comprensione dell'interazione sociale nel suo complesso. (Phillips et alii, 1972, p. 189)

L'intervistatore, dunque, gioca un ruolo fondamentale nell'intervista. Non è solamente colui che sollecita la discussione, ma è parte integrante, attore protagonista di essa. La buona riuscita di un'intervista, infatti, non dipende solamente dal comportamento individuale delle due parti, ma anche dall'interazione che si crea (Roulston et alii, 2003). In questa interazione e in particolar modo nel lavoro presentato nei paragrafi successivi, alla luce del ruolo istituzionale del gruppo di intervistati di riferimento, si è optato per dare all'intervista un'apertura, una flessibilità e una dialogicità tali da avere una traccia comune per tutte le interviste, ma anche da lasciare spazio a possibili approfondimenti suscitati da tematiche trattate dall'intervistato durante il colloquio. Il principio seguito nella conduzione dell'intervista è quello descritto da Losito (2015): la traccia dell'intervista si è configurata come lo strumento e il mezzo per dare omogeneità alle rilevazioni, ma le domande-stimolo, soprattutto nella seconda parte del colloquio, non sono state standardizzate e rese a priori uguali per tutti, non ritenendo che a una stessa formulazione dovesse corrispondere una uguale reazione da parte degli intervistati, visti anche i modelli di valutazione e i contesti nazionali differenti, riportati nel secondo capitolo della seconda parte, in cui si situano i processi valutativi.

Il compito dell'intervistatore non è quello di semplice lettore di domande e di passivo classificatore e codificatore di risposte. L'intervistatore, infatti:

a) deve formulare la domanda nel modo che ritiene più adeguato caso per caso, per far sì che effettivamente tutti la comprendano in funzione di ciò che essa è destinata a rilevare;

b) deve ascoltare attentamente la risposta e, se necessario, intervenire per dare e/o ricevere chiarimenti e per sollecitare approfondimenti;

c) deve interpretare la risposta, sintetizzarla e classificarla riconducendola a una delle categorie che costituiscono le alternative che "chiudono" la domanda. Questa sorta di istantanea analisi del contenuto di un testo orale (cioè la risposta, per come essa risulta non solo da un'affermazione più o meno

pertinente dell'intervistato, ma anche dal lavoro semiotico che questa stessa affermazione può attivare tra i due interlocutori) richiede competenza, esperienza, disposizione al dialogo, capacità ermeneutica esattamente come accade o dovrebbe accadere per le interviste senza questionario. (Losito, 2015, p. 2)

Nel caso specifico, le interviste come situazioni comunicative in cui si instaura una relazione tra l'intervistato e l'intervistatore sono state condotte telematicamente a causa della situazione pandemica che ha coinvolto anche il mondo della ricerca a partire dalla primavera del 2020. L'interazione sociale nella quale chi scrive si è trovata per condividere e scambiare conoscenze e significati rispetto alle tematiche oggetto di approfondimento è avvenuta in un ambiente virtuale realizzato all'interno della piattaforma per videotelefonia ZOOM (Di Fraia, 2004). A causa delle criticità sanitario-emergenziali dovute al diffondersi del virus SARS-CoV-2 si è dovuto ricorrere obbligatoriamente alla comunicazione mediata da computer in rispetto delle restrizioni per gli spostamenti fisici delle persone. La rete web ha permesso forme di interazione mediate che, anche se maggiormente depersonalizzate rispetto a quelle in presenza inizialmente preventivate prima dell'inizio della pandemia, hanno messo in comunicazione soggetti residenti in quartieri diversi della stessa città o, addirittura, in paesi diversi. L'ambiente virtuale, quindi, si è venuto ad affiancare a quello reale dando vita a una socialità di secondo livello in cui la virtualità ha fatto sì che la rete esistesse in quanto "luogo" (Paccagnella, 2000) e come "terzo spazio" (Oldenburg & Brissett, 1982; Metitieri & Manera, 2000; Ducheneaut et alii, 2007; Memarovic et alii, 2014; Timm-Bottoms & Reilly, 2015).

Nonostante l'obbligata revisione del piano di raccolta dei dati inizialmente progettato, la realizzazione dell'intervista in modalità sincrona mediante il supporto del personal computer e previo appuntamento fissato nei giorni precedenti al colloquio telematico, ha permesso l'incontro e il dialogo, pur essendo distanti nello spazio. L'assenza di comunicazione non verbale e l'"anonimato visivo", dovuto alla presenza di uno schermo "protettore" (Lea & Spears, 1992), hanno

portato chi scrive a doversi dotare di un sistema di “lettura a distanza” delle reazioni paraverbali e non verbali degli intervistati. Ciò ha comportato il prendere appunti e note molto particolareggiate durante l’intervista, audioregistrata previo consenso dei partecipanti, sulle quali poter tornare nella fase di analisi dei dati che verrà illustrata nei paragrafi successivi. A dispetto di ciò, l’interazione personale mediata, il senso di maggiore libertà di espressione dovuto all’assenza della sensazione di “essere messo a nudo” in una relazione *face-to-face* (Boshier, 1990) e l’abitudine ad essere sottoposti a interviste da parte dei testimoni privilegiati, hanno permesso di indagare in profondità le tematiche oggetto di studio con vantaggi di tempo e di costo e con un sostanziale alto coinvolgimento dei partecipanti (Di Fraia, 2004).

3.2. La trascrizione e l’analisi delle interviste

Dopo aver costruito la traccia comune delle interviste, aver predisposto la possibilità di porre alcune domande suppletive sulla base dei temi emersi durante il colloquio (Bichi, 2002) e dopo aver condotto le interviste telematicamente mediante il supporto della piattaforma ZOOM (Di Fraia, 2004; Losito, 2015), si è passati alla trascrizione di queste ultime giungendo, in una fase successiva, all’analisi dei dati qualitativi.

Per la trascrizione ci si è avvalsi del supporto del software Trint, un programma di trascrizione automatica online di audio e video: dopo aver importato il file audio, il software ha convertito la registrazione in testo. Tuttavia, consapevoli dei livelli di accuratezza che si possono raggiungere nelle trascrizioni assistite, si è passati a una revisione del documento prodotto riascoltando le registrazioni audio. Le parole degli intervistati sono state trascritte integralmente *verbatim* facendo riferimento ad alcune dimensioni paralinguistiche e non verbali e soddisfacendo gli obiettivi conoscitivi (Pallotti, 1999; Oliver et alii, 2005; Fatigante, 2006; Barbour, 20014; Pagani, 2020). Il testo convertito ha funzionato da canovaccio di partenza nel quale sono state corrette manualmente le imprecisioni e gli errori di trascrizione. Se la trascrizione completamente manuale è un lavoro molto lungo e

minuzioso, il metodo “misto” manuale e automatico offre molti vantaggi. Da una parte velocizza il processo e permette di passare poi all’analisi dei dati sbobinati, dall’altra dà al ricercatore la possibilità di “immergersi nei dati” (Patton, 2002), acquisendo familiarità con essi e generando nuove idee e intuizioni. La trascrizione, come descritto da Creswell (2012) nel modello a spirale, diventa così una forma preliminare di analisi dei dati della quale bisogna dare conto per l’influenza che ha sul processo di analisi successivo (Lapadat, 2000; Tilley & Powick, 2002; McLellan et alii, 2003; Gibbs, 2018; Pagani, 2020).

Tradurre in forma testuale una performance orale presenta limiti analoghi a quelli che si incontrano realizzando un ritratto. Per quanto abile possa essere il disegnatore, non potrà mai rendere in tutta la sua complessità e fluida dinamicità il soggetto che intende raffigurare: qualche dettaglio andrà inevitabilmente perso, la tridimensionalità appiattita, il movimento imbrigliato nella staticità del foglio. Riconosciuti i limiti intrinseci del processo di trascrizione, per evitare di restare involuppati in un processo di progressiva “limatura” del trascritto che potrebbe potenzialmente procedere all’infinito, occorre capire quando è il momento di fermarsi. Come sottolinea Silverman (2006), “le trascrizioni possono sempre essere migliorate e la ricerca della perfezione è illusoria e richiede molto tempo. Piuttosto, lo scopo è arrivare a una trascrizione concordata, adeguata al compito che dobbiamo svolgere” (ivi, p. 288). Il nostro obiettivo, dunque, è raggiungere un livello di accuratezza sufficiente per procedere con l’approccio analitico prescelto senza tradire i dati e i significati che essi portano con sé. (Pagani, 2020, p. 52)

Per l’analisi dei dati si è scelto di far ricorso alla tecnica dell’analisi tematica. Si tratta, infatti, di un metodo interpretativo per i dati qualitativi che permette di individuare pattern di significato in quanto caratterizzato dalla flessibilità (Braun & Clarke, 2006). Terry e colleghi (2017) riconoscono differenti approcci di analisi tematica che Braun e Clarke (20019) sistematizzano in una “classificazione tripartita” in base alle concettualizzazioni di ciò che è il “tema”, ai sistemi di codifica e all’identificazione e sviluppo dei temi (Pagani, 2020). L’analisi tematica cui si fa riferimento nel lavoro dottorale è quella che le stesse Braun e Clarke (2019) hanno definito “riflessiva”.

Nell'AT riflessiva, la soggettività non viene rimossa dal processo analitico, ma riconosciuta e accolta. [...] Come conseguenza di questa diversa visione, la codifica, da rigida applicazione di un *coding framework* - predefinito, rigido e inflessibile - all'intero dataset, viene intesa come un processo fluido, flessibile, organico, che evolve man mano che l'analisi procede. La presenza di più codificatori indipendenti, per quanto utile a confrontare e discutere le osservazioni analitiche, non è più una necessità. La credibilità dell'analisi deriva dal modo rigoroso e sistematico con cui viene affrontato il processo analitico, non da misure statistiche [...]: del resto, anche un alto livello di accordo tra codificatori si limita a dirci che essi sono stati addestrati a codificare allo stesso modo, non offre garanzie sulla qualità dell'analisi (Yardley, 2008). In questo approccio, l'obiettivo non è quello di pervenire a una codifica il più possibile "attendibile" [...], rivelando i codici "sepolti" nei dati; ma piuttosto di far leva sulle competenze e la riflessività del ricercatore per giungere a una codifica plausibile e coerente (metaforicamente, il passaggio è da un ricercatore-archeologo a un ricercatore-scultore). [...] L'assenza di un'affiliazione teorica specifica rende l'AT estremamente duttile e adatta ad affrontare un ampio range di domande di ricerca, permettendo, ad esempio, di indagare le esperienze e i vissuti dei partecipanti; le rappresentazioni e costruzioni di determinati fenomeni; le pratiche messe in atto dagli individui e il modo in cui i partecipanti danno un senso alle proprie azioni; i fattori individuali e sociali alla base di un fenomeno. (Pagani, 2020, p. 58-63)

Data la sua flessibilità e la possibilità di essere applicata sia a campioni ampi che a campioni ristretti (Clarke & Braun, 2013), come nel caso della ricerca dottorale, l'analisi tematica riflessiva è stata dunque applicata ai dati raccolti dell'intero dataset delle interviste telematiche condotte con i National Project Manager dell'indagine PISA dell'OCSE, partecipanti descritti nel primo capitolo della seconda parte. Essendo in presenza di riflessioni ricche, si è optato per una comprensione approfondita dei significati veicolati individuando tre temi principali in essi presenti e aderenti alle domande di ricerca, in particolare alla seconda (le istituzioni pubbliche come rispondono alla valutazione di sistema?) (Terry et al., 2017): valutazione standardizzata degli apprendimenti; valutazione della scuola e per la scuola; valutazione e policy processes. Alla luce della ridotta ampiezza del corpus testuale

dovuta al modesto numero di intervistati e alla luce della loro elevata qualificazione sull'oggetto di indagine del progetto di ricerca dottorale, l'analisi tematica riflessiva è stata condotta manualmente utilizzando un approccio carta/matita con il supporto del pacchetto Office 365. La posizione di rilievo ricoperta dai National Project Manager dell'indagine PISA dell'OCSE ha richiesto, infatti, una lettura testuale *verbatim* di quanto dichiarato nelle interviste qualitative. Ciò ha dissuaso chi scrive dall'utilizzo di un software di analisi computer assistita dei dati qualitativi per mantenere un approccio di approfondimento più critico e riflessivo (Pagani, 2020).

Nel concreto dell'analisi dei dati, come indicato da Braun e Clarke (2006) nell'articolazione dell'analisi tematica riflessiva in fasi iterative e ricorsive, si è proceduto dapprima con la familiarizzazione dei dati: sono stati letti i testi e sono state annotate le riflessioni suscitate dalla lettura. Successivamente sono stati generati dei codici iniziali, ovvero nei trascritti sono state individuate quelle parti che potevano rispondere alle domande di ricerca e ai bisogni conoscitivi esposti nel primo capitolo della seconda parte di questo lavoro, conferendogli, attraverso il processo di etichettatura, un'organizzazione interna. Dalla definizione dei codici si è passati alla definizione dei temi.

Se i codici costituiscono i mattoni dell'analisi, i temi ne rappresentano i muri portanti (Braun & Clarke, 2013). Un buon codice cattura un'idea. Un tema, invece, ha una portata più ampia: riunisce in sé diversi codici, racchiudendo un concetto organizzatore centrale declinato nelle sue molteplici sfaccettature (ciascuna delle quali potrebbe essere associata a un codice). La domanda di ricerca può offrire al ricercatore una guida in questa fase, aiutandolo a non perdersi nella ricchezza dei dati qualitativi. Volgendo lo sguardo a essa, il ricercatore può più facilmente determinare quali tra i possibili pattern di significato identificati sono salienti e assicurarsi così che i temi raccontino una storia sui dati coerente e rilevante. [...] Ciò che è davvero determinante è quanto il possibile tema risulti significativo rispetto alla domanda di ricerca (Braun, Clarke, 2013). (Pagani, 2020, p. 72-73)

Infine, sono stati rivisti i temi individuati inizialmente, dandogli una definizione definitiva. Nei paragrafi 3.4, 3.5 e 3.6 verranno presentati i

risultati dell'analisi tematica riflessiva condotta. Le riflessioni dei National Project Manager dell'indagine PISA dell'OCSE, infatti, sono state organizzate intorno alle tre tematiche emerse (valutazione standardizzata degli apprendimenti; valutazione della scuola e per la scuola; valutazione e policy processes) (Braun & Clarke, 2013; Terry et al., 2017; Pagani, 2020).

3.3. Nota operativa: la traccia delle interviste e il consenso informato

Prima di illustrare le tematiche emerse dall'analisi delle interviste semi-strutturate e telematiche, per completezza si riporta la traccia del consenso informato e si descrive lo strumento costruito ad hoc per i colloqui con i National Project Manager dell'indagine PISA dell'OCSE, rispettivamente per l'Italia Laura Palmerio, per la Spagna Lis Cercadillo e per la Finlandia Arto Ahonen.

Per questioni etiche e istituzionali, dal momento che gli intervistati, testimoni privilegiati, hanno partecipato alla ricerca in qualità di rappresentanti istituzionali e responsabili nazionali dell'indagine PISA dell'OCSE, è stato preparato, prima dello svolgimento dell'intervista semi-strutturata telematica, un consenso informato da far firmare. Nello svolgere la ricerca, infatti, bisogna sempre informare i partecipanti e acquisire da parte loro verbalmente o per iscritto il consenso alla partecipazione e al trattamento dei dati. Chi scrive si è rivolta, in tale direzione, al Comitato Etico di Dipartimento compilando un modulo autocertificativo e dando informazioni sul protocollo e sullo svolgimento della ricerca. La documentazione prodotta è stata posta al vaglio del Comitato Etico e una volta approvata e protocollata è stata inviata ai partecipanti alle interviste. Trattandosi di un progetto di ricerca dottorale che ha visto il coinvolgimento di stakeholders provenienti da varie realtà estere, si è dovuto rendere accessibile l'informativa a tutti i partecipanti. La parte sul Regolamento UE circa il trattamento dei dati è rimasta invariata, in quanto è uniforme in tutti i paesi parte dell'Unione Europea, mentre la parte descrittiva degli obiettivi, degli strumenti e dei partecipanti alla ricerca è stata tradotta in

lingua inglese prima di essere inviata ai National Project Manager di Spagna e Finlandia. Nonostante alcuni aspetti della ricerca siano cambiati nel tempo a causa del perdurare dell'emergenza pandemica, si riporta di seguito il consenso informato per la partecipazione all'intervista, per l'audioregistrazione di quest'ultima e per la diffusione dei dati, tradotto e adattato da chi scrive alla lingua inglese.

Dear participant,

the Department of Social and Developmental Psychology at Sapienza, University of Rome is conducting research in the context of the Doctorate in Social and Developmental Psychology and Educational Research (XXXV cycle, Ph.D student Eleonora Mattarelli, tutor Prof. xxx and Prof. xxx). This study aims to investigate practices and perceptions related to educational system assessment and to impacts that the pandemic due to Covid-19 has had and is having in the educational field in five countries: Italy, Finland, France, Germany and Spain.

The reason we are conducting this research is to find out and analyze the status of school external assessment and self-evaluation. If the purpose is to ensure the quality of educational processes, is important to know the ways in which learning outcomes are monitored and focus the attention on didactic solutions adopted, available resources, investments made and resilience shown by the school system.

The research design is comparative. We are inviting to participate institutional entities engaged in the education sector in Italy, Finland, France, Germany and Spain with the aim of comparing their points of view. For greater completeness of the descriptions, alongside the interviews, we will examine websites and documentation produced by various institutions.

Before you decide freely if you want to participate, please read this informed consent carefully. Your participation in this research is entirely voluntary. Participation does not involve risks. If you have any question you may ask them to the interviewer, Dr. Eleonora Mattarelli, or later you may contact any of the following: Eleonora Mattarelli, xxx@xxx.it; Prof. xxx, xxx@xxx.it; Prof. xxx, xxx@xxx.it.

We are asking your availability to carry out a one-to-one interview which, due to the contingent situation, will be administered through the Zoom software, lasting about 60 minutes. The interview will focus on the relevance of the educational system assessment, on the effects of the pandemic and on the dimensions that are central to the construction of a quality educational system in which evidence-based decisions are made.

Our discussion will be audiotaped to help the interviewer accurately transcribe and capture your insights in your own words. The tapes will only be heard by the interviewer for the purpose of this study. Data analysis will not have any diagnostic or evaluative purpose and results dissemination will take place in academic publications, conferences and feedback events.

Sensitive data collected for this research will be audiotaped with a voice recorder, organized, consulted, selected, filed anonymously using an alphanumeric code and processed for the achievement of the objectives of the research project in accordance with the European General Data Protection Regulation (GDPR 2016/679). Data collected will be stored on a personal computer protected by a password for the time necessary to complete the research and in any case for a period of 5 years from the end of the study.

Ethical procedures for academic research require that interviewees explicitly agree to be interviewed. This consent form is necessary for us to ensure that you understand the purpose of your involvement and that you agree to the conditions of your participation. Sign this form to certify that you approve the following:

1. The interview will be recorded and a transcript will be produced. The transcribed interview or extracts from it may be used as described above;
2. I have been able to ask any questions I might have and I understand that I am free to contact the researcher with any questions I may have in the future;
3. All or part of the content of this interview may be used in academic papers, spoken presentations and in other feedback events.

By signing this form I declare that I have read the Information sheet, I agree and I am voluntarily taking part in this project.

Place and date
SIGNATURE

Per raggiungere gli obiettivi descritti nel primo capitolo della seconda parte del presente lavoro e per rispondere agli interrogativi di ricerca si è deciso di utilizzare l'intervista qualitativa semi-strutturata come strumento per comprendere più da vicino le tematiche oggetto di indagine.

La finalità principale di questo strumento è allora quella di permettere una conoscenza approfondita e qualitativamente ricca dei fenomeni, attivando un processo dialogico e un'interazione conoscitiva. [...] Esiste una serie di possibilità per costruire e gestire questa tecnica nella modalità più funzionale alla raccolta di dati in relazione al contesto, al gruppo di riferimento, alle risorse e intenzioni di analisi. (Benvenuto, 2015, p. 220)

Tra le fasi preliminari della stesura della traccia dell'intervista rientra l'analisi della letteratura sulle tematiche oggetto di indagine. Leggendo di studi che si sono occupati precedentemente degli aspetti della valutazione dei sistemi educativi sono emerse, infatti, alcune questioni, trasformate poi in domande per i colloqui, che si è ritenuto interessante indagare per corroborare i dati raccolti tramite l'analisi documentale (Bottani, 2009; Cardone et alii, 2010; Muzzioli et alii, 2012; OECD, 2013; Agnelli, 2014; Eurydice, 2015; Landri & Maccarini, 2016; OECD, 2019abc). Dopo aver precisato le tematiche e aver definito il livello di strutturazione e di direttività, si è passati alla costruzione della traccia dell'intervista. Le tracce per ciascuna intervista con i National Project Manager dell'indagine PISA dell'OCSE di Italia, Finlandia e Spagna hanno previsto da una parte una traccia comune per la comparazione delle risposte nell'ottica di uno studio di caso multiplo e comparativo con approccio esplorativo e descrittivo (Yin, 2009; Bray et alii, 2009), dall'altra è stata lasciata la possibilità agli intervistati di esprimere liberamente le proprie argomentazioni prevedendo domande suppletive

e affondi specifici emersi durante l'interazione (Bichi, 2002; Benvenuto, 2015). La formulazione delle domande ha comportato la definizione del loro ordine e in un primo momento la strutturazione in sei sezioni differenti (Eurydice, 2015; Save the Children, 2020):

- Percezione della valutazione;
- Esiti delle indagini nazionali e internazionali;
- Diffusione degli esiti della valutazione;
- Status della valutazione esterna;
- Status dell'autovalutazione;
- Conseguenze del momento pandemico.

Essendo il gruppo di riferimento particolarmente ristretto e privilegiato (per ciascun contesto nazionale è nominato per ogni ciclo dell'indagine PISA dell'OCSE un solo responsabile nazionale, come illustrato nel primo capitolo della seconda parte), non si è potuti procedere ad un vero e proprio try-out su larga scala. Tuttavia, nei contatti e nei colloqui preliminari, precedenti l'intervista effettiva, con la National Project Manager italiana si è provveduto ad assicurarsi che le domande, così come formulate, fossero valide, chiare e comprensibili, chiedendo di esprimere osservazioni sulla struttura dell'intervista e sulla possibilità di modificarne alcune. Il lavoro di *labor limae* ha permesso di avere la versione definitiva della traccia dell'intervista e di approssimarne la durata a circa un'ora.

La prima parte dell'intervista è stata dedicata ai dati emersi dall'ultima edizione dell'indagine PISA, quella del 2018 (OECD, 2019abc). Dovendo intervistare i responsabili nazionali dell'indagine realizzata da due decenni dall'OCSE sulle competenze dei quindicenni e volendo metterli a proprio agio, si è pensato di avviare il colloquio dai risultati rilevati sia dalla prova cognitiva dello studio, sia dal questionario di sfondo che viene somministrato agli stakeholders coinvolti. La valutazione degli apprendimenti, come è stato illustrato nel capitolo tre della prima parte del presente lavoro, rappresenta una delle possibili forme che può assumere la valutazione dei sistemi scolastici (OECD, 2013). Durante il colloquio sono stati fatti rimandi non solo allo studio sui quindicenni, ma anche alle indagini standardizzate realizzate all'interno dei confini nazionali. Tuttavia, si è preferito focalizzarsi

maggiormente sugli studi internazionali, non solo perchè rappresentano il lavoro quotidiano dei testimoni privilegiati ascoltati, ma anche perchè restituiscono la possibilità di fare una comparazione più completa, dal momento che tutte e tre le realtà nazionali di norma prendono parte alle indagini proposte sia dall'OCSE che dall'IEA. Va ricordato, infatti, che in Finlandia, come emerso dall'analisi documentale, approfondita nel secondo capitolo della seconda parte del presente lavoro, non c'è una vera e propria tradizione legata ai test sugli apprendimenti degli studenti somministrati annualmente e in modo censuario. L'istituto di valutazione del sistema scolastico finlandese, KARVI, realizza indagini tematiche e sulle competenze degli studenti, ma adotta una prospettiva e un framework teorico più simile a quello dell'indagine PISA dell'OCSE in cui la valutazione è intesa non tanto come un momento diagnostico, quanto come un momento formativo per l'implementazione delle azioni future (OECD, 2019abc; KARVI-FINEEC, 2022).

Nella tabella 3 sono riportate a titolo esemplificativo le domande poste alla National Project Manager dell'indagine PISA dell'OCSE della Spagna circa alcune evidenze emerse dallo studio sulle competenze dei quindicenni.

Analyzing PISA 2018 results, it emerges that Spain scored below the OECD average in both mathematics and science. These data are different from those reported by the IEA-TIMSS and IEA-PIRLS studies in which younger students achieve better school performance. In your opinion, what is the reason for this difference in results?
Are the differences between Autonomous Communities, between male and female students and related to socio-economic status (eg. natives vs immigrant students) highlighted by PISA results also confirmed by national assessments? How do you think an education system can mitigate these differences?
We know from several studies that the family and cultural background of the young person and the preparation of the teacher affect the student's results. What contribution does your work make in trying to reduce the strong impact of family and cultural background on student achievement?

Tabella 3. Traccia domande sull'indagine PISA dell'OCSE.

La seconda parte dell'intervista è stata dedicata alla percezione che si ha della valutazione del sistema educativo nei contesti nazionali di riferimento, vista dal punto di vista di chi è coinvolto nella valutazione esterna e standardizzata degli apprendimenti. L'utilizzo di indagini standardizzate su larga scala per monitorare gli esiti negli apprendimenti e quindi valutare i sistemi scolastici è divenuta, a partire dagli anni 2000, una pratica sempre più adottata, nonostante le controversie e le critiche che sono state mosse al *Large Scale Assessment* e all'utilizzo dei numeri e dei dati da esso provenienti per le scelte politiche riguardanti l'educazione (Pons, 2001; Adams, 2003; Chung, 2010; Baird et alii, 2011; Viteritti & Giancola, 2015; Volante, 2015; Gorur, 2016; Baird et alii, 2016; Araujo et alii, 2017; Hopfenbeck et alii, 2018; Rautalin, 2018; Fischman et alii, 2019). La valutazione dei sistemi educativi dovrebbe basarsi su una "politica informata" dalle *best practices*, da adeguate evidenze empiriche e dalla costruzione del consenso tra i vari stakeholders. La percezione della valutazione tra gli attori coinvolti e la costruzione di consenso tra i dirigenti scolastici, gli insegnanti, gli studenti, le famiglie e tutto il personale educativo assicura una buona riuscita della valutazione come momento di confronto e riflessione per il miglioramento (OECD, 2013).

Nella tabella 4 sono riportate a titolo esemplificativo le domande poste alla National Project Manager dell'indagine PISA dell'OCSE della Spagna circa le percezioni legate alla valutazione del sistema scolastico e le risposte date a livello centrale.

What do you think and how do you welcome the ideological criticisms of the OECD-PISA program? In Spain are there also criticisms against the national assessment (eg. Evaluacion de tercer/sexta curso de Educacion Primaria)?
In your opinion, how is external evaluation seen by those who are involved in it?
Do you think that these assessment and evaluation systems can have negative "side effects" on teachers' behaviour? Do they have a clear purpose for the evaluation system?

Feeling an active part of a process is undoubtedly a factor that promotes collaboration and interest of those who live it. In what way and through which mechanisms, in your opinion, does this model of system evaluation involve the assessed in the evaluation processes?
How is the information coming from the external evaluation received by the central institution - Ministry of Education? In your opinion, is the system assessment and evaluation linked to improvement interventions?
In Spain, how is the presence of INEE perceived from your point of view by school principals, teachers, students and families?
Do you think that schools use the data from the system evaluation to make decisions?

Tabella 4. Traccia domande sulla percezione della valutazione.

L'ultima parte dell'intervista ha approfondito il rapporto tra la spesa pubblica per l'educazione e la valutazione, l'autovalutazione e la valutazione esterna delle scuole e la situazione pandemica che ha messo duramente alla prova i sistemi educativi di tutto il mondo. La questione della spesa pubblica così come quella del momento emergenziale è particolarmente rilevante. Le decisioni che i governi nazionali prendono in merito agli investimenti interni, nello specifico in merito alla spesa pubblica per l'istruzione, dipendono da una serie di priorità che ogni contesto si dà e dal supporto che il settore privato offre. L'investimento pubblico per l'istruzione è tra le voci di spesa di ogni nazione, infatti l'obiettivo comune è assicurare a tutti le stesse opportunità di accesso. Vi sono, tuttavia, alcune differenze tra i paesi oggetto di studio negli investimenti del prodotto interno lordo per l'educazione. Ciò si ripercuote sul budget di fondi pubblici messo a disposizione per gli istituti di valutazione nazionale e per la ricerca educativa e valutativa in senso lato (OECD, 2022). La pandemia da Covid-19, inoltre, non solo ha colpito economicamente le industrie, ma ha anche portato a un decremento di investimenti pubblici per l'educazione che ha influenzato la qualità dell'istruzione fornita dalle istituzioni educative già provate dall'implementazione della didattica a distanza per sopperire alle

restrizioni imposte dalla diffusione del virus (Agrusti et alii, 2008; Agrusti, 2010; Save the Children, 2020; Reimers & Schleicher, 2020; OECD, 2020; 2021; 2022; Mattarelli, 2021ab; Di Donato, Mattarelli, Reggimenti, 2021; Agrusti & Corsini, 2021; Di Genova & Mattarelli, 2021).

Nella tabella 5 sono riportate a titolo esemplificativo le domande poste alla National Project Manager dell'indagine PISA dell'OCSE della Spagna circa gli investimenti pubblici per l'istruzione, nello specifico per la valutazione, l'autovalutazione delle istituzioni educative e le conseguenze della pandemia da Covid-19 e il possibile supporto offerto dalla valutazione dei sistemi educativi. La traccia integrale è riportata anche nell'Appendice I.

From the data published annually by the OECD (*Education at a Glance 2020*) it appears that in 2017 Spain invested just over 4% of its national GDP in education in line with the average of European countries. At an economic level, what is the budget available to INEE? Does it receive funds from the state?

In most cases, the external evaluation of the educational system is accompanied by the self-evaluation of schools as a moment of reflection and improvement. Is it mandatory in your country? How often does it happen?

How are the system assessment studies, like the OECD-PISA program, including the important variable of the pandemic? In such a delicate scenario, what, in your opinion, can be the contribution of the system evaluation for restarting?

Tabella 5. Traccia domande sul rapporto investimenti pubblici-valutazione, sull'autovalutazione e sulla situazione emergenziale-pandemica.

3.4. Alcune riflessioni dalle interviste agli NPM dell'indagine OCSE-PISA - valutazione standardizzata degli apprendimenti

I paesi oggetto di analisi, Italia, Finlandia e Spagna, oltre a far ricorso a modelli di valutazione del sistema scolastico differenti, riportano anche esiti diversi nell'indagine PISA dell'OCSE, sia in termini di competenze degli studenti quindicenni misurate attraverso la prova

cognitiva, sia in termini di variabili di sfondo che influiscono sui processi di insegnamento-apprendimento (OECD, 2019abc). L'indagine, come ricordato nel capitolo tre della prima parte del presente lavoro, ha l'obiettivo di rilevare, con cadenza triennale, il grado con cui coloro che si apprestano a concludere l'obbligo scolastico sono capaci di applicare e utilizzare nella vita adulta quanto appreso tra i banchi di scuola. Gli ambiti indagati sono la lettura, la matematica e le scienze. A rotazione, ogni tre anni, il focus si concentra su uno di questi tre ambiti, mentre gli altri due vengono misurati attraverso un set di domande ridotto. Nel 2018 il focus principale era sulla lettura, mentre il *dominio innovativo* era sulle competenze globali (OECD, 2019abc). Nel presente paragrafo verrà dato conto delle principali riflessioni emerse durante le interviste con i National Project Manager dell'indagine in merito alla risonanza che la valutazione standardizzata degli apprendimenti ha all'interno della società, prima tematica emersa nell'analisi dei dati. Per facilitare la contestualizzazione delle opinioni e delle percezioni emerse verrà dapprima ricostruito il profilo del paese nella più estesa indagine sui livelli di apprendimento, successivamente verranno riportati alcuni estratti significativi delle interviste.

Gli studenti finlandesi, fin dalla prima edizione dello studio (OECD, 2001), hanno raggiunto punteggi superiori alla media dei paesi partecipanti sia in lettura, che matematica e scienze (tradizionalmente la media OCSE è stabilita a 500 punti, tuttavia negli ultimi anni, a causa del peggioramento delle performance degli studenti, quest'ultima si è abbassata). Nel 2018 il punteggio medio finlandese in lettura è stato di 520 punti, in matematica 507 e nelle scienze 522. Nonostante in lettura gli studenti totalizzino un punteggio che si inserisce tra i primi cinque paesi al mondo, in matematica tra i primi dodici e nelle scienze tra i primi nove, le competenze dei quindicenni sono protagoniste di un lento declino che riguarda tutti e tre gli ambiti e che ha iniziato a manifestarsi dopo il 2006 (OECD, 2007). Pur rimanendo sopra la media degli altri paesi partecipanti, la differenza tra i punteggi medi del 2015 e quelli del 2018 è particolarmente significativa per le scienze, mentre rimane più contenuta per la matematica e la lettura. Nello specifico dei livelli di competenza, quelli indicati dall'OCSE per descrivere cosa lo studente è

in grado di fare (livello inferiore a 2 - 1a, 1b, 1c, livello 2, livello 3, livello 4, livello 5, livello 6), negli ultimi anni in matematica il declino è stato uniforme in tutti i livelli. Tuttavia, in particolar modo in lettura, ma anche nelle scienze, il declino è stato registrato soprattutto tra i *lower performers* (livello 2 o inferiore). Una delle possibili letture di tale fenomeno potrebbe essere attribuita al divario di genere: nel 2018 è stata riportata una differenza di 52 punti tra la performance degli studenti e quella delle studentesse, uno dei divari più ampi dell'ultima edizione disponibile. Al contrario, il gap dovuto a differenze di status socio-economico è più contenuto rispetto alla media degli altri paesi (79 punti vs 89 punti). I buoni risultati raggiunti da più di venti anni dagli studenti finlandesi sono da ricondursi, a parere del National Project Manager intervistato, al significato dato socialmente all'educazione (aspetto emerso dall'analisi dei dati qualitativi). Quest'ultimo può essere declinato in vari modi: da una parte vi è la formazione specialistica degli insegnanti sia da un punto di vista disciplinaristico che da un punto di vista psico-pedagogico, dall'altra vi è l'omogeneità tra scuole nella misura in cui l'offerta formativa di ogni istituzione scolastica non differisce sostanzialmente dalle altre. A variare è la varianza interna a ciascuna scuola dovuta al fatto che gli studenti hanno livelli di competenza e background differenti, ma non quella tra scuole poiché il sistema basato sull'equità permette ad ogni studente di avvalersi dello stesso insegnamento di qualità indipendentemente dall'istituto scolastico frequentato. Bisogna sottolineare che, ciononostante, negli ultimi anni le differenze dovute al background socio-economico-culturale sono aumentate.

Questo aspetto è divenuto argomento di dibattito a livello nazionale. Anche in Finlandia il background socio-economico delle famiglie ha iniziato ad influire di più sul successo degli studenti rispetto al passato. In questi ultimi anni c'è stato uno sviluppo in negativo. Nel 2009 sembrava ancora che il sistema scolastico finlandese fosse in grado più degli altri paesi OCSE di livellare questo gap, ma in questi ultimi dieci anni siamo arrivati quasi allo stesso livello degli altri. L'influenza del background familiare è aumentata anche in Finlandia ed è diventata un fattore di preoccupazione. Stiamo perdendo qualcosa? Questo era uno degli aspetti che rendevano il sistema scolastico finlandese di successo.

Sembra che al crescere dell'influenza del background socio-economico, sia diminuito il livello di competenze medio. Non so dire se ci siano altri modi per arrestare questa inversione, noi lo abbiamo sottolineato più volte nei nostri rapporti: c'è un trend del genere anche qui da noi. (NPM finlandese, traduzione a cura dell'autrice)

Similmente in Finlandia un altro aspetto che desta preoccupazione è il già citato gap di genere relativo alle competenze degli studenti e delle studentesse in aumento negli ultimi anni.

Se compariamo, per esempio, i risultati finlandesi in lettura con quelli internazionali, ci rendiamo conto che nell'indagine PISA abbiamo uno tra i gap di genere più elevati, cioè la differenza tra le competenze degli studenti e delle studentesse è di 50 punti in media a favore di queste ultime, o meglio nei punteggi PISA la loro competenza in lettura è di 50 punti superiore. [...] Questo è un aspetto che ha catalizzato l'attenzione: la differenza di genere nelle competenze è stata sempre a favore delle ragazze, ma negli ultimi dieci anni abbiamo iniziato a notare queste differenze anche in matematica e nelle scienze sempre a favore delle studentesse. In poche parole le competenze degli studenti finlandesi sono peggiorate di più di quelle delle studentesse. (NPM finlandese, traduzione a cura dell'autrice)

Gli studenti spagnoli, invece, in PISA 2018 hanno raggiunto un punteggio medio inferiore alla media OCSE sia in matematica (481 punti) che nelle scienze (483 punti) avvicinandosi alla performance degli studenti lituani e ungheresi e mantenendo il trend temporale costante in tutti gli ambiti oggetto di indagine. Diversamente dalla Finlandia e dalla media OCSE, tra i quindicenni spagnoli non ci sono molti *top performers* (livello 6) e il livello dei *low performers* (livello 2 o inferiore) è in linea con la media degli altri paesi. Il background socio-economico dello studente spiega il 12% (vs 14% media OCSE) della varianza nei risultati in matematica e il 10% (vs 13% media OCSE) della varianza dei risultati nelle scienze. Questo risultato, vicino alla media degli altri paesi, è dovuto a una serie di riforme centrali che sono state implementate negli ultimi anni.

Mi stupisce che il numero di libri che si hanno a casa sia ancora un buon indicatore e predittore dei risultati dello studente. Certo... L'indicatore più rilevante è il background familiare, specialmente quello della madre. Il titolo di studio materno, l'occupazione materna. Nell'educazione formale, negli ultimi due decenni, stiamo cercando di mitigare questi effetti, tuttavia l'impatto che hanno è la causa di tutte queste differenze nei risultati. Certo che lo è. Sicuramente conoscerai il sistema educativo spagnolo e saprai che è molto diviso a seconda della regione di riferimento. Per questo abbiamo cercato di fare del nostro meglio, a partire dal ministro e dal ministero, nell'aver una politica comune per le valutazioni. Infatti, in Spagna ci sono, da un certo punto di vista, 17 sistemi educativi differenti. Ad esempio a Madrid c'è più supporto... La politica regionale di Madrid è molto concentrata sulle scuole private, laddove in altre regioni non è così. Qui a Madrid c'è più segregazione tra le scuole, nelle scuole e tra gli studenti rispetto ad altre regioni come l'Andalusia o Castiglia e Leon. Spesso questa diversità non è propriamente legata al colore del partito che è al governo poichè la regione di Madrid e di Castiglia e Leon hanno lo stesso partito conservativo al potere, però loro non hanno la stessa... Così c'è più segregazione a Madrid e più equità a Castiglia e Leon. Tuttavia, stiamo cercando di... Dal ministero vengono distribuiti i fondi, derivanti dalle tasse, in tutto il paese. Noi cerchiamo di dirigere la politica verso la richiesta di borse di studio e sostegno per gli studenti più svantaggiati. Dobbiamo cercare di incrementare questo aspetto. Tuttavia, uno dei fattori che influisce sugli insegnanti e sugli studenti, entrando nei processi di insegnamento-apprendimento, è, per esempio, il numero di studenti presenti in ogni classe. Dal ministero c'è una raccomandazione circa il numero di studenti per classe, però ogni regione interpreta questa raccomandazione in modi diversi. Credo, per questo motivo, che a partire dal ministero dovremmo essere più assertivi e sottolineare che si tratta di una legge per tutti. Tuttavia... Si sa, cerchiamo di discutere di questi aspetti per raggiungere il consenso di tutti. Uno degli aspetti che ritengo... Sono un'insegnante, sono un'insegnante di scuola secondaria e una docente accademica e penso che il numero di studenti per classe sia cruciale per dare il supporto e l'educazione individualizzata di cui hanno bisogno. (NPM spagnola, traduzione a cura dell'autrice)

Nonostante la descrizione della National Project Manager della Spagna, gli indicatori internazionali evidenziano che nel 2019, ultimi dati disponibili, il numero di studenti per classe nell'istruzione pubblica era di 20.9 unità, mentre nell'istruzione privata di 24.1 unità. Se il dato

sull'istruzione finanziata dallo stato è in linea con la media OCSE (21.1 unità), quello dell'istruzione privata è ben al di sopra della media OCSE (19.7 unità) e, come riportato dalla responsabile nazionale dell'indagine PISA dell'OCSE, è un fattore che va attenzionato soprattutto in quelle regioni in cui l'istruzione privata è molto comune (OECD, 2021b).

Le differenze di genere sono più ridotte rispetto a quelle registrate per la Finlandia. Tra gli studenti e le studentesse vi sono 6 punti di differenza in matematica a favore dei primi (5 punti in media OCSE) e 2 punti di differenza nelle scienze a favore delle seconde similmente alla media OCSE (OECD, 2019ab).

Stiamo registrando in media gli stessi risultati anche se c'è ancora una differenza di genere in lettura, performance migliore per le studentesse e peggiore per gli studenti, e in matematica, performance leggermente migliore per gli studenti. Si tratta però di un problema della nostra società, molto grande per le nostre studentesse, soprattutto in matematica. Abbiamo fatto alcune analisi confermatrice dei risultati PISA attraverso un'altra valutazione esterna, una valutazione nazionale e... È un dato di fatto, ma abbiamo anche delle ottime notizie per quanto riguarda gli studenti immigrati e l'integrazione in quanto i risultati per gli studenti immigrati di seconda generazione sono molto molto simili a quelli dei nativi. Questo aspetto è emerso sia nelle indagini internazionali che in quelle nazionali. Sembrerebbe che stiamo facendo un buon lavoro con gli studenti non nativi. (NPM spagnola, traduzione a cura dell'autrice)

Nell'edizione 2018 di PISA la Spagna ha riscontrato un problema non irrilevante nella misurazione delle competenze in lettura. L'OCSE ha deciso di non diffonderne i risultati, né a livello nazionale né a livello subregionale poiché, nonostante vengano soddisfatti i criteri tecnici, dall'analisi dei dati è emerso un comportamento non plausibile da parte degli studenti nell'inserire la risposta, fattore che inficia la comparabilità dei risultati. Lo stesso comportamento è stato rilevato anche in matematica e scienze, ma con un peso significativamente inferiore (OECD, 2019ab). Gli studenti quindicenni spagnoli, così come quelli italiani e finlandesi, raggiungono risultati peggiori rispetto a quelli delle indagini IEA, in particolare nelle indagini PIRLS e TIMSS che

interessano gli studenti della scuola primaria e della scuola secondaria di primo grado. Questa differenza, a parere della National Project Manager della Spagna, potrebbe essere dovuta non solamente ad ambiti differenti testati nei diversi studi ma anche ad attitudini diverse da parte degli studenti.

Probabilmente i nostri studenti di scuola secondaria sono stanchi delle valutazioni e di queste indagini esterne. Hanno già molti esami a scuola. Infatti, abbiamo avuto un problema molto molto grande con i risultati in lettura. Non so se hai letto l'ultimo report contenente i dati del 2018 in cui c'è una nota che spiega alcuni di questi nostri problemi. Alcuni studenti di alcune regioni, soprattutto quelli tra i *best achiever*, non erano molto entusiasti di fare questo test poiché si sovrapponeva con altri test scolastici. Ci abbiamo perso un po' perchè non potevamo prevederlo e dal ministero non potevamo reagire così velocemente come avremmo dovuto. Così... Alcuni dei nostri studenti migliori non hanno dato il loro massimo nel test in lettura di PISA. Nelle scienze non c'era molta differenza, mentre in matematica... Beh, ci sono stati alti e bassi in questi 20 anni di PISA. Dovremmo approfondire questo aspetto di più... Di sicuro c'è la fatica. C'è un problema di affaticamento nei nostri studenti di scuola secondaria. Quelli della scuola primaria è più facile metterli sul compito, ma per i quindicenni... (NPM spagnola, traduzione a cura dell'autrice)

Gli studenti italiani, infine, raggiungono un punteggio inferiore alla media OCSE in lettura e nelle scienze (476 punti vs 487 media OCSE in lettura; 468 punti vs 489 media OCSE in scienze), mentre rientrano nella media OCSE per la matematica (487 punti vs 498 media OCSE). Contrariamente ai due paesi analizzati fino ad ora, se i punteggi medi in lettura e nelle scienze sono peggiorati dopo il 2012, questo peggioramento è attribuibile in particolar modo alle studentesse per la lettura (la performance degli studenti è rimasta pressoché stabile) e agli studenti con performances migliori (livelli 5 e 6) sia maschi che femmine per le scienze. Il livello raggiunto dai quindicenni italiani è simile a quello dei quindicenni spagnoli in matematica e inferiore per lettura e scienze. Così come riportato da anni, anche nell'edizione 2018 di PISA sono state rilevate differenze tra macroaree e indirizzi di studio. Per quanto riguarda le macroaree, oltre alle differenze preesistenti tra Nord

e Sud, è stato possibile condurre analisi differenziate per quattro regioni/province in cui l'ampiezza delle scuole campionate lo ha permesso. Nello specifico della lettura, le province autonome di Trento e Bolzano hanno raggiunto risultati sopra la media nazionale, simili a quelli di Slovenia e Germania, la regione Toscana in media nazionale, mentre la regione Sardegna al di sotto della media nazionale, simili a quelli di Turchia e Grecia. In matematica, invece, le province autonome di Trento e Bolzano hanno raggiunto risultati simili ai paesi *best performers* europei (livello 6) (Estonia, Paesi Bassi, Svizzera e Polonia), lasciando davanti a sé solamente le economie asiatiche (OECD, 2019abc).

Quando i nostri studenti si trovano ad affrontare quel tipo di prove non è che non sono in grado per niente, ma non lo sono totalmente. E questo infatti... In matematica siamo migliorati nel tempo, quindi abbiamo raggiunto la media OCSE. Però anche la matematica ha questa caratteristica... Probabilmente trainati dalle prove nazionali si è cominciato a lavorare di più su questi test che non erano i classici problemi che si vedono sui libri di testo. Invece lettura e scienze... Diciamo le scienze, beh, non si fanno nelle prove nazionali, quindi gli studenti non hanno usufruito di questo aspetto, di questo vantaggio qui. La lettura, probabilmente, ecco, il modo in cui si usano i testi a scuola, è completamente diverso da quella tipologia di prova proposta. Anche qui sarebbe interessante fare un confronto con altri paesi e analizzare, per esempio, la Finlandia, dove invece in lettura si raggiungono buonissimi punteggi... Analizzare, per esempio, se c'è un tipo di valutazione che prende in considerazione anche questo aspetto del transfer cognitivo oppure, in ogni caso, cosa si fa di diverso da noi a scuola. (NPM italiana)

Da anni, come accennato, in Italia sono registrate differenze nel rendimento tra studenti e studentesse, tra nativi e immigrati, tra macroaree e tra indirizzi di studio. Per quanto riguarda i risultati in lettura, le differenze di genere nel 2018 erano inferiori a quelle della media degli altri paesi e a quella registrata nel 2009, tuttavia, l'assenza di questo gap non è dovuta ad un restringimento della forbice, bensì al peggioramento della performance delle studentesse in linea con la performance degli studenti rimasta stabile. Il rapporto nativi-immigrati, invece, è cambiato negli ultimi anni: se nel 2009 gli studenti non nativi

erano il 6% della popolazione totale, nel 2018 sono diventati il 10%. La differenza media tra la performance degli studenti nativi e quella degli studenti immigrati è di 43 punti, tuttavia, se si tiene conto del background socio-economico-culturale di questi ultimi la differenza scende a 22 punti (OECD, 2019ab).

Il fatto di avere individuato queste differenze è proprio, penso, il motivo e il valore principale della valutazione standardizzata, cioè il fatto di utilizzare un metro comune di valutazione, permette di mettere in evidenza le differenze che altrimenti non sarebbe possibile evidenziare. La valutazione standardizzata esterna ha questo vantaggio che non è possibile in nessun altro modo. Magari si sapeva già, no?! Si aveva quella impressione vivendolo e sentendone parlare [...] però i dati ci confermano che c'è quel determinato elemento, quindi è un'altra cosa. Insomma, possiamo ragionare su dati oggettivi. Le differenze che ha citato lei sono diverse. Un conto è parlare di differenze fra indirizzi di studio, per esempio, perché ce le possiamo aspettare. È anche naturale che ci siano differenze tra indirizzi di studio, perché è chiaro che lo studente che è particolarmente dotato nelle materie matematiche e scientifiche, non si iscriverà all'istituto professionale, si iscriverà al liceo e così invece viceversa. [...] La diversa utenza che arriva nei diversi indirizzi di studio è la principale motivazione della differenza nei risultati. [...] Ciò si va a intersecare anche con lo status socio-economico e culturale della famiglia. Questo è già meno naturale perché è ovvio che... Il figlio della famiglia benestante, colta, istruita, anche se non è molto brillante a scuola, magari si iscrive al liceo. [...] Insomma, diciamo che le differenze fra indirizzi di studio sono abbastanza naturali. Invece un'altra questione sono le differenze fra macroaree e quelle di genere. Lì il discorso è più antipatico perché sono dovuti... Cioè a quel punto si ha proprio un senso di iniquità del sistema educativo che non dà evidentemente a tutti le stesse opportunità e per motivi che nulla hanno a che vedere con le persone. Perché il fatto che io nasca al Nord o al Sud non dovrebbe incidere sulle mie opportunità, oppure nasco femmina invece che maschio non dovrebbe incidere. Il discorso nativi e immigrati, secondo me è molto legato sia all'aspetto socio-economico-culturale che a quello territoriale. Più che al discorso di essere immigrati, perché in alcune analisi si è visto che quando si tiene sotto controllo lo status socio-economico e culturale della famiglia, si elimina una bella fetta di differenza. Quindi in realtà è più quello... Poi ovviamente è chiaro che l'immigrato che magari non parla bene la lingua è un po' più in svantaggio, però pure lì se c'è la motivazione... Spesso sono più motivati loro dei nativi. Superano tutta una serie

di difficoltà. Insomma sì, gli interventi quali potrebbero essere? Questa è la domanda da 1.000.000 di dollari. A mio parere qui servirebbero riforme, riforme strutturali forti. [...] Immagino che creare opportunità maggiori per i giovani sia a livello sociale, culturale ed educativo, ma anche occupazionale, trainerebbe anche il rendimento scolastico. Io la penso così. È chiaro che qui ci sono interventi grandi che andrebbero fatti. Insomma, penso che sia molto difficile intaccare questa realtà. Per quanto riguarda le differenze di genere... Anche lì sono radicate nella cultura e anche nel come si fa scuola perché evidentemente... La matematica, per dire, viene probabilmente insegnata già nella scuola primaria in un modo che è più consono ai bambini che alle bambine. Viceversa, per la lettura. [...] è anche una questione di formazione degli insegnanti. Gli insegnanti andrebbero formati per un'istruzione di genere, tra virgolette, cioè che sia sensibile a questo aspetto e quindi non porti allo stereotipo della bambina che odia la matematica o del bambino che odia leggere, che odia le storie. Perché è chiaramente una cosa costruita culturalmente. Però ecco, sono interventi, che è difficile... Richiedono proprio di andare in profondità in aspetti culturali radicati. [...] Mai si comincia, mai si fa nulla. Si potrebbe iniziare da piccole cose... Però per prima cosa sarebbero necessari ulteriori studi per comprendere meglio quali sono i fattori, soprattutto nel discorso maschi-femmine, quali sono i fattori che producono queste differenze e, dopodiché, quando sappiamo meglio quali sono i fattori, si può anche intervenire. (NPM italiana)

Sia in Italia che negli altri paesi oggetto di studio la valutazione standardizzata degli apprendimenti non è e non è stata priva di riflessioni sia di tipo pratico-metodologico che ideologico. Zhao (2020) individua tre assi principali lungo i quali si muovono le critiche nei confronti dell'indagine PISA dell'OCSE: la visione dell'educazione che sottostà all'indagine, la sua attuazione e interpretazione e l'impatto che produce sull'istruzione a livello globale (Bracey, 2005; Eivers, 2010; Fischbach et alii, 2013; Auraujo, 2017; Rautalin, 2018). La questione è ben chiara anche ai National Project Manager intervistati, tuttavia l'interpretazione data sia dalla comunità scientifica che dall'opinione pubblica alla valutazione esterna degli apprendimenti varia a seconda del contesto di riferimento.

Probabilmente grazie alla storia di successo che abbiamo alle spalle nell'indagine PISA, posso dire che i media e il dibattito pubblico sono abbastanza interessanti allo studio. Ogni volta che vengono diffusi i risultati, se ne parla. Se ne parla anche durante i tre anni tra un'edizione e un'altra, infatti, quando vengono pubblicati ulteriori report, l'attenzione si sofferma su un aspetto o su un altro. È un'indagine che interessa molto, anche se non saprei dire quale fattore interessa di più e su quale si scrivono più articoli. Ci sono critiche? Certo che ci sono. Le principali critiche ideologiche ruotano per l'appunto attorno all'influenza che l'indagine può avere sulle decisioni politiche riguardanti il sistema scolastico. Si sta iniziando a modificare e guidare la didattica verso la direzione che porta al miglioramento degli esiti? In poche parole, i framework di PISA influiscono sulla politica nazionale... In questo senso sono state mosse alcune critiche. Tuttavia, la mia personale opinione è che finché i quadri di riferimento dell'OCSE sono sistematici, sono gestiti da persone competenti e c'è una riflessione corale con i rappresentanti dei paesi partecipanti... [...] La valutazione non deve diventare un modo per guardare solamente alla vita professionale di un individuo e alle competenze che questa richiede... Poi costruiamo la valutazione per rispondere a questa domanda. In quel caso può succedere che la valutazione subisce un'inversione di tendenza e porta a uno sguardo molto limitato nel tempo. Il lavoro valutativo deve essere lungimirante e deve avere sempre presenti quali sono le competenze importanti. (NPM finlandese, traduzione a cura dell'autrice)

Diversa è la situazione in Spagna in cui non mancano le critiche mosse verso l'OCSE, ma dove la percezione ideologica della valutazione standardizzata degli apprendimenti dipende anche da altri fattori.

Per alcuni anni, dopo la grande crisi finanziaria, si è discusso molto della valutazione esterna. Al governo in quel periodo c'era il partito popolare conservatore, tra il 2011 e il 2019. [...] Ora abbiamo un governo di coalizione di sinistra. Ci sono punti di vista ideologici differenti in relazione alla valutazione esterna. Per i conservatori la valutazione esterna è tutto. Loro vorrebbero fare sempre esami perchè... Quando il partito conservatore era al potere, hanno provato a implementare test per tutti, non test a campione, ma per tutti. Abbiamo dovuto persuaderli dicendo che era impossibile, siamo in crisi, non possiamo spendere tutti questi soldi per i test. [...] Gli approcci di sinistra sono più per l'integrazione. [...] È un bilanciamento difficile tra due prospettive. In Spagna, poichè abbiamo questo crogiolo di regioni, utilizziamo le indagini

internazionali per fare una possibile comparazione tra le comunità autonome. Le regioni accettano meglio questo approccio, l'approccio internazionale rispetto a quello nazionale del ministero centrale. Questo è il motivo per cui... Per avere dei risultati comparabili internazionalmente dovremmo creare un tavolo di lavoro con le regioni dal punto di vista dell'indagine PISA, non dal punto di vista nazionale poiché in passato è stato molto contestato, spero tu possa capire la situazione. In un certo senso la misurazione internazionale è rispettata da tutti i partiti ideologici. Anche se arriva PISA, l'OCSE e così via, è ben rispettato anche nella coalizione di sinistra. Noi facciamo team con loro, loro fanno team con noi, ci sono anche l'Unesco e l'Unione Europea, organismi sovranazionali. Penso che questa sia una delle ragioni per cui gli studi OCSE non sono molto contestati in Spagna... Tuttavia, ci sono molte critiche nei confronti dell'OCSE. PISA vuole essere uno studio universale. Sono la National Project Manager dell'indagine, ma non posso non vedere i limiti dello studio. I team di PISA stanno facendo molto lavoro, per esempio, sulle competenze socio-emotive. Cercano di integrare sempre più aspetti legati all'educazione nello studio PISA. Abbiamo bisogno di più studi per comparare. È un peccato, secondo me, che l'Unione Europea non stia favorendo sempre più le proprie indagini. (NPM spagnola, traduzione a cura dell'autrice)

In Italia lo scenario è differente rispetto a quello spagnolo e finlandese in relazione alle critiche nei confronti della valutazione standardizzata degli apprendimenti.

La valutazione PISA o anche quelle nazionali che fa l'Invalsi chiaramente non sono cose perfette. Insomma, sono delle cose che sono migliorabili e nel tempo sono migliorate. Se prendiamo le prove nazionali, rispetto a quando sono partite, sono stati fatti molti miglioramenti, molti... Il lavoro è diventato più raffinato, più preciso, più rigoroso. Non è una macchina statica che è rimasta ferma nel tempo, indipendentemente da quello che è successo all'esterno. Sia le indagini internazionali che quelle nazionali cercano di seguire l'evoluzione della società. Pensiamo ai framework nelle indagini internazionali che vengono sempre aggiornati. È chiaro però che dei difetti li hanno. Insomma, non è che sono perfetti. Prendendo PISA si vede molto forte l'impronta dell'OCSE di tipo economicista che magari va a trascurare altre questioni. [...] Secondo me la repulsione per la valutazione esterna dipende dal fatto che si ha paura in qualche modo, cioè si scatena quella paura naturale della valutazione che abbiamo tutti, perché la valutazione la viviamo come una bocciatura o una

promozione. La valutazione è solo intesa, spesso anche inconsciamente, come qualcosa che o boccia o promuove. Quindi, fundamentalmente, abbiamo paura della bocciatura. Abbiamo paura del no e trascuriamo invece tutto l'aspetto formativo della valutazione, ovvero l'aspetto migliorativo che invece è il principale scopo della valutazione esterna standardizzata, perché la valutazione esterna non va a bocciare né a promuovere nessuno e soprattutto non porta né a premi né a punizioni. Molti insegnanti, per esempio, la vivono in questo modo. Però l'esperienza ormai di anni ha insegnato che non è così. Nessun insegnante in Italia ha avuto una sanzione perché la sua classe era andata male alle prove Invalsi. È che noi non siamo abituati a vedere la valutazione come una fonte di informazioni per poter migliorare. (NPM italiana)

Nonostante le critiche, nonostante le visioni politiche o le reazioni che la valutazione esterna degli apprendimenti può suscitare, si tratta di una valutazione formativa che mette in relazione il fare ricerca e gli esiti che da essa derivano con i processi decisionali (Calvani & Vivanet, 2014).

3.5. Alcune riflessioni dalle interviste ai NPM dell'indagine OCSE-PISA - valutazione della scuola e per la scuola

Nel presente paragrafo verranno illustrate le percezioni, riportate dai National Project Manager dell'indagine PISA dell'OCSE durante le interviste, in merito all'utilità che la valutazione delle istituzioni scolastiche ha per le medesime istituzioni scolastiche, seconda tematica emersa nell'analisi dei dati. Come avvenuto nel paragrafo precedente, per facilitare la contestualizzazione delle opinioni e delle percezioni emerse verranno presentati alcuni estratti significativi delle interviste.

Le reazioni che i processi valutativi provocano nei valutatori e nei valutati possono variare. Il modo in cui la valutazione viene vissuta, infatti, può dipendere dallo sviluppo storico che quest'ultima ha avuto, dagli obiettivi che attraverso essa si vogliono raggiungere e dal senso di coinvolgimento vissuto dagli attori coinvolti (OECD, 2013). Si tratta di un approccio non solamente scientifico, ma anche culturale nei confronti dei processi valutativi, come confermato dai National Project Manager.

In Finlandia, a parere del responsabile nazionale dell'indagine PISA dell'OCSE, la valutazione esterna di tanto in tanto è percepita come

gravosa poichè impegna molto tempo, tuttavia, quella degli apprendimenti, ad esempio, non coinvolge tutte le istituzioni scolastiche dal momento che si basa su un piano di campionamento.

Certamente se la valutazione è percepita come gravosa, allora ci si chiede anche l'importanza che ha. In generale però posso dire che l'approccio alla valutazione qui è positivo. Per la mia personale esperienza anche l'atteggiamento nei confronti dell'indagine PISA è positivo, gli studenti partecipano e non abbiamo problemi con i tassi di risposta. Senza dubbio c'è anche chi è critico e negativo nei confronti della valutazione esterna, ma si tratta di una minoranza. (NPM finlandese, traduzione a cura dell'autrice)

In Spagna, invece, il consenso nei confronti della valutazione è aumentato sempre di più negli ultimi anni. Dirigenti scolastici, docenti, studenti e famiglie hanno familiarizzato maggiormente con queste forme di valutazione, percependola come un qualcosa che non si può evitare, tuttavia vorrebbero essere più partecipi.

Loro vorrebbero essere più coinvolti nei risultati, perchè, si sa, quando facciamo una valutazione di sistema, realizziamo degli studi che non sono rappresentativi per ogni scuola. Spesso le scuole vorrebbero avere più feedback. In PISA, per esempio, è molto difficile fare un discorso del genere perchè abbiamo a che fare con una popolazione studentesca, non con il singolo studente. Sì, c'è stata una familiarizzazione con la valutazione. Il problema, tuttavia, è che negli ultimi 10-20 anni abbiamo moltiplicato troppo le nostre valutazioni e i nostri monitoraggi. [...] Soprattutto le indagini internazionali sono come una grande ruota sempre in movimento che non possiamo arrestare. (NPM spagnola, traduzione a cura dell'autrice)

L'Italia attualmente vive una situazione simile a quella spagnola in cui la valutazione esterna è divenuta, con il passare degli anni, sempre più familiare.

Sicuramente la situazione è un po' cambiata perché forse si sono abituati, di conseguenza la inseriscono [la valutazione] nella routine quotidiana dei vari adempimenti. [...] Sicuramente molti hanno anche capito che il senso di minaccia è infondato. Molti magari hanno anche apprezzato, ci arrivano molti

apprezzamenti. Insomma, gli insegnanti e i dirigenti scolastici apprezzano l'informazione che arriva dalle prove. Però sicuramente c'è ancora tanto... Tanto rifiuto di questo. Come dire? Intolleranza rispetto a questa valutazione che viene percepita come una invadenza, come un'invasione indebita dentro il proprio territorio. Questa percezione ancora c'è, anche se è sicuramente diminuita. (NPM italiana)

Avere chiaro il significato della valutazione è senza dubbio tra quei fattori che ne favoriscono l'accettazione. In Finlandia, da quanto riportato dal National Project Manager, gli insegnanti comprendono l'importanza dei processi valutativi, sanno che non si tratta della valutazione di un singolo studente, di un singolo insegnante o di una singola scuola, ma di un monitoraggio più ampio del sistema scolastico. Questo non comporta "effetti collaterali negativi" dovuti alla valutazione. Diversa è la percezione degli stakeholders spagnoli.

Per il sistema pubblico, per le scuole pubbliche... Non ripongono particolari speranze in questo tipo di valutazione. Hanno i loro problemi quotidiani. Tuttavia, dalla prospettiva delle scuole target... La prendono più seriamente. Di fatto non abbiamo dati su questi aspetti. Intendo che... Credo che gli insegnanti non abbiano tempo di praticare il *teaching to the test*, per esempio per PISA. Al momento hanno tante cose da fare, sono abbastanza sopraffatti. (NPM spagnola, traduzione a cura dell'autrice)

In Italia, invece, la questione dei comportamenti opportunistici affonda le sue radici nel tempo (Trincherò, 2014c; Falzetti, 2019).

Perché c'è il *teaching to the test*? Perché si vuole andare bene, ma non è quello lo scopo. Lo scopo della prova è monitorare, ma soprattutto per te [insegnante]. Tu devi capire se il tuo insegnamento ordinario porta a quei risultati oppure no. Se tu addestri il bambino e il ragazzo a fare le prove chiaramente introduci un bias, a quel punto il dado che ti torna indietro non ti dice realmente qual è la situazione. (NPM italiana)

Il coinvolgimento degli attori inseriti nei processi valutativi è un altro degli aspetti della valutazione delle istituzioni scolastiche che fa di quest'ultima una valutazione per la scuola, a servizio del miglioramento.

È importante, per questo motivo, costruire una comunità all'interno della quale il coinvolgimento e l'accoglimento della valutazione sono visti come responsabilità condivise (OECD, 2013).

In Finlandia, in questo senso, si cerca di lavorare molto sulla partecipazione e sul coinvolgimento degli studenti, in primis, nei momenti valutativi che li riguardano, come quelli degli apprendimenti, ma anche sulla percezione del senso comune, non solamente di chi vive la scuola, in merito a chi realizza la valutazione.

Riteniamo che la partecipazione ai momenti valutativi sia parte del lavoro scolastico dello studente. [...] Non abbiamo particolari strategie motivazionali. Cerchiamo solo di... Inviando il materiale agli studenti campionati, documenti che spiegano l'indagine, e li sproniamo a partecipare, incoraggiandoli, sottolineando che si tratta di una buona occasione di apprendimento poichè le prove PISA sono molto diverse da ciò che vedono quotidianamente a scuola. [...] Gli istituti di ricerca educativa sono visti come partner esperti e validi, come implementatori affidabili delle ricerche in corso. [...] Un'autorità affidabile. Collaboriamo molto con le scuole perchè sono loro che raccolgono i dati per noi. Gli insegnanti sono responsabili di queste raccolte dati, noi li supportiamo e inviamo i materiali. Interagiamo molto con loro. La percezione è che si avvicinano a noi positivamente, anche se a volte ci dobbiamo anche un po' scusare con loro per le richieste di collaborazione che gli inviamo. Li contattiamo e questo vuol dire per loro lavoro in più sia preparatorio che di partecipazione. (NPM finlandese, traduzione a cura dell'autrice)

In Spagna la situazione è differente. Le valutazioni che avvengono nelle comunità autonome sono censuarie, tuttavia le istituzioni scolastiche non sempre hanno chiare queste ultime. Inoltre, vedono gli istituti in cui si fa ricerca come posti lontani da loro.

Il sistema di valutazione regionale è per tutti. Loro hanno la valutazione esterna organizzata e sviluppata dalle autorità regionali. Inoltre, ci sono le indagini internazionali o alcuni altri studi nazionali del ministero. A volte le scuole non hanno ben chiaro di che tipo di test si tratta, se proviene dall'autorità regionale, dal ministero o dall'unità internazionale. Dovremmo essere più chiari sugli obiettivi di ogni valutazione. Uno dei problemi che abbiamo in Spagna è che non possiamo andare direttamente nelle scuole. Mi piacciono molto i paesi

come l'Estonia, per esempio. La mia collega estone conosce ogni scuola del paese. Loro vanno lì a fare visita, parlano con i dirigenti scolastici, con gli studenti, con gli insegnanti. È tutto molto più diretto, c'è molto più coinvolgimento. In Spagna però abbiamo questi step differenti. Dobbiamo passare dalle autorità regionali per arrivare agli studenti e agli insegnanti. Credo che questa sia una delle difficoltà che riscontriamo nel nostro lavoro al ministero. [...] Noi siamo visti come in un posto lontano. [...] L'INEE non è percepito vicino alle scuole, tuttavia non siamo due cose diverse e separate. Possiamo collaborare insieme alle regioni. Molte persone in Spagna criticano questo stato regionale, federale, ma è una scusa per evitare il confronto, il confronto ideologico. Una delle cose che più apprezzo in questo sistema è che possiamo lavorare con ogni regione e non possiamo prendere nessuna decisione se non abbiamo prima raggiunto il consenso unanime. È un buon forum per l'educazione. (NPM spagnola, traduzione a cura dell'autrice)

In Italia la valutazione del sistema educativo si compone sia della valutazione esterna che dell'autovalutazione, ma non mancano le azioni di miglioramento e la rendicontazione sociale, come indicato nel D.P.R n. 80 del 2013 (D.P.R 80/13, 2013).

Ovviamente la parte che coinvolge di più è quella dell'autovalutazione per definizione, perché la valutazione standardizzata non deve coinvolgere troppo, altrimenti non sarebbe più standardizzata se coinvolgesse attivamente, attivamente nel senso di come avviene la valutazione. Però anche in quella standardizzata, prima della valutazione si può informare bene. E questo è uno degli sforzi che si stanno facendo, si stanno facendo dei webinar alle scuole, spiegando bene alcuni elementi. [...] Secondo me sulle famiglie bisognerebbe fare uno sforzo maggiore di coinvolgimento e sicuramente anche con gli studenti, perché molti studenti stessi non capiscono perché devono sottoporsi a queste prove. (NPM italiana)

Le scuole utilizzano i dati provenienti dai processi valutativi per prendere decisioni? Se in Finlandia la risposta sembrerebbe essere affermativa, in Spagna bisognerebbe incentivare questo utilizzo poiché le scuole beneficiano dei dati provenienti dalla valutazione del sistema scolastico meno di quanto potrebbero. Per questo motivo si potrebbe

avviare un intervento di sensibilizzazione su questo aspetto, a parere della National Project Manager intervistata.

Inviando alle scuole un report contenente gli esiti. Sono informazioni confidenziali che non vengono diffuse pubblicamente. I risultati delle scuole non vengono pubblicati da nessuna parte. Noi li restituiamo a loro, indichiamo per ciascun istituto qual è il livello di competenze specifico, come si sono distribuiti gli esiti degli studenti, qual è il punteggio degli studenti e quale quello delle studentesse e altre analisi. È un report molto specifico... In un certo qual modo è il nostro "premio" per aver partecipato allo studio. È importante, dal mio punto di vista, che le scuole ne facciano buon uso a seconda delle proprie necessità. Certo, in alcune scuole ho visto che le informazioni sono state successivamente pubblicate sul sito web [...], soprattutto in quei casi in cui i punteggi sono soddisfacenti. Sembra che da parte delle scuole ci sia molto interesse e gli esiti della valutazione vengono utilizzati con beneficio. Inoltre, se qualche esito in qualche area è stato più debole, allora concentriamo la nostra attenzione su quello. (NPM finlandese, traduzione a cura dell'autrice)

In Italia tutto è demandato alla referenzialità e alle competenze di chi dovrebbe utilizzare quei dati.

È anche un discorso di competenze e di saper leggere i dati, saperli interpretare e di saperli utilizzare per introdurre dei miglioramenti. Molti dirigenti lo fanno, molti altri no. Anche qui è lasciato al volere o alle buone intenzioni o anche alle abilità dei singoli dirigenti e docenti. Per esempio, il docente della prova può avere delle informazioni utili, magari si rende conto che in quella determinata abilità i suoi bambini e ragazzi hanno delle difficoltà e quindi può andare a lavorare un po' di più su quello. Banalmente già fare questo vorrebbe dire utilizzare le prove. (NPM italiana)

L'approccio critico e consapevole all'utilizzo dei dati è una delle tematiche emerse dall'analisi delle interviste qualitative. Tenere conto di dati affidabili per prendere decisioni può, infatti, migliorare lo stesso processo decisionale (Miceli, 2004).

3.6. Alcune riflessioni dalle interviste ai NPM dell'indagine OCSE-PISA - valutazione e policy processes

Negli ultimi anni è aumentata la richiesta da parte dei cittadini di servizi pubblici efficaci, efficienti e di qualità, che sappiano andare incontro alle esigenze, quelle educative comprese. Di fronte all'aumento della complessità e della frammentazione, i governi stanno facendo sempre più ricorso a processi decisionali informati da dati concreti. Le evidenze empiriche, infatti, svolgono un ruolo cruciale nella progettazione, implementazione e attuazione di politiche pubbliche migliori. Nonostante la capacità di informare le decisioni pubbliche con le evidenze empiriche raccolte sia parte essenziale di una buona governance pubblica, il collegamento tra le prove scientifiche e i processi decisionali non è sempre semplice, tanto è vero che la mancanza di competenze nella lettura dei dati, come ricordato in chiusura del paragrafo precedente, le lacune istituzionali e la mancanza di un'intermediazione effettiva stanno ostacolando lo sviluppo dell'*evidence informed policy-making* (OECD, 2020g). Per questi motivi nel presente paragrafo verranno sintetizzate le percezioni, riportate dai National Project Manager dell'indagine PISA dell'OCSE durante le interviste, in merito al rapporto tra la valutazione dei sistemi scolastici e la politica, terza tematica emersa nell'analisi dei dati. Come avvenuto nei paragrafi precedenti, per facilitare la contestualizzazione delle opinioni e delle percezioni emerse verranno menzionati alcuni estratti significativi delle interviste.

In Finlandia, a parere del National Project Manager, c'è molta attenzione da parte delle istituzioni centrali nei confronti degli esiti delle valutazioni, in particolare quelli delle indagini standardizzate internazionali.

Il Ministero dell'Educazione e della Cultura segue i risultati delle valutazioni molto attentamente. Per quanto riguarda per esempio PISA, noi realizziamo l'indagine, ma l'incarico lo riceviamo dal ministero che prende tutte le decisioni del caso sullo studio. In un certo senso noi siamo gli esecutori, mentre il ministero guida il processo, elargisce i fondi e pubblica i rapporti. Siamo noi a scrivere i rapporti, ma il Ministero dell'Educazione e della Cultura

li diffonde. Inoltre, vengono organizzate conferenze stampa. Quando viene elaborato il programma del governo... Per esempio, durante l'ultima edizione di PISA... Nell'ultimo governo e nell'ultimo programma era riportato tra gli obiettivi che bisognava migliorare i risultati di PISA. Ora che i risultati sono peggiorati, nel programma del governo è stato riproposto nuovamente di raggiungere un determinato benchmark, bisognerebbe migliorare ancora i punteggi complessivi. L'indagine PISA ha avuto una grande risonanza sia nelle politiche scolastiche che negli obiettivi del ministero. (NPM finlandese, traduzione a cura dell'autrice)

In Spagna sono elaborati dei rapporti informativi, tuttavia, a causa dell'assetto politico-federale si riscontrano difficoltà nell'implementazione delle politiche centrali a livello regionale.

Inviando raccomandazioni, per esempio abbiamo alcuni programmi/campagne per promuovere la lettura nella scuola primaria e in particolar modo nella scuola secondaria. Però, quando cerchiamo di realizzare queste politiche e queste misure direttamente nelle scuole, dobbiamo persuadere le regioni a collaborare e dobbiamo stanziare dei fondi per le regioni per implementare questo tipo di politiche nelle scuole reali. Gli stiamo andando incontro, intendo che stiamo realizzando alcuni programmi sulle discipline STEM, in particolar modo per l'accesso delle ragazze alle discipline STEM. Potremmo fare di più. Potremmo fare di più. Dovremmo spremere i nostri dati di più perchè investiamo molto lavoro nell'intero processo: amministrazioni differenti, codifica, analisi dei dati etc. E poi qual è il feedback che ricevono le scuole? Come possiamo migliorarlo? Dovremmo lavorarci di più. (NPM spagnola, traduzione a cura dell'autrice)

In Italia, invece, la valutazione di sistema potrebbe essere più collegata a interventi di miglioramento centrale, a parere della National Project Manager. Il Ministero dell'Educazione e del Merito, così come nelle altre realtà nazionali prese in esame, richiede, in virtù dell'autonomia scolastica, di realizzare dei monitoraggi e delle valutazioni periodiche sia all'Invalsi, l'istituto di valutazione nazionale, sia all'unità internazionale che assicura la partecipazione italiana alle indagini standardizzate realizzate da organismi sovranazionali quali l'OCSE e l'IEA.

Il ministero richiede queste valutazioni e quindi dovrebbe accoglierle bene. [...] Quanto è collegata a interventi di miglioramento? La valutazione, secondo me, poco. [...] Purtroppo questo è un problema grande in Italia ed è tipicamente italiano perché in molti altri paesi, invece, hanno utilizzato i risultati di PISA, per esempio, per fare delle riforme. La Germania, anche la Polonia. La Polonia ha riformato tutto il sistema educativo sulla base dei risultati. PISA. Ora forse questa è un'esagerazione. Diciamo però che i dati sono stati utilizzati. Quindi sì, volendo si possono utilizzare, cioè si può tenere conto di alcuni aspetti. [...] Ahimè, spesso ti viene anche da chiedere cosa le facciamo a fare le ricerche. Insomma spesso, perché se poi i dati non vengono utilizzati... Potenzialmente, invece, potrebbero essere utilizzati. Danno tanti spunti e anche tante idee, soprattutto quelle che derivano dalla comparazione con altri Paesi o altri sistemi. Non abbastanza, sicuramente i dati non vengono utilizzati abbastanza. (NPM italiana)

Uno degli indicatori che esplicitano l'impatto che la ricerca scientifica ha sulle politiche è quello economico. L'investimento nell'educazione e nella ricerca favorisce la crescita economica, la produttività, il benessere personale e comunitario, inoltre riduce le disuguaglianze sociali. Nel 2019 i paesi OCSE, Italia, Finlandia e Spagna compresi, hanno investito in media il 4.9% del Prodotto Interno Lordo (PIL) per le istituzioni educative dal livello primario, al livello terziario, ricerca compresa (il 3.4% del PIL per l'istruzione primaria e secondaria, 1.5% del PIL per l'istruzione terziaria). La spesa per l'educazione può essere influenzata da molti fattori: a livello primario e secondario influiscono il numero di studenti, la durata del percorso scolastico e l'effettiva assegnazione dei fondi, mentre a livello terziario influiscono i criteri di accesso all'istruzione accademica, il numero di studenti iscritti nei diversi settori di studio e gli investimenti nell'attività di ricerca. In Finlandia nel 2019 il 5.2% del PIL è stato investito nell'educazione, ripartito tra il 3.7% per l'istruzione primaria e secondaria e l'1.5% per l'istruzione terziaria. In Spagna l'investimento complessivo è stato del 4.4% del PIL, il 3.1% per l'istruzione primaria e secondaria e l'1.3% per l'istruzione terziaria. In Italia, invece, la percentuale di PIL investita per l'educazione è stata

globalmente del 3.8%, il 2.9% per l'istruzione primaria e secondaria e lo 0.9% per l'istruzione terziaria (OECD, 2022).

In Finlandia sia le istituzioni scolastiche pubbliche che gli atenei ricevono finanziamenti dallo stato, ma, mentre nelle prime il budget messo a disposizione non dipende dai buoni risultati raggiunti dagli studenti, a livello terziario negli ultimi anni si è andati incontro a un sistema di gestione delle risorse basato sui risultati raggiunti in termini di numero di laureati, progetti di ricerca e pubblicazioni.

Avremmo bisogno di più fondi di base per l'università. Se si è costretti ad andare a cercare i fondi tramite la partecipazione a bandi di ricerca, la ricerca di base potrebbe risentirne. Bisognerebbe incrementare la parte dei finanziamenti base dello Stato e ridurre la quota derivante dalla partecipazione a bandi concorrenziali. In futuro bisognerebbe concentrarsi sulla revisione del sistema di gestione dei fondi per la ricerca. Certo, sono stati fatti dei tagli anche all'istruzione primaria e secondaria. Tagli abbastanza forti durante gli ultimi due governi. Ci sono stati momenti di crisi e anche la scuola ne ha risentito. Quando i risultati di PISA erano ai livelli migliori, all'inizio degli anni 2000, anche i finanziamenti pubblici per l'istruzione erano migliori di quello che sono ora. Qualche relazione potrebbe esserci. Bisognerebbe investire nell'educazione, non tagliare ancora. È chiaro. (NPM finlandese, traduzione a cura dell'autrice)

In Spagna l'INEE, l'istituto di valutazione nazionale, descritto nel secondo capitolo della seconda parte del presente lavoro, riceve fondi dallo Stato per realizzare le attività di cui è responsabile.

Abbiamo un budget, ma è molto ridotto. Nell'INEE, l'istituto in cui svolgo le mie attività, non lavorano molte persone al momento. C'è molta ricerca da fare. A volte siamo costretti ad assumere ricercatori esterni per svolgere alcune analisi o redigere alcuni report perchè non abbiamo abbastanza personale per fare quello che dovremmo fare. (NPM spagnola, traduzione a cura dell'autrice)

In Italia un maggior stanziamento di risorse economiche potrebbe facilitare alcune attività che hanno anche a che fare con la valutazione del sistema scolastico nel suo complesso.

Se avessimo le risorse potremmo fare anche tante più cose per coinvolgere e per informare di più, per far capire meglio che in realtà queste informazioni possono essere utili a tutti. [...] Tornando al discorso dei finanziamenti, noi abbiamo effettivamente un problema grande di finanziamenti. Abbiamo veramente... Abbiamo pochi soldi. Da una parte il ministero ci chiede tante cose. [...] Dipende dallo stato, dipende dal governo. Fondamentalmente dipende da chi c'è al governo in quel momento. E poi? Allora in Italia la ricerca è poco finanziata in generale e dentro la ricerca, la ricerca educativa, è proprio la Cenerentola. (NPM italiana)

I finanziamenti per l'istruzione generalmente non rispondono alle fluttuazioni a breve termine del PIL, bensì a quelle a lungo termine. Il 2020, l'anno dello scoppio della pandemia, ha visto in tutti i paesi OCSE una diminuzione del PIL a seguito della diffusione del COVID-19 e della crisi socio-sanitaria. Dal 2008 al 2020, infatti, la spesa media per l'educazione, dal livello primario a quello terziario, rapportata al PIL di ciascuna nazione è diminuita del 9% in Italia e del 10,8% in Spagna (OECD, 2022). L'impatto della pandemia sui finanziamenti a lungo termine per l'istruzione è ancora incerto, tuttavia tutti i paesi e le economie OCSE hanno attuato misure finanziarie immediate per supportare non solo i sistemi sanitari, ma anche quelli sociali come la scuola. Si è passati da una didattica in presenza a una didattica a distanza in un primo momento. In un secondo momento, quando sono state riaperte le scuole, si è tornati nelle aule che spesso erano mal attrezzate per consentire il rispetto dei requisiti di distanziamento sociale e di sicurezza di studenti, insegnanti, dirigenti scolastici e famiglie. Anche le pratiche valutative, sia esterne che interne alle istituzioni scolastiche, hanno dovuto prendere in considerazione queste limitazioni. Le valutazioni esterne e le visite ispettive, là dove previste, hanno subito una battuta di arresto e le valutazioni standardizzate degli apprendimenti, l'indagine PISA dell'OCSE in primis, sono state posticipate di un anno (Save the Children, 2020; OECD, 2021a, 2022).

Ci sono in ballo così tanti paesi, così tante situazioni diverse nelle scuole che bisogna tenerne conto. Noi siamo riusciti abbastanza bene a realizzare il *field trial* di PISA, anche se alcune scuole hanno dovuto adottare la didattica a

distanza all'inizio della rilevazione. Al momento siamo tornati alla tradizionale didattica in presenza. [...] Senza dubbio ci è voluta più organizzazione e flessibilità: la rilevazione è stata realizzata in momenti diversi, abbiamo allungato la deadline della prova sul campo perchè in alcuni casi è accaduto che gli studenti non erano presenti a scuola. Questa prima rilevazione di PISA, prima del *main study*, è molto vicina temporalmente all'inizio della pandemia, sarà interessante analizzarne i risultati. Forse tra un anno la pandemia apparterrà già alla storia. Ora invece... È qui e ora, dati aggiornatissimi. [...] Sarà interessante anche osservare se ci sono stati dei cedimenti qui o lì. Saremo in grado di monitorare anche le correlazioni, nelle scuole ci possono essere delle differenze, anche all'interno della Finlandia stessa. Sicuramente nel mondo ci sono scuole in cui si è riusciti a continuare con la didattica in presenza e scuole in cui, invece, ci si è dovuti spostare per molto tempo sulla didattica a distanza. Ci sono delle differenze? Come reagiscono le competenze a queste differenze? È molto interessante, si apriranno nuove domande di ricerca. (NPM finlandese, traduzione a cura dell'autrice)

In Spagna, invece, la variabile pandemia è stata presa in considerazione anche nelle ricerche valutative nazionali.

Abbiamo preso in considerazione la pandemia nelle nostre valutazioni nazionali. O meglio, abbiamo aggiunto alcune domande sulla pandemia e sulle differenti misure che le scuole hanno adottato per far fronte a questa situazione. Anche negli studi internazionali è stata presa in considerazione. Stiamo aggiungendo molte domande circa questi aspetti in PISA. È tutto così grande. Sono 30 domande su... Su come le scuole sono riuscite ad organizzarsi nel bel mezzo di queste problematiche. Una delle risposte che abbiamo ricevuto è che hanno combattuto per fermare il prima possibile la didattica digitale a distanza perché, in particolar modo per gli studenti di scuola secondaria, intendo i gradi otto e nove, ma anche il grado sette, è stato terribile per tutte le conseguenze che ciò ha portato. Uno degli aspetti positivi della pandemia è che la società e le famiglie hanno appreso sulla loro pelle che gli insegnanti sono necessari per i bambini, i media digitali non sono abbastanza. [...] Gli studenti stessi non hanno interiorizzato di essere a scuola. Loro pensavano di essere a casa a fare i compiti perchè se non ricevi abbastanza supporto da parte dei genitori... La maggior parte dei genitori hanno capito che devono essere più coinvolti nei processi di apprendimento dei loro figli. Credo che non torneremo alla situazione pre-

pandemica, siamo in una situazione nuova ora. [...] Dobbiamo apprendere da questa terribile esperienza. (NPM spagnola, traduzione a cura dell'autrice)

In una situazione eccezionale come quella pandemica che ha modificato le nostre vite, il contributo della ricerca valutativa per la ripartenza può essere fondamentale.

Quale può essere il contributo per la ripartenza? Anche qui è un contributo sempre conoscitivo, sempre informativo. Tuttavia, non è poco, perché essendo un problema generale, quindi non un problema di pochi, ma un problema di tutti... Valutazioni di questo tipo, indagini di questo tipo sono quello che ci vuole per avere dei punti di partenza. Le informazioni e i dati sono fondamentali per questo e ovviamente questo è il contributo principale che può dare la valutazione per la ripartenza. (NPM italiana)

La crisi socio-sanitaria ed educativa ha messo a nudo le vulnerabilità dei sistemi formativi, tanto in Italia quanto nel resto del mondo. Numerosi studi sono stati condotti durante e dopo le prime fasi della pandemia sui limiti e sulle potenzialità del nuovo assetto. La ricerca, infatti, può configurarsi in questo scenario come impulso per la consapevolezza e l'implementazione di piani di miglioramento all'interno di processi politici e decisionali informati dalle evidenze empiriche e valutative.

Capitolo quarto

Il questionario

4.1. Il questionario ai dirigenti scolastici, verso la triangolazione italiana

Denzin e Lincoln (1994) definiscono gli strumenti che compongono la ricerca qualitativa un bricolage.

Le molteplici metodologie della ricerca qualitativa possono essere viste come un bricolage, e il ricercatore come bricoleur. [...] Il bricoleur produce un bricolage, cioè un insieme compatto di pratiche che forniscono soluzioni a un problema in modo concreto. “La soluzione (bricolage) che è il risultato del metodo del bricoleur è una costruzione [emergente]” (Weinstein & Weinstein, 1991, p. 161) che cambia e prende nuove forme man mano che diversi strumenti, metodi e tecniche vengono aggiunti al puzzle. Nelson et al. (1992) descrivono la metodologia dei cultural studies “come un bricolage. È una scelta di pratiche pragmatiche, strategiche e autoriflessive” (p. 2). Questa comprensione può essere applicata similmente alla ricerca qualitativa. (Denzin e Lincoln, 1994, p. 2, traduzione a cura dell’autrice)

L’approccio del bricolage (Denzin & Lincoln, 1994) è stato applicato nella ricerca dottorale nell’utilizzo di strumenti e tecniche di raccolta dei dati differenti. All’analisi documentale descritta nel capitolo due della seconda parte, si è affiancata la realizzazione delle interviste qualitative semi-strutturate ai National Project Manager dell’indagine PISA dell’OCSE descritte nel capitolo tre della seconda parte e la somministrazione del questionario semi-strutturato ad un gruppo di dirigenti scolastici italiani. Riprendendo Mertens e Hesse-Biber (2013), infatti, l’intento è stato quello di utilizzare strumenti di raccolta dei dati misti per avere una comprensione più approfondita della valutazione dei sistemi scolastici in Italia, Finlandia e Spagna, oggetto di studio del progetto di ricerca qui presentato (Doyle et al., 2009; Trincherò, 2014a; Trincherò & Robasto, 2019): strumenti di raccolta dei dati di varia

natura, infatti, permettono di sostenere la triangolazione e, di conseguenza, la validità della ricerca (Hesse-Biber, 2010).

Se l'intervista rientra tra le tecniche che permettono di entrare in diretto contatto con le persone per indagarne i pensieri, le riflessioni, le esperienze e le percezioni attraverso interazioni sociali, il questionario permette di raggiungerne un numero più cospicuo contemporaneamente con un minore impiego di tempo (Benvenuto, 2015). Nonostante la mancanza di unanimità nella definizione dello studio di caso, nonostante per alcuni sia un disegno di ricerca, per altri uno stile e per altri ancora una strategia o un metodo (Yin, 2018) e nonostante sia stato considerato spesso come una metodologia di ricerca più qualitativa, quindi lontana da strumenti di ricerca come i questionari, tipicamente più quantitativi (Sena, 2021), si è deciso di ricorrere al questionario semi-strutturato per la maneggevolezza nella diffusione e nella somministrazione. Inoltre, seppur non previsto nel disegno iniziale della ricerca, è stato fatto ricorso al questionario, affiancandolo alle interviste, per un duplice motivo. Da un punto di vista metodologico poichè garantisce l'anonimato, dunque fa in modo che i rispondenti, nel caso specifico i dirigenti scolastici italiani, siano meno condizionati da aspetti di desiderabilità sociale, e lascia a questi ultimi, notoriamente sovraccarichi di lavoro, la possibilità di decidere in quale momento compilarlo (Coggi & Ricchiardi 2005; Lucisano & Salerni, 2009). Da un punto di vista scientifico-di disegno di ricerca, invece, poichè, venuta meno la possibilità di interloquire con alcuni contesti nazionali a causa della mancanza di disponibilità da parte dei testimoni privilegiati per il sovraccarico di lavoro contingente (l'edizione 2021 di PISA è stata posticipata al 2022 a causa della pandemia), si è pensato di esplorare con un approccio bottom-up il punto di vista di un gruppo di dirigenti scolastici italiani che nella pratica quotidiana si trovano coinvolti nei processi di valutazione del sistema scolastico, nello specifico di autovalutazione delle istituzioni scolastiche. La rilevazione tramite il questionario si è aggiunta, dunque, in corso d'opera per sopperire al mancato contatto con alcuni National Project Manager dell'indagine PISA dell'OCSE e per approfondire il contesto italiano,

quello più vicino geograficamente a chi scrive, mentre in Italia, così come nel resto del mondo, gli spostamenti fisici erano limitati.

I questionari sono strumenti di raccolta delle informazioni, definiti come un insieme strutturato di domande e relative categorie di risposta definite a priori da chi lo costruisce, ovvero di domande cosiddette “chiuse” dove all’intervistato (inteso come colui che risponde alle domande scritte del questionario) viene richiesto di individuare tra le risposte presentate quella che più si avvicina alla propria posizione, e/o di domande “aperte”, che non prevedono cioè delle risposte predeterminate. (Zammuner, 1996, p. 39)

Tra i limiti del questionario spesso è annoverata la sua “caratterizzazione fortemente oggettivante” (Benvenuto, 2015): è uno strumento che non sempre si presta a scandagliare in profondità realtà specifiche. Per questo motivo, tenendo conto della duplice natura delle risposte che tramite esso possono essere richieste (Zammuner, 1996), sono state costruite sia domande strutturate che domande non strutturate con la possibilità di rispondere in modo aperto, favorendo così un’esplorazione più qualitativa. Inoltre, per andare incontro alle esigenze personali e professionali dei rispondenti, i dirigenti scolastici italiani, è stata lasciata la possibilità di somministrare lo strumento in una modalità duplice, sia in forma scritta come questionario, che in forma orale come intervista. La struttura del questionario-intervista verrà descritta nel paragrafo successivo.

4.2. La struttura del questionario

Lo studio di caso multiplo e comparativo condotto durante il percorso dottorale si è articolato su due livelli (Yin, 2009; Zanazzi 2014; Scheerens, 2019). Come descritto da Zanazzi (2014), il livello internazionale ha previsto un’analisi macro-documentale dei report e delle linee guida messe a punto comunitariamente e riportate nel capitolo due della seconda parte e in apertura del presente lavoro (OECD, 2004, 2013; Eurydice, 2015), il livello nazionale, invece, ha indagato più da vicino nei contesti micro-nazionali le implementazioni di queste indicazioni e le percezioni legate ai processi valutativi,

ascoltando la voce dei National Project Manager dell'indagine PISA dell'OCSE e di un gruppo di dirigenti scolastici italiani. A tal proposito, le interviste qualitative sono state affiancate dal questionario semi-strutturato, somministrato anche telefonicamente, per rispondere in particolar modo alla terza domanda di ricerca (quale supporto è fornito alle scuole nei processi di autovalutazione?). Quest'ultima, che ha previsto anche un'analisi documentale delle pratiche di autovalutazione scolastica così come riportate nelle carte ufficiali (capitolo secondo della seconda parte, paragrafo 2.6), unitamente alla somministrazione del questionario è stata aggiunta in un secondo momento, quando, a seguito della diffusione del virus SARS-CoV-2 e delle restrizioni negli spostamenti fisici, è divenuta obbligatoria la revisione dell'iniziale progetto di ricerca che aveva previsto soggiorni all'estero, nello specifico nei paesi oggetto di studio. Inoltre, come già ricordato, propagatasi la pandemia, anche gli organismi sovranazionali come l'OCSE e l'IEA, punti di riferimento per la valutazione standardizzata degli apprendimenti a livello internazionale, sono stati chiamati a prendere decisioni in merito allo svolgimento delle indagini previste per il biennio 2020-2021. È per questo motivo che il *main study* dell'indagine PISA è stato posticipato di un anno, andando ad aggiungere carico di lavoro sui National Project Manager, i responsabili nazionali, che, di conseguenza, in alcuni casi hanno preferito non dare la propria disponibilità nel prendere parte alla ricerca. Alla luce di questa situazione, si è deciso quindi di rimanere nel contesto nazionale italiano, senza dubbio più facile da raggiungere rispetto a quello internazionale, e di indagare le percezioni e i bisogni formativi dei dirigenti scolastici coinvolti nella loro professione nei processi di autovalutazione.

La scelta è stata quella di utilizzare un questionario semi-strutturato con domande aperte per indagare più in profondità il punto di vista dei rispondenti. Le modalità di somministrazione utilizzate sono state quelle telematiche e telefoniche per la facilità nella diffusione e a causa della situazione contingente. La compilazione manuale del questionario è avvenuta tramite Google Moduli, mentre la somministrazione verbale della traccia, la medesima del questionario, è avvenuta telefonicamente per chi lo ha richiesto tramite mail.

Preliminarmente alla costruzione del questionario, è stata svolta l'analisi della letteratura sull'ambito indagato e si è dialogato informalmente con un gruppo di dirigenti scolastici incontrati ad un convegno circa il loro punto di vista sulla valutazione delle singole istituzioni scolastiche e sui bisogni formativi percepiti. Le informazioni qualitative raccolte in quella situazione informale e spontanea hanno rappresentato il punto iniziale per l'elaborazione degli item delineando le aree da indagare attraverso lo strumento. Così, dopo aver definito le tematiche e gli obiettivi conoscitivi, esplorativi e descrittivi dello strumento, si è passati alla predisposizione delle domande, alla scelta delle modalità di risposta e all'elaborazione della lettera di accompagnamento contenente la dichiarazione di assenso del rispondente per la raccolta e l'analisi dei dati, riportata di seguito nella tabella 6.

Gentile Dirigente Scolastico,

sono Eleonora Mattarelli, dottoranda di ricerca presso la scuola di Dottorato in Psicologia Sociale, dello Sviluppo e Ricerca Educativa della Sapienza, Università di Roma. Nell'ambito di un progetto di ricerca dottorale comparativo-internazionale sto conducendo un'indagine sull'autovalutazione delle scuole, supervisionata dal professore xxx.

Dopo aver preso in considerazione le pratiche e le percezioni relative alla valutazione del sistema scolastico in tre importanti realtà educative europee quali Italia, Finlandia e Spagna, l'obiettivo ora è quello di indagare la visione dei Dirigenti Scolastici Italiani dell'autovalutazione scolastica, attraverso un questionario-intervista.

Il Suo punto di vista è importante per comprendere i processi di autovalutazione attuati nella scuola che dirige. Pertanto, è importante rispondere con la massima sincerità facendo riferimento alla sua scuola e alla sua esperienza personale.

Tutti i dati verranno raccolti in modo anonimo, nel rispetto del Regolamento generale UE sulla protezione dei dati (General Data Protection Regulation, GDPR) n. 679/2016, e al D.Lgs. n. 196/2003 "Codice in materia di protezione dei dati personali" (modificato dal D.Lgs. n. 101 del 10/08/2018). I dati saranno trattati in forma aggregata. Nessun questionario-intervista sarà riconducibile ai singoli partecipanti alla ricerca. Le analisi dei dati hanno uno scopo conoscitivo

rispetto al tema della ricerca e non hanno alcuno scopo diagnostico o valutativo. I risultati del questionario-intervista verranno comunicati e pubblicati in sedi scientifiche, garantendo comunque la non riconoscibilità personale di chi ha partecipato alla ricerca.

Se è interessato/a a rimanere informato/a sugli esiti della ricerca può mettersi in contatto con me in qualità di Referente per il progetto, ed è anche possibile richiedere informazioni e porre qualsiasi domanda concernente l'indagine: dott.ssa Eleonora Mattarelli (xxx@xxx.it).

La ringrazio per la collaborazione.

Dott.ssa Eleonora Mattarelli

Cliccando sul tasto "acconsento" che segue attesta che:

1. Ha ricevuto e letto le informazioni sullo scopo e le procedure della ricerca;
2. Dà il consenso a partecipare alla ricerca.

Tabella 6. Lettera introduttiva al questionario-intervista somministrato a un gruppo di dirigenti scolastici.

Dopo una prima stesura del questionario, si è passati a un try-out qualitativo dello strumento. Attraverso la tecnica del *thinking aloud* il questionario è stato analizzato con un collega dottorando, dirigente scolastico di un liceo, e con il gruppo dei suoi più stretti collaboratori (Payne, 1994; Boren & Ramey, 2000; Clemmensen et al., 2008). Gli item sono stati letti, sono state discusse le possibili interpretazioni di alcuni di essi, è stata verificata la comprensibilità del linguaggio utilizzato. Inoltre, è stato chiesto di esprimere un parere complessivo sul contenuto dello strumento, considerando la possibilità di aggiungere, eliminare o modificare gli item. Sono state analizzate anche le modalità di risposta più appropriate per alcuni di essi. Grazie alle osservazioni raccolte, la prima versione dello strumento è stata rivista: sono stati eliminati due item ridondanti, ne sono stati aggiunti tre di approfondimento e per quattro item è stata modificata la modalità di risposta, da aperta a chiusa e/o viceversa.

La versione definitiva del questionario, per facilitare la conoscenza del background di riferimento dei rispondenti e per facilitare le analisi successive, è stata suddivisa in due sezioni a cui è stata aggiunta

un'ultima domanda aperta di commenti e suggestioni personali (Zanazzi, 2014): la sezione introduttiva di anagrafica e la seconda di approfondimento. La durata della compilazione è stata stimata in trenta minuti circa.

Nella prima sezione sono state inserite domande relative all'età, al genere, alla provincia di residenza e all'istituto diretto. Si tratta di informazioni che definiscono le variabili di sfondo e socio-ambientali (Cecalupo, 2021). Consapevoli del sovraccarico di lavoro dei dirigenti scolastici, si è deciso di rendere breve la prima sezione per approfondire soprattutto la seconda. Gli item anagrafici sono riportati di seguito nella tabella 7.

A1. Età (in cifre)	...
A2. Genere	Donna Uomo Preferisco non specificarlo
A3. Provincia di residenza	...
A4. L'istituto in cui attualmente presta servizio è	Istituto comprensivo Scuola primaria Scuola secondaria di I grado Scuola secondaria di II grado Centro Provinciale per l'Istruzione degli Adulti (CPIA) Altro
A5. Provincia di riferimento dell'istituto	...
A6. Nell'istituto in cui attualmente presta servizio è	Titolare dell'incarico Reggente
A7. Da quanti anni dirige l'istituto in cui attualmente presta servizio?	...
A8. Da quanti anni ricopre l'incarico di	...

dirigente scolastico?	
-----------------------	--

Tabella 7. Sezione anagrafica questionario dirigenti scolastici.

Nella seconda sezione sono state operazionalizzate le questioni emerse tramite il confronto con il gruppo di dirigenti scolastici incontrati informalmente a un convegno. Inoltre, è stata effettuata una revisione della letteratura già esistente composta da report, riflessioni e ricerche svolte sulla figura del dirigente scolastico nei processi di autovalutazione (McBeath, 1999; Barzanò, 2000; Schildkamp, 2007; Schildkamp & Visscher, 2009; Vanhoof et alii, 2009; Castoldi, 2011; Moretti, 2011; Epifani et alii, 2016; Muzzioli et alii, 2016; Giampietro et alii, 2016; Sette et alii, 2018; Benvenuto et alii, 2019; Tomba, 2019a; De Santis & Asquini, 2020). Gli item risultanti dal lavoro di ascolto dei dirigenti scolastici, di analisi della letteratura e di studio di ricerche simili sono riportati di seguito nella tabella 8. La traccia integrale è riportata anche nell'Appendice II.

B1. Secondo lei i processi di autovalutazione d'istituto es. compilazione del Rapporto di Autovalutazione, così come riportati nel Sistema Nazionale di Valutazione, sono sostenibili in relazione alle risorse umane e finanziarie disponibili e alle tempistiche richieste?	Si No Soltanto in parte
B1.1. Se, secondo lei, i processi di autovalutazione d'istituto sono sostenibili soltanto in parte, può spiegarne il motivo?	...
B2. Nell'istituto che dirige quale organizzazione è stata messa a punto per la redazione del Rapporto di Autovalutazione (es. chi elabora, chi approva, chi pubblica il RAV)?	...

Il questionario

B3. La compilazione del Rapporto di Autovalutazione è vissuto da lei come	Un adempimento formale Uno strumento funzionale per la riflessione e il miglioramento della scuola Entrambe le precedenti opzioni di risposta Altro
B4. Quanto, secondo lei, gli indicatori proposti nel Rapporto di Autovalutazione sono adeguati per descrivere la situazione del singolo e specifico contesto scolastico?	1 Per nulla 2 3 4 5 Moltissimo
B5. Quanto, secondo lei, i processi di autovalutazione d'istituto e la compilazione del Rapporto di Autovalutazione sono utili nell'ottica del miglioramento della scuola?	1 Per nulla 2 3 4 5 Moltissimo
B5.1. Può motivare la risposta alla domanda precedente?	...
B6. Nell'istituto che dirige si utilizzano altre forme non strutturate di autovalutazione d'istituto?	Sì No
B6.1. Se sì, come sono organizzate?	Consigli di classe Consigli di istituto Dipartimenti e aree disciplinari Amministrazione Altro
B7. Nel corso della sua carriera ha mai partecipato a momenti di formazione sull'autovalutazione d'istituto?	Sì No
B8. Sente di voler essere maggiormente formato sui temi dell'autovalutazione d'istituto?	Sì No

B8.1. Se sì, secondo lei quali argomenti sono più urgenti e utili?	...
B9. Gli esiti dei processi di autovalutazione d'istituto vengono diffusi e condivisi nella sua scuola (tra docenti, genitori e studenti)?	Si No
B9.1. Se sì, in che modo?	...
B10. Quando si trova a prendere decisioni per la sua scuola, su queste ultime influiscono gli esiti dei processi di autovalutazione?	1 Per nulla 2 3 4 5 Moltissimo
B10.1. Se prende decisioni per la sua scuola anche sulla base degli esiti dei processi di autovalutazione, può argomentare meglio la sua risposta fornendo degli esempi concreti?	...
B11. Quanto, secondo lei, i processi di autovalutazione nella sua scuola sono efficaci ed efficienti (es. hanno delle ricadute concrete)?	1 Per nulla 2 3 4 5 Moltissimo
B11.1. Se hanno delle ricadute concrete, può riportare un esempio?	...
B12. Se vuole aggiungere qualcos'altro sulla sua personale percezione o esperienza rispetto all'autovalutazione d'istituto, può scrivere liberamente nello spazio sottostante.	...

Tabella 8. Sezione finale questionario dirigenti scolastici.

4.3. I risultati del questionario - le variabili di fondo

Il questionario è stato somministrato ai dirigenti scolastici italiani in due modalità, manuale-digitale e telefonica, in base al modo preferito da ciascun rispondente. La versione digitale è stata diffusa attraverso Google Moduli, ha permesso di abbattere i costi e di suddividere graficamente gli item nelle due sezioni, preceduti da una schermata introduttiva di spiegazione del progetto di ricerca e di descrizione della privacy (Cecalupo, 2021). Tutti gli item, ad eccezione di quelli di approfondimento concatenati gli uni agli altri, ovvero dipendenti in larga parte dalla risposta fornita alla domanda precedente (es. item B1.1, B5.1, B6.1 etc.), sono stati contrassegnati da un asterisco rosso indicante l'obbligatorietà della risposta per evitare la perdita di dati preziosi. La versione telefonica, invece, è stata somministrata verbalmente e audioregistrata. All'inizio della telefonata chi scrive ha letto il consenso informato al trattamento dei dati, dopodiché si è passati alla raccolta delle informazioni. Quest'ultima versione, seppur scelta da un numero molto esiguo di partecipanti, ha permesso di scendere più in profondità nelle percezioni dei rispondenti che, grazie all'interazione sociale che si crea durante un colloquio-intervista (Zammuner, 1996), hanno argomentato i propri punti di vista.

La diffusione del questionario, con i contatti di chi scrive, è avvenuta tramite canali social e telematici. Il link è stato condiviso in pagine Facebook e nelle chat di Whatsapp di dirigenti scolastici, nonché diffuso attraverso passaparola e conoscenze. Per questo motivo il gruppo di dirigenti scolastici partecipanti non è da considerarsi come un campione rappresentativo poichè coloro che si iscrivono ai gruppi social e partecipano alle chat di messaggistica istantanea rappresentano nella maggior parte dei casi un gruppo auto-selezionato (Vitalini, 2010; Lombi, 2015).

La somministrazione è iniziata nei primi giorni del mese di agosto 2022 ed è terminata nei primi giorni del mese di ottobre 2022, periodo di passaggio tra la ripresa delle attività per i dirigenti scolastici e l'inizio del nuovo anno. In questo paragrafo e nel successivo verranno presentate le analisi dei dati di sfondo per delineare le caratteristiche di

chi ha compilato il questionario o ha risposto all'intervista telefonica. Per le analisi dei dati è stato utilizzato il programma statistico IBM SPSS nella versione 26 e il pacchetto Office, nello specifico l'applicazione Microsoft 365 Excel.

I dirigenti scolastici che hanno risposto al questionario sono 69, coloro che hanno preferito l'intervista telefonica sono 3, due dirigenti scolastici della provincia di Roma e un dirigente scolastico della provincia di Benevento. Per la partecipazione non vi erano particolari requisiti, infatti l'anzianità di servizio dei rispondenti si distribuisce da 0 anni (dirigente scolastico appena entrato in servizio) a 28 anni. Nel grafico 1 sono sintetizzate le risposte dei dirigenti scolastici circa gli anni di dirigenza. A questi ultimi è stata data la possibilità di inserire manualmente il valore in cifre, nelle analisi successive i valori sono stati raggruppati per fasce di anzianità visibili sull'asse delle ascisse del grafico. L'81.60% dei dirigenti scolastici rispondenti non ricopre l'incarico da più di 10 anni.

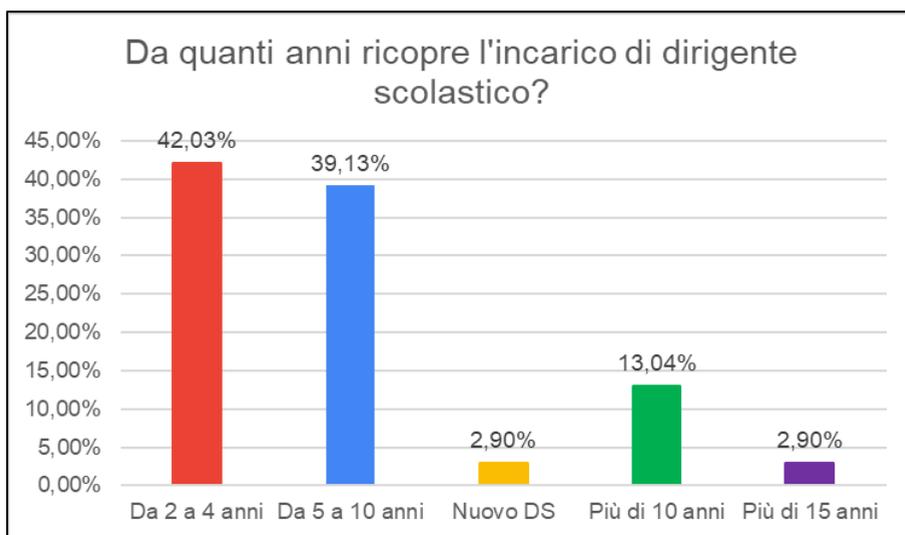


Grafico 1. Anzianità di servizio dei partecipanti. Fonte: elaborazione propria.

Prendendo in considerazione i gradi scolastici, i rispondenti si distribuiscono al loro interno come riportato nel grafico 2.

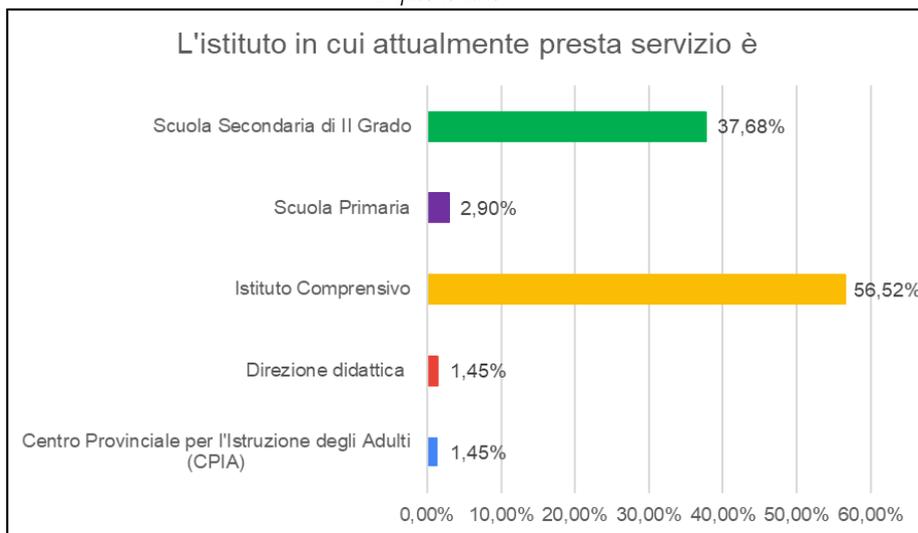


Grafico 2. Grado scolastico di riferimento dei partecipanti. Fonte: elaborazione propria.

Tra i rispondenti nessun dirigente scolastico presta servizio nella scuola secondaria di primo grado, ricompresa nella maggior parte dei casi negli istituti comprensivi, mentre il 5,80% di loro presta servizio presso scuole primarie o circoli didattici, ovvero sedi scolastiche di uno stesso territorio che possono comprendere scuole dell'infanzia e scuole primarie, o centri provinciali per l'istruzione degli adulti.

Il genere, tolto il caso specifico di un rispondente che ha preferito non indicarlo, è rappresentato nell'82,61% da donne e nel 15,94% da uomini, dato in linea con i risultati di TALIS 2018, l'indagine svolta con cadenza quinquennale sulla professionalità dei docenti e dei dirigenti scolastici dall'OCSE (OECD, 2019d, 2020e). Le donne, infatti, rappresentano la maggior parte degli insegnanti in tutti i paesi partecipanti allo studio, eccetto il Giappone, mentre rappresentano la minoranza nella dirigenza scolastica nella metà dei paesi considerati. L'Italia, se da una parte appartiene a quei paesi in cui la professionalità docente è scelta perlopiù da donne, dall'altra rientra anche tra quei paesi in cui più del 60% dei dirigenti scolastici sono donne insieme a Brasile, Bulgaria, Ungheria, Lettonia, Romania, Russia, Slovenia e Svezia. In Spagna poco meno del 50% dei dirigenti scolastici sono donne, mentre le insegnanti donne sono

più del 60%. Al contrario, in Finlandia meno della metà dei dirigenti scolastici sono donne, mentre le insegnanti donne sono circa il 70% (OECD, 2019d).

L'età media dei rispondenti è di 54,5 anni. Nel 2018 l'età media dei dirigenti scolastici nei paesi OCSE partecipanti all'indagine TALIS era di 52 anni, otto anni in più rispetto all'età media degli insegnanti. Per l'Italia l'età media dei dirigenti scolastici era di 56 o più anni, superiore a quella dei rispondenti al questionario. Ciò dipende in larga misura dal fatto che i dirigenti scolastici, in Italia e nel resto del mondo, sono spesso selezionati tra coloro che sono già insegnanti: per diventare dirigente scolastico bisogna avere un certo numero di anni di esperienza pregressa, una formazione specifica e, nel caso del nostro paese, passare attraverso un concorso pubblico. L'Italia, secondo le statistiche internazionali, è anche tra quei paesi che detengono il primato di dirigenti scolastici prossimi al pensionamento (60 o più anni). In questo senso, l'invecchiamento della popolazione dei dirigenti scolastici dovrebbe spingere i decisori politici in ambito educativo a formare e promuovere una nuova generazione di dirigenti scolastici o a riorganizzare la dirigenza delle istituzioni scolastiche, per esempio con la creazione di reti di scuole dirette da un solo dirigente scolastico, una reggenza di istituzioni scolastiche (ovvero prive di dirigente scolastico titolare) messa però a sistema - nel gruppo dei rispondenti nessun dirigente scolastico ha risposto di essere reggente, tutti sono titolari dell'incarico (OECD, 2019d, 2020e).

Nel questionario è stato chiesto di indicare anche la provincia di riferimento dell'istituto in cui è prestato servizio. Durante la fase di analisi le province sono state raggruppate in macroaree, così come definite dall'ISTAT. Per facilitare le analisi successive, le macroaree Nord-Ovest e Nord-Est sono state unite in un'unica macroarea. Il grafico 3 riporta la distribuzione dei dirigenti scolastici per macroarea.

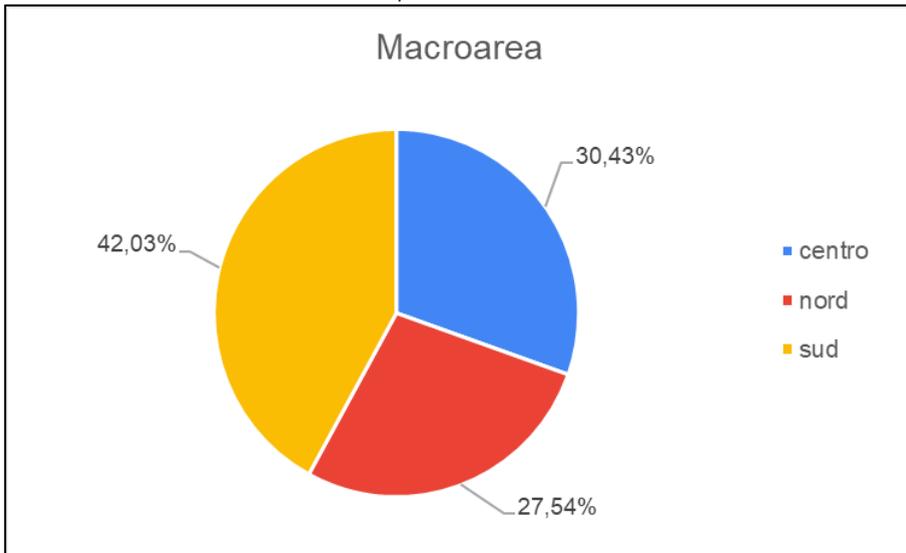


Grafico 3. Distribuzione dei rispondenti per macroarea dell'istituto di provenienza. Fonte: elaborazione propria.

Nel gruppo dei dirigenti scolastici in esame la macroarea più rappresentata è quella del sud, seguita dal centro e dal nord. Ai rispondenti è stato chiesto di indicare, oltre alla provincia di riferimento dell'istituto scolastico diretto, anche la provincia di residenza. L'intento è stato quello di indagare il fenomeno migratorio nelle professioni educative che spingono sempre più professionisti nell'ambito scolastico e nell'ambito educativo in generale a spostarsi, nella migliore delle ipotesi all'interno della stessa regione per cercare una stabilità lavorativa, nella peggiore delle ipotesi da una parte all'altra del paese (Colagiovanni, 2003; Colucci & Gallo, 2017). Il 17,39% dei rispondenti ha riportato casi di mobilità interprovinciale o interregionale. Esempi concreti sono dirigenti scolastici che risiedono a Napoli, ma prestano servizio a Roma o di coloro che risiedono a Pordenone, ma esercitano a Udine. Vi è anche chi ha la residenza a Caltanissetta, Ragusa o Viterbo e presta servizio a Massa Carrara, Monza-Brianza o Genova.

4.4. I risultati del questionario - l'autovalutazione delle istituzioni scolastiche

Dopo aver delineato le caratteristiche dei rispondenti, in questo paragrafo verranno riportate le evidenze emerse dalla seconda sezione del questionario, quella riguardante le pratiche e le percezioni dei dirigenti scolastici italiani legate all'autovalutazione delle scuole. Come descritto nei paragrafi precedenti, il questionario è stato suddiviso in due parti. Nella seconda sono state inserite alcune domande filtro che, in base alla risposta fornita, hanno richiesto un ulteriore approfondimento da parte del rispondente. La modalità di risposta prescelta per l'approfondimento è stata quella aperta così da permettere l'argomentazione del proprio punto di vista e la raccolta di informazioni preziose sull'organizzazione delle pratiche di autovalutazione e sulle differenti esperienze.

L'analisi delle risposte aperte è avvenuta attraverso l'individuazione di macrocategorie e microcategorie con un approccio di tipo bottom-up, facendo riferimento alla flessibilità del *framework method* (Braun & Clarke, 2006; Gale et alii, 2013). Le categorie sono emerse tramite la lettura e la rilettura delle evidenze fornite dai dirigenti scolastici italiani (Batini et alii, 2020). Lo sforzo è stato quello di individuare categorie neutre che non fossero influenzate dai preconcetti di chi scrive, per questo il confronto costante con altri colleghi è stato prezioso. A una prima lettura delle risposte disponibili, è seguita l'individuazione delle tematiche ricorrenti che, raggruppate in base ai nuclei tematici, hanno permesso di delineare le categorie di riferimento. Come evidenziato da Batini e colleghi (2020) per la costruzione del modello di categorie è stato applicato il criterio dell'eshaustività e della non sovrapposibilità. Tuttavia, visto il corpus di risposte non particolarmente esteso, si è cercato di mantenere un numero adeguato di categorie per prendere in considerazione tutte le tematiche emerse senza però disperdere le informazioni.

La prima domanda della seconda sezione del questionario, quella sulle pratiche e sulle percezioni legate all'autovalutazione, chiedeva ai rispondenti di segnalare se i processi di autovalutazione, così come

indicati nel Sistema Nazionale di Valutazione, siano sostenibili in relazione alle risorse umane e finanziarie e alle tempistiche richieste. Il 21.74% dei dirigenti scolastici ha risposto che non sono sostenibili, il 39.13% che sono sostenibili e il 39.13% sostenibili solamente in parte. Tra questi ultimi, in egual numero di chi ha risposto positivamente, la non completa sostenibilità dei processi di autovalutazione è dovuta da una parte a *problematiche organizzative*, dall'altra a *carenze strutturali*. Per la prima categoria individuata, *problematiche organizzative*, sono state rilevate 21 occorrenze, per la seconda, *carenze strutturali*, 18 occorrenze (il numero delle occorrenze coincide con i rispondenti alla domanda di approfondimento che non rappresentano la totalità dei rispondenti al questionario). Tra le problematiche organizzative riportate dai dirigenti scolastici ci sono quelle legate alle tempistiche troppo stringenti, a un carico di lavoro eccessivo alla luce delle attività da seguire e monitorare nella quotidianità e a uno scarso coinvolgimento da parte del personale della scuola. Al contrario, tra le carenze strutturali si annoverano la mancanza di risorse finanziarie e umane dedicate ai processi di autovalutazione di istituto e la mancanza di formazione delle funzioni strumentali che gestiscono questi processi. Quest'ultimo aspetto è disconfermato dalle parole di una delle dirigenti scolastiche che hanno risposto all'intervista telefonica.

Si tratta di capire che cosa si va a fare all'interno del RAV. Se il dirigente è consapevole bastano i tempi e bastano le risorse. Io ho formato un gruppo di persone che collaborano con me in questo lavoro che è il nucleo interno di valutazione. Sono diventate molto brave quindi riusciamo a fare un ottimo lavoro nei tempi necessari ed è utilissimo per la progettazione delle attività dell'istituto. (dirigente scolastico della provincia di Roma)

L'organizzazione dei processi di autovalutazione di istituto è un aspetto delicato nell'ambito del coordinamento interno delle scuole.

Alle istituzioni scolastiche si richiederà di emanciparsi dalla cultura burocratica per proiettarsi in una cultura organizzativa nuova, che veda la logica dell'adempimento formale progressivamente sostituita dalla logica della collaborazione, delle intese, della responsabilità. Saranno, perciò, chiamate a

pensare e a gestire progetti secondo modalità soggettive, specifiche di ogni istituto autonomo, dalle quali dovrebbero derivare il successo formativo degli alunni e il conseguimento degli obiettivi che l'istituzione scolastica persegue. La qualità di un sistema scolastico si misura, infatti, sia dal tipo di interventi posti in essere per realizzare i processi che porteranno ai risultati attesi, sia, anche, da quelle azioni che riguardano la gestione adeguata del quotidiano e dell'ordinario. (Serra, 2000, p. 7)

Alla luce di quanto evidenziato da Serra (2000), ai rispondenti è stato chiesto in che modo viene vissuta la compilazione del Rapporto di Autovalutazione. La tabella 9 riporta una sintesi delle risposte.

Un adempimento formale	11.76%
Uno strumento funzionale per la riflessione e il miglioramento della scuola	47.06%
Entrambe le precedenti opzioni di risposta	41.18%

Tabella 9. Percezioni legate alla compilazione del RAV da parte dei dirigenti scolastici. Fonte: elaborazione propria.

Della stessa opinione del 47.06% dei rispondenti è una delle dirigenti scolastiche della provincia di Roma intervistate telefonicamente.

Sicuramente è essenziale per la gestione, perché, se non faccio una autovalutazione, come faccio a progettare e a capire quali sono i punti di forza e di debolezza e in che modo devo, di conseguenza, adoperare le risorse finanziarie? La scuola è una azienda, se non si fa il bilancio, la pianificazione non si può fare. (dirigente scolastico della provincia di Roma)

Per quanto riguarda, invece, la suddivisione dei compiti per la compilazione del Rapporto di Autovalutazione, dalle risposte dei dirigenti scolastici emerge un quadro variegato: in alcuni istituti è responsabilità dei gruppi di lavoro con compiti predefiniti come i Nuclei Interni di Valutazione, lo staff di presidenza, la commissione PTOF, in altri è il dirigente in prima persona ad occuparsene o delega alla funzione strumentale responsabile dell'autovalutazione. Spesso è il

collegio dei docenti ad approvare il lavoro svolto, mentre a rendere visibile esternamente il RAV può essere il dirigente o qualche collaboratore tecnico-amministrativo da lui/lei delegato. Viceversa, è rilevata più unanimità nell'utilizzo degli indicatori del RAV. Il 43.48% dei rispondenti è concorde nell'affermare che gli indicatori proposti nel Rapporto di Autovalutazione sono molto adeguati per descrivere la situazione del singolo e specifico contesto scolastico, mentre solamente l'1.45% dei rispondenti li ritiene per nulla adeguati (tabella 10).

1 – per nulla	1.45%
2	10.14%
3	40.58%
4	43.48%
5 - moltissimo	4.35%

Tabella 10. Grado di accordo dei dirigenti scolastici circa l'adeguatezza degli indicatori proposti nel RAV. Fonte: elaborazione propria.

Il legame tra le pratiche di autovalutazione di istituto e la spinta al miglioramento è una delle tematiche più dibattute negli ultimi decenni (Paletta, 2005; Corsini, 2010; Domenici, 2012; Muzzioli et al, 2016; Perla & Vinci, 2016; Epifani et al, 2016). Il 28.99% dei rispondenti ritiene che l'autovalutazione di istituto e la compilazione del RAV non sono né utili né inutili per il miglioramento, mentre il 49.28%, circa la metà, vede questo strumento e questi momenti di riflessione come abbastanza utili per l'avanzamento delle pratiche didattiche (il 14.49% li ritiene per nulla o poco utili, il 7.25% utilissimi). La media delle risposte (3,48) date da tutti i rispondenti alla scala Likert a cinque passi proposta nell'item (*Quanto, secondo lei, i processi di autovalutazione d'istituto e la compilazione del Rapporto di Autovalutazione sono utili nell'ottica del miglioramento della scuola?*) non è significativamente differente da quella data dai dirigenti scolastici con meno di 5 anni di esperienza (3,46) o da quelli con più di 5 anni di esperienza (3,47), da ciò si può dedurre che l'anzianità di servizio

nel gruppo di dirigenti scolastici in questione non rappresenta una discriminante per la percezione dell'utilità o della non utilità dell'autovalutazione d'istituto. Nella domanda aperta sull'utilità/inutilità dei processi di autovalutazione di istituto e della compilazione del RAV sono state rilevate 62 occorrenze, 39 afferenti alla macrocategoria *utile* e 23 afferenti alla macrocategoria *non utile*. Significativi sono le microcategorie rubricate nella macrocategoria *non utile*: i dirigenti scolastici riportano che le pratiche di autovalutazione di istituto non sono utili per il miglioramento della scuola a causa, ancora una volta, dello scarso coinvolgimento del personale (*Basso coinvolgimento, nonostante gli sforzi "divulgativi"*), dell'assenza di ricadute pratiche (*Servirebbero azioni più concrete e meno teoriche*) e delle criticità gestionali, come la mancanza di tempo (*La quantità infinita di incombenze burocratiche rende la compilazione del RAV una formalità da sbrigare in fretta*). Al contrario nella macrocategoria *utile* sono rubricate le microcategorie legate all'acquisizione della consapevolezza e della riflessione critica (27 occorrenze) e all'autovalutazione come volano per il miglioramento (12 occorrenze). Dello stesso avviso è la dirigente scolastica di un liceo della provincia di Roma che ha risposto all'intervista telefonica.

Non so cosa hanno risposto i miei colleghi, comunque un buon dirigente se non compila il RAV e quindi non va a fare un lavoro di autovalutazione di quello che ha fatto e di ciò che ancora deve fare, di ciò che è andato bene e di ciò che non è andato bene, poi i motivi possono essere vari, non può fare un piano di miglioramento perchè non ha chiaro che cosa deve fare e ciò non si può fare ad occhio. Inoltre, il piano di miglioramento va declinato nel POF che poi diventa il piano finanziario che poi diventa il piano delle attività dei docenti, quindi sono tutti legati. Uno è legato all'altro, è fondamentale e propedeutico. (dirigente scolastico della provincia di Roma)

I bisogni formativi circa la valutazione del sistema scolastico e nello specifico l'autovalutazione delle istituzioni scolastiche rappresentano un altro aspetto affrontato nel questionario attraverso una sottosezione apposita. La formazione necessaria al dirigente scolastico si è andata delineando negli ultimi anni di pari passo con il dibattito sull'autonomia scolastica. Alla leadership scolastica, infatti, sempre più spesso è

richiesto un nucleo di competenze da poter mettere al servizio della scuola per la buona riuscita dei processi di autovalutazione (Robasto, 2018; Tomba, 2019b; Notti, 2020). La quasi totalità dei rispondenti al questionario (91.30%) afferma di aver preso parte a occasioni di formazione sull'autovalutazione di istituto (l'8.70% afferma di non aver mai partecipato a momenti di formazione simili). Il 64.71% afferma anche che vorrebbe essere maggiormente formato su queste tematiche a fronte del 35.29% che, invece, non ritiene proficua un'ulteriore formazione. Solamente uno dei rispondenti ha specificato nel campo libero lasciato appositamente nel form di Google Moduli di non volere essere propriamente formato, ma di volersi avvalere, se fosse possibile, di un servizio di consulenza. Gli argomenti ritenuti più urgenti si concentrano intorno al polo della *gestione dei processi*, prima macrocategoria individuata, e al polo della *formazione contenutistica*, seconda macrocategoria individuata nelle risposte. Nella gestione dei processi rientrano competenze quali la condivisione dei momenti di riflessione, il *middle management*, il *problem solving* e la programmazione, il coinvolgimento del personale della scuola e il miglioramento delle strategie di diffusione e di motivazione dei docenti. Nell'area dei contenuti da approfondire, invece, rientrano aspetti legati all'autovalutazione di dirigenti scolastici e insegnanti, alla rendicontazione sociale, alla valutazione standardizzata degli apprendimenti, all'analisi degli indicatori, alle tecniche di raccolta dei dati e alla costruzione di buoni piani di miglioramento. A proposito della formazione, uno dei dirigenti scolastici intervistati della provincia di Roma è anche parte dei nuclei esterni di valutazione che si recano nelle scuole per le visite esterne, come descritto nel capitolo due della seconda parte del presente lavoro, mentre un altro durante l'intervista telefonica ha riportato di aver predisposto un piano di formazione per i docenti della scuola in cui attualmente presta servizio.

Assolutamente sì. Quando ero docente, come vicepresidente, abbiamo fatto formazione. La formazione è stata fatta dal dirigente perchè per il primo RAV che compilammo fu necessario un lavoro di staff molto forte dal momento che non riuscivamo a capire di cosa si trattasse. Successivamente la grande formazione l'ho avuta dall'Invalsi quando è stata fatta la selezione per entrare

nei nuclei esterni di valutazione. Chiaramente lì c'è stata una formazione massiccia, anche in situazione quando siamo andati a visitare le scuole con l'ispettore. Ero capitata con un ispettore veramente eccellente il quale ci ha formato, mentre ci faceva capire in che modo lavorare nello specifico, è stato molto bravo. Da lui ho imparato tantissime cose, come leggere i dati, quali dati, leggerli nel contesto, perchè un dato negativo dipende dal contesto dal quale proviene. Questo è stato per me un tesoro. In questo modo ho imparato anche a fare il dirigente in un certo modo. La formazione l'ho fatta come docente quando ero parte del NIV e come parte del NEV. [...] Vorrei dire a tutti di seguire il mio percorso perchè mi ha aperto davvero un mondo, però, devo dire che durante questa settimana ho partecipato a due nuovi corsi di formazione su autovalutazione e rendicontazione sociale, nonostante io faccia formazione, io stessa sia formatrice nelle scuole su queste tematiche. Secondo me bisogna sempre fare formazione perchè è circolare, no?! Un po' impari e un po' insegni. (dirigente scolastico della provincia di Roma)

Lo avevo già fatto nella scuola tre anni fa e lo farò anche adesso in questa nuova, già lo ho proposto informalmente al collegio poi lo approverò: una formazione con un esperto di valutazione di sistema che nel caso specifico di tre anni fa fu un professore dell'Università di Salerno. Lo ho contattato anche ora per fare formazione ai miei docenti e mi ha dato la sua disponibilità nei mesi di ottobre e novembre. Credo sia importante (dirigente scolastico della provincia di Salerno)

La diffusione degli esiti della valutazione rappresenta non solo uno dei bisogni formativi espressi dal gruppo di dirigenti scolastici in esame, ma anche un approfondimento che è stato loro richiesto nella compilazione del questionario. Il 7.25% dei rispondenti ha affermato che nella propria scuola gli esiti dei processi di autovalutazione di istituto non vengono diffusi tra docenti, famiglie e studenti. Nel restante 92.75% dei casi la condivisione tra gli stakeholders esiste, ma è tutt'altro che sistematica. Dalle risposte date alla domanda aperta di dettaglio, infatti, si evince che esistono principalmente due modalità di divulgazione degli esiti, *esterna* e *interna* (macrocategorie che hanno guidato l'analisi), che possono essere tra loro mescolate, ma che non esiste un iter di diffusione univoco che accomuna la maggioranza. In alcuni istituti la divulgazione avviene tramite collegio di istituto/collegio dei docenti

(Attraverso report analitico al collegio ed al consiglio d'istituto), in altri tramite la redazione di report e circolari ad opera del dirigente scolastico, delle funzioni strumentali o del nucleo interno di valutazione (A livello di docenti c'è un report annuale della figura strumentale, a nome dell'intero gruppo di lavoro; Attraverso assemblee o circolari). In altri istituti gli esiti dei processi di autovalutazione sono restituiti alle famiglie tramite assemblee e consigli di classe (Con il collegio docenti nelle sedute ordinarie, con il consiglio di istituto nelle sedute dedicate. Con i rappresentanti delle famiglie nelle occasioni possibili), tramite il sito istituzionale della scuola o su altri portali (Pubblicati su sito e comunicati attraverso incontro di rendicontazione sociale; Con la pubblicazione su Scuola in Chiaro del RAV), tramite aree riservate di quest'ultimo a cui poter accedere (Un'analoga comunicazione è fatta al consiglio di Istituto e pubblicata in un'area del sito riservata alle famiglie) oppure sono caricati nel cloud della scuola per la l'analisi da parte dei valutatori esterni (Condivisione collegiale - documenti predisposti e condivisi in area cloud di sistema - rapporti in occasione di visite di sorveglianza Accreditamento Regione Piemonte).

Il *decision making evidence based and informed* è un altro degli aspetti molto dibattuti negli ultimi anni quando si parla di politiche educative (OECD, 2013). L'OCSE (2020g) ha definito il *decision making evidence based and informed* come un processo decisionale basato e informato dalle evidenze in cui più fonti di informazioni, tra cui statistiche, dati, *best practices* ed esiti di ricerca, vengono consultate prima di prendere una decisione per pianificare, attuare e (se pertinente) modificare lo *status quo*. La buona governance si basa, infatti, proprio sulle competenze di chi è chiamato a prendere decisioni nel consultare i dati a disposizione per raggiungere obiettivi ampi e migliorare in un'ottica costruttiva. Ciò, tuttavia, richiede un approccio sistemico che coinvolga non solo le figure apicali di un governo, ma anche chi, nella sua attività quotidiana, è chiamato a prendere decisioni per far funzionare una piccola comunità come quella scolastica. È il caso dei dirigenti scolastici. Nel gruppo di riferimento alla domanda *Quando si trova a prendere decisioni per la sua scuola, su queste ultime influiscono gli esiti dei processi di autovalutazione?* le risposte si sono distribuite come mostrato nel grafico 4 in cui 1 indica *per nulla* e 5 indica *moltissimo*.

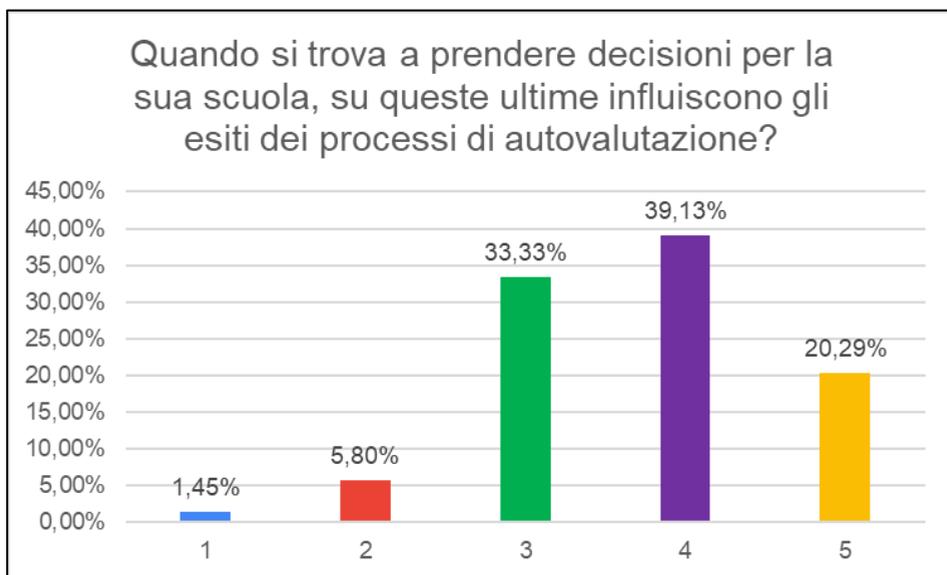


Grafico 4. Risposte alla domanda *Quando si trova a prendere decisioni per la sua scuola, su queste ultime influiscono gli esiti dei processi di autovalutazione?* Fonte: elaborazione propria.

Una *best practice* esemplificativa in tal senso è l'esperienza riportata da uno dei dirigenti scolastici intervistati telefonicamente che a partire dai dati emersi dall'autovalutazione di istituto, nello specifico dalle criticità, ha realizzato e sta realizzando alcuni progetti ad hoc.

Quando sono arrivata, noi avevamo delle criticità sulle strumentazioni che, partecipando a vari progetti e a vari PON, stiamo risolvendo. Poi avevamo delle fasce di alunni in gravi difficoltà, tra cui ragazzi e bambini di etnia rom. Ho notato che i loro esiti a distanza erano abbastanza scarsi, nel senso che alcuni molto spesso si perdevano e non proseguivano. Quindi stiamo seguendo questo aspetto. Per esempio, abbiamo incrementato molto il lavoro sulle fasce deboli, in particolare sugli studenti di etnia rom, infatti quest'anno abbiamo avuto due ragazzi che si sono diplomati nella nostra scuola e si sono iscritti alla scuola superiore. Seguirò questa cosa perchè è un primo risultato importante, sai, non sono tanti, ma non è nemmeno poco. In questo senso stiamo cercando di sostenere questi ragazzi in questo modo. [...] Un altro progetto in collaborazione con Save the Children mette a disposizione un supporto pomeridiano online ad alcuni studenti, quindi un tutoraggio a distanza per i compiti a casa. Abbiamo

stretto un accordo anche con Sant'Egidio per la collaborazione nei campi nomadi. Abbiamo elaborato dei progetti ad hoc di finanziamenti per i bambini per gli strumenti, per permettergli di partecipare ad uscite didattiche nell'ottica dell'inclusione. (dirigente scolastico della provincia di Roma)

Contrariamente l'efficienza e l'efficacia dei processi di autovalutazione non è percepita né positivamente né negativamente dai rispondenti per il 50% dei casi. Per il 4.35% di loro i processi di autovalutazione nella scuola non sono per nulla efficaci ed efficienti, per l'11.59% lo sono poco, per il 30.43% lo sono abbastanza, mentre per il 2.90% sono davvero molto efficaci ed efficienti (grafico 5, 1 - *per nulla*, 5 - *moltissimo*).

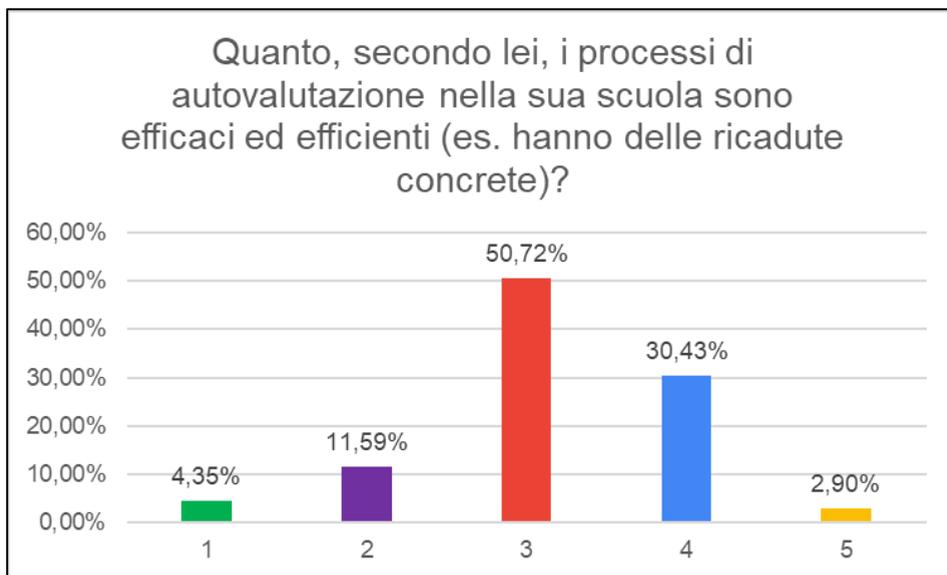


Grafico 5. Risposte alla domanda *Quanto, secondo lei, i processi di autovalutazione nella sua scuola sono efficaci ed efficienti (es. hanno delle ricadute concrete)?* Fonte: elaborazione propria.

Emblematica a tal proposito è la risposta di uno dei dirigenti scolastici intervistati.

Mi piacerebbe molto approfondire questo discorso, però lei sa bene che nella nostra professione, con tutte le cose che ci sono da fare, purtroppo ci troviamo a

rincorrere disposizioni a volte anche vuote, a dover rispondere di problematiche strutturali di cui non siamo causa. Quindi sostanzialmente tutta la parte che dovrebbe riguardare il lavoro di un dirigente scolastico, cioè l'aspetto didattico, le ricadute sugli apprendimenti, la pedagogia didattica, tutto ciò che riguarda il percorso didattico e di valutazione... Le indagini sulla valutazione di sistema dell'Invalsi, dell'OCSE-PISA le apprezzo. [...] Nonostante alcune associazioni siano troppo critiche sulla valutazione di sistema, secondo me è molto utile come base di partenza. (dirigente scolastico della provincia di Salerno)

L'ultimo item del questionario, facoltativo e a risposta aperta, lasciava ai rispondenti la possibilità di aggiungere altro sulla propria percezione o esperienza rispetto all'autovalutazione di istituto. A titolo esemplificativo si riportano di seguito alcune suggestioni emerse.

Ci vuole formazione obbligatoria per passare dalla logica dell'adempimento a quella della comunità che si fa carico dei problemi. (Rispondente 1)

La valutazione dovrebbe essere centrale ma non c'è tempo per occuparsene. (Rispondente 2)

Il format del RAV andrebbe migliorato, gli interventi dei nuclei esterni di valutazione estesi a un maggior numero di scuole. L'intero processo dovrebbe essere gestito dalle scuole nella loro autonomia con maggiore consapevolezza. (Rispondente 3)

La scuola tra rendicontazione sociale, RAV e PTOF è diventata una azienda che produce documenti che nessuno legge e conosce ad eccezione del dirigente scolastico e dei quattro docenti che ci lavorano. (Rispondente 4)

L'autovalutazione è uno strumento molto potente e molto importante; occorrerebbe che il dirigente scolastico fosse affiancato da un middle management per una gestione più efficace di questa area di presidio. (Rispondente 5)

L'autovalutazione mette in campo un processo complesso che richiederebbe un forte investimento di tempo e di risorse umane. Non sempre si possiede

capacità critica e metacognitiva sulle strategie e azioni adottate. Si assiste a una grande autoreferenzialità, soprattutto nei docenti. (Rispondente 6)

Sembrerebbe che avere consapevolezza della comunità scolastica e del territorio in cui si opera sia uno degli aspetti che suscitano riflessioni nei dirigenti scolastici raggiunti. La non capillare diffusione della cultura della valutazione fa sì che non tutti gli attori coinvolti abbiano chiari i meccanismi valutativi in atto nelle scuole. Inoltre, la pandemia ha aumentato il carico di lavoro di chi la scuola la vive e, dunque, a parere di molti rispondenti bisognerebbe far guadagnare nuovamente all'autovalutazione di istituto una posizione centrale nei processi gestionali. La mancanza di tempo e di risorse umane, infatti, sembrano essere due degli ostacoli più imponenti per la corretta gestione delle pratiche valutative, anche se, l'autovalutazione può configurarsi come un'occasione di "apprendimento e di riflessione a tutti i livelli" (Rispondente 7) in cui la cultura della valutazione non è appannaggio dei soli dirigenti scolastici, ma è una cultura propria di tutto il sistema.

Conclusioni

La tesi di dottorato si inserisce nel dibattito sulla valutazione dei sistemi scolastici. Perché è importante parlarne oggi? La valutazione della scuola, per la scuola e nella scuola si configura ormai da qualche decennio a questa parte come una “sfida istituzionale” e una “questione pedagogica” rilevante (Cerini, 2010). A tal proposito, nel fare chiarezza nel *mare magnum* della valutazione, Argentin (2016) individua tre possibili termini inglesi che ne chiariscono il significato. L'*accountability* fa riferimento alla valutazione dei prodotti realizzati dall'implementazione delle procedure. L'*assessment* fa riferimento alla valutazione dei risultati raggiunti o alle procedure seguite per raggiungerli nell'erogazione di un servizio. Infine, l'*evaluation* fa riferimento alla valutazione degli effetti derivanti dalla messa a sistema delle procedure e delle politiche pubbliche (Landri & Maccarini, 2016). Nonostante sia intesa in senso lato, la valutazione è uno strumento conosciuto dalle molte sfaccettature che ancora fatica ad amalgamarsi nelle e con le politiche educative e nelle e con le pratiche di miglioramento. La *Global Society* e la *Knowledge Society*, tuttavia, hanno evidenziato l'importanza dell'*evidence informed policy-making* e l'importanza che in questa direzione può assumere la valutazione dei servizi pubblici, istituzioni scolastiche comprese (OECD, 2020g). In particolare, negli ultimi decenni e con tempistiche differenti, le agende educative nazionali hanno iniziato a popolarsi di misure di decentralizzazione, feedback e *accountability* (Faubert, 2009; Bray et alii, 2015). L'autonomia scolastica ha evidenziato la necessità di mettere a punto dispositivi volti alla valutazione delle istituzioni educative e, contemporaneamente alla diffusione della tecnologia, ha permesso la realizzazione di analisi su larga scala. Il *Large Scale Assessment* ha così iniziato ad essere sempre più utilizzato per monitorare i sistemi educativi e la loro organizzazione, si pensi per esempio alla valutazione standardizzata degli apprendimenti o alle pratiche di valutazione esterna e autovalutazione delle scuole (Jullien, 1817; Brickman, 1960; OECD 2001, 2011, 2020b; Martini, 2014; Agrusti, 2014; Mattarelli, 2022). Parallelamente si è venuta affermando la questione della qualità

dell'istruzione, della formazione e dell'educazione nel suo complesso. Come stabilire il buono o il cattivo funzionamento di un sistema scolastico? Sulla base di quali criteri valutare gli input e gli output educativi? Definire il concetto di qualità non è semplice poiché è strettamente correlato con gli obiettivi da realizzare e con il contesto di riferimento. Per questo motivo molto spesso alla qualità si affiancano l'efficacia, l'efficienza e l'equità come "termometro" di quest'ultima (Dell'Anna, 2021). A questo proposito, Vertecchi (2003) segnala quattro funzioni che caratterizzano la valutazione: la valutazione didattica dell'organizzazione scolastica, la valutazione formativa di monitoraggio, la valutazione sommativa attraverso benchmark standardizzati e la valutazione della qualità. La quarta sembra essere quella più insidiosa dal momento che le riunisce tutte e tre in vista del miglioramento. Dell'Anna (2021, p. 20), infatti, precisa che si tratta di "un processo ricorsivo tra programmazione, realizzazione, verifica, e miglioramento".

A partire da questo scenario e facendo proprio questo framework di riferimento, il lavoro di ricerca ha avuto lo scopo di tracciare i confini e le caratteristiche della valutazione dei sistemi scolastici in tre realtà europee, l'Italia, la Finlandia e la Spagna. I due problemi conoscitivi da cui ha preso piede il progetto dottorale sono da scovare da una parte nelle spinte valutative che si sono affermate sempre di più e che non sempre sono state correlate a concreti interventi di sviluppo (Allulli, 2017), dall'altra nella diffusione della cultura della valutazione che si è alternata tra adempimenti burocratici e impulsi per il miglioramento (Giampietro et al, 2016). Le conseguenti domande di ricerca hanno guidato lo studio. Se sulle finalità e sugli oggetti della valutazione dei sistemi scolastici c'è accordo, grazie anche alla promulgazione di linee guida comunitarie, sui processi valutativi sussiste ancora divergenza (Fondazione Agnelli, 2014). Per questo motivo, si è ritenuto utile e interessante tentare di dare una risposta alle tre domande di ricerca emerse a seguito dell'analisi della letteratura preesistente: quali modelli di valutazione del sistema scolastico vengono attuati in Italia, Finlandia e Spagna? Le istituzioni pubbliche come rispondono alla valutazione di sistema? Quale supporto è fornito alle scuole nei processi di autovalutazione? In aggiunta all'Italia, la Finlandia e la Spagna sono stati scelti come casi "emblematici e rilevanti" (Flyvbjerg, 2006) poiché

importanti realtà educative europee (OECD, 2004). La Finlandia in quanto è ritenuto uno dei sistemi scolastici migliori al mondo di cui si è iniziato a parlare subito dopo la prima edizione dell'indagine PISA dell'OCSE (OECD, 2001), la Spagna dal momento che è un sistema scolastico con un trascorso simile a quello italiano, ma con un assetto geo-politico federale differente (OECD, 2019abc). Nonostante la "palude classificatoria" in cui affonda e la scarsa considerazione che spesso lo colpisce (Gerring, 2004; Sena, 2021), è stato fatto ricorso allo studio di caso avendo come obiettivo la sua estendibilità (Khan, 2007), ovvero la volontà di arricchire l'esperienza di chi legge dando la possibilità di scendere nella profondità della valutazione dei sistemi scolastici. L'approccio è stato esplorativo-descrittivo e si è avvalso della comparazione tra casi (Yin, 2009). Rifacendosi al modello multi-livello di Bray e Thomas (1955), la comparazione si è esplicitata a partire da un livello *geografico/locazionale* che ha tenuto conto dei paesi oggetto di analisi, è proseguita attraverso i *gruppi demografici non locazionali*, ovvero gli individui interessati, ed è terminata nel livello degli *aspetti legati all'educazione* e alla società, cioè le strutture di gestione della valutazione della scuola. All'interno di questi multi-livelli teorici, si sono avvicinati due ulteriori livelli di analisi. Come individuato da Zanazzi (2014), il livello internazionale ha indagato i documenti prodotti a guida della valutazione dei sistemi educativi, il livello nazionale, invece, ha corroborato i dati raccolti nel panorama internazionale andando ad interrogare alcuni testimoni privilegiati di ciascun contesto. L'enucleazione dei cinque passi flessibili per l'implementazione di uno studio di caso descritta da Merriam (1998) ha fatto da sfondo alla definizione del quadro teorico della ricerca, ma, diversamente da quanto presupposto dalla studiosa statunitense, se i metodi di raccolta dei dati sono stati perlopiù qualitativi, non è mancato l'utilizzo di tecniche quantitative proprie dei *mixed methods* (Trinchero & Robasto, 2019). Riprendendo il *progressive focusing* di Perlett e Hamilton (1976), infatti, lo studio di caso si è sviluppato man mano che esso si realizzava.

L'analisi documentale qualitativa, descritta nel secondo capitolo della seconda parte, ha previsto l'analisi dei contenuti organizzati per temi e categorie. La revisione sistematica e critica dei documenti disponibili, sia in formato cartaceo che digitale, attraverso un iniziale processo di

skimming è stata seguita dall'individuazione dei dati e dei passaggi pertinenti, rilevanti e aderenti alle domande di ricerca e dalla rilettura approfondita delle fonti (Atkinson & Coffey, 1997; Bowen, 2009). La restituzione è avvenuta secondo il modello proposto da Cardone e colleghi (2010) che ha dato la possibilità di organizzare il materiale raccolto seguendo alcune domande guida. Le interviste qualitative rivolte ai National Project Manager dell'indagine PISA, i responsabili e rappresentanti nazionali dello studio dell'OCSE, descritte nel terzo capitolo della seconda parte, hanno approfondito e avvalorato quanto sistematizzato attraverso l'analisi documentale. L'intervista, infatti, essendo una conversazione tra l'intervistatore e l'intervistato all'interno della quale ha luogo una dinamica sociale di negoziazione di significati (Corbetta, 1999), ha permesso di cogliere sfumature e percezioni riguardanti la valutazione dei sistemi educativi ascoltando la voce di testimoni molto privilegiati. La traccia costruita ad hoc ha assicurato omogeneità alla raccolta dei dati, ma la non standardizzazione degli stimoli, soprattutto nella seconda parte del colloquio, ha permesso di cogliere le reazioni degli intervistati alla luce dei contesti nazionali differenti all'interno dei quali si situano i processi valutativi. La conduzione online su piattaforma ZOOM se da una parte ha "raffreddato" la costruzione dell'interazione tra chi scrive e l'esponente istituzionale intervistato, dall'altra ha garantito la raccolta dei dati anche durante i primi mesi della pandemia (Di Fraia, 2004; Lombi, 2015). L'analisi tematica riflessiva ha guidato l'analisi dei dati. Grazie alla sua flessibilità, ha permesso di esaminare le ricche riflessioni e suggestioni emerse durante i colloqui telematici e di enucleare tre temi principali aderenti alle domande di ricerca: valutazione standardizzata degli apprendimenti; valutazione della scuola e per la scuola; valutazione e policy processes (Clarke & Braun, 2013; Terry et al., 2017). Il questionario, descritto nel quarto capitolo della seconda parte, aggiunto al sopraggiungere del diniego alla partecipazione alla ricerca da parte di alcuni National Project Manager e al perdurare delle criticità sanitarie che hanno imposto la revisione dell'originale progetto di ricerca, ha triangolato le fonti, in particolar modo nel contesto italiano (Denzin, 1970). Le risposte dei dirigenti scolastici sono state raccolte attraverso una duplice modalità: compilazione manuale tramite Google Moduli e

risposta telefonica. Visto il noto sovraccarico di lavoro dei dirigenti scolastici è stata data la possibilità di scegliere la tipologia di somministrazione. Quella telefonica, essendo un'intervista, ha permesso di approfondire meglio alcuni aspetti, quella manuale, invece, ha permesso di raggiungere un maggior numero di rispondenti (Zammuner, 1996; Lombi, 2015). Le domande a risposta aperta, che si sono aggiunte a quelle che prevedevano una risposta chiusa, hanno consentito di esplorare i punti di vista dei dirigenti scolastici offrendo dati qualitativi (Sena, 2021). L'analisi di queste ultime è avvenuta tramite l'individuazione di macrocategorie e microcategorie con un approccio di tipo bottom-up, facendo riferimento al *framework method* (Batini et alii, 2020).

La sfida del lavoro di ricerca è stata quella di indagare in ottica comparativa un argomento tanto delicato quanto complesso come le pratiche e le percezioni legate alla valutazione dei sistemi scolastici. Senza la pretesa di essere esaustivi, la rilevanza del lavoro di ricerca consiste nell'aver ricostruito i quadri attuali della valutazione dei sistemi scolastici attraverso l'analisi documentale, nell'aver ascoltato la voce dei National Project Manager dell'indagine PISA dell'OCSE, massimi esperti nazionali di valutazione degli apprendimenti, e nell'aver gettato un fascio di luce su un gruppo di dirigenti scolastici italiani e sulle loro opinioni ed esperienze in merito alla valutazione di istituto e all'autovalutazione.

La prima parte del presente lavoro si è concentrata sul framework teorico della valutazione nella sua accezione più ampia e della valutazione dei sistemi educativi nello specifico. Dopo aver approfondito l'epistemologia della valutazione educativa, primo capitolo della prima parte, sono state descritte le possibili forme che la valutazione dei sistemi educativi può assumere. È stata ricostruita la storia del lento sviluppo degli indicatori dell'OCSE per l'educazione, secondo capitolo della prima parte, e delle indagini comparative standardizzate sugli apprendimenti. Sono stati presentati i principali studi dell'IEA, *International Association for the Evaluation of Educational Achievement*, e dell'OCSE, Organizzazione per la Cooperazione e lo Sviluppo Economico, con particolare rimando all'indagine PISA sulle competenze dei quindicenni, terzo capitolo della prima parte. Infine, è

stata delineata la valutazione delle istituzioni scolastiche, intesa come valutazione esterna e autovalutazione di istituto, quarto capitolo della prima parte. La seconda parte del lavoro, empirica, ha tracciato il percorso compiuto all'interno del progetto di ricerca dottorale. È stato descritto il disegno di ricerca nel primo capitolo della seconda parte e sono stati riportati tanto i risultati dell'analisi documentale quanto quelli delle interviste qualitative e del questionario nei capitoli due, tre e quattro della seconda parte.

In generale i tre paesi oggetto di studio, Italia, Finlandia e Spagna, hanno implementato tre modelli differenti per la valutazione dei propri sistemi scolastici che rispecchiano l'assetto sociale e politico di ciascuna realtà. L'Italia ha un sistema valutativo più centralizzato, la Finlandia più decentralizzato, mentre la Spagna frammentato-di tipo federale. L'analisi documentale qualitativa, discussa nel secondo capitolo della seconda parte, ha messo in evidenza che in tutti i contesti di studio le autorità politiche centrali hanno centrato la valutazione dei sistemi educativi attraverso atti legislativi. In Finlandia le leggi n. 1295 del 2013 e n. 582 del 2015 definiscono i ruoli e i compiti di KARVI-FINEEC, *Finnish Education Evaluation Centre*, l'istituto di ricerca valutativa incaricato di valutare il sistema educativo (Laki, 1295/13, 2013, 582/15, 2015) (cfr. § 2.3.1). In Spagna le Ley Organiche n. 2 del 2006 e n. 3 del 2020 regolamentano la valutazione del sistema educativo dando compiti precisi all'INEE, *Instituto Nacional de Evaluacion Educativa*, compresa la realizzazione di indicatori nazionali per l'educazione, aspetto della valutazione del sistema scolastico e della realizzazione dell'equità educativa tipicamente spagnolo (Ley Organica, 2/06, 2006, 3/20, 2020) (cfr. § 2.3.2). In Italia il D.P.R n. 80 del 2013 istituisce il Sistema Nazionale di Valutazione a cui capo si trovano l'Invalsi, l'Indire e il Contingente ispettivo che concretizzano la valutazione in quattro passi: autovalutazione delle istituzioni scolastiche, valutazione esterna, azioni di miglioramento e rendicontazione sociale (D.P.R 80/13, 2013) (cfr. § 2.3.3). Gli stessi enti, KARVI, INEE e Invalsi, sono responsabili, ciascuno per le proprie competenze attribuitegli dalle leggi, della valutazione del sistema scolastico e degli apprendimenti su scala nazionale (cfr. § 2.4.1, 2.4.2 e 2.4.3). In Finlandia questi ultimi vengono misurati su base campionaria e longitudinale, poiché l'obiettivo è quello di rilevare le

tendenze storiche. Inoltre, su scala nazionale è prevista una valutazione specifica degli apprendimenti per chi conclude la scuola dell'obbligo. In Spagna, al contrario, gli apprendimenti sono testati sia su base censuaria, come in Italia, che campionaria, a seconda che si tratti dell'*evaluacion de diagnostico* o dell'*evaluacion general del sistema educativo*, e i risultati che da essi ne derivano confluiscono nelle azioni di miglioramento contenute all'interno del Programma Generale Annuale. Per quanto riguarda l'oggetto della valutazione degli istituti nazionali, invece, esistono alcune divergenze. Mentre in Spagna e Italia vengono realizzate annualmente indagini censuarie nazionali, in Spagna all'interno dell'*evaluacion de diagnostico*, volte a testare le competenze, soprattutto nella lingua madre e in matematica (e in inglese in Italia), degli studenti di alcune annualità – in Italia le annualità coinvolte sono più numerose rispetto a quelle spagnole –, in Finlandia le rilevazioni sono di tipo campionario e non vengono effettuate con cadenza regolare (Ferrer, 2014; Invalsi, 2018ab) (cfr. § 2.5.2 e 2.5.3), bensì in base a quanto definito dai piani di valutazione elaborati per ogni triennio (KARVI-FINEEC, 2022) (cfr. § 2.5.1). Un'altra discrepanza emersa tramite l'analisi documentale condotta è da rintracciarsi nelle pratiche di valutazione esterna delle istituzioni scolastiche (cfr. § 2.6.1, 2.6.2, 2.6.3): queste ultime se, con diversi protocolli, sono presenti sia in Spagna che in Italia, sono del tutto assenti in Finlandia dove non è previsto che i valutatori esterni si rechino nelle scuole per valutarne il funzionamento. Al contrario, l'autovalutazione delle istituzioni scolastiche è presente in tutti i contesti oggetto di studio, anche se i modelli sottesi variano da paese a paese. In Finlandia, per esempio, l'autovalutazione degli erogatori di servizi educativi è obbligatoria così come lo è quella proposta da KARVI secondo il piano di valutazione. Per l'autovalutazione delle scuole sono state emanate linee guida nazionali, i criteri di qualità dell'istruzione di base, tenendo conto delle differenti "culture operative" (Opetusministeriö, 2010), ma sono le scuole stesse in sinergia e collaborazione con le autorità locali a decidere concretamente come realizzare l'autovalutazione di istituto costruendo un framework ad hoc che tiene conto della qualità delle strutture e dei processi e che si realizza attraverso descrittori e criteri guida. In Spagna, invece, la frammentazione geografica in comunità autonome lascia a ciascuna di

esse la possibilità di decidere autonomamente come, dove e quando richiedere alle istituzioni scolastiche di autovalutarsi. Indipendentemente da quanto stabilito regionalmente, si pensi alle comunità autonome di Aragon, Navarra e Castiglia-La Mancha (Comunidad autonoma de Castilla-La Mancha, 2003; Centro de Profesorado de Sevilla, 2021; Servicio de inspeccion educativa, 2021; Gobierno de Aragon, 2022), i processi di autovalutazione mettono in luce punti di forza e criticità tenendo molto conto degli esiti delle visite esterne, se ricevute, e della valutazione standardizzata degli apprendimenti nazionale e internazionale. Diversamente dall'Italia, le scuole sono supportate in questi momenti da consulenti scolastici e da un coordinatore presente in ogni istituto, figura simile al nucleo interno di valutazione italiano. Le istituzioni coinvolte avviano, diversamente dal nostro paese, un processo partecipato a cui prendono parte sia attori interni che esterni alla scuola, il territorio e la comunità. La valutazione esterna, invece, ricade nelle competenze del Ministero dell'Educazione e della Formazione Professionale, delle Comunità Autonome e dell'Ispettorato. I valutatori realizzano rilevazioni e report periodici e si attengono alle funzioni per loro stabilite da ciascuna comunità autonoma: le visite non avvengono in tutte le scuole, ma solo in quelle segnalate. Il protocollo seguito è molto simile a quello italiano, infatti, i valutatori prima consultano i documenti, poi si recano sul posto, infine realizzano dei colloqui e redigono il rapporto finale per la scuola, per le autorità regionali e per il Dipartimento dell'istruzione centrale (Boja, 2019; BOCM, 2021). In Italia, per quanto riguarda l'autovalutazione delle istituzioni scolastiche, la presenza del Rapporto di Autovalutazione da compilare e diffondere ogni anno uniforma i processi autovalutativi delle scuole dal nord al sud del paese. Tuttavia, la presenza di indicatori e descrittori guida al suo interno, compresa la nuova versione per il triennio 2022-2025, rende lo strumento italiano simile nella struttura a quello di Finlandia e Spagna. La valutazione esterna è regolamentata, oltre che dal D.P.R 80/2013, anche dalla Conferenza per il coordinamento funzionale del Sistema Nazionale di Valutazione e dall'Invalsi che ne stabilisce il protocollo. Un aspetto che nella valutazione esterna accomuna l'Italia e la Spagna è la presenza dei valutatori che si recano nelle scuole, ma, mentre in Italia le unità sono

tre, in Spagna sono due (Opetusministeriö, 2010; Eurydice, 2015; Invalsi, 2016ab; Boja, 2019; BOCM, 2021) (cfr. § 2.6.1, 2.6.2 e 2.6.3). L'analisi documentale ha permesso di scattare un'istantanea sullo stato dell'arte della valutazione dei sistemi scolastici al giorno d'oggi. Oltre a status differenti di avanzamento, è emersa una compartecipazione diversa ai processi valutativi, l'autovalutazione delle istituzioni scolastiche, ad esempio, è risultata essere più partecipata in Finlandia e Spagna che in Italia. È emersa anche una interazione differente tra autovalutazione e valutazione esterna, così come stabilito dalle norme e dalle pratiche che la rendono concreta nei contesti scolastici, e una disponibilità differente di documenti consultabili. Se per alcuni aspetti e per alcuni contesti è stato meno complesso consultare la documentazione pubblicata, per altri non lo è stato affatto. La documentazione sui contenuti delle prove nazionali per la valutazione degli apprendimenti degli studenti spagnoli, ad esempio, è stata di difficile reperimento, al contrario di quella italiana cui si è prestata particolare attenzione anche alla luce delle critiche che hanno caratterizzato l'impostazione e la validità dei test Invalsi (Trinchero, 2014c; Corsini & Losito, 2014). Ciò che si è rilevato, inoltre, è che mentre in alcuni contesti nazionali, la Finlandia per esempio, sono molto chiari i modelli adottati nella valutazione, si pensi al Ciclo di Deming o al *Common Assessment Framework* dell'autovalutazione delle istituzioni scolastiche, per altri, ad esempio la Spagna, è stato difficile rintracciare il modello alla base della valutazione. Senza dubbio quest'ultimo paese sconta, come già accennato, un'elevata frammentazione delle pratiche valutative a causa dell'assetto federalistico. Non da ultimo, la non conoscenza particolarmente approfondita della lingua spagnola ha reso più onerosa l'interpretazione della documentazione. Un'altra difficoltà incontrata è legata al reperimento di una bibliografia critica circa la valutazione di sistema, che avrebbe potuto arricchire l'analisi documentale condotta. Le fonti critiche nazionali sulle tematiche oggetto di studio sono state difficilmente individuate, ad eccezione dell'Italia e in parte della Spagna (Ferrer, 2014). Ciò non ha permesso una comparazione complessiva dei "problemi della valutazione" ma ha permesso, nell'ottica del *progressive focusing* (Parlett & Hamilton, 1976), di concentrarsi su questi aspetti durante i colloqui con i National Project Manager dell'indagine PISA

dell'OCSE che hanno aiutato a ricostruire, attraverso le loro parole, la delicata faccenda delle critiche ideologiche mosse dagli attori coinvolti e dall'opinione pubblica alle pratiche di valutazione (Mattarelli & Zanazzi, 2021).

Le riflessioni dei National Project Manager dell'indagine PISA dell'OCSE, discusse nel terzo capitolo della seconda parte, oltre ad aver argomentato i risultati raggiunti dagli studenti nelle indagini standardizzate e internazionali sugli apprendimenti (cfr. § 3.4), hanno sottolineato alcuni aspetti della valutazione scolastica di sistema e del rapporto che questa ha con le autorità pubbliche. In Finlandia, da parte di chi vive la scuola l'approccio alla valutazione è positivo anche se le pratiche valutative di tanto in tanto sono percepite come gravose e le maggiori critiche vengono mosse alla rilevanza che gli esiti di queste ultime hanno nel guidare le scelte politiche. In Spagna, al contrario, gli stakeholders, cioè dirigenti scolastici, insegnanti, studenti e famiglie, preferirebbero essere più coinvolti nella diffusione dei risultati e avere più feedback. I giudizi negativi nei confronti della valutazione di sistema si ancorano alla possibilità che le indagini standardizzate possano essere uno strumento universale per la valutazione di qualsiasi aspetto educativo. Diversamente, in Italia, la paura nei confronti della valutazione sembra essere diminuita negli ultimi anni, anche se permangono retaggi e critiche circa la possibilità di poter "essere bocciati". Il senso di coinvolgimento e la familiarizzazione dei valutati ha contribuito ad aumentare l'accettazione delle pratiche valutative, ma la percezione di queste ultime da parte dei valutati stessi come un qualcosa di scervo e lontano ha notevoli ripercussioni. Ancora una volta la suddivisione spagnola in comunità autonome implica una serie di conseguenze nell'implementazione della valutazione e nel trasferimento di quest'ultima dalle autorità centrali ai contesti scolastici regionali e locali. L'utilizzo dei dati da parte delle scuole è stato un altro aspetto approfondito durante le interviste poiché è un passaggio e un fattore molto importante nella valutazione di sistema. In Finlandia gli esiti delle valutazioni sugli apprendimenti sono restituiti alle scuole attraverso report che, a parere dell'intervistato, sono utilizzati come base di partenza per prendere decisioni. In Spagna, ma soprattutto in Italia, a parere della National Project Manager, a volte la mancanza di

competenze nella lettura degli esiti lascia all'autoreferenzialità l'utilizzo delle evidenze nei processi di *decision making*. Se l'utilizzo dei dati da parte delle scuole rappresenta uno degli effetti della valutazione a livello micro, l'utilizzo della valutazione da parte delle autorità pubbliche locali o centrali è uno degli effetti a livello meso e macro. Così come le scuole tengono conto degli esiti, anche il Ministero dell'Educazione e della Cultura finlandese segue con molta attenzione gli sviluppi delle valutazioni e le include nei programmi politici. In Spagna, così come avvenuto nel livello micro, l'assetto federale ostacola le politiche centrali che, prima di arrivare nelle istituzioni scolastiche, devono passare dai vertici delle comunità autonome. Questo comporta che i dati potrebbero essere utilizzati di più e in modo più proficuo. In Italia, invece, a parere della National Project Manager, non sembra esserci particolare considerazione degli esiti né a livello micro, né a livello meso o macro. Spesso, infatti, sembra che le politiche educative ed economiche non siano costruite a partire dalle evidenze empiriche. Questo dato è in contrapposizione con quanto dichiarato dai 69 rispondenti al questionario, discusso nel quarto capitolo della seconda parte, a cui si sono aggiunti 3 dirigenti scolastici che hanno preferito l'intervista telefonica (cfr. § 4.4). Questi ultimi, infatti, nel 47,06% dei casi hanno risposto che la valutazione di istituto e la compilazione del Rapporto di Autovalutazione sono strumenti funzionali per la riflessione e il miglioramento della scuola, sebbene il 41,18% abbia aggiunto che si tratta anche di un adempimento formale. Emblematica, in questo senso è la risposta all'item *Quando si trova a prendere decisioni per la sua scuola, su queste ultime influiscono gli esiti dei processi di autovalutazione?*. In una scala Likert a cinque passi in cui 1 indica *per nulla* e 5 *moltissimo*, l'1,45% dei dirigenti scolastici ha affermato che gli esiti dei processi di autovalutazione di istituto non influiscono sui processi decisionali, il 5,80% che influiscono poco e il 33,33% che influiscono solo in parte. Il 39,13% ha affermato che gli esiti dei processi di autovalutazione hanno una modesta influenza sui processi decisionali, mentre il 20,29% che influiscono moltissimo. I bisogni formativi legati all'autovalutazione di istituto rappresentano una tematica complessa e delicata. Il 91,30% dei rispondenti ha preso parte ad occasioni di formazione su tematiche

inerenti la valutazione di sistema, ma il 64,71% afferma di voler essere ulteriormente aggiornato e formato.

Il momento emergenziale e pandemico ha indubbiamente imposto alcuni limiti al lavoro di ricerca dottorale. Lo studio era stato inizialmente pensato per essere realizzato attraverso visite in loco. Il progetto di ricerca iniziale, infatti, prevedeva la realizzazione di interviste e focus group *face to face* con esperti nazionali in ambito valutativo e con i direttori degli istituti di ricerca valutativa di ciascun paese. Il sopraggiungere della pandemia a cinque mesi dall'inizio del dottorato ha richiesto una modifica sostanziale dei piani iniziali. Si è cercato di riprendere questi ultimi dopo il primo lockdown del 2020, ma il perdurare delle criticità sanitarie e l'istituzione di limitazioni negli spostamenti tra regioni italiane e stati esteri non lo ha concesso. L'effetto della pandemia sulla ricerca valutativa, nello specifico sulla valutazione standardizzata degli apprendimenti, si è reso evidente anche quando alcune indagini internazionali calendarizzate per il biennio 2020-2021 sono state posticipate al 2022. Le conseguenze che ciò ha avuto sulla ricerca dottorale sono palpabili. Essendo stato scelto come gruppo di testimoni privilegiati quello formato dai National Project Manager dell'indagine PISA, il posticipo del *main study* sulle competenze dei quindicenni al 2022 ha fatto sì che gran parte del secondo anno del percorso di dottorato, tradizionalmente dedicato alla raccolta dei dati sul campo, si sovrapponesse con l'indagine principale OCSE. La ripercussione è stata notevole: inizialmente nella rosa dei contesti nazionali da indagare attraverso la ricerca di dottorato erano stati inseriti anche altri paesi, Germania e Francia in primis. Tuttavia, alla luce della situazione critica e del carico di lavoro caratterizzato da imprevisti e didattica a distanza, sottolineato dagli stessi National Project Manager nelle interviste qualitative, i responsabili nazionali dell'indagine di Germania e Francia non sono riusciti ad accordare la loro disponibilità nel prendere parte alla ricerca. In un primo momento si è cercato di correre ai ripari aggiungendo alla rosa dei paesi, oltre a Italia, Finlandia e Spagna di cui si aveva avuto già la disponibilità, anche i Paesi Bassi, modello di efficacia scolastica secondo Scheerens (Scheerens et alii, 2011; Scheerens, 2018), ma al sopraggiungere di un ulteriore diniego, si è deciso di mantenere lo *status quo*. Questo, se da

una parte non ha permesso l'ampliamento del numero dei paesi oggetto di indagine, dall'altra ha concesso l'opportunità di focalizzare l'attenzione su determinati aspetti della valutazione dei sistemi scolastici su cui probabilmente, con un numero di paesi più esteso, non sarebbe stato possibile fare un affondo, si pensi per esempio al questionario rivolto ai dirigenti scolastici somministrato nel contesto nazionale. Un limite che, dunque, è divenuto un punto di forza. Oltre alle questioni emergenziali-pandemiche, sono presenti altri limiti strutturali del lavoro di ricerca condotto (Zanazzi, 2014). In primis le tempistiche stringenti del dottorato che, nonostante la concessione di una proroga di tre mesi a causa della pandemia, non hanno permesso di estendere il periodo di somministrazione del questionario rivolto ai dirigenti scolastici italiani. La finestra temporale loro concessa per la compilazione del questionario è stata ridotta a poco meno di due mesi, a ridosso della fine della pausa estiva e della ripresa delle attività didattiche: per assicurare il giusto tempo e la validità all'analisi dei dati non è stata, infatti, presa in considerazione la possibilità di prolungare la finestra temporale di compilazione. Non si esclude che, se ciò fosse avvenuto, si sarebbero potute raccogliere altre risposte, ampliando così il campione (Zanazzi, 2014). In secundis, la distanza geografica dagli altri due paesi coinvolti, Finlandia e Spagna, non ha permesso di somministrare il questionario ai rispettivi dirigenti scolastici. Per non incorrere in un ulteriore rifiuto vista la situazione concitata creata dal sopraggiungere della pandemia, chi scrive ha desistito dal cercare contatti con dirigenti scolastici oberati di lavoro nelle altre due realtà oggetto di indagine, lasciando questo come un possibile sviluppo futuro della ricerca ora che, con cautela, si sta tornando alla normalità. Il terzo limite, come evidenziato da Zanazzi (2014), è intrinseco alla realizzazione di studi di caso comparativi e riguarda la scelta dei casi rivelatori (Sena, 2021). Dovendo realizzare degli studi di caso, che per definizione prevedono un'analisi approfondita dei fenomeni (Yin, 2009), non è stato possibile esplorare tutti i paesi europei e/o quelli coinvolti nella valutazione dei sistemi scolastici e dei risultati negli apprendimenti. Bisognava fare una scelta "geografica" e di ricerca. La selezione finale è ricaduta sulla Finlandia e sulla Spagna, oltre all'Italia, poiché ritenute nazioni che per motivi differenti, illustrati nel primo capitolo della seconda parte, potevano dar

vita a una comparazione significativa. Questo, tuttavia, ne ha escluse altre che per motivi metodologici e tempistico-organizzativi non sono state prese in considerazione. A spirale, all'interno delle realtà educative selezionate, sono stati individuati alcuni esponenti istituzionali che, grazie alla loro eccellente esperienza nel campo della valutazione dei sistemi scolastici e degli apprendimenti, potevano corrispondere agli obiettivi della ricerca. Il punto di forza, ma indubbiamente anche di debolezza di una scelta simile, risiede nella visione che gli intervistati hanno restituito "dal loro punto di vista istituzionale". Per cercare di ovviare a quest'ultimo aspetto è stato costruito e somministrato il questionario a un gruppo di dirigenti scolastici italiani, inserendo l'ampliamento della platea tra i possibili sviluppi futuri della ricerca.

In termini di triangolazione, ovvero di commistione di due o più fonti, tecniche di raccolta dei dati e teorie (Denzin, 1970), le evidenze raccolte tramite l'analisi documentale sono state corroborate dalle riflessioni emerse durante le interviste qualitative e dalle evidenze del questionario. La triangolazione nella ricerca qualitativa, come quella presentata nel presente lavoro, è caratterizzata, infatti, dal rigore nell'espletamento della stessa che le conferisce validità se le procedure seguite e i risultati emersi rispondono alle domande di ricerca e agli obiettivi prefissati (Flick, 2007; Guion et alii, 2011). Per rispondere alla prima domanda di ricerca è stata realizzata l'analisi documentale che ha descritto le pratiche di valutazione dei sistemi scolastici. Per rispondere alla seconda domanda di ricerca sono state condotte le interviste qualitative volte ad esplorare il rapporto tra valutazione, percezioni ad essa legate e impiego delle evidenze da parte delle autorità pubbliche. Per rispondere alla terza domanda di ricerca sull'autovalutazione delle istituzioni scolastiche, quella aggiunta con il sopraggiungere della pandemia, è stata svolta un'analisi documentale per i tre contesti nazionali di studio ed è stato somministrato il questionario rivolto ai dirigenti scolastici nella realtà italiana. La comparazione ha permesso una *environmental triangulation* intesa come l'identificazione di fattori ambientali, nel caso specifico valoriali, che possono influire sul fenomeno oggetto di studio (Guion et alii, 2011). Una serie di limiti temporali, geografici e pandemici non hanno permesso di procedere con la medesima triangolazione delle fonti per tutti e tre i contesti nazionali:

mentre per l'Italia è stato possibile somministrare il questionario con l'obiettivo di indagare con un approccio bottom-up il punto di vista di chi è coinvolto nei processi scolastici e valutativi quotidianamente, per gli altri due paesi non è stato possibile replicare la triangolazione. Per la Finlandia, tuttavia, la triangolazione delle fonti intesa in questo modo è stata impostata, anche se è rimasta in fase embrionale. Durante l'ultima fase del percorso dottorale, infatti, chi scrive ha avuto l'opportunità di incontrare dal vivo un dirigente scolastico finlandese e di visitare la scuola da lui diretta.

Rosario Fina è il primo italiano a dirigere una scuola finlandese. Il primo contatto è avvenuto a seguito della messa in onda del programma Report del 9 gennaio 2023 in cui, in un servizio/reportage, si discuteva della scuola finlandese come uno dei sistemi educativi migliori al mondo. Dal primo scambio di e-mail, ad un successivo incontro telematico, si è giunti ad un incontro dal vivo a Ruovesi. Ruovesi è un comune finlandese di circa 4.000 abitanti situato nella Finlandia Occidentale dove si trova l'istituto comprensivo diretto da Fina che conta poco meno di 240 studenti di scuola primaria e circa 30 insegnanti. Annessa vi è anche una scuola secondaria superiore (liceo con diversi indirizzi). Nella visita è stata osservata una giornata scolastica interagendo con il dirigente scolastico e alcuni insegnanti. Si è trattato di uno *shadowing*, così come definito da Sclavi (2005), "osservando tutto quello che viene in mente per contrasto con la scuola italiana" (Sclavi, 2005).

Se dei pesci intelligenti si proponessero di sviluppare una scienza del loro habitat, probabilmente incomincerebbero analizzando e classificando gli altri pesci, la vegetazione subacquea, le condizioni del fondo marino degli oceani e altro. Ma c'è una cosa che forse solo da ultimo o forse mai studierebbero: l'acqua. Insegnanti, studenti, genitori che si occupano della scuola si trovano in condizioni analoghe: le loro percezioni e analisi si inscrivono entro un più ampio "campo di possibilità" che risulta visibile solo all'occhio o di uno straniero o di chi riesca ad estraniarsi rispetto al carattere ordinario della "vita come è usuale". Ci sono aspetti fondamentali e "troppo ovvi" delle coreografie messe quotidianamente in atto nella scuola, che, come l'acqua per i pesci, sono dati per scontati e sfuggono a chi ci vive dentro. Questo libro propone al lettore, sia americano che italiano, proprio di fare questa esperienza: ri-vivere una

giornata di scuola, guardandola “con occhi nuovi”, attraverso il filtro di aspettative e cornici interpretative diverse, appartenenti “all’altra” cultura. [...] L’idea dello *shadowing* mi è venuta leggendo *Musica per camaleonti*, di Truman Capote. Il libro è una raccolta di brevi racconti narrati in prima persona. In uno di questi l’autore descrive una giornata di *shadowing* dietro la propria donna delle pulizie, nel suo giro quotidiano per gli appartamenti dei clienti. [...] A differenza della “osservazione partecipante”, nello *shadowing* l’osservatore assume anche sè stesso, le proprie emozioni, le proprie abitudini di pensiero, la continua ricerca e contrattazione sulla propria identità come parte fondamentale della dinamica interattiva studiata. Lo *shadowing* mi è sembrato uno strumento da sperimentare in campo sociologico e particolarmente adatto alla ricerca sulla giornata di scuola. (Sclavi, 2005, p. 1-3)

Così come accaduto a Sclavi (2005), durante la visita nella scuola comprensiva di Ruovesi tutto ha iniziato a funzionare “alla rovescia” rispetto alle ricerche tradizionali. Chi scrive non conosceva nel dettaglio cosa avrebbe osservato, ma sapeva che avrebbe sperimentato una realtà nuova. La giornata è iniziata già alle ore 7.30, quando due telefonate di due insegnanti assenti per quel giorno hanno inaugurato il tragitto in macchina dalla stazione ferroviaria di Tampere a Ruovesi. Organizzate le supplenze e scambiate alcune osservazioni con Fina, si è arrivati nel piazzale innevato della scuola. Una volta entrati nell’edificio è stata la volta di alcune presentazioni di rito agli insegnanti presenti nella sala docenti lungo il percorso per la presidenza e di un caffè condiviso. La didattica è iniziata alle ore 8.50 per attendere l’arrivo dello scuolabus che trasporta gratuitamente tutti gli studenti ubicati nel vasto territorio comunale. La prima attività della giornata è stata l’osservazione di una lezione di musica degli studenti di una prima classe della scuola secondaria inferiore intenti, con il loro insegnante, a riadattare una nota canzone finlandese nel testo cantato e nella melodia suonata con gli strumenti musicali (chitarre elettriche, chitarre classiche, bassi, batterie e tastiere) messi a disposizione dall’istituto. Un’ora di lezione e 10 minuti di pausa. Successivamente è stata la volta dell’inaugurazione della festa dei diplomandi, coloro che celebrano l’ultimo giorno di scuola prima dello stop delle lezioni di un mese per lasciare spazio allo studio in vista degli esami di stato della primavera. Terminata la prima parte dei festeggiamenti, chi scrive ha incontrato l’orientatrice scolastica, una

figura professionale non ancora presente nelle scuole italiane, che orienta gli studenti in vista degli studi presenti e futuri e del mondo del lavoro. La pausa pranzo è stata trascorsa nella mensa della scuola insieme agli studenti, mangiando cibo tipico finlandese. Il primo pomeriggio si è aperto con la seconda parte dei festeggiamenti dei diplomandi che, travestiti per l'occasione con costumi carnevaleschi, hanno coinvolto gli insegnanti in momenti celebrativi e commemorativi. Durante l'ultima parte della giornata non è mancata la possibilità di sostare nella sala docenti, di scambiare qualche suggestione con chi era presente e di incontrare Fina che, dopo aver svolto riunioni con le autorità locali e fatto colloqui di lavoro per nuovi insegnanti da immettere in servizio, ha accompagnato chi scrive in visita negli spazi della scuola (il materiale fotografico raccolto è disponibile nell'Appendice III). Contestualmente si ha avuta la possibilità di approfondire alcune tematiche e di realizzare un'intervista.

Fina è arrivato in Finlandia da studente di lingua e letteratura finlandese all'allora Istituto Orientale di Napoli. Il sogno di viaggiare e un percorso accademico che lo ha portato a soggiornare da borsista in Finlandia più di una volta prima della laurea, fino alla scelta decisiva presa già prima della discussione della tesi: tornare per intraprendere una carriera. Dopo le prime collaborazioni con l'IIC di Helsinki ed i corsi ministeriali per i bambini italiani all'estero, iniziano diverse collaborazioni con enti pubblici e privati. Una volta acquisite le abilitazioni necessarie all'insegnamento in Finlandia attraverso percorsi universitari ad hoc, Fina inizia ad insegnare nella scuola pubblica finlandese con una lunghissima parentesi nella scuola europea di Helsinki dove ricopre i primi ruoli amministrativi e di gestione. Quest'esperienza suscita interesse verso la leadership educativa e lo porta a intraprendere il percorso accademico di formazione e abilitazione per i dirigenti scolastici. Questa scelta nasce in primis dalla passione, poi dalle collaborazioni avute con i dirigenti scolastici delle scuole in cui ha insegnato e con i sindacati scolastici. La direzione di una scuola è, secondo lui, sia un'abilità alla quale si viene formati che una competenza che si acquisisce sul campo con gli anni di servizio, l'umiltà e l'esperienza. Tra i "segreti" del sistema scolastico finlandese, infatti, annovera non solo la formazione continua dei docenti e dei dirigenti

scolastici, ma anche il *modus operandi* di questi ultimi che si fanno ricercatori con e dei propri studenti.

Questo percorso è importante, che sia chiaro, perché avere l'esperienza della materia e conoscerla bene non rende insegnante. [...] La didattica e la pedagogia entrano in gioco perché il fatto di essere esperti della materia, secondo me e secondo il sistema finlandese, è l'*entry level*, il minimo. [...] Un'altra cosa importante è il percorso di supplenza che qui inizia presto. Quando sono arrivato come studente laureando non ci ho nemmeno pensato a fare richiesta di supplenza, essendo abituato all'Italia. La discriminante all'inizio del mio percorso lavorativo è stata proprio questa: quando ero già laureato e avevo le carte in regola per poter accedere alla professione... Finché non ho fatto i miei anni di esperienza, un pezzettino alla volta, ho avuto una concorrenza che arrivava già alla laurea con 3 o 4 anni di supplenza. (Rosario Fina, dirigente scolastico finlandese)

La questione dei fondi economici destinati all'educazione interessa tanto i dirigenti scolastici italiani quanto quelli finlandesi.

È vero che il governo Marin ha puntato molto sull'educazione, resta il fatto però che i tagli del governo precedente sono stati grandi, specialmente alla ricerca, oltre che all'istruzione. [...] I fondi di cultura, istruzione e ricerca non si possono tagliare. Sono troppo importanti per la vita di un paese, a mio avviso. Chiaramente è anche vero che il momento economico era ed è ancora delicato. [...] I fondi sono diminuiti, ma c'è la possibilità, attraverso una serie di progetti ministeriali, di accrescere le risorse a disposizione della scuola, ovvero quelle che stanziava il comune che riceve anche dallo stato attraverso una serie di aiuti e finanziamenti vari. Partecipando a questi progetti si hanno risorse ulteriori con le quali sviluppare e portare avanti un istituto. Portare avanti una scuola è sempre una storia di qualche milione di euro, anche in un posto piccolo. Gli stipendi, l'elettricità, la manutenzione dell'immobile... (Rosario Fina, dirigente scolastico italiano)

L'autovalutazione di istituto avviene a livello locale attraverso la collaborazione dei dirigenti scolastici e dei loro team con le autorità municipali, l'autonomia didattica e pedagogica si riflette in quella valutativa.

Ci sono una serie di questionari su vari aspetti statistici che elaborano i vari enti nazionali e che arrivano a te come dirigente scolastico che li compili fornendo dei numeri, dei dati. Questo serve a inviare dati che vanno al ministero dal territorio per poi tornare al territorio. Inoltre, ci sono una serie di questionari fatti all'interno. È sì vero però che c'è grande autonomia. Direi che questa è un'altra parola chiave. Il livello di autonomia è molto alto. [...] Allo stesso modo il ministero dà delle linee guida alle quali tu ti attieni, dalle quali poi il comune tira fuori le sue linee guida che dà alla scuola, la quale a sua volta ha dei regolamenti interni approvati ancora una volta dal comune. C'è questo rimbalzo di "io lo faccio, ti dico come fare, tu lo fai in maniera autonoma e ti fai dire da me se è accettabile che tu lo faccia così". Si è autonomi all'interno di un certo frame. Dunque, la libertà c'è, ma è situata in un frame. Diciamo che è uno spazio di movimento più o meno ampio, non è una libertà assoluta. [...] Facciamo molta autovalutazione con gli studenti. Facciamo autovalutazione anche a livello individuale, è presente proprio nel curriculum scolastico: nella scuola primaria la prima valutazione del semestre è un'autovalutazione fatta insieme all'insegnante. È un qualcosa che esiste ed è proprio del sistema. Ancora una volta anche quella è una ricerca perché si stabiliscono insieme dei criteri che un po' ce li hanno dati attraverso le linee guida e attraverso i criteri di valutazione dell'Agenzia Nazionale per l'Educazione, ma di cui poi abbiamo deciso l'applicazione e la declinazione a livello locale. Intendo anche a livello di unità scolastica di istituto, abbiamo discusso tra di noi e stabilito cosa significhi concretamente dal punto di vista della valutazione nella nostra scuola raggiungere o meno determinati obiettivi. Per questo dobbiamo rilevare come queste cose le stiamo facendo, abbiamo bisogno di dati e di analisi. Quindi ci diamo un feedback analizzando, sviluppando, ascoltandoci, è importante. Porta a una scuola migliore, a una partecipazione migliore nei casi in cui tutti capiscono lo spirito di questo processo e lo attuano. [...] Le persone vengono coinvolte perché quando tu ti autovaluti, le chiami in gioco, fai il questionario con le famiglie, fai incontri con i genitori, cerchi di avere un feedback, raccogli i dati, anche empirici, che danno sostegno o meno rispetto a certe tesi. Se non lo fai, fai male, dirigi male. A volte capita di non farlo, ma di norma il numero di questionari fatti ogni anno è abbastanza elevato. È comunque un momento in cui si riflette insieme su dove si sta andando. (Rosario Fina, dirigente scolastico finlandese)

I questionari che arrivano nelle scuole finlandesi dall'esterno sono elaborati dall'Istituto di Statistica, dal Ministero dell'Educazione e della Cultura e dall'Agenzia Nazionale per l'Educazione per raccogliere dati

su aspetti particolari, si pensi ad esempio all'accoglienza dei bambini ucraini vittime della guerra, alla pandemia, all'organizzazione del supporto. Sulla base dei risultati vengono elaborate le linee guida da diffondere nelle scuole. L'autovalutazione interna, invece, avviene a partire dalle linee guida elaborate centralmente e condivise localmente e viene supportata da una serie di strumenti gratuiti o a pagamento, come le indagini sugli esiti negli apprendimenti in una determinata disciplina acquistabili da KARVI, l'istituto di valutazione scolastica nazionale. Il servizio di consulenza esterna messo a disposizione delle istituzioni scolastiche da KARVI va ad aggiungersi alla valutazione prossimale realizzata quotidianamente dall'insegnante e all'autovalutazione dello studente. Su richiesta, infatti, gli insegnanti possono acquistare una batteria di test disciplinaristici per avere un feedback esterno e, conseguentemente, fornire dati standardizzati a chi si occupa di politiche educative e di elaborazione di linee guida. I risultati delle valutazioni, così come anticipato dal National Project Manager dell'indagine PISA per la Finlandia, sono utilizzati nei momenti decisionali anche a livello micro.

Anche il sindacato dei dirigenti scolastici ci aiuta in questo, ci sono corsi su come utilizzare i dati, su come fare queste cose, appunto su come dirigere un istituto scolastico utilizzando i dati che ci vengono forniti o su come imparare a fare, da dirigente, ricerca all'interno del proprio istituto. Ce lo insegna l'università, se frequenti percorsi universitari, e ce lo insegnano i vari enti che offrono, sia a insegnanti che a dirigenti scolastici e funzionari che si occupano di educazione, corsi a cadenza regolare. Noi studiamo tanto e continuamente. (Rosario Fina, dirigente scolastico finlandese)

Considerando i possibili sviluppi futuri della ricerca e proiettandosi nel lungo termine, sarebbe sicuramente interessante continuare nella direzione già intrapresa con il dirigente scolastico Fina. Un approfondimento del punto di vista dei dirigenti scolastici in altre realtà diverse dall'Italia, porterebbe a un ampliamento delle prospettive raccolte impreziosendo i dati provenienti da un approccio bottom-up che indaga le percezioni legate all'autovalutazione di istituto a partire dal basso, cioè da chi la vive e vi è coinvolto quotidianamente. Un'altra possibile strada da intraprendere, ora che il momento pandemico più

acuto sembra essere superato, è quella che riporta al progetto di ricerca dottorale originario. Includere altri paesi potrebbe essere una pista di ricerca interessante che permetterebbe di estendere la comparazione. Un altro possibile percorso da intraprendere potrebbe prevedere l'approfondimento dei bisogni formativi dei dirigenti scolastici e degli insegnanti. Volutamente è stata tralasciata la valutazione di questi ultimi, anche se è un aspetto attuale di cui al momento si sta dibattendo molto (OECD, 2013), soprattutto nel contesto italiano in cui si sta andando incontro a una riforma della formazione iniziale dei docenti. La valutazione di dirigenti scolastici e insegnanti affiancata alle competenze professionali valutative richieste a chi lavora all'interno della scuola potrebbe essere un ulteriore sviluppo futuro della ricerca iniziata durante il percorso di dottorato. In realtà, l'ultimo punto è stato già in parte approfondito da chi scrive grazie al lavoro svolto negli ultimi anni come Collaboratrice Tecnica presso l'Invalsi. Il progetto PON Valu.E (Valutazione/Autovalutazione Esperta delle scuole), in cui si ha avuto la possibilità di collaborare, ha avuto l'obiettivo di riflettere su quale tipo di competenze costruire per fare sì che il processo di valutazione del sistema scolastico italiano, così come delineato dal Sistema Nazionale di Valutazione, possa esplicarsi, ingaggiando le istituzioni scolastiche in un percorso oneroso ma che ha come obiettivo ultimo il miglioramento. Attraverso il progetto pluriennale coordinato da Invalsi e finanziato dai fondi europei, sono stati colti alcuni stimoli provenienti dal contesto nazionale e internazionale. Il disegno della ricerca si è articolato in tre azioni: valutare la valutazione, sostenere l'autovalutazione delle scuole e delineare le competenze decentrate per la valutazione. La prima azione è stata finalizzata al miglioramento degli strumenti e dei protocolli, la seconda alla sperimentazione empirica di modelli innovativi di valutazione e autovalutazione delle scuole, mentre la terza, ancora in corso, è stata finalizzata all'implementazione di un ambiente di apprendimento virtuale dedicato a insegnanti, dirigenti scolastici e valutatori esperti precedentemente selezionati da Invalsi. Il percorso formativo Valu.Elearn, frutto di quest'ultima azione, si è svolto all'interno dell'ambiente virtuale di apprendimento Moodle, ha accolto circa 500 partecipanti e ha permesso loro di interagire all'interno della piattaforma visionando un pacchetto di videolezioni registrate da

docenti accademici, svolgendo un test di autovalutazione sui contenuti proposti, svolgendo alcune *e-tivities* con feedback collettivo di un tutor universitario e conducendo una riflessione finale (Romiti, Fabbro, Mattarelli, 2022). La necessità di sviluppare competenze decentrate per l'espletamento dei processi di valutazione delle istituzioni scolastiche ha portato all'elaborazione di un progetto che si è svolto orizzontalmente con un investimento sugli attori sociali della scuola. La relazione tra valutazione e miglioramento, due dei passi descritti all'interno del D.P.R n. 80 del 2013, non è di tipo causale, tuttavia il coinvolgimento degli attori e la conoscenza e la comprensione degli strumenti, dei fini e degli usi possono essere fattori che promuovono il miglioramento e la qualità a seguito della valutazione (Paloma et alii, 2020; Poliandri et alii, 2022; Romiti, Fabbro, Mattarelli, 2022; Romiti, Fabbro, Mattarelli, 2023).

La questione della qualità è dibattuta anche a livello universitario. Da qualche decennio a questa parte sia gli atenei italiani che esteri si stanno interrogando su quale tipo di formazione offrire agli studenti iscritti e come valutare non solo quest'ultima, ma anche il resto dei servizi offerti. Sul finire del secolo scorso, infatti, è stato istituito il Processo di Bologna, una rete di università nazionali e internazionali riunite con l'obiettivo di collaborare alla riforma dell'istruzione terziaria (Perotti, 2010; Del Gobbo, 2018). Similmente al versante scolastico, anche in quello universitario gli esperti hanno a lungo dibattuto su come potesse essere declinata la qualità dell'istruzione, cosa si potesse migliorare e come si potesse valutare (Binetti & Cinque, 2016).

L'esperienza della crisi e la consapevolezza che in gran parte potesse essere attribuita a una insufficienza del sistema formativo di alta qualità, nel settembre 2003 spinse i ministri dei paesi aderenti al processo di Bologna, tra i quali l'Italia, a chiedere all'ENQA (*European Network for Quality Assurance in Higher Education*), il network europeo che raggruppa le agenzie nazionali di valutazione e accreditamento, di sviluppare delle linee guida che implementassero il livello della qualità universitaria. Si trattava di elaborare una serie di criteri, strutturati sotto forma di linee guida, che consentissero non solo di proporre iniziative per migliorare la qualità educativa delle università, ma di fare un passo ulteriore per assicurare concretamente un salto di qualità alle istituzioni universitarie e alle agenzie preposte alla loro verifica esterna. (Binetti & Cinque, 2016, p. 7-8)

Le linee guida sono state approvate nel 2005 e sono state inserite all'interno del documento *Standards and Guidelines for Quality Assurance in the European Higher Education Area* (ENQA, 2015). Per l'elaborazione è stato seguito un lavoro di confronto tra paesi in una visione *unitaria* (Binetti & Cinque, 2016). Oggi l'ENQA, divenuta *European Association for Quality Assurance in Higher Education*, continua ad avere il compito di migliorare la qualità dell'istruzione terziaria assicurando il corretto funzionamento del processo di assicurazione della qualità in tutti i paesi (ENQA, 2021).

Per l'Italia risponde della valutazione della qualità dell'istruzione superiore l'ANVUR, Agenzia Nazionale di Valutazione del Sistema Universitario e della Ricerca, l'ente pubblico che, sotto la sorveglianza del ministero, ha il compito di valutare periodicamente gli istituti di istruzione e formazione terziaria. L'agenzia è nata nel 2006, dal 2019 è parte dell'ENQA e tra le sue attività principali annovera l'ASN, Abilitazione Scientifica Nazionale dei professori universitari, l'AVA, Autovalutazione, Valutazione e Accredimento delle istituzioni universitarie, e la VQR, Valutazione della Qualità della Ricerca (Mattarelli & Zanazzi, 2021). L'AVA, guidata dai principi di indipendenza, imparzialità, professionalità, riservatezza e trasparenza che caratterizzano l'operato dell'ANVUR (ANVUR, 2014), è un processo mirato a garantire la qualità e l'efficienza accademica sulla base di un modello di valutazione che si fonda su criteri definiti *ex ante* (D.L. 240/10, 2010). Così come avviene per le istituzioni scolastiche italiane tramite i processi di autovalutazione e valutazione esterna, l'Assicurazione della Qualità, AQ, garantisce attraverso verifiche condotte da personale esterno all'ateneo la correttezza e l'adeguatezza dei processi accademici. Il secondo ciclo del modello AVA, nello specifico AVA 3, si basa su una struttura che ha come variabili principali di indagine la *qualità della didattica e dei servizi agli studenti*, *l'assicurazione della qualità* e la *qualità della ricerca e della terza missione/impatto sociale* e che ha come destinatari principali i corsi di laurea, le scuole di dottorato e i dipartimenti (ANVUR, 2022ab). Affinchè possa aver luogo *l'accreditamento iniziale e periodico delle sedi e dei corsi di studio*, durante le visite esterne l'ateneo deve possedere alcuni requisiti che sono sintetizzabili attraverso indicatori: per ogni variabile oggetto di

monitoraggio, ovvero gli ambiti di valutazione, infatti, sono indicati punti di attenzione, aspetti da considerare e indicatori (ANVUR, 2022b). Come si può notare, la valutazione italiana del sistema universitario fa proprie componenti più avanzate. Le modalità di certificazione della qualità, difatti, pur poggiandosi sempre sulla dicotomia autovalutazione-valutazione esterna, sono diverse rispetto a quelle dei livelli di istruzione e formazione scolastici (Damiani et alii, 2016). Se in Spagna è redatto un sistema nazionale di indicatori per la scuola dell'obbligo, *Sistema Estatal de Indicadores de la Educación* (Ministerio de Educación y Formación Profesional, 2022), in Italia questo sistema nazionale non è stato ancora implementato a livello scolastico. Al contrario, a livello universitario la messa a punto di indicatori è costante e avviene attraverso il lavoro svolto dall'ANVUR (ANVUR, 2014, 2022ab; Rumiati et alii, 2018; Rumiati et alii, 2019). Spostandosi sul livello terziario, questa potrebbe rappresentare un'ulteriore pista di approfondimento futuro del lavoro di ricerca dottorale presentato.

Appendice I: Esempio di traccia di intervista ai National Project Manager dell'indagine PISA dell'OCSE

INTERNATIONAL AND NATIONAL LEARNING OUTCOMES ASSESSMENT

1. Analyzing PISA 2018 results, it emerges that Spain scored below the OECD average in both mathematics and science. These data are different from those reported by the IEA-TIMSS and IEA-PIRLS studies in which younger students achieve better school performance. In your opinion, what is the reason for this difference in results?
2. Are the differences between Autonomous Communities, between male and female students and related to socio-economic status (eg. natives vs immigrants students) highlighted by PISA results also confirmed by national assessments? How do you think an education system can mitigate these differences?
3. We know from several studies that the family and cultural background of the young person and the preparation of the teacher affect the student's results. What contribution does your work make in trying to reduce the strong impact of family and cultural background on student achievement?

PERCEPTIONS ABOUT EVALUATION

1. What do you think and how do you welcome the ideological criticisms of the OECD-PISA programme? In Spain are there also criticisms against the national assessment (eg. Evaluacion de tercer/sexta curso de Educacion Primaria)?
2. In your opinion, how is external evaluation seen by those who are involved in it?

Appendice

3. Do you think that this assessment and evaluation systems can have negative "side effects" on teachers' behavior? Do they have clear the purpose of the evaluation system?
4. Feeling an active part of a process is undoubtedly a factor that promotes collaboration and interest of those who live it. In what way and through which mechanisms, in your opinion, does this model of system evaluation involve the assessed in the evaluation processes?
5. How is the information coming from the external evaluation received by the central institutions - Ministry of Education? In your opinion, is the system assessment and evaluation linked to improvement interventions?
6. In Spain, how is the presence of INEE perceived from your point of view by school principals, teachers, students and families?
7. Do you think that schools use the data from the system evaluation to make decisions?

GOVERNMENT'S FUNDINGS

1. From the data published annually by the OECD (*Education at a Glance 2020*) it appears that in 2017 Spain invested just over 4% of national GDP in education in line with the average of European countries. At an economic level, what is the budget available to INEE? Does it receive funds from the state?

SELF-ASSESSMENT AND EVALUATION

1. In most cases, the external evaluation of educational system is accompanied by the self-evaluation of schools as a moment of reflection and improvement. Is it mandatory in your country? How often does it happen?

COVID-19 PANDEMIC

1. How are the system assessment studies, like the OECD-PISA programme, including the important variable of the pandemic? In such a delicate scenario, what, in your opinion, can be the contribution of the system evaluation for restarting?

Appendice II: Questionario rivolto ai dirigenti scolastici italiani

Questionario-intervista sui processi di autovalutazione di istituto somministrato attraverso Google Moduli e telefonicamente a chi lo ha richiesto.

Gentile Dirigente Scolastico,

sono Eleonora Mattarelli, dottoranda di ricerca presso la scuola di Dottorato in Psicologia Sociale, dello Sviluppo e Ricerca Educativa della Sapienza, Università di Roma. Nell'ambito di un progetto di ricerca dottorale comparativo-internazionale sto conducendo un'indagine sull'autovalutazione delle scuole, supervisionata dal professore Giorgio Asquini.

Dopo aver preso in considerazione le pratiche e le percezioni relative alla valutazione del sistema scolastico in tre importanti realtà educative europee quali Italia, Finlandia e Spagna, l'obiettivo ora è quello di indagare la visione dei Dirigenti Scolastici Italiani dell'autovalutazione scolastica, attraverso un questionario-intervista.

Il Suo punto di vista è importante per comprendere i processi di autovalutazione attuati nella scuola che dirige. Pertanto, è importante rispondere con la massima sincerità facendo riferimento alla sua scuola e alla sua esperienza personale.

Tutti i dati verranno raccolti in modo anonimo, nel rispetto del Regolamento generale UE sulla protezione dei dati (General Data Protection Regulation, GDPR) n. 679/2016, e al D.Lgs. n. 196/2003 "Codice in materia di protezione dei dati personali" (modificato dal D.Lgs. n. 101 del 10/08/2018). I dati saranno trattati in forma aggregata. Nessun questionario-intervista sarà riconducibile ai singoli partecipanti

alla ricerca. Le analisi dei dati hanno uno scopo conoscitivo rispetto al tema della ricerca e non hanno alcuno scopo diagnostico o valutativo.

I risultati del questionario-intervista verranno comunicati e pubblicati in sedi scientifiche, garantendo comunque la non riconoscibilità personale di chi ha partecipato alla ricerca.

Se è interessato/a a rimanere informato/a sugli esiti della ricerca può mettersi in contatto con me in qualità di Referente per il progetto, ed è anche possibile richiedere informazioni e porre qualsiasi domanda concernente l'indagine: dott.ssa Eleonora Mattarelli (eleonora.mattarelli@uniroma1.it).

La ringrazio per la collaborazione.

Dott.ssa Eleonora Mattarelli

Cliccando sul tasto "acconsento" che segue attesta che:

1. Ha ricevuto e letto le informazioni sullo scopo e le procedure della ricerca;
2. Dà il consenso a partecipare alla ricerca.

Ai sensi del Decreto legislativo n. 101 del 10/08/2018, dopo aver letto la parte informativa su scopi e procedura della ricerca, esprimo il consenso a partecipare.

- Acconsento

SEZIONE A

A1. Età (in cifre)

.....

A2. Genere

- Uomo
- Donna
- Preferisco non specificarlo

A3. Provincia di residenza

.....

A4. L'istituto in cui attualmente presta servizio è

- Istituto Comprensivo
- Scuola primaria
- Scuola Secondaria di I Grado
- Scuola Secondaria di II Grado
- Centro Provinciale per l'Istruzione degli Adulti (CPIA)
- Altro

A5. Provincia di riferimento dell'istituto

.....

A6. Nell'istituto in cui attualmente presta servizio è

- Titolare dell'incarico
- Reggente

A7. Da quanti anni dirige l'istituto in cui attualmente presta servizio?

.....

A8. Da quanti anni ricopre l'incarico di dirigente scolastico?

.....

SEZIONE B

B1. Secondo lei i processi di autovalutazione d'istituto es. compilazione del Rapporto di Autovalutazione, così come riportati nel Sistema Nazionale di Valutazione, sono sostenibili in relazione alle risorse umane e finanziarie disponibili e alle tempistiche richieste?

- Sì
- No

- Soltanto in parte

B1.1. Se, secondo lei, i processi di autovalutazione d'istituto sono sostenibili soltanto in parte, può spiegarne il motivo

.....

B2. Nell'istituto che dirige quale organizzazione è stata messa a punto per la redazione del Rapporto di Autovalutazione (es. chi elabora, chi approva, chi pubblica il RAV)?

.....

B3. La compilazione del Rapporto di Autovalutazione è vissuta da lei come

- Un adempimento formale
- Uno strumento funzionale per la riflessione e il miglioramento della scuola
- Entrambe le precedenti opzioni di risposta
- Altro

B4. Quanto, secondo lei, gli indicatori proposti nel Rapporto di Autovalutazione sono adeguati per descrivere la situazione del singolo e specifico contesto scolastico?

- 1 – per nulla
- 2
- 3
- 4
- 5 – moltissimo

B5. Quanto, secondo lei, i processi di autovalutazione d'istituto e la compilazione del Rapporto di Autovalutazione sono utili nell'ottica del miglioramento della scuola?

- 1 – per nulla
- 2

- 3
- 4
- 5 – moltissimo

B5.1. Può motivare la risposta alla domanda precedente?

.....

B6. Nell'istituto che dirige si utilizzano altre forme non strutturate di autovalutazione d'istituto?

- Sì
- No

B6.1. Se sì, come sono organizzate?

- Consigli di classe
- Consigli di istituto
- Dipartimenti e aree disciplinari
- Amministrazione
- Altro

B7. Nel corso della sua carriera ha mai partecipato a momenti di formazione sull'autovalutazione d'istituto?

- Sì
- No

B8. Sente di voler essere maggiormente formato sui temi dell'autovalutazione d'istituto?

- Sì
- No
- Altro

B8.1. Se sì, secondo lei quali argomenti sono più urgenti e utili?

.....

B9. Gli esiti dei processi di autovalutazione d'istituto vengono diffusi e condivisi nella sua scuola (tra docenti, genitori e studenti)?

- Sì
- No

B9.1. Se sì, in che modo?

.....

B10. Quando si trova a prendere decisioni per la sua scuola, su queste ultime influiscono gli esiti dei processi di autovalutazione?

- 1 – per nulla
- 2
- 3
- 4
- 5 – moltissimo

B10.1. Se prende decisioni per la sua scuola anche sulla base degli esiti dei processi di autovalutazione, può argomentare meglio la sua risposta fornendo degli esempi concreti?

.....

B11. Quanto, secondo lei, i processi di autovalutazione nella sua scuola sono efficaci ed efficienti (es. hanno delle ricadute concrete)?

- 1 – per nulla
- 2
- 3
- 4
- 5 – moltissimo

B11.1. Se hanno delle ricadute concrete, può riportare un esempio?

.....

B12. Se vuole aggiungere qualcos'altro sulla sua personale percezione o esperienza rispetto all'autovalutazione d'istituto, può scrivere liberamente nello spazio sottostante.

.....

Appendice III: Reportage fotografico dell'Istituto Comprensivo di Ruovesi (Finlandia)

Fotografie scattate durante la visita all'Istituto Comprensivo di Ruovesi.



Figura 1. Istituto Comprensivo di Ruovesi, veduta dall'esterno. Fonte propria.



Figura 2. Sala docenti dell'Istituto Comprensivo di Ruovesi. Fonte propria.

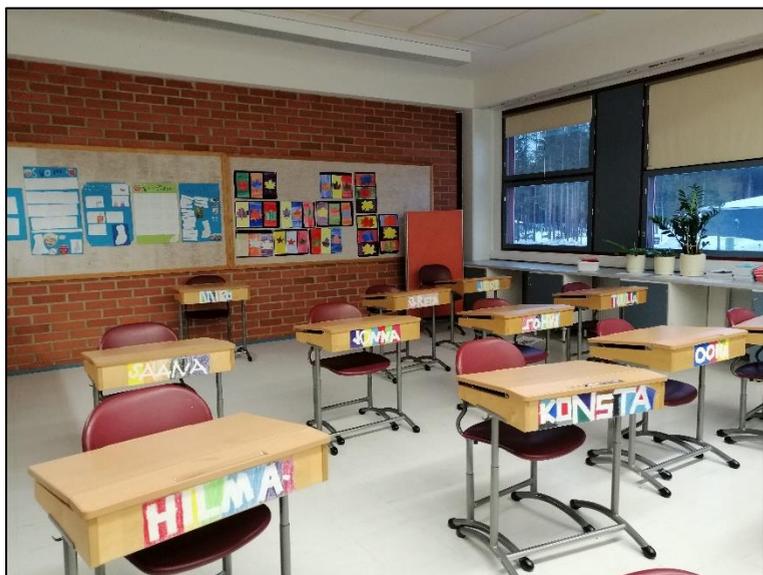


Figura 3. Aula della scuola prima dell'Istituto Comprensivo di Ruovesi. Fonte propria.



Figura 4. Aula di musica dell'Istituto Comprensivo di Ruovesi. Fonte propria.

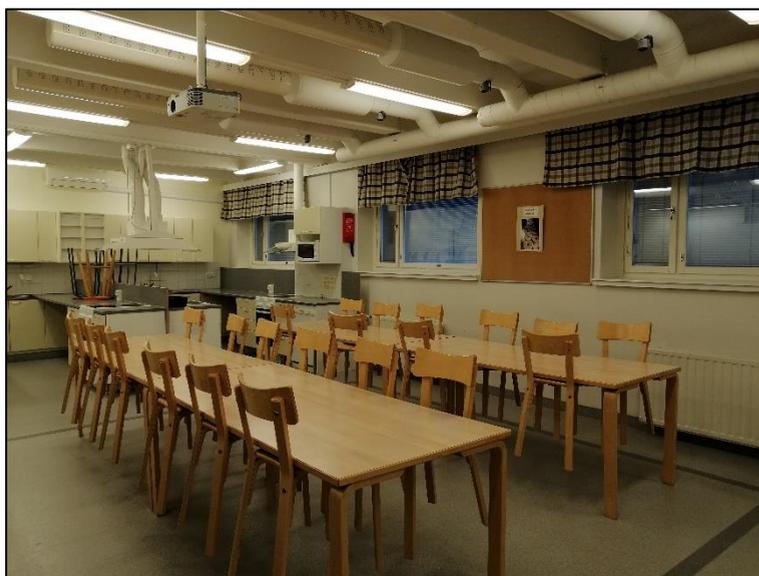


Figura 5. Aula di cucina dell'Istituto Comprensivo di Ruovesi. Fonte propria.



Figura 6. Aula di cucito dell'Istituto Comprensivo di Ruovesi. Fonte propria.

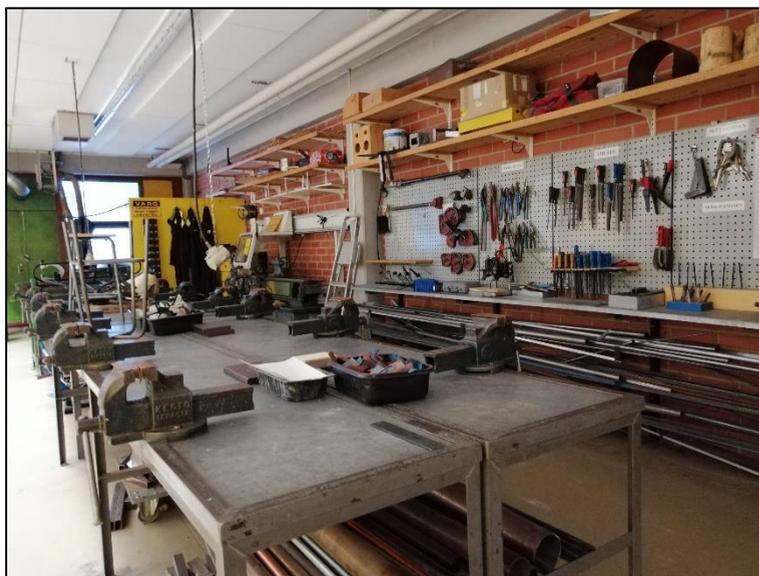


Figura 7. Aula di lavorazione del ferro dell'Istituto Comprensivo di Ruovesi. Fonte propria.



Figura 8. Aula per la lavorazione del legno dell'Istituto Comprensivo di Ruovesi. Fonte propria.



Figura 9. Decorazione murale nell'Istituto Comprensivo di Ruovesi. Fonte propria.



Figura 10. Corridoio con armadietti dell'Istituto Comprensivo di Ruovesi. Fonte propria.



Figura 11. Mensa scolastica dell'Istituto Comprensivo di Ruovesi. Fonte propria.



Figura 12. Dispositivi tecnologici nell'Istituto Comprensivo di Ruovesi. Fonte propria.



Figura 13. Decorazione murale nell'Istituto Comprensivo di Ruovesi. Fonte propria.

Bibliografia

- Adams, R. J. (2003). Response to 'Cautions on OECD's Recent Educational Survey (PISA)'. *Oxford Review of Education*, 377-389.
- Agasisti, T., Falzetti, P., & Freddano, M. (2015). L'uso dei risultati delle Rilevazioni Nazionali per l'autovalutazione delle scuole. *Rassegna Italiana di Valutazione*, 19(61), 28-48.
- Agnelli, F. G. (2014). La valutazione della scuola: A che cosa serve e perché è necessaria all'Italia. Gius. Laterza & Figli Spa.
- Agrusti, F. (2010). Come la tecnologia qualifica le interazioni. In Vertecchi, B. (a cura di), *Orbis Dictus. Un ambiente adattivo multilingue per l'istruzione in rete* (pp. 65-79). MILANO: FrancoAngeli.
- Agrusti, F. (2014). I dati valutativi italiani dal 1971 ad oggi: progettazione e realizzazione di un archivio inferenziale longitudinale (Vol. 11). Edizioni Nuova Cultura.
- Agrusti, F., & Corsini, C. (2021). Indagine esplorativa sugli esami universitari a distanza. *CADMO*, (2021/1).
- Agrusti, F., Kramer, B., Mileva, N. (2008). Research methodology and approach. In *The impact of technology on distance learning students* (pp. 16-27). DUBLINO: Ericsson.
- Aguilar, L. E. (2018). Three Models of Comparative Analysis: Time, Space, and Education. *Annual Review of Comparative and International Education* 2017.
- Alcaide, A. M. (2021). La Inspección educativa en la LOMLOE. Proyecto, enmiendas y ley. Significación de los cambios. *Avances en Supervisión Educativa*, (35).
- Alexander, R., Broadfoot, P., & Phillips, D. (Eds.). (1999). *Learning from Comparing: new directions in comparative education research: Volume 1: contexts, classrooms and outcomes*. Symposium Books Ltd.
- Allmendinger, J., & Leibfried, S. (2003). Education and the welfare state: the four worlds of competence production. *Journal of European social policy*, 13(1), 63-81.
- Allodola, V. F. (2014). Metodi di ricerca qualitativa in Medical Education: approcci, strumenti e considerazioni di rigore scientifico. *Education Sciences & Society*, 5(1).
- Allport, F. H., (1955). *Theories of Perception and the Concept of Structure*. New York: Wiley.
- Allulli, G. (2007). La valutazione della scuola: un problema di governance. *Economia dei servizi*, 2(3), 453-470.
- Allulli, G. (2010). *Dalla strategia di Lisbona a Europa 2020. Politiche Europee della Formazione*, Roma.

- Allulli, G. (2014). Una valutazione in cerca di strategia. *Scuola democratica*, (1), 215-222.
- Altheide, D., Coyle, M., DeVriese, K., & Schneider, C. (2008). Emergent qualitative document analysis. *Handbook of emergent methods*, 2008, 127-151.
- Ammermüller, A. (2004). PISA: What makes the difference? Explaining the gap in PISA test scores between Finland and Germany, 04-004.
- Ammermüller, A. (2005). Poor background or low returns? Why immigrant students in Germany perform so poorly in PISA, 05-018.
- Anderson, R. A., Crabtree, B. F., Steele, D. J., & McDaniel Jr, R. R. (2005). Case study research: The view from complexity science. *Qualitative health research*, 15(5), 669-685.
- Angrosino, M. V., & Mays de Pérez, K. A. (2000). Rethinking observation: From method to context. *Handbook of qualitative research*, 2, 673-702.
- ANVUR (2014). Codice etico. Retrived from: https://www.anvur.it/wp-content/uploads/2014/10/Codice_etico_Anvur2014.pdf.
- ANVUR (2022a). Accreditamento periodico delle sedi e dei corsi di studio universitari. Retrived from: <https://www.anvur.it/wp-content/uploads/2022/06/AVA-3-Modello-2022-05-26.pdf>.
- ANVUR (2022b). Modello di accreditamento periodico delle sedi e dei corsi di studio universitari. Retrived from: https://www.anvur.it/wp-content/uploads/2023/02/AVA3_Requisiti_2022_09_08-1.pdf.
- Araujo, L., Saltelli, A., & Schnepf, S. V. (2017). Do PISA data justify PISA-based education policy?. *International Journal of Comparative Education and Development*, 19(1), 20-34.
- Argentin, G. (2016). Non è un Paese per l'evidenza... e pur si muove. Peculiarità, limiti e potenzialità delle sperimentazioni controllate nel sistema scolastico italiano, in Landri, P., Maccarini, M. (ed.), *Uno specchio per la valutazione della scuola. Paradossi, controversie, vie d'uscita*, Franco Angeli, Milano, pp. 72- 95, [<http://hdl.handle.net/10807/99371>].
- Arias, P. (2020). *Análisis resultados TIMSS 2019*. Santiago de Chile: Fundación acciónEDUCAR.
- Arnove, R. F. (2001). Comparative and International Education Society (CIES) facing the twenty-first century: Challenges and contributions. *Comparative Education Review*, 45(4), 477-503.
- Arriazu, R., & Solari, M. (2015). The role of education in times of crisis: A critical analysis of the Europe 2020 Strategy. *KEDI Journal of Educational Policy*, 12(2).
- Asquini, G. (2007). Che cosa ci dicono le indagini internazionali sull'istruzione. *Economia & lavoro*, 41(1), 57-0.
- Asquini, G. (2018). Non solo pezzi di carta. Gli esiti sociali dell'istruzione. Corsini, C. (a cura di) *Rileggere Visalberghi*, 57-67.

- Atkinson, P. A. & Coffey, A. (1997). Analysing documentary realities. In D. Silverman (Ed.), *Qualitative research: Theory, method and practice*, London: Sage, 45-62.
- Baird, J. A., Johnson, S., Hopfenbeck, T. N., Isaacs, T., Sprague, T., Stobart, G., & Yu, G. (2016). On the supranational spell of PISA in policy. *Educational Research*, 58(2), 121-138.
- Baird, J., Isaacs, T., Johnson, S., Stobart, G., Yu, G., Sprague, T., & Daugherty, R. (2011). *Policy effects of PISA*. Oxford University Centre for Educational Assessment.
- Ballantine, J. H., Stuber, J., & Everitt, J. G. (2021). *The sociology of education: A systematic analysis*. Routledge.
- Ballou, D. (2001). Pay for performance in public and private schools. *Economics of education review*, 20(1), 51-61.
- Barbier, J.M. (1989), *La valutazione nel processo formativo*, Torino, Loescher.
- Barbieri, N. S., Gaudio, A., & Zago, G. (Eds.). (2016). *Manuale di educazione comparata: insegnare in Europa e nel mondo*. La Scuola.
- Barbour, R. S. (2014). Quality of data analysis. *The SAGE handbook of qualitative data analysis*, 496-510.
- Barzanò, G. (Ed.). (2000). *L'autovalutazione nella scuola. Teorie, strumenti, esempi*. Pearson Italia Spa.
- Bassey, M. (1999). *Case study research in educational settings*. McGraw-Hill Education (UK).
- Batini, F., Barbisoni, G., Pera, E., Toti, G., Sposetti, P., Szpunar, G., ... & Scipione, L. (2020). Un modello di analisi delle domande aperte nell'indagine nazionale SIRD sulla didattica a distanza durante l'emergenza Covid-19. *RicercaAzione*, 12(2), 47-72.
- Battista, M. T. (1999). Research, Reflection, Practice: Geometry results from the third international mathematics and science study. *Teaching Children Mathematics*, 5(6), 367-373.
- Baxter, J., & Clarke, J. (2013). Farewell to the tick box inspector? Ofsted and the changing regime of school inspection in England. *Oxford Review of Education*, 39(5), 702-718.
- Baxter, P., & Jack, S. (2008). Qualitative Case Study Methodology: Study Design and Implementation for Novice Researchers. *The Qualitative Report*, 13(4), 544-556.
- Beaton, A. E. (1996a). *Mathematics Achievement in the Middle School Years. IEA's Third International Mathematics and Science Study (TIMSS)*. Boston College, Center for the Study of Testing, Evaluation, and Educational Policy, Campion Hall 323, Chestnut Hill, MA 02167.
- Beaton, A. E. (1996b). *Science Achievement in the Middle School Years. IEA's Third International Mathematics and Science Study (TIMSS)*. Boston College, Center

- for the Study of Testing, Evaluation, and Educational Policy, TIMSS International Study Center, Champion Hall 323, Chestnut Hill, MA 02167.
- Beaton, A. E., & Robitaille, D. F. (1999). An overview of the third international mathematics and science study. *International comparisons in mathematics education*, 11, 30-47.
- Bechger, T. M., van Schooten, E., De Gloppe, C., & Hox, J. J. (1998). The validity of international surveys of reading literacy: The case of the IEA reading literacy study. *Studies in Educational Evaluation*, 24(2), 99-125.
- Bechis, M., & Tomatis, L. (2014). Strumenti per la responsabilità sociale in ambito scolastico. *La Tecnica della Scuola* (23-24), 124-125.
- Beeby C.E. (1977), *The meaning of evaluation*, «Current Issues in Education», vol. 4, pp. 68-78.
- Bell, T. H. (1993). Reflections one decade after a nation at risk. *Phi Delta Kappan*, 74(8), 592-597.
- Benadusi, L. (2009). *Lo stato dell'arte dell'autonomia scolastica*. Programma Education FGA.
- Benadusi, L. (2014). Indicatori di efficacia e di equità in prospettiva comparativa. In C. Roverselli (a cura di), *Comprendere l'alterità. Comparazione, intercultura, Europa* (pp. 173-194). Roma: Aracne.
- Benvenuto, G. (2008). La docimologia: dalla critica agli esami all'analisi dei principi della valutazione. In G. Benvenuto & A. Giacomantonio (Eds.), *Un po' di storia della valutazione scolastica: letture e riflessioni* (pp. 1-6). Nuova Cultura.
- Benvenuto, G. (2015). *Stili e metodi della ricerca educativa*. Carocci.
- Benvenuto, G., & Giacomantonio, A. (2008). *Un po' di storia della valutazione scolastica: letture e riflessioni*. Roma: Centro Stampa Nuova Cultura.
- Benvenuto, G., Consoli, G., & Fattorini, O. (2019). Luci e ombre nei processi di elaborazione del RAV (Rapporto di Autovalutazione): un'indagine sul punto di vista dei Dirigenti scolastici. In P. Lucisano & A.M. Notti, *Training actions and evaluation processes*. Atti del Convegno Internazionale SIRD, 163-173.
- Berliner, D. C. & Biddle, B. J. (1995). *The Manufactured Crisis: Myths, Fraud, and the Attack on America's Public Schools*. Addison Wesley Publishing Company.
- Bezzi C. (2000), *La valutazione dei servizi alla persona*. Perugia, Giada.
- Bianchini, P. (2010). *Le origini delle materie. Discipline, programmi e manuali scolastici in Italia* (Vol. 1, pp. 1-257). Sei.
- Bichi, R. (2002). *L'intervista biografica. Una proposta metodologica*. Milano: Vita e pensiero.
- Binetti, P., & Cinque, M. (2016). *Valutare l'Università & Valutare in Università. Per una "cultura della valutazione"*. Milano: FrancoAngeli.
- Binkley, M. R., & Williams, T. (1996). *Reading literacy in the United States: Findings from the IEA reading literacy study*. US Department of Education, Office of

- Educational Research and Improvement, National Center for Education Statistics.
- Black P., William D. (1998). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan*, 80(2): 139-148.
- Blake, D. W. (1982). The purposes and nature of comparative education: the contribution of IL Kandel. *Comparative Education*, 18(1), 3-13.
- BOCM (2021). Plan General Plurianual de actuación de la Inspección Educativa. Cursos 2021-2025. Retrived from: https://www.bocm.es/boletin/CM_Orden_BOCM/2021/08/26/BOCM-20210826-4.PDF.
- Boja (2019). Plan General de Actuación 2019-2023. Retrived from: https://www.juntadeandalucia.es/boja/2019/145/BOJA19-145-00030-11604-01_00159874.pdf.
- Bonaiuti, G., Calvani, A., & Picci, P. (2012). Tutorship e video annotazione: il punto di vista degli insegnanti/Tutorship and video annotation: teachers' perspectives on innovation. *Italian Journal of Educational Research*, V, 246–258.
- Bondioli, A. (Ed.). (2000). *Manuale di valutazione del contesto educativo. Teorie, modelli, studi per la rilevazione della qualità nella scuola* (Vol. 51). FrancoAngeli.
- Bondioli, A., & Savio, D. (2021). Pandemia e realtà educative 0-6: uno studio esplorativo sulla "ripartenza". *ITALIAN JOURNAL OF EDUCATIONAL RESEARCH*, 138-149.
- Boren, T., & Ramey, J. (2000). Thinking aloud: Reconciling theory and practice. *IEEE transactions on professional communication*, 43(3), 261-278.
- Borri, A., & Bovi, O. (2007). *Educazione comparata* (Vol. 7). Morlacchi Editore.
- Boshier, R. (1990). Socio-psychological factors in electronic networking. *International Journal of Lifelong Education*, 9(1), 49-64.
- Bottani N., & Tuijnman, A. (1994). International Education Indicators: Framework, Development and Interpretation. In Ceri-Oecd, *Making Education Count. Developing And Using International Indicators*. Paris: Oecd Publishing. 21:35. (Ed.It. *Valutare L'insegnamento*, Roma, Armando Editore, 1994, 27:42).
- Bottani, N. & Pegoraro, R. (2005). La situation de l'enseignement secondaire dans les pays de l'OCDE. In N. Bottani, C. Magnin, & E. Zotto (Eds.). *L'enseignement secondaire à l'échelle mondiale: bilans et perspectives. Actes du colloque de Genève organisé à l'occasion de la 47ème Conférence internationale de l'éducation du BIE*. Genève, SRED.
- Bottani, N. & Vrignaud, P. (2005): *La France et les évaluations internationales*. Haut Conseil de l'évaluation de l'école, Quaderno.16, pp. 167, Ministère de l'Education Nationale, Paris.

- Bottani, N. (1990), The Background of The Ceri/Oecd Project on International Educational Indicators. *International Journal of Educational Research*, 14-4, pp. 335-342.
- Bottani, N. (2003). Le indagini comparate internazionali sul profitto scolastico. Cinquant'anni di storia che hanno cambiato il mondo dell'educazione. In N. Bottani & A. Cenerini (Eds.), *Una pagella per la scuola. La valutazione tra autonomia e equità*. (pp. 265-288). Gardolo-Trento, Edizioni Erickson.
- Bottani, N. (2009). Il difficile rapporto tra politica e ricerca scientifica sui sistemi scolastici. Fondazione Giovanni Agnelli, Working Paper n. 17.
- Bottani, N., & Tuijman, A. (1990). Indicatori internazionali dell'educazione: struttura, sviluppo e interpretazione. OCSE, *Valutare l'insegnamento*.
- Bowen, G. A. (2009). Document analysis as a qualitative research method. *Qualitative research journal*.
- Bracey, G. (2005). Put out over PISA. *Phi Delta Kappan*, 86(10), 797.
- Bracey, G. (2007). The First Time 'Everything Changed': The 17th Bracey Report on the Condition of Public Education. *Phi Delta Kappan*, Vol. 89, No. 02, 119-136.
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative research in psychology*, 3(2), 77-101.
- Braun, V., & Clarke, V. (2019). Reflecting on reflexive thematic analysis. *Qualitative research in sport, exercise and health*, 11(4), 589-597.
- Bray, M., & Thomas, R. M. (1995). Levels of comparison in educational studies: Different insights from different literatures and the value of multilevel analyses. *Harvard educational review*, 65(3), 472-491.
- Bray, M., Adamson, B., & Mason, M. (2009). *Educazione comparata. Approcci e metodi di ricerca*. FrancoAngeli.
- Bray, M., Adamson, B., & Mason, M. (2014). Different models, different emphases, different insights. In *Comparative education research* (pp. 417-436). Springer, Cham.
- Brickman, W. W. (1960). A historical introduction to comparative education. *Comparative education review*, 3(3), 6-13.
- Brinkerhoff R.O. (2003), *The success case method. Find out quickly what's working and what's not*, San Francisco (CA), Berrett-Koehler.
- Brown, M. (1996). FIMS and SIMS: the first two IEA International Mathematics Surveys. *Assessment in education: principles, policy & practice*, 3(2), 193-212.
- Brown, P. A. (2008). A review of the literature on case study research. *Canadian Journal for New Scholars in Education/Revue canadienne des jeunes chercheurs et chercheurs en education*, 1(1).
- Bubb, S., Earley, P., Ahtaridou, E., Jones, J., & Taylor, C. (2007). The self-evaluation form: is the SEF aiding school improvement?. *Management in education*, 21(3), 32-37.

- Burgess, S., Greaves, E., Vignoles, A., & Wilson, D. (2015). What parents want: School preferences and school choice. *The Economic Journal*, 125(587), 1262-1289.
- Calenda, M., & Milito, F. (2020). L'attualità degli studi docimologici. *Italian Journal of Educational Research*, (24), 102-119.
- Callegari, C. (2013). Prospettive dell'educazione comparata oggi. *Annali online della Didattica e della Formazione Docente*, 5(6), 31-43.
- Callegari, C. (2017). L'educazione comparata nell'epoca globale: la tradizione italiana e le prospettive future. *STUDIUM EDUCATIONIS-Rivista semestrale per le professioni educative*, (2), 93-104.
- Callegari, C. (2019). Ricerca storica ed educazione comparata. *STUDIUM EDUCATIONIS-Rivista semestrale per le professioni educative*, (3), 47-56.
- Callegari, C., & Gaudio, A. (2018). Problemi e prospettive dell'educazione comparata oggi. *Rivista di Storia dell'Educazione*, 5(2), 7-13.
- Calvani, A., & Vivianet, G. (2014). Evidence Based Education e modelli di valutazione formativa per le scuole. *Journal of Educational, Cultural and Psychological Studies (ECPS Journal)*, 1(9), 127-146.
- Cameron, J., Cowen, R., Holmes, B., Hurst, P., & McLean, M. (1983). *International Handbook of Education Systems*. 3 volumes, Chichester: John Wiley & Sons.
- Campbell D., & Russo J. (1999), *Social Experimentation*, Thousand Oaks, CA, Sage.
- Campbell, D. (1969), *Reforms as Experiments*, *American Psychologist*, 4.
- Campbell, D. (1975). III. "Degrees of freedom" and the case study. *Comparative political studies*, 8(2), 178-193.
- Capperucci, D. (2018). Valutazione e autovalutazione delle scuole: dall'accountability al miglioramento dei risultati. In *Scuola, Democrazia, Educazione. Formare ad una nuova società della conoscenza e della solidarietà* (pp. 1018-1022). Pensa Multimedia.
- Capperucci, D. (2020). Didattica a distanza in contesti di emergenza: le criticità messe in luce dalla ricerca. *Studi sulla Formazione/Open Journal of Education*, 23(2), 13-22.
- Cardarello, R., & Pintus, A. (2019). La comprensione del testo nella scuola italiana: un bilancio storico e critico. Strategie efficaci per la comprensione del testo. *Il Reciprocal Teaching*, 47-75.
- Cardno, C. (2018). Policy document analysis: A practical educational leadership tool and a qualitative research method. *Educational Administration: Theory and Practice*, 24(4), 623-640.
- Cardone M., Muzzioli P., Poliandri D. E, Romiti S. (2010). La valutazione delle scuole: alcune idee utili alla luce delle esperienze europee. *Orientamenti Pedagogici*, 57 (4), 697-713.

- Carli, R., Dolcetti, F., Giovagnoli, F., Gurrieri, R., & Paniccia, R. M. (2015). La cultura locale del Servizio di assistenza specialistica nelle scuole della Provincia di Roma. *Quaderni di Psicologia Clinica*, (2), 16-32.
- Carro, R., & Nencioni, P. (2016). Il progetto VALeS: il punto di vista dei consulenti per il miglioramento. *Scuola democratica*, 7(2), 507-522.
- Casadellà, M., Massip-Sabater, M., González-Monfort, N., Dias-Gomes, A., & Barroso-Hortas, M. J. (2022). Imaginación, educación para el futuro y cultura democrática: Políticas educativas en la Península Ibérica. *Comunicar: Revista Científica de Comunicación y Educación*, 30(73), 57-67.
- Case, P., Case, S., & Catling, S. (2000). Please show you're working: a critical assessment of the impact of OFSTED inspection on primary teachers. *British Journal of Sociology of Education*, 21(4), 605-621.
- Castellana, G., & Corsini, C. (2018). Valutazione formativa vs accountability: l'impiego del Valore Aggiunto nella Ricerca-Formazione. *Lifelong Lifewide Learning*, 14(31), 56-78.
- Castellana, G., & Rossi, L. (2021). La didattica a distanza secondo docenti, alunni e famiglie: il percorso di autovalutazione in un istituto comprensivo. *ITALIAN JOURNAL OF EDUCATIONAL RESEARCH*, 203-222.
- Castoldi, M. (2005). L'indagine OCSE-PISA: un'opportunità per la scuola italiana. *Rassegna CNOS*, 3, 25-45.
- Castoldi, M. (2008). Si possono valutare le scuole? Il caso italiano e le esperienze europee. Torino: SEI.
- Castoldi, M. (2009). *Valutare le competenze*, Roma, Carocci.
- Castoldi, M. (2011). Autovalutazione di Istituto. In *Il Dirigente scolastico fra tradizione, innovazione, management* (pp. 122-132). Armando.
- Castoldi, M. (2012). *Valutare a scuola. Dagli apprendimenti alla valutazione di sistema*. Roma, Carocci editore.
- Castoldi, M. (2013). Valutazione. *Dizionario di Didattica*, 417-430. Brescia: La Scuola. <http://hdl.handle.net/2318/136298>.
- Cecalupo, M. (2021). DADA (Didattiche per Ambienti Di Apprendimento): la percezione degli insegnanti a seguito dell'introduzione del modello DADA. [Tesi di dottorato, Sapienza Università di Roma]. Iris Sapienza <https://iris.uniroma1.it/handle/11573/1517341?mode=complete>.
- Cecconi, L. (2002). *La ricerca qualitativa in educazione*. Milano: FrancoAngeli.
- Centro del Profesorado de Sevilla (2021). Propuesta de protocolo para trabajar en los centros educativos la elaboración de la memoria de autoevaluación, el plan de mejora y el plan de formación anuales. Retrieved from: <https://www.juntadeandalucia.es/educacion/portals/ishare-servlet/content/395b3059-c05c-4e55-a0ad-33ed5bdaa9dc>.

- Cerini, G. (2010), *La valutazione tra autonomia delle scuole e standard nazionali*, in *Educazione e scuola*. Retrived from: <https://www.edscuola.it/archivio/riformeonline/valutazione.htm>.
- CERI-OCSE (a cura di) (1994), *Valutare l'insegnamento: per una scuola che conti*. Roma: Armando Editore.
- Cesareni, D., Montebello, M., & Scippo, S. (2020). *Apprendimento delle materie STEM in Italia: analisi delle principali indagini nazionali e internazionali*. I *PROBLEMI DELLA PEDAGOGIA*, 2.
- Chapman, C. (2001). *Unlocking the potential: Inspection as a mechanism for school improvement*. *Improving Schools*, 4(3), 41-50.
- Chapman, C. (2002). *Ofsted and School Improvement: teachers' perceptions of the inspection process in schools facing challenging circumstances*. *School leadership & management*, 22(3), 257-272.
- Cecchi, D., & Zollino, F. (2001). *Struttura del sistema scolastico e selezione sociale*. *Rivista di politica economica*, 8, 43-84.
- Chelimsky, E. & Shadish, W. (1999), *Evaluation for the 21st Century: A Handbook*, London, Sage.
- Chen, H. T. (1990), *Theory-Driven Evaluation*, Newbury Park, CA, Sage.
- Christie, C. A., Ross, R. M., & Klein, B. M. (2004). *Moving toward collaboration by creating a participatory internal-external evaluation team: A case study*. *Studies in Educational Evaluation*, 30(2), 125-134.
- Chung, J. H. (2010). *Finland, PISA, and the implications of international achievement studies on education policy*. In *The impact of international achievement studies on national education policymaking*. Emerald Group Publishing Limited.
- Clarke, V., & Braun, V. (2013). *Successful qualitative research: A practical guide for beginners*. *Successful qualitative research*, 1-400.
- Clemmensen, T., Hertzum, M., Hornback, K., Shi, Q. & Yammiyavar, P. (2008). *Cultural Cognition in the Thinking-Aloud Method for Usability Evaluation*. *Twenty Ninth International Conference on Information Systems. ICIS 2008 Proceedings*.
- Coggi, C. & Notti, A.M. (a cura di) (2002). *Docimologia*, Lecce, Pensa Multimedia.
- Coggi, C., & Ricchiardi, P. (2005). *Progettare la ricerca empirica in educazione*. Carocci.
- Cohen, S., Manion, L., & Morrison, K. (2011). *Research methods in education*. Routledge, London, 7th edition.
- Colagiovanni, P. (2003). *Il Contratto della Dirigenza Scolastica*. *Studi Urbinati, A-Scienze giuridiche, politiche ed economiche*, 54(3), 363-404.
- Çolak, M. S., & Ege, A. (2013). *An assessment of EU 2020 strategy: Too far to reach?*. *Social Indicators Research*, 110(2), 659-680.

- Colucci, M., & Gallo, S. (2017). In cattedra con la valigia: Gli insegnanti tra stabilizzazione e mobilità. Donzelli Editore.
- Comunidad Autónoma de Castilla-La Mancha (2003). Resolución de 30 de mayo de 2003, de la Dirección General de Coordinación y Política Educativa. Retrived from: <https://www.educa.jccm.es/es/normativa/resolucion-30-mayo-2003-direccion-general-coordinacion-poli>.
- Corbetta, P. (1999). Metodologie e tecniche della ricerca sociale. Bologna: Il Mulino.
- Corbetta, P. (2003). Social research: Theory, methods and techniques. Sage.
- Corbin, J. & Strauss, A. (2008). Basics of qualitative research: Techniques and procedures for developing grounded theory (3rd ed.). Thousand Oaks, CA: Sage.
- Corbin, J. M., & Strauss, A. (2011). Grounded theory methodology. Handbook of Qualitative Research, 273, 285.
- Corriere della Sera (2008). Il passato e il buon senso. Retrived from: https://www.mef.gov.it/ufficio-stampa/articoli/2008_2013-giulio_tremonti/documenti/19650_200808221309rassegna.pdf.
- Corsini, C. (2010). Valutazione come classifica e autovalutazione come ricerca. Italian Journal of Educational Research, (5), 41-47.
- Corsini, C. (2012). La lezione americana: l'impiego del Valore Aggiunto nella valutazione di scuole e insegnanti. In "Scuola Democratica: Learning for Democracy", 3, Milano, pp. 108-116.
- Corsini, C. (2013). La validità di contenuto delle prove INVALSI di comprensione della lettura. *Italian Journal of Educational Research*, (10), 46-61.
- Corsini, C. (2015). Valutare scuole e docenti: Un'indagine sul punto di vista di chi insegna. Edizioni Nuova Cultura.
- Corsini, C. (2018a), Una valutazione per il successo formativo. In: I. D. M., Scierra, M., Bartolucci, R., Salvato. Lettura e dispersione. Processi e linguaggi dell'apprendimento, p. 193-210, MILANO: FrancoAngeli.
- Corsini, C. (2018b). Sull'utilità e il danno di "misurazione e valutazione" in educazione. In: Cristiano Corsini (a cura di). Rileggere Visalberghi. p. 13-28, ROMA: Nuova Cultura.
- Corsini, C. (2018c). Inclusione e culture valutative. In S. Polenghi, M. Fiorucci, L. Agostinetto (a cura di), Diritti, Cittadinanza, Inclusione, Lecce: Pensa, pp. 85-94.
- Corsini, C. (2019). Accountability e ricerca educativa: una riflessione sulla validità delle prove Invalsi somministrate fino al 2017. In G. Elia, S. Polenghi & V. Rossini (Eds.), La scuola tra saperi e valori etico-sociali. Politiche culturali e pratiche educative (pp. 993-1005). Pensa MultiMedia.
- Corsini, C. (2020a). Valutazione: controllo o sviluppo?. *Numero monografico Lend-IMI Istanbul Giugno 2020 anno XLIX*, 20.

- Corsini, C. (2020b). Tra accountability e valutazione: il caso delle prove Invalsi. In "Scholé : rivista di educazione e studi culturali", LVIII, 1, 2020, Brescia, pp. 45-59.
- Corsini, C. (Ed.). (2018). *Rileggere Visalberghi*. Edizioni Nuova Cultura.
- Corsini, C., & Asquini, G. (2010). L'evoluzione dei risultati di lettura nelle diverse edizioni di PISA. In "Pisa 2006: approfondimenti tematici e metodologici". Roma: Armando, pp. 1000-1025.
- Corsini, C., & Asquini, G. (2012). Come evolve la competenza in lettura dei quindicenni: indicazioni dalle indagini PISA (2000-2006). in "Linguistica educativa". Atti del XLIV Congresso internazionale di studi della Società di linguistica italiana (SLI), Viterbo, 27-29 settembre 2010, pp. 311-327.
- Corsini, C., & Gueli, C. (2022). Dal voto alla valutazione per l'apprendimento. *Journal of Educational, Cultural and Psychological Studies (ECPS Journal)*, (26), 163-178.
- Corsini, C., & Losito, B. (2014). Le rilevazioni Invalsi: a che cosa servono?. *Cadmo*, (2013/2).
- Corsini, C., Scierri, I. D. M., & Scionti, A. (2017). La validità delle prove INVALSI di comprensione della lettura (2010-2016). La funzione educativa della valutazione, 335-352.
- Cotter, K. E., Centurino, V. A., & Mullis, I. V. (2019). Developing the TIMSS 2019 mathematics and science achievement instruments. *Methods and procedures: TIMSS*, 1-1.
- Council of Europe (2001). *Common European Framework of Reference for Languages: learning, teaching, assessment*. Retrived from: https://invalsi-areaprove.cineca.it/docs/2019/QCER_2001.pdf.
- Cramer, J. F., & Browne, G. S. (1956). *Contemporary Education: A Comparative Study of National Systems*. New York: Harcourt, Brace & World.
- Creswell, J. W. (2002). *Educational research: Planning, conducting, and evaluating quantitative* (Vol. 7). Prentice Hall Upper Saddle River, NJ.
- Cronbach, L. J., Ambron, S. R., Dornbusch, S. M., Hess, R. D., Hornik, R. C., Phillips, D. C., ... & Weiner, S. S. (1980). *Toward reform of program evaluation*. San Francisco: Jossey-Bass.
- Crosswhite, F. J. (1985). *Second international mathematics study: Summary report for the United States*. National Center for Education Statistics.
- Cumming, A. (1996). IEA's studies of language education: their scope and contributions. *Assessment in Education: Principles, Policy & Practice*, 3(2), 179-192.
- Cutcliffe, J. R. (2000). Methodological issues in grounded theory. *Journal of advanced nursing*, 31(6), 1476-1484.
- D. L. 240/10 (2010). Norme in materia di organizzazione delle università, di personale accademico e reclutamento, nonché delega al Governo per incentivare

- la qualità e l'efficienza del sistema universitario. Retrived from: <https://www.normattiva.it/uri-res/N2Ls?urn:nir:stato:legge:2010;240>.
- D.L 258/99 (1999). Riordino del Centro europeo dell'educazione, della biblioteca di documentazione pedagogica e trasformazione in Fondazione del museo nazionale della scienza e della tecnica "Leonardo da Vinci". Retrived from: https://www.gazzettaufficiale.it/atto/serie_generale/caricaDettaglioAtto/originario?atto.dataPubblicazioneGazzetta=1999-08-04&atto.codiceRedazionale=099G0330&elenco30giorni=false.
- D.L 62/17 (2017). Norme in materia di valutazione e certificazione delle competenze nel primo ciclo ed esami di Stato. Retrived from: <https://www.gazzettaufficiale.it/eli/id/2017/05/16/17G00070/sg>.
- D.Lgs 286/04 (2004). Istituzione del Servizio nazionale di valutazione del sistema educativo di istruzione e di formazione, nonche' riordino dell'omonimo istituto, a norma degli articoli 1 e 3 della legge 28 marzo 2003, n. 53. Retrived from: https://www.gazzettaufficiale.it/atto/serie_generale/caricaDettaglioAtto/originario?atto.dataPubblicazioneGazzetta=2004-12-01&atto.codiceRedazionale=004G0320&elenco30giorni=false.
- D.P.R 419/74 (1974). Sperimentazione e ricerca educativa, aggiornamento culturale e professionale ed istituzione dei relativi istituti. Retrived from: https://www.edscuola.it/archivio/norme/decreti/dpr419_74.pdf.
- D.P.R 80/13 (2013). Regolamento sul sistema nazionale di valutazione in materia di istruzione e formazione. Retrived from: https://www.istruzione.it/valutazione/allegati/DPR_%2028_03_13.pdf.
- Dahlgard, J. J., Kristensen, K., & Kanji, G. K. (1995). Total quality management and education. *Total Quality Management*, 6(5), 445-456.
- Damiani, V., Agrusti, F., & Ciolfi, A. (2016). Le prove a risposta aperta della sperimentazione TECO 2015: esplorazione dell'archivio testuale. *CADMO*, 24(2), 105-111.
- Damiano, E. (Ed.) (2007). *Il mentore. Manuale di tirocinio per insegnanti in formazione*. Milano: FrancoAngeli.
- De Santis, C. (2022). *Competenze e valutazione nel modello DADA. Didattiche per Ambienti Di Apprendimento, pratiche educative e punto di vista genitoriale: uno studio di caso*. Stamen: Roma.
- De Santis, C., & Asquini, G. (2020). L'autovalutazione delle Scuole DADA: i RAV tra competenze chiave e ambiente di apprendimento. *RicercAzione*, 12(2), 207-223.
- Del Gobbo, G. (2018). Potenzialità del Corso "Educatore professionale socio-pedagogico" per il miglioramento della QA dei Corsi L-19 nel quadro del Processo di Bologna. *Form@re*, 18(3).

- Dell'Anna, S. (2021). *Modelli di valutazione di un sistema scolastico inclusivo: prospettive di dialogo tra implementazione, ricerca e (auto) miglioramento*. Milano: FrancoAngeli.
- deMarrais, K. B. (2004). *Qualitative Interview Studies: Learning through Experience*. In K. B. deMarrais & S. D. Lapan (Eds.) (2004), *Foundations for research: Methods of inquiry in education and the social sciences* (pp. 51-68). Lawrence Erlbaum Associates Inc. Publisher.
- deMarrais, K. B., & Lapan, S. D. (Eds.). (2003). *Foundations for research: Methods of inquiry in education and the social sciences*. Routledge.
- Deming, E. W. (1994). *The new economics for industry, government, education*. Cambridge, MA: MIT Press.
- Denzin N. K., Lincoln Y. S. (2005b). *Handbook of qualitative research*. Thousand Oaks (California): Sage.
- Denzin, N. K. (1970). *The research act: A theoretical introduction to sociological methods*. Aldine.
- Denzin, N. K., & Lincoln, Y. S. (Eds.). (1994). *Handbook of qualitative research*. sage. Thousand Oaks: The Sage.
- Dewey, J. (1938), *Logic, the Theory of Inquiry*, Holt and Co, New York, (trad. it. *Logica, teoria dell'indagine*, Einaudi, Torino, 1949).
- Dewey, J. (1939), *Theory of Valuation*, University press, Chicago (trad. it. *Teoria della valutazione*, La Nuova Italia, Firenze, 1960).
- Dewey, J. (1951). *Le fonti di una scienza dell'educazione*. La Nuova Italia.
- Di Donato, D., Mattarelli, E., & Reggimenti, I. (2021). *Insegnare capovolti prima e durante la pandemia. Un'esperienza di RicercaFormazione nell'IC San Giorgio a Liri (FR). Tecnologie digitali e ambienti di apprendimento oltre l'emergenza*. Book of Abstract (pp. 5- 6). Congresso nazionale CKBG.
- Di Fraia, G. (2004). *Le interviste on line. e-Research: Internet per la ricerca sociale e di mercato*. Bari: Editori Laterza.
- Di Genova, N., Mattarelli, E. (2021). *Valutazione di sistema in emergenza: quali sono le strategie per contrastare la povertà educativa?. QTimes Journal of Education Technology and Social Studies*, 13(2), 58-69.
- Di Nunzio, D., Toscano, E., Pirro, F., & Pedaci, M. (2021). *Il lavoro "da remoto" degli insegnanti: riflessioni da un'indagine sulla didattica a distanza durante la pandemia*. In *"Sociologia del lavoro"*, 160, 2, Milano: FrancoAngeli. Pp. 247-256.
- Direction de l'évaluation et de la prospective (D.E.P), (1991). *L'Etat de l'école*. Ministère de l'Education Nationale, Paris.
- DOGV (2015). *ORDEN de 22 de marzo de 2005, de la Conselleria de Cultura, Educación y Deporte, por la que se regula el procedimiento para la implantación de sistemas de gestión de calidad en los centros educativos no universitarios de*

- la Comunidad Valenciana. Retrived from: https://dogv.gva.es/datos/2005/04/15/pdf/2005_X3903.pdf.
- Domenici, G. (2001). *Manuale della valutazione scolastica*, Bari, Laterza.
- Domenici, G. (2012). Valutazione e autovalutazione come risorse aggiuntive nei processi di istruzione. *Education Sciences & Society*, 2(2).
- Donmoyer, R. (1990). Generalizability and the single case study. In E. W. Eisner, & A. Peshkin (Eds.), *Qualitative inquiry in education: The continuing debate* (pp.175-200). New York: Teachers College Press.
- Doyle, L., Brady, A. M., & Byrne, G. (2009). An overview of mixed methods research. *Journal of research in nursing*, 14(2), 175-185.
- Ducheneaut, N., Moore, R. J., & Nickell, E. (2007). Virtual “third places”: A case study of sociability in massively multiplayer games. *Computer Supported Cooperative Work (CSCW)*, 16(1), 129-166.
- EACEA/Eurydice/Eurostat (2012). *Key Data on Education in Europe 2012 Edition*. Brussels: EACEA P9 Eurydice.
- Earl L.M. (2013). *Assessment as Learning: Using Classroom Assessment to Maximize Student Learning*. Thousand Oaks: Corwin Press.
- Eckstein, H. (2002). Case study and theory in political science. In R. Gomm, M. Hammersley, & P. Foster (Eds.), *Case study method: Key issues, key texts* (pp. 119-163). London: Sage.
- Eco U. (2004), *La combinatoria della creatività*, conferenza per la Nobel Foundation, Firenze 15 settembre 2004, www.umbertoeco.it.
- Ehren, M. C., & Visscher, A. J. (2006). Towards a theory on the impact of school inspections. *British journal of educational studies*, 54(1), 51-72.
- Ehren, M., Paterson, A., & Baxter, J. (2020). Accountability and trust: two sides of the same coin?. *Journal of Educational Change*, 21(1), 183-213.
- Eivers, E. (2010). PISA: Issues in implementation and interpretation. *The Irish Journal of Education/Iris Eireannach an Oideachais*, 94-118.
- Ellerani, P. G. (2006). Per una valutazione «autentica». *Innovazione educativa*, 2, 50-56.
- Elley, W. B. (1992). How in the world do students read? IEA study of reading literacy. Retrieved from: <https://files.eric.ed.gov/fulltext/ED360613.pdf>.
- Elley, W. B. (1998). An insider's view of the IEA reading literacy study. *Studies in Educational Evaluation*, 24(2), 127-136.
- ENQA (2015). *Standards and Guidelines for Quality Assurance in the European Higher Education Area (ESG)*. Brussels, Belgium.
- ENQA (2021). *Statutes Of The European Association For Quality Assurance In Higher Education*. Retrived from: <https://www.enqa.eu/wp-content/uploads/ENQA-Statutes-2021.pdf>.

- Epifani, G., Giampietro, L., Poliandri, D., Pranterà, E., & Sette, S. (2016). Il ruolo dei processi di autovalutazione. Il dirigente scolastico per il miglioramento della scuola. *Scuola democratica*, 7(2), 439-450.
- Ertl, H. (2006). Educational standards and the changing discourse on education: The reception and consequences of the PISA study in Germany. *Oxford Review of Education*, 32(5), 619-634.
- Espinoza, V. M. (1983). La educación comparada: breve estudio documental. *Educar*, 3, 169-181.
- European Commission (2020). Supporting school self-evaluation and development through quality assurance policies: key considerations for policy makers. Summary of the Report by ET2020 Working Group Schools. Retrieved from <https://euagenda.eu/upload/publications/supporting-school-self-evaluation-and-development.pdf>.
- European Commission/EACEA/Eurydice (2015). Assuring Quality in Education: Policies and Approaches to School Evaluation in Europe. Eurydice Report. Luxembourg: Publications Office of the European Union.
- European Commission/EACEA/Eurydice (2022). The structure of the European education systems 2022/2023: schematic diagrams. Eurydice Facts and Figures. Luxembourg: Publications Office of the European Union.
- Eurydice (2004). Evaluation of schools providing compulsory education in Europe. Brussels: Eurydice.
- Eurydice (2019). I sistemi scolastici al traguardo del 2020. Retrieved from: https://eurydice.indire.it/wp-content/uploads/2020/05/Quaderno_44_file-per-WEB.pdf.
- Falanga, M. (2014). La valutazione del personale docente, delle Istituzioni scolastiche e l'autovalutazione: considerazioni di carattere giuridico (Vol. 3). Programma Education FGA Working Paper, Fondazione Giovanni Agnelli, 53.
- Falcinelli F. (2014), Categorie e funzioni della valutazione. In L. Galliani (a cura di), *L'agire valutativo*, Brescia, La Scuola, pp. 69-80.
- Falzetti, P. (2019). Un sistema scolastico in difficoltà e un preoccupante dualismo territoriale: i risultati delle prove INVALSI. *Social Policies*, 6(3), 527-532.
- Fatigante, M. (2006). Teoria e pratica della trascrizione in analisi conversazionale. L'irriducibilità interpretativa del sistema notazionale. Bürki, Y. & De Stefani, E. (a cura di): *Trascrivere la lingua. Dalla filologia all'analisi conversazionale*. Bern (Peter Lang), 219-256.
- Fereday, J. & Muir-Cochrane, E. (2006). Demonstrating rigor using thematic analysis: A hybrid approach of inductive and deductive coding and theme development. *International Journal of Qualitative Methods*, 5(1), 80-92. Retrieved from: http://www.ualberta.ca/~iiqm/backissues/5_1/pdf/fereday.pdf.

- Ferguson, N., Earley, P., Fidler, B., & Ouston, J. (2000). *Improving schools and inspection: The self-inspecting school*. Sage.
- Fernig, L. (1959). The global approach to comparative education. *International review of education*, 343-355.
- Ferrari, M. (2000). Differenti approcci al problema della qualità nella scuola. A. Bondioli, M. Ferrari (a cura di) (2000). *Manuale di valutazione del contesto educativo: teorie, modelli, studi per la rilevazione della qualità nella scuola*, 66-77.
- Ferrer, A. T. (2014). Veinte años de políticas de evaluación general del sistema educativo en España. *Revista de evaluación de programas y políticas públicas*, (2), 1-21.
- Fetterman D.M. (2001), *Foundations of empowerment evaluation*, Thousand Oaks (CA), Sage.
- Fideli, R., & Marradi, A. (1996), Intervista, in *Enciclopedia delle Scienze Sociali*, vol. V, Roma, Istituto dell'Enciclopedia Italiana, pp. 71-82.
- Field, J. (2000). *Lifelong learning and the new educational order*. Trentham Books, Ltd., Westview House, 734 London Road, Stoke on Trent, ST4 5NP, United Kingdom UK.
- Fiorucci, M., Riccardi, V., & Stillo, L. (2022). La formazione interculturale di insegnanti ed educatori: alcune riflessioni sul Sistema Integrato 0-6. *FOR-Rivista per la formazione*, (3), 33-35.
- Firestone, W. A., Schorr, R. Y., & Monfils, L. F. (Eds.). (2004). *The ambiguity of teaching to the test: Standards, assessment, and educational reform*. Routledge.
- Fischbach, A., Keller, U., Preckel, F., & Brunner, M. (2013). PISA proficiency scores predict educational outcomes. *Learning and Individual Differences*, 24, 63-72.
- Fischman, G. E., Topper, A. M., Silova, I., Goebel, J., & Holloway, J. L. (2019). Examining the influence of international large-scale assessments on national education policies. *Journal of education policy*, 34(4), 470-499.
- Fishbein, B., Martin, M. O., Mullis, I. V., & Foy, P. (2018). The TIMSS 2019 item equivalence study: Examining mode effects for computer-based assessment and implications for measuring trends. *Large-scale Assessments in Education*, 6(1), 1-23.
- Flick, U. (2007). *Designing qualitative research*. Sage.
- Flyvbjerg, B. (2004). Five misunderstandings about case-study research. *Sociologisk tidsskrift*, 12(2), 117-142.
- Flyvbjerg, B. (2011). Case study. *The Sage handbook of qualitative research*, 4, 301-316.
- Fong, K. (2019). Subject to evaluation: How parents assess and mobilize information from social networks in school choice. In *Sociological Forum* (Vol. 34, No. 1, pp. 158-180).

- Galliani, L. & Notti, A.M. (a cura di) (2014), *Valutazione Educativa*, Lecce, Pensa Multimedia.
- Galliani, L. (2009), *Web ontology della valutazione educativa*, Lecce, Pensa Multimedia.
- Galliani, L. (a cura di) (2015), *L'agire valutativo*, Brescia, La Scuola.
- García Crespo, F. J., & Galián Nicolás, B. (2018). Equidad educativa en Asturias, España y Europa a partir de los resultados PIRLS 2016. Retrived from: https://redined.educacion.gob.es/xmlui/bitstream/handle/11162/181484/Evaluacion-informe_N17.pdf?sequence=1&isAllowed=y.
- García Crespo, F. J., Suárez Álvarez, J., Fernández Alonso, R., & Muñoz Fernández, J. (2022). Academic resilience in Mathematics and Science: Europe TIMSS-2019 data. *Psicothema*.
- Garden, R. A. (1987). The second IEA mathematics study. *Comparative Education Review*, 31(1), 47-68.
- Gardner, D. R. (1983). "A Nation at Risk": The Imperative for Educational Reform. A Report to the Nation and the Secretary of Education United States Department of Education. Washington D.C.
- Garrido, J. L. G. (2000). *Diagnosis of the educational system*. Ministerio de Educación.
- Garrido, J. L. G., RUIZ, M. J. G., & Starkie, E. G. (2012). *La educación comparada en tiempos de globalización* (Vol. 6302307). Editorial UNED.
- Gattullo M. (1989), *Quantitativo e qualitativo in educazione e nella ricerca educativa*, *Scuola e Città*, 4: 158-162.
- George, A. L., & Bennett, A. (2005). *Case studies and theory development in the social sciences*. mit Press.
- German Federal Ministry of Education and Research (2003). *Vertiefender Vergleich der Schulsysteme ausgewählter PISA-Teilnehmerstaaten*. Retrieved from: https://www.telekom-stiftung.de/sites/default/files/ackeren_isabell_van_-_2003_-_vertiefender_vergleich_der_schulsysteme_ausgewaehlter_pisa-teilnehmerstaaten.pdf.
- Gerring, J. (2004). What is a case study and what is it good for?. *American political science review*, 98(2), 341-354.
- Giacomantonio, A. (2008). *Approcci e funzioni della valutazione*. In G. Benvenuto & A. Giacomantonio, *Un po' di storia della valutazione scolastica: letture e riflessioni* (pp. 41-50). Nuova Cultura.
- Giampietro, L., Poliandri, D., Quadrelli, I., & Romiti, S. (2016). *L'autovalutazione in Italia. Istanza di rendicontazione o sfida per migliorarsi?*. *Scuola democratica*, 7(2), 467-480.
- Gibbs, G. R. (2018). *Analyzing qualitative data*. Sage.

- Giovannini M.L. (2014), Il dibattito su valutazione ed evidenze: per un processo valutativo credibile e trasparente, *ECPS - Educational, Cultural and Psychological Studies*, 9, 1: 101-126.
- Glaser, B., Strauss, A. (1967), *The discovery of Grounded Theory: strategies for qualitative research*, Chicago, IL, Aldine Pub Co.
- Glass, G. (2008). *Fertilizers, Pills, and Magnetic Strip. The Fate of Public Education in America*. Charlotte, North Carolina, Information Age Publishing.
- Gobierno de Aragon (2022). Instrucciones para los institutos de educación secundaria de la comunidad autónoma de Aragón en relación con el curso 2022/2023. Retrieved from: <https://educa.aragon.es/documents/20126/2707664/IES.pdf/0d1a4fc4-ceee-ec18-81cc-8054280195ec?t=1658124729957>.
- Goldstein, A. E., & Reiboldt, W. (2004). The multiple roles of low income, minority women in the family and community: A qualitative investigation. *The qualitative report*, 9(2), 241-265.
- Gomm, R., Hammersley, M., & Foster, P. (2000). Case study and generalization. *Case study method*, 98-115.
- Gore, A. (1993), *Creating a Government that Works Better and Costs Less*. Report of the National Performance Review, New York, Random House.
- Gorur, R. (2016). Seeing like PISA: A cautionary tale about the performativity of international assessments. *European Educational Research Journal*, 15(5), 598-616.
- Greenstein L. (2010). *What Teachers Really Need to Know About Formative Assessment*. Alexandria: Association for Supervision & Curriculum Development.
- Grek, S. (2012). What PISA knows and can do: Studying the role of national actors in the making of PISA. *European Educational Research Journal*, 11(2), 243-254.
- Guba E.G., & Lincoln Y.S. (1989). *Fourth Generation Evaluation*. Newbury Park.
- Guba, E.G. & Lincoln, Y.S. (1982). Epistemological and methodological bases of naturalistic inquiry, *Educational Communications and Technology Journal*, vol. 30 (4).
- Guion, L. A., Diehl, D. C., & McDonald, D. (2011). Triangulation: establishing the validity of qualitative studies. FCS6014/FY394, Rev. 8/2011. *Edis*, 2011(8), 3-3.
- Gustafsson, J. (2017). Single case studies vs. multiple case studies: A comparative study. Retrieved from: <https://www.diva-portal.org/smash/get/diva2:1064378/FULLTEXT01.pdf>.
- Gustafsson, J. E., & Rosén, M. (2006). The dimensional structure of reading assessment tasks in the IEA reading literacy study 1991 and the Progress in International Reading Literacy Study 2001. *Educational Research and Evaluation*, 12(5), 445-468.

- Hanna, G. (2000). Declining gender differences from FIMS to TIMSS. *Zentralblatt für Didaktik der Mathematik*, 32(1), 11-17.
- Hans, N. (2012). *Comparative education: A study of educational factors and traditions*. Routledge.
- Hansen, W. L. (1993). Bringing total quality improvement into the college classroom. *Higher Education*, 25(3), 259-279.
- Hanushek E. A. & Woessmann L. (2015), *The knowledge capital of nations: education and the economics of growth*, The MIT Press, Cambridge, Massachusetts.
- Hanushek E. A. (2011), The economic value of higher teacher quality, *Economics of Education Review*, 30, 3: 466-479.
- Härkönen, R., Pihlava, M., & Terva, I. (2019). Lapsen arvoista varhaiskasvatusta: apuväline sosionomiopiskelijoille varhaiskasvatuksen laadun arviointiin. Retrieved from: https://www.theseus.fi/bitstream/handle/10024/267619/Harkonen_Roosa,%20Pihlava_Meri,%20Terva_Iida.pdf?sequence=2.
- Härnqvist, K. (1987). The IEA revisited. *Comparative Education Review*, 31(1), 129-136.
- Harris, D. N. & Herrington, C. D. (2006). Accountability, Standards, and the Growing Achievement Gap: Lessons from the Past Half-Century. *American Journal of Education*, 112(2), pp. 209-238.
- Harvey, L., & Green, D. (1993). Defining quality. *Assessment & evaluation in higher education*, 18(1), 9-34.
- Hayden, M., Thompson, J., & Walker, G. (Eds.). (2003). *International education in practice: Dimensions for schools and international schools*. Routledge.
- Hesse-Biber, S. (2010). Qualitative approaches to mixed methods practice. *Qualitative inquiry*, 16(6), 455-468.
- Hoepfl, M. C. (1997). *Choosing qualitative research: A primer for technology education researchers*. Volume 9 Issue 1 (fall 1997).
- Honig, M. I., & Coburn, C. (2008). Evidence-based decision making in school district central offices: Toward a policy and research agenda. *Educational policy*, 22(4), 578-608.
- Hooper, M., Mullis, I. V., & Martin, M. O. (2016). PIRLS 2016 context questionnaire framework. *PIRLS*, 31-53.
- Hopfenbeck, T. N., Lenkeit, J., El Masri, Y., Cantrell, K., Ryan, J., & Baird, J. A. (2018). Lessons learned from PISA: A systematic review of peer-reviewed articles on the programme for international student assessment. *Scandinavian Journal of Educational Research*, 62(3), 333-353.
- Hopkins, D. (1988). *Doing School Based Review. Instruments and Guidelines* (Technical report n. 5). Leuven, Acco.

- Hopkins, D. (1990). The International School Improvement Project (ISIP) and Effective Schooling: towards a synthesis. *School Leadership & Management*, 10(2-3), pp. 179-194.
- House, J. D. (2006). Mathematics beliefs and achievement of elementary school students in Japan and the United States: Results from the Third International Mathematics and Science Study. *The Journal of genetic psychology*, 167(1), 31-45.
- Howie, S., Combrinck, C., Roux, K., Tshele, M., Mokoena, G., & Palane, N. M. (2017). PIRLS 2016. Pretoria: Centre for Evaluation and Assessment.
- Hubbard, D. L. (1994). Total quality management in higher education: Learning from the factories. *Developing quality systems in education*, 135-148.
- Hutchinson, S., & Wilson, S. H. (1986). Grounded theory. *Nursing research: A qualitative perspective*, 111-130.
- Hyman, H. H. (1967), Evaluating Social Action Programm, in Lazarsfeld, P. F. (a cura di) (1967), *The Uses of Sociology*, New York, Basic Book.
- Ignas, E., & Corsini, R. J. (eds.) (1981). *Comparative Educational Systems*. Itasca: F. E. Peacock Publishers.
- Invalsi (2016a). La valutazione esterna delle scuole in Italia: a cosa serve, com'è realizzata? Retrieved from: https://www.invalsi.it/snv/docs/ccfsnv/Documento_divulgativo_DEF.pdf.
- Invalsi (2016b). Programma e protocollo per le visite di valutazione esterna nel Sistema Nazionale di Valutazione. Retrieved from: https://www.invalsi.it/snv/docs/ves/Programma_Protocollo_Valutazione_estern_a_2016.pdf.
- Invalsi (2017). Statuto (ai sensi del Decreto Legislativo 25 novembre 2016, n. 218, articoli 3, 4 e 19). Retrieved from: https://www.invalsi.it/amm_trasp/documenti/attigenerali/Statuto_Invalsi_Post_Decreto_218.pdf.
- Invalsi (2018a). Quadro di Riferimento delle Prove Invalsi di Matematica. Retrieved from: https://invalsi-areaprove.cineca.it/docs/file/QdR_MATEMATICA.pdf.
- Invalsi (2018b). Quadro di Riferimento delle Prove Invalsi di Italiano. Retrieved from: https://invalsi-areaprove.cineca.it/docs/file/QdR_ITALIANO.pdf.
- Invalsi (2022a). RAV 2022-2025 Guida all'autovalutazione per le scuole statali e paritarie del I e del II ciclo. Retrieved from: https://www.miur.gov.it/documents/20182/0/GUIDA+ALL%27AUTOVALUTAZIONE+RAV_def.pdf/658e8482-b55a-a0cc-661a-e41eddb5295d?t=1670254031400.
- Invalsi (2022b). RAV 2022-2025 Mappa degli indicatori per l'autovalutazione delle scuole statali e paritarie del I e del II ciclo. Retrieved from: <https://www.miur.gov.it/documents/20182/0/Mappa+degli+indicatori.pdf/e1b5fd e3-7492-4943-fc0a-c26101136b4f?t=1670254688590>.

- Izzo, D., & Ciurnelli, B. (2020). L'impatto della pandemia sulla didattica: percezioni, azioni e reazioni dal mondo della scuola. *Lifelong Lifewide Learning*, 16(36), 26-43.
- J., van Petegem, P., & Visscher, A. (2012). The use of school selfevaluation results in the Netherlands and Flanders. *British Educational Research Journal*, 38,125–52. doi:10.1080/01411926.2010.528556.
- Juan, M. B., & Alfageme-González, M. B. (2021). Inicio del proceso de calidad en un centro público de Educación Secundaria. *Revista de educación, innovación y formación: REIF*, (5), 55-73.
- Jullien, M. A. (1817). *Esquisse et vues préliminaires d'un ouvrage sur l'éducation comparée*. chez L. Colas, imprimeur.
- Kandel, I. L. (1933). *Studies in Comparative Education*. Boston: Houghton Mifflin.
- Kandel, I. L. (1936). *Comparative education*. *Review of Educational Research*, 6(4), 400-416.
- Karjalainen, T., Korhonen, K., Kytölä, T., Vikberg, T., Alila, K., & Vuorinen, L. (2016). *Kasvatus- ja koulutuspalvelut. Peruspalvelujen tila-raportti 2016, osa II, 5*. KARVI-FINEEC (2022). *Koulutuksen arviointisuunnitelma 2020–2023*. Retrived from: https://karvi.fi/wp-content/uploads/2020/03/KARVI_arviointisuunnitelma_paivitetty_2022_web.pdf
- Khan, S. (2002). *Teaching chemistry using guided discovery and an interactive computer tool*. Unpublished doctoral dissertation, University of Massachusetts, Amherst.
- Khan, S. (2007). The case in case-based design of educational software: A methodological interrogation. *Educational Technology Research and Development*, 55, 1-25.
- Klinke, A., & Renn, O. (2002). A new approach to risk evaluation and management: Risk-based, precaution-based, and discourse-based strategies 1. *Risk Analysis: An International Journal*, 22(6), 1071-1094.
- Kreber, C., Brook, P., & Policy, E. (2001). Impact evaluation of educational development programmes. *International Journal for Academic Development*, 6(2), 96-108.
- Kyriakides L., Creemers B. & Charalambous E. (2018), *Equity and quality dimensions in educational effectiveness*, Springer International Publishing AG, Cham, Switzerland. <https://doi.org/10.1007/978-3-319-72066-1>.
- Kyriakides L., Creemers B. & Panayiotou A. (2020), *Developing and testing theories of Educational Effectiveness addressing the dynamic nature of education*, in Hall J., Lindorff A., Sammons P., Eds., *International Perspectives in Educational Effectiveness Research* (pp. 33-69), Springer Nature Switzerland AG. <https://doi.org/10.1007/978-3-030-44810-3>.

- Kyriakides, L., & Campbell, R. J. (2004). School self-evaluation and school improvement: A critique of values and procedures. *Studies in Educational Evaluation*, 30(1), 23-36.
- Kyriakides, L., & Charalambous, C. Y. (2014). Educational effectiveness research and international comparative studies: Looking back and looking forward. *Educational policy evaluation through international comparative assessments*, 33-50.
- Labuschagne, A. (2003). Qualitative research: Airy fairy or fundamental. *The qualitative report*, 8(1), 100-103.
- Laki 1295/13 (2013). Laki Kansallisesta koulutuksen arviointikeskuksesta. Helsinki, 30 joulukuuta 2013.
- Laki 582/15 (2015). Laki Kansallisesta koulutuksen arviointikeskuksesta annetun lain 1 ja 2 §:n muuttamisesta. Helsinki, 8 toukokuuta 2015.
- Laki 628/98 (1998). Perusopetuslaki. Retrived from: <https://www.finlex.fi/fi/laki/ajantasa/1998/19980628>.
- Lamma Valentini, M. T., Lovato, M. T., & Zirotti Rossi, M. (1964). Un anno di scuola media unica. *il Mulino*, 13(11), 1122-1135.
- Landri, P., & Maccarini, A. M. (2016). Uno specchio per la valutazione della scuola. *Paradossi, controversie, vie d'uscita*. Milano: FrancoAngeli.
- Lapadat, J. C. (2000). Problematizing transcription: Purpose, paradigm and quality. *International journal of social research methodology*, 3(3), 203-219.
- Lastrucci E. (2017a), Variabili, indicatori ed indici, in Lastrucci E., a cura di, *Valutazione di sistema e Autovalutazione di istituto* (pp. 85-103), Editoriale Anicia, Roma.
- Lastrucci E. (2017b), Qualità, efficienza ed efficacia dell'attività formativa, in Lastrucci E., a cura di, *Valutazione di sistema e Autovalutazione di istituto* (pp. 133-168), Editoriale Anicia, Roma.
- Lea, M., & Spears, R. (1992). Paralanguage and social perception in computer-mediated communication. *Journal of Organizational Computing and Electronic Commerce*, 2(3-4), 321-341.
- Legge 10/11 (2011). Conversione in legge, con modificazioni, del decreto-legge 29 dicembre 2010, n. 225, recante proroga di termini previsti da disposizioni legislative e di interventi urgenti in materia tributaria e di sostegno alle imprese e alle famiglie. (11G0052). Serie Generale n.47 del 26-02-2011. Retrived from: <https://www.gazzettaufficiale.it/gunewsletter/dettaglio.jsp?service=1&datagu=2011-02-26&task=dettaglio&numgu=47&redaz=11A02788&tmstp=1300194646567>.
- Legge 176/07 (2007). Conversione in legge, con modificazioni, del decreto-legge 7 settembre 2007, n. 147, recante disposizioni urgenti per assicurare l'ordinato avvio dell'anno scolastico 2007-2008 ed in materia di concorsi per ricercatori

- universitari. Retrived from:
https://www.gazzettaufficiale.it/atto/serie_generale/caricaDettaglioAtto/originario?atto.dataPubblicazioneGazzetta=2007-10-26&atto.codiceRedazionale=007G0195&elenco30giorni=false.
- Lehmann, R. (2011). The impact of IEA on educational policy making in Germany. In IEA (Vol. 958, No. 2008, p. 50).
- Lehtonen, M. (2013). Perusopetuksen laatutyö. Erilaisia tapoja ottaa laatukriteerit hallintaan. Retrived from:
<https://julkaisut.valtioneuvosto.fi/bitstream/handle/10024/75284/okm07.pdf?sequence=1&isAllowed=y>.
- Leiber, T., Stensaker, B., & Harvey, L. (2015). Impact evaluation of quality assurance in higher education: methodology and causal designs. *Quality in Higher Education*, 21(3), 288-311.
- Leino, K., Nissinen, K., Puhakka, E. & Rautopuro, J. (2017). Lukutaito luodaan yhdessä: kansainvälinen lasten lukutaitotutkimus (PIRLS 2016), Finnish Institute for Education, University of Jyväskylä, Jyväskylä, Finland (<http://urn.fi/URN:ISBN:978-951-39-7292-9>).
- Ley Orgánica 1/90 (1990), de 3 de octubre, de Ordenación General del Sistema Educativo (LOGSE). BOE núm. 238.
- Ley Orgánica 9/95 (1995), de 20 de noviembre, de la Participación, la Evaluación y el Gobierno de los Centros (LOPEG). BOE núm. 278.
- Ley Orgánica 2/06 (2006), de 3 de mayo, de Educación (LOE). BOE núm. 106.
- Ley Orgánica 8/13 (2013), de 9 de diciembre, para la mejora de la calidad educativa (LOMCE). BOE núm. 95.
- Ley Orgánica 3/20 (2020), de 29 de diciembre, por la que se modifica la Ley Orgánica 2/2006, de 3 de mayo, de Educación. BOE núm. 340.
- Lichtner, M. (2016). Pedagogia e filosofia in Democrazia e educazione. *Scuola democratica*, 7(3), 783-790.
- Lincoln, Y. S., & Guba, E. G. (2000). The only generalization is: There is no generalization. *Case study method*, 27, 44.
- Llorent Bedmar, V., & Cobano-Delgado Palma, V. (2018). Reformas en la legislación educativa española durante el actual período democrático: Una perspectiva crítica. *Archivos analíticos de políticas educativas*, 26 (125), 1-23.
- Lombi, L. (2015). *Le web survey*. Milano: FrancoAngeli.
- López Cabanes, A., & Ruiz Gimeno, J. (2004). Gestión de calidad en centros educativos no universitarios¿ qué es?¿ para qué vale?¿ cómo se puede aplicar?. *Educación en el 2000: revista de formación del profesorado*.
- Losito, B. (2010). I metodi. In B. Vertecchi, G. Agrusti & B. Losito, *Origini e sviluppi della ricerca valutativa* (pp. 91-166). Franco Angeli.
- Losito, G. (2015). *L'intervista nella ricerca sociale*. Gius. Laterza & Figli Spa.

- Love, A. J. (1991). *Internal evaluation: Building organizations from within* (Vol. 24). Sage.
- Love, P. (2013). Document analysis. In *Research in the college context* (pp. 99-112). Routledge.
- Lucisano P. & Corsini C. (2015), *Docenti e valutazione di scuole e insegnanti*, *Giornale Italiano della ricerca educativa - Italian Journal of Educational Research*, 15: 97-109.
- Lucisano P. (1989), *Misurare perché e come*, in P. Lucisano, a cura di, *Lettura e comprensione*, pp. 43-68, Loescher, Torino.
- Lucisano P. (2015), "Comprendere la valutazione", *I problemi della pedagogia*, 2: 335-343.
- Lucisano, P. (2020). Fare ricerca con gli insegnanti. I primi risultati dell'indagine nazionale SIRD "Per un confronto sulle modalità di didattica a distanza adottate nelle scuole italiane nel periodo di emergenza COVID-19". *Lifelong Lifewide Learning*, 16(36), 3-25.
- Lucisano, P., & Salerni, A. (2009). *Metodologia della ricerca in educazione e formazione*. Carocci.
- Lundberg, I., & Linnakyla, P. (1993). *Teaching Reading around the World*. IEA Study of Reading Literacy. Hamburg, IEA.
- MacBeath, J. (2008), «Leading Learning in the Self-Evaluating School», *School Leadership & Management: Formerly School Organisation*, 28 (4), 385-99.
- Maccario, D. (2012). Valutare per promuovere il successo scolastico. *Education Sciences & Society*, 2(2).
- MacDonald, B., & Walker, R. (1977). Case-study and the social philosophy of educational research. *Cambridge journal of Education*, 5(1), 2-11.
- Maestro Martín, C. (2005). Las evaluaciones nacionales e internacionales. La mejora de la calidad del sistema educativo: el éxito de todos los alumnos como objetivo. VI Seminario de Primavera: La enseñanza de las Matemáticas y el informe PISA. 2005, pp. 5-20.
- Maestro Martín, C. (2006). La evaluación del sistema educativo. *Transatlántica de educación*, (1), 55-68.
- Malen, B., Rice, J. K., Matlach, L. K., Bowsher, A., Hoyer, K. M., & Hyde, L. H. (2015). Developing organizational capacity for implementing complex education reform initiatives: Insights from a multiyear study of a teacher incentive fund program. *Educational Administration Quarterly*, 51(1), 133-176.
- Malizia G. (2008). *Educazione comparata*. In J.M. Prellezo, G. Malizia, C. Nanni (a cura di), *Dizionario di scienze dell'educazione* (pp. 379-382). Roma: LAS.
- Manzon, M. (2014). Comparing places. In *Comparative education research* (pp. 97-137). Springer, Cham.

- Marlow-Ferguson, R. (2002). *World education encyclopedia: A survey of educational systems worldwide*. Gale Cengage.
- Martens, K., & Niemann, D. (2013). When do numbers count? The differential impact of the PISA rating and ranking on education policy in Germany and the US. *German Politics*, 22(3), 314-332.
- Martin, M. O. (1996). *Third international mathematics and science study: An overview*. Third international mathematics and science study (TIMSS) Technical report, 1. Retrieved from: <https://timss.bc.edu/timss1995i/TIMSSPDF/TRall.pdf#page=8>.
- Martin, M. O., & Kelly, D. L. (1996). *Third International Mathematics and Science Study. Technical Report. Volume I: Design and Development*. Boston College, Center for the Study of Testing, Evaluation, and Educational Policy, Campion Hall 323, Chestnut Hill, MA 02167; <http://wwwwcsteep.bc.edu/timss>.
- Martin, M. O., Mullis, I. V., & Hooper, M. (2017). *Methods and Procedures in PIRLS 2016*. International Association for the Evaluation of Educational Achievement.
- Martin, M. O., Mullis, I. V., & Foy, P. (2020). *Diseño de la evaluación TIMSS 2019*. TIMSS 2019. Marcos de la evaluación, 80.
- Martin, M. O., von Davier, M., & Mullis, I. V. (2020). *Methods and Procedures: TIMSS 2019 Technical Report*. International Association for the Evaluation of Educational Achievement.
- Marzano, A. (2019). *Formazione per il cambiamento della scuola. Piani di miglioramento, azione e riflessione: un circolo teorico-pratico da ricomporre. Strategie efficaci per la comprensione del testo*. *Il Reciprocal Teaching*, 127-151.
- Mathison, S. (1991). What do we know about internal evaluation?. *Evaluation and Program Planning*, 14(3), 159-165.
- Mathison, S. (2011). Internal evaluation, historically speaking. *New Directions for Evaluation*, 2011(132), 13-23.
- Mattarelli, E. (2021a). *Gli studenti universitari e la Didattica a Distanza: resilienza e percezione del carico di lavoro*. *ITALIAN JOURNAL OF EDUCATIONAL RESEARCH*, 087-094.
- Mattarelli, E. (2021b). *The Finnish Educational Community at the Time of COVID-19, Perceptions and Experiences*. In *VOLUME II Proceedings*, 19. Second edition of the International Conference of the journal "Scuola Democratica", pp. 19-29.
- Mattarelli, E., & Zanazzi, S. (2021). *La valutazione dei sistemi universitari: un confronto tra Italia e Finlandia*. *Formazione, lavoro, persona*, 33, 231-248.
- Mattarelli, E. (2022). *Tra pratiche e percezioni sulla valutazione di sistema: confronto tra l'Italia e altri paesi europei*. In *La ricerca educativa e didattica nelle scuole di dottorato in Italia*. Quaderni del dottorato SIRD. Lecce: Pensa MultiMedia.

- Matthews, P., & Smith, G. (1995). OFSTED: inspecting schools and improvement through inspection. *Cambridge Journal of Education*, 25(1), 23-34.
- McBeath, J. (1999). *Schools must speak for themselves*. London: Routledge.
- McGuinn, P. (2012). Stimulating reform: Race to the Top, competitive grants and the Obama education agenda. *Educational Policy*, 26(1), 136-159.
- McKnight, C. C. (1987). *The Underachieving Curriculum: Assessing US School Mathematics from an International Perspective. A National Report on the Second International Mathematics Study*. Stipes Publishing Co., 10-12 Chester St., Champaign, IL 61820.
- McLellan, E., MacQueen, K. M., & Neidig, J. L. (2003). Beyond the qualitative interview: Data preparation and transcription. *Field methods*, 15(1), 63-84.
- McNamara, G., & O'Hara, J. (2008). The importance of the concept of self-evaluation in the changing landscape of education policy. *Studies in educational evaluation*, 34(3), 173-179.
- Melchiori, R. (2002). *Adas. Il laboratorio della valutazione. Aspetti concettuali*. Franco Angeli.
- Memarovic, N., Fels, S., Anacleto, J., Calderon, R., Gobbo, F., & Carroll, J. M. (2014). Rethinking third places: Contemporary design with technology. *The Journal of Community Informatics*, 10(3).
- Merriam, S. B. (1988). *Case study research in education: A qualitative approach*. Jossey-Bass.
- Merriam, S. B. (1998). *Qualitative research and case study applications in education*. Jossey-Bass.
- Mertens, D. M., & Hesse-Biber, S. (2013). Mixed methods and credibility of evidence in evaluation. *New directions for evaluation*, 2013(138), 5-13.
- Metitieri, F., & Manera, G. (2000). *Dalla email al chat multimediale*. Milano: FrancoAngeli.
- Meuret, D. (2001). De la contribution des indicateurs au débat sur l'éducation. Une étude de cas: L'Etat de l'Ecole. *Politiques d'éducation et de formation. Analyses et comparaisons internationales*, 13, 13-27.
- Miceli, R. (2004). *Numeri, dati, trappole*. Carocci.
- Mikkola, A., Repo, L., Vlasov, J., Paananen, M., & Mattila, V. (2017). *Varhaiskasvatuksen arvioinnin nykytila*. (Julkaisut/Kansallinen koulutuksen arviointikeskus; No. 22:2017). Kansallinen koulutuksen arviointikeskus (KARVI).
- Mincu, M. E. (2020). *Sistemi scolastici nel mondo globale. Educazione comparata e politiche educative* (pp. XI-195). Mondadori.
- Ministerio de Educación y Formación Profesional (2022). *Sistema Estatal de Indicadores de la Educación*. Retrived from:

- <https://www.educacionyfp.gob.es/inee/dam/jcr:1f4a7f5c-a3bd-40f8-9433-b400e38e2629/seie-2022.pdf>.
- MIUR (2022). Rapporto di Autovalutazione. Guida operativa. Retrieved from: https://www.miur.gov.it/documents/20182/0/Guida_operativa_RAV_2022.pdf/956b372d-512a-dc2f-d5ad-a36bac1fbd1d?t=1664277243325.
- Moehlman, A.H. (1963). *Comparative Educational Systems*. Washington DC: The Center for Applied Research in Education.
- Monarca, H., & Vallev, J. (2014). Evaluaciones nacionales e internacionales como política de mejora de la educación en España en el contexto europeo. ATAS do VI Encontro do CIED–I Encontro Internacional em Estudos Educacionais. Avaliação: Desafios e Riscos, 3-15.
- Moretti, G. (2011). Dirigenza scolastica e competenze di leadership. In "Leadership educativa e autonomia scolastica: il governo dei processi formativi e gestionali nella scuola di oggi". Roma: Armando. Pp. 31-67.
- Mullis, I. V. (1997). *Mathematics Achievement in the Primary School Years*. IEA's Third International Mathematics and Science Study (TIMSS). TIMSS International Study Center, Center for the Study of Testing, Evaluation, and Educational Policy, Champion Hall, School of Education, Boston College, Chestnut Hill, MA 02167; world wide web: <http://wwwwcsteep.bc.edu/timss>.
- Mullis, I. V., & Martin, M. O. (2007). TIMSS in perspective: Lessons learned from IEA's four decades of international mathematics assessments. *Lessons learned: What international assessments tell us about math achievement*, 9-36. Retrieved from: <https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=6dd44fc957f10434935ccc43c8dda923026f247f2>.
- Mullis, I. V., & Martin, M. O. (2015). *PIRLS 2016 Assessment Framework*, 2nd Edition. Retrieved from: https://timss.bc.edu/pirls2016/downloads/P16_Framework_2ndEd.pdf.
- Mullis, I. V., & Martin, M. O. (2017). *TIMSS 2019 Assessment Frameworks*. International Association for the Evaluation of Educational Achievement. Herengracht 487, Amsterdam, 1017 BT, The Netherlands.
- Mullis, I. V., & Martin, M. O. (2019). *PIRLS 2021 Assessment Frameworks*. Retrieved from: https://pirls2021.org/frameworks/wp-content/uploads/sites/2/2019/04/P21_FW_Ch1_Assessment.pdf.
- Mullis, I. V., Martin, M. O., & Sainsbury, M. (2016). *PIRLS 2016 reading framework*. PIRLS, 11-29.
- Mullis, I. V., Martin, M. O., Beaton, A. E., Gonzalez, E. J., Kelly, D. L., & Smith, T. A. (1998). *Mathematics and Science Achievement in the Final Year of Secondary School: IEA's Third International Mathematics and Science Study (TIMSS)*. TIMSS International Study Center, Boston College, School of Education,

- Campion Hall, Chestnut Hill, MA 02167; World Wide Web: <http://wwwwcsteep.bc.edu/timss>.
- Mullis, I. V., Martin, M. O., Foy, P., & Drucker, K. T. (2012). TIMSS & PIRLS International Study Center. Chestnut Hill, MA: Boston College.
- Mullis, I. V., Martin, M. O., Foy, P., Kelly, D. L., & Fishbein, B. (2020). TIMSS 2019 international results in mathematics and science. Boston College, TIMSS & PIRLS International Study Center website: <https://timssandpirls.bc.edu/timss2019/international-results>.
- Musa, N., & Agbaire, J. J. (2013). Methodological issues in comparative education studies: An exploration of the approaches of Kandel and Holmes. *Journal of educational review*, 6(3), 349-356.
- Muzzioli, P., Perazzolo, M., Poliandri, D., & Quadrelli, I. (2016). La qualità del percorso di autovalutazione. *Scuola democratica*, 7(2), 421-438.
- Muzzioli, P., Poliandri, D., Quadrelli, I., & Romiti, S. (2012). La valutazione nelle scuole in Italia: indicazioni e tendenze da alcuni progetti sperimentali. *Autonomie locali e servizi sociali*, 35(3), 463-478.
- Nelson, C., Treichler, P. A., & Grossberg, L. (1992). Cultural studies: An introduction. *Cultural studies*, 1(5), 1-19.
- Nelson, R., Ehren, M., & Godfrey, D. (2015). Literature review on internal evaluation. London: Institute of Education.
- Neumann, K., Fischer, H. E., & Kauertz, A. (2010). From PISA to educational standards: The impact of large-scale assessments on science education in Germany. *International Journal of Science and Mathematics Education*, 8(3), 545-563.
- Nevo, D. (2001). School evaluation: internal or external? *Studies in Educational Evaluation*, 27, 95-106.
- Niemann, D., Martens, K., & Teltemann, J. (2017). PISA and its consequences: Shaping education policies through international comparisons. *European Journal of Education*, 52(2), 175-183.
- Niemi, A. M., & Jahnukainen, M. (2018). Tuen tarve, työelämäpainotteisuus ja itsenäisyyden vaatimus ammatillisen koulutuksen kontekstissa. *Ammattikasvatuksen aikakauskirja*, 20(1), 9-25.
- Nokelainen, P., Asplund, R., Juujärvi, S., & Kovalainen, A. (2018). Muutokset haastavat ammatillisen koulutuksen. *Ammattikasvatuksen aikakauskirja*, 20(1), 4-8.
- Notti A.M. (2010). Valutazione e contesto educativo, Lecce, Pensa Editore.
- Notti A.M. (a cura di) (2014). A scuola di valutazione, Lecce, Pensa MultiMedia.
- Notti, A. M. (2020). La formazione del dirigente scolastico. Pp. 373-383. In *L'insegnante e il dirigente scolastico nella scuola dell'autonomia tra didattica, governance e progetto culturale*. Roma: Armando editore.

- Oakes, J. (1986): Educational Indicators: A guide for policymakers. CPRE (Occasional Paper Series).
- OECD (1992a). Education at a Glance Oecd Indicators. Regards sur l'éducation les indicateurs de l'Ocde. Paris: Oecd Publishing.
- OECD (1992b). The Oecd International Education Indicators: A Framework for Analysis. Paris: Oecd Publishing.
- OECD (2000), Measuring Student Knowledge and Skills: The PISA 2000 Assessment of Reading, Mathematical and Scientific Literacy, PISA, OECD Publishing, Paris, <https://doi.org/10.1787/9789264181564-en>.
- OECD (2001), Knowledge and Skills for Life: First Results from PISA 2000, PISA, OECD Publishing, Paris, <https://doi.org/10.1787/9789264195905-en>.
- OECD (2004), What Makes School Systems Perform? Seeing School Systems through the Prism of PISA, PISA, OECD Publishing, Paris, <https://doi.org/10.1787/9789264017726-en>.
- OECD (2007), PISA 2006: Science Competencies for Tomorrow's World: Volume 1: Analysis, PISA, OECD Publishing, Paris, <https://doi.org/10.1787/9789264040014-en>.
- OECD (2009a), "The Usefulness of PISA Data for Policy Makers, Researchers and Experts on Methodology", in PISA Data Analysis Manual: SAS, Second Edition, OECD Publishing, Paris, <https://doi.org/10.1787/9789264056251-2-en>.
- OECD (2009b). OECD Review on Evaluation and Assessment Frameworks for Improving School Outcomes – Design and Implementation Plan for the Review. <https://www.oecd.org/education/school/44568070.pdf>.
- OECD (2011), Lessons from PISA for the United States, Strong Performers and Successful Reformers in Education, France, OECD Publishing.
- OECD (2013). Synergies for Better Learning: An International Perspective on Evaluation and Assessment. OECD Reviews of Evaluation and Assessment in Education. OECD Publishing. <https://doi.org/10.1787/9789264190658-en>.
- OECD (2013b), OECD Skills Outlook 2013: First Results from the Survey of Adult Skills, OECD Publishing, Paris, <https://doi.org/10.1787/9789264204256-en>.
- OECD (2013c), The Survey of Adult Skills: Reader's Companion, OECD Publishing, Paris, <https://doi.org/10.1787/9789264204027-en>.
- OECD (2015), Adults, Computers and Problem Solving: What's the Problem?, OECD Skills Studies, OECD Publishing, Paris, <https://doi.org/10.1787/9789264236844-en>.
- OECD (2016a), The Survey of Adult Skills: Reader's Companion, Second Edition, OECD Skills Studies, OECD Publishing, Paris, <https://doi.org/10.1787/9789264258075-en>.

- OECD (2016b), *Skills Matter: Further Results from the Survey of Adult Skills*, OECD Skills Studies, OECD Publishing, Paris, <https://doi.org/10.1787/9789264258051-en>.
- OECD (2019a), *PISA 2018 Results (Volume I): What Students Know and Can Do*, PISA, OECD Publishing, Paris, <https://doi.org/10.1787/5f07c754-en>.
- OECD (2019b), *PISA 2018 Results (Volume II): Where All Students Can Succeed*, PISA, OECD Publishing, Paris, <https://doi.org/10.1787/b5fd1b8f-en>.
- OECD (2019c), *PISA 2018 Results (Volume III): What School Life Means for Students' Lives*, PISA, OECD Publishing, Paris, <https://doi.org/10.1787/acd78851-en>.
- OECD (2019d), *TALIS 2018 Results (Volume I): Teachers and School Leaders as Lifelong Learners*, TALIS, OECD Publishing, Paris, <https://doi.org/10.1787/1d0bc92a-en>.
- OECD (2019e), *The Survey of Adult Skills: Reader's Companion, Third Edition*, OECD Skills Studies, OECD Publishing, Paris, <https://doi.org/10.1787/f70238c7-en>.
- OECD (2019f), *Skills Matter: Additional Results from the Survey of Adult Skills*, OECD Skills Studies, OECD Publishing, Paris, <https://doi.org/10.1787/1f029d8f-en>.
- OECD (2020a), *PISA 2018 Results (Volume IV): Are Students Smart about Money?*, PISA, OECD Publishing, Paris, <https://doi.org/10.1787/48ebd1ba-en>.
- OECD (2020b), *PISA 2018 Results (Volume V): Effective Policies, Successful Schools*, PISA, OECD Publishing, Paris, <https://doi.org/10.1787/ca768d40-en>.
- OECD (2020c), *PISA 2018 Results (Volume VI): Are Students Ready to Thrive in an Interconnected World?*, PISA, OECD Publishing, Paris, <https://doi.org/10.1787/d5f68679-en>.
- OECD (2020d). *PISA National Project Manager Manual*. Retrieved from <https://search.oecd.org/pisa/pisaproducts/PISA-2022-National-Project-Manager-NPM-Manual.pdf>.
- OECD (2020e), *TALIS 2018 Results (Volume II): Teachers and School Leaders as Valued Professionals*, TALIS, OECD Publishing, Paris, <https://doi.org/10.1787/19cf08df-en>.
- OECD (2020f). *Learning remotely when schools close: How well are students and schools prepared? Insights from PISA*. Retrieved from: <https://www.oecd.org/coronavirus/en/policy-responses>.
- OECD (2020g), *Building Capacity for Evidence-Informed Policy-Making: Lessons from Country Experiences*, OECD Public Governance Reviews, OECD Publishing, Paris, <https://doi.org/10.1787/86331250-en>.
- OECD (2021a). *The State of School Education: One Year into the COVID Pandemic*. Retrieved from: <https://doi.org/10.1787/201dde84-en>.

- OECD (2021b), *Education at a Glance 2021: OECD Indicators*, OECD Publishing, Paris, <https://doi.org/10.1787/b35a14e5-en>.
- OECD (2021c), *The Assessment Frameworks for Cycle 2 of the Programme for the International Assessment of Adult Competencies*, OECD Skills Studies, OECD Publishing, Paris, <https://doi.org/10.1787/4bc2342d-en>.
- OECD (2022), *Education at a Glance 2022: OECD Indicators*, OECD Publishing, Paris, <https://doi.org/10.1787/3197152b-en>.
- Oktaý, J. S. (2012). *Grounded theory*. Oxford University Press.
- Oldenburg, R., & Brissett, D. (1982). The third place. *Qualitative sociology*, 5(4), 265-284.
- Oldham, E. (2003). From FIMS to PISA to our classrooms: Can large-scale cross-national studies help us to teach and learn differently?. *Irish Educational Studies*, 22(1), 1-27.
- Oliver, D. G., Serovich, J. M., & Mason, T. L. (2005). Constraints and opportunities with interview transcription: Towards reflection in qualitative research. *Social forces*, 84(2), 1273-1289.
- Olson Beal, H. K., & Hendry, P. M. (2012). The ironies of school choice: Empowering parents and reconceptualizing public education. *American Journal of Education*, 118(4), 521-550.
- Opetusministeriö (2010). *Perusopetuksen laatukriteerit*. Retrived from: <https://julkaisut.valtioneuvosto.fi/bitstream/handle/10024/75622/opm06.pdf>.
- Owen, G. T. (2014). Qualitative methods in higher education policy analysis: Using interviews and document analysis. *The qualitative report*, 19(26), 1.
- Paaso, A. (2012). Osaava ammatillinen opettaja 2020. *Ammattikasvatuksen aikakauskirja*, 14(3), 46-56.
- Paccagnella, L. (2000). *La comunicazione al computer*. Sociologia delle reti telematiche. Bologna: Il Mulino.
- Pagani, V. (2020). *Dare voce ai dati*. Edizioni Junior.
- Paletta, A. (2005). *Autovalutazione e gestione strategica delle attività educative*. Paper presentato al seminario nazionale sul tema "La valutazione del sistema scolastico", Rimini, 3-4.
- Paletta, A. (2013). *Valutazione dell'istruzione e miglioramento: il contributo degli studi di management*. Torino: Fondazione Agnelli.
- Pallotti, G. (1999). *I metodi della ricerca*. Galatolo R., Pallotti G.(a cura di), *La conversazione. Un'introduzione allo studio dell'interazione verbale*, R. Cortina editore, Milano, 365-407.
- Paloma, F. G., Poliandri, D., & Giampietro, L. (2020). Il Progetto Value for Schools: Ricerca pedagogica e Learning Analytics per l'autovalutazione delle scuole. *Formazione & insegnamento*, 18(1 Tome I), 294-307.

- Palomba, D. (2011). Gli studi comparativi in educazione: un'introduzione storico-critica. *ECPS Journal*, 4(2011), 29-45.
- Palomba, D. (2016). Studi comparativi in educazione: la dimensione teorica. In C. Callegari (a cura di), *L'educazione comparata tra storia ed etnografia* (pp. 23-49). Roma: Anicia.
- Palumbo M. (2001), *Il processo di valutazione. Decidere, programmare, valutare*, Milano, FrancoAngeli.
- Palumbo, M., & Pandolfini, V. (2016). Dalla valutazione al miglioramento scolastico: processi e attori. Evidenze dal progetto VALeS. *Scuola democratica*, 7(2), 351-368.
- Paolone, R.A. (2016). Accademica o applicata?. Educazione comparata, convergenze-divergenze globali ed etnografiche nella contemporaneità. In N.S. Barbieri, A. Gaudio, G. Zago (a cura di), *Manuale di educazione comparata. Insegnare in Europa e nel mondo* (pp. 43-77). Brescia: La Scuola.
- Parlett, M., & Hamilton, D. (1972). Evaluation as illumination: A new approach to the study of innovatory programs: in Occasional paper, Centre for Research in the Educational Sciences, University of Edinburgh. Edinburgh.
- Parlett, M., & Hamilton, D. (1976). Evaluation as illumination. In *Curriculum Evaluation Today: Trends and Implications*, Tawney, D. Schools Council.
- Parlett, M. (1976). Illuminative evaluation and research in education. In Appendix I of *Towards Integration: A Study of Blind and Partially Sighted Children in Ordinary Schools*, Jamieson, M., Parlett, M., & Pocklington, K. (1977), NFER, Windsor.
- Pasimeni, P. (2013). The Europe 2020 Index. *Social indicators research*, 110(2), 613-635.
- Pastori, G., & Guadagni, V. (2016). Le prove INVALSI nella scuola: processi organizzativi, comunicativi e informativi nell'esperienza di dirigenti, insegnanti e studenti. Uno studio qualitativo in Lombardia. In P. Landri, A. Maccarini (a cura di), *Uno specchio per la valutazione della scuola. Paradossi, controversie, vie d'uscita*, FrancoAngeli, Milano, 157-181.
- Patton M. Q. (1997). *Utilization Focused Evaluation*. The New Century Text, Beverly Hills (CA), Sage.
- Patton, M. Q. (2002). Two decades of developments in qualitative inquiry: A personal, experiential perspective. *Qualitative social work*, 1(3), 261-283.
- Pavan G. (2014). Come valutare la lettura nella scuola primaria. Una pratica virtuosa delle Indagini IEA. In C. Roverselli (a cura di), *Comprendere l'alterità. Comparazione, intercultura, Europa* (pp. 195-221). Roma: Aracne.
- Payne, J. W. (1994). Thinking aloud: Insights into information processing. *Psychological Science*, 5(5), 241-248.
- Pellerey, M. (2006). *Dirigere il proprio apprendimento*, Brescia, La Scuola.

- Peltola, N. (2021). Valmentavan kulttuurin johtaminen hajautetun organisaation etätyössä: Case Kansallinen koulutuksen arviointikeskus. Retrived from: https://www.theseus.fi/bitstream/handle/10024/506472/Opinn%C3%A4ytety%C3%B6_B6_Peltola_Niina_L8814.pdf?sequence=2.
- Perazzolo, M. (2019). Strumenti automatici a sostegno della lettura dei rav. In P. Lucisano, & A. Notti (Eds.), Training actions and evaluation processes. Atti del Convegno Internazionale SIRD, pp. 265–272. Pensa Multimedia. ISBN: 978-88-6760-634-4.
- Perla, L., & Vinci, V. (2016). Rapporti di Autovalutazione e Piani di Miglioramento: analisi e mappatura dei bisogni di formazione della Scuola pugliese. Primi step di una ricerca documentativa. *Italian Journal of Educational Research*, (17), 191-218.
- Perotti, L. (2010). Riforme universitarie e processo di Bologna. *Polis*, 24(1), 121-142.
- Pfau, B. N., & Kay, I. T. (2002). The five key elements of a total rewards and accountability orientation. *Benefits Quarterly*, 18(3), 7.
- Phillips, B. S., Baccianini, M., & Mortara, V. (1972). Metodologia della ricerca sociale. Il mulino.
- Plowright, D. (2007). Self-evaluation and Ofsted inspection: developing an integrative model of school improvement. *Educational Management Administration & Leadership*, 35(3), 373-393.
- Poliandri, D. (2016). Quattro contributi sull'autovalutazione delle scuole: un'introduzione. *Scuola democratica*, 7(2), 415-420.
- Poliandri, D., & Romiti, S. (2012). Osservare le scuole, valutarle e accompagnarle verso il miglioramento: il modello "Valutazione e miglioramento" proposto dall'invalsi. *Economia & lavoro*, 46(1), 61-0.
- Poliandri, D., Muzzioli, P., Quadrelli, I., & Romiti, S. (2013). Valutare per migliorare: un'esperienza da cui partire. *Italian Journal of Educational Research*, (10), 92-106.
- Poliandri, D., Perazzolo, M., Pillera, G. C., & Giampiero, L. (2022). Un'opportunità di formazione sull'autovalutazione delle scuole: il punto di vista di insegnanti e dirigenti. *Annali online della Didattica e della Formazione Docente*, 14(23), 169-184.
- Pollitt, C. (1995). Justification by works or by faith? Evaluating the new public management. *evaluation*, 1(2), 133-154.
- Pomante, L. (2016). La larga noche de la educación española. La larga noche de la educación española. In "History of Education and Children's Literature", XI, 2. EUM-Edizioni Università di Macerata, 439-450.
- Pons, X. (2011). What do we really learn from PISA? The sociology of its reception in three European countries (2001–2008) 1. *European Journal of Education*, 46(4), 540-548.

- Postlethwaite, T. N. (ed.) (1988). *The Encyclopedia of Comparative Education and National Systems of Education*. Oxford: Pergamon Press.
- Postlethwaite, T. N. (1995). International empirical research in comparative education. An example of the studies of the International Association for the Evaluation of Educational Achievement (IEA). *Tertium comparationis*, 1(1), 1-19.
- Purves, A. C. (1987a). The evolution of the IEA: A memoir. *Comparative Education Review*, 31(1), 10-28.
- Purves, A. C. (1987b). IEA an agenda for the future. *International Review of Education/Internationale Zeitschrift für Erziehungswissenschaft/Revue Internationale de l'Education*, 33(1), 103-107.
- Purves, A. C. (1992). Reflections on research and assessment in written composition. *Research in the Teaching of English*, 108-122.
- Putnam, R. D. (2004). Education, Diversity, Social Cohesion and "Social Capital". Note for discussion. Paper presented at the Meeting of OECD Education Ministers, Dublin.
- Quadrelli, I., Giampietro, L., Poliandri, D., & Romiti, S. (2018). L'autovalutazione delle scuole VALeS: il punto di vista degli attori. *VALUTAZIONE. STUDI E RICERCHE*, 115-138.
- Raffe, D., Branner, K., Croxford, L., & Martin, C. (1999). Comparing England, Scotland, Wales and Northern Ireland: The Case for "Home Internationals" in Comparative Research. *Comparative Education*, Vol. 35, No. 1, pp. 9-25.
- Ragin, C. C., & Becker, H. S. (Eds.). (1992). *What is a case? Exploring the foundations of social inquiry*. Cambridge university press.
- Raimondi, E., De Luca, S., & Barone, C. (2013). Origini sociali, risorse culturali familiari e apprendimenti nelle scuole primarie: un'analisi dei dati Pirls 2006. *Quaderni di Sociologia*, (61), 34-49.
- Rautalin, M. (2018). PISA and the criticism of Finnish education: justifications used in the national media debate. *Studies in Higher Education*, 43(10), 1778-1791.
- Ravecca, A. (2009). *Studiare nonostante*. Capitale sociale e successo scolastico degli studenti di origine immigrata nella scuola superiore. Milano: FrancoAngeli.
- Reding, V. (2002). Prefazione. In European Commission (Ed.), *Questione chiave dell'istruzione in Europa. La professione docente in Europa: profili, tendenze e sfide. Rapporto 1. Formazione iniziale e passaggio alla vita professionale*, vol.3. Bruxelles: Eurydice.
- Reimers, F. M., Schleicher, A. (2020). *Schooling Disrupted, Schooling Rethought: How the Covid-19 pandemic is changing education*. Retrieved from: https://read.oecd-ilibrary.org/view/?ref=133_133390-1rtuknc0hi&title=Schooling-disrupted-schooling-rethought-How-the-Covid-19-pandemic-is-changing-education.

- Repo, L., Paananen, M., Mattila, V., Lerkkanen, M. K., Eskelinen, M., Gammelgård, L., ... & Marjanen, J. (2018). Varhaiskasvatussuunnitelman perusteiden 2016 toimeenpanon arviointi: Varhaiskasvatussuunnitelmien käyttöönotto ja sisällöt. Julkaisut/Kansallinen koulutuksen arviointikeskus, (16) 2018.
- Reynolds, D., Sammons, P., De Fraine, B., Van Damme, J., Townsend, T., Teddlie, C., & Stringfield, S. (2014). Educational effectiveness research (EER): A state-of-the-art review. *School effectiveness and school improvement*, 25(2), 197-230.
- Reynolds, K., Khorramdel, L., & von Davier, M. (2022). Can students' attitudes towards mathematics and science be compared across countries? Evidence from measurement invariance modeling in TIMSS 2019. *Studies in Educational Evaluation*, 74, 101169.
- Robasto, D. (2017). *Autovalutazione e piani di miglioramento a scuola*. Roma: Carocci.
- Robasto, D. (2018). Dai rapporti di autovalutazione ai piani di formazione: una possibile analisi dei bisogni formativi inespressi. In *Scuola, Democrazia, Educazione. Formare ad una nuova società della conoscenza e della solidarietà* (pp. 979-985). Pensa Multimedia.
- Robitaille, D. F. (1990). Achievement comparisons between the first and second IEA studies of mathematics. *Educational Studies in Mathematics*, 21(5), 395-414.
- Rogge, N. (2019). EU countries' progress towards 'Europe 2020 strategy targets'. *Journal of Policy Modeling*, 41(2), 255-272.
- Roldán, M. C., & Giménez, R. J. C. (2022). Las leyes de educación en España vs resultados de evaluación del Informe Pisa. *Educatio Siglo XXI*, 40(1), 9-30.
- Romei, P. (1995). *Autonomia e progettualità*. Scandicci: La Nuova Italia.
- Romiti, S., Fabbro, F., & Mattarelli, E. (2022). Un modello formativo per le competenze valutative. Dalle teorie della formazione online e della valutazione alla piattaforma di e-learning. In *Book of Abstract* (pp. 197-199). CONVEGNO LA RICERCA EDUCATIVA PER LA FORMAZIONE DEGLI INSEGNANTI, Centro di Ricerca Educativa sulla Professionalità dell'Insegnante.
- Romiti, S., Fabbro, F., & Mattarelli, E. (2023). Collective feedback as a formative assessment practice in an e-learning platform for teachers' professional development. *QTimes Journal of Education Technology and Social Studies*, 15(1), 563-576.
- Rondini, R. (1995). *Per un servizio nazionale di valutazione*. Roma: Anicia.
- Rosenthal, L. (2004). Do school inspections improve school quality? Ofsted inspections and school examination results in the UK. *Economics of education review*, 23(2), 143-151.
- Rossi, P. H., Lipsey, M. W., & Henry, G. T. (2018). *Evaluation: A systematic approach*. Sage publications.

- Roulston, K., deMarrais, K., & Lewis, J. B. (2003). Learning to interview in the social sciences. *Qualitative inquiry*, 9(4), 643-668.
- Rumiati, R., Checchi, D., Ancaiani, A., Ciolfi, A., Sabella, M., Infurna, M. R., & Di Benedetto, A. (2019). Il problem solving come competenza trasversale. Inquadramento e prospettive nell'ambito del progetto TECO. *Scuola democratica*, 10(1), 239-258.
- Rumiati, R., Ciolfi, A., Di Benedetto, A., Sabella, M., Infurna, M. R., Ancaiani, A., & Checchi, D. (2018). Key-competences in higher education as a tool for democracy. *Form@ re-Open Journal per la formazione in rete*, 18(3), 7-18.
- Rusconi, M. (1995). La carta della scuola, in Aa. Vv., *La qualità della scuola, Annali della pubblica istruzione*, 5-6, 112-122.
- Rust, V. D., Soumaré, A., Pescador, O., & Shibuya, M. (1999). Research strategies in comparative education. *Comparative Education Review*, 43(1), 86-109.
- Sadler sir M. (1900). How far can we learn anything of practical value from the study of foreign systems of education?. Reprinted 1964 in *Comparative Education Review*, Vol. 7, No. 3, pp. 307-314. Sage.
- Sánchez, P. A., Abellán, C. M. A., & Sanz, M. P. G. (2015). Evaluación de planes de mejora en centros educativos de orientación inclusiva. *Profesorado, Revista de Currículum y Formación del Profesorado*, 19(3), 326-346.
- Sanmartín, O. (2014). Andalucía, País Vasco, Asturias, Cataluña y Canarias se resistirán a aplicar la Lomce. *El Mundo*. Retrived from: <http://www.elmundo.es/espana/2014/01/29/52e8fab122601d695c8b4572.html>.
- Saura, G., & Luengo, J. (2015). Política global más allá de lo nacional. Reforma educativa (LOMCE) y el régimen de estandarización (OCDE). *Bordón*, 67(1), 135-148. <http://dx.doi.org/10.13042/Bordon.2015.67109>.
- SAVE THE CHILDREN (2020), *La scuola che verrà*. Retrived from: https://s3.savethechildren.it/public/files/uploads/pubblicazioni/la-scuola-che-verra_0.pdf.
- Scheerens J. (2011a), Il controllo dei sistemi scolastici caratterizzati da una gerarchia a legami deboli. in Scheerens J., Mosca S., Bolletta R., a cura di, *Valutare per gestire la scuola. Governance, leadership e qualità educativa* (pp. 3-14), Pearson Italia, Milano-Torino.
- Scheerens, J. (2011b). Valutazione scolastica interna ed esterna. *Valutare per Gestire la Scuola. Governance, leadership e qualità educativa*, 59-82.
- Scheerens, J. (2019). *Efficacia ed Inefficacia Educativa, esame critico della Knowledge Base*. Springer.
- Scheerens, J., Mosca, S. & Bolletta, R. (2011), *Valutare per gestire la scuola. Governance, leadership e qualità educativa*, Milano, Bruno Mondadori.

- Scheerens, J., van Ravens, J., & Luyten, H. (2011). Un'applicazione di indicatori, basata sulla ricerca, per misurare la qualità dell'istruzione. *Journal of Educational, Cultural and Psychological Studies (ECPS Journal)*, 2(3), 53-75.
- Schildkamp, K. (2007). The utilisation of a self-evaluation instrument for primary education. *Enschede: University of Twente*, 186.
- Schildkamp, K., & Visscher, A. (2009). Factors influencing the utilisation of a school self-evaluation instrument. *Studies in Educational Evaluation*, 35, 150-159. doi:10.1016/j.stueduc.2009.12.001.
- Scholz, R. W., & Tietje, O. (2002). *Embedded case study methods: Integrating quantitative and qualitative knowledge*. Sage.
- Schön, D. A. (1993). *Il professionista riflessivo: per una nuova epistemologia della pratica professionale* (Trad. di A. Barbanente). Bari: Dedalo (Original work published. 1983).
- Sclavi, M. (2005). A una spanna da terra. Una giornata di scuola negli Stati Uniti e in Italia e i fondamenti di una metodologia umoristica. Pearson Italia Spa.
- Scotto Di Luzio, A. (2012). Vent'anni di storia della scuola, 1990-2010. Vent'anni di storia della scuola, 1990-2010, 35-50.
- Scriven M. (1967). The methodology of evaluation. In R.W. Tyler, R.M. Gagne, M. Scriven (eds.), *Perspectives of curriculum evaluation* (Vol. I, pp. 39-83). Chicago: Rand McNally.
- Scriven M. (1972), Pros and Cons about Goal Free Evaluation, «Evaluation comments», vol. 3, pp. 1-4.
- Scriven, M. (1982), *Evaluation Thesaurus*, Beverly Hills, Ca, Sage.
- Semeraro, R. (2011). L'analisi qualitativa dei dati di ricerca in educazione. *Italian Journal of educational research*, (7), 97-106.
- Sena, B. (2016). L'approccio del case study nella ricerca socio-economica. *SOCIOLOGIA E RICERCA SOCIALE*, (2016/111).
- Sena, B. (2021). Il case study nella ricerca sociale. Carrocci.
- Serpieri, R. (2016). Anarcheologia del canone valutativo. *Scuola democratica*, 7(2), 299-208.
- Serra, L. (2000). *Cultura organizzativa e autovalutazione nella scuola che cambia*. Armando Editore.
- Servicio De Inspeccion Educativa (2012). Sistema De Indicadores Para Valorar Los Planes De Mejora, Su Implantación Y Seguimiento. Retrived from: https://www.educacion.navarra.es/documents/57308/57761/Sistema_indic_sgto_planes_mejora.pdf/353bab4b-6f4d-435f-acca-cb1a19903f87.
- Sette, S., Bianco, P., Russo, S., Poliandri, D. & Epifani, G. (2018). Atteggiamiento verso l'autovalutazione e percezione degli effetti sulla scuola: uno studio sui membri dei nuclei di autovalutazione. *RIV Rassegna Italiana di Valutazione*, 71-72, pp. 258-275, DOI: 10.3280/RIV2018-071013.

- Shank, M. D., Walker, M., & Hayes, T. (1996). Understanding professional service expectations: do we know what our students expect in a quality education?. *Journal of Professional Services Marketing*, 13(1), 71-89.
- Shavelson, R. J., McDonnell, L., & Oakes, J. (1990a). Steps in designing an indicator system. *Practical Assessment, Research, and Evaluation*, 2(1), 12.
- Shavelson, R. J., McDonnell, L., & Oakes, J. (1990b). What are educational indicators and indicator systems?. *Practical Assessment, Research, and Evaluation*, 2(1), 11.
- Sicurello, R. (2020). Ricerca e analisi dei dati nell' Educazione Comparata. *STUDIUM EDUCATIONIS-Rivista semestrale per le professioni educative*, (2), 39-49.
- Siippainen, A., Paananen, M., Metsämuuronen, J., Repo, L., Sarkkinen, T., Alasuutari, M., ... & Kirjavainen, T. (2020). Viisivuotiaiden maksuttoman varhaiskasvatuskokeilun arvioinnin toinen vaihe: varhaiskasvatukseen osallistuminen, kokeilun kustannukset ja järjestäminen. *Julkaisut/Kansallinen koulutuksen arviointikeskus*.
- Silverman, D. (2000). Analyzing talk and text. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (2nd ed.), Thousand Oaks, CA: Sage, 821-834.
- Silverman, D. (2015). *Interpreting qualitative data*. Sage.
- Simons, H. (2013). Enhancing the quality of education through school self-evaluation In M. K. Lai & S. Kushner (Eds.), *A Developmental and Negotiated Approach to School Self-evaluation*. *Advances in Program Evaluation Volume 14*. Bingley, UK: Emerald Group Publishing.
- Simons, H. (2014). Case study research: In-depth understanding in context. *The Oxford handbook of qualitative research*, 455-470.
- Simpson, R. L., Lacava, P. G., & Sampson Graner, P. (2004). The no child left behind act: Challenges and implications for educators. *Intervention in school and clinic*, 40(2), 67-75.
- Sjoberg, G., Williams, N., Vaughan, T., & Sjoberg, A. (1991). The case study approach. In *A Case for the Case Study*. The University of North Carolina Chapel Hill.
- Sobe, N. W., & Kowalczyk, J. A. (2013). Exploding the Cube: Revisioning "Context" in the Field of Comparative Education. *Current Issues in Comparative Education*, 16(1), 6-12.
- Spallacci, A. (2015). Autavalutazione delle scuole e valutazione di sistema. Ruolo e competenze di chi valuta dall'interno. *Rassegna Italiana di Valutazione*, 19(61).
- Stake, R. (1975). *Evaluating the Arts in Education: A Responsive Approach*, Columbus, Ohio, Merril.
- Stake, R. (1980). Program Evaluation, particularly Responsive evaluation, in W. B. Dockrell, D. Hamilton (Eds.), *Rethinking educational research*, Hodder and Stoughton, London, pp. 72-87.

- Stake, R. (1983). Program Evaluation, particularly Responsive evaluation, in G. F. Madaus, M. Scriven, D. Stufflebeam, Evaluation models: viewpoints on educational and human services evaluation, Kluwer-Nijhoff Publishing, Boston, pp. 287-310.
- Stake, R. (1995). The art of case research. Sage.
- Stake, R. (2005). Qualitative case studies. In N. K. Denzin & Y. S. Lincoln (Eds.), The Sage handbook of qualitative research (3rd ed., pp. 433-466). Thousand Oaks, CA: Sage.
- Stake, R. E. (2013). Multiple case study analysis. Guilford press.
- Stame, N. (1998), L'esperienza della valutazione, Roma, SEAM.
- Stame, N. (2001), Tre approcci alla valutazione: distinguere e combinare. In M. Palumbo, Il processo di valutazione. Decidere, programmare, valutare, Milano, FrancoAngeli.
- Stame, N. (2018). Per una cultura della valutazione come strumento di miglioramento e di progresso di un sistema democratico. Rivista giuridica del Mezzogiorno, 32(2), 457-464.
- Stenhouse, L. (1985). A note on case study and educational practice. Field methods in the study of education, 263-271.
- Steyn, H. J., & Wolhuter, C. C. (2000). Education Systems of Emerging Countries: Challenges of the 21st Century. Noordbrug, South Africa: Keurkopie Uitgewers.
- Stiggins R. J. (2002). Assessment crisis: The absence of assessment for learning. Phi Delta Kappan. 83(10): 758-765.
- Stornaiuolo, R. (2021). L'orizzonte di senso del Sistema Integrato "zerosei": una sfida educativa e sociale. IUL Research, 2(4), 225-235.
- Strauss, A. & Corbin, J. (1998). Basics of qualitative research: Techniques and procedures for developing grounded theory (2nd ed.). London: Sage.
- Strauss, A., & Corbin, J. M. (1997). Grounded theory in practice. Sage.
- Stufflebeam D.L. (1973), Toward a Science of Educational Evaluation, Englewood Cliffs (NJ), Educational Technology Publications.
- Stumbriene, D., Camanho, A. S., & Jakaitiene, A. (2020). The performance of education systems in the light of Europe 2020 strategy. Annals of Operations Research, 288(2), 577-608.
- Suchman E. A. (1967), Evaluative Research: Principles and Practice in Public Service and Social Action Programs, New York, Russel sage Foundation.
- Sundberg, D. (2019). Three waves of education standardisation: How the curriculum changed from a matter of concern to a matter of fact. In New Practices of Comparison, Quantification and Expertise in Education (pp. 50-65). Routledge.

- Sutcliffe, W., & Pollock, J. (1992). Can the total quality management approach used in industry be transferred to institutions of higher education?. *The Vocational Aspect of Education*, 44(1), 11-27.
- Swaffield S. (2011), "Getting to the heart of authentic Assessment for Learning", *Assessment in Education: Principles, Policy and Practice*, 18, 4: 443-449.
- Sykes, G., Schneider, B., & Plank, D. N. (Eds.). (2012). *Handbook of education policy research*. Routledge.
- Tammaro, R., Calenda, M., & Iannotta, I. S. (2017). La valutazione: modelli teorici e pratiche. *Pedagogia più didattica, Rivista Erickson*, 3, 1.
- Tang, A., Li, W., & Liu, D. (2022). The Impact of Teachers' Professional Development in Science Pedagogy on Students' Achievement: Evidence from TIMSS 2019. *Journal of Baltic Science Education*, 21(2), 258-274.
- Taube, K., & Mejdung, J. (1996). A nine country study: what were the differences between the low-and high-performing students in the IEA Reading Literacy Study. *Reading literacy in an international perspective. Collected papers from the IEA reading literacy study*, 63-100.
- Terry, G., Hayfield, N., Clarke, V., & Braun, V. (2017). Thematic analysis. *The SAGE handbook of qualitative research in psychology*, 2, 17-37.
- Tight, M. (2017). *Understanding case study research: Small-scale research with meaning*. Sage.
- Tilley, S. A., & Powick, K. D. (2002). Distanced data: Transcribing other people's research tapes. *Canadian Journal of Education/Revue canadienne de l'éducation*, 291-310.
- Timm-Bottos, J., & Reilly, R. C. (2015). Learning in third spaces: Community art studio as storefront university classroom. *American journal of community psychology*, 55(1), 102-114.
- Tomba, B. (2019a). La responsabilità dirigenziale nella valutazione di sistema. *Atti del Convegno Internazionale SIRD. Training actions and evaluation processes*, 135-140.
- Tomba, B. (2019b). Il Dirigente scolastico promotore dell'innovazione. *Formazione & insegnamento*, 17(1 Suppl.), 187-192.
- Trethewey, A. R. (2014). *Introducing comparative education*. Elsevier.
- Tribus, M. (1994). *Total quality management in education. Developing quality system in education*. London: Routledge.
- Trincherò, R. (2002). *Manuale di ricerca educativa*. Milano: Franco Angeli.
- Trincherò, R. (2014a), *Metodi quantitativo-sperimentali, qualitative-ermeneutici e misti nella valutazione educativa e nella ricerca educativa*. In L. Galliani (a cura di), *L'agire valutativo*, Brescia, La Scuola, pp. 183-199.
- Trincherò, R. (2014b). Five research principles to overcome the dualism quantitative-qualitative. *Education Sciences & Society*, 5(1).

- Trincherò, R. (2014c). Il Servizio Nazionale di Valutazione e le prove Invalsi. Stato dell'arte e proposte per una valutazione come agente di cambiamento. *Form@re-Open Journal per la formazione in rete*, 14(4), 34-49.
- Trincherò, R. (2018). Attivare cognitivamente con la valutazione formante, in Notti A.M., a cura di, *La funzione educativa della valutazione*, pp. 73-90, Pensa, Lecce.
- Trincherò, R., & Robasto, D. (2019). I mixed methods nella ricerca educativa (pp. 1-138). Mondadori Education spa.
- Tyler R.W. (1942), General Statement on Evaluation, «The Journal of Educational Research», vol. 35, n. 7, pp. 492-501.
- Ukkola, A., & Metsämuuronen, J. (2021). Kehittyvät taidot, kehittyvä opetussuunnitelma. *Kieli, koulutus ja yhteiskunta*, 12(3).
- Van Velzen, W.G., Miles, M.B., Ekholm, M., Hameyer, U., & Robin, D. (Eds.) (1985). *Making school improvement work. A conceptual guide to practice (ISIP Book No.1)*, Leuven, Acco.
- Vanhoof, J., Van Petegem, P., & De Maeyer, S. (2009). Attitudes towards school self-evaluation. *Studies in Educational Evaluation*, 35, 21-28. doi:10.1016/j.stueduc.2009.01.004.
- Vannoni, M. (2015). What are case studies good for? Nesting comparative case study research into the lakatosian research program. *CrossCultural Research*, 49(4), 331-357.
- VanWynsberghe, R., & Khan, S. (2007). Redefining case study. *International journal of qualitative methods*, 6(2), 80-94.
- Varsori, A. (Ed.). (2007). *Alle origini del presente: l'Europa occidentale nella crisi degli anni Settanta (Vol. 101)*. Milano: FrancoAngeli.
- Venäläinen, S., Saarinen, J., Johnson, P., Cantell, H., Jakobsson, G., Koivisto, P., ... & Viitala, M. (2020). Näkymiä OPS-matkan varrelta: Esi- ja perusopetuksen opetussuunnitelmien perusteiden 2014 toimeenpanon arviointi. Retrieved from: https://helda.helsinki.fi/bitstream/handle/10138/314749/KARVI_0520.pdf?sequence=1.
- Vertanen, I. (2002). Ammatillinen opettajuus vuonna 2010: toisen asteen ammatillisen koulutuksen opettajan työn muutokset vuoteen 2010 mennessä. Hämeen ammattikorkeakoulu.
- Vertecchi, B. (2003). *Manuale della valutazione*. Milano: FrancoAngeli.
- Vertecchi, B. (2007). Aldo Visalberghi: un'idea della scuola e della società, in CADMO, anno 2007, numero 1, pp. 1-6.
- Vertecchi, B., Agrusti, G., & Losito, B. (2010). *Origini e sviluppi della ricerca valutativa*. Milano: FrancoAngeli.
- Vettenranta, J., Hiltunen, J., Kotila, J., Lehtola, P., Nissinen, K., Puhakka, E., ... & Ström, A. (2020a). Perustaidoista vauhtia koulutielle: Neljännen luokan

- oppilaiden matematiikan ja luonnontieteiden osaaminen: kansainvälinen TIMSS 2019-tutkimus Suomessa. Julkaisuja.
- Vettenranta, J., Hiltunen, J., Kotila, J., Lehtola, P., Nissinen, K., Puhakka, E., ... & Ström, A. (2020b). Tulevaisuuden avaintaidot puntarissa: Kahdeksannen luokan oppilaiden matematiikan ja luonnontieteiden osaaminen: kansainvälinen TIMSS 2019-tutkimus Suomessa. Julkaisuja.
- Vincent-Lacrin, S. (2019). *Measuring Innovation in Education 2019: What Has Changed in the Classroom?* Educational Research and Innovation. Paris: OECD Publishing.
- Visalberghi, A. (1951). La storia della filosofia nel pensiero di Dewey. *Rivista di Storia Della Filosofia*, 6(4).
- Visalberghi, A. (1955). *Misurazione e valutazione nel processo educativo*, Milano: Edizioni di Comunità.
- Visalberghi, A. (1958), *Esperienza e valutazione*. Torino: Taylor.
- Visalberghi, A. (1977a). Rapporto generale sui risultati IEA in Italia e sulle ricerche connesse. Valutazione complessiva dei risultati. In *Misurazione del rendimento scolastico. Indagini IEA e situazione italiana. Quaderni degli Annali della Pubblica Istruzione*, 5 (1977), pp. 125-150.
- Visalberghi, A. (1977b). *Filosofia dell'educazione. Scuola e Città*, 28(12). *La Nuova Italia*, 521-535.
- Visalberghi, A. (1978). *Pedagogia e scienze dell'educazione*. Milano: Mondadori.
- Visalberghi, A. (1988). *Insegnare ed apprendere. Un approccio evolutivo*. *La Nuova Italia*.
- Vitalini, A. (2010). L'uso delle reti sociali per la costruzione di campioni probabilistici: possibilità e limiti per lo studio di popolazioni senza lista di campionamento. *Studi di Sociologia*, 383-398.
- Viteritti, A., & Giancola, O. (2015). Il ruolo delle grandi survey in campo educativo. L'indagine PISA e il governo dell'educazione tramite i numeri. *Rassegna Italiana di Sociologia*, 56(3-4), 555-580.
- Vlasov, J., Sarkkinen, T., & Harkoma, S. (2020). *Varhaiskasvatuksen laadun arvioinnin kehittäminen ja järjestäjien tukeminen vuosina 2020–2023*. Retrieved from: https://karvi.fi/wp-content/uploads/2021/08/Vaka_Hankesuunnitelma_Laadunhallinta.pdf.
- Volante, L. (2004). Teaching to the Test: What Every Educator and Policy-Maker Should Know. *Canadian Journal of Educational Administration and Policy*.
- Volante, L. (2015). The Impact of PISA on Education Governance: Some Insights from Highly Reactive Policy Contexts. *International Studies in Educational Administration (Commonwealth Council for Educational Administration & Management (CCEAM))*, 43(2).

- Wach, E., & Ward, R. (2013). Learning about qualitative document analysis. Retrieved from: <https://opendocs.ids.ac.uk/opendocs/bitstream/handle/20.500.12413/2989/PP%20InBrief%2013%20QDA%20FINAL2.pdf?sequence=4&isAllowed=y>.
- Wahba, M. A., & Bridwell, L. G. (1976). Maslow reconsidered: A review of research on the need hierarchy theory. *Organizational behaviour and human performance*, 15(2), 212-240.
- Waldow, F. (2009). What PISA did and did not do: Germany after the 'PISA-shock'. *European Educational Research Journal*, 8(3), 476-483.
- Walker, D. A. (1976). The IEA six subject survey: An empirical study of education in twenty-one countries. Halsted.
- Walker, D., & Myrick, F. (2006). Grounded theory: An exploration of process and procedure. *Qualitative health research*, 16(4), 547-559.
- Warner-Griffin, C., Liu, H., Tadler, C., Herget, D., & Dalton, B. (2017). Reading Achievement of US Fourth-Grade Students in an International Context: First Look at the Progress in International Reading Literacy Study (PIRLS) 2016 and ePIRLS 2016. NCES 2018-017. National Center for Education Statistics.
- Watkins A. (2007). Assessment in inclusive settings: key issues for policy and practice, Odense: European Agency for Development in Special Needs.
- Weeden, P., Winter, J., & Broadfoot, P. (2009). Valutazione per l'apprendimento nella scuola. Strategie per incrementare la qualità dell'offerta formativa. Edizioni Erickson.
- Weinstein, D., & Weinstein, M. A. (1991). Georg Simmel: sociological flaneur bricoleur. *Theory, Culture & Society*, 8(3), 151-168.
- Weiss A.J., Gruber, E. J. (1987). The Managed Irrelevance of Federal Education Statistics. In: Alonso W. & S. P. (Eds.). *The Politics of Numbers*. (pp. 363 - 391). New York, Russel Sage Foundation.
- Weiss, C. (1972), *Evaluation Researches*, Englewood Cliffs, NJ, Prentice Hall.
- Wiggins G. (1998). *Educative Assessment: Designing Assessments to Inform and Improve Student Performance*. San Francisco: Jossey-Bass.
- William D. (2000). Integrating formative and summative functions of assessment. Paper presented to Working Group 10 of the International Congress on Mathematics Education.
- William D. (2011). What is assessment for learning. *Studies in Educational Evaluation*, 37: 3-14.
- Wilkins, D., & Antonopoulou, V. (2020). Do performance indicators predict Ofsted ratings? An exploratory study of children's services in England. *Journal of Children's Services*.

- Wolf R.L. (1979), The Use of Judicial Evaluation methods in the Formulation of Educational Policy, «Educational Evaluation and Policy Analysis», vol. 1, pp. 19-28.
- Wolf, R. M. (1987). Educational evaluation: the state of the field. *International Journal of Educational Research*, 11(1), 1-143.
- Yardley, L. (2008). Demonstrating validity in qualitative psychology. *Qualitative psychology: A practical guide to research methods*, 2, 235-251.
- Yazan, B. (2015). Three approaches to case study methods in education: Yin, Merriam, and Stake. *The qualitative report*, 20(2), 134-152.
- Yin, L., & Fishbein, B. (2019). Creating and interpreting the TIMSS 2019 context questionnaire scales. *Methods and procedures: TIMSS*, 16-1.
- Yin, R. K. (1994a). Designing Single-and Multiple-Case. *Improving educational management: Through research and consultancy*, 135.
- Yin, R. K. (1994b). Discovering the future of the case study. *Method in evaluation research. Evaluation practice*, 15(3), 283-290.
- Yin, R. K. (2003a). Design and methods. *Case study research*, 3(9.2).
- Yin, R. K. (2003b). Designing case studies. *Qualitative research methods*, 5(14), 359-386.
- Yin, R. K. (2009). *Case study research: Design and methods*. Sage.
- Yin, R. K. (2018). *The Case Study Research and Application. Design and Methods* (6a ed.). Sage.
- Zambelli, A., De Carlo, E., Mazzetti, G., & Guglielmi, D (2021). Il ruolo dei genitori nel processo di scelta della scuola secondaria. *Counseling, Riviste Digitali Erickson*, (14), 41-61, doi: 10.14605/CS1422104.
- Zammuner, V.L. (1996). *Interviste e questionari*. Roma, Borla Edizioni.
- Zammuner, V.L. (1998). *Tecniche dell'intervista e del questionario*. Bologna: Il Mulino.
- Zanazzi, S. (2014). Evaluating and financing university research. A comparative case study: Italy, France, Spain and Germany. Roma: Nuova Cultura.
- Zanazzi, S., & Corsini, C. (2015). Valutare scuola e università: approccio emergente, interventi e criticità. *I problemi della pedagogia*, 2, 305-334.
- Zanchin, M. R. (Ed.). (2002). *I processi di apprendimento nella scuola dell'autonomia: analisi disciplinare e personalizzazione dei talenti*. Armando Editore.
- Zanniello, G. (2016). La valutazione e l'autovalutazione delle scuole italiane. *Pedagogia e Vita*, 74, 92-112.
- Zanniello, G. (2017). Autovalutazione e valutazione d'Istituto. *VALUTARE*, 10, 91-102.
- Zhao, Y. (2020). Two decades of havoc: A synthesis of criticism against PISA. *Journal of Educational Change*, 21(2), 245-266.

Bibliografia

- Zimmerman, R. (1986). The management of risk. In V. T. Covello, J. Menkes, & J. Munpower (Eds.), *Risk evaluation and management* (pp. 435–460). New York: Plenum.
- Zizioli E. (2009). Luigi Volpicelli. Un idealista “fuori dalle formule”. Roma: Anicia.

Sitografia

(ultima consultazione 21 febbraio 2023)

<https://boe.es/buscar/act.php?id=BOE-A-2006-7899&p=20220727&tn=0#tvi>

https://en.wikipedia.org/wiki/A_Nation_at_Risk

https://it.wikipedia.org/wiki/Organizzazione_per_la_cooperazione_e_lo_sviluppo_economico

https://it.wikipedia.org/wiki/Qualit%C3%A0_totale

<https://karvi.fi/>

<https://okm.fi/etusivu>

https://op.europa.eu/webpub/eac/education-and-training-monitor-2020/countries/finland_fi.html#five

<https://op.europa.eu/webpub/eac/education-and-training-monitor-2020/countries/spain.html#five>

<https://www.educacionyfp.gob.es/inee/portada.html>

<https://www.educacionyfp.gob.es/portada.html>

<https://www.etimo.it/?term=valutare>

<https://www.gov.uk/government/organisations/ofsted/about>

<https://www.iea.nl/about/org/history>

<https://www.iea.nl/studies/iea/pirls>

<https://www.invalsi.it/invalsi/index.php>

<https://www.invalsi.it/invalsi/ri/vales/index.php>

<https://www.merriam-webster.com/dictionary/case%20study>

<https://www.miur.gov.it/>

Ringraziamenti

Se il punto in cui ti immergi in un fiume è il presente, pensai, allora il passato è l'acqua che ti ha superato, quella che va verso il basso e dove non c'è più niente per te, mentre il futuro è l'acqua che scende dall'alto, portando pericoli e sorprese.

Il passato è a valle, il futuro a monte.

~ Le otto montagne, Paolo Cognetti

Siamo giunti alle ultime pagine di questo lavoro e alle ultime battute del percorso di dottorato. È ora di scendere per un momento dalla bicicletta, quella che quando si hanno ossigeno, polmoni e un buon equipaggiamento può portare ovunque lo si desideri, e guardare indietro alla strada percorsa. È una strada di montagna, battuta, caratterizzata da salite e discese, da una pandemia mondiale che mai avrei immaginato di incrociare lungo il sentiero, ma anche da panorami bellissimi e fontane di acqua fresca per dissetarsi e ristorarsi. Una strada che, lungo le pedalate, mi ha dato la possibilità di incontrare alcune persone che è doveroso ringraziare. Alcune sono state presenti fin dall'inizio del cammino, altre le ho conosciute a metà, quando eravamo tutti chiusi nelle nostre case a combattere con il diffondersi di un virus, altre ancora verso la fine.

Il primo ringraziamento va a Laura Palmerio, Arto Ahonen e Lis Cercadillo, i National Project Manager dell'indagine PISA dell'OCSE, e ai dirigenti scolastici, incontrati dal vivo, tramite un form di Google Moduli o nella scuola finlandese di Ruovesi, senza i quali questo lavoro di ricerca oggi non esisterebbe. Thank you for sharing your experiences with me.

Il secondo ringraziamento va alla Sapienza, mamma Sapienza, e al Dipartimento di Psicologia dei Processi di Sviluppo e Socializzazione che, nei corsi di laurea triennale e magistrale prima e nella scuola di dottorato poi, mi hanno accolta che non ero nemmeno ventenne e mi hanno formata, fino a diventare la persona che sono oggi, fuori e dentro l'università, dentro e fuori la ricerca.

Ringrazio in particolar modo il mio tutor, il prof. Asquini, con il quale da qualche anno ormai studiamo e discutiamo delle questioni educative

che tanto ci appassionano. Per la sua solarità e per la guida sapiente e saggia dal primo giorno in cui, durante l'esame orale che sostenevo da studentessa del corso di laurea triennale, mi disse "quindi tu saresti italo-finlandese... Cosa ci fai qui allora?". Quella non risposta e quel sorriso abbozzato mi hanno portata dove sono oggi.

Ringrazio il mio co-tutor, il prof. Lastrucci, che nel momento in cui mi sono fermata a un bivio, come può accadere su qualsiasi strada, ha saputo mettersi al mio fianco per continuare il cammino.

Ringrazio tutti i professori e le professoresse di Villa Mirafiori, l'accogliente Villa Mirafiori, dove è sempre bello e stimolante tornare anche solo per incontrarvi.

Ringrazio i colleghi dottorandi e dottori di ricerca, presenti e passati. È stato prezioso condividere con voi gioie e dolori di chi si avvicina a questo mondo. Ci auguro un bellissimo proseguimento di strada.

Ringrazio i colleghi di Invalsi, in particolar modo dell'Area 5 – Innovazione e Sviluppo, con i quali ho condiviso piacevoli e arricchenti giornate lavorative. Grazie per il tempo trascorso assieme a fare ricerca e per i caffè durante le pause.

Ringrazio i miei amici, quelli di sempre e quelli che si sono aggiunti durante questo viaggio. Grazie perché mi supportate e sopportate, sapete ascoltarmi e farmi ridere.

Ringrazio il mio compagno di avventure, esempio di tenacia e determinazione e portatore sano di disordine meraviglioso. Senza di te la mia vita, soprattutto quella degli ultimi mesi, non sarebbe stata così dinamica e movimentata.

Ringrazio la mia amata famiglia, italiana e finlandese, fonte di intraprendenza e punti di riferimento. Alle mie nonne, ai nonni che non ci sono più. Ai miei genitori da sempre e per sempre miei sostenitori. Grazie dal più profondo del cuore, tuhannet kiitokset!

Finito di stampare nel mese di maggio 2023
presso il Centro Stampa Pioda, Roma