



Evaluating Educational Quality

A cura di
Cristiano Corsini,
Giuseppe C. Pillera,
Christopher H. Tienken,
Maria Tomarchio



Processi
e Linguaggi
dell'Apprendimento

FrancoAngeli
OPEN  ACCESS



Processi
e Linguaggi
dell'Apprendimento

Direzione: Roberto Trincherò

Comitato direttivo

Funzioni: accoglienza delle proposte di pubblicazione e prima scrematura

Barbara Bruschi, Renato Grimaldi, Roberto Farné, Alberto Parola, Daniela Robasto, Barbara Sini, Simona Tirocchi

Comitato Scientifico

Funzioni: referaggio anonimo, con doppio cieco, mediante sistema on line

Michele Baldassarre, Federico Batini, Guido Benvenuto, Giovanni Bonaiuti, Vincenzo Bonazza, Antonio Calvani, Gianna Cappello, Lucia Chiappetta Cajola, Cristina Coggi, Barbara Demo, Luciano Di Mele, Piergiuseppe Ellerani, Ivan Enrici, Damiano Felini, Adelaide Gallina, Marco Gui, Sara Nosari, Alessandro Perissinotto, Maria Ranieri, Paola Ricchiardi, Emanuela Torre, Carla Tinti, Giuliano Vivinet, Tamara Zappaterra.

La Collana accoglie studi teorici, storico-comparativi ed empirico-sperimentali riguardanti i processi e i linguaggi dell'apprendimento dalla primissima infanzia alla "grande anzianità". I testi proposti sono volti a indagare "come si apprende" nelle varie età della vita e come è possibile mettere in atto processi di formazione efficaci nel promuovere apprendimento, tenendo conto del dibattito contemporaneo in pedagogia, didattica, psicologia cognitiva, neuroscienze. In quest'ottica, i testi proposti esplorano i metodi, le strategie, le tecniche e gli strumenti efficaci nei percorsi di educazione, istruzione e formazione, scolastica ed extrascolastica, lungo tutto l'arco della vita.

Oggetti di interesse sono quindi l'educazione e la formazione improntate dall'evidenza quantitativa e qualitativa, l'apprendimento esperienziale in diversi contesti - dal gioco spontaneo del bambino all'interazione mediata dai social network -, i linguaggi medialti per l'apprendimento e le tecnologie in grado di promuoverlo, il potenziamento cognitivo come strumento per affrontare un vasto spettro di bisogni educativi, la *gamification*, la robotica educativa, la giocomotricità e le sinergie tra apprendimento cognitivo e motorio, lo *storytelling*, i prodotti mono e multimediali per l'infanzia e il gioco educativo nelle sue varie forme e accezioni.

La collana accoglie contributi di studiosi italiani e di altri paesi, sotto forma di monografie, volumi collettanei, rapporti di ricerca, traduzioni, descrizioni di esperienze e sperimentazioni in contesti scolastici ed extrascolastici.

Il Comitato direttivo e il Comitato scientifico intendono promuovere attraverso la collana un ampio, aperto e proficuo dibattito tra ricercatori, insegnanti, educatori e tutti gli studiosi che siano interessati ai processi e ai linguaggi dell'apprendimento nelle varie età della vita.

Ogni volume è sottoposto a referaggio con modello "doppio cieco".



Il presente volume è pubblicato in open access, ossia il file dell'intero lavoro è liberamente scaricabile dalla piattaforma **FrancoAngeli Open Access** (<http://bit.ly/francoangeli-oa>).

FrancoAngeli Open Access è la piattaforma per pubblicare articoli e monografie, rispettando gli standard etici e qualitativi e la messa a disposizione dei contenuti ad accesso aperto. Oltre a garantire il deposito nei maggiori archivi e repository internazionali OA, la sua integrazione con tutto il ricco catalogo di riviste e collane FrancoAngeli massimizza la visibilità, favorisce facilità di ricerca per l'utente e possibilità di impatto per l'autore.

Per saperne di più:

http://www.francoangeli.it/come_publicare/publicare_19.asp

I lettori che desiderano informarsi sui libri e le riviste da noi pubblicati possono consultare il nostro sito Internet: www.francoangeli.it e iscriversi nella home page al servizio "Informatemi" per ricevere via e-mail le segnalazioni delle novità.

Evaluating Educational Quality

A cura di

Cristiano Corsini,
Giuseppe C. Pillera,
Christopher H. Tienken,
Maria Tomarchio

FrancoAngeli
OPEN  ACCESS

Volume stampato con il contributo dell'Università degli Studi "G. d'Annunzio" di Chieti-Pescara – Dipartimento di Scienze filosofiche, pedagogiche ed economico-quantitative.

Isbn elettronico 9788835103738

Copyright © 2020 by FrancoAngeli s.r.l., Milano, Italy.

Pubblicato con licenza *Creative Commons Attribuzione-Non Commerciale-Non opere derivate 4.0 Internazionale* (CC-BY-NC-ND 4.0)

L'opera, comprese tutte le sue parti, è tutelata dalla legge sul diritto d'autore. L'Utente nel momento in cui effettua il download dell'opera accetta tutte le condizioni della licenza d'uso dell'opera previste e comunicate sul sito

<https://creativecommons.org/licenses/by-nc-nd/4.0/deed.it>

Copyright © 2020 by FrancoAngeli s.r.l., Milano, Italy. ISBN 9788835103738

Indice

Prefazione. Sulla differenza tra la valutazione in educazione e la valutazione educativa , di <i>Cristiano Corsini</i>	Pag.	7
Ipercomplessità della valutazione educativa: contesti, pratiche, ricerche , di <i>Cristiano Corsini, Giuseppe C. Pillera, Christopher H. Tienken, Maria Tomarchio</i>	»	11
Dentro i processi della didattica. Osservare e valutare per la professionalità dell'insegnante della scuola dell'infanzia , di <i>Ira Vannini</i>	»	17
La <i>global data infrastructure</i> di PISA e la politica OCSE nascosta in piena vista: <i>data infrastructures</i> alternative per valutare la qualità dell'educazione , di <i>Camilla Addey</i>	»	29
Valutazione dei docenti e valutazione del merito in Italia: lezioni dall'America , di <i>Christopher H. Tienken</i>	»	39
Emerging policies and practices of teaching evaluation in the Norwegian school system , di <i>Silje Gløppen e Guri Skedsmo</i>	»	55

La valutazione delle competenze nei CPIA: un caso studio , di <i>Germana Patti e Giuseppe M. Tomaselli</i>	»	77
L'esperienza di tirocinio esterno degli studenti fuoricorso in Scienze pedagogiche dell'Università di Catania. Comparazione con studenti regolari e strategie di miglioramento , di <i>Giuseppe C. Pillera</i>	»	83
Postfazione. A conti fatti... , di <i>Maria Tomarchio</i>	»	111
Gli autori	»	115

Prefazione. Sulla differenza tra la valutazione in educazione e la valutazione educativa

di *Cristiano Corsini*

Nell'ultimo ventennio si sono sviluppati apparati, attività e istituti appositamente deputati alla valutazione dell'efficacia, dell'efficienza, della qualità del lavoro svolto da scuole e università. Un fugace e distratto sguardo su questo fenomeno sembrerebbe confermare l'impressione che qualcosa si stia muovendo sotto il cielo della valutazione educativa. E in effetti, quando leggiamo le diffuse analisi che generalmente i media riservano all'uscita di rapporti sulle rilevazioni INVALSI, su indagini internazionali come PISA o sui *ranking* di scuole e università non è facile resistere alla tentazione di inquadrare tanto sfoggio di numeri, indicatori "oggettivi" e classifiche su istituti e accademie entro l'ampia e apparentemente scontata locuzione di "valutazione educativa".

Tuttavia, non appena aguzziamo se non l'ingegno perlomeno la vista, il fenomeno tende ad assumere un altro aspetto. Se provassimo a prendere in esame riflessioni, esperienze e analisi relative alla valutazione educativa succedutesi nell'ultimo secolo – iniziando dai lavori di Piéron sull'inaffidabilità delle commissioni d'esame in Francia e dalla teoria deweyana della valutazione per procedere con la valutazione formativa di Scriven, declinata più recentemente, attraverso la mediazione della psicologia dell'apprendimento, in *assessment for learning* e in *assessment as learning* o, a livello di valutazione di programma, nella quarta generazione di Guba e Lincoln e nella *developmental evaluation* di Patton – faremmo fatica a riconoscere alcunché di rigorosamente valutativo ed educativo in gran parte dei processi che sono stati prima richiamati. Infatti, pur tenendo conto delle differenze tra scuole e oggetti di riflessione (*assessment* ed *evaluation* non sono certo termini del tutto sovrapponibili), è possibile riscontrare una definizione comune di valutazione educativa trasversale alle riflessioni e alle esperienze messe in campo in questi decenni. Possiamo definire la valutazione educativa un giudizio di valore, espresso sulla distanza riscontrata tra obiettivi e realtà, utile alla riduzione di tale distanza.

Emergono dunque tre dimensioni: la valutazione come giudizio, come accertamento/misurazione e come riscontro in grado di informare un processo di cambiamento. Ed è rispetto a queste tre dimensioni che buona parte di quella che oggi è definita valutazione educativa segna il passo, come rilevato nei diversi contributi raccolti in questo volume. Un volume che restituisce una diversità di sguardi che si sono incontrati in un Convegno che ha rappresentato la tappa di un processo e un dialogo agevolati, all'interno dell'Università di Catania, dall'integrazione, fortemente voluta dalla prof.ssa Tomarcho, tra l'insegnamento di *Modelli di progettazione pedagogica e politiche educative* e quello di *Metodi e tecniche di valutazione*.

In primo luogo, il fatto di concepire la valutazione come giudizio esclude che essa possa essere oggettiva. Il processo valutativo è una forma di gestione del potere e chi valuta dà risposte alle questioni che lo informano (perché? come? cosa? chi?). L'oggettività, brandita come arma nel discorso pubblico sulla valutazione di scuole e università, nasconde in realtà meccanismi di mistificazione e di rimozione propri di una visione autoritaria, non partecipata né democratica. È sorprendente come questa retorica dell'oggettività – fondata sull'incapacità o sul rifiuto di rendere conto del complesso intreccio di rapporti tra i dati che vengono diffusi e gli strumenti, le ipotesi, le scelte metodologiche e valoriali che inevitabilmente li hanno prodotti – alimenti, agendo in modo così opaco, il mito della trasparenza.

In secondo luogo, misurare è necessario, ma lo è perché il giudizio di valore va espresso sulla distanza che noi riscontriamo tra le cose come sono e le cose come vorremmo che fossero. La misura di questa distanza, che dovrebbe caratterizzare ogni processo valutativo, va raccolta con senso della misura e deve farsi carico di conferire affidabilità al giudizio, ma anche di rappresentare i diversi valori in campo, perché altrimenti risulterebbe iniqua, soprattutto se imposta dall'alto, dettata da logiche di standardizzazione imposte da scarsa sensibilità ai contesti, da una cieca fiducia nella razionalità assoluta o da esigenze di contenimento dei costi. Se è vero che la valutazione conserva una sua soggettività è vero anche che soggettività non significa arbitrio: condividere e negoziare non solo i fini della valutazione ma anche i mezzi della misurazione consente di strutturare in maniera democratica e partecipata il processo valutativo, legando tra loro affidabilità ed equità.

Infine, la terza caratteristica della valutazione, che è quella più importante: il giudizio di valore emesso sulla distanza tra realtà e obiettivi deve servire a ridurre questo divario, altrimenti della valutazione non ce ne facciamo niente ed essa diventa una inutile e burocratica perdita di tempo, che tende a riprodurre la realtà più che a trasformarla. La messe di indicatori rovesciata su un'opinione pubblica ansiosa di stilare classifiche non ha nulla a che vedere col riscontro valutativo: questo, infatti, deve farsi carico di

informare in maniera affidabile strategie di miglioramento e deve perciò essere criteriiale e non meramente normativo. Conoscere la distanza tra il punteggio di uno studente e il punteggio medio della sua classe, della sua scuola, della sua regione, della media OCSE, della Finlandia o di Saturno di per sé non aiuta a informare i processi educativi.

Misuriamo per valutare, valutiamo per educare. I processi di elaborazione del giudizio che non rendono conto del complesso rapporto tra misurazione e valutazione non meritano di essere definiti valutativi, così come non possiamo definire educative le valutazioni che non hanno un impatto trasformativo e migliorativo sui soggetti e i contesti che chiamano in causa. I lavori ospitati nel presente volume accettano la sfida della valutazione educativa restituendo esperienze diverse, analisi critiche, punti di forza e di debolezza di un processo che, ambendo a mettere in discussione l'agire educativo, non può continuare a sottrarsi al giudizio: valutare la valutazione è scelta sostenibile e doverosa dal punto di vista educativo e valutativo.

Ipercomplessità della valutazione educativa: contesti, pratiche, ricerche

di *Cristiano Corsini, Giuseppa C. Pillera, Christopher H. Tienken,
Maria Tomarchio.*

I contributi qui ospitati sono stati sollecitati dal “Gruppo di ricerca sugli effetti dell’accountability sulla qualità dei processi educativi”, formato originariamente da Corsini, Tienken e Tomarchio. Il gruppo, al quale Pillera si è unito dando un contributo fondamentale nell’organizzazione del convegno catanese e nella curatela di questo volume, si occupa di sottoporre ad analisi critica l’incidenza di fenomeni nazionali e sovranazionali di rendicontazione e standardizzazione di procedure volte al controllo della qualità, dell’efficacia e dell’efficienza in educazione. Il convegno del 2017 ha rappresentato la prima tappa di un percorso di raccolta di una pluralità di voci e contesti sulla tematica.

Il contributo di Ira Vannini, seguendo le orme del volume di Visalberghi *Misurazione e valutazione nel processo educativo* (1955), riflette criticamente sul rapporto tra didattica e valutazione, intendendo quest’ultima come imprescindibile strumento di ricerca per l’avanzamento della prima. Le considerazioni sul testo visalberghiano vengono poste in dialogo con le modalità di sperimentazione di una scala a supporto dell’osservazione delle prassi didattiche dei maestri nella scuola dell’infanzia (PraDISI), un’ipotesi di lavoro aperta nella direzione di una valutazione autenticamente formativa, posta cioè al servizio dello sviluppo della professionalità docente. Una valutazione così concepita si muove all’interno del circuito virtuoso tra misurazione e giudizio critico, in cui il secondo potrà tanto più rappresentare un fecondo tentativo di risposta valida alle questioni educative, quanto la prima saprà essere rigorosa, scevra da pregiudizi e attagliata all’oggetto e al contesto che vuole saggiare. Proprio come realizzato all’interno del percorso di validazione dello strumento presentato, si tratta di tracciare direttrici di collaborazione tra scuola e università che invitino a un protagonismo del docente, chiamato alla responsabilità di un impegno individuale e collegiale proprio a partire dal vaglio e dalla definizione di indicatori e strumenti per l’osservazione, la misurazione e dunque la valutazione delle proprie pratiche didattiche.

Prospettive di Ricerca-Formazione (Asquini, 2018) come quelle tratteggiate da Vannini si pongono in diretta antitesi con le tendenze alle valutazioni su larga scala dei sistemi educativi nazionali, basate su test standardizzati e portate avanti da organizzazioni internazionali, come l'OCSE, interessate soltanto in chiave strumentale alle questioni dell'istruzione e dell'educazione. Se è vero che PISA prescinde dai curricula, bisogna vedere quanto i curricula prescindano da PISA, ossia quanto la semplificazione che ogni misurazione porta con sé non semplifichi a sua volta la didattica, impoverendo la qualità dell'insegnamento e dell'apprendimento. Alla luce della tesi avanzata da Sellar e Lingard (2013), Camilla Addey, che da anni riflette su tali questioni muovendosi come ricercatrice tra diversi punti di osservazione privilegiata su diversi continenti, porta a segno nel suo contributo una lucida critica alla pervasività dimostrata dalla *global education governance* dell'OCSE, che, attraverso il programma PISA for Development, sta puntando a omologare la valutazione dei sistemi educativi dei paesi emergenti al modello PISA, ideato per i paesi membri (soggetti a situazioni, risorse e contesti di istruzione totalmente differenti) e peraltro non esente da profili critici. Per tale via – dimostra Addey – non solo si vanno colonizzando, appiattendole, le ricchezze e le specificità nazionali e territoriali, ma si punta ad egemonizzare e condizionare deliberatamente e pesantemente lo sviluppo delle politiche nazionali e internazionali dell'istruzione e dell'educazione.

Pur soggetta alle stesse pressioni (e spesso conseguenti compressioni e compromissioni del diritto all'educazione e all'istruzione), la normativa italiana garantisce ancora una certa autonomia da questo punto di vista: le scuole possono sviluppare curricula verticali legati al territorio e ai propri studenti ed elaborare Rapporti di Auto Valutazione e Piani di Miglioramento personalizzati a livello di singola istituzione. Al contempo, però, si assiste al progressivo perfezionamento di un sistema di valutazione nazionale che sta costruendo, con riferimento a framework di indagini internazionali, delle tendenze che sono opposte all'autonomia. Christopher H. Tienken si occupa di valutazione da una prospettiva lucidamente critica e interna agli Stati Uniti d'America, in cui il *new public management*, cioè la tendenza ad utilizzare in termini di accountability i processi valutativi per giustificare l'allocazione di risorse (dal nostro punto di vista è soprattutto questo), ha raggiunto, possiamo dire, livelli di guardia. Nel suo contributo, Tienken passa in rassegna critica le principali tappe e teorie riguardanti l'utilizzo degli esiti delle prove degli studenti nei test standardizzati come indicatori per la valutazione dell'insegnamento e per l'implementazione di politiche di premialità rivolte ai docenti. Alla luce degli studi intorno ai deludenti risultati di tali politiche sul miglioramento dell'insegnamento e dell'apprendimento, il monito che arriva dall'altra sponda dell'Atlantico ci restituisce il rischio di una

restrizione *de facto* dei curricula e della didattica ai contenuti e alla metodologia dei test standardizzati proposti alle scuole dalle agenzie di valutazione (negli USA principalmente private). Al fine di trarre dall'esperienza statunitense elementi da indirizzare al miglioramento dei processi di valutazione e premialità degli insegnanti previsti ne "La Buona Scuola", Tienken propone di basarsi su teorie, come quella della motivazione a due stadi Herzberg (1966; Herzberg et al., 1959) e dell'autodeterminazione di Ryan e Deci (2000), che chiariscono come le persone siano professionalmente motivate più da elementi che conferiscono una qualità percepibile al lavoro svolto, quali innovazione, competenza, cooperazione, creatività, autonomia professionali, che dai soli bonus economici erogati in una riprovevole (psico)logica comportamentista di stimolo-risposta.

Sulla stessa linea, Silje Gløppen e Guri Skedsmo dell'Università di Oslo ricostruiscono approfonditamente il dibattito intorno a strumenti e pratiche di valutazione dell'insegnamento emergenti nelle scuole norvegesi e di come istanze di controllo ed elementi di sviluppo vi siano variamente intrecciati. Se talvolta gli input normativi, come l'utilizzo dell'unico strumento proposto a livello nazionale e relativo ad una valutazione dell'insegnamento da parte di tutti gli studenti delle scuole superiori, possono essere ricondotti a prospettive di dialogo con e di centratura sugli studenti, tuttavia il sistema-paese sta sperimentando in anni recenti un disarticolato e pericoloso scivolamento dalla "fiducia nella professione" verso la "fiducia nei risultati" (Uljens et al., 2013), testimoniata da un uso crescente di tecniche di valutazione degli insegnanti basate sui risultati degli studenti o sull'ispezione dell'insegnamento da parte dei superiori e da una maggiore standardizzazione delle procedure di lavoro, della regolamentazione esterna e delle misure di responsabilità individuale. Il contributo delle due studiosi norvegesi racconta un Paese che sta vivendo "sulla pelle" del proprio sistema educativo la tensione tra insegnanti, preoccupati riguardo alla conservazione di una propria autonomia e della fiducia di cui godono, e decisori politici e rappresentanti dei sistemi di valutazione esterna, con le loro aspettative sul miglioramento della qualità e i loro strumenti di controllo. Il modo in cui questa tensione si va manifestando sembra dipendere da come elementi di rendicontazione e di responsabilità, di controllo e di sviluppo, di premialità e di motivazione vengono individuati, dosati, legittimati, proposti e percepiti tra i professionisti dell'istruzione. Si delinea quindi il quadro di una sfida aperta: sebbene gli insegnanti norvegesi, anche per tramite delle rappresentanze sindacali, mantengano una posizione relativamente forte nell'elaborazione delle politiche educative, la valutazione dell'insegnamento viene tuttavia sempre più assorbita nella sfera di politiche e pratiche di educazione globalizzata, che

rischiano di produrre serie ripercussioni sulla motivazione dei docenti, sulle pratiche di insegnamento e, in definitiva, sull'apprendimento degli studenti.

Tornando in Italia, il contributo di Germana Patti e Giuseppe M. Tomaselli si contestualizza all'interno di una tipologia particolare di istituzione scolastica che, per la novità della sua introduzione e per la particolarità dell'utenza cui è destinata, si propone come modello formativo e organizzativo dai tratti particolarmente innovativi. I Centri Provinciali per l'Istruzione degli Adulti (CPIA), infatti, sono individuati quali soggetti pubblici di riferimento per la costruzione delle reti territoriali per l'apprendimento permanente, coordinando e realizzando azioni di accoglienza, orientamento e istruzione rivolte alla popolazione adulta (e soprattutto ai gruppi svantaggiati come ad esempio i migranti) e caratterizzate da un alto grado di flessibilità nella formulazione dell'offerta didattica e da uno stretto raccordo con i principali attori locali pubblici e privati. Il contributo, oltre ad approfondire l'assetto organizzativo-didattico dei CPIA sulla scorta della normativa di recente istituzione, si sofferma su un caso di studio relativo ad un progetto di didattica per competenze realizzato con la finalità di sperimentare ipotesi e strutture proposte per una loro valutazione. L'attività, difatti, si inserisce appieno nel quadro di ampia autonomia che, proprio in funzione della loro mission, il legislatore ha previsto per queste scuole, sia in termini di ricerca didattica, attraverso l'istituzione delle reti regionali di CPIA organizzate in Centri Regionali di Ricerca, Sperimentazione e Sviluppo (CRRS), sia relativamente alla valutazione, ad esempio con la costruzione e sperimentazione di un RAV specifico, che è attualmente in via di definizione proprio con un lavoro coordinato a livello nazionale tra i CRRS regionali e le università, tra cui quella di Catania.

Proprio nell'ambito universitario, nonostante sia stato soggetto negli ultimi decenni a forti pressioni in direzione di una valutazione della ricerca e della didattica principalmente basata su misurazione e ranking finalizzati a sempre più inique allocazioni differenziali delle risorse, permangono spazi di manovra per costruire forme di valutazione più sensibile ai contesti, più centrata sull'ascolto degli studenti e in definitiva più utile agli studenti stessi e a chi di loro deve farsi carico, come testimonia il contributo di Giuseppe C. Pillera. Il suo lavoro, presentato a conclusione del volume, è frutto di un'indagine svolta presso il Corso di laurea in Scienze pedagogiche e progettazione educativa dell'Università di Catania, mediante somministrazione di un questionario semi-strutturato agli studenti che hanno svolto il tirocinio formativo curriculare previsto presso aziende, enti pubblici e soggetti no profit esterni. Lo strumento – messo a punto proprio insieme agli studenti del dipartimento in contesti di didattica laboratoriale extracurricolare – mira a riconoscere, fin dalla sua costruzione, la soggettività dei processi di

valutazione, focalizzandosi sulla loro natura situata e sul ruolo dei soggetti valutati, in una prospettiva che guarda alla valutazione come strumento di apprezzamento del valore e di ricerca della qualità, piuttosto che come mera accountability. Lo studio presentato, parte di un ampio e duraturo impegno di ricerca valutativa del Corso di laurea sul suo tirocinio curricolare, si focalizza in chiave comparativa su quella particolare categoria di studenti non tradizionali rappresentata dai cosiddetti fuoricorso, dando voce a quelle persone che, irregolari negli studi per i più svariati motivi spesso prescindenti dalle volontà o capacità individuali, il sistema di valutazione delle università – per com'è oggi costruito – tende a considerare come anomalie, note dissonanti, elementi da ricollocare al più presto sui nastri di percorsi formativi cui, specie da ottiche nazionali e sovranazionali, a volte si guarda come a catene di montaggio, giudicabili su parametri di efficienza e produttività. Studenti non tradizionali, non frequentanti, fuoricorso, con bisogni educativi speciali, allora, appaiono come ferite aperte nel corpo *sanificato* di percorsi formativi che si vogliono prestabiliti e omogeneizzanti, ferite forse così brucianti, soggettività così *scomode* proprio perché portano a evidenza le contraddizioni di cui i nostri sistemi educativi, sociali, economici sono latori.

Ciononostante, nelle pieghe della normativa, delle circolari ministeriali, dei regolamenti di ateneo, a volte neanche troppo nascosti, emergono impulsi verso l'utilizzo di una valutazione formativa e autentica di pratiche, apprendimenti, contesti educativi. La narrazione per anni dominante nel campo della valutazione, in ambito educativo come in quello più generale della programmazione sociale, ha indotto ad associare la rendicontazione con la responsabilità, il che posiziona coloro che sostengono le *ragioni* della responsabilità e della complessità contro l'aridità e la pretesa oggettività della rendicontazione quali sostenitori di una prospettiva irresponsabile (Biesta, 2010).

Al contrario, una valutazione autenticamente formativa (Wiggins, 1993) e sensibile (Guba e Lincoln, 1987; Patton, 1997), centrata sui contesti e sui bisogni dei suoi utilizzatori, esige altrettanto se non maggiore impegno e responsabilità, ammette la pluralità dei punti di osservazione e delle conclusioni, accetta la soggettività e persino l'incertezza, configurandosi come strumento *formativo* proprio perché costitutivo dell'azione educativa stessa, non separato/separabile da essa; è, in poche parole, una valutazione più complessa, anzi caratterizzata da quella «ipercomplessità propria dei problemi umanamente impegnativi che implicano *anche* e principalmente l'analisi valutativa circa le motivazioni di decisioni umane che intervengano [...]. Non è possibile cioè compiere valutazioni serie che non coinvolgano largamente conoscenze ed esperienze relative all'andamento naturale dei processi sui quali si preveda di poter influire» (Visalberghi, 1987). Pratiche ad essa ispirate possono rappresentare vie maestre per ridare autonomia nella

responsabilità (e responsabilità nell'autonomia) alle istituzioni educative e agli insegnanti, restituendo centralità all'unicità della relazione educativa ed esaltando e valorizzando le specificità dei contesti formativi, rispetto ai quali il fine di qualsivoglia azione valutativa dovrebbe essere interpretato come “mezzo procedurale” e “mezzo materiale” di sviluppo.

Una domanda niente affatto retorica è se noi, docenti-ricercatori scolastici e universitari, sapremo dominare tale “ipercomplessità”, saggiando criticamente l'innovazione in campo educativo e valutativo e assumendoci la responsabilità di scegliere, di volta in volta, cosa cambiare e quando restare o resistere.

Bibliografia

- Asquini G., a cura di (2018), *La Ricerca-Formazione. Temi, esperienze, prospettive*, FrancoAngeli, Milano.
- Biesta G.J.J. (2010) “Why ‘What Works’ Still Won’t Work: From Evidence-Based Education to Value-Based Education”, *Studies in Philosophy and Education*, 29: 491-503, doi: 10.1007/s11217-010-9191-x.
- Guba E.G. e Lincoln Y.S. (1987), *Fourth Generation Evaluation*, in Palumbo D.J., ed., *The Politics of Program Evaluation*, Sage, Thousand Oaks, CA.
- Herzberg F. (1966), *Work and the Nature of Man*, World Publishing, Cleveland.
- Herzberg F., Mausner B. and Snyderman B.B. (1959), *The motivation to work* (2nd ed.), John Wiley & Sons, New York.
- Patton M.Q. (1997), *Utilization-focused Evaluation*, Sage, Thousand Oaks, CA.
- Ryan R.M. and Deci E.L. (2000), “Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being”, *American Psychologist*, 55: 68-78.
- Sellar S. and Lingard B. (2013), “The OECD and the expansion of PISA: New global modes of governance in education”, *British Educational Research Journal*, 40, 6: 917-936.
- Uljens M., Møller J., Ärlestig H. and Frederiksen L.F. (2013), *The Professionalisation of Nordic School Leadership*, in Moos L., ed., *Transnational Influences on Values and Practices in Nordic Educational Leadership. Is there a Nordic Model?*, Springer, Dordrecht.
- Visalberghi A. (1955), *Misurazione e valutazione nel processo educativo*, Edizioni di Comunità, Milano.
- Visalberghi A. (1987), “Educare alla complessità del reale”, *Scuola e Città*, 1: 1-7.
- Wiggins G. (1993), *Assessing student performance: Exploring the purpose and limits of testing*, Jossey-Bass, San Francisco, CA.

Dentro i processi della didattica. Osservare e valutare per la professionalità dell'insegnante della scuola dell'infanzia

di *Ira Vannini*

1. Introduzione

Nel volume “Misurazione e valutazione”, tutta l’attenzione di Aldo Visalberghi (1955) è posta sulla complessa problematica della valutazione degli apprendimenti che, come afferma l’Autore, «è cosa seria», evidenziandone fin dall’inizio le molteplici relazioni con i diversi piani delle tematiche dell’insegnamento.

Appare a questo proposito interessante aprire lo sguardo su quanto sembra non esserci nel volume, ma che invece – osservando attentamente – è particolarmente presente: Visalberghi incentra infatti tutta la sua riflessione sulle questioni valutative senza mai disgiungerle (esplicitamente o implicitamente) da tre temi fondamentali, tra loro strettamente interconnessi:

- il tema dell’azione didattica (e dell’individualizzazione in particolare);
- il discorso sulla funzione sociale della scuola, osservata costantemente come il mezzo fondamentale per promuovere qualità ed equità degli apprendimenti degli studenti;
- l’esigenza di contribuire alla professionalità della funzione docente, nel senso di sostenere prioritariamente negli insegnanti un atteggiamento profondamente e intrinsecamente scientifico, di chi è abituato a lavorare con il pensiero riflessivo del dubbio e del ragionamento argomentativo, attraverso interrogativi, ipotesi, verifiche sul campo e, di nuovo, riflessività.

Il discorso sul misurare e il valutare di Visalberghi si appoggia prima di tutto su un abito mentale di pensiero scientifico (profondamente e naturalmente deweyano) ed è solo in questo senso che può essere affrontato il tema della valutazione degli apprendimenti scolastici, così come – volgendo lo sguardo al campo suo corrispettivo e più ampio – della valutazione della didattica. Diviene a tal proposito interessante leggere quest’opera visalberghiana cercando possibili parallelismi tra *la valutazione degli alunni e la valutazione di sistema* (Corda Costa e Visalberghi, 1994), analizzando come la distinzione e il rapporto tra il misurare e il valutare costituiscano gli elementi

portanti per le valutazioni scolastiche in generale, affinché siano serie, aderenti ad un pensiero scientifico, che non banalizza, bensì richiama a un pensiero analitico e critico, antidogmatico, di trasparenza e responsabilità partecipativa di ciascun insegnante.

Ma vediamo le ragioni.

2. La didattica come oggetto di misurazioni e valutazioni

Per parlare di valutazione degli apprendimenti, Visalberghi afferma che occorre considerare modificabili i comportamenti e gli apprendimenti degli allievi; allo stesso modo, e ancor più, per parlare di valutazione della didattica è necessario considerare modificabili i comportamenti degli insegnanti, che questi apprendimenti promuovono.

In un saggio di alcuni anni or sono, Vertecchi (1999) scrive che l'apprendimento, da Comenio in poi, richiede mediazione didattica realizzata attraverso un artificium, «un'attività consapevolmente organizzata per conseguire un intento» (p. 8); in questo senso, si stabilisce un rapporto mezzi-fini che, lungi dal poter essere considerato in modo semplificato o esaustivo rispetto al profitto degli studenti, però ci interroga sulla necessità di analizzare con estrema attenzione sia gli apprendimenti, sia le azioni di insegnamento. Se dunque sono importanti le procedure e gli strumenti che danno conto di quanto lo studente ha appreso, altrettanto lo saranno quelli capaci di fare luce sui comportamenti della mediazione didattica degli insegnanti. Scrive ancora Vertecchi (1999): «se si opta per una interpretazione forte della “mediazione didattica” [...] aumenta il numero di variabili cui collegare i risultati, ma aumenta anche la possibilità di orientare i processi al conseguimento di determinati intenti. Infatti, se è difficile (almeno in tempi brevi), e in alcuni casi impossibile, modificare le caratteristiche personali dei soggetti coinvolti nell'attività didattica, gli spazi di libertà si dilatano se si considerano modificabili i comportamenti degli insegnanti. Al limite, tale modifica può consistere nell'attribuire la funzione di insegnare a un soggetto collettivo invece di considerarla una sorta di prerogativa posseduta a titolo personale» (p. 10).

La mediazione didattica dunque è da considerarsi come una variabile fondamentale da studiare, da conoscere nel minimo dettaglio, al fine di poterla orientare, sostenere, accompagnare; così da incidere sempre più sui risultati di apprendimento e la crescita intellettuale di tutti i bambini delle nostre scuole. Le pratiche di mediazione didattica degli insegnanti divengono pertanto un oggetto di studio di enorme fascino per la ricerca educativa, capace di conoscere, descrivere, verificare, ma anche di valutare e orientare verso un fine (di valore) queste stesse pratiche. Quest'ultimo aspetto – della ricerca

valutativa appunto sulla didattica – costituisce un punto fondamentale, in quanto pone in rilievo nuovamente il tema della valutazione quale mezzo imprescindibile per analizzare i fenomeni dell'insegnamento proponendone un confronto con elementi valoriali e indicatori di qualità. Nel caso della mediazione didattica, tale elemento valoriale di riferimento dovrebbe poter corrispondere a pratiche capaci di modificarsi e differenziarsi in funzione delle variabili di ingresso degli allievi, al fine di una «progressiva costruzione di un profilo culturale che possa permanere e sostanziare il soggetto (dargli gli attrezzi per...) nel seguito della vita» (Vertecchi, 1999, pp. 7-8).

«Omnia omnibus», afferma ancora Vertecchi, cioè dare tutto a tutti. Si ripropone qui dunque il fine di valore della didattica, oggi ancor più attuale ed urgente: «dare ogni cosa a ciascuno» e, si aggiunge, «a livello elevato».

2.1. L'interesse delle pratiche didattiche nella scuola dell'infanzia

In questo discorso sull'importanza delle misurazioni e valutazioni nell'ambito della didattica, è di rilievo individuare la scuola dell'infanzia come un contesto scolastico particolarmente interessante al fine di esplorare i processi didattici. La scuola dell'infanzia intesa dunque come una vera e propria protagonista, insieme alle sue e ai suoi insegnanti, nella ricerca valutativa sulla qualità delle pratiche di insegnamento, allo scopo di innescare processi di modificazione e miglioramento.

Per chi scrive, la scuola dell'infanzia costituisce un ordine di scuola privilegiato per realizzare percorsi di ricerca valutativa rivolti ai temi della mediazione didattica; si tratta infatti innanzitutto di un contesto dove i processi di insegnamento-apprendimento sono “liberi” dalle pressioni normative della votazione attribuita al profitto conseguito dagli alunni. Tale ambito di scuola, inoltre, ha ampiamente evidenziato nei decenni una forte apertura all'innovazione e alla sperimentazione didattica, spesso di esempio e fermento per gli altri ordini di scuola. Occorre inoltre rimarcare che, per la sua struttura organizzativa aperta e dinamica, i contesti cosiddetti “tre-sei” costituiscono in Italia il luogo dove più rilevante è lo scambio collegiale fra insegnanti, il confronto progettuale quotidiano sulle questioni connesse alle scelte, appunto, educative e didattiche da realizzare in sezione.

Per tutte queste ragioni, si tenterà di utilizzare una recente ricerca valutativa sull'osservazione delle pratiche didattiche nella scuola dell'infanzia allo scopo di mettere in evidenza l'importanza dei concetti del misurare e del valutare nell'ambito dell'insegnamento e, ancor meglio, dello sviluppo della professionalità dell'insegnante. Si tratta in qualche misura di una scusa per intrecciare i fili del discorso di Visalberghi con risultati di ricerca attuali sui

quali sarebbe davvero molto bello poter avere ancora il suo parere e il suo consiglio lungimirante.

La ricerca in questione è relativa all'uso di uno strumento per l'osservazione e la valutazione formativa delle prassi didattiche delle maestre e dei maestri dei contesti tre-sei anni denominato PraDISI (Prassi Didattiche dell'Insegnante della Scuola dell'Infanzia italiana). Nato da una ricerca condotta da Rossella D'Ugo e Ira Vannini (2015) nell'ambito di una collaborazione tra il Dipartimento di Scienze dell'Educazione "G. M. Bertin" dell'Alma Mater Studiorum Università di Bologna e il Coordinamento Pedagogico dello stesso Comune bolognese, la Scala PraDISI si concentra sulle prassi dell'insegnante che, messe in atto all'interno delle varie situazioni didattiche quotidiane, favoriscono e sostengono lo sviluppo delle competenze del bambino e della bambina: quali saperi e strategie mette in campo l'insegnante in classe/sezione per gestire i processi di insegnamento-apprendimento e promuovere in modo equo una buona padronanza di abilità fondamentali per tutti gli allievi della scuola dell'infanzia italiana? (D'Ugo, Vannini, 2015).

La Scala si propone come uno strumento di osservazione sistematica delle pratiche dell'insegnante validato sul territorio nazionale; essa tuttavia risponde anche ad un secondo obiettivo, quello di essere un mezzo per la valutazione formativa della professionalità didattica dell'insegnante, capace quindi non solo di rilevare dati, descrivere e analizzare, ma anche di esprimere giudizi di valore in relazione a specifici ed espliciti indicatori di qualità. Questo secondo scopo, prettamente valutativo, è orientato ad un'ottica "formativa", di un uso cioè della valutazione in funzione del miglioramento e della crescita professionali dei soggetti coinvolti. In questo senso, il PraDISI viene posto all'attenzione del dibattito scientifico e dell'analisi critica degli insegnanti come un'ipotesi di lavoro ancora aperta: uno strumento la cui utilità nel sostenere e orientare la professionalità docente dovrà essere via via messa alla prova. Ma vediamo con ordine le diverse questioni.

2.2. Misurare e valutare le prassi didattiche

Innanzitutto il PraDISI mira a introdurre nelle scuole dell'infanzia procedure sistematiche di osservazione dell'insegnante all'interno della sezione, laddove egli sia impegnato nella gestione dei processi di insegnamento-apprendimento (sia quindi durante le routine della giornata educativa, sia durante le attività più strettamente curricolari).

In Italia, occorre rimarcare che l'osservazione dell'insegnante in classe è stata fino a pochi anni fa una pratica pressoché sconosciuta, nel senso di

estranea al modo di fare scuola. Curioso, a questo proposito, che anche in merito al testing, il Visalberghi del 1955 (p. 7) si ritrovasse a rilevare la stessa situazione di «pratica sconosciuta» e, in relazione ad una sua introduzione nelle scuole, lo studioso ne descrivesse i possibili rischi:

- di fraintendimento per una cultura anglo-sassone molto lontana;
- di reazione respingente dovuta ad una impressione di tecnicismo (che assale soprattutto «chi non abbia maturato per altra via un vivo interesse all'argomento») (Visalberghi, 1955).

Appunto, e soprattutto, per i rischi di un potenziale impatto distorto sulle percezioni degli insegnanti, è molto importante mantenere il PraDISI (e altri strumenti di osservazione e valutazione che sono stati e si vanno ancora costruendo in molteplici percorsi di ricerca nel nostro Paese) come una ipotesi di lavoro aperta, così da controllarne progressivamente – con un atteggiamento critico e problematico – le sue possibili ricadute all'interno dei contesti scolastici, e discuterle così con gli stessi insegnanti:

- da un lato mantenendo aperto e vivace il dibattito;
- dall'altro lato, provando ad evitare e a resistere alle spinte che spesso portano le tecniche e gli strumenti ad essere utilizzati su un piano di realtà inferiore e banalissima (Visalberghi, 1955, p. 9), grettamente utilitaristica (con fini troppo spesso non espliciti o non condivisi con gli insegnanti, che dovrebbero invece essere protagonisti di un uso formativo di queste tecniche).

Ma in quale modo uno strumento di osservazione e valutazione formativa può essere ritenuto coerente e utile per analizzare il discorso visalberghiano sulla misurazione e la valutazione?

Lo sfondo metodologico che fa da cornice a strumenti quali il PraDISI si appoggia sull'approccio della *formative educational evaluation* che, da Scriven (1967, 1981) in poi, si caratterizza come quell'azione valutativa che agisce durante lo svolgimento dei processi educativi e didattici, con l'obiettivo di individuarne punti di forza e di debolezza e, conseguentemente, apportare miglioramenti. È una valutazione che regola i processi nel loro svolgersi e ne sostiene e orienta il miglioramento; proprio per questa sua capacità di agire dentro i processi, ha una forte connotazione analitica, di attenzione ai dettagli e ha dunque la possibilità di cogliere bisogni specifici e individuare soluzioni mirate. Si tratta insomma di una potente funzione valutativa in vista della presa di decisioni opportune e “ricostruttive” del percorso; generalmente essa comporta anche scelte metodologiche di tipo sistemico, di coinvolgimento attivo dei soggetti “valutati”, attente alle molteplici variabili contestuali. Gli attori che partecipano ai processi valutativi di tipo formativo sono coinvolti essi stessi in percorsi di analisi, riflessione e riprogettazione

delle loro azioni, nell'ottica anche dell'*empowerment evaluation*, con il risultato di una progressiva crescita e sviluppo della propria professionalità (Fetterman, 1996, 2007).

Come sappiamo, questa prospettiva in Italia ha preso avvio con gli studi del gruppo di ricerca di Egle Becchi (Bondioli e Ferrari, 2004), che ritrova i suoi riferimenti teorici in Beeby (1977), in Wolf (1987) e nelle definizioni del Joint Committee (1981). Rispetto a tali riferimenti, è interessante rileggere la definizione che Egle Becchi e Vega Scalera attribuiscono al “valutare” in senso formativo: un «accertamento intersoggettivo di più dimensioni formative e organizzative di un contesto educativo e indicazione della loro misurabile distanza da espliciti livelli considerabili ottimali da un gruppo di riferimento, in vista di un'incidenza concreta sull'esperienza educativa» (Bondioli e Ferrari, 2004, p. 30).

Come possiamo vedere, si tratta di una definizione che connette strettamente due dimensioni fondamentali: l'esigenza di rigore nel momento della misurazione – escludendo dal processo di rilevazione del dato «qualsiasi giudizio impressionistico» – e la collegialità, negoziale e democratica, nel momento della interpretazione del dato. In effetti, lo stesso PraDISI, che sposa questo tipo di approccio, mira a enfatizzare una metodologia valutativa che si sviluppa su due momenti distinti: il misurare e il valutare. Due azioni e, allo stesso tempo, due modalità di pensiero che si pongono una accanto all'altra in una logica di sequenzialità o, meglio, di contiguità ciclica:

- dapprima si raccolgono dati, con l'abito mentale della misura, con l'attenzione alla validità e all'affidabilità del dato; con lo scopo di delimitare un qualcosa di tangibile, dicibile, condivisibile... sul quale valga poi la pena fare un'analisi critica, discuterne, costruire;
- in seguito si valuta, e in questa operazione ci si immerge nel contesto, si chiamano a raccolta i suoi protagonisti, si promuovono tra essi transazioni e riflessività critica. Sulla base del dato (anche “smontandolo”) è possibile costruire qualcosa di nuovo, perché si parte da un punto in comune e si va verso un'ideale di qualità condiviso fra più soggetti, diretti interessati del contesto valutato.

Ma come approfondire questi concetti attraverso il pensiero di Aldo Visalberghi? Nel suo discorso sulla valutazione degli apprendimenti, Visalberghi afferma appunto che occorre considerare il misurare e il valutare. Proviamo a farlo anche noi in merito alla valutazione della didattica.

3. Osservare-Misurare

Il misurare è un'operazione di conteggio e confronto, ma è anche abito di equilibrio e discrezione: un abito mentale che rifiuta il giudizio affrettato, un abito di pensiero riflessivo.

Nel testing si devono misurare “operazioni mentali”; nell'osservazione si misurano azioni, prassi: si tratta di uno strumento più diretto, in una scala di oggettività forse l'osservazione è più oggettiva del testing, tuttavia oggi altrettanto lontana dalla pratica scolastica quanto lo era il testing negli anni '50.

L'osservazione è insomma lontana dalle abitudini diffuse di chi pratica l'insegnamento; ma l'idea di misura che essa porta con sé può avere un interesse particolare per gli insegnanti (e comunque per qualsiasi soggetto che abbia in mente un'idea di crescita professionale intesa come “mettersi in gioco”): quello della possibilità di analizzare nel dettaglio, senza giudizio, le proprie prassi quotidiane nella relazione di insegnamento-apprendimento con le bambine e i bambini.

Tutto ciò è inoltre ancor più degno di interesse oggi, immersi in contesti sociali dove sempre più spesso il giudizio affrettato, “di pancia”, il pregiudizio e gli stereotipi la fanno da padrone. Analizzare nel dettaglio vuol dire contrastare questa postura mentale dominante, soprattutto all'interno di luoghi educativi quali sono le scuole. Esso significa ricercare una obiettività possibile, che nulla toglie al fatto che nell'analisi sono comunque fusi «atti di vera e propria valutazione» (Visalberghi, 1955, p. 16); tuttavia, se lo scopo dell'obiettività è poter disporre di materiale utile in vista di «giudizi impegnativi, motivati, ricchi di conseguenze pratiche importanti» (Visalberghi, 1955, p. 16), allora la ricerca dell'obiettività analitica forse vale la pena.

Se è vero che l'atto valutativo, interpretativo, è “decisivo” nelle azioni educative e didattiche, esso potrà però essere considerato importante in ragione del suo fondarsi su un precedente atto misurativo, di sospensione del giudizio per dedicarsi ad una seria ed umile raccolta di dati. Quel che è importante del dato, afferma Visalberghi (1955), è che sia valido in relazione a quanto si vuole saggiare: «Se ci ha da essere valutazione, allora la misurazione porta serietà», o quantomeno – per riprendere in analogia un pensiero di Mario Gattullo (1984) sul rapporto tra sperimentazione e innovazione, la misurazione «dà la vista» alla valutazione (p. 67).

Ancora, sostiene Visalberghi (1955, p. 53): «Non si abbia paura delle ambizioni del testing, si pretenda solo che esso proceda con metodi e mentalità scientifica: non esiste migliore garanzia a che non perda mai il senso dei limiti, il senso della “misura”, senza bisogno che gli sian poste barriere artificiali dall'esterno, in nome di non so quale saccente supervisione che pretenda di conoscere prima di sperimentare».

Le misurazioni divengono pertanto le basi per «veri e propri giudizi, che sono veri e propri programmi d'azione», esse dunque acquisiscono senso all'interno di «ipotesi interpretative orientate a migliorare la situazione» (Visalberghi, 1955, p. 96).

4. Valutare

Il valutare chiama in causa il pensiero interpretativo, critico; un pensiero che ritroviamo ampiamente all'interno delle procedure delineate dal PraDISI quando, a partire dal dato rilevato, si inizia una fase di confronto intersoggettivo tra gli insegnanti osservati, occasione di riflessione transazionale che pone il dato in relazione a criteri di qualità definiti: gli insegnanti giudicano la distanza tra un punteggio rilevato e gli indicatori di qualità, riflettono sulle motivazioni contestuali delle evidenze raccolte, discutono – in una prospettiva progettuale – su azioni da intraprendere per il miglioramento. Si tratta insomma di usare un pensiero di giudizio critico in modo costruttivo, o ricostruttivo.

Tornando a Visalberghi (1955), in merito ai giudizi egli riflette su quelli che formula l'insegnante e afferma: «[...] i giudizi che un insegnante formula durante l'anno sul conto dei singoli allievi sono cosa seria soltanto se comportano l'indicazione dell'opera educativa, sempre in qualche misura individualizzata, di cui ogni allievo ha bisogno per il suo ulteriore progresso» (p. 96).

Anche qui dunque il giudizio acquista significato solo se diviene una possibilità di soluzione per quella particolare situazione, il giudizio ha nuovamente una finalità ri-costruttiva nel suo aiutare il singolo individuo ad affrontare e superare le sue difficoltà. Si evidenzia dunque come le valutazioni trovino senso se sono utili per il soggetto “valutato”. Questa affermazione la possiamo confermare sia nel caso delle valutazioni degli apprendimenti degli allievi, sia nelle valutazioni che si concentrano sulla didattica: la valutazione deve poter essere utile per l'insegnante che ne è protagonista. Deve cioè risultare degna di interesse per gli insegnanti, “proficua” in termini di possibilità di crescita e sviluppo professionale.

L'atto valutativo assume dunque in quest'ottica il suo pieno significato di azione “di valore”, nel senso intrinseco di quello che davvero conta in educazione, di ciò che porta alla crescita, al cambiamento in termini di costruzione di situazioni intellettuali e di vita sempre migliori. Visalberghi su questo fa un discorso chiarissimo: la valutazione non può essere al servizio di motivazioni estrinseche all'apprendimento, essa vale in sé stessa come strumento per l'apprendimento. Egli parla di una educazione e di una valutazione «[...] liberata dagli aspetti competitivi, di emulazione malintesa, che sono un

portato ineliminabile dell'uso dei voti, delle classifiche, delle graduatorie, e di quant'altre tecniche estrinseche di confronto [che] immiseriscono le ambizioni e tarpano le ali ai veri interessi» (Visalberghi, 1955, p. 20).

Con gli stessi significati si ritiene fondamentale pensare alla valutazione della didattica, una valutazione che può e deve dunque essere formativa, nel senso di aiutare a ri-costruire la professionalità dell'insegnante e a migliorare nel tempo la didattica, nell'ottica dell'apprendimento per tutti e di tutti.

Questa affermazione che propone la possibilità di una valutazione formativa va oggi vista come un auspicio, ma anche come un interrogativo di ricerca, da analizzare in una prospettiva storica, diacronica, nell'ottica di una scuola che – come variabile indipendente – mira a contribuire alla costruzione di valori sociali di collaborazione e di solidarietà (prima fra gli insegnanti e poi con gli allievi). La valutazione diventa, in questo senso, un programma di lavoro, un impegno, un impegno partecipato. E occorre sottolineare che è proprio il momento misurativo – che sostanzia la valutazione – che rappresenta l'opportunità e lo stimolo ad un impegno “collaborativo e sociale” tra gli insegnanti («oggi così isolati e scarsamente comunicativi rispetto a quella che è la sostanza del loro insegnamento», Visalberghi, 1955, p. 93). Solo una valutazione capace di appoggiarsi su pratiche misurative rigorose porta a situazioni di confronto e condivisione di un pensiero critico collegiale e collettivo: gli insegnanti del PraDISI, per esempio, si impegnano nell'apprendere le caratteristiche metodologiche e procedurali dello strumento, ne analizzano indicatori e descrittori, imparano a focalizzare lo sguardo sulle situazioni da osservare e, finalmente, vanno sul campo e raccolgono dati. Si tratta di un lavoro che nasce dal confronto (su termini, significati, procedure, azioni...) e confluisce nel confronto, fino ad arrivare ad una interpretazione dei dati da condividere, negoziare. Ben differente è la pratica di una valutazione che poggia sul “niente”, su sensazioni impressionistiche, vissuti emotivi non detti, parole ambigue... dove ciascun insegnante rischia quasi sempre di ritrovarsi solo, anche nel confronto con gli altri.

Il rapporto tra misurazione e valutazione richiama dunque la relazione tra oggettività (o ricerca seria di una intersoggettività dicibile) e pensiero critico. L'oggettività da accettare è quella «relativa e critica, ... mai assoluta e dogmatica» (Visalberghi, 1955, p. 84), una oggettività che si fa strumento di un pensiero scientifico e sperimentale, che fa del dubbio e della curiosità per la verifica di ipotesi la propria traiettoria abituale, il proprio abito mentale. Una oggettività che è sempre storicamente e culturalmente determinata, da negoziare “dentro” la cultura delle scuole, in senso intersoggettivo, rendendo attivi e protagonisti coloro che stanno dentro il processo valutativo e che, grazie a buoni processi di misurazione, possono impegnarsi in interpretazioni e riflessioni ri-costruttive e progettuali.

5. Per non concludere... in merito a un'opera ancora attuale

La ricerca valutativa sulla didattica, e in particolare sulle pratiche didattiche nella scuola dell'infanzia, trova dunque ampi e ricchi spunti di riferimento e preziosi orientamenti nell'opera visalberghiana "Misurazione e valutazione".

Come abbiamo potuto vedere, relativamente all'esperienza del PraDISI, l'insegnante osservatore fa esperienza della pratica misurativa, si impraticisce nell'uso di strumenti e procedure di raccolta dei dati. Si immerge in un vero e proprio percorso di Ricerca-Formazione (Asquini, 2018), accompagnato ad acquisire l'abito mentale del ricercatore all'interno degli stessi contesti didattici in cui agisce.

L'invito, a questo proposito, è che gli insegnanti facciano davvero esperienza diretta di questi percorsi, che entrino nel merito degli aspetti procedurali dell'osservazione sul campo, che smontino gli strumenti "oggettivi" e li criticino, affinché si eviti che questi stessi strumenti possano, ad un certo punto «piovere dall'alto come antipatico modo di controllo non desiderato, e mezzo di invito al conformismo nella direzione prediletta da chi detiene il potere politico e facilmente perciò può assumere i tecnici al suo servizio» (Asquini, 2018, p. 92).

Allo stesso modo e sempre nell'ottica della Ricerca-Formazione (così come definita dal Centro CRESPI <https://centri.unibo.it/crespi/it>), gli insegnanti che decidono di coinvolgersi in un percorso di misurazione e valutazione della didattica, fanno esperienza del momento specificamente valutativo e si impegnano in un pensiero riflesso condiviso: il valutare diviene così quel programma di azione, quel giudizio per il cambiamento: un impegno a pensare, a pensare consapevolmente, a cambiare e innovare.

Si tratta insomma di un abito che porta con sé una spinta di pensiero quasi a dire rivoluzionario (necessario ieri, ma tanto più oggi, nel troppo frequente abbassamento delle aspirazioni e delle attese "di valore" di chi fa scuola!).

Rileggendo le parole di Visalberghi (1955): distinguere «e mettere accanto» misurazione e valutazione è, in sé, «un impegno a battersi per una radicale revisione delle strutture e delle tendenze attuali, e non giustifica certo il comodo adattarsi nelle prassi tradizionali» (p. 97).

E ancora: «Il testing avrà la funzione che gli insegnanti sapranno conferirgli» (Visalberghi, 1966, p. 98).

La domanda resta la stessa anche per la valutazione della didattica: quanto e a quali condizioni potrà contribuire a rinnovare metodi e clima sociale dentro le scuole e tra gli insegnanti?

Bibliografia

- Asquini G., a cura di (2018), *La Ricerca-Formazione. Temi, esperienze, prospettive*, FrancoAngeli, Milano.
- Beeby C.E. (1977), "The Meaning of Evaluation", *Current Issues in Education*, 4: 68-78.
- Bondioli A. e Ferrari M., a cura di (2004), *Verso un modello di valutazione formativa. Ragioni, strumenti e percorsi*, Edizioni Junior, Bergamo.
- Corda Costa M. e Visalberghi A. (1994), "Dalla valutazione degli alunni alla valutazione di sistema", *Scuola e Città*, 7: 273-277.
- D'Ugo R. e Vannini I. (2015), *PraDISI. La valutazione formativa delle prassi didattiche dell'insegnante di scuola dell'infanzia: osservare per riprogettare*, FrancoAngeli, Milano.
- Fetterman D.M., Kaftarian S.J. and Wandersman A., eds. (1996). *Empowerment Evaluation: Knowledge and Tools for Self-Assessment & Accountability*, SAGE Publishing, Thousand Oaks, CA.
- Fetterman D.M., Kaftarian S.J. and Wandersman A. (2007), "Empowerment Evaluation. Yesterday, Today, and Tomorrow", *American Journal of Evaluation*, 28, 2: 179-198.
- Gattullo M. (1984), *Sperimentare e decidere*, in Becchi E. e Vertecchi B., a cura di, *Manuale critico della sperimentazione e della ricerca educativa*, FrancoAngeli, Milano.
- Joint Committee on Standards for Educational Evaluation (1981), *Standards for evaluations of Educational Programs, Projects and Materials*, SAGE Publishing, Thousand Oaks, CA.
- Scriven M. (1967), *The methodology of evaluation*, Rand McNally, Chicago.
- Scriven M. (1981), *Evaluation Thesaurus*, Edgepress, Inverness, CA.
- Vertecchi B. (1999), 'Riscolarizzare' la società. *Lezione tenuta il 29 aprile 1999 a Covilha, presso l'Universidade de Beira Interior*, testo disponibile al sito: https://www.academia.edu/18769475/_Riscolarizzare_la_societ%C3%A0.
- Visalberghi A. (1955), *Misurazione e valutazione nel processo educativo*, Edizioni di Comunità, Milano.
- Wolf R.M. (1987), "The nature of educational evaluation", *International Journal of Educational Research*, 11, 1: 7-20.

La *global data infrastructure* di PISA e la politica OCSE nascosta in piena vista: *data infrastructures alternative* per valutare la qualità dell'educazione

di *Camilla Addey*

1. Introduzione

Perché le prove nazionali assomigliano sempre più a PISA? Perché il quarto Obiettivo per lo Sviluppo Sostenibile (fornire un'educazione di qualità, equa e inclusiva, e opportunità di apprendimento per tutti) è così ben allineato con la *global education governance* dell'OCSE? Nonostante gli innumerevoli obiettivi avanzati e gli indicatori disponibili per misurare gli sviluppi, c'è una sorprendente simmetria tra PISA e il quarto Obiettivo per lo Sviluppo Sostenibile (conosciuto come SDG4). Questo articolo ripercorre la storia recente della *global education governance* dell'OCSE e analizza in maniera critica il crescente dominio dell'OCSE nel determinare come misurare e concepire l'educazione. Si tratta di una analisi essenziale, in quanto il modo in cui si misura l'educazione determina lo scopo dell'educazione e di conseguenza le politiche e la pianificazione dell'educazione.

Per analizzare il rapporto tra le prove nazionali, la metrica di PISA e il quarto Obiettivo per lo Sviluppo Sostenibile, questo articolo si basa su dati raccolti nell'ambito del progetto di ricerca "PISA for Development for Policy" (Addey, 2016, 2017, 2018, 2019). Il progetto analizza come l'OCSE realizza progetti – in questo caso PISA for Development – nel campo dell'educazione in contesti in cui l'OCSE non aveva lavorato precedentemente e come questi strumenti vengano utilizzati dai governi in loco. PISA for Development (PISA-D) è un'estensione di PISA, il test sviluppato dall'OCSE nel 1997 per rilevare le prestazioni in matematica, lettura e scienze dei quindicenni in diversi paesi. Essendo stato sviluppato per contesti educativi come quelli dei paesi membri dell'OCSE, la metrica PISA è poco adatta a contesti in cui le abilità rilevate tendono a raggrupparsi nei livelli più bassi (OECD, 2013b; BLOEM, 2013) e dove ci sono alte percentuali di studenti che non frequentano la scuola. PISA-D permette di rilevare le prestazioni degli studenti sotto il livello minimo di PISA, di includere i quindicenni non più frequentanti la scuola

e di analizzare un contesto educativo potenzialmente diverso da quello dei paesi OCSE (per esempio, includendo diversi indicatori per misurare lo status socioeconomico degli studenti).

Ai fini del progetto di ricerca “PISA for Development for Policy”, PISA-D rappresenta un caso particolarmente interessante in quanto è il primo progetto nel campo dell’educazione che l’OCSE sviluppa per paesi a basso e medio reddito, molti dei quali non hanno mai stabilito rapporti formali con l’OCSE. PISA-D ha dunque rappresentato il primo passo dell’OCSE nello sviluppo del campo dell’educazione a un livello internazionale non limitato agli stati membri (Addey, 2017; Auld et al., 2018), ma anche la prima opportunità per i paesi a basso e medio reddito di stabilire un rapporto formale con l’OCSE (Addey, 2019).

I dati per il progetto di ricerca “PISA for Development for Policy” sono stati raccolti nel 2015 e 2016 durante lo sviluppo di PISA-D, tramite interviste con funzionari dell’OCSE, le aziende contrattate dall’OCSE per sviluppare PISA-D, e alti funzionari del governo Ecuadoriano e Paraguayano (entrambi paesi partecipanti a PISA-D). I dati raccolti si basano anche su osservazioni etnografiche di eventi e riunioni PISA-D e di analisi di documenti di lavoro e pubblicazioni dell’OCSE.

Analizzando questi dati attraverso la lente di analisi proposta da Sellar e Lingard (2013) sulla *global governance* infrastrutturale ed epistemologica (esposta in seguito), questo articolo avanza tre tesi. Primo, che l’OCSE sta lavorando sull’allineamento delle prove nazionali con PISA. Secondo, che lo sviluppo di PISA-D è stato limitato ai fini di rendere i risultati paragonabili ai risultati di PISA ma soprattutto al fine di rendere le prove nazionali nei paesi PISA-D allineabili alla metrica PISA. Terzo, che l’OCSE ha sviluppato PISA-D in maniera da rafforzare la strategia dell’Organizzazione e posizionarsi in maniera centrale rispetto al quarto Obiettivo per lo Sviluppo Sostenibile – in questo modo la strategia di sviluppo dell’organizzazione stessa è supportata dall’intera comunità mondiale raccolta e impegnata nell’avanzare SDG4. I dati raccolti e la sorprendente simmetria tra PISA e il quarto Obiettivo per lo Sviluppo Sostenibile logicamente portano a chiedersi se l’OCSE non abbia proposto PISA e PISA-D come metro per misurare gli sviluppi verso il raggiungimento mondiale di SDG4 oppure se SDG4 non sia stato formulato come risultato delle pressioni dell’OCSE. Prima di analizzare queste tesi, è necessario ripercorrere la storia della recente espansione della *global education governance* dell’OCSE, analizzata attraverso il concetto di *global governance* infrastrutturale ed epistemologica.

2. L'OCSE e la *global education governance* infrastrutturale ed epistemologica

Nelle ultime due decadi, dall'inizio del programma PISA, gli studiosi hanno analizzato il crescente ruolo dell'OCSE nelle politiche educative a livello nazionale ma anche il suo ruolo nella *global education governance*. In particolare, la ricerca si è centrata sulla sempre maggiore fiducia nei dati e uso dei rankings internazionali (Gorur 2011; Martens, 2007; Grek, 2009) e il modo in cui la costruzione di indicatori e metriche per analizzare i sistemi educativi e l'apprendimento incanalano un determinato modo di concepire e agire sull'educazione (Lingard et al., 2013; Sellar e Lingard, 2013).

Gli studiosi Sellar e Lingard, ispirati dalla ricerca di Woodward sull'OCSE (2009), hanno avanzato una tesi per cui l'OCSE esercita una *global education governance* che descrivono come infrastrutturale ed epistemologica. Sellar e Lingard descrivono come *governance* infrastrutturale il processo per cui l'OCSE crea e attiva una infrastruttura globale che permette di quantificare l'apprendimento, così creando un unico spazio globale di misurazione. Questa infrastruttura globale decontestualizza, standardizza e trasforma complesse e diverse forme di apprendimento in fatti quantificati che previamente esistevano in modo né misurabile né paragonabile. L'infrastruttura a cui si fa riferimento trasforma l'apprendimento, complesso e multiplo nei diversi contesti, in fatti quantificabili e paragonabili attraverso prove e pratiche standardizzate. Questa infrastruttura è descritta come una *global data infrastructure* che produce dati oggettivi sull'apprendimento che possono essere interpretati nello stesso modo, a prescindere dal contesto di apprendimento e dell'utilizzo delle capacità.

Numerosi ricercatori che si occupano della sociologia della quantificazione (Espeland e Stevens, 1998, 2008; Gorur, 2015, 2017) e della emergente area di ricerca *infrastructure studies* (Star, 1999; Larkin, 2013; Sellar, 2017, 2014; Lingard et al. 2013; Easterling, 2014), mettono in luce come le infrastrutture nascondono in piena luce le decisioni politiche che sono prese e impresse nella creazione delle infrastrutture (Star, 1999; Lingard et al., 2013). In particolar modo, Easterling (2014) ci ricorda che non dobbiamo guardare a quello che una infrastruttura ci dice, ma a quello che fa. Le decisioni su come sviluppare PISA, la *global data infrastructure* dell'OCSE, sono prese da gruppi di esperti selezionati e chiusi in riunioni non accessibili a persone esterne. La *global data infrastructure* viene presentata come oggettiva ma incanalata una certa concezione dell'educazione, come per esempio lo scopo dell'educazione. L'adozione e l'applicazione di questa *global data infrastructure*, nello specifico PISA, allinea molteplici e diversi attori nel loro modo di concepire l'educazione. Per esempio, nell'epoca di PISA, la

qualità dell'educazione è concepita sempre più come il raggiungimento di un determinato livello di prestazioni in scienza, matematica e lettura. La scrittura, per esempio, che non è facilmente quantificabile né standardizzabile attraverso diversi paesi e culture, è esclusa dagli obiettivi cui viene conferito valore nelle politiche educative. Questo modellare e allineare il modo in cui viene intesa l'educazione come risultato della *global data infrastructure* è quel che Sellar e Lingard (2013) chiamano *governance* epistemologica. La *global education governance* infrastrutturale ed epistemologica mette in luce perché è importante capire come l'OCSE cerca di inserire la metrica PISA nelle prove nazionali (portando a una conseguente mancanza di metriche e metodi di valutazione dell'educazione alternativi) e come utilizza gli Obiettivi per lo Sviluppo Sostenibile per avanzare potenzialmente a livello globale il modo in cui l'educazione è concettualizzata in PISA, la sua *global data infrastructure*.

3. Sinergie tra le prove nazionali e la metrica PISA

Nonostante sia difficile osservare un impatto diretto delle International Large-Scale Assessments (ILSAs) sulle politiche educative, molti studiosi e organizzazioni internazionali cercano di distillare questa relazione (Fischman et al., 2018; Breakspeare, 2012). Quel che si osserva, e che è più semplice distillare come impatto diretto come PISA, è l'aumento delle prove nazionali allineate a PISA. Uno studio di Baird et al. (2016) mostra una serie di esempi. La Cina ha introdotto una prova nazionale, così simile da essere stata ribattezzata "PISA Cinese"; la Norvegia ha introdotto una valutazione della comprensione della lettura che è interamente basata sul modello PISA; mentre le nuove prove canadesi, come quelle Invalsi, sono strutturalmente molto simili a PISA. Verger et al. (2020) hanno mostrato che nella riforma delle prove catalane nel 2011, si rileva un allineamento al modello PISA. Nel caso dell'Ecuador, le prove sono state allineate con PISA ancora prima di adottare PISA, dimostrando così che PISA ha un impatto sulle prove ancor prima o anche senza che il programma sia stato adottato. Questi sono solo alcuni degli esempi che sono stati studiati.

Nonostante l'OCSE abbia dichiarato che è sempre stato l'obiettivo di PISA quello di arricchire le prove nazionali, solo recentemente questo aspetto di PISA è stato preso più seriamente. Nel 2011, l'OCSE ha commissionato uno studio su come collegare le prove nazionali con PISA (OECD, 2011). Lo studio, che mirava a sviluppare una metodologia e delle linee guida da mettere a disposizione dei governi, ha stabilito che esistono metodi che permettono di collegare le prove nazionali con PISA che danno risultati validi (OECD, 2011, p. 26). Poco dopo, nel 2015, l'OCSE ha pubblicato la

strategia a lungo termine di PISA (OECD, 2013a) in cui uno dei sei obiettivi era quello di rafforzare il rapporto tra le prove nazionali e PISA. L'importanza di questo obiettivo è confermata nei documenti che descrivono come le imprese private contattate dall'OCSE per sviluppare PISA-D sono tenute ad analizzare come le prove nazionali nei paesi partecipanti a PISA-D possano essere allineate con PISA-D, con lo scopo di proporre approcci e opzioni per massimizzare le sinergie tra le prove nazionali e PISA-D. Alle imprese contattate è stato inoltre specificato che nello sviluppo di PISA-D, non potevano divergere in alcuna parte dal design metodologico di PISA (OECD, 2014b, p. 10 e p. 19). Questo ovviamente ha creato tensioni visto che lo scopo originale di PISA-D era quello di essere adattato alle realtà dei paesi a basso e medio reddito. Ma come descritto da funzionari Ecuadoriani, nessun paese partecipante a PISA-D era disposto a sacrificare la possibilità di far parte di PISA, tramite l'adozione di PISA-D. Nonostante i rappresentanti dei paesi fossero interessati ad avere dei dati che riflettessero il loro contesto, tutti i rappresentanti e lo staff dell'OCSE hanno concordato che era più importante mantenere la metrica PISA intatta e permettere ai paesi partecipanti in PISA-D di paragonarsi con PISA.

È inoltre interessante osservare come, nonostante l'OCSE abbia specificato che l'attività di *capacity building* organizzata mirasse ad agevolare la partecipazione in PISA-D, queste capacità dovessero essere rafforzate a livello istituzionale. In altre parole, le capacità a cui punta l'OCSE potrebbero essere mirate a diffondere le metodologie di analisi delle competenze nelle istituzioni responsabili per riformare le prove o svilupparne dove non esistono, piuttosto che ad agevolare l'implementazione di PISA-D.

Questo approccio va di pari passo con il discorso promosso dall'OCSE riguardo alle attuali prove nazionali: i funzionari OCSE le hanno quasi tutte bocciate come metodologicamente invalide e povere in confronto agli standard di PISA. Non è difficile immaginare come questo discorso sia quello che viene promosso negli istituti o negli enti che sviluppano prove nazionali e a cui viene insegnata la metodologia PISA. Ritornando alla *global data infrastructure*, presentata come metodologicamente la più solida e valida ma anche come oggettiva e apolitica, vediamo come nel collegare o allineare le prove nazionali a PISA, un certo modo di concepire e valutare l'educazione venga incanalato e diffuso a livello prima globale, e poi a livello nazionale. L'educazione valutata attraverso la *global data infrastructure* produce numeri che in realtà sono modellati da decisioni politiche nascoste dentro l'infrastruttura di PISA. È più facile ascoltare cosa dice la *global data infrastructure* piuttosto che capire cosa fa.

4. PISA e l'OCSE allineati con Obiettivi per lo Sviluppo Sostenibile. O viceversa?

Nel 2012, in seguito a una riunione dei paesi membri dell'OCSE, l'Organizzazione ha dichiarato che, visti i cambiamenti economici a livello globale, si sarebbe trasformata in una Organizzazione globale, includendo così anche i paesi a basso e medio reddito. Nella "OECD Strategy on Development" (2012), l'Organizzazione prevede di rivedere i suoi strumenti per informare le politiche in modo da adattarele a nuovi contesti dove non ha previamente assistito i governi. Nel campo dell'educazione, la Strategia stabilisce che PISA giocherà un ruolo di primaria importanza. È in questo periodo, nel 2013, che l'OCSE riconosce ufficialmente che i risultati di PISA sono poco informativi nei paesi a basso e medio reddito (OECD, 2013b; Bloem, 2013). Ma è anche questo il periodo in cui si riuniscono gruppi di lavoro di esperti e si creano sistemi per dare voce a milioni di persone che vogliono partecipare nella definizione degli obiettivi del post-2015. L'OCSE si piazza in prima linea per promuovere l'idea di un obiettivo non più centrato sull'accesso all'educazione, ma sull'acquisizione di abilità. I funzionari dell'OCSE dichiarano che alla domanda su come misurare l'acquisizione delle abilità, la risposta era ovvia per tutti, in particolare visto la decisione di sviluppare PISA-D. I documenti che l'OCSE pubblica nel periodo precedente al 2015, l'anno della dichiarazione degli Obiettivi per lo Sviluppo Sostenibile, promuovono PISA come metrica universale per lo sviluppo di un obiettivo centrato sull'acquisizione di abilità (OECD, 2014). Nel 2015, il quarto Obiettivo, che prevede una educazione primaria e secondaria di qualità, equa e gratuita che comporti apprendimenti efficaci e significativi, è chiaramente allineata con PISA. Nella formulazione dell'Obiettivo, PISA non è definita come la metrica, ma nasce spontaneamente un quesito: perché l'attuale quarto Obiettivo per lo Sviluppo Sostenibile è così ben allineato con la metrica di PISA? Nonostante le infinite proposte avanzate e gli indicatori disponibili per misurarne gli sviluppi, c'è una sorprendente simmetria tra PISA e il quarto Obiettivo per lo Sviluppo Sostenibile.

Si può dire che PISA-D è stato creato per estendere la metrica di PISA e analizzare potenziali progressi futuri nel contesto degli Obiettivi per lo Sviluppo Sostenibile. Oppure il quarto Obiettivo è così formulato come risultato della pressione dell'OCSE? Questo resta un punto interrogativo che meriterebbe ulteriore ricerca. L'unica risposta attualmente disponibile è la soddisfazione espressa dai funzionari OCSE nell'osservare quanto siano allineati gli Obiettivi per lo Sviluppo Sostenibile con la Strategia per lo Sviluppo dell'OCSE e gli obiettivi OCSE (OECD, 2012). Un ulteriore argomento in favore della pressione esercitata dall'OCSE sulla formulazione del quarto

Obiettivo è la storia emersa sulla partecipazione della Cambogia in PISA-D, in cui l'attuale ministro dell'educazione è stato forzato alla partecipazione sotto falsa promessa che PISA è la metrica con cui viene misurato il progresso verso il quarto Obiettivo (Auld et al., 2018) – nonostante questo sia lontano dall'essere vero.

5. Conclusioni

L'articolo ha dimostrato che, attraverso un'analisi del progetto PISA for Development, è possibile osservare come l'OCSE stia lavorando all'allineamento tra PISA e le prove nazionali – sia nei paesi a basso e medio reddito che in quelli ad alto reddito. L'articolo ha inoltre dimostrato che, nonostante rimanga un grande quesito sulla formulazione del quarto Obiettivo per lo Sviluppo Sostenibile, quest'ultimo è espresso in maniera da rafforzare il modo in cui l'OCSE punta a riformarsi e a sostenere le sue attività di sviluppo internazionale nel campo dell'educazione. Analizzare questi cambi nelle attività dell'OCSE alla luce della tesi avanzata da Sellar e Lingard, mette in luce come l'OCSE stia cercando di radicare la propria *global data infrastructure* nelle *data infrastructures* dei singoli paesi, portando così non solo alla scomparsa di metodi alternativi di valutazione dell'educazione, ma anche a una colonizzazione – nascosta nell'infrastruttura – del modo di concepire l'educazione e agire su di essa.

L'autorità acquisita dall'OCSE nel campo dell'educazione, insieme all'impatto della comunità mondiale raccolta e impegnata nell'avanzare SDG4, ma anche alla natura nascosta di quello che l'infrastruttura fa, rendono la metrica PISA egemonica. Davanti a un tale dominio e una tale unità nel modo di valutare e concepire l'educazione, come possono emergere proposte alternative? Questo articolo invita a prendere in considerazione la politica delle organizzazioni internazionali, la formulazione delle agende globali per lo sviluppo e delle infrastrutture, piuttosto che accettare le attuali *global data infrastructures* come i migliori strumenti disponibili per valutare la qualità dell'educazione. Le *infrastructures* alternative vanno allineate agli scopi dell'educazione, piuttosto che alle politiche nascoste nelle infrastrutture dominanti.

Bibliografia

- Addey C. (2016), "PISA for Development and the sacrifice of policy-relevant data", *Educação & Sociedade*, 37, 136: 685-706, doi: 10.1590/es0101-73302016166001.
- Addey C. (2017), "Golden relics & historical standards: how the OECD is expanding global education governance through PISA for Development", *Critical Studies in Education*, 58, 3: 311-325, doi: 10.1080/17508487.2017.1352006.
- Addey C. (2018), *Researching inside the international testing machine: PISA parties, midnight emails & red shoes*, in Maddox B., ed., *International Large-Scale Assessments in Education. Insider Research Perspectives*, Bloomsbury, London, testo disponibile al sito: <https://bloomsburycp3.codemanttra.com/viewer/5bca026e713c090001a4068f>.
- Addey C. (2019), "The appeal of PISA for Development in Ecuador and Paraguay: Theorising and applying the global ritual of belonging", *Compare: A Journal of Comparative and International Education*, doi: 10.1080/03057925.2019.1623653.
- Auld E., Rappleye J. and Morris P. (2018), "PISA for Development: how the OECD and World Bank shaped education governance post-2015", *Comparative Education*, 55, 2: 197-219, doi: 10.1080/03050068.2018.1538635.
- Baird J.-A. et al. (2016), "On the supranational spell of PISA in policy". *Educational Researcher* 58, 2: 121-138.
- Bloem S. (2013), *PISA in Low and Middle Income Countries*, OECD Publishing, Paris.
- Breakspear S. (2012), "The policy impact of PISA: an exploration of the normative effects of international benchmarking in school system performance", *OECD Education Working Papers*, 71, OECD Publishing, Paris.
- Easterling K. (2014), *Extrastatecraft: The Power of Infrastructure Space*, Verso Books, New York.
- Espeland W.N. and Stevens M. (1998), "Commensuration as a Social Process", *Annual Review of Sociology*, 24: 313-343.
- Espeland W.N. and Stevens M. (2008), "A sociology of quantification", *Archives of European Sociology*, 49, 3: 401-436.
- Fischman G., Marcetti Topper A., Silova I., Goebel J. and Holloway J.L. (2018), "Examining the influence of international large-scale assessments on national education policies", *Journal of Education Policy*, doi: 10.1080/02680939.2018.1460493.
- Gorur R. (2011), "ANT on the PISA Trail: Following the Statistical Pursuit of Certainty", *Educational Philosophy and Theory*, 43, 1:76–93.
- Gorur R. (2012), "The invisible infrastructure of standards", *Critical Studies in Education*, 54, 2: 132-142, doi: 10.1080/17508487.2012.736871.
- Gorur R. (2015), *Assembling a Sociology of Numbers*, in Hamilton M., Maddox B. and Addey C., eds., *Literacy as Numbers*, Cambridge University Press, Cambridge.
- Gorur R. (2017), "Towards productive critique of large-scale comparisons in education", *Critical Studies in Education*, 58, 3:341-355.
- Grek S. (2009), "Governing by Numbers: The PISA 'Effect' in Europe", *Journal of Education Policy*, 24, 1: 23-37.

- Larkin B. (2013), "The Politics and Poetics of Infrastructure", *Annual Review of Anthropology*, 42: 327-343.
- Leigh Star S. (1999), "The ethnography of infrastructure", *American Behavioural Scientist*, 43: 377-391.
- Lingard B., Martino W. and Rezaei-Rashti G. (2013), "Testing regimes, accountabilities and education policy: commensurate global and national developments", *Journal of Education Policy*, 28, 5: 539-556, doi: 10.1080/02680939.2013.820042.
- Martens K. (2007), *How to become an influential actor - the 'comparative turn' in* Martens K., Rusconi A. and Leuze K., eds., *New Arenas in Education Governance. The Impact of International Organizations and Markets on Educational Policy Making*, Palgrave Macmillan, New York.
- OECD (2011), *Using PISA to Internationally Benchmark State Performance Standards*, OECD, Paris.
- OECD (2012), *OECD Strategy on Development*, OECD, Paris.
- OECD (2013a), *Beyond PISA 2015: A Longer-Term Strategy for PISA*, OECD, Paris.
- OECD (2013b), *The OECD's contribution on education to the post-2015 framework: PISA for Development. OECD and post-2015 reflection series*, OECD, Paris.
- OECD (2014a), *Report on the Implementation of the OECD Strategy on Development*, OECD, Paris.
- OECD (2014b), *PISA-D Call for Tender*, OECD, Paris.
- Sellar S. (2014), "Data infrastructure: a review of expanding accountability systems and large-scale assessments in education", *Discourse*, 36, 5: 765-777.
- Sellar S. (2017), "Making network markets in education: the development of data infrastructure in Australian schooling", *Globalisation, Societies and Education*, 15, 3: 341-351, doi: 10.1080/14767724.2017.1330137.
- Sellar S. and Lingard B. (2013), "The OECD and the expansion of PISA: New global modes of governance in education", *British Educational Research Journal*, 40, 6: 917-936.
- Verges A., Prieto M., Pagès M. and Villamor P. (2020), "Common standards, different stakes: A comparative and multi-scalar analysis of accountability reforms in the Spanish education context", *European Educational Research Journal*, 19, 2: 142-164.
- Williamson B. (2017), *Big Data in Education: The digital future of learning, policy and practice*, SAGE Publications, Thousand Oaks, CA.
- Woodward R. (2009), *The Organisation for Economic Cooperation and Development (OECD)*, Routledge, Abingdon.

Valutazione dei docenti e valutazione del merito in Italia: lezioni dall'America

di *Christopher H. Tienken*

1. Contesti della valutazione del merito in Italia

Il 9 luglio 2015, il Parlamento italiano ha approvato la Legge n. 107 di riforma del sistema educativo, conosciuta come “La Buona Scuola”. La legge ha normato una serie di riforme del sistema dell’educazione pubblica italiana della scuola dell’obbligo nelle aree del curriculum, dell’istruzione, dell’infrastruttura scolastica, della valutazione docente, del reclutamento dei docenti e della valutazione dell’istituzione scolastica. Le disposizioni contenute ne “La Buona Scuola” influenzano i macro- e micro-curricula, come gli insegnanti distribuiscono tali curricula e come essi vengono valutati.

Lo scopo di questo capitolo è quello di dare informazioni per lo sviluppo della policy futura del programma di valutazione del merito degli insegnanti promulgato da “La Buona Scuola” con lo scopo di fornire indicazioni per l’adozione dei criteri utilizzabili per determinare il merito. Fornisco una rassegna dei risultati dei grandi programmi di premialità del merito emanati nelle scuole pubbliche negli Stati Uniti dalla creazione del Fondo federale per l’incentivazione dei docenti (Federal Teacher Incentive Fund, TIF) avviato dall’amministrazione del Presidente George W. Bush nel 2006. L’amministrazione Bush ha stabilito il finanziamento iniziale di 99 milioni di dollari e i fondi sono aumentati durante la presidenza di Obama a oltre 435 milioni di dollari (Marsh, 2012). Dal 2006, varie città e stati in America hanno sperimentato proposte di premialità del merito ed esiste una raccolta di studi empirici di alta qualità che ne documentano i risultati. Uso i quadri teorici delle aree del comportamentismo, della motivazione e dell’autodeterminazione per criticare i programmi di premialità del merito e fornire raccomandazioni generali di policy per valutare l’insegnamento che sono supportate dalla teoria e dalla ricerca.

2. Valutazione del merito in Italia

“La Buona Scuola” ha emanato procedure di valutazione che hanno un impatto diretto sul modo in cui gli insegnanti vengono valutati e pagati. In particolare, il concetto di premialità del merito è introdotto nel paragrafo 126 della Legge come parte del processo di valutazione annuale degli insegnanti. I funzionari del Ministero dell’Istruzione (allora MIUR) non hanno fornito specifiche linee guida d’azione del docente o criteri per l’assegnazione del bonus del merito.

Alla fine dell’anno scolastico 2017/2018, gli unici criteri disponibili per guidare il nuovo sistema di valutazione del merito degli insegnanti stabilivano che fosse il capo d’istituto a indicare gli insegnanti con le migliori prestazioni insieme a un comitato scolastico di valutazione che deve essere composto: (a) dal dirigente scolastico, (b) da un valutatore esterno come un ispettore governativo o dirigente di un’altra scuola, (c) da tre insegnanti della scuola e (d) da 1-2 rappresentanti dei genitori a seconda del tipo di scuola. Sono richiesti due rappresentanti dei genitori per gli istituti che accolgono gli alunni nei gradi di scuola primaria e secondaria di primo grado e un genitore nelle scuole secondarie superiori.

A ogni comitato di valutazione della scuola veniva concessa la possibilità di individuare i criteri per l’esame dei suoi insegnanti e per l’assegnazione della premialità di merito. In pratica, il dirigente della scuola era colui che sviluppava i criteri poi utilizzati dai comitati scolastici. L’autonomia dei comitati scolastici per determinare i criteri può essere considerata sia positivamente che negativamente. L’autonomia è positiva perché il personale scolastico ha l’opportunità di sviluppare criteri personalizzati per le scuole specifiche e per gli studenti ai quali insegnano. Tuttavia, l’assenza di criteri ufficiali aumenta la possibilità che il merito venga assegnato a seguito dei “favori” fatti al dirigente o venga dato a coloro che il dirigente ammira, indipendentemente da quanto l’insegnante si impegni per gli studenti e la scuola. Sulla base di conversazioni informali con insegnanti in diverse province in tutta Italia, sembra che la seconda ipotesi sia vera nella maggior parte dei casi.

Si prevede che il Ministero dell’Istruzione italiana pubblichi linee guida più precise in merito ai criteri di valutazione degli insegnanti che le commissioni scolastiche dovranno seguire per determinare quali insegnanti riceveranno un premio di merito. Il Ministero sembrava incline a includere i risultati dei test degli studenti come parte dei criteri per la valutazione del merito, indipendentemente dalle prove che suggeriscono che non si dovrebbe usare tale sistema. Gli educatori e i responsabili politici italiani hanno senz’altro bisogno di riprogettare la valutazione degli insegnanti basandola sulle

evidenze, invece che sull'ideologia, come punto di partenza per la loro revisione del programma di premialità del merito.

3. Quadro teorico per la premialità del merito. Teoria dell'aspettativa e Teoria dei processi

Diverse teorie sono alla base della maggior parte dei programmi di premialità del merito utilizzati dal 2006 negli Stati Uniti: la Teoria dell'Aspettativa e la Teoria del Processo sono due riferimenti comuni a molti dei maggiori. La Teoria dell'Aspettativa propone che una persona prenda decisioni razionali su quanti sforzi fare nel suo lavoro basandosi su (a) l'aspettativa che lo sforzo porti a un risultato personalmente soddisfacente e su (b) quanto la persona apprezzi il risultato previsto. I sostenitori della valutazione del merito sulla base dei risultati dei test degli studenti ritengono che gli insegnanti lavoreranno maggiormente e produrranno migliori punteggi nei test standardizzati degli studenti in cambio di una ricompensa monetaria per tali punteggi più alti.

Il modello di Porter e Lawler (1968), inserito nella Teoria del Processo, è la base sulla quale la Teoria dell'Aspettativa viene resa operativa. Il modello di Porter-Lawler fornisce la prova che è meglio dare uno stimolo esterno, in questo caso la ricompensa per il merito, che spingerà le persone a lavorare di più o a mettere più impegno per raggiungere gli obiettivi predeterminati correlati all'ottenimento della ricompensa per il merito. La maggior parte degli esperimenti su larga scala per la premialità del merito per gli insegnanti negli Stati Uniti si concentra su un bonus *tantum* per gli insegnanti che producono punteggi più alti nei test standardizzati obbligatori per gli studenti. L'obiettivo dei punteggi dei test più alti è impostato da un ente esterno, non dagli insegnanti, in base agli obiettivi posti dall'ente esterno stesso. Tutti i programmi di premialità del merito su vasta scala tentati negli Stati Uniti si basano sul modello di Porter-Lawler, in cui gli obiettivi e gli stimoli sono impostati esternamente e gli insegnanti sono spinti a raggiungere gli obiettivi e ad accettare gli stimoli.

4. Comportamentismo

Negli Stati Uniti, la teoria comportamentista del condizionamento operante (Skinner, 1938, 1953) è correlata al modello di Porter-Lawler, perché gli stimoli esterni sono usati per cambiare il comportamento degli insegnanti e per concentrare i loro sforzi su un preciso obiettivo organizzativo:

aumentare i punteggi dei test standardizzati. Il condizionamento operante è dunque alla base dei programmi di valutazione del merito che si fondano su premi *tantum* per gli insegnanti. Il condizionamento operante descrive i processi utilizzati per modificare i comportamenti, basati sul rinforzo positivo e negativo: il fulcro del modello di Porter-Lawler. Attraverso il condizionamento operante le persone adattano i loro comportamenti alla condotta prevista per ricevere una ricompensa o per evitare la punizione (Skinner, 1938, 1953). Nel caso della valutazione del merito, gli insegnanti adeguano la loro azione d'insegnamento al raggiungimento degli obiettivi esterni stabiliti dai responsabili politici o dagli amministratori scolastici. Ad esempio, se il premio di valutazione del merito si basa sui punteggi dei test standardizzati degli studenti, gli insegnanti concentreranno la loro azione d'insegnamento sui contenuti che più probabilmente si trovano nel test standardizzato e trascorreranno più tempo a preparare gli studenti per il test con l'obiettivo di ricevere il premio.

Il condizionamento operante è una forma di psicologia dello stimolo-risposta, ossia una "scienza per il controllo degli altri" (Bredo, 2002, p. 25), basata in parte sulla premessa che la punizione e la mancanza di punizione motiveranno le persone a raggiungere gli obiettivi prefissati (Rescorla e Wagner, 1972). Il ricevere una ricompensa per il merito o la mancanza di tale ricompensa agiscono come stimoli di ricompensa e punizione. Il condizionamento operante è radicato nel lavoro sperimentale con gli animali e nel controllo del loro comportamento (Skinner, 1938, 1953).

La Teoria X di McGregor (1960) è un'altra teoria che spiega perché alcuni decisori politici promuovono programmi di premialità del merito e dà un'idea del motivo per cui sono state create alcune proposte di valutazione del merito: i pregiudizi di alcuni responsabili politici sul fatto che gli insegnanti siano pigri. I sistemi di premialità del merito, infatti, sono creati per motivare gli insegnanti a lavorare di più, perché c'è un'opinione di fondo che gli insegnanti siano pigri e lavorerebbero di più per più soldi.

Tuttavia, in alcuni sistemi di valutazione del merito gli insegnanti pensano che non sia possibile per loro avere l'opportunità di raggiungere il successo perché considerano il sistema come soggettivo, strutturato ingiustamente o sanno che fattori extrascolastici come il contesto familiare e il contesto comunitario influenzano i punteggi dei test standardizzati (Tienken, 2016). Quindi, la Teoria della reattanza (Brehm, 1972) è una parte di ogni sistema di valutazione di merito, come evidenziato dal modo di agire di alcuni docenti che si ritirano dalla partecipazione attiva alla vita scolastica, cercano di sovvertire il sistema del merito criticando i test standardizzati, si demoralizzano o smettono di collaborare con i loro colleghi. Gli effetti negativi della valutazione del merito riducono l'efficacia e la collaborazione

complessiva e creano un ciclo negativo di rendimento didattico inferiore e un restringimento dei curricula a ciò che è più probabile venga testato o si allinea ai criteri di valutazione del merito. Dewey (1939) ha avvertito delle possibili conseguenze dei sistemi della valutazione del merito in base ai punteggi dei test: «“Do not put your hand to the plow until the cost has been counted” are but a few of the many maxims. [...] Only the exigencies of a preconceived theory (in all probability one seriously infected by the conclusions of an uncritically accepted “subjectivistic” psychology) will ignore the concrete differences that are made in the content of “likings” and “prizings”, and of desires and interests, by evaluating them in their respective causal capacities when they are taken as means» (pp. 32-33).

Comprendere le varie teorie e ricerche sugli schemi della premialità del merito esistenti negli Stati Uniti può fornire importanti informazioni agli educatori e ai responsabili politici italiani in modo che non debbano ripetere i fallimenti del sistema degli Stati Uniti. Comprendere gli errori del passato può aiutare i responsabili politici italiani a concentrarsi sulla costruzione di un sistema che abbia maggiori probabilità di produrre risultati educativi positivi mediante strumenti più efficaci.

5. Esperienza precedente con la premialità del merito negli Stati Uniti

I responsabili delle politiche negli Stati Uniti hanno sperimentato programmi su larga scala di premialità del merito per insegnanti per oltre un decennio. La forma più comune di premialità utilizzata nei programmi di valutazione degli insegnanti negli Stati Uniti è il bonus annuale individuale, pagato agli insegnanti con l'obiettivo di aumentare i punteggi degli studenti nei test standardizzati. Il bonus annuale individuale è un pagamento una tantum che gli insegnanti possono ricevere e pertanto non aumenta in modo permanente il loro stipendio. L'insegnante deve continuare a soddisfare i criteri per la valutazione del merito ogni anno al fine di ricevere un bonus.

Ci sono stati cambiamenti nei processi di valutazione degli insegnanti in quasi tutti gli stati dall'adozione della riforma federale dell'istruzione, nota come “No Child Left Behind Act” nel 2002 (NCLB PL 107-110, 2002) e la creazione del TIF nel 2006. Le modifiche ai programmi di valutazione degli insegnanti sono state accelerate dal programma di finanziamento federale “Race to the Top” nel 2009 e 2010, che prevedeva la formazione di funzionari statali istruiti sulla valutazione dei docenti e dei dirigenti in base ai risultati dei test degli studenti, come responsabili dell'ammissione del personale docente e dirigente a concorrere per il premio in denaro fornito dal

programma di finanziamento federale. La firma della legge federale “Every Student Sings Act”, introdotta dal presidente Barack Obama nel dicembre 2015, ha dato ai funzionari della pubblica istruzione la libertà di basare le valutazioni degli insegnanti e degli amministratori scolastici sulla base dei risultati dei test standardizzati degli studenti.

La maggior parte delle 50 agenzie educative pubbliche negli Stati Uniti d’America hanno sperimentato alcune forme di premialità del merito o programmi di valutazione dei docenti basati sui risultati degli alunni nei test standardizzati già dal 2006. I risultati delle scelte di policy basate sui programmi di premialità del merito riferita a risultati dei test standardizzati degli studenti negli Stati Uniti non sono stati positivi in relazione al raggiungimento dello scopo stabilito di innalzare lo stesso livello dei risultati ottenuti dagli studenti nei test standardizzati e in alcuni casi ha avuto come risultato una riduzione del curriculum in relazione a ciò che viene principalmente testato (Au, 2006). Non mancano risultati di studi su vasti programmi della premialità del merito effettuati a New York City, Nashville, Tennessee, Round Rock, Texas, Austin, Newark, New Jersey e sul Programma di Avanzamento degli Insegnanti (Teacher Advancement Program, TAP) che è stato posto in essere in 15 stati (e.g. Fryer, 2011; Marsh, 2012; Springer, Ballou & Peng, 2008; Springer et al., 2009, 2012a, 2012b).

Programmi di valutazione del merito basati sui risultati dei test standardizzati per gli studenti sono stati anche implementati in numerose città in Colorado, Florida, Michigan, Minnesota, South Carolina e nella città di Washington D.C. In quasi tutti i casi i programmi di premialità del merito basati sui risultati dei test standardizzati degli studenti non hanno avuto effetti statistici significativi sui risultati degli studenti misurati dai test standardizzati o solo risultati trascurabili: meno dello 0,15%. Tuttavia, centinaia di milioni di dollari e un numero inestimabile di ore di lavoro dei docenti sono stati spesi per queste iniziative. Sia se la valutazione del merito fosse stata fatta su base individuale, come nei programmi a Nashville, Tennessee, o sulla base di gruppi, come a Round Rock, Texas, o sulla base degli istituti, come a New York City, i risultati non sono stati statisticamente rilevanti e molte risorse economiche votate a tale sforzo avrebbero potuto essere usate per scopi migliori.

Per esempio, New York City ha speso 75 milioni di dollari per i programmi di merito per le istituzioni scolastiche durante i tre anni di sperimentazione dal 2007 al 2010 e non si sono avuti risultati statistici rilevanti nelle valutazioni in matematica e in lingua inglese quando misurati con test standardizzati (Fryer, 2011). Intere istituzioni scolastiche rispondevano ai requisiti per la valutazione del merito se i punteggi dei test standardizzati nelle scuole miglioravano. L’idea di base era di collegare l’insegnamento nelle

istituzioni scolastiche all'obiettivo di alzare i risultati dei test standardizzati. I 75 milioni di dollari spesi per il programma avrebbero potuto essere usati per sostenere quasi 8.000 delle famiglie più povere di New York City per uscire dallo stato di povertà per almeno un anno, fornendo opportunità aggiuntive di vita per i bambini di queste famiglie.

Sono stati anche documentati effetti negativi sulle esperienze educative degli studenti quando i loro insegnanti sono parte di un programma di premialità del merito basato sui test standardizzati dei risultati degli studenti. Au (2011) ha riportato nei risultati del suo studio, basato su una meta-analisi, che un'importante reazione alla standardizzazione inclusiva della valutazione docenti basata sui punteggi dei test degli studenti era il restringimento del curriculum agli argomenti che più probabilmente venivano testati. Il curriculum delle esperienze degli studenti veniva impoverito e i docenti uniformavano il loro insegnamento ai criteri che determinavano la loro premialità. Invece di migliorare e arricchire le esperienze di educazione per gli studenti la premialità del merito sembra perciò restringere e impoverire le esperienze educative.

6. Sistemi attuali di valutazione degli insegnanti negli USA

L'educazione negli Stati Uniti è decentralizzata. Non essendo inclusa nella costituzione nazionale, l'educazione è una delle prerogative dei singoli stati e così in effetti, gli Stati Uniti possono avere 50 differenti sistemi di educazione. Tuttavia, ci sono regolamentazioni nazionali per alcuni aspetti dell'educazione pubblica, come la protezione per i bambini disabili e la regolamentazione che proibisce la discriminazione. Ci sono, però, poche regolamentazioni nazionali per la valutazione degli insegnanti e delle scuole, anche se ne esistono alcune. Perciò, i sistemi di valutazione dei docenti differiscono in qualche modo nei 50 stati. Tra gli stati, inoltre, ci sono differenze in alcuni aspetti della valutazione docente, dato che nelle singole città possono essere aggiunte regolamentazioni e scelte di policy al sistema statale. Ma in ogni stato è previsto che gli insegnanti vengano osservati dal dirigente scolastico o da altri amministratori scolastici, mentre insegnano in classe.

Gli insegnanti devono venire osservati mentre insegnano almeno da una a tre volte all'anno. Alcuni stati specificano i criteri che devono essere usati durante l'osservazione dell'insegnante, altri stati hanno un menù di criteri tra i quali le scuole del distretto possono scegliere, mentre altri Stati consentono alle singole istituzioni scolastiche di determinare i criteri per l'osservazione dell'insegnante e le valutazioni finali. Gli insegnanti sono valutati in base ai risultati ottenuti dalle loro valutazioni. Ci sono quattro comuni livelli di

valutazione: (a) insoddisfacente (*unsatisfactory*), (b) base (*basic*), (c) esperto (*proficient*), (d) distinto (*distinguished*). In alcuni stati gli insegnanti ricevono punteggi per le loro valutazioni, una valutazione numerica per ogni anno scolastico. Il punteggio numerico corrisponde a un livello di valutazione, ad esempio *proficient*.

Dal 2002, un numero crescente di stati include i risultati degli studenti nei test annuali standardizzati in matematica e nella lingua inglese come parte del punteggio di valutazione degli insegnanti. A partire dal 2015, 40 stati hanno utilizzato i risultati degli studenti nei test standardizzati come parte del sistema di punteggio di valutazione degli insegnanti e 29 stati hanno utilizzato i risultati dei test degli studenti come parte delle valutazioni per i dirigenti scolastici (Doherty e Jacobs, 2015).

7. Il profitto al di sopra delle persone

La pratica della premialità degli insegnanti in base ai risultati dei test degli studenti viene presa in prestito dal mondo degli affari for profit. La valutazione del merito è più comunemente utilizzata in lavori che valutano i dipendenti in base a un tipo di output chiaramente quantificabile. La valutazione del merito è spesso utilizzata nelle industrie e nei lavori di fabbrica che includono il lavoro a tempo, come assemblare l'elettronica di consumo o riparare le trasmissioni di automobili. Tuttavia, la valutazione del merito viene anche utilizzata nelle professioni dei servizi umani come l'assistenza infermieristica, ma il merito si basa sul portare a termine compiti facilmente quantificabili e di breve durata, come rispondere a un campanello di chiamata del paziente entro un periodo di tempo stabilito (Glassman, Glassman, Champagne e Zugelder, 2010). L'infermiere non viene giudicato in base a quanto bene il paziente risponde all'intervento medico o se il paziente segue il consiglio dell'infermiere dopo che questi ha risposto alla chiamata, mentre le proposte di merito per le scuole pubbliche negli Stati Uniti giudicano l'insegnante su quanto bene lo studente risponde all'intervento didattico.

L'aumento dei profitti aziendali è un fattore motivante per l'implementazione dei sistemi di valutazione del merito, come evidenziato dall'uso sempre più frequente di tali sistemi nel settore delle imprese a scopo di lucro rispetto al settore non profit. È mia opinione che il settore for profit abbia desideri e interessi diversi legati alla valutazione della produzione dei dipendenti rispetto al settore non profit, e soprattutto rispetto all'istruzione. Le imprese a scopo di lucro sono interessate a fare soldi. Questo è il loro obiettivo principale. L'istruzione pubblica non è un business. Gli studenti sono esseri umani, non parte di un margine di profitto. Gli insegnanti lavorano per

sviluppare il bambino nella sua interezza in termini di orientamenti socio-civici e aspetti professionali, non solo accademicamente basati sugli obiettivi strettamente definiti di un test standardizzato. Gli obiettivi dell'insegnamento e lo sviluppo del bambino sembrano in contrasto con le ipotesi di business for profit che sono alla base della premialità del merito.

Dewey (1939) ha avvertito che gli interessi e i desideri relativi alla valutazione degli insegnanti devono essere allineati ad altri fattori coinvolti nella valutazione delle prestazioni scolastiche: «The chief weakness of current theories of valuation which relate the latter to desire and interest is due to failure to make an empirical analysis of concrete desires and interests as they actually exist. When such an analysis is made, certain relevant considerations at once present themselves» (p. 29).

8. Perché gli insegnanti scelgono di insegnare

La Teoria della motivazione a due stadi di Herzberg può essere utilizzata per iniziare a comprendere alcuni potenziali effetti negativi che possono verificarsi con la valutazione del merito dell'insegnante basata sui risultati dei test standardizzati degli studenti (Herzberg, 1966; Herzberg, Mausner e Snyderman, 1959). Herzberg ha teorizzato che i fattori interni relativi al lavoro di una persona, come una maggiore autonomia rispetto alla pianificazione, all'implementazione e alla valutazione del lavoro, aiutino a motivare la persona e ad aumentare la soddisfazione sul lavoro. Incoraggiare la creatività, l'assunzione di rischi e l'innovazione da parte dei dipendenti attraverso una maggiore autonomia sul loro lavoro porta a una maggiore produttività, più creatività, più elevati livelli di qualità e soddisfazione complessiva del lavoro. Fattori come la paga non aumentano la motivazione. La retribuzione può ridurre la motivazione quando essa non è paragonabile a lavori simili, tuttavia, i dipendenti non saranno maggiormente motivati da una maggiore retribuzione, ma sono demotivati dalle retribuzioni basse. I dipendenti sono motivati dalla qualità del lavoro.

L'idea generale di Herzberg secondo cui la qualità del lavoro motiva i dipendenti è stata dimostrata nella professione insegnante negli Stati Uniti e nei paesi dell'Unione Europea. I risultati delle inchieste sul perché gli insegnanti scelgono la professione docente suggeriscono che è il valore intrinseco del lavoro d'insegnamento il fattore motivante, mentre il salario è un fattore estrinseco e non motiva gli insegnanti a lavorare più duramente (Webb et al., 2004). Sebbene la retribuzione sia necessaria e il risentimento possa aumentare quando la retribuzione è percepita come troppo bassa, una maggiore retribuzione non genera necessariamente una maggiore

soddisfazione sul lavoro o un migliore risultato d'insegnamento se misurato ai risultati dei test degli studenti.

I risultati di un'indagine negli Stati Uniti (Tienken, 2009) e altri studi pubblicati in un libro di Sandra Chistolini (2009) dell'Università di Roma Tre, si allineano con la teoria della motivazione a due fattori di Herzberg. In quegli studi, insegnanti provenienti da Italia, Stati Uniti, Polonia, Belgio, Cipro, Libia, Slovacchia e Turchia hanno risposto a un'indagine sui fattori correlati al motivo per cui hanno scelto l'insegnamento come professione e alle ragioni per cui continuano nella loro professione. I risultati in tutti i paesi studiati hanno indicato che gli insegnanti non insegnano per la ricompensa finanziaria. Le persone scelgono la professione d'insegnante perché trovano il lavoro significativo e credono di poter fare la differenza nel mondo.

Gli insegnanti hanno risposto che erano intrinsecamente motivati dal loro lavoro con gli studenti, dal valore sociale della professione e dalle opportunità di crescita personale, non da una ricompensa monetaria esterna (Chistolini, 2009). Gli insegnanti degli otto paesi dello studio hanno indicato di essere interessati alla loro professionalità e all'aiutare gli studenti a maturare in modo cognitivo ed emotivo, non a stimoli economici. I risultati di questi studi del 2009 sostengono l'ipotesi che la teoria di Herzberg è applicabile all'istruzione in vari paesi. Tuttavia, le ragioni per cui gli insegnanti scelgono la professione raramente sono prese in considerazione quando vengono elaborate e messe in atto le proposte di premialità del merito.

Allo stesso modo, Corsini (2015) ha riportato, in uno studio su 733 insegnanti in Sicilia, che il sistema di valutazione genera opposizione quando gli insegnanti si sentono minacciati dalla valutazione o può essere un fattore motivante quando gli obiettivi della valutazione sono in linea con gli scopi d'insegnamento del docente. I risultati di Corsini (2015), insieme a quelli di Chistolini (2009), affermano che la Teoria della reattanza, la Teoria dell'autodeterminazione e la Teoria della motivazione di Herzberg sono applicabili al contesto italiano e dovrebbero essere considerati in qualsiasi politica di valutazione dell'insegnamento.

Le ragioni fornite dagli insegnanti sul motivo per cui scelgono l'insegnamento si allineano maggiormente con la Teoria dell'autodeterminazione (Ryan e Deci, 2000) che con la Teoria dello stimolo-risposta o con la Teoria del processo. La Teoria dell'autodeterminazione suggerisce che le persone hanno tre principali bisogni psicologici che devono essere soddisfatti al fine di produrre una motivazione intrinseca: (a) competenza, (b) relazione e (c) autonomia. Le persone crescono professionalmente e personalmente quando questi bisogni sono soddisfatti. Nel contesto dell'istruzione, i bisogni riguardano: (a) la competenza professionale, (b) la collegialità e (c) l'autonomia professionale.

Mulder (2014) ha definito la competenza professionale come un insieme di conoscenze, abilità, disposizioni e atteggiamenti necessari per agire efficacemente secondo degli standard professionali ed etici e secondo norme e usi. La competenza professionale nell'insegnamento può includere la conoscenza e l'uso della scienza pedagogica, l'applicazione dei principi dello sviluppo infantile basati sulla ricerca e lo sviluppo di una cultura della cura sociale ed emotiva in classe. La collegialità è rappresentata da insegnanti che collaborano per migliorare la loro efficacia con gli studenti. La cooperazione può assumere la forma di collaborazione a progetti pedagogici, di condivisione delle risorse, di attività di lavoro in comunità di apprendimento professionale o altre forme di lavoro di squadra (Shah, 2016), come spesso viene praticato nella medicina moderna.

L'autonomia professionale nell'istruzione è rappresentata dal fatto che agli insegnanti è concesso di determinare e sviluppare curricula, valutazioni per gli studenti e sviluppare criteri di valutazione per sé stessi. Dewey (1924) ha equiparato l'autonomia professionale all'emancipazione degli insegnanti e ha collegato la creatività e l'autonomia degli studenti all'autonomia professionale dei docenti, avvertendo che non è possibile aspettarsi creatività negli studenti e apprendimento indipendente senza autonomia professionale per gli insegnanti: «Finally, I saw how inconsistent it was to expect this greater amount of creative, independent work from the student when the teachers were still unemancipated; when the teachers were still shackled by too many rules and prescriptions and too much of a desire for uniformity of method and subject matter» (p. 470).

Le teorie della motivazione e dell'autodeterminazione a due fattori di Herzberg forniscono supporto all'osservazione di Dewey e aiutano anche a spiegare coerentemente i risultati delle indagini degli insegnanti negli Stati Uniti e nei paesi dell'Unione europea. Le teorie suggeriscono che gli insegnanti devono essere i principali attori nel processo di valutazione e che i criteri di valutazione devono riflettere il lavoro dell'insegnamento e non tentare di quantificare risultati di lavoro.

9. Un'opportunità di miglioramento

I responsabili delle politiche negli Stati Uniti hanno ignorato la scienza sulla valutazione del merito. Milioni di dollari di denaro pubblico e innumerevoli ore di insegnamento e di apprendimento sono stati sprecati nel tentativo di costruire sistemi basati su teorie viziate e ideologie scorrette. Spero che politici e educatori italiani scelgano un percorso diverso.

Essi hanno un'opportunità storica di fare ciò che i responsabili politici degli Stati Uniti non sono riusciti a realizzare: creare un sistema di valutazione che porti a migliori esperienze di istruzione per gli studenti e una maggiore crescita professionale e autonomia per gli insegnanti. I politici italiani possono evitare gli errori del passato se presteranno attenzione alla ricerca sulla motivazione e sul perché le persone scelgono di insegnare e se coinvolgeranno attivamente docenti di scuole pubbliche e professori universitari che si interessano di educazione nel processo decisionale in modi autentici. Un risultato potrebbe essere una politica che aiuti gli studenti a diventare più creativi, innovativi, socialmente consapevoli e a consentire agli insegnanti di esercitare il loro giudizio professionale senza timore di punizioni.

10. Idee per l'implementazione

La domanda su come valutare al meglio l'insegnamento deve essere risolta dalla risposta che si darà a una domanda di proporzioni storiche: quali sono gli scopi della scuola pubblica? Docenti, pedagogisti e politici italiani devono riflettere sulla risposta a questa domanda nel contesto italiano.

Essi devono proporre un sistema di valutazione basato innanzitutto sulle funzioni della scuola pubblica in Italia, sulle teorie dell'apprendimento e dell'insegnamento, sui fattori motivanti e sulle teorie riguardo al perché le persone scelgono la professione d'insegnante. La forma della valutazione dell'insegnamento deve seguire la funzione della scuola, non dettarla. La valutazione deve servire la scuola e la professionalità anziché la valutazione della professionalità e del servizio scolastico. La questione della valutazione non può essere risolta fino a quando gli scopi della scuola pubblica in Italia non saranno chiaramente definiti.

La funzione socio-civica dell'educazione pubblica è quella che sembra trascurata dai redattori de "La Buona Scuola". L'Italia ha una tradizione di pensiero democratico e di pedagogia sperimentale in educazione, come esemplificato dall'educatore siciliano Michele Crimi e altri (Tomarchio, 2008). Gli educatori e i decisori politici italiani devono attingere alla loro ricca storia nell'area dell'educazione socio-civica per contribuire ad informare la valutazione dell'insegnamento. Dewey (1952) ha spiegato il ruolo socio-civico dell'educazione pubblica nel favorire la democrazia: «For the creation of democratic society we need an educational system where the process of moral-intellectual development is in practice as in theory a cooperative transaction of inquiry engaged in by free, independent human beings who treat ideas and the heritage of the past as means and methods for the

further enrichment of life, quantitatively and qualitatively, who use the good attained for the discovery and establishment of something better» (p. xi).

Alcuni principi comuni aggiuntivi per un sistema nazionale di valutazione dell'insegnamento in Italia potrebbero richiedere alle commissioni locali di valutazione dell'insegnamento di sviluppare criteri propri, basati su fattori sui quali gli insegnanti e il personale scolastico hanno più controllo, piuttosto che sui risultati degli studenti nei test standardizzati: (a) sulla qualità delle opportunità che insegnanti e altro personale scolastico creano, progettano e implementano affinché gli studenti possano sperimentare e partecipare attivamente alle funzioni della scuola pubblica, (b) sui tipi di opportunità che gli insegnanti creano e implementano affinché gli studenti seguano le loro passioni e interessi ed esprimano la loro creatività e innovatività, (c) su come gli insegnanti organizzano e personalizzano il curriculum in modo che gli studenti possano partecipare attivamente alla loro istruzione, (d) sui modi in cui gli insegnanti costruiscono i rapporti con i loro studenti, (e) sui tipi di opportunità che gli insegnanti e altro personale scolastico creano per gli studenti perché possano diventare studenti per tutta la vita e (f) sulle opportunità che gli insegnanti e altro personale scolastico forniscono agli studenti per esplorare gli interessi imprenditoriali nel contesto di una risoluzione dei problemi socialmente consapevole.

La personalizzazione locale può anche avvenire attraverso opportunità per le scuole di aggiungere criteri relativi agli obiettivi scolastici territoriali. Tali criteri saranno decisi totalmente dalla volontà degli insegnanti e richiederanno la maggioranza del consenso del personale scolastico. Il dirigente potrebbe anche dare degli indirizzi d'azione per il raggiungimento di un obiettivo scolastico proprio dell'istituzione e che verrà incluso nel sistema di valutazione. Invece di osservare gli insegnanti insegnare e valutarli, il sistema italiano di valutazione dell'insegnamento dovrebbe concentrarsi sui criteri di valutazione, osservando come gli studenti reagiscono ai modi in cui gli insegnanti rendono a loro volta operativi i criteri di valutazione nelle loro classi; l'indirizzo d'azione, le reazioni degli studenti ad esso e il modo in cui gli insegnanti cambiano approccio in base alle reazioni degli studenti.

Un insegnante può realizzare azioni tecniche, come dichiarare chiaramente un obiettivo di apprendimento o una miriade di altre azioni, ma tuttavia non offrire mai agli studenti opportunità di partecipare attivamente alla loro istruzione. L'insegnante potrebbe essere tecnicamente competente nella produzione di azioni didattiche appropriate, ma gli studenti potrebbero non avere mai l'opportunità di seguire i propri interessi e passioni. Agli insegnanti potrebbe essere richiesto di presentare i loro risultati ogni due o tre anni a tutta il collegio docenti e al comitato di valutazione locale. Ciò renderebbe il processo di valutazione scaglionato in modo che non tutti gli

insegnanti vengano valutati ogni anno. L'intervallo di tempo esteso consentirebbe loro, inoltre, di sviluppare un percorso lavorativo professionale. Forse il quarto, l'ottavo e il dodicesimo anno di valutazione potrebbero costituire occasioni per presentazioni più ampie e approfondite, con effetti su meriti, promozioni o altri riconoscimenti. Gli insegnanti possono ricevere feedback collegiali e feedback dai gruppi di valutazione locali insieme alle proprie riflessioni professionali, per creare una serie triangolare di dati di valutazione.

Un sistema basato sugli esempi presentati sopra riguarda più le motivazioni degli insegnanti, ma si allinea anche alle funzioni generali della scuola pubblica in una società democratica e a come gli studenti apprendono. Si allinea alla Teoria dell'autodeterminazione e alla Teoria della motivazione a due fattori di Herzberg e include gli insegnanti nello sviluppo del sistema di valutazione.

11. Pensieri di chiusura

Come affermò Dewey (1924), come possono gli studenti imparare a partecipare attivamente a una democrazia, a criticare le idee e pensare in modo creativo, complesso e innovativo se i loro insegnanti sono prigionieri di un sistema standardizzato per promuovere la conformità e governato da forze autocratiche? Qualsiasi sistema di valutazione dell'insegnamento deve essere di natura anticonformista, basato su un paradigma comune sviluppato a partire dallo scopo della scolarizzazione in quel particolare contesto culturale, dalla ricerca e dalle teorie su come gli studenti apprendono meglio e sui fattori motivanti per gli insegnanti.

Qualsiasi sistema di valutazione deve liberare gli insegnanti per creare e innovare, non costringerli. Deve fornire criteri chiari, personalizzati per la situazione specifica della scuola, in modo che essi abbiano l'opportunità di dimostrare i loro risultati in modo trasparente, democratico e centrato sullo studente.

Bibliografia

Au W. (2011), "Teaching under the new Taylorism: high-stakes testing and the standardization of the 21st century curriculum", *Journal of Curriculum Studies*, 43, 1: 25-45, doi: 10.1080/00220272.2010.521261.

- Bredo E. (2002). *The Darwinian Center to the vision of William James*, in Garrison J., Poedeschi P. and Bredo E., eds., *William James and Education*, Teachers College Press, New York.
- Brehm J.W. (1972), *Response to loss of freedom: A theory of psychological reactance*, General Learning Press, Morristown, NJ.
- Chistolini S., a cura di (2009), *Insegnanti. Identità ed etica della professione in Italia, Stati Uniti d'America, Polonia, Belgio, Cipro, Libia, Slovacchia, Turchia*. Edizioni Kappa, Roma.
- Corsini C. (2015), *Valutare scuole e docenti*, Edizione Nuova Cultura, Roma.
- Dewey J. (1924), "The Classroom Teacher", *General Science Quarterly*, 7: 463-472.
- Dewey J. (1939), *Theory of valuation*, in Neurath O., Carnap R. and Morris C., eds., *International Encyclopedia of Unified Science*, 2,4, University of Chicago Press, Chicago.
- Dewey J. (1952), *Introduction*, in Clapp E.R. (1952) *The Use of Resources in Education*, Harper & Brothers, New York.
- Doherty K. M. and Jacobs S. (2015), *State of the States 2015: Evaluating Teaching, Leading and Learning*, National Council on Teacher Quality, Washington DC, testo disponibile al sito: <https://eric.ed.gov/?id=ED581451>.
- Fryer R.G. (2011), *Teacher incentives and student achievement: Evidence from New York City public schools*, Working paper 16850, National Bureau of Economic Research, Cambridge, MA, testo disponibile al sito: <http://www.nber.org/papers/w16850>.
- Glassman M., Glassman A., Champagne P.J. and Zuegelder M.T. (2010), "Evaluating pay for performance systems: Critical issues for implementation", *Compensation & Benefits Review*, 42, 4: 231-238.
- Herzberg F. (1966), *Work and the Nature of Man*, World Publishing, Cleveland.
- Herzberg F., Mausner B. and Snyderman B.B. (1959), *The motivation to work* (2nd ed.), John Wiley & Sons, New York.
- Legge n. 107 del 13 luglio 2015, *Riforma del sistema nazionale di istruzione e formazione e delega per il riordino delle disposizioni legislative vigenti*.
- Marsh J. (2012), "The Micropolitics of implementing a school-based bonus policy: The case of New York City's compensation committees", *Educational Evaluation and Policy Analysis*, 34, 2: 164-184.
- McGregor D. (1960), *The Human Side of Enterprise*, McGrawHill, New York.
- Mulder M. (2014), *Conceptions of Professional Competence*, in Billett S., Harteis C., Gruber H., eds., *International Handbook of Research in Professional and Practice-based Learning*, Springer, Dordrecht.
- No Child Left Behind (NCLB) Act of 2001, Pub. L. No. 107-110, § 115, Stat. 1425 (January 8, 2002).
- Porter L.W. and Lawler E.E. (1968), *Managerial attitudes and performance*, Dorsey Press and Richard D. Irwin, Homewood, IL.
- Rescorla R.A. and Wagner A.R. (1972), *A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement*, in Black A.H. and Prokasy W.F., eds., *Classical Conditioning II*, Appleton-CenturyCrofts, New York.

- Ryan R.M. and Deci E.L. (2000), “Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being”, *American Psychologist*, 55: 68-78.
- Shah D. (2016), “Transformational Teamwork: The Virtues and Values of Collaboration”, *Marshall Journal of Medicine*, 2, 3, doi: 10.18590/mjm.2016.vol2.iss3.1.
- Skinner B.F. (1938), *The Behavior of organisms: An experimental analysis*, Appleton-Century, New York.
- Skinner B.F. (1953), *Science and human behavior*, Simon and Schuster, New York.
- Springer M.G., Ballou D. and Peng X. (2008), *Impact of the Teacher Advancement Program on Student Test Score Gains: Findings from an Independent Appraisal*, National Center on Performance Incentives at Vanderbilt University, Nashville, TN.
- Springer M.G., Lewis J.L., Podgursky M.J., Ehlert M.W., Taylor L.L., Lopez O.S. and Peng A. (2009), *Governor’s Educator Excellence Grant (GEEG) Program: Year three evaluation report*, National Center on Performance Incentives, Nashville, TN.
- Springer M.G., Ballou D., Hamilton L., Le V., Lockwood J.R., McCaffrey D., Pepper M. and Stecher B. (2012a), *Final Report: Experimental Evidence from the Project on Incentives in Teaching*, National Center on Performance Incentives at Vanderbilt University, Nashville, TN.
- Springer M.G., Pane J.F., Le V., McCaffrey D.F., Burns S.F., Hamilton S. and Stecher, B. (2012b), *Team Pay for Performance: Experimental Evidence From the Round Rock Pilot Project on Team Incentives*, National Center on Performance Incentives, Nashville, TN.
- Tienken C.H. (2016), *Standardized test results can be predicted, so stop using them to drive education policymaking*, in Tienken C. and Mullen C., eds., *Education policy perils: Tackling the tough issues*, Taylor Francis Routledge, Philadelphia, PA.
- Tienken C.H. (2009), *Professional growth opportunities in the USA: What are important to teachers?*, in Chistolini S., a cura di, *Insegnanti. Identità ed etica della professione in Italia, Stati Uniti d’America, Polonia, Belgio, Cipro, Libia, Slovacchia, Turchia*, Edizioni Kappa, Roma.
- Tomarchio M. (2008), *Lo sperimentalismo pedagogico in Sicilia e Michele Crimi*, Anicia, Roma.
- Vroom V. H. (1964), *Work and motivation*, Wiley, New York, NY.
- Webb R., Vulliamy G., Hamalaninen S., Sarja A., Kimonen E. and Nevalainen R. (2004), “A comparative analysis of primary teacher professionalism in England and Finland”, *Comparative Education*, 40, 1: 83-107.

Author’s note: I would like to thank Lucia M. Collerone for translating this article into Italian.

Emerging policies and practices of teaching evaluation in the Norwegian school system

di *Silje Gløppen e Guri Skedsmo*

1. Introduction

Since the late 1990s, a series of worldwide education reform initiatives have aimed to modernize and improve the quality of education. Inspired by governing logics from New Public Management and influenced by transnational agencies like the Organization for Economic Co-operation and Development (OECD), policy makers across the world have developed and re-invented output-oriented tools for education evaluation and assessment (Gunter et al., 2016; Pons, 2017). Some tools increasingly focus on assessing student performance on standardized tests as a measure of educational quality which is used as benchmarks to compare schools and regions with nations (Skedsmo, 2009). Other tools, such as school inspection and supervision and various evaluation procedures that mainly focus on the school organization also aim to take processes and context into account (Ehren et al., 2013). In addition, teaching evaluation has become a prominent tool in education reform systems for governing and developing the education sector (Hallinger et al., 2014).

2. Teaching Evaluation

The term teaching evaluation used in this chapter is somewhat different than the more commonly used terms “teacher evaluation”, “teacher assessment” or “teacher appraisal”. The reason for using this specific term is based on recognition of the importance of how concepts in use shape content and meaning, and also contribute to legitimizing tools and practices. The term teaching evaluation uses the verb teaching, which signals the interactivity between teachers and students in the learning process, and that the objective for evaluation is the process of instruction itself, not the teacher.

A majority of recent tools and practices for teaching evaluation rests upon external control strategies, and the underlying assumption that performative accountability will lead to enhanced instructional practice, student learning and academic performance (Hallinger et al., 2014; Rowan & Miller, 2007). Some scholars state teaching evaluation has become a part of the global “edu-business” (Ball et al., 2012). Out of 28 countries surveyed in the OECD “Review on Evaluation and Assessment Frameworks for Improving School Outcomes” (2013), 22 reported having national- or state-level policy frameworks for teaching evaluation. Responses from Norway and the five remaining countries from the survey sample, indicate that practices to provide feedback on teachers’ work are designed and implemented locally.

3. Testing and Teaching Evaluation

In the Norwegian context, output-oriented policies and government tools influenced by performative accountability have been introduced during the 2000s, but not without controversy. National testing, which represents an important initiative aimed to improve educational quality, was tried out for the first time in 2004 as a precursor for the National Quality Assessment System (NQAS) which was introduced in 2005. Teacher unions and researchers criticized the tests at that time for not being good enough, and for not being clear on what they were supposed to measure. The country’s largest teachers’ union mobilized, when the leadership of the former ministry of education wanted the results to be published. The controversy about the publication of results led to a delay in implementation, but revised national tests were implemented in 2007.

The NQAS includes a mixture of old and new tools for education quality evaluation (Skedsmo, 2009). The national tests and the international comparative achievement studies, such as the Programme for International Student Assessment (PISA), Trends in International Mathematics and Science Study (TIMSS) and Progress in International Reading Literacy Study (PIRLS) represented new inventions, while the existing School-leaving Examination and the Craft Certificate have constituted the examination system in secondary school for quite some time in Norway and are categorized as traditional tools. Later on, the school inspection was added to the tool kit in Norway.

Currently in Norway, students in grade 5, 8 and 9 must take the national tests in literacy, numeracy and English. There are clear expectations stated in national policies with regard to using the test results and also other information provided by the NQAS to improve practice on different levels of the

school system (Skedsmo, 2009). However, it has been challenging for teachers to use the results for formative purposes to improve teaching and learning (Allerup et al., 2009; Mausestagen et al., 2016; Seland et al., 2013; Skedsmo, 2009, 2011). At local levels, the test results are often used for benchmarking of school quality and to hold key actors, such as teachers and school leaders to account (Skedsmo & Møller, 2016). The extent to which and how the test results are included in accountability practices vary widely among the 426 municipalities in charge of comprehensive schooling.

Teaching evaluation is relatively new in the Norwegian education policy context, and so far, there are not many empirical studies on how such practices are perceived by teachers. Norwegian principals and teachers have until recent years experienced a working context characterized by a relatively high degree of autonomy, and government-backed trust in that professionals do their jobs in the best possible ways. Along with a development towards more output-oriented approaches to govern schools, there has been a slight shift from “trust in the profession” to “trust in results” (Uljens et al., 2013). At the same time, depending on how the quality assessment system in the municipalities are designed, information gathered are also used to hold teachers to account and to set development aims (Skedsmo & Møller, 2016). Some of the most common practices are student evaluation of teachers’ instruction (survey in upper secondary schools), the school walkthrough (observation of teaching by superiors, e.g.: principals, leaders at intermediate level or internal or external “experts”), use of student performance results, employee appraisals, peer counselling, and/or self assessment (portfolios) (Lillejord et al., 2014)

So far, teaching evaluation in Norway is an unexplored field and we have little knowledge about the extent to which and in what ways various tools and practices promote professional development and how the teachers respond to accountability policies (Mausestagen, 2013). A key question linked to this is the degree of trust teachers’ have in practices of teaching evaluation (Lillejord et al., 2014).

In this chapter, we aim to systematize existing knowledge on teaching evaluation, discuss what characterizes tools in emerging practices of teaching evaluation in Norwegian schools and how elements of control and development are played out within practices. This implies an exploration of what tools, methods and information sources are used and how they are legitimized in evaluation procedures.

4. Teaching evaluation in Norwegian policies

Until the beginning of 2000, the working context of teachers was dominated by a professional understanding characterized by autonomy, trust and loyalty to students. A national curriculum provided general as well as subject oriented aims. A process-oriented view of learning was emphasized as well as broader social and humanistic aims in education (Slagstad, 1998; Telhaug et al., 2006). Teachers traditionally had substantial autonomy and were trusted by the public and the state education bureaucracy to teach the curriculum on the basis of their knowledge and values (Mausethagen & Mølsted, 2014; Sivesind, 2008). Local reports formed the basis for decisions regarding infrastructure including buildings, equipment, material, personnel resources as well as compulsory requirements for subjects and courses (Karlsen, 1993; Telhaug, 1994). Based on the information provided, national authorities (i.e., the Ministry of Education) were responsible for resolving problems and making improvements to education. Student achievement was measured in terms of a traditional examination system, and not used to provide educational statistics for governing purposes (Skedsmo, 2009). The teachers were responsible for developing and enhancing instructional practices within the schools and within the profession, and subject to minimal scrutiny from the state or the public.

In 1988-89 the OECD published a report that questioned Norway's lack of documentation on educational quality. This can be seen as a starting point for discussions around quality assurance systems and how to measure and monitor educational quality in terms of student performance (Skedsmo, 2009). During the 1990s, such systems were developed mainly under the umbrella of "school-based evaluation", focusing on developing local schools as organizations (Nyhus & Monsen, 2012).

Parallel to discussions around how to measure and document educational quality, various strategies were implemented to modernize the education system, known as New Public Management (NPM). The Norwegian version of NPM came in two phases (Christensen & Læg Reid, 2011). The first phase emphasized decentralization and delegation and was accompanied by the introduction of management by objectives (MBO) during the 1990s. MBO was proposed to solve problems related to strongly divided and inflexible governing structures, unclear division of tasks and responsibilities, and a lack of national coordination and consistency, all of which encourage disclaiming responsibility (Møller & Skedsmo, 2013). National testing, which was used for the first time in 2004, and NQAS, introduced in 2005, can be seen as initiatives representing the second phase (Christensen & Læg Reid, 2011). National testing and NQAS can be seen as central coordination strategies aimed

to solve problems not yet resolved by the first phase dominated by MBO initiatives. In addition, the policy rhetoric emphasized strong leadership and accountability (Møller & Skedsmo, 2013).

As a *policy theory of action*, accountability can be related to the introduction of new systems for governing with the overall ambition to enhance quality in education (Fuller, 2008; Langfeldt, 2008). The underlying idea of this prevailing theory is that holding municipalities, schools and teachers accountable to certain measures – typically tests and evaluations – will increase students’ performance because teachers will work harder and schools will adopt more effective methods (Heilig, 2011 in Mausethagen, 2013). Skedsmo & Mausethagen (2016) refer to this type of accountability as *performance-oriented*. Individuals are accountable for results achieved in accordance with objectives, standards and tasks. Dependent on practices of use, there is greater emphasis on incentives or sanctions, competition, control and practices for performance monitoring (Skedsmo & Mausethagen, 2016). Based on the focus on performance and outcomes, such practices relate to teaching evaluation and in many municipalities they are part of performance management included in the local quality assessment system (Skedsmo & Møller, 2016).

In addition to new assessment tools, a new curriculum reform for primary, lower and upper secondary schools was launched in 2006: The Knowledge Promotion. This reform contained important strategic areas found in the European Qualifications Framework (EQF), such as guidelines formulated in terms of competency aims. This can be seen as a part of the development towards a more output-oriented means of governing schools which formed the basis for accountability practices to emerge both on municipal and school level (Skedsmo, 2009; Skedsmo & Møller, 2016).

In the wake of the Knowledge Promotion reform, teacher training programs also underwent reforms in direction of output-orientation and MBO. From the study year 2017/2018, students who want to become teachers in primary and lower-secondary education in Norway have to complete a five-year teacher education programme (master degree) or subjects equivalent at various university colleges or universities, in addition to a teacher-training program. Teachers who completed general teacher education before 2005 must supplement any missing credits in Norwegian, English, Mathematics and Natural Sciences within 2025, to be certified as a “general teacher”. This has led to a heated debate as “general teachers” who lack credits in given subjects but have nearly 30 years of classroom experience now seem to be at risk of being condemned as not qualified.

Inspired by EQF, the Royal Ministry of Education and Research introduced the Norway National Qualifications Framework for lifelong learning

(NQF) in 2011. Within this framework, Norwegian educational authorities wanted to map the status of teaching evaluation practices in Norwegian schools. The OECD's report "Reviews of Evaluation and Assessment in Education" in Norway gave examples of practices (Nusche et al., 2011). The report recommended developing teaching evaluation systems as an important step in establishing a comprehensive system of evaluation and assessment in Norway, and as means to «further professionalize the teaching profession». The report also pointed out that so far, very few municipalities have put in place adequate systems for quality evaluation which might include evaluating teachers' practice. The OECD emphasizes that there is not necessarily a direct relationship between teacher appraisal, feedback (OECD's terms), and students' achievement results, but argues that a culture of systematic evaluation and feedback practices bear possibilities for contributions to improvement of instructional practices. Evaluation and feedback as single events will not have the same effect as comprehensive and systematic practices (Nusche et al., 2011). Earlier, in 2009, the Norwegian national education authorities assigned the GNIST partnership – a committee set to increase the quality and status of the teaching profession, teacher education, and school leadership. The committee was led by Professor Eyvind Elstad, Department of Teacher Education and School Research, University of Oslo, and consisted of several stakeholders. In addition to the national educational authorities¹ the committee included representatives from teacher education institutions, school owners (municipalities and counties) and unions for teachers and school leaders. The committee was asked in 2013 to explore the conditions for establishing a system of teaching evaluation. The results were reported in two phases, and suggested four elements that might be fruitful when developing systems for teaching evaluation in Norway (GNIST-arbeidsgruppen, 2014):

1. students' feedback on instructional practices;
2. peer counselling and feedback;
3. observation and follow-up by school leaders;
4. information on students' learning, Bildung and development.

The report emphasized that if these four elements were used formatively, and seen in relation to each other, it could potentially lead to development of the individual teacher as well as the entire teaching staff in terms of professional development, enhancement of overall evaluation culture in schools, and development of schools as learning organizations (Elstad et al., 2015). In addition, one important finding from this work was that based on solid

¹Observers from the Ministry of Education, secretariat from the Norwegian Directorate for Teaching and Training.

international and national empirical support (Bryk & Schneider, 2002; Christophersen et al., 2011, 2012; Elstad et al., 2012, 2015; Tschannen-Moran & Gareis, 2015), the prerequisite for implementation of teacher evaluation is a party cooperation between the assessed and the assessor, built on mutual, relational trust. In Norwegian context, this confidence will affect how and whether evaluation practices will work as intended, which is to contribute to better instructional practice for better learning outcomes (GNIST-arbeidsgruppen, 2014).

At the same time as the GNIST-committee worked on their report, the Knowledge Centre for Education was asked by the Norwegian national education authorities to develop a systematic knowledge synthesis of research on teaching evaluation. The report was published in 2014 and it presented findings mainly from the OECD-countries and China. The report followed to a large extent the same recommendations for policies and practices as reported by the GNIST-committee; improvement of practice is the goal, but success is dependent on mutual agreement on the basis for improvement (Lillejord et al., 2014). The report points out that important prerequisites to consider before implementing tools for evaluation in the Norwegian context, are developing standards and clear criteria. The systematic review also reported that systems for teaching evaluation often have double functions: they are to be both formative; as to contribute to professional development, and summative to measure and control results. In all the systems examined in the review (from Chile, China, Belgium and Portugal) there was a tendency that the summative assessment forms became dominant, and that the formative sides of the assessment systems were pushed aside. In other words, assessment systems were likely to be used solely for bureaucratic purposes instead of contributing to the intended professional development (Lillejord et al., 2014).

Both the earlier mentioned OECD-report from 2011 (Nusche et al., 2011) and results from the OECD Teaching and Learning International Survey (TALIS) have been referred to when Norwegian education authorities argue for the need to implement teaching evaluation. Norway is approximately equal to the international TALIS average in terms of proportion of teachers who have not received feedback on the work they carry out (Caspersen, 2016). Only 16 percent of a representative selection of Norwegian teachers said in 2013 that they had never received any formal feedback on their work. More than half of the teachers had received feedback from colleagues, whereas less than half reported to have received formal feedback from the principal or someone in the school leadership team and 44% said that feedback had been given by someone in the leadership team. Only 10% of teachers had feedback from external experts (Caspersen, 2016). Importantly, it is pointed out that the follow-up after assessment was found to be weaker in Norway compared to

other countries (Caspersen, 2016). Only 13% of the principals in the TALIS survey reported that development and training plans were discussed often or always following assessment of teachers, a result that is reported to be below the average in all the other Nordic countries participating in TALIS (Sweden, Finland and Iceland) (Caspersen, 2016). Systematic and comprehensive actions for improvement of the individual teachers' instructional practices was also rare in Norway, compared to other countries (Caspersen, 2016). Hybertsen et al. (2014) found however, that principals who have attended school leadership training, to a greater degree assess teachers more often than principals who have no formal school leader education.

Teaching evaluation as a tool for measuring educational quality or controlling of teachers' work can be seen as a break with the normative interpretation of the teacher professions mandate, thus creating a tension between the experience of external control or accountability, and professional identity, responsibility and development (Mausethagen, 2013; Skedsmo, 2009). As pointed out by members of the GNIST committee and the Knowledge Centre for Education, mutual trust seems to be one of the most important prerequisites for a successful implementation of teaching evaluation systems. Trust is earned and cannot be easily produced or managed. It is also a multi-faceted construct, and what elements people put into the construct depend on the context or nature of the trust relationship (Tschannen-Moran & Gareis, 2015). As mentioned above, the explicit and main justification by the OECD and the Norwegian educational authorities for implementing teaching evaluation systems is to develop and «further professionalize the teaching profession» (Nusche et al., 2011) in order to possibly develop better teaching practices for better student learning outcomes. One of the core issues of this chapter is therefore to give an outline of how professional development is addressed in emerging tools and practices of Norwegian teaching evaluation.

5. Practices of teaching evaluation

Norwegian schools have a tradition of peer counseling that is a method characterized by a symmetric relationship between the colleagues involved, where reflections on practice are put at the core of the activity (Lauvås et al., 2004). This method is integrated in teacher education and teachers' practice and aims to further develop their "theory of practice". Theory of practice refers to a general capacity to deal with challenges in the daily work and comprises experience, observations, theoretical knowledge and understanding, values and ideals. As such, peer counselling can be described as learning

oriented and it generates knowledge, norms and ethics for the teacher profession (Lauvås et al., 2004).

So far, there are no national, comprehensive strategies or frameworks for conducting teaching evaluation in Norway. The specific details are left up to local educational authorities; counties and municipalities to develop, establish, and conduct systems for teaching evaluation as part of their local quality assurance system. Norway has 426 municipalities and 18 counties, and very different preconditions for systematic and comprehensive evaluation systems. Teaching evaluation practices are locally developed and the tools in use normally include student surveys and national testing as well as examination results. The extent to which the information from these tools are embedded in accountability practices vary. The use of evaluation technologies both by managers at the municipal or county level and principals to monitor students' outcomes can be read as a shift towards what has been termed organizational professionalism (Evetts, 2009). In the Norwegian school system, this is exemplified by increased standardization of work procedures, external regulation and accountability measures (Skedsmo & Møller, 2016).

The only nationally developed document specifically describing practices of teaching evaluation is a Guideline for Student Evaluation of Instruction, developed in 2011 by the Norwegian Directorate for Education and Training, the Student Organization, the Norwegian Association of Local and Regional Authorities, and the Union of Education Norway (2011). The document is regarded as a new tool in a framework for teaching evaluation and is mainly used in upper secondary education. The guideline is not considered suitable for teaching evaluation in primary education (GNIST-arbeidsgruppen, 2014). The document is marked by a high degree of student participation and is to a large extent influenced by the Students Organization's agenda. The guideline draws some general principles that shine light upon intentions for evaluation of instruction. First, student evaluation of instruction is supposed to be carried out by the teacher and the students – in a dialogue. The guide states explicitly that evaluation should “develop the dialogue between teacher and student”, and that it should concern “issues that they together or individually can do something about”. The assessment should be directly related to the training in the individual subject, and deal with learning objectives, working methods, learning strategies, content and organization in the classes. Based on these principles, students and teachers follow four phases in quality development.

- Phase 1. Students and teachers prepare for the assessment work. They are familiar with the purpose of assessing the instruction in subjects, who to involve and how. Furthermore, they must determine the procedure and how to follow up the results.

- Phase 2. The students and the teacher will gather information about what parts of the instruction they will evaluate. There are different ways to do this: using dialogue, writing personal logs, observation and questionnaires.
- Phase 3. Information is systematized and followed up by dialogue between the teacher and the student. This in order to form a common basis for further developing the instruction.
- Phase 4. Follow-up points, responsibilities and deadlines are needed, to achieve successful development work.

To a certain degree, the guideline reflects an important characteristic of evaluation practices in Norway today. The Student Organization's impact on the framework, as well as the focus on dialogue, can be interpreted as a direction toward a student-centered perspective on teaching evaluation, as well as a framework for formative assessment. Evaluation is conducted at the "lowest level", as the students in the classrooms are given the opportunity to give direct feedback on instructional practices, and the teacher is there to consider and discuss results. Emphasis is on dialogue and methods of formative character, and the aim is to form a common basis for further developing the instruction.

During the work on developing the guidelines for student evaluation of instruction and also the work on the GNIST-reports, questions about anonymity and students' competence of evaluation instructional practices were raised and discussed. The Students Organization argued that students should have the opportunity to stay anonymous, due to concerns of being subjected to sanctions if critical judgments were raised. The Norwegian Association of Graduate Teachers however, stated that principles within the Knowledge Promotion document emphasize education for democratic understanding and participation, and that anonymity could compromise these underlying values (GNIST-arbeidsgruppen, 2014). The discussion about students' evaluation skills and maturity also brought up that some students might use anonymity as an excuse to not give accurate answers, or even try to rate the teacher negatively as a response to dissatisfaction with the teacher's grading or other personal attributes. The latter arguments seem to be overlooked, as current practices seem to be that the students anonymously fill out surveys and questionnaires.

The purpose of teaching evaluation raises fundamental questions. Evaluation practices imply that the teacher's work is evaluated systematically. This affects the interests of employee groups, employers, national authorities and students (GNIST-arbeidsgruppen, 2014). The severity increases significantly when assessments are tied to consequences, like we have seen in other countries, for instance the USA, England and Chile. Although one of the most

commonly used tools for teaching evaluation in Norway can be regarded as relatively “low stakes”, and to a certain degree in line with formative aspects and a professional orientation, there has been tensions and controversies in the development of teaching evaluation in Norway.

6. Tensions in emerging teaching evaluation practices

In the case of Norway, one overall challenge is in the tension between teachers’ perception of autonomy and trust, and national and local educational authorities’ systems of external quality control and expectations about improvement. How this tension manifests itself depends on how elements of control and accountability are locally played out and perceived among practitioners. The following section presents some examples that relate to conceptual issues and terms used about evaluation procedures, as well as possible misuse of performance data linked to individual persons.

The GNIST commission describes controversies, especially connected to the terms “teacher assessment”, “teacher evaluation” or “teacher appraisal”. The OECD uses the term “teacher appraisal”. These terms, often used in international literature, have been subject to heated discussions between teacher unions and interest groups in Norway. Findings from The Knowledge Center for Education, and Phase 1 by the working group of the GNIST-partnership, suggest that assessment must be limited to concentrate on specific aspects of the work and that it should be considered part of a larger context (GNIST-arbeidsgruppen, 2014; Lillejord et al., 2014). If the assessment is to have a positive impact on the quality of schools, it must also be based on a combination of several methods (Lillejord et al., 2014). Further, the GNIST working group concludes that «teacher assessment/evaluation/appraisal is not only insufficient to describe the content of an assessment system; it can also be perceived as misleading» (GNIST-arbeidsgruppen, 2014, authors’ translation). It is not the personified teacher that is to be evaluated, but his or her instructional practices. Based on this, the term “teacher assessment” has potential to be poorly received, even though evaluation systems prove to be consistently good (GNIST-arbeidsgruppen, 2014).

The Union of Education Norway and The Norwegian Association of Graduate Teachers argue that the term “teaching assessment” or “teaching evaluation” should be used, in accordance with this statement by a Union representative: «Confidence and credibility, legitimacy and relevance were circumscribed by the working group as the basis for further work. The Union of Education Norway believes the term Teaching Assessment best meets this criterion. The Union of Education Norway believes Teaching Assessment

includes both students' feedback on the learning situation, colleague/peer observation and counselling and the managements' observations and follow-up» (GNIST-arbeidsgruppen, 2014, p. 14).

The Student Organization suggests “teaching evaluation”, whereas the Education Science Students in the Union of Education Norway prefer “learning development” and argue that the name choice is as important as the main purpose of the system is professional development. The Association of School Leaders uses the term “professional development assessment system”, while The Norwegian Association of Local and Regional Authorities uses the term “system for professional development in schools” (GNIST-arbeidsgruppen, 2014). The work group also points out that there already exist local systems for teaching evaluation, and that the name for a comprehensive Norwegian system has to take former and/or new facets of practices into account (*ibidem*).

Another important question relates to purposes of data use and The Norwegian Personal Data Act. If personal information is collected in writing and stored either electronically or in a manual personal register at the school when obtaining students' feedback, the Personal Data Act will apply (GNIST-arbeidsgruppen, 2014). In such cases there must be a legal basis for the treatment of data in the form of a voluntary, explicit and informed consent. Consent must be provided by the students who evaluate and from the teachers being evaluated. The declaration of consent must specify the information to be collected, the purpose of the collection, who should have access to the information and when the information is to be deleted. It must also be stated that the consent may be called back at any time and that the consequence of such a revocation is that the information collected must be deleted (GNIST-arbeidsgruppen, 2014). Should the students' evaluation of the teachers be done through, for example, an electronic questionnaire administered by a technical provider (data processor), the person responsible for the processing (either school or school owner) must ensure that there is a written agreement with the data processor. Furthermore, the processor must ensure that information security and internal control procedures are established for the processing of personal data (GNIST-arbeidsgruppen, 2014).

With regards to national tests, mapping tests and other surveys that are authorized by regulation by the Education Act section 2-4, it follows that the purpose of the provision is to map the students' learning outcomes and basic skills. It is not within the scope of the collection of personal data by regulation to the Education Act §2-3 or §2-4 to map and evaluate teachers at the individual level – such use can therefore not be said to be within the original purpose of the collections. Students and teachers must consent to use should personal information be collected via pupil examination, national examinations and

mapping tests used as an element in a model for teaching evaluation. This implies that a declaration of consent must be drawn up, specifying what information has already been collected that will be used in the assessment scheme, what purpose of the use is, who will have access to the information and when the information is to be deleted. It must also be stated that the consent can be called back at any time (GNIST-arbeidsgruppen, 2014).

Current practices indicate that teachers are not particularly concerned with aspects of the Norwegian Personal Data Act when it comes to the use of aggregated data such as national test results for evaluating instructional practices. The reasons for this have not been explored, but one possible explanation could be that the Norwegian teachers have little awareness and knowledge about the law, and how it might apply to the teachers' everyday life. Another might be that current practices of teaching evaluation are not necessarily directly connected to the use of these types of data but used as basis by school leaders to reflect on how teachers' practice can enhance student learning. These types of reflections take place in meetings for teacher teams and employee appraisals (Hjemmen & Hopøy, 2016; Mausethagen et al., 2017).

In international research on teacher evaluation, the role of principals, their competencies to evaluate teachers' practice and how evaluation procedures are embedded in instructional or learning-centred leadership are problematized (Briggs & Dadey, 2017; Hallinger et al., 2014; Lavigne & Chamberlain, 2017; Skedsmo & Huber, 2017). So far, there is a lack of research when it comes to how school leaders in Norway carry out teaching evaluation. Nevertheless, as practices for teaching evaluation are emerging in Norway, and is increasingly being tied to performative accountability, we might experience increased awareness about these issues.

One example of how Norwegian teachers are becoming more aware of changes in the understanding of their professionalism is the teacher strike in autumn 2014. Norway experienced an extensive long-term teacher strike, which in many ways can be interpreted precisely as an expression of a tension between external control and professional trust. The largest teacher union – the Union of Education Norway (Utdanningsforbundet) – stated that the cause for the strike was a long built-up frustration and mistrust, because of the great distance between the central authorities' visions and goals for the school, and the reality that members experienced in everyday life (Lied & Fjørtoft, 2014). While national politicians have formulated ambitious goals for the school, teachers have experienced that the opportunities to concentrate on core tasks of teaching has deteriorated. The teacher strike was explained as a struggle to defend the teacher professions uniqueness and to let teachers be teachers (Lied & Fjørtoft, 2014). Some of the explanations for

this conflict/tension may be found in the relatively swift change in understanding of the teachers' role due to new policies influenced by the international trends mentioned in the introduction.

7. Purposes of teaching evaluation

In the following section we outline some theoretical aspects of the debates found within the research field of teaching evaluation and connect our findings to what we have presented from the Norwegian context. First, we briefly consider formative and summative perspectives on assessment and evaluation and relate these to purposes of teaching evaluation. We then discuss how possible discrepancies between formulated purposes of teaching evaluation and practices of teaching evaluation initiated, especially when those practices are integrated in governing systems emphasizing elements such as performance management linked to certain forms of accountability.

Decades of research on assessment practices have concluded with some general principles that form a basis of modern assessment theory used in frameworks for assessing students' learning worldwide. These principles include the work by the British research group Assessment Reform Group (ARG), and especially Black and Wiliam (1998, 2009), Stobart (2008) and Hayward and Spencer (2010). The principles emphasize formative aspects of assessment and are situated within a sociocultural view on teaching and learning. Hence, assessment systems ought to undertake all available information sources within and surrounding schools. Other important principles are that (a) the objects of assessment must be clear to both the assessor and the assessed; (b) the methods for conducting the assessment must be agreed upon; and (c) the purposes of the assessment should be clear and mutually acknowledged by both parties (Black et al., 2002).

The principles are primarily directed towards the teacher-student setting, but are considered universal and could be just as applicable in governing good assessment practices in schools as well as in governing municipalities or corporations interested in developing assessment systems which promote learning and development (Lillejord et al., 2014). Mutual trust between parties is also considered a prerequisite for professional learning and development in schools (Tschannen-Moran, 2009). In addition, intrinsic motivation theories form the basis for both development of trusting professional organizations and formative assessment practices.

Teaching evaluation occurs under different labels, and connotations of the concepts are related to the context in which they appear and are being used. Hallinger, Heck and Murphy (2014) contrast "teacher performance

evaluation” and “instructional supervision”. The first concept is defined as «the formal assessment of a teacher by an administrator, conducted with the intention of drawing conclusions about his/her instructional performance for the purpose of making employment decisions» (Castetter in Hallinger, Heck and Murphy, 2014). In this definition, an external discretionary power is to evaluate the teacher’s work. This model of assessment relies on external motivation theory, with focus on economics and extrinsic incentives (Firestone, 2014). It can be understood in line with NPM and performance management, where no improvement will take place unless the performance is assessed and the person held to account (Clarke & Newman, 1997; Ranson, 2003). We also find that performance-based evaluation often relies on the use of summative forms of assessment, such as for instance student performance data.

The concept “instructional supervision” is seen as a more growth-oriented type of coaching, conducted by administrators, supervisors or peers, where the aim is to develop teaching capacity through a process of observation and feedback (Hallinger et al., 2014). It corresponds more with the use of formative forms of assessment, like peer-counselling or student feedback. Concepts where focus is growth-oriented and procedural can be seen as relying on internal motivation theory which uses psychology and intrinsic incentives (Firestone, 2014).

Although purposes of evaluation reforms often focus on formative aspects and developmental purposes, the results of assessment reforms have showed some tensions and discrepancies between government assessment policies and classroom assessment practices (Berry & Adamson, 2011). Berry and Adamson advocate that the teachers’ classroom practices of assessment still strongly associate assessment with controlling instead of linking assessment to developing processes related to teaching and learning. They claim this might be due to the increase of externally initiated, international comparative tests or extensive testing within a high-stakes setting, which undermine processes needed for keeping focus on assessment *for* learning, giving more focus to the outcomes to be measured and monitored (Berry, 2011). This also seems to be the challenge in teaching evaluation.

Governing systems that rely on practices described above are usually and broadly influenced by NPM (Christensen & Læg Reid, 2011; Gregory, 2003) and performance-based managerialism (Clarke & Newman, 1997). In the context of NPM, standardized testing is often coupled with elements of competition (for example, publication of student results), privatization and contractualism, where the underlying assumption is that free choice will grant development of best practice. Countries vary regarding the role that market mechanisms play, such as parental choice, and whether incentives are connected to student test scores, such as performance-related pay for teachers

(e.g. Jabbar, 2013). Such contexts are usually described as “high-stakes” accountability contexts.

Within high-stakes contexts we often find teacher evaluation models that build on the use of Value Added Measures (VAM). VAMs are supposed to represent the individual teacher’s impact on student learning during the prior year, and describe important dimensions of the teachers “effectiveness” (e.g. Hanushek & Rivkin, 2010; Harris & McCaffrey, 2010). As such, these models aim to provide a fairer approach to evaluate teachers’ work than students’ performance on standardized tests. VAMs have been subject to research for more than two decades, and there is a significant body of research questioning the validity, reliability and fairness in the use of these models (e.g. Amrein-Beardsley, 2008; Amrein-Beardsley et al., 2016; Berliner, 2014; Hallinger et al., 2014; Kelly & Downey, 2010).

As described above, one important challenge of teaching evaluation is for what purposes it is used and also how the practices and tools used to carry out the evaluation processes align with proclaimed purpose. Marzano (2012) distinguishes the contrasting purposes of assessment, as “measuring” teachers and “developing” teachers have different implications. Although both measurement and development are important aspects of teaching evaluation, an evaluation system designed primarily for measurement will look quite different from a system designed primarily for development. When measurement is the primary purpose, a small set of elements is sufficient to determine a teacher’s skill in the classroom. However, if the emphasis is on teacher development, the model needs to be both comprehensive and specific and focus on the teacher’s growth in various instructional strategies (Marzano, 2012). If the purpose of evaluation is professional development, one important prerequisite is the teachers’ motivation to develop. Firestone (2014) found that intrinsic and extrinsic motivation theories and available evidence raise doubts about for instance performance based pay but not necessarily the use of other extrinsic incentives. Motivation theories also suggest that to maintain effective intrinsic incentives, policies to remove “ineffective teachers” must not reduce autonomy or trust among “effective teachers” (Firestone, 2014).

This way of managing is considered to be the dominant form of accountability within the field of education in many countries, and is often described as hierarchic, administrative or management-oriented (Murphy et al., 2013; Romzek & Dubnick, 1987; Sinclair, 1995). Biesta (2010) even argues that the rhetoric in education has caused individuals to associate accountability with responsibility, which positions those arguing against accountability as proponents of irresponsible action.

8. Discussion and recommendations

Teaching evaluation in Norway has not yet been sufficiently empirically studied for us to be able to identify possible outcomes of teaching evaluation practices. However, there are some findings from research on implementation and enactment of assessment policies that seem relevant to highlight. First, we find that there are tensions between teachers and teacher unions and the education authorities when it comes to implementing new tools for evaluating teachers' work. The main tension we have addressed, that is between external control versus professional trust and development, can also be connected to many of the most prominent debates in the many discourses of education quality worldwide.

What determines “good quality education” is a question of legitimacy, power, ethics, values, moral and the many different aspects of knowledge and knowledge production itself. What constitutes “the good teacher” is subject to scrutiny in the field of teaching evaluation as well as in the overall education debate. However, conclusions we can draw from the outline above is that teaching evaluation has become a part of the globalized education field of policies and practices which will influence emerging practices of teaching evaluation in Norway.

What we also find, is that the Norwegian teachers to a certain degree seem to have a relatively strong position in educational policy making, as they, to some extent, are invited to participate in both the process of examining prerequisites for implementation of teaching evaluation (through the GNIST-committee) and the making of guidelines and frameworks for practices (guidelines for student evaluation of instruction). As an example, reports from the GNIST committee clearly voice the teachers' and the unions' perspectives when it comes to overarching challenges connected to how evaluation practices might cohere or not cohere with the terms used to describe desired outcomes of teaching evaluation. Teachers' voices are communicated in the discussion of the term “teacher evaluation/assessment” versus “teaching evaluation”, and thus raise interesting and fundamental questions to the understanding of concepts and what teaching evaluation really is supposed to be. Also, when it comes to issues related to democracy and participation that are emphasized in the national curriculum in Norway, teachers question current practices of evaluating instruction where students fill out questionnaires anonymously. These findings are interesting with regard to what we have learned from implementation processes in other countries; the balance of power and participation seems to be rather unique in the Norwegian context.

Teaching evaluation might contribute to professional development for teachers, the teaching profession and schools, if purposes and practices are aligned and supported by recent research. Based on theories and findings from international research conducted during the last two decades, we argue that precautions must be considered when nations implement systems for teaching evaluation. From what we have learned, the risk of turning teaching evaluation into tools for performance management with a strong focus on controlling and monitoring teachers according to student performance on standardized tests and other types of outcomes may have severe implications for teachers' motivation for professional development, with possible further impact on instructional practices and student learning.

It is therefore important to critically and thoroughly examine findings and knowledge from countries that have experienced effects of evaluation policies implemented, with a focus on intended as well as unintended consequences. We also suggest that such empirical research should be combined with, and seen in relation to evaluation theory, theories of motivation, and knowledge about how to create trustful, cooperative and learning organizations. In addition, when getting inspired by policies and concepts from other countries, culture, traditions and context need to be carefully considered to develop national or local practices for teaching evaluation that are perceived as relevant and useful by teachers, but also other key actors such as school leaders and students.

References

- Allerup P., Velibor K., Kvåle G., Langfeldt G. and Skov P. (2009), "Evaluering av det nasjonale kvalitetsvurderingssystemet for grunnsopplæringen (Evaluation of the national quality assessment system for the comprehensive education system)", FoU-rapport, 8, Kristiansand.
- Amrein-Beardsley A. (2008), "Methodological concerns about Education Value-Added Assessment System", *Educational Researcher*, 37: 65-75.
- Amrein-Beardsley A., Polasky S., and Holloway-Libell J. (2016), "Validating 'value added' in the primary grades: one district's attempts to increase fairness and inclusivity in its teacher evaluation system", *Educational Assessment, Evaluation and Accountability*, 28, 2: 139-159, doi: 10.1007/s11092-015-9234-5.
- Ball S., Maguire M. and Braun A. (2012), *How schools do policy. Policy enactments in secondary schools*, Routledge, Oxon.
- Berliner D. (2014) "Exogenous variables and value-added assessments: A fatal flaw", *Teachers College record*, 116, 1: 1-31.
- Berry R. (2011), *Assessment reforms around the world*, in Berry R. and Adamson B., eds., *Assessment reform in education: Policy and practice*, Springer, New York.

- Berry R. and Adamson, B., eds. (2011), *Assessment reform in education: Policy and practice*, Springer, New York.
- Biesta G.J.J. (2010) “Why ‘What Works’ Still Won’t Work: From Evidence-Based Education to Value-Based Education”, *Studies in Philosophy and Education*, 29: 491-503, doi: 10.1007/s11217-010-9191-x.
- Black P., Harrison C., Lee C., Marshall B., and Wiliam D. (2002), *Working inside the Black Box: assessment for learning in the classroom*, GL Assessment, London.
- Black P. and Wiliam D. (1998), “Assessment and classroom learning”, *Assessment in Education: Principles, Policy & Practice*, 5, 1: 7-74, doi: 10.1080/0969595980050102.
- Black P. and Wiliam D. (2009), “Developing the theory of formative assessment”, *Educational Assessment, Evaluation and Accountability*, 21, 1: 5-31, doi: 10.1007/s11092-0089068-5.
- Briggs D.C. and Dadey N. (2017), “Principal holistic judgments and high-stakes evaluations of teachers”, *Educational Assessment, Evaluation and Accountability*, 29, 2: 155-178, doi: 10.1007/s11092-016-9256-7.
- Bryk A. and Schneider B. (2002), *Trust in Schools. A Core Resource for Improvement*, Russell Sage Foundation, New York.
- Caspersen J. (2016), *Can collegial work and school leader feedback improve teachers' self-efficacy in Nordic classrooms?*, in S. Ludvigsen, S., Nortvedt G., Pettersen A., Pettersson A., Sollerman S., Ólafsson R., Taajamo M., Caspersen J., Nysström P. and Braeken J., eds., *Northern Lights on PISA and TALIS*. Copenhagen: Nordic Council of Ministers.
- Christense T. and Læg Reid P. (2011), *The Ashgate Research Companion to New Public Management*, Ashgate Research Companion, Farnham.
- Christophersen K.-A., Elstad E. and Turmo A. (2011). “The nature of social practice among school professionals: consequences for the school academic pressure exerted by teachers in their teaching”, *Scandinavian Journal of Educational Research*, 55, 6: 639-654, doi: 10.1080/00313831.2011.594605.
- Christophersen K.-A., Elstad E. and Turmo A. (2012), “The strength of accountability and teachers’ organisational citizenship behaviour”, *Journal of Educational Administration and History*, 50, 3: 612-628, doi: 10.1108/09578231211249844.
- Clarke J. and Newman J. (1997), *The Managerial State: power, politics and ideology in the remaking of social welfare*, Sage Publications Ltd, London.
- Ehren M.C.M., Altrichter H., McNamara G. and O’Hara J. (2013), “Impact of school inspections on improvement of schools—describing assumptions on causal mechanisms in six European countries”, *Educational Assessment, Evaluation and Accountability*, 25, 1: 3-43, doi: 10.1007/s11092-012-9156-4.
- Elstad E., Christophersen K.-A. and Turmo A. (2012), “Antecedents of school teaching intensity within two different management regimes: an assessment-based accountability regime and regime without external pressure on results”, *International Journal of Education Policy and Leadership*, 7, 6.
- Elstad E., Lejonberg E. and Christophersen K.-A. (2015), “Teaching evaluation as a contested practice: Teacher resistance to teaching evaluation schemes in Norway”, *Education Inquiry*, 6, 4:375- 399, doi: 10.3402/edui.v6.27850.

- Firestone W. (2014), "Teacher evaluation policy and conflicting theories of motivation", *Educational Researcher*, 43, 2: 100-107.
- Fuller B. (2008), *Overview: Liberal learning in centralising states*, in Fuller B., Henne M. K. and Hannum E., eds., *Strong states, weak schools: The benefits and dilemmas of centralised accountability*, Emerald Group, Bingley.
- GNIST-arbeidsgruppen. (2014), *Rapport til drøfting i GNIST-partnerskapet fra arbeidsgruppen som har utredet forutsetninger for en ordning med lærervurdering*, Oslo.
- Gregory R. (2003), *Accountability in Modern Government*, in Peters B. G. and Pierre J., eds., *Handbook of Public Administration*, Sage, London.
- Gunter H., Hall D., Serpieri R. and Grimaldi, E., eds. (2016), *New Public Management and the Reform of Education: European lessons for policy and practice*, Routledge, London.
- Hallinger P., Heck R. and Murphy J. (2014), "Teacher evaluation and school improvement: An analysis of the evidence", *Educational Assessment, Evaluation and Accountability*, 26, 1: 5-28.
- Hanushek E.A. and Rivkin S.G. (2010), "Generalizations about using value-added measures of teacher quality", *American Economic Review*, 100: 267-271.
- Harris D.N. and McCaffrey D.F. (2010), *Value-added: Assessing teachers' contributions to student achievement*, in Kennedy M., ed., *Teacher assessment and the quest for teacher quality*, Jossey-Bass, San Francisco, CA.
- Hayward L. and Spencer E. (2010), "The complexities of change: formative assessment in Scotland", *Curriculum Journal*, 21, 2: 161-177, doi: 10.1080/09585176.2010.480827.
- Hjemmen S.S. and Hopøy M. (2016), *Medarbeidersamtalen - Et ledelsesverktøy for læring og utvikling i skolen?*, Master Thesis, University of Oslo, Oslo.
- Hybertsen I.D., Stensaker B., Federici R.A., Olsen M.S., Solem, A. and Aamodt, P.O. (2014), *Ledet til endring, Nasjonal rektorutdanning i grunn- og videregående skole; endringer på skolene, måloppnåelse og anbefalinger, Sluttrapport fra evalueringen av den nasjonale rektorutdanningen*, Nordisk institutt for studier av innovasjon, forskning og utdanning (NIFU), Oslo.
- Jabbar H. (2013), "The case of 'payment-by-results': re-examining the effects of an incentive programme in nineteenth-century English schools", *Journal of Educational Administration and History*, 45, 3: 220-243.
- Karlsen G. (1993), *Desentralisering - løsning eller oppløsning?*, Ad Notam Gyldendal, Oslo.
- Kelly A. and Downey C. (2010), "Value-added measures for schools in England: looking inside the 'black box' of complex metrics", *Educational Assessment, Evaluation and Accountability*, 22, 3: 181-198, doi: 10.1007/s11092-010-9100-4.
- Langfeldt G. (2008), *Ansvarsstyring i utdanningssektoren - statlig ambisjon og ideologi*, in Langfeldt G., Elstad E. and Hopmann S., eds., *Ansvarlighet i skolen. Politiske spørsmål og pedagogiske svar*, Cappelen Akademisk Forlag, Fagernes.
- Lauvås P., Lycke K. and Handal G. (2004), *Kollegaveiledning i skolen.*, Cappelen Akademisk Forlag, Oslo.

- Lavigne A.L. and Chamberlain R.W. (2017), "Teacher evaluation in Illinois: school leaders' perceptions and practices", *Educational Assessment, Evaluation and Accountability*, 29, 2: 179-209, doi: 10.1007/s11092-016-9250-0.
- Lied R. and Fjørtoft T. (2014), *Letter to Members of the Teacher Union*, available online: <https://www.utdanningsforbundet.no/nyheter/2014/til-vare-ledermedlemmer>.
- Lillejord S., Kristin B., Ruud E., Hauge T.E., Hopfenbeck T.N., Astrid T. et al. (2014), *Former for lærervurdering som kan ha positiv innvirkning på skolens kvalitet. En systematisk kunnskapsoversikt*, Kunnskapssenter for utdanning, Oslo.
- Marzano R.J. (2012), "The Two Purposes of Teacher Evaluation", *Education Leadership*, 70, 3: 14-19.
- Mausethagen S. (2013), "Accountable for what and to whom? Changing representations and new legitimization discourses among teachers under increased external control", *Journal of educational change*, 14, 4: 423-444, doi: 10.1007/s10833-013-9212-y.
- Mausethagen S. and Mølstad C.E. (2014), *Licence to teach? Læreplananalyse og profesjonsutvikling*, in Elstad E. and Helstad K., eds., *Profesjonsutvikling i skolen*, Universitetsforlaget, Oslo.
- Mausethagen S., Skedsmo G. and Prøitz T.S. (2016), "Ansvarliggjøring og nye organisasjonsrutiner i skolen – rom for læring?", *Nordiske organisasjonsstudier*, 18, 2: 79-97.
- Mausethagen S., Skedsmo G. and Prøitz T.S. (2017), *Bruk av elevresultater for å styrke og lede profesjonsutvikling – utfordringer og muligheter*, in Aas M. and Paulsen J.M., eds., *Ledelse i fremtidens skole*, Fagbokforlaget, Oslo.
- Murphy J., Hallinger P. and Heck R.H. (2013), "Leading via teacher evaluation: the case of the missing clothes?", *Educational Researcher*, 42, 6: 349-354.
- Møller J. and Skedsmo G. (2013), "Modernizing education - NPM reform in the Norwegian education system", *Journal of Educational Administration and History*, 45, 4: 336-355, doi: 10.1080/00220620.2013.822353.
- Norwegian Directorate for Education and Training (2011), *Undervisningsvurdering - en veileder for elever og lærere (Teaching evaluation - guidelines for students and teachers)*, available online: https://www.udir.no/globalassets/upload/skoleutvikling/veiledere/udir_veileder_undervisningsvurdering_9nov2011.pdf.
- Nusche D., Earl L., Maxwell W. and Shewbridge C. (2011), *OECD Reviews of Evaluation and Assessment in Education*, OECD, Norway.
- Nyhus L. and Monsen L. (2012), "School development as a communication process", *Policy Futures in Education*, 10, 4: 461-474.
- Pons X. (2017), "Fifteen years of research on PISA effects on education governance: A critical review", *European Journal of Education*, 52: 131-144.
- Ranson S. (2003), "Public Accountability in the Age of Neo-Liberal Governance", *Journal of Education Policy*, 18, 5: 459-480.
- Romzek B.S. and Dubnick M.J. (1987), "Accountability in the Public Sector: Lessons from the Challenger Tragedy", *Public Administration Review*, 47, 3: 227-238.
- Rowan B. and Miller R.J. (2007), "Organizational Strategies for Promoting Instructional Change: Implementation Dynamics in Schools Working With Comprehensive School Reform Providers", *American Educational Research Journal*, 44, 2: 252-297.

- Seland I., Vibe N. and Hovdhaugen E. (2013), *Evaluering av nasjonale prøver som system (Evaluation of the system of national standardised testing)*, NIFU, Oslo
- Sinclair A. (1995) "The Chameleon of Accountability: Forms and Discourses", *Accounting Organizations and Society*, 20, 2/3: 219-237.
- Sivesind K. (2008), *Teacher education in the light of the assessment model in curriculum making*, Dissertation lecture for the degree of Doctor Philos, Faculty of Education, University of Oslo.
- Skedsmo G. (2009), *School Governing in Transition? Perspectives, Purposes and Perceptions of Evaluation Policy*, (PhD), University of Oslo.
- Skedsmo G. (2011), "Formulation and realisation of evaluation policy: inconsistencies and problematic issues", *Journal of Educational Assessment, Evaluation and Accountability*, 23, 1: 5-20.
- Skedsmo G. and Huber S.G. (2017), "Evaluation of educators' performance-balancing various measures to improve practice", *Journal of Educational Assessment, Evaluation and Accountability*, 29, 2: 107-110, doi: 10.1007/s11092-017-9262-4.
- Skedsmo G. and Mausesthaugen S. (2016), *Accountability policies and educational leadership - a Norwegian perspective*, in Easley J. and Tulowitzki P., eds., *Accountability and Educational Leadership – country perspectives*, Routledge, London.
- Skedsmo G. and Møller J. (2016), *Governing by new performance expectations in Norwegian schools*, in Gunter H., Hall D., Serpieri R. and Grimaldi E., eds., *New Public Management and the Reform of Education: European lessons for policy and practice*, Routledge, London.
- Slagstad R. (1998), *De nasjonale strateger (The national strategies)*, Pax, Oslo.
- Stobart G. (2008), *Testing Times: The Uses and Abuses of Assessment*, Routledge, Abingdon.
- Telhaug A.O. (1994), *Norsk skoleutvikling etter 1945 (Norwegian school development after 1945)*, Didakta, Oslo.
- Telhaug A.O., Mediås O.A. and Aasen P. (2006), "The Nordic Model in Education: Education as part of the political system in the last 50 years", *Scandinavian Journal of Educational Research*, 51, 3: 245-283.
- Tschannen-Moran M. (2009), "Fostering Teacher Professionalism in Schools. The Role of Leadership Orientation and Trust", *Educational Administration Quarterly*, 45, 2: 217-247, doi: 10.1177/0013161X08330501.
- Tschannen-Moran M. and Gareis C.R. (2015), "Principals, Trust, and Cultivating Vibrant Schools", *Open Access, Societies*, 5: 256-276, doi: 10.3390/soc5020256.
- Uljens M., Møller J., Årlestig H. and Frederiksen L.F. (2013), *The Professionalisation of Nordic School Leadership*, in Moos L., ed., *Transnational Influences on Values and Practices in Nordic Educational Leadership. Is there a Nordic Model?*, Springer, Dordrecht.

La valutazione delle competenze nei CPIA: un caso studio

di Germana Patti e Giuseppe M. Tomaselli

«L'apprendimento degli adulti è un elemento essenziale del ciclo dell'apprendimento permanente che copre l'intera gamma di attività di apprendimento formale, non formale e informale, sia generale che professionale, intraprese da adulti dopo aver lasciato il ciclo dell'istruzione e della formazione iniziali» (Risoluzione del Consiglio Europeo 2011/c 372/01 pubblicata sulla GUCE del 20 dicembre 2011).

I Centri Provinciali per l'Istruzione degli Adulti (da qui in avanti CPIA), istituiti ai sensi del d.p.r. n. 263/2012, sono il soggetto pubblico di riferimento per la costituzione delle reti territoriali per l'apprendimento permanente e rappresentano un punto di riferimento istituzionale stabile, strutturato e diffuso per coordinare e realizzare – per quanto di competenza – azioni di accoglienza, orientamento e accompagnamento rivolte alla popolazione adulta, e soprattutto ai gruppi svantaggiati.

I Centri hanno la medesima autonomia attribuita alle istituzioni scolastiche, sono dotati di un proprio organico; hanno i medesimi organi collegiali delle istituzioni scolastiche, con gli opportuni adattamenti; sono organizzati in modo da stabilire uno stretto raccordo con le autonomie locali, il mondo del lavoro e delle professioni; realizzano un'offerta formativa.

L'offerta formativa realizzata dai Centri, caratterizzata da un alto grado di flessibilità, è strutturata per livelli di apprendimento in:

- percorsi di alfabetizzazione e apprendimento della lingua italiana finalizzati al conseguimento di un titolo attestante il raggiungimento di un livello di conoscenza della lingua italiana non inferiore al livello A2 del Quadro comune europeo di riferimento per la conoscenza delle lingue, elaborato dal Consiglio d'Europa;
- percorsi di primo livello suddivisi in due periodi didattici; il primo permette di conseguire il titolo di studio conclusivo del primo ciclo; il secondo permette di conseguire la certificazione attestante l'acquisizione delle competenze di base connesse all'obbligo di istruzione relative alle

attività e insegnamenti generali comuni a tutti gli indirizzi degli istituti professionali e degli istituti tecnici.

Il CPIA, in quanto Rete Territoriale di Servizio, svolge non solo le attività di istruzione, ma anche attività di Ricerca, Sperimentazione e Sviluppo (RS&S) in materia di istruzione degli adulti.

Ai Centri possono iscriversi:

- gli adulti, anche stranieri, che non hanno assolto l'obbligo di istruzione o non sono in possesso del titolo di studio conclusivo del primo ciclo di istruzione (i NEET, i drop-out, gli analfabeti di ritorno, gli inoccupati e i disoccupati, così come i soggetti con provvedimenti di cautela personale);
- gli adulti stranieri che vogliono iscriversi ai percorsi di alfabetizzazione e di apprendimento della lingua italiana e alle sessioni di formazione civica e di informazione nonché ai test linguistici e culturali.

Sono ammessi anche studenti minorenni non in possesso del titolo di studio conclusivo del primo ciclo di istruzione, purché abbiano compiuto il sedicesimo anno di età.

Il CPIA prevede che i percorsi di istruzione siano organizzati in modo da consentire la personalizzazione del percorso sulla base di un Patto Formativo Individuale, definito previo riconoscimento iniziale dei saperi e delle competenze formali, informali e non formali posseduti dall'adulto. Il processo che porta al riconoscimento dei crediti è articolato in tre fasi: identificazione, valutazione, attestazione, in coerenza con le indicazioni europee contenute nella Raccomandazione del Consiglio del 20 dicembre 2012 sulla convalida dell'apprendimento non formale e informale. Elementi fondamentali per la personalizzazione dei percorsi sono l'organizzazione per gruppi di livello e la progettazione per unità di apprendimento delle competenze, intese come insieme autonomamente significativo di conoscenze, abilità e competenze correlate ai livelli e ai periodi didattici.

Nell'Allegato 1 delle Linee guida di cui al Decreto Interministeriale 12 marzo 2015 sono declinati i risultati di apprendimento, descritti in competenze, conoscenze e abilità da acquisire al termine del percorso, aggregati nei seguenti assi culturali: asse dei linguaggi, asse storico-sociale, asse matematico e asse scientifico-tecnologico; sono stati elaborati riferendoli al livello 2 del Quadro europeo delle qualifiche, così come definiti dai descrittori dell'allegato II alla Raccomandazione del Parlamento Europeo e del Consiglio del 23 aprile 2008, e tenendo conto anche delle otto competenze chiave per l'apprendimento permanente.

Per il primo periodo didattico, che ha un orario complessivo di 400 ore, sono state individuate 22 competenze da acquisire al termine del percorso (tab. 1).

Tab. 1 - Primo livello - primo periodo didattico nei CPIA.

<i>Assi culturali</i>	<i>Ore</i>
Asse dei linguaggi	198
Asse storico-sociale	66
Asse matematico	66
Asse scientifico-tecnologico	66
Totale	400
di cui Accoglienza	40

Date la novità organizzativa e strutturale del CPIA e le peculiarità del bacino di utenti cui esso si rivolge, nell'a.s. 2016/2017 gli autori, rispettivamente docenti di lettere e matematica presso il CPIA "Alberto Manzi" di Siracusa, hanno pensato di avviare un progetto sperimentale di didattica per competenze che avesse come finalità ultima la strutturazione di una valutazione delle medesime tramite compito autentico esperto.

Si è scelto di attuare la sperimentazione nel punto di erogazione di Lentini, dove gli autori erano in servizio nell'a.s. 2016/2017, e di avviare il progetto all'interno di un gruppo di livello costituito da 16 studenti, tutti italo-foni, cui non erano stati riconosciuti crediti formali, non formali o informali in fase di accoglienza.

Il progetto è stato organizzato in tre fasi, progettazione, realizzazione e rielaborazione, suddivise tramite diagramma di Gantt all'interno delle 400 ore previste dal d.p.r. n. 263 del 2012 per il raggiungimento delle competenze del primo periodo-primo livello.

Nella fase di progettazione si è innanzitutto lavorato ad un ripensamento delle Unità di Apprendimento (UdA) adottate dal CPIA di Siracusa in un'ottica inter- e multidisciplinare: si è infatti individuata una competenza focus, cui correlare altre competenze previste dall'Allegato 1 delle Linee guida di giugno 2015.

La scelta tra le 22 competenze previste dal d.p.r. non è stata particolarmente impervia, in quanto, a nostro avviso, diverse competenze ruotano attorno all'ampio e complesso concetto della trasformazione del territorio, indagandolo sotto molteplici aspetti: analisi della complessità del presente sulla base dei cambiamenti storici, geografici e sociali del passato (competenza 9); effetti dell'azione dell'uomo sui sistemi territoriali (competenza 10); analisi e descrizione dei fenomeni appartenenti alla realtà naturale e artificiale (competenza 17); oppure osservazione delle modifiche degli ecosistemi per cause naturali o antropiche (competenza 19). La decisione è quindi

caduta sulla competenza 10 dell'asse storico-sociale come competenza focus, per diverse ragioni: innanzitutto, perché ad essa potevano essere correlate diverse altre competenze previste dalle Linee guida non immediatamente legate al concetto di trasformazione del territorio, come la competenza 4 dell'asse dei linguaggi sulla tutela e conservazione dei beni del patrimonio artistico e culturale locale; e, in secondo luogo, perché essa ci avrebbe consentito un'osservazione del raggiungimento delle competenze in un contesto reale e una loro valutazione mediante compito autentico esperto.

È nata così l'UdA "Una visita a Catania": trasformazione e cambiamenti nel territorio, un percorso sul tema delle trasformazioni del paesaggio urbano e periurbano in un'ottica fortemente interdisciplinare, inteso come presupposto indispensabile per la lettura, la comprensione e l'interpretazione dei fenomeni di trasformazione degli spazi urbani e periurbani.

Per la costruzione di quest'UdA abbiamo associato prima alla competenza focus, la C10, le competenze ad essa immediatamente collegabili, ossia le già citate C4, C9, C17 e C19, e poi altre competenze previste dalle Linee guida che potevano contribuire al raggiungimento della C10, come ad esempio le competenze C2 e C3 dell'asse dei linguaggi afferenti alla lettura, comprensione e interpretazione, nonché alla stesura di testi scritti, sia alcune tra le 8 competenze chiave di cittadinanza, il cui raggiungimento è previsto anche per i CPIA nella Circolare n. 9 del 3 novembre 2017 relativa agli Esami di Stato nei CPIA .

Selezionate le competenze, abbiamo estrapolato dal monte ore previsto da ogni Dipartimento per singola competenza le ore utili per il raggiungimento degli obiettivi dell'UdA e scelto i contenuti adatti al raggiungimento delle conoscenze e delle abilità previste per ogni singola competenza: questa parte di lavoro è stata la più ardua in quanto non esistono libri scolastici per i CPIA pensati e organizzati secondo UdA inter- e multidisciplinari e i contenuti dei testi scolastici per la scuola secondaria di primo grado si rivolgono a un pubblico di adolescenti e non di adulti e spesso non sono "adatti". Questa iniziale difficoltà ci ha spinto a cercare materiale nella quotidianità (articoli di giornale, brochure turistiche, dati di rapporti regionali e provinciali sui cambiamenti del territorio catanese, racconti su Catania antica e moderna ecc.), ma ci ha anche consentito di mettere in atto una "reale" didattica per competenze basata su esperienze e situazioni concrete e quotidiane: ad esempio la lettura di un articolo di giornale, l'ascolto di una trasmissione televisiva, l'impiego delle operazioni di aritmetica per il calcolo del budget per un'attività, la scelta del registro linguistico adatto per scrivere o per parlare con differenti destinatari, la ricerca consapevole, ragionata e critica su Internet.

A conclusione del monte ore previsto per ogni competenza abbiamo elaborato delle prove di realtà a risposta chiusa, a risposta costruita, svolte

singolarmente o in gruppo in modo tale da mobilitare nel processo di verifica almeno due delle tre dimensioni coinvolte nella manifestazione della competenza, ossia la dimensione oggettiva e quella soggettiva.

La fase di realizzazione dell'UdA si è conclusa con la visita alla città di Catania e in questa fase è stato possibile valutare le competenze nelle varie dimensioni coinvolte in relazione alla loro manifestazione.

La scelta dell'itinerario, il tempo necessario da dedicare alla visita del luogo di interesse scelto, la spiegazione dei luoghi di interesse, dell'attuale organizzazione urbana della città di Catania e il confronto con il passato sono stati totale appannaggio degli alunni della classe sperimentale: questi studenti hanno dovuto fare da ciceroni ad altri studenti del punto di erogazione di Lentini, iscritti al secondo periodo, cercando di valorizzare sul campo il lavoro svolto in aula e di dimostrare la loro competenza sull'argomento, rispondendo a continue domande dei compagni e trovandosi ad affrontare anche situazioni impreviste, come un incontro con dei turisti inglesi.

Durante lo svolgimento della visita i docenti accompagnatori hanno riempito una griglia di osservazione sistematica sul comportamento degli alunni coinvolti, i cui indicatori riguardavano autonomia, partecipazione, responsabilità, flessibilità e consapevolezza.

Durante l'ultima fase dell'UdA, la rielaborazione, gli studenti sono stati invitati a riflettere sull'esperienza svolta mediante la realizzazione di un'autobiografia che ripercorresse il percorso svolto e li portasse a compiere una riflessione metacognitiva sul lavoro effettuato.

Per la valutazione delle prove di verifica e dell'autobiografia abbiamo elaborato delle nostre rubriche valutative, di cui proponiamo in tab. 2 un breve esempio.

Dalle prove di realtà, dalle griglie di osservazione e dalle autobiografie si è potuto concretamente osservare come i risultati raggiunti nel corso della sperimentazioni nel punto di erogazioni di Lentini siano stati positivi per gli alunni coinvolti: essi hanno avuto la possibilità di realizzare pienamente le loro capacità non solo nel modo di comprendere e riferire ma, soprattutto, nel dimostrare, attraverso i comportamenti oggettivi che hanno adottato, quello che effettivamente avevano imparato a fare, a pensare e a dire nel percorso formativo. Alcuni di questi alunni hanno ripreso fiducia nell'istituzione scolastica e con maggiore maturità hanno deciso di continuare gli studi iscrivendosi alla scuola secondaria di secondo grado o in percorsi di formazione professionale. Anche i docenti coinvolti hanno ottenuto un arricchimento della loro formazione, in quanto hanno imparato a progettare un percorso che, partendo dalla valorizzazione dell'apprendimento centrato sull'esperienza, potesse offrire agli alunni gli strumenti per sviluppare competenze chiave utili nella vita, dando senso a ciò che apprendono.

Tab. 2 - Esempio di rubrica valutativa ideata per la sperimentazione.

<i>Dimensioni</i>	<i>Livelli</i>			
	<i>Iniziale</i>	<i>Base</i>	<i>Intermedio</i>	<i>Avanzato</i>
Padronanza risorse cognitive	Utilizza solo se guidato alcune risorse cognitive.	Utilizza le risorse cognitive di base proposte.	Utilizza la maggior parte delle risorse cognitive proposte.	Utilizza con sicurezza e proprietà le risorse cognitive proposte.
Identificazione scopo comunicativo	Mette a fuoco la situazione problematica, se guidato.	Mette a fuoco la situazione problematica in base alle indicazioni date.	Mette a fuoco la situazione problematica in maniera autonoma e prefigura il risultato atteso.	Mette a fuoco in maniera autonoma la situazione problematica e prefigura il risultato atteso, trovando soluzioni originali.
Produzione	Se guidato, produce semplici messaggi.	Produce semplici messaggi sulla base di linee guida.	Produce messaggi complessi e aderenti al contesto, seguendo alcune linee guida.	Produce in maniera autonoma messaggi complessi, originali e coerenti con il contesto e l'interlocutore.
Atteggiamento verso il lavoro	Richiede sollecitazioni continue.	Lavora se inizialmente stimolato.	Lavora in modo autonomo, richiedendo solo qualche conferma.	Lavora in modo autonomo.

Unica criticità che si è evidenziata durante la realizzazione, nonché la predisposizione del progetto, è stata una certa fragilità di alcuni docenti sulle tematiche della didattica per competenze, della valutazione delle competenze tramite rubriche valutative e di prestazione: insegnare nei CPIA è una sfida che molti docenti sottovalutano, ritenendo che insegnare agli adulti problematici sia come insegnare a degli adolescenti, e applicano delle metodologie poco funzionali e inadatte.

Resta da dire che il progetto non si sarebbe avviato, né avrebbe ottenuto tali risultati se fosse mancato il lucido e continuo supporto del dirigente scolastico del CPIA di Siracusa, prof. Simonetta Arnone, e della vicepresidente, prof. Grazia Rita Fisicaro, nonché dello staff e dei docenti tutti, e, soprattutto, del prof. Cristiano Corsini, il cui supporto è stato fondamentale per la stesura delle rubriche di valutazione e delle griglie di prestazione.

L'esperienza di tirocinio esterno degli studenti fuoricorso in Scienze pedagogiche dell'Università di Catania. Comparazione con studenti regolari e strategie di miglioramento

di *Giuseppe C. Pillera*

1. Contesto e strumento della ricerca

La Raccomandazione del Consiglio dell'UE su un Quadro di Qualità per i Tirocini del 10 marzo 2014, a partire dal riconoscimento della «esistenza di un rapporto tra la qualità del tirocinio e gli esiti occupazionali» (*considerata*, punto 5), si preoccupa di suggerire, mediante un articolato in 28 punti, standard di qualità comunitari per favorire un'amalgama delle legislazioni degli stati membri e far fronte a una serie di rischi e problematiche che, allora come oggi, rischiano di minare l'utilità del tirocinio per i giovani in formazione, come ad esempio la sua utilizzazione, anche ripetuta, in sostituzione di rapporti di lavoro, contenuti di apprendimento scadenti e basso profilo delle mansioni assegnate ai tirocinanti. Appena due anni dopo, nel Documento di lavoro sull'attuazione del Quadro di qualità per i tirocini del 4 ottobre 2016, la Commissione europea esprime un giudizio lusinghiero sull'impianto normativo italiano: non solo ammette che le "Linee-guida in materia di tirocini" frutto dell'Accordo Stato-Regioni del 24 gennaio 2013 (poi aggiornate il 25 maggio 2017) risultavano già largamente conformi al sopra citato Quadro di qualità, ma addirittura ne riconosce lo status di modello per la Raccomandazione del 2014. È da rilevare che nella legislazione italiana in materia di tirocinio la presenza del soggetto promotore è un principio non solo qualificante ma addirittura inderogabile dell'istituto¹, che così si differenzia

¹ Iuzzolino et al. (2018) sottolineano proprio «la divaricazione concettuale tra la Commissione (e con essa la gran parte dei Paesi europei) e l'Italia su un aspetto – la presenza del soggetto promotore – di assoluto rilievo. Mentre l'assunto della Commissione è che l'assenza del soggetto promotore qualifichi una specifica tipologia di tirocini (quelli nel libero mercato), il presupposto su cui poggia l'impianto legislativo italiano è che il tirocinio in quanto tale debba invece necessariamente prevedere la presenza di un soggetto promotore: la sua assenza configurerebbe un rapporto non qualificabile come tirocinio. Due prospettive diverse che probabilmente sottendono due modi differenti di intendere il concetto stesso di tirocinio» (p. 35).

nettamente da altre tipologie di ingresso nel mondo del lavoro caratterizzate da scopi formativi, come il contratto di apprendistato.

Nel nostro Paese, il tirocinio formativo per gli studenti universitari è incoraggiato a partire dal Decreto MIUR n. 270 del 22 ottobre 2004, che all'art. 10 c. 5-d recita «i corsi di studio dovranno prevedere [...] attività formative, non previste dalle lettere precedenti, volte ad acquisire ulteriori conoscenze linguistiche, nonché abilità informatiche e telematiche, relazionali, o comunque utili per l'inserimento nel mondo del lavoro, nonché attività formative volte ad agevolare le scelte professionali, mediante la conoscenza diretta del settore lavorativo cui il titolo di studio può dare accesso, tra cui, in particolare, i tirocini formativi e di orientamento di cui al decreto 25 marzo 1998, n. 142, del Ministero del lavoro». Nel solco del rinnovamento della didattica universitaria tracciato a livello europeo già a partire dal *processo di Bologna*, il tirocinio formativo, insieme ai laboratori, ha guadagnato spazio nei curricula dei corsi di studio e attenzione nel panorama della ricerca sul campo, rappresentando per gli studenti universitari una preziosa opportunità di arricchimento, integrazione e sperimentazione del bagaglio di conoscenze e competenze sviluppate durante il percorso di studi e costituendosi come «momento persino decisivo perché i giovani possano stabilire un ponte efficace fra la teoria e la pratica, fra le esigenze di verifica e quelle di autoverifica e di auto orientamento, fra dimensioni culturali e competenze operative del proprio profilo professionale» (Paparella e Perrucca, 2006, p. 9-10).

Il tirocinio, spostatosi progressivamente al centro delle politiche attive del lavoro nazionali e internazionali e incorporato nei curricula universitari, ha attirato su di sé un rinnovato interesse degli studi di area psico-pedagogica e didattica, che hanno in questi anni focalizzato l'attenzione sui suoi risvolti non solo nell'ambito della formazione docenti, dove una quasi ininterrotta tradizione di riflessione affonda le radici nei primi anni del XX secolo² e giunge fino ai giorni nostri (Bondioli et al., 2006; Maccario, 2014), e non tanto in termini trasmissivo-conoscitivi, quanto autoriflessivi e orientativi, progettuali e motivazionali (Paparella e Perucca, 2006; Perucca, 2005; Salerno, 2007). Tali declinazioni di analisi risultano particolarmente rilevanti per le esperienze di tirocinio di allieve e allievi che si preparano a fare il loro ingresso in mondi professionali incentrati sulla relazione umana e, in

² Nonostante il fallimento politico, dovuto allo scoppio della I guerra mondiale e al conseguente mutare in senso autocratico del clima politico italiano, il progetto portato avanti dal Ministro dell'Istruzione Luigi Credaro tra il 1911 e il 1914 «rappresentò un serio tentativo di porre mano ad una riforma globale dell'istruzione magistrale e servì a delineare e chiarire i punti nodali sui quali era ormai ineluttabile intervenire» (Di Pol, 2014), tra cui appunto, una riforma del tirocinio dei maestri nella durata e nella sostanza. Riguardo alle implicazioni della riforma dei Regi Ginnasi Magistrali voluta da Credaro e alle conseguenti sperimentazioni portate avanti in quegli anni in Sicilia, si veda anche Tomarchio (2008).

particolare, sulla relazione di cura educativa (Bartolini e Riccardini, 2006; Bastianoni e Spaggiari, 2014), dove l'applicazione delle conoscenze acquisite in via teorica, per diventare sapere (saper fare e saper essere) è costretta a confrontarsi, aggiornarsi e ripensarsi quotidianamente nell'incontro con l'estrema complessità, variabilità e problematicità dell'ambiente e del materiale umano *per cui e insieme* a cui si lavora. Inoltre, non ci appare di secondaria importanza il fatto che, preparandosi ad affrontare simili contesti professionali, gli allievi possano trarre grande giovamento da un tempo e da uno spazio, quello del tirocinio, in cui l'esperienza professionale trova occasione di travaso per così dire non artificiale, bensì all'interno del *terreno* stesso culturale/culturale della pratica educativa, a confronto e sotto la guida di professionisti che, in funzione di tutor, rivestono un ruolo cruciale.

Essendo il tirocinio universitario inserito in un continuum esperienziale che vuole e suole definirsi formativo, molte delle ricerche accademiche in questo campo si preoccupano di mettere a punto – prevalentemente in maniera inferenziale, mediante la rilevazione e l'analisi dell'esperienza degli studenti tirocinanti – indicatori di qualità dell'esperienza stessa (Perucca, 2005; Salerno, 2007). Così, nell'a.a. 2013/2014, il Corso di laurea magistrale (CdLM) in Scienze pedagogiche e progettazione educativa (LM-85) dell'Università di Catania, avendo valutato la necessità di mettere a punto procedure di monitoraggio del tirocinio esterno³, ha affidato al *Laboratorio di ricerca educativa*, allora curato da Cristiano Corsini, la costruzione di uno strumento finalizzato a raccogliere le opinioni degli studenti sull'attività svolta. La rilevazione dei punti di forza e di debolezza delle varie dimensioni dell'esperienza dei suoi tirocinanti non ha trascurato quelle maggiormente dipendenti dalle scelte più direttamente riferite al progetto formativo del CdLM, senza timore di metterne in discussione anche prassi da tempo consolidate: il questionario, infatti, coerentemente con gli scopi cui obbediva, chiedeva non solo di valutare il lavoro e i contesti del tirocinio, ma anche di indicare se e come lo studio universitario desse significato alla formazione sul campo e viceversa, tenendo presente la relazione dialettica tra le due esperienze.

Lo strumento è stato sviluppato dalle studentesse e studenti che frequentavano i corsi di Pedagogia sperimentale (Corso triennale L-19) e di Metodi e tecniche di valutazione (Corso di LM-85), adattando quello ideato da Anna Salerno (2007) presso l'Università "La Sapienza" di Roma. Gli studenti universitari hanno analizzato lo strumento utilizzato a Roma, proponendo, sulla base della loro precedente esperienza formativa, molte delle modifiche

³ Nel CdLM analizzato in questo contributo, il tirocinio, che può essere attivato dagli studenti già a partire dal primo anno di corso, è obbligatorio e suddiviso in almeno 25 ore interne (seminari di approfondimento con docenti e incontri con professionisti presso le sedi accademiche) e minimo 100 ore da svolgersi presso enti convenzionati con l'ateneo.

apportate nella versione attuale, collegata alla pagina web del corso di laurea magistrale (www.disfor.unict.it/it/corsi/lm-85/tirocini-e-stage). La scelta di coinvolgere attivamente gli studenti nella costruzione e validazione del questionario è stata dettata da due ragioni: in primo luogo, si è reputato che gli studenti rappresentino propriamente i testimoni privilegiati, in grado di assicurare allo strumento validità e aderenza al contesto specifico; in secondo luogo, il lavoro di ricerca necessario per costruire e validare lo strumento ha costituito un'attività tematica basata su compiti, secondo approcci legati alla didattica per competenze e alla valutazione autentica, rappresentando per i partecipanti al laboratorio un'occasione per migliorare conoscenze specifiche sulla ricerca educativa, applicate in contesti veramente sfidanti come risorse per risolvere problemi e non come nozioni da imparare con l'obiettivo di superare un esame.

I 114 item del questionario (di cui 5 a risposta aperta) sono organizzati in cinque sezioni (Scierri, 2017, pp. 19-36):

- 1) informazioni generali sul tirocinante;
- 2) motivazioni personali e vissuti relativi all'esperienza di tirocinio;
- 3) conoscenze e competenze, sia pregresse e richieste che sviluppate;
- 4) relazioni all'interno del contesto lavorativo;
- 5) rapporti tra il tirocinio e lo studio universitario.

Attraverso questo strumento gli studenti possono valutare: loro stessi, cioè le competenze che hanno acquisito, i problemi emersi, il valore complessivo dell'esperienza e la sua utilità per le future scelte lavorative; l'ente ospitante, e dunque le attività assegnate, la formazione ricevuta, il supporto offerto dal tutor aziendale e il tipo di relazioni stabilite; l'Università, ovvero l'organizzazione del tirocinio e il supporto ricevuto, nonché la coerenza tra i contenuti del tirocinio e il curriculum di studi (Salerni, 2007).

In un quadro generalmente positivo, le risposte fornite dagli studenti, dall'a.a. 2013/2014 all'a.a. 2016/2017, presentano tuttavia alcuni nodi critici, trasversali agli enti specifici e ai settori di attività, tra i quali si evidenziano alcune dinamiche che sostanziano da un lato la relazione tra operatori degli enti e studenti, dall'altro, soprattutto, quella tra insegnamento universitario ed esperienza di tirocinio: in particolare alcune difficoltà nel rapporto tra ciò che viene acquisito in aula e l'esperienza nei contesti pratici caratterizza alcune delle risposte degli intervistati. L'analisi dei dati, condotta attraverso l'utilizzo di tecniche qualitative e quantitative e pubblicata nel volume collettaneo curato da Corsini e Strongoli (2017), può essere sintetizzata in tre punti, che rappresentano le principali questioni emergenti:

- contenuti e qualità del curriculum accademico, con specifico rimando alla natura della didattica universitaria reputata troppo teorica, anche laddove essa offre attività laboratoriali;
- continuità e coerenza delle attività di tirocinio con il percorso accademico, con un richiamo ad una maggiore oculatezza nella selezione degli enti autorizzati ad ospitare i tirocinanti;
- clima generale dell'ambiente di lavoro, preparazione professionale dei lavoratori dell'ente ospitante e qualità della loro relazione con i tirocinanti, con particolare riferimento alla figura del tutor, il cui apporto assume un ruolo chiave nella valutazione degli studenti rispetto agli esiti formativi del tirocinio.

2. Scopi, metodologie e strumenti di analisi

Nonostante i progressi compiuti negli ultimi decenni per ridurre l'eccessiva durata degli studi, il tasso di studenti fuoricorso (ossia gli iscritti da un numero di anni superiore alla durata regolare del corso) rimane un problema significativo del nostro sistema universitario. Secondo i dati Anvur riportati dal Nucleo di valutazione dell'Università di Catania (2018, p. 14), nell'a.a. 2016/2017 erano il 27% a livello nazionale e il 32% nel Sud del Paese. Per lo stesso a.a., con riguardo al CdLM oggetto di questo contributo (LM-85), il Nucleo di valutazione d'ateneo (2017, pp. 28-35) riporta appena il 30,6% di laureati entro la durata regolare del corso (in linea, tuttavia, con la media di ateneo relativa alle lauree vecchio ordinamento, magistrali e a ciclo unico, pari al 31%) e il 59,2% di immatricolati che si laurea entro un anno oltre la durata regolare (sensibilmente meno della media di ateneo del 69,1%), sebbene vada tenuto in debito conto che nel corso della carriera universitaria in *Scienze pedagogiche* quasi tre studenti su quattro hanno avuto esperienze di lavoro in larga parte (44,4%) coerenti col settore di studi (Alma Laurea, 2019).

Richiamandosi alla teoria della *social integration* di Tinto (1988, 1993), che identifica tra i fattori in gioco nel problema della dispersione non solo l'influenza del contesto universitario ma anche quella dell'ambiente esterno su un processo longitudinale di interazione tripolare individuo, sistema accademico e sistema sociale, Pastore e altri (2015) evidenziano come «Esperienze positive, e quindi integrative, rafforzano la persistenza nel sistema universitario attraverso il loro impatto su intenzioni e impegno sia nei confronti dell'obiettivo della laurea (*goal commitment*), sia nei confronti dell'istituzione (*institutional commitment*)» (p. 343). In questo contributo, dunque, cercheremo di mettere in luce le componenti chiave per un'esperienza di tirocinio soddisfacente, le difficoltà, le criticità e i limiti riscontrati

in maniera peculiare dai 48 studenti fuoricorso coinvolti nella ricerca (44 femmine e 4 maschi), verificando o smentendo i risultati emersi per la totalità dei 110 tirocinanti che hanno partecipato alla rilevazione. Di conseguenza, i dati dell'indagine esposti nel volume sopra citato sono stati rianalizzati isolando gli studenti fuoricorso, così da riflettere soprattutto sulle somiglianze e differenze tra la loro opinione e quella degli studenti in regola (62 studenti, di cui 60 femmine) con i tempi di completamento previsti per il CdLM.

Le metodologie di analisi utilizzate sono miste: nel quadro della comparazione tra i due gruppi di studenti, ci avvarremo della statistica descrittiva per le domande chiuse e, per le domande aperte, della categorizzazione, oltre ad alcune tecniche di *text data mining* già da noi utilizzate, sia in seno al volume sul tirocinio sopra citato, sempre per l'analisi delle domande aperte del questionario (Pillera, 2017b, pp. 101-119), sia nello studio di altri argomenti e contesti (Pillera, 2020). Per elaborare le risposte alle domande aperte prese in considerazione (quelle con un numero non troppo esiguo di rispondenti, tale da garantire una quantità sufficiente di testo), abbiamo utilizzato il software di analisi testuale KH Coder (<http://khcoder.net>), creato da Koichi Higuchi (2015) presso la Ritsumeikan University di Kyoto (Giappone) e distribuito sotto licenza open source per i principali sistemi operativi (Linux, Windows, Apple OS). Il programma è in grado di estrarre i morfemi alla base di ogni parola⁴, catalogarli in un database MySQL, quindi restituire misurazioni, eseguire calcoli ed elaborare i risultati in forma grafica.

Dopo aver sottoposto il corpus di risposte aperte a un semplice pretrattamento⁵, abbiamo utilizzato diverse tecniche – dalla semplice analisi delle occorrenze (f), alla *keyword in context analysis*, per focalizzare e contestualizzare concetti particolarmente ambigui o significativi – concentrandoci, in particolare, sull'analisi delle co-occorrenze mediante diagrammi di rete, che consentono di visualizzare l'articolazione della rete semantica del discorso,

⁴Il metodo di lemmatizzazione utilizzato è chiamato “Stemming with snowball”, purtroppo l'unico disponibile in KH Coder per lingue diverse dal giapponese o dall'inglese (per le quali è possibile utilizzare lo “Stanford POS Tagger”, un metodo coerente con la grammatica di quelle lingue): «Stemming is a process that cuts ends of words, according to simple rules. For example, it can extract ‘says’ or ‘saying’ as the original form ‘say’, but extracts ‘said’ as just ‘said’. [...] it has a faster processing time and its extraction is accurate enough to analyze semantic content» (Higuchi, 2015, pp. 11-12). In questo lavoro, segnaleremo i morfemi esattamente come sono stati estratti dal software (di solito senza desinenze finali e suffissi), distinguendoli con l'utilizzo del carattere corsivo (viceversa nelle didascalie).

⁵Il pretrattamento del testo è stato limitato a pochi semplici interventi: gli errori grammaticali sono stati corretti, filtrate le cosiddette parole vuote (articoli, preposizioni semplici e articolate, congiunzioni, pronomi, aggettivi possessivi e dimostrativi, forme di essere e di avere), alcuni nomi e verbi, presenti nella domanda, anch'essi filtrati quando compaiono regolarmente come formula di risposta rituale.

ovvero quali sono i campi di significato più comunemente condivisi e come sono associati.

I diagrammi di rete – utilizzati con successo sin dai primi passi dell’analisi del contenuto (Osgood, 1959) e fino a tempi più recenti (Danowski, 1993) – chiariscono la struttura delle co-occorrenze rispetto ad altre tecniche, come il *multi-dimensional scaling* o la *cluster analysis*⁶, e possono essere interpretati come una «mappatura delle isotopie» (Lancia, 2016), da intendersi come temi generali o specifici caratterizzati dalla ricorrenza di tratti semantici (Rastier et al., 2002, p. 204)⁷. L’identificazione delle isotopie, come risultato non di un mero conteggio ma di un processo interpretativo (Rastier, 1987), facilita l’interpretazione del discorso, poiché ognuna di esse mette a fuoco un contesto condiviso da più parole. In questo modo un diagramma di rete delle co-occorrenze, fornendo una sorta di mappa concettuale ponderata di un corpus testuale, «consente di ricostruire ‘un filo’ del discorso all’interno della trama complessiva costituita dal corpus o da un suo sottoinsieme» (Lancia, 2020, pp. 90-91).

3. Analisi e discussione dei dati

3.1. Motivazioni personali e vissuto dell’esperienza di tirocinio

Le due domande sui motivi che hanno ispirato la scelta di un settore e di un ente in cui svolgere il tirocinio sono precedute da altrettanti item dicotomici che invitano a chiarire la possibilità di libera scelta sia del settore che dell’organizzazione ospitante. In questa sezione, nonostante tendenze tutto sommato sovrapponibili, si può sottolineare come gli studenti fuoricorso dichiarino più ampie possibilità di scegliere il percorso di tirocinio,

⁶Per determinare la localizzazione dei nodi nel diagramma, KH Coder applica un metodo sviluppato da Fruchterman & Reingold (1991) e orientato alla massima chiarezza. Tuttavia, a differenza che in un *multi-dimensional scaling*, due nodi mostrati vicini non sono co-occorrenti se non sono collegati da una linea (*edge* o arco), la cui intensità (da punteggiato a grassetto) indica la frequenza della co-occorrenza di quei dati morfemi. La frequenza di una co-occorrenza esprime il numero di volte in cui due o più unità lessicali (nel nostro caso morfemi) sono contemporaneamente presenti all’interno della porzione di testo presa come unità di analisi (nel nostro caso le risposte degli studenti a una data domanda aperta).

⁷La nozione di isotopia (dal greco antico: stesso posto) – introdotta in linguistica dal semiologo Greimas (1966) per descrivere la ricorrenza di più parole con tratti semantici comuni all’interno delle unità sintagmatiche (frasi o testi) – si riferisce a una concezione gestaltica del significato: esso non risiede tanto nelle singole parole quanto nelle loro relazioni, secondo il principio che un insieme è qualcosa di diverso e superiore rispetto alla semplice somma delle sue parti.

specialmente per quanto riguarda il settore di attività (96% contro l'87% degli studenti regolari). Circa le motivazioni della scelta di un determinato campo per l'esperienza (fig. 1), sottolineiamo che gli studenti fuoricorso mostrano uno scetticismo lievemente inferiore e un pragmatismo leggermente maggiore, rivelando (in maniera solo apparentemente contraddittoria): da un lato, ragioni più profonde, legate a interesse e passione, insieme a una più forte speranza di lavorare nello stesso ramo dell'ente scelto; dall'altro, minore disponibilità a mettersi alla prova e, coerentemente con una scelta che appare piuttosto informata, minore interesse nell'osservare e conoscere un settore che, appunto, sembrano scegliere in maniera più consapevole già a monte.

In merito alla coppia di item riguardanti la selezione dell'ente ospitante, restituiamo una sostanziale conferma dei dati precedenti, con una specificazione sul ruolo che tutor e docenti, da un lato, e altri studenti, dall'altro, giocano riguardo a tale scelta: come ci si potrebbe aspettare, infatti, gli studenti regolari sono moderatamente influenzati da altri studenti (5% contro 0,9% degli studenti fuoricorso), mentre gli studenti fuoricorso sono lievemente influenzati da tutor e docenti (3,7% contro lo 0% degli studenti regolari).

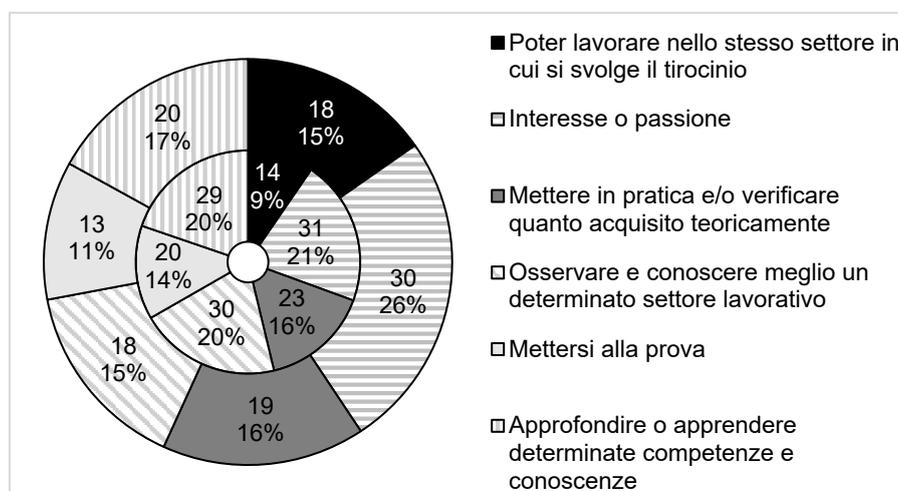


Fig. 1 – Motivazioni della scelta del settore di lavoro per il tirocinio: confronto tra studenti fuoricorso (cerchio esterno) e studenti regolari (cerchio interno). Il totale delle risposte per ciascun gruppo è superiore al numero dei rispondenti perché era possibile fornire più di una risposta.

Se nei dati appena esposti, e ancor meno nel successivo item sulle motivazioni generali, le risposte non divergono eccessivamente tra i due gruppi del campione, la domanda riguardante il momento del percorso universitario in cui inizia l'interesse per il tirocinio presenta invece distribuzioni delle

risposte notevolmente differenziate (fig. 2): il 78,4% degli studenti non tradizionali è interessato al tirocinio solo dopo aver completato quasi tutti gli esami, contro il 26,2% degli studenti regolari. Questo item, inoltre, conferma la scarsa influenza dell'opinione dei colleghi sugli studenti fuoricorso ma anche quella dei corsi frequentati all'università.

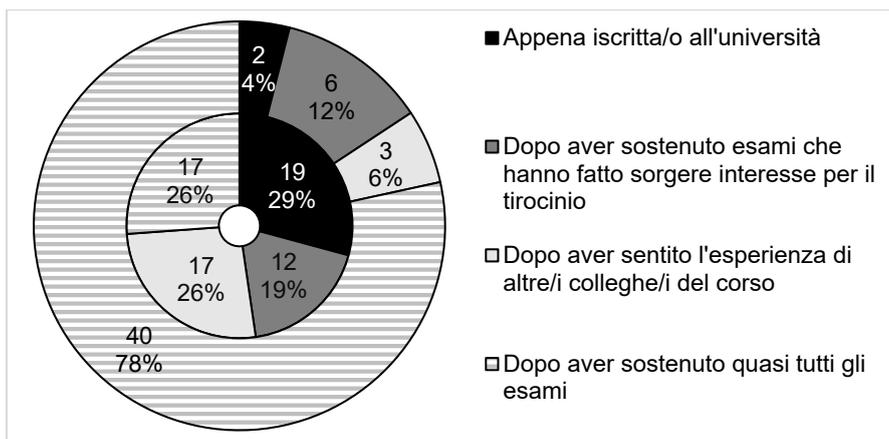


Fig. 2 – Individuazione del momento del percorso universitario in cui gli studenti iniziano a interessarsi al tirocinio: confronto tra studenti fuoricorso (cerchio esterno) e studenti regolari (cerchio interno). Il totale delle frequenze per ciascun gruppo è superiore al numero dei rispondenti perché era possibile fornire più di una risposta.

Dall'articolata sezione del questionario dedicata alla descrizione dell'ambiente di lavoro si riportano in fig. 3 gli item con distribuzioni maggiormente diversificate tra i due gruppi sotto esame. Innanzitutto, gli studenti fuoricorso tendono a valutare gli spazi di lavoro in modo nettamente più positivo; in secondo luogo, soffrono in misura minore della sensazione di non essere tenuti in considerazione e dichiarano in misura maggiore di svolgere compiti appropriati al proprio ruolo; di conseguenza, avvertono un minore senso di inadeguatezza e si sentono maggiormente a proprio agio. La valutazione generalmente più positiva sull'esperienza di tirocinio mostrata dai fuoricorso è confermata anche per quanto concerne: la serenità sul luogo di lavoro, l'adeguatezza degli orari in relazione alle proprie esigenze e degli spazi rispetto ai compiti richiesti, l'adeguatezza dell'ente ospitante con riguardo allo svolgimento dei compiti assegnati e quella del personale rispetto all'accoglienza dei tirocinanti.

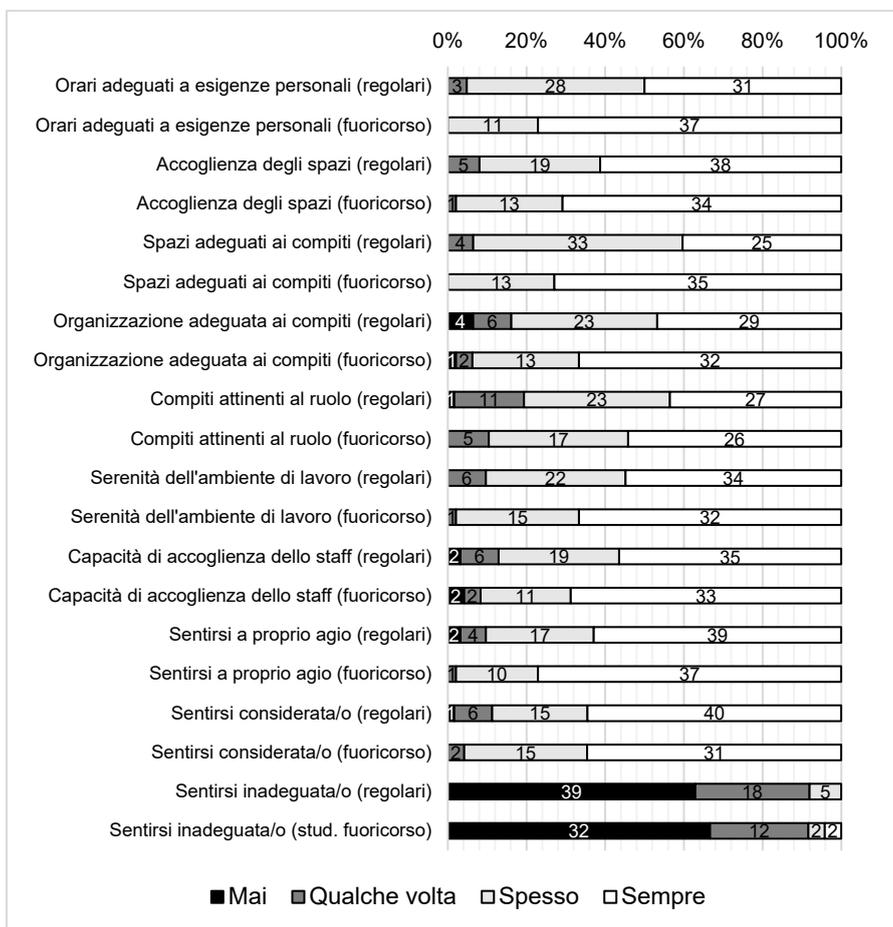


Fig. 3 – Item relativi al vissuto del tirocinio: confronto degli item con distribuzioni percentuali maggiormente differenziate tra le risposte degli studenti regolari e di quelli fuoricorso (con indicazione dei valori assoluti per ciascuna opzione di risposta).

3.2 Relazioni sul luogo di lavoro

Come previsto e confermato da entrambi i gruppi di rispondenti, l'interfaccia principale tra loro stessi e l'organizzazione ospitante è rappresentata dal tutor aziendale in oltre otto casi su dieci, sebbene circa il 35% del campione affermi di essere assistito anche da altri membri del personale. Il rapporto con il tutor è largamente descritto come abbastanza o molto continuativo, amichevole, informale, pacifico, costruttivo e stimolante. Ciononostante, registriamo alcune differenze tra studenti regolari e fuoricorso, come mostrato in fig. 4,

dove si può notare che questi ultimi appaiono inclini a considerare la relazione con il tutor pacifica, costruttiva, stimolante e costante in una percentuale più elevata e, viceversa, conflittuale, non costruttiva, demotivante ed episodica in misura minore rispetto agli studenti regolari.

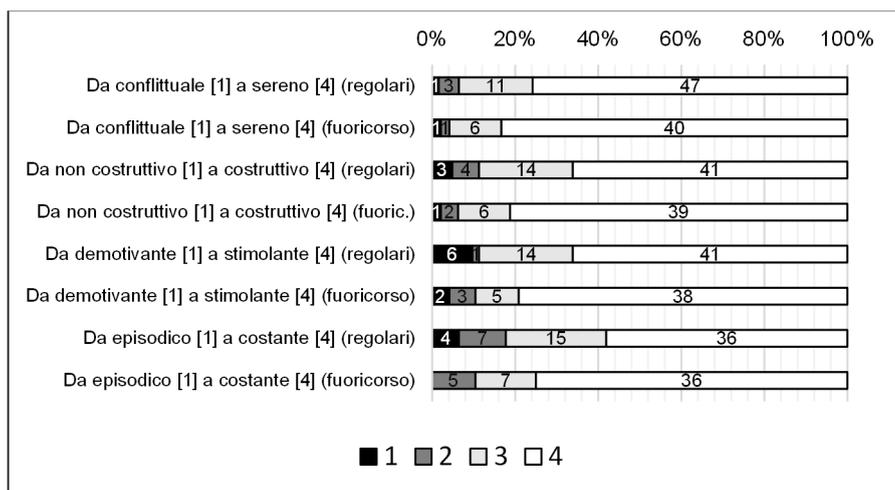


Fig. 4 – Valutazione del rapporto con il tutor aziendale confronto degli item con distribuzioni percentuali maggiormente differenziate tra le risposte degli studenti regolari e di quelli fuoricorso (con indicazione dei valori assoluti per ciascuna opzione di risposta).

Il contesto di positività sopra menzionato è confermato dalle risposte alle domande aperte sui problemi incontrati con il resto del personale e con gli utenti: i rapporti sul posto di lavoro appaiono basati su una collaborazione serena e fruttuosa – che viene spesso descritta (in particolare riferendosi al tutor) come dotata di un alto valore umano e professionale – e i problemi sembrano essere molto limitati e circoscritti a casi specifici per entrambi i gruppi di studenti. Tuttavia, citiamo in tab. 1 alcune critiche evidenziate dai rispondenti, che, emergendo a macchia di leopardo anche in altri punti del questionario, potrebbero essere riassunte nei seguenti tre punti chiave:

1. richiesta di svolgere compiti poco coerenti o del tutto irrilevanti per lo stage (per entrambi i gruppi);
2. incompatibilità di carattere tra studente e tutor (solo per studenti regolari);
3. difficoltà nel contatto con gli utenti (per entrambi i gruppi).

Tab. 1 - Problemi con tutor e utenti evidenziati dai rispondenti.

	Studenti regolari	Studenti fuoricorso
Problemi con il tutor	Assegnazione di compiti o richiesta di svolgere attività poco o per niente rilevanti rispetto al ruolo assunto come tirocinante.	
	Completa ignoranza dei bisogni e delle strategie educative.	
	Costanti lamentele sul lavoro svolto dal tirocinante.	
	Incompatibilità di carattere.	
Problemi con gli utenti	Umore del tutor dipendente da problemi familiari e personali e conseguente relazione conflittuale con i tirocinanti.	
	Alcuni problemi di "primo contatto" con una parte di utenti socialmente ed economicamente svantaggiati.	Ambiente "chiuso", difficile terreno di relazione.

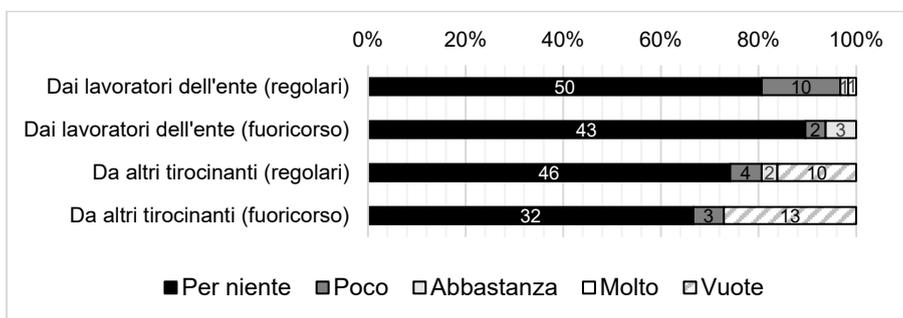


Fig. 5 – Impresione di essere percepite/i come rivali dai lavoratori dell'ente e da altri tirocinanti: confronto distribuzioni percentuali nelle risposte di studenti regolari e fuoricorso (con indicazione dei valori assoluti per ciascuna opzione di risposta).

Sempre per quanto riguarda l'aspetto relazionale, notiamo lievi differenze tra i due gruppi di studenti nella sfumata sensazione di essere percepiti come rivali dai lavoratori dell'ente e dagli altri tirocinanti nello stesso ente (fig. 5): nel primo caso i fuoricorso soffrono in misura minore tale disagio (circa uno su dieci, contro circa il doppio degli studenti regolari); nel secondo caso, invece, la percentuale di disagio è simile tra i due gruppi (circa uno su dieci), tuttavia l'item presenta un elevato tasso di reticenza, specialmente tra i fuoricorso.

3.3 Progetto formativo e attività di tirocinio

In generale, il tirocinante sembra essere impegnato nello sviluppo di abilità professionali principalmente legate all'organizzazione del proprio lavoro (primariamente un lavoro di squadra), con attività di osservazione, pianificazione e coordinamento (presumibilmente in ruoli di assistente) praticate in una significativa quantità di casi (Pillera, 2017a, p. 77). Senza pretendere di derivare dall'analisi testuale conclusioni incontrovertibili su analogie e distinzioni nette tra i due gruppi di studenti circa il concreto svolgimento delle attività, limitiamo i nostri commenti a un'interpretazione ermeneutica di ciò che è sottolineato più frequentemente dagli studenti fuoricorso, vale a dire di ciò a cui questo gruppo sembra attribuire maggiore importanza.

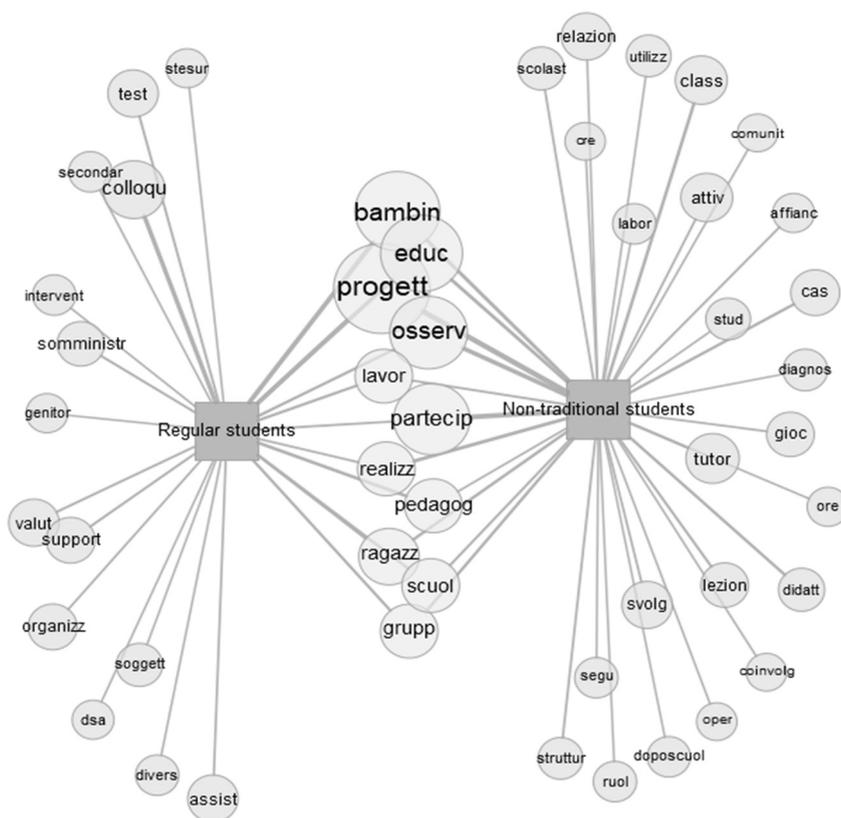


Fig. 6 – Diagramma di rete (words-to-variable) delle prime 60 co-occorrenze dei morfemi con $f \geq 3$ (esclusi attivit, svolt, tirocin e alcuni aggettivi e avverbi modali) nelle risposte alla domanda aperta sulla descrizione delle attività del tirocinio, incrociate con lo status di studente regolare (regular) o fuoricorso (non-traditional).

I morfemi nel diagramma di rete delle associazioni *words-to-variable* in fig. 6 identificano alcune aree di attività comuni (*cluster* interni) e altre differenti (*cluster* esterni): interpretiamo e riprendiamo quelle più significative in tab. 2, concentrandoci su contenuti, contesti e utenti delle attività di tirocinio. In fig. 7, il diagramma di rete delle co-occorrenze *words-to-words* si riferisce invece alle sole risposte degli studenti fuoricorso, mappando le relazioni più strette tra i concetti più frequentemente espressi.

Tab. 2 - Analisi della rete delle co-occorrenze sulla base del diagramma di rete delle co-occorrenze in fig. 6.

Attività tirocinio	Studenti regolari	Studenti fuoricorso
Contenuti e metodologie in comune	Progettazione e realizzazione di attività (<i>progett, realizz</i>); osservazione e osservazione partecipante (<i>partecip, osserv</i>); supporto e affiancamento (<i>assist, support, affianc</i>); lavoro e attività di gruppo (<i>grupp</i>); pedagogia e educazione (<i>pedagog, educ</i>).	
Contenuti e metodologie differenti	Somministrare test, condurre colloqui e valutare le risposte (<i>test, somministr, colloqu, valut</i>); organizzazione e programmazione di interventi (<i>organizz, intervent, stesur</i>); disturbi specifici dell'Apprendimento (DSA).	Aspetti relazionali (<i>relazion</i>); supporto ai docenti (<i>lezion, didatt</i>); laboratorio (<i>labor</i>); gioco e creatività (<i>gioc, crea</i>); affiancamento (<i>affianc</i>); coinvolgimento (<i>coinvolg</i>); caso di studio e diagnosi funzionale (<i>stud, cas, diagnos</i>).
Contesti in comune	Scuola (<i>scuol</i>).	
Contesti differenti	Scuola secondaria (<i>secondar</i>).	Comunità (<i>comunit</i>); classe scolastica (<i>class, scolast</i>); doposcuola (<i>doposcuol</i>).
Utenti in comune	Bambini e ragazzi (<i>bambin, ragazz</i>).	
Utenti differenti	Genitori (<i>genitor</i>); persone disabili (<i>divers</i>).	Ludopatici (<i>gioc</i>).

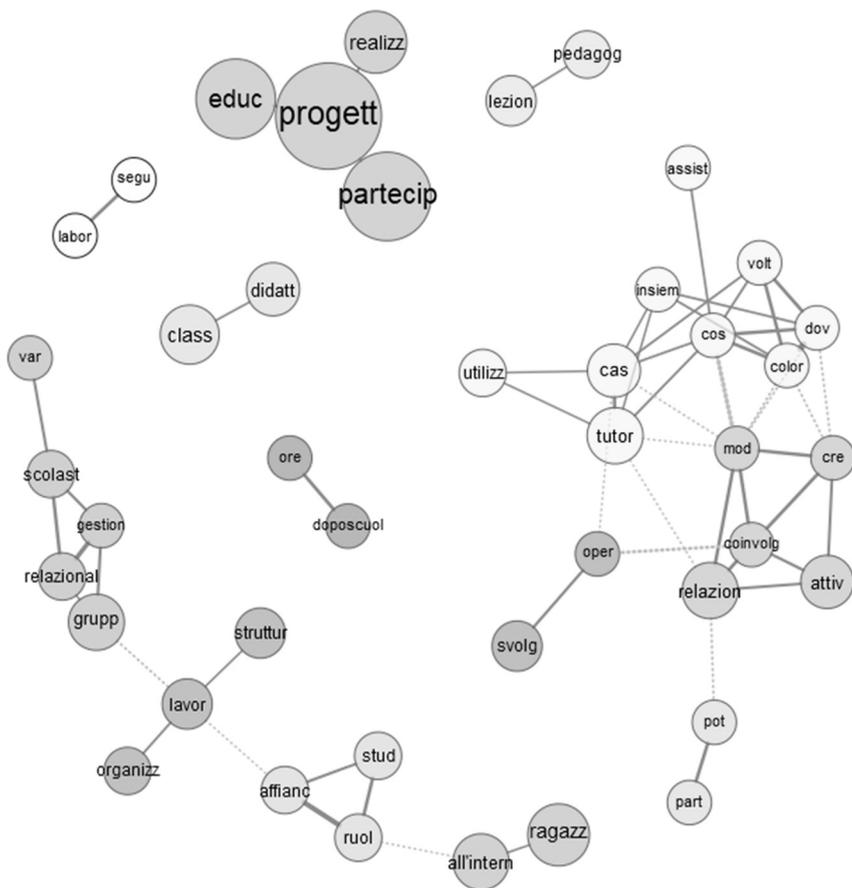


Fig. 7 – Diagramma di rete (words-to-words) delle prime 60 co-occorrenze dei morfemi con $f \geq 3$ (esclusi attivit, svolt, tirocin e alcuni aggettivi e avverbi modali) nelle risposte degli studenti fuoricorso alla domanda aperta sulla descrizione delle attività del tirocinio.

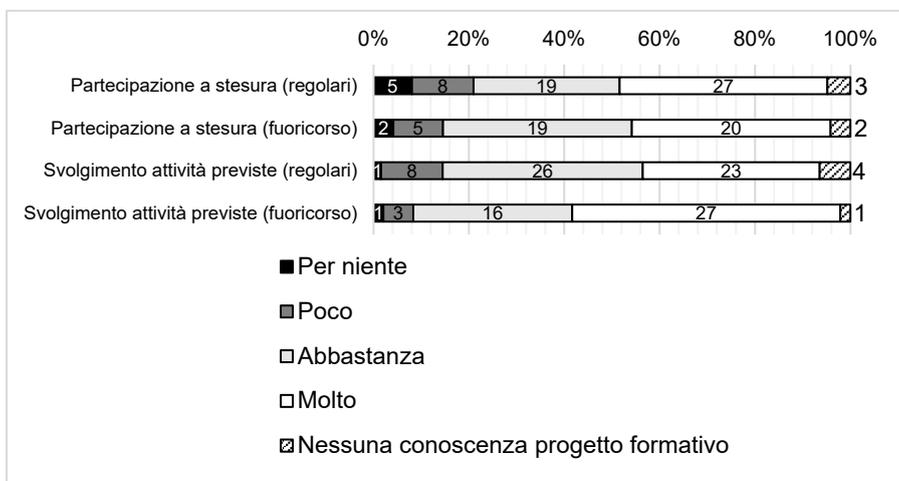


Fig. 8 – Il progetto formativo del tirocinio: partecipazione attiva alla stesura e svolgimento effettivo delle attività previste. Confronto distribuzioni percentuali nelle risposte di studenti regolari e fuoricorso (con indicazione dei valori assoluti per ciascuna opzione di risposta).

Le attività svolte durante il tirocinio procedono (o dovrebbero procedere) sulla scorta del relativo progetto formativo. A questo proposito, segnaliamo che i fuoricorso si sentono leggermente più coinvolti nella sua elaborazione e lo reputano più coerente con il suo effettivo svolgimento rispetto agli studenti regolari (fig. 8).

Inoltre, gli studenti regolari, rispetto a quelli fuoricorso, soffrono maggiormente le modalità organizzative del tirocinio, riportando una totale mancanza di pianificazione in quasi un caso su cinque, una variabilità del contributo richiesto sulla base delle esigenze quotidiane in tre casi su 10 e una chiara pianificazione delle attività di tirocinio in meno della metà dei casi (fig. 9). Riguardo alla valutazione di adeguatezza del numero di ore di tirocinio per lo svolgimento del progetto formativo, sono sempre gli studenti regolari a lamentare in misura maggiore un'eccessiva brevità dell'esperienza (fig. 10).

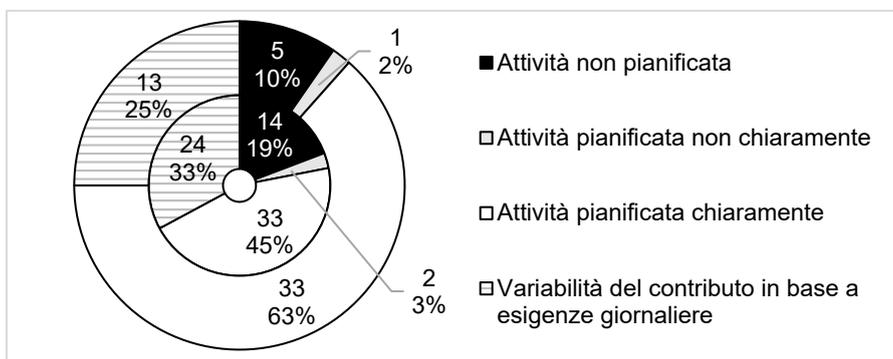


Fig. 9 – Modalità organizzative del tirocinio: confronto tra studenti fuoricorso (cerchio esterno) e studenti regolari (cerchio interno). Il totale delle risposte per ciascun gruppo è superiore al numero dei rispondenti perché era possibile fornire più di una risposta.

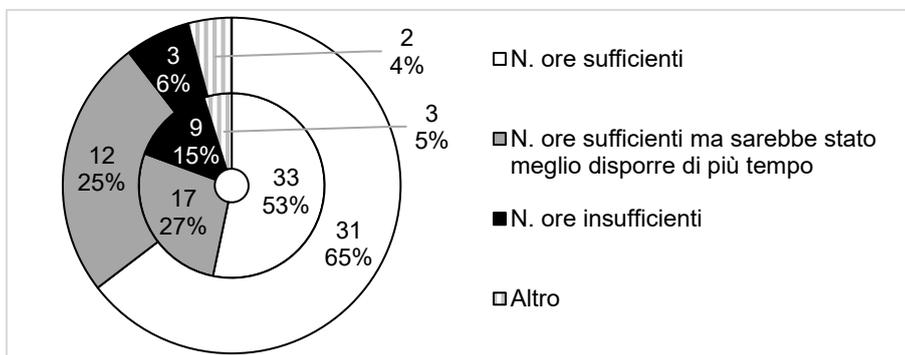


Fig. 10 – Adeguatezza del numero di ore di tirocinio per lo sviluppo del progetto formativo: confronto tra studenti fuoricorso (cerchio esterno) e studenti regolari (cerchio interno).

3.4 Competenze richieste e competenze sviluppate

Approfondendo l'analisi delle attività, è stato chiesto agli studenti di indicare quali settori di competenze pregresse fossero richiesti per svolgere il tirocinio e quali siano stati sviluppati durante l'esperienza stessa. In generale, il tirocinio sembra richiedere e promuovere (e in media richiede un po' più di quanto promuova) la capacità di gestire lo stress, ma anche l'identificazione del tirocinante in un ruolo professionale all'interno dell'organizzazione ospitante, ruolo che appare caratterizzato da prerequisiti e sviluppi importanti legati alle capacità di problem solving, nonché all'attivazione di risorse in termini di abilità comunicative e relazionali collegate all'interazione

in presenza (Pillera, 2017a, pp. 59-77). In tab. 3 presentiamo il confronto percentuale dettagliato tra le due categorie di studenti, per le quali evidenziamo, nell'elenco a seguire, le principali differenze in merito all'autovalutazione delle competenze richieste e sviluppate in funzione del tirocinio.

- *Abilità comunicative.* Per quanto riguarda la necessità pregressa di conoscenza di una lingua straniera, gli studenti fuoricorso hanno una più equa distribuzione delle risposte nei tre gradini inferiori della scala e, allo stesso tempo, sembrano averla sviluppata meglio, sebbene L2 e competenze informatiche rappresentino le due aree con le valutazioni più basse in entrambi i gruppi. Riguardo alla competenza nella lingua madre (capacità comunicative orali e scritte, da considerarsi fattori chiave come precedentemente sottolineato), i fuoricorso le attribuiscono una maggiore centralità, sia in termini di sviluppo che, soprattutto, in termini di requisito. Inoltre, quasi la metà dei fuoricorso valuta l'espressione orale abbastanza ben sviluppata, a differenza degli studenti regolari, che si dividono quasi omogeneamente sui tre passi più alti della scala.
- *Autonomia.* Gli studenti fuoricorso considerano competenze quali l'organizzazione del lavoro, la scelta e la gestione dei materiali di lavoro, le capacità critiche e autocritiche come maggiormente necessarie e percepiscono il loro sviluppo (specialmente quello delle prime due) in maniera significativamente più ampia.
- *Capacità organizzative.* In un quadro di sostanziale omogeneità, notiamo che i fuoricorso ritengono in più larga misura estremamente importante l'identificazione nel proprio ruolo e, allo stesso tempo, dichiarano tale capacità poco o per nulla sviluppata in misura minore rispetto agli studenti regolari. Sottolineiamo inoltre che gli studenti fuoricorso considerano le loro capacità di progettazione molto più sviluppate e sembrano aver imparato di più in merito ad attività che includono la gestione di gruppo.
- *Competenza sociale.* Come abbiamo appena visto, la competenza nel gestire un gruppo sembra rappresentare un'abilità pregressa in misura maggiore fra gli studenti regolari, tuttavia la capacità nel lavoro di gruppo e nel relazionarsi con altre persone (specialmente quest'ultima) appaiono molto più apprezzate dai fuoricorso, sia come prerequisito del tirocinio che come uno dei suoi migliori risultati.
- *Abilità trasversali.* Sebbene gli studenti fuoricorso considerino in misura minore la capacità di sopportare lo stress quale prerequisito molto importante, allo stesso tempo, rispetto agli studenti regolari, la reputano molto più sviluppata grazie all'esperienza di tirocinio, così come l'abilità nel problem solving.

In sintesi, le distribuzioni nelle risposte suggeriscono che, di norma, gli studenti fuoricorso considerano le competenze delineate nel questionario necessarie e sviluppate in misura maggiore rispetto agli studenti regolari.

Tab. 3 – Autovalutazione delle competenze richieste e sviluppate nel tirocinio (dati in percentuale su 62 studenti regolari e 48 fuoricorso).

	Competenza richiesta				Competenza sviluppata				
	Per niente	Poco	Abb.	Molto	Per niente	Poco	Abb.	Molto	
Lingua straniera	29,2	33,3	31,3	6,3	52,1	31,3	6,3	10,4	Fuor.
	27,4	53,2	16,1	3,2	59,7	32,3	4,8	3,2	Reg.
Abilità informatiche	18,8	35,4	35,4	10,4	35,4	45,8	10,4	8,3	Fuor.
	12,9	37,1	38,7	11,3	43,5	29	21	6,5	Reg.
Espressione orale	0	0	20,8	79,2	8,3	14,6	47,9	29,2	Fuor.
	0	3,2	33,9	62,9	8,1	22,6	33,9	35,5	Reg.
Espressione scritta	2,1	16,7	31,3	50	12,5	31,3	33,3	22,9	Fuor.
	4,8	9,7	46,8	38,7	16,1	35,5	30,6	17,7	Reg.
Organizzazione lavoro	0	2,1	31,3	66,7	2,1	12,5	33,3	52,1	Fuor.
	1,6	4,8	32,3	61,3	6,5	9,7	43,5	40,3	Reg.
Scelta e gestione materiale di lavoro	0	8,3	35,4	56,3	4,2	10,4	33,3	52,1	Fuor.
	0	3,2	43,5	53,2	6,5	11,3	43,5	38,7	Reg.
Capacità critica e autocritica	0	0	33,3	66,7	0	6,3	35,4	58,3	Fuor.
	1,6	4,8	30,6	62,9	4,8	11,3	29	54,8	Reg.
Capacità progettuali	0	6,3	27,1	66,7	4,2	10,4	20,8	64,6	Fuor.
	1,6	1,6	30,6	66,1	8,1	11,3	37,1	43,5	Reg.
Gestione gruppo di lavoro	2,1	10,4	39,6	47,9	2,1	16,7	31,3	50	Fuor.
	4,8	6,5	33,9	54,8	11,3	14,5	30,6	43,5	Reg.
Identificazione nel ruolo	0	2,1	22,9	75	2,1	2,1	33,3	62,5	Fuor.
	0	8,1	24,2	67,7	4,8	12,9	21	61,3	Reg.
Relazione con altre persone	0	0	2,1	97,9	2,1	2,1	16,7	79,2	Fuor.
	0	1,6	9,7	88,7	1,6	9,7	21	67,7	Reg.
Lavoro in gruppo	0	4,2	37,5	58,3	0	12,5	33,3	54,2	Fuor.
	0	3,2	45,2	51,6	4,8	16,1	40,3	38,7	Reg.

	<i>Competenza richiesta</i>				<i>Competenza sviluppata</i>				
	<i>Per niente</i>	<i>Poco</i>	<i>Abb.</i>	<i>Molto</i>	<i>Per niente</i>	<i>Poco</i>	<i>Abb.</i>	<i>Molto</i>	
Capacità di apprezzare diversità e multiculturalità	2,1	2,1	20,8	75	0	4,2	35,4	60,4	Fuor.
	1,6	8,1	16,1	74,2	4,8	14,5	27,4	53,2	Reg.
Problem solving	0	0	27,1	72,9	0	2,1	33,3	64,6	Fuor.
	0	3,2	21	75,8	6,5	4,8	32,3	56,5	Reg.
Sopportazione stress	0	8,3	35,4	56,3	0	12,5	35,4	52,1	Fuor.
	0	6,5	27,4	66,1	3,2	9,7	40,3	46,8	Reg.

3.5 Retrospective e prospettive: la relazione tra esperienza di studio e di tirocinio

Una serie di item del questionario suggerisce una riflessione sul rapporto tra tirocinio e università, cercando di esplorare la coerenza delle due esperienze di apprendimento, l'impatto reciproco, l'utilità del tirocinio in termini di orientamento e collocamento professionale. In questa sezione, gli studenti fuoricorso, rispetto a quelli regolari, dichiarano in percentuale maggiore di ricevere materiali di studio (rispettivamente 50% e 43,8%), peraltro giudicati più utili (fig. 11). Anche sulla stima dell'utilità dei corsi universitari rispetto al tirocinio le differenze tra studenti regolari e fuoricorso sembrano piuttosto marcate, con gli ultimi che esprimono un parere più chiaramente positivo (fig. 11). In tendenza opposta, l'influenza del tirocinio sulla tesi di laurea è meno marcata tra i fuoricorso (fig. 12). Considerando il contributo complessivo del tirocinio in rapporto al percorso formativo universitario degli studenti (fig. 11), le differenze tra i due gruppi appaiono più sfocate ma comunque apprezzabili, suggerendo che gli studenti fuoricorso tendono a considerare il tirocinio più coerente con i propri studi, interpretazione confermata da un ulteriore item in cui emergono in misura maggiore tra gli studenti regolari le lacune del background universitario durante lo svolgimento del tirocinio (24,2% contro 14,6% segnato dai fuoricorso).

Con riferimento agli studenti fuoricorso, nell'analisi della rete delle co-occorrenze illustrata in fig. 13 – che incrocia il giudizio quantitativo relativo al contributo del tirocinio al proprio percorso formativo (secondo item della terza coppia in fig. 11) con la sua esplicazione qualitativa – emerge ancora la centralità della figura del tutor, collegata sia a valutazioni alte che basse, mentre le motivazioni principali delle valutazioni più alte appaiono collegate al valore

che sembrerebbe essere stato attribuito all'esperienza in termini di conoscenza (*conosc*, *conoscent*) e partecipazione (*partecip*).

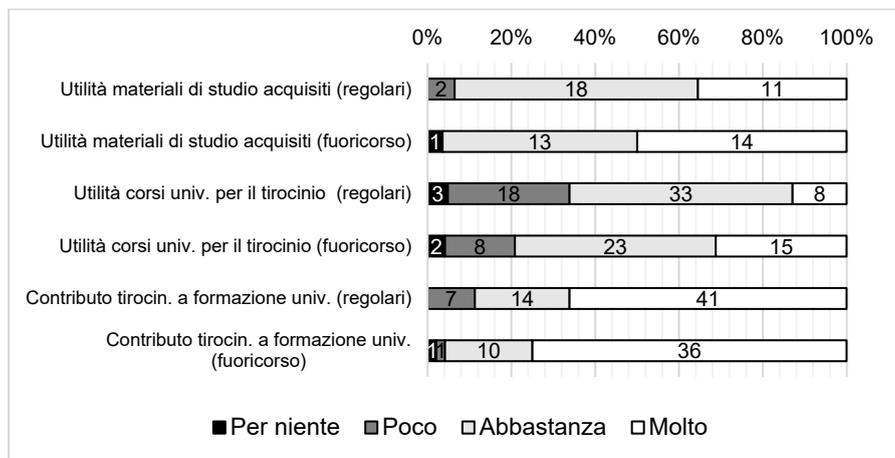


Fig. 11 – Utilità dei materiali di studio eventualmente forniti dall'ente (escluse risposte vuote); utilità dei corsi universitari seguiti rispetto all'esperienza di tirocinio; contributo complessivo del tirocinio al percorso formativo universitario: confronto distribuzioni percentuali nelle risposte di studenti regolari e fuoricorso (con indicazione dei valori assoluti per ciascuna opzione di risposta).

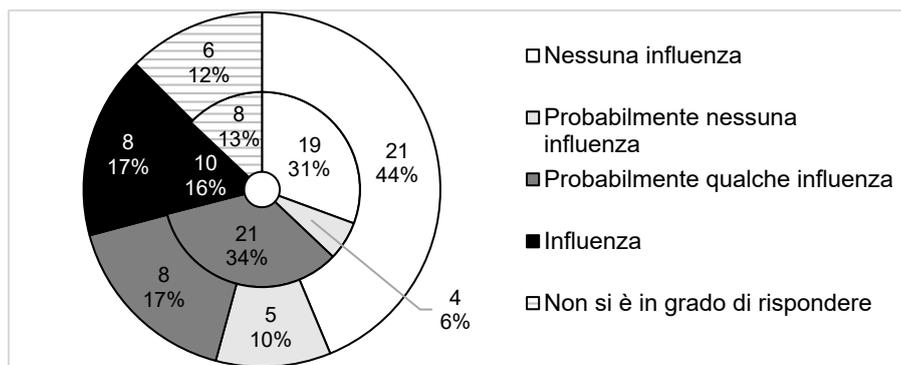


Fig. 12 – Influenza riscontrata o stimata dell'esperienza di tirocinio sul lavoro di tesi: confronto tra studenti fuoricorso (cerchio esterno) e studenti regolari (cerchio interno).

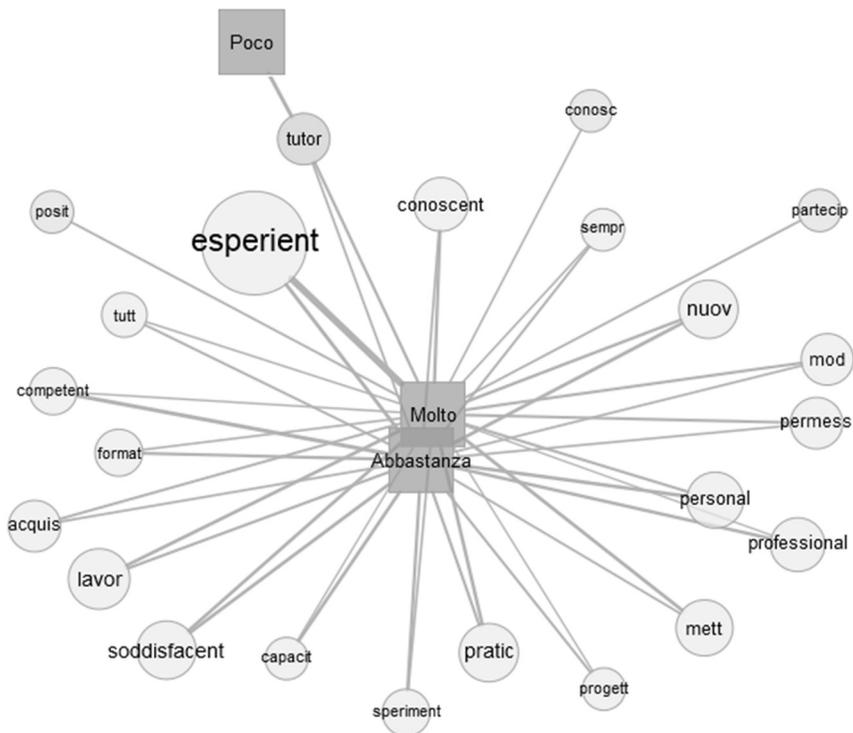


Fig. 13 – Diagramma di rete (words-to-variable) delle prime 50 co-occorrenze di morfemi con $f \geq 5$ (esclusi attivit, tirocin) nell'esplicazione del contributo del tirocinio al percorso formativo degli studenti fuoricorso, incrociate con i giudizi quantitativi su tale contributo (secondo item della terza coppia in fig. 11).

Infine, nei suggerimenti per risolvere possibili problemi e difficoltà, sebbene le risposte siano esigue, i due gruppi presentano diverse somiglianze ma anche alcune apprezzabili differenze (fig. 14): gli studenti fuoricorso, da un lato, danno un peso minore alla preparazione universitaria orientata a specifiche tipologie di utenza (in particolare tecniche di intervista e pedagogia clinica) e, soprattutto, una minore importanza al miglioramento del supporto da parte del tutor, dimostrando in qualche modo maggiore integrazione, flessibilità e adattamento rispetto ai contesti; da un altro, assegnano un peso nettamente superiore al miglioramento della comunicazione e della condivisione degli obiettivi tra università (ufficio tirocinio, docenti) e aziende – cui attribuiscono l'insorgere di problemi burocratici e, ancora una volta, il rischio di incoerenza tra tirocinio e percorso universitario – e richiedono in misura assai maggiore un ricalibrazione del rapporto tra il tempo dedicato al tirocinio e agli esami curriculari.

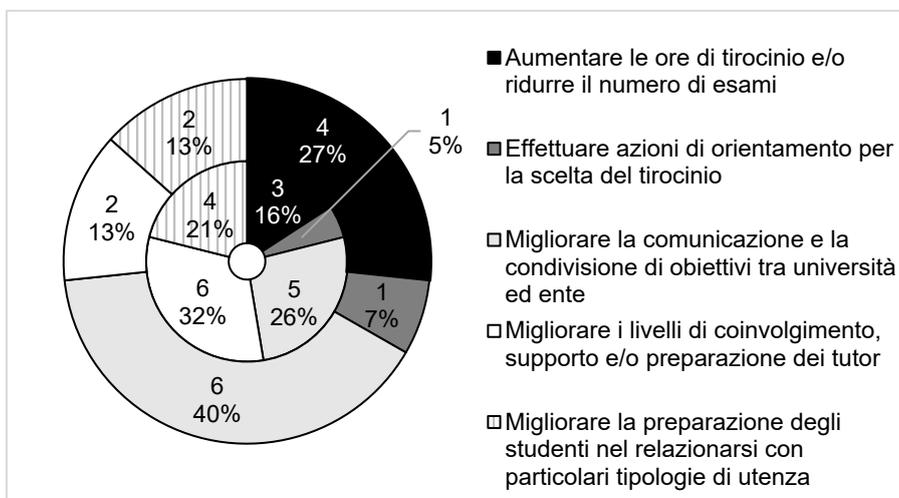


Fig. 14 – Categorizzazione e rappresentazione percentuale delle soluzioni individuate dagli studenti alle difficoltà riscontrate durante il tirocinio: confronto tra studenti fuoricorso (cerchio esterno) e studenti regolari (cerchio interno). Il totale delle risposte per ciascun gruppo è inferiore al numero dei rispondenti perché non tutti hanno fornito una risposta.

4. Conclusioni e sviluppi futuri

Sebbene le differenze tra i due gruppi di studenti esaminate nei paragrafi precedenti alcune volte non emergano distintamente, nel complesso appare abbastanza chiaramente che gli studenti fuoricorso affrontano il tirocinio con più ottimismo e lo vivono più serenamente, probabilmente usufruendo anche di un più ampio background di esperienze personali e professionali rispetto agli studenti regolari. Questo sembra condurre a risvolti positivi sui rapporti instaurati nel contesto aziendale, sulle competenze sviluppate e il loro riconoscimento, sulla sensazione di trovarsi in un percorso ininterrotto e coerente tra l'università e il mondo del lavoro. Il dato trova conferma, seppure debole, in uno degli item sul collocamento post-tirocinio: il 29,2% dei fuoricorso, infatti, riceve una proposta di lavoro dall'ente ospitante, contro 22,6 % degli studenti regolari.

Naturalmente, il panorama sopra tratteggiato influisce sul livello di gratificazione: ad es. solo il 4,2% degli studenti fuoricorso (contro il 12,9% di quelli regolari) sconsiglierebbe l'esperienza di tirocinio ai colleghi e il diagramma in fig. 15 evidenzia inequivocabilmente la maggiore soddisfazione generale dei fuoricorso relativamente all'esperienza svolta. Nel diagramma a rete delle co-occorrenze in fig. 16, utilizzato per analizzare l'esplicazione qualitativa del

giudizio quantitativo sulla soddisfazione generale, emerge ancora una volta, in positivo come in negativo, il ruolo chiave del tutor aziendale, inserito quale snodo cruciale nella complessità dei fattori in gioco che abbiamo cercato di mettere in luce in questo contributo.

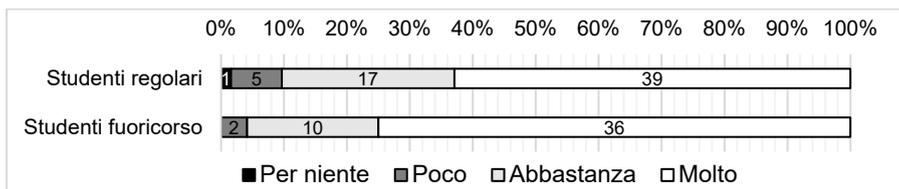


Fig. 15 – Soddisfazione complessiva rispetto all’esperienza di tirocinio: confronto distribuzioni percentuali nelle risposte di studenti regolari e fuoricorso (con indicazione dei valori assoluti per ciascuna opzione di risposta).

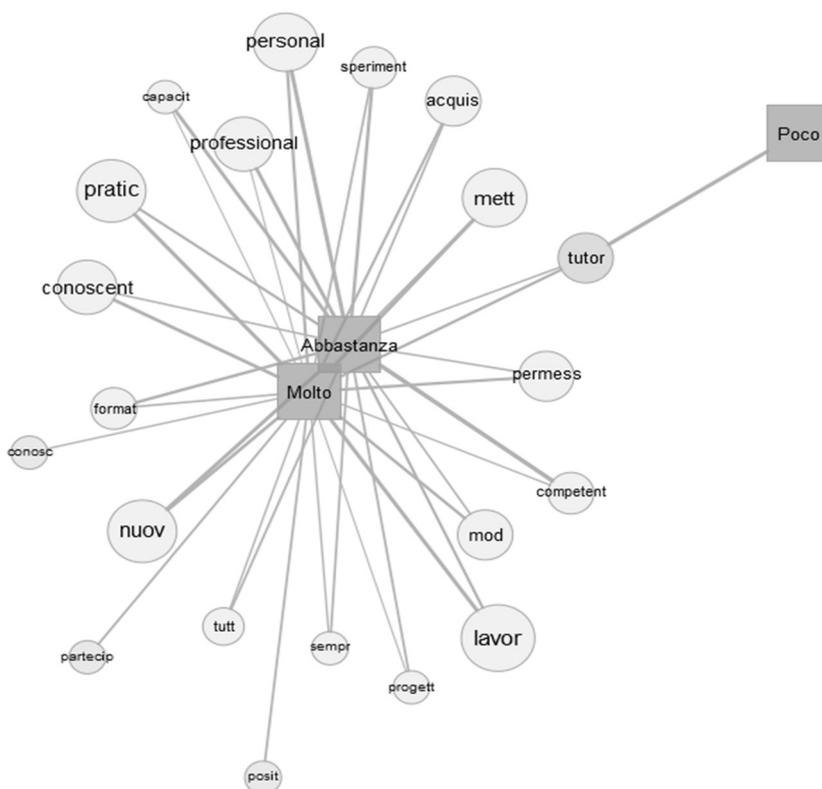


Figura 16 – Diagramma di rete (words-to-variable) delle prime 60 co-occorrenze di morfemi con $f \geq 5$ (esclusi esperient, tirocin, soddisfacent) nell’esplicazione della soddisfazione complessiva degli studenti fuoricorso sul tirocinio, incrociate con i giudizi quantitativi sulla stessa (fig. 15).

Sulla base del materiale analizzato, possiamo riassumere nei punti che seguono le principali proposte per migliorare ulteriormente l'esperienza di tirocinio, in particolare per gli studenti fuoricorso.

1. Affinare la coerenza tra il corso di laurea e i contesti di lavoro, favorendo l'approfondimento delle conoscenze sugli ambiti in cui spendere le proprie competenze e utilizzando un più ampio ventaglio di strategie di insegnamento (case history, simulazioni, laboratorio, attività basate su compiti, ecc.) in grado di ridurre il *gap* teoria-pratica segnalato, soprattutto per quanto riguarda gli strumenti diagnostici (tecniche di intervista, somministrazione di test) e le tecniche cliniche (per disabilità mentali, autismo, difficoltà di apprendimento), ma anche le capacità di comunicazione e relazione con particolari gruppi di utenti (ludopatici, tossicodipendenti, persone indigenti).
2. Migliorare i servizi di orientamento per una scelta più consapevole sia del settore che dell'organizzazione specifica in cui svolgere il tirocinio e coinvolgere maggiormente gli studenti nella condivisione degli obiettivi e dei metodi dell'attività, promuovendo, così, responsabilità e impegno.
3. Approfondire i collegamenti tra l'Università e gli enti ospitanti, prestando particolare attenzione alla relazione con (e possibilmente alla formazione delle) persone che assumeranno il ruolo di tutor, e mantenere una comunicazione costante con enti e studenti, controllando che le attività effettivamente svolte siano coerenti con il ruolo dei tirocinanti e corrispondano al progetto formativo siglato.
4. Promuovere servizi per lo sviluppo della carriera, così da supportare gli studenti nella loro ricerca di lavoro dopo il tirocinio sfruttando le competenze, le risorse e i contatti sviluppati durante questa esperienza.

Paradossalmente i fuoricorso, che dovrebbero rappresentare una categoria di studenti problematica, non solo sembrano apprezzare maggiormente e fruire meglio il tirocinio, ma forniscono anche spunti interessanti (che potrebbero essere ulteriormente esplorati attraverso casi studio e interviste in profondità) per migliorare l'esperienza di quelli regolari. In effetti, successive indagini qualitative e quantitative potrebbero chiarire, al di là degli elementi già individuati, se l'orientamento più positivo degli studenti fuoricorso sia imputabile alle differenze nelle aspettative o nella propensione critica oppure ancora, come ipotizzato sopra, a una diversa maturità, frutto di un più articolato background di percorsi formativi, lavorativi ed esistenziali.

Approfondimenti di questo genere potrebbero prendere a base per l'analisi, valorizzandole, le relazioni finali prodotte dai tirocinanti alla fine del percorso, troppo spesso e troppo a lungo concepite dagli studenti e relegate dai docenti al ruolo di mero mezzo burocratico di attestazione del percorso,

come ad esempio recentemente realizzato, con metodologia qualitativa, da Perla (2012) o da Salerni et al. (2013). Su tali corpora, sicuramente vasti e potenzialmente preziosi, potrebbero incentrarsi anche ricerche basate su tecniche di *text data mining* come quelle illustrate in questo contributo.

Bibliografia

- AlmaLaurea (2019), *XXI indagine (2019) - Profilo dei Laureati 2018*, testo disponibile al sito: <https://www.almalaurea.it/universita/profilo/profilo2018>.
- Bartolini A. e Riccardini M.G. (2006), *Il tirocinio nella professionalità educativa*, Il Segno dei Gabrielli editori, Negarine di S. Pietro in Cariano (VR).
- Bastianoni P. e Spaggiari E. (2014), *Apprendere per educare. Il tirocinio nelle lauree di scienze dell'educazione*, Carocci Editore, Roma.
- Bondioli A., Ferrari M., Marsilio M. Tacchini I., a cura di (2006), *I saperi del tirocinio. Formare gli insegnanti nelle SSIS*, FrancoAngeli, Milano.
- Danowski J.A. (1993), *Network Analysis of Message Content*, in Richards W.D. Jr. and Barnett G.A., eds., *Progress in communication sciences IV*, Ablex, Norwood.
- Decreto MIUR n. 270 del 22 ottobre 2004, recante *Modifiche al regolamento recante norme concernenti l'autonomia didattica degli atenei, approvato con decreto del Ministro dell'università e della ricerca scientifica e tecnologica 3 novembre 1999, n. 509*.
- Di Pol R.S. (2014), *La formazione del maestro italiano tra istanze pedagogiche e scelte politico-sociali. Un profilo storico*, in Maccario D., a cura di, *Insegnare a insegnare. Il tirocinio nella formazione dei docenti: il caso di Torino*, FrancoAngeli, Milano.
- Fruchterman T.M.J. and Reingold E.M. (1991), "Graph Drawing by Force-directed Placement", *Software-Practice and experience*, 21, 1: 1129-1164.
- Greimas A.J. (1966), *Sémantique structural*, Larousse, Paris.
- Higuchi K. (2015), *KH Coder 2.x Reference Manual*, Ritsumeikan University, Kyoto, testo disponibile al sito: <https://pdfs.semanticscholar.org/ff23/bda-deaa0ef027f1bfa2ba05526dbc8f69227.pdf>.
- Iuzzolino G., Lotito S., Sofronici B. e Tosi G. (2018), *L'attuazione della Raccomandazione del Consiglio dell'Unione Europea su un Quadro di Qualità per i tirocinanti*, INAPP, Roma.
- Lancia F. (2020), *T-Lab plus 2020. Manuale d'uso. Strumenti per l'analisi dei testi*, testo disponibile al sito: <http://tlab.it/it/download.php>.
- Maccario D., a cura di (2014), *Insegnare a insegnare. Il tirocinio nella formazione dei docenti: il caso di Torino*, FrancoAngeli, Milano.
- Nucleo di valutazione dell'Università di Catania (2017). *Valutazione del Sistema di Qualità dell'Ateneo e dei Corsi di Studio (CdS)*, testo disponibile al sito: www.unict.it/it/ateneo/relazioni-del-nucleo.
- Nucleo di valutazione dell'Università di Catania (2018). *Relazione annuale 2018*. testo disponibile al sito: www.unict.it/it/ateneo/relazioni-del-nucleo.

- Osgood C.E. (1959), *The Representational Model and Relevant Research Methods*, in De Sola Pool I., ed., *Trends in Content Analysis*, Urbana, University of Illinois Press.
- Paparella N. e Perucca A. (2006), *Le attività di laboratorio e di tirocinio nella formazione universitaria. Vol. II. Indagini e strumenti*, Armando Editore, Roma.
- Pastore S., Manuti A., Falcicchio G., Rossini V. e Gemma C. (2015), “Interventi di recupero studenti fuoricorso e inattivi. Riflessioni a margine di un progetto pilota”, *Orientamenti Pedagogici*, 62, 2: 339-350.
- Perla L., a cura di (2012), *Scrittura e tirocinio universitario. Una ricerca sulla documentazione*, FrancoAngeli, Milano.
- Perucca A., a cura di (2005), *Le attività di laboratorio e di tirocinio nella formazione universitaria, vol. I. Identità istituzionale, modello organizzativo, indicatori di qualità*, Armando editore, Roma.
- Pillera G.C. (2017a), *Il tirocinio e l'autovalutazione delle competenze*, in Corsini C. e Strongoli R.C., a cura di, *Valutare il tirocinio universitario. L'esperienza del CdLM in Scienze pedagogiche e progettazione educativa dell'Università di Catania*, CUECM, Catania.
- Pillera G.C. (2017b), *Analisi testuale delle risposte aperte*, in Corsini C. e Strongoli R.C., a cura di, *Valutare il tirocinio universitario. L'esperienza del CdLM in Scienze pedagogiche e progettazione educativa dell'Università di Catania*, CUECM, Catania.
- Pillera G.C. (2020), “La cultura degli orti scolastici in Sicilia. Analisi testuale sugli obiettivi di progetti didattici di coltura della terra come contributo per l'analisi del discorso sull'educazione all'ambiente e alla sostenibilità”, *Ricerche di Pedagogia e Didattica - Journal of Theories and Research in Education*, 15, 1: 127-201.
- Rastier F. (1987), *Sémantique interprétative*, Presses Universitaires de France, Paris.
- Rastier F., Cavazza M., Abeillé A. (2002), *Semantics for Descriptions*, CSLI, Stanford.
- Salerni A., a cura di (2007), *Apprendere tra università e lavoro. Un modello per la gestione del tirocinio universitario*, Homolegens, Roma.
- Salerni A., Sposetti P. e Szpunar G. (2013), “La narrazione scritta come elemento di valutazione del tirocinio universitario”, *Ricerche di Pedagogia e Didattica - Journal of Theories and Research in Education*, 8, 2: 9-26, doi: 10.6092/issn.1970-2221/4034.
- Scierra I.D.M. (2017), *Il questionario per la valutazione del tirocinio*, in Corsini C. e Strongoli R.C., a cura di, *Valutare il tirocinio universitario. L'esperienza del CdLM in Scienze pedagogiche e progettazione educativa dell'Università di Catania*, CUECM, Catania.
- Tinto V. (1988), “Stages of student departure: Reflections on the longitudinal character of student leaving”, *Journal of Higher Education*, 59, 4: 438-455.
- Tinto V. (1993), *Leaving college. Rethinking the causes and cures of student attrition*, University of Chicago Press, Chicago, IL.
- Tomarchio M., a cura di (2008), *Lo sperimentalismo pedagogico in Sicilia e Michele Crimi*, Anicia, Roma.

Postfazione.

A conti fatti...

di *Maria Tomarchio*

Mi accingo a restituire alcune considerazioni di sintesi sul volume che accoglie i contributi del seminario di studi sul tema *Evaluating educational quality* (25-26 maggio 2017) e, senza particolare intenzione, la memoria rievoca il ricordo di altri due convegni che, a distanza di pochi anni, hanno avuto luogo presso la medesima sede del Dipartimento di Scienze della Formazione dell'Università di Catania: *Misurazione e valutazione di scuole e università. Inattualità di Aldo Visalberghi* (Catania, 18-19 gennaio 2016) e *Pedagogia militante Diritti, culture, territori* (Catania, 6-7-8 novembre 2014).

L'impressione che ricavo da quella che può apparire una semplice associazione d'idee pressoché inconsapevole è che, effettivamente, una linea di continuità ideale tenga assieme un buon numero di iniziative di ricerca della sede catanese e che al suo interno si collochi coerentemente il seminario di studi oggetto del presente volume. Iniziative, più o meno immediatamente prossime in ordine di tempo, volte ad un confronto e ad un approfondimento spesso, e non a caso, di respiro internazionale, che richiamano e pongono in dialogo, direttamente o indirettamente, terreni di studio e di interesse della pedagogia generale e sociale e della pedagogia sperimentale assieme.

Un significativo segmento temporale, in tale ideale percorso, ha visto la presenza stabile a Catania, in veste di ricercatore di Pedagogia sperimentale, di Cristiano Corsini, con il quale è nato e si è sviluppato nel tempo un dialogo sul terreno del lavoro di ricerca e un rapporto di collaborazione che, proprio perché sospinto da due interpreti, rappresentanti di settori scientifico-disciplinari diversi, è divenuto via via interessante e significativo, fino al punto da indurci ad avviare, all'interno del piano degli studi del corso di laurea magistrale *Scienze pedagogiche e progettazione educativa* da me presieduto, un'esperienza di co-docenza in un corso integrato di 12 crediti formativi composto da due moduli, uno di *Modelli di progettazione pedagogica e politiche educative* (6 cfu M-Ped/01) e l'altro di *Metodi e tecniche di*

valutazione (6 cfu M-Ped/04). Un corso integrato che, peraltro, dall'a.a. 2015/2016 è l'unico insegnamento propedeutico all'avvio delle attività di tirocinio esterno.

Didattica e ricerca, come sappiamo, specialmente in università, non viaggiano dentro comparti separati; tuttavia è indubbio che a dare senso e coesione al dialogo scientifico avviato con Corsini, al di là delle iniziative condivise, è stata anche la convinzione, ferma e profonda da entrambe le parti, che la pedagogia, oggi più che mai, non può darsi fini in una prospettiva di libertà e democrazia se guarda ai processi valutativi e alla didattica puntando, più o meno dichiaratamente, all'omologazione.

Tornando alla presente raccolta di contributi, è possibile rievocare, se non proprio i motivi ispiratori, certamente la filosofia che opera sullo sfondo della sopramenzionata ideale linea di continuità della ricerca, richiamando accenti ricorrenti nei saggi proposti al lettore in questa sede.

Un primo richiamo va all'indirizzo delle responsabilità dei saperi pedagogici in età contemporanea nella prospettiva di un definitivo superamento di ogni rapporto di separazione, o peggio di estraneità, tra dimensioni empirico-sperimentali legate ai processi valutativi, riflessioni di carattere teorico-fondativo e aspetti socio-educativi; l'invito è ad operare sempre scelte consapevoli delle ricadute rispetto ai fini ai quali devono corrispondere gli strumenti adottati, a far propri, come intrinsecamente legati alla logica pedagogica, gli aspetti più problematici della contemporaneità dal punto di vista dei diritti dei soggetti, delle società multiculturali, del rapporto tra vissuto delle persone e dei territori, e di farlo a prescindere da diffuse e generiche vesti ammantate di facile solidarismo. In tale direzione Christopher Tienken si chiede, ad esempio, se New York City non avrebbe potuto utilizzare meglio i 75 milioni di dollari investiti su inadeguati programmi di merito per i docenti durante tre anni di sperimentazione, dal 2007 al 2010, magari a sostegno delle quasi 8.000 famiglie più povere, fornendo loro opportunità aggiuntive di vita per la crescita dei figli.

Un ulteriore richiamo è ad allargare ad una dimensione internazionale il panorama del confronto entro cui condurre le nostre considerazioni, più di quanto già non accada; non vi è ambito di studi pedagogici che oggi non abbia ad arricchirsi da un tale decentramento e incrocio di prospettive, come il lettore ha avuto modo di appurare grazie al punto di vista restituito da Silje Gløppen e Guri Skedsmo, che ricostruisce il dibattito norvegese degli ultimi anni intorno alle questioni relative alla valutazione degli insegnanti e dell'insegnamento; un dibattito in cui assume una posizione centrale la domanda sul come possa definirsi la qualità in educazione alla luce dei suoi profondi legami con questioni di legittimità e potere, etica e valori, aspetti diversi della conoscenza e della produzione stessa di conoscenza.

I tempi sono maturi, è evidente, anche per analisi di carattere comparativo che non circoscrivano il terreno d'indagine esclusivamente a studi di ambito storico. Indicazioni estremamente interessanti in tal senso giungono, tra gli altri, dal contributo di Camilla Addey che, conducendo una consapevole critica alla pervasività dimostrata dalla *global education governance* dell'OCSE, consegna un interrogativo sul quale vale riflettere: perché l'attuale quarto Obiettivo per lo Sviluppo Sostenibile è così ben allineato con la metrica di PISA?

Non ultima per rilevanza la sollecitazione che proviene dal contributo di Giuseppe Pillera, in cui lo sguardo riflessivo di studenti universitari regolari e fuoricorso sulla propria esperienza di tirocinio curriculare ci interroga su quanto l'università possa e debba ancora migliorare: non solo nei termini di un'offerta didattica che potenzi coerenza e continuità con i contesti professionali attraverso la diversificazione dei metodi e degli strumenti di insegnamento (oltre che dei soli contenuti disciplinari), ma anche nei termini dell'intensificazione di un dialogo e di un radicamento sul territorio che la metta nelle condizioni di assolvere pienamente le sue funzioni orientative nei confronti dello studente. Lo studio, di rilievo anche per gli aspetti metodologici dell'analisi, è peraltro basato su uno strumento già da anni in adozione presso il corso di laurea magistrale che presiedo, uno strumento che, messo a punto proprio dagli studenti e dalle studentesse del dipartimento in seno ai laboratori extracurricolari allora coordinati dal prof. Corsini, è orientato a riconoscere e valorizzare la soggettività dei processi di valutazione e la loro natura situata.

Su questioni analoghe, del resto, Ira Vannini consegna al volume una relazione approfondita che, incentrata sul contesto della formazione docenti in campo valutativo, raccomanda che gli insegnanti sviluppino una concreta esperienza di percorsi di ricerca educativa. La postura cui essi sono invitati – ma la sollecitazione è diretta direttamente anche all'università, investita del compito di formare e aggiornare le loro competenze – è quella del docente-ricercatore, come delineato nel modello di Ricerca-Formazione del CRESPI (Centro di Ricerca Educativa sulla Professionalità dell'Insegnante). Facendo esperienza del momento valutativo, entrando nel merito degli aspetti procedurali dell'osservazione sul campo e della costruzione e decostruzione degli strumenti strutturati, i docenti possono trovare occasione insostituibile per una riflessione collegiale sui fini, i mezzi e il senso stesso dell'educare, ma anche evitare che questi stessi strumenti, come avverte Visalberghi, sulle cui orme si muove il contributo di Vannini, possano “piovere dall'alto” come strumento di controllo e uniformazione ai *desiderata* degli apparati di potere politico-economico.

A conti fatti, ciascuno dei contributi, e il volume nel suo insieme, consegna all'attenzione dati, riflessioni critiche, considerazioni su importanti aspetti di metodo, ma soprattutto richiama il lettore su questioni aperte di rilievo che sollecitano ulteriore approfondimento. Sebbene appaia ormai quasi un luogo comune sottolineare quale importanza rivesta la qualità delle domande ai fini del progresso nella ricerca, non vi è dubbio che, tra le righe delle riflessioni che il testo consegna, troviamo interrogativi ai quali è legata la qualità del futuro della ricerca educativa.

Gli autori

Camilla Addey, Marie Curie research fellow presso il GEPS – Centro di ricerca Globalisation, Education and Social Policies dell’Universitat Autònoma de Barcelona (Spagna). Precedentemente, è stata docente di *Educazione internazionale e comparata* presso il Teachers College, Columbia University (USA), e ricercatrice presso l’Università Humboldt di Berlino (Germania), dove ha conseguito un PhD sulla valutazione internazionale su larga scala dell’UNESCO in Mongolia e Laos. Conduce ricerche sulle valutazioni internazionali su larga scala, sulla politica educativa globale e sulle privatizzazioni nel settore dell’istruzione. Attualmente è impegnata nel progetto di ricerca ILSA INC, sul coinvolgimento del settore privato nelle valutazioni internazionali su larga scala.

Cristiano Corsini, professore associato presso il Dipartimento di Scienze della formazione dell’Università degli studi Roma Tre, dove ricopre gli insegnamenti di *Pedagogia sperimentale, Pedagogia sperimentale e valutazione scolastica, Docimologia e valutazione dei servizi socio-educativi*. È stato ricercatore all’Università degli studi di Catania e professore associato all’Università degli studi “G. D’Annunzio” di Chieti-Pescara, nonché visiting professor presso la Seton Hall University (New Jersey, USA). Collabora al comitato editoriale di diverse riviste, tra cui *I Problemi della Pedagogia* e *Cadmo*, e co-dirige *EducAzione – Collana di Pedagogia e Didattica* di A.&G. – CUECM. È autore e curatore di monografie e numerosi contributi scientifici, ha partecipato come valutatore e formatore all’indagine internazionale OCSE-PISA. La sua monografia *Valutare scuole e docenti* ha vinto il Premio Italiano di Pedagogia – SIPED 2016. La sua attività di ricerca è finalizzata all’analisi dei paradigmi, degli approcci e delle scelte metodologiche che caratterizzano le prassi valutative di scuole, università, docenti e istituti preposti alla valutazione educativa (INVALSI, ANVUR, OCSE, IEA).

Silje Gloppen, PhD in valutazione delle pratiche didattiche degli insegnanti in Norvegia, lavora presso il Department of Teacher Education and School Research, University of Oslo (Norvegia), dove è stata anche research fellow in seno all’Institute of Teacher Education and School Development (ILS). I suoi principali interessi di ricerca comprendono lo sviluppo scolastico, la politica dell’istruzione, la governance e l’internazionalizzazione dell’istruzione, nonché lo sviluppo e la professionalizzazione

degli insegnanti. È anche coinvolta nel progetto Erasmus + “Aiding Culturally Responsive Assessment in Schools” (ACRAS).

Germana Patti, docente di lettere presso il CPIA “Alberto Manzi” di Siracusa, dove è stata Funzione strumentale Area 4, docente responsabile della programmazione delle attività legate alla realizzazione del Piano di Miglioramento, responsabile del progetto “Un percorso sperimentale di didattica per competenze nei CPIA” e coordinatrice o partecipante in numerose attività dell’istituto. È coinvolta nelle attività del Centro Regionale di Ricerca, Sperimentazione e Sviluppo dei CPIA siciliani.

Giuseppe C. Pillera, dottore di ricerca in “Fondamenti e metodi dei processi formativi” presso il Dipartimento di Scienze della formazione dell’Università degli studi di Catania, dove ricopre l’insegnamento di *Progettazione e valutazione di interventi formativi sul territorio* ed è stato assegnista di ricerca. Dal 2015 al 2016 visiting researcher presso l’Universidad de Sevilla (Spagna), è componente di gruppi di ricerca universitari, formatore docenti e valutatore esterno in diversi progetti scolastici. È revisore di riviste scientifiche internazionali e componente del comitato scientifico della collana Tascabili Bonanno - Pedagogia. I suoi interessi di ricerca si collocano nel campo delle nuove tecnologie per la formazione e l’apprendimento collaborativo (in particolare LMS Moodle), la progettazione e la valutazione nei contesti socioeducativi e della formazione professionale, scolastica e universitaria, ambiti di studio sui quali ha scritto una monografia, numerosi contributi in volume e articoli su riviste nazionali e internazionali.

Guri Skedsmo, già professoressa associata presso il Department of Teacher Education and School Research, University of Oslo (Norvegia) e professoressa all’Institute for the Management and Economics of Education, University of Teacher Education Zug (Svizzera), insegna attualmente presso l’Institute for Research on Professions and Professional Learning (IPP) di Goldau (Svizzera). I suoi principali interessi di ricerca riguardano i settori della governance e della leadership educativa, dello sviluppo e del cambiamento scolastico e dello sviluppo professionale dei dirigenti scolastici. È editor-in-chief per la rivista scientifica *Educational Assessment, Evaluation and Accountability (EAEA) - International Journal of Policy, Practice and Research*.

Christopher H. Tienken, professore associato di *Education Administration* presso il College of Education and Human Services, Department of Education Leadership, Management & Policy, Seton Hall University (New Jersey, USA). Vanta una vasta esperienza nell’amministrazione della scuola pubblica, dove, iniziando la sua carriera come maestro elementare, ha in seguito ricoperto i ruoli di assistente sovrintendente, direttore del curriculum e dell’istruzione, preside in differenti ordini scolastici. È l’ex editor dell’*American Association of School Administrators Journal of Scholarship and Practice* e l’attuale editor della *Kappa Delta Pi Record*. Collabora con colleghi delle università di Roma Tre e di Catania, dove è stato visiting professor. I suoi interessi di ricerca includono questioni di riforma della scuola come la standardizzazione, l’influenza della qualità del curriculum sui risultati degli studenti

e la validità di test standardizzati ad alto rischio come strumenti decisionali. È autore di oltre 80 pubblicazioni, tra cui monografie, contributi in volume e articoli su riviste, sia in sede nazionale che internazionale.

Maria Tomarchio, professoressa ordinaria di Pedagogia Generale e Sociale presso il Dipartimento di Scienze della Formazione dell'Università degli studi di Catania, dove è Presidente del Consiglio del Corso di Laurea Magistrale *Scienze Pedagogiche e Progettazione Educativa* e vice-coordinatrice del Dottorato di Ricerca *Processi formativi, modelli teorico-trasformativi e metodi di ricerca applicati al territorio*. Componente designata del Comitato Tecnico-scientifico del Centro Regionale di Ricerca, Sperimentazione e Sviluppo dei CPIA siciliani e del Gruppo di consulenza scientifica universitaria della Rete Nazionale dei Centri Regionali di Ricerca, Sperimentazione e Sviluppo per l'Istruzione degli Adulti. Versanti privilegiati nell'attività di ricerca scientifica: Educazione Nuova e Scuola Attiva in Italia e in Europa. Modelli, temi, figure, con particolare riferimento alle multiformi esperienze di ricerca empirico-sperimentale condotte tra ottocento e Novecento; Educazione e Scuole all'aperto; sperimentazioni educativo-didattiche con particolare riferimento al contesto siciliano; Educazione, Ambiente, territorio e cittadinanza attiva democratica.

Giuseppe M. Tomaselli, docente di matematica presso il CPIA "Alberto Manzi" di Siracusa, dove è stato responsabile del progetto "Un percorso sperimentale di didattica per competenze nei CPIA", coordinatore del progetto di ampliamento dell'offerta didattica "Porte aperte al CPIA Lentini" e docente formatore in numerose altre iniziative presso l'istituto. È inoltre coinvolto nelle attività del Centro Regionale di Ricerca, Sperimentazione e Sviluppo dei CPIA siciliani e membro volontario del gruppo di ricerca del Centro di Ricerca, Sperimentazione e Sviluppo della rete dei CPIA del Lazio.

Ira Vannini, professore ordinario in Pedagogia Sperimentale presso il Dipartimento di Scienze dell'Educazione "G.M. Bertin" dell'Alma Mater Studiorum Università di Bologna, dove è Coordinatrice del Corso di Laurea Magistrale a ciclo unico in *Scienze della formazione primaria* e dove insegna *Metodologia della ricerca empirica e Teorie e metodi di progettazione e valutazione*, in particolare nei corsi di laurea per la formazione iniziale degli insegnanti. In sinergia con il Centro CRESPI (<https://centri.unibo.it/crespi/it>), le sue recenti attività di ricerca (in ambito nazionale e internazionale) si sono rivolte a temi quali: i percorsi di Ricerca-Formazione e la ri-progettazione curricolare in ambito scolastico e universitario; l'analisi delle metodologie di formazione dei docenti anche attraverso le tecniche di video-analisi; il monitoraggio di sperimentazioni didattiche nella scuola dell'infanzia, primaria e secondaria; il rapporto tra insegnanti e valutazione scolastica e le pratiche di formative assessment. Le principali metodologie di ricerca utilizzate sono quelle dell'*educational evaluation research* e della ricerca empirica su grande campione.

Questo 
LIBRO

 ti è piaciuto?

Comunicaci il tuo giudizio su:
www.francoangeli.it/latuaopinione.asp



VUOI RICEVERE GLI AGGIORNAMENTI
SULLE NOSTRE NOVITÀ
NELLE AREE CHE TI INTERESSANO?



SEGUICI IN RETE



SOTTOSCRIVI
I NOSTRI FEED RSS



ISCRIVITI
ALLE NOSTRE NEWSLETTER

FrancoAngeli

La passione per le conoscenze

Copyright © 2020 by FrancoAngeli s.r.l., Milano, Italy. ISBN 9788835103738

Vi aspettiamo su:

www.francoangeli.it

per scaricare (gratuitamente) i cataloghi delle nostre pubblicazioni

DIVISI PER ARGOMENTI E CENTINAIA DI VOCI: PER FACILITARE
LE VOSTRE RICERCHE.



Management, finanza,
marketing, operations, HR

Psicologia e psicoterapia:
teorie e tecniche

Didattica, scienze
della formazione

Economia,
economia aziendale

Sociologia

Antropologia

Comunicazione e media

Medicina, sanità



Architettura, design,
territorio

Informatica, ingegneria

Scienze

Filosofia, letteratura,
linguistica, storia

Politica, diritto

Psicologia, benessere,
autoaiuto

Efficacia personale

Politiche
e servizi sociali



FrancoAngeli

La passione per le conoscenze

Copyright © 2020 by FrancoAngeli s.r.l., Milano, Italy. ISBN 9788835103738

Nell'ultimo ventennio si sono sviluppati apparati, attività e istituti appositamente deputati alla valutazione dell'efficacia, dell'efficienza, della qualità del lavoro svolto da scuole e università. I lavori ospitati nel presente volume accettano la sfida della valutazione educativa restituendo esperienze diverse, analisi critiche, punti di forza e di debolezza di un processo che, ambendo a mettere in discussione l'agire educativo, non può continuare a sottrarsi al giudizio: valutare la valutazione è scelta sostenibile e doverosa dal punto di vista educativo e valutativo.

Contributi di: Camilla Addey, Cristiano Corsini, Silje Gloppen, Germana Patti, Giuseppe C. Pillera, Guri Skedsmo, Christopher H. Tienken, Maria Tomarchio, Giuseppe M. Tomaselli, Ira Vannini.

Cristiano Corsini, professore associato presso il Dipartimento di Scienze della Formazione dell'Università degli Studi Roma Tre, autore e curatore di monografie e numerosi contributi scientifici, ha partecipato come valutatore e formatore all'indagine internazionale OCSE-PISA. Nel 2016, la sua monografia *Valutare scuole e docenti* ha vinto il Premio Italiano di Pedagogia – SIPED.

Giuseppe C. Pillera, dottore di ricerca e docente a contratto presso il Dipartimento di Scienze della Formazione dell'Università degli Studi di Catania, è stato visiting researcher presso l'Universidad de Sevilla (Spagna) e componente di gruppi di ricerca universitari, formatore docenti e valutatore esterno in diversi progetti scolastici.

Christopher H. Tienken, professore associato presso il College of Education and Human Services, Department of Education Leadership, Management & Policy della Seton Hall University (New Jersey, USA). Vanta una vasta esperienza nell'amministrazione della scuola pubblica ed è stato visiting professor presso le università di Roma Tre e di Catania.

Maria Tomarchio è professoressa ordinaria di Pedagogia generale e sociale presso il Dipartimento di Scienze della Formazione dell'Università degli Studi di Catania, dove presiede il CdLM *Scienze Pedagogiche e Progettazione Educativa* ed è vice-coordinatrice del Dottorato di Ricerca *Processi formativi, modelli teorico-trasformativi e metodi di ricerca applicati al territorio*.