

# Il dibattito su valutazione ed evidenze: per un processo valutativo credibile e trasparente

Maria Lucia Giovannini

*Alma Mater Studiorum - Università di Bologna, Dipartimento di Scienze dell'Educazione*

doi: 10.7358/ecps-2014-009-giov

marialucia.giovannini@unibo.it

---

## THE DEBATE ON EVALUATION AND EVIDENCE: FOR A CREDIBLE AND TRANSPARENT EVALUATION PROCESS

### ABSTRACT

*This paper presents the main issues concerning the lively national and international debate on the evidence in the evaluation process, with specific attention to the American context in which the debate is particularly prominent. Initial consideration is given to the fundamental steps and elements of reflection and debate associated with the theme of Evidence Based Education in its broadest sense. Then, the specificity of evidence in relation to research and decision-making processes is examined. Particular attention is given to key aspects relating to the development of a theory to describe the role of evidence in evaluation that takes account of its multifaceted nature, the implied value attributed to it, the context in which it is acquired and used, the judgement process that attributes it value and the use of value judgments in the decision-making process. Finally, some issues which make the debate on this theme also relevant to the Italian context are presented in order to assure that the provision or acquisition of evidence is supportive of an evaluation process that is both credible and transparent.*

*Keywords:* Credible evidence in evaluation, Evaluation, Evidence, Evidence based education, Evidence based evaluation.

## 1. ALCUNE PRECISAZIONI E INTERROGATIVI INIZIALI

Valutazione è oggi un termine con cui si deve continuamente fare i conti; ovunque lo si usi sembra essere proposto come «parola d'ordine, sinonimo di prova, rigore, trasparenza, democrazia» (Damiano, 2011, p. 10). Evidenza è anch'esso un termine sempre più diffuso e utilizzato: basti pensare alle espressioni «cultura dell'evidenza» o «società globale basata sull'evidenza» (Donaldson, 2009) o, più specificamente in ambito educativo, al movimento dell'*Evidence Based Education* (EBEd).

Considerato che un aspetto costitutivo della valutazione è «l'attribuzione di un giudizio di valore» (per es. Visalberghi, 1955; Gattullo, 1967; Scriven, 1991; Domenici, 1993) e che a nostro avviso una possibilità da non sottovalutare è che le evidenze non sono garanzia di democrazia ma possono «rappresentare la giustificazione per politiche educative autoritarie, anche con riduzione di spazi di autonomia e di pluralismo» (Calvani, 2013, p. 96), l'obiettivo di questo contributo è quello di cercare di mettere in luce i principali elementi del dibattito sulla valutazione basata sulle evidenze cercando di allargarlo non solo al contesto europeo ma anche a quello americano.

Come abbiamo più volte ribadito, un aspetto a nostro avviso importante consiste nel sottolineare le varie fasi/operazioni individuabili nella valutazione, che andrebbero possibilmente tenute distinte e non «schiacciate» sull'operazione finale di espressione di un giudizio di valore. D'altro canto, il fatto che la valutazione richiami degli aspetti valoriali, e pure degli schemi concettuali sottesi, ci induce a evidenziare che essa non può essere identificata soltanto con la sua dimensione tecnica o metodologica. Facendo riferimento, per esempio, alla valutazione di un progetto scolastico innovativo, non possiamo trascurare il quadro dei valori educativi presi implicitamente o esplicitamente a riferimento (Gattullo, 1967) e dunque non possiamo disgiungerla «da una più generale visione della scuola pubblica e dalla sua intenzionalità formativa, e, più in generale, dalla visione della società» (Giovannini & Marcuccio, 2012, p. 17). Sebbene il termine valore costituisca il nucleo del concetto di valutazione, è però, come afferma House (2003, p. 14), «la natura dei valori e delle affermazioni di valore che resta controversa [...] anche tra i principali teorici della valutazione».

Continuando a far riferimento, a titolo esemplificativo, alla tematica della valutazione dei percorsi e delle innovazioni educative, si può affermare che sono stati compiuti dei passi in avanti rispetto alla situazione della seconda metà degli anni '80 (Becchi & Vertecchi, 1986; Giovannini, 1988); tuttavia, non possiamo parlare, ancora, di una articolata e *diffusa* cultura della valutazione. Inoltre, gli sviluppi che possono essere considerati tali in questo ambito – si pensi, per esempio, alla nascita e allo sviluppo del Sistema nazio-

nale di valutazione e alla più intensa partecipazione dell'Italia alle indagini internazionali – fanno riferimento a una specifica cultura della valutazione rispetto alla quale solo in pochi casi si è proposta una cultura della valutazione «alternativa» (per es. le reti di autovalutazione delle scuole). Per questo motivo, c'è il rischio che gli echi del dibattito sulla valutazione basata sulle evidenze che arrivano nel contesto italiano da altri Paesi siano percepiti come rilevanti senza una piena consapevolezza «esperienziale» delle differenze a livello di clima e motivi che li hanno generati e delle possibili implicazioni per i diversi ambiti della società. A questo contribuisce, e non poco, la retorica delle politiche educative presenti nei documenti delle istituzioni europee e nazionali, che «leggono» il fenomeno valutativo e la sua rilevanza alla luce delle logiche del *new public management* divenute ormai trasversali ai diversi contesti nazionali (Kushner & Norris, 2007). Il tratto denotativo nella valutazione dell'attribuzione del giudizio di valore comporta la distinzione, sostenuta da molte voci nell'ambito della valutazione dei progetti/programmi, tra ricerca e valutazione; proprio per questo nel presente contributo verrà analizzata la specificità del dibattito sulla valutazione basata sulle evidenze.

Nel processo di costruzione di quelle che vengono denominate evidenze, gli elementi empirici che si incontrano nella «realtà» possono diventare oggetto di attività tramite le esperienze personali che si fanno nella vita quotidiana o tramite le pratiche professionali oppure tramite le attività di ricerca scientifica. Tale processo di elaborazione produce degli «esiti» che consistono, in estrema sintesi, in enunciati dichiarativi o procedurali sulla realtà. Nel caso della ricerca, questi enunciati sono definibili, secondo le parole di Dewey (1938; trad. it., p. 18), come «asserzione giustificata» ossia come enunciati formulati seguendo un processo rigoroso e trasparente. A questo punto, gli *esiti* possono entrare all'interno di due processi distinti: il processo di formulazione del giudizio di valore oppure il processo decisionale. Il giudizio di valore, una volta espresso, può portare alla formulazione di «raccomandazioni» o può anch'esso essere utilizzato in un processo decisionale, una fase che però non rientra tra quelle della valutazione. Il tema dell'*Evidence Based Education* ci sembra spostato – secondo anche l'interpretazione di Calvani (2011) – sul processo decisionale.

In relazione al processo di costruzione al quale si è fatto riferimento, evidenziamo qui alcune domande presenti nel dibattito sulle evidenze: Che cosa si intende per evidenza? Che cosa costituisce un'evidenza? Quali elementi possono essere considerati evidenze? È il materiale empirico originario che può essere considerato un'evidenza? Oppure è l'esito della ricerca scientifica che può essere considerato un'evidenza? Se gli esiti di una ricerca scientifica non sono già di per sé evidenze, è necessario che essi siano sottoposti a un qualche trattamento ulteriore? In quest'ultimo caso, a quale elaborazione devono essere sottoposti i dati per potersi trasformare in una evidenza?

Cosa intendiamo per valutazione basata sulle evidenze? Una valutazione che utilizza gli esiti di una ricerca empirica per formulare il proprio giudizio di valore? Oppure una valutazione che nel processo di formulazione del giudizio rende trasparente tutto il percorso di costruzione del giudizio stesso esplicitando in che modo gli esiti di ricerca siano interconnessi con altri elementi (per es. esigenze degli *stakeholders*, criteri, standard)?

## 2. ALCUNI MOMENTI DEL DIBATTITO SU «EVIDENCE BASED» E VALUTAZIONE NEGLI STATI UNITI D'AMERICA

Nel cercare di rispondere agli interrogativi proposti, non si può fare a meno di considerare lo sviluppo negli Stati Uniti del movimento culturale dell'*Evidence Based Education* e dell'*Evidence Based Evaluation* (EBEv), che hanno fatto dello slogan *What works* il *Leitmotiv* del rapporto tra politiche educative, ricerca e valutazione.

Il dibattito circa il metodo valutativo da privilegiare per supportare le scelte politiche si riaccende ciclicamente nei contesti anglofoni, e statunitense in particolare. Dapprima si è assistito al confronto tra la posizione della valutazione a supporto di «riforme come esperimenti» (Campbell, 1969) e quella che criticava l'uso dei metodi di ricerca da laboratorio per supportare le riforme politiche (Cronbach *et al.*, 1980); poi si è scatenata la «guerra dei paradigmi» (Guba, 1990; Scriven, 1993; Datta, 1994), espressione che racchiude l'animato dibattito sull'uso dei metodi quantitativi e qualitativi nell'ambito della valutazione che di fatto non si è ancora spento, anche se i toni polemici si sono attenuati.

Negli ultimi due decenni il dibattito si è riaperto in relazione al nascere e allo svilupparsi del movimento dell'*Evidence Based Education* che ha catalizzato attorno a sé alcuni dei principali temi già oggetto di riflessione da anni nell'ambito della comunità della ricerca empirica e della valutazione in educazione, di cui però ha tenuto conto soltanto parzialmente. Esso è stato innescato dalle sollecitazioni fornite dalle scelte politiche del governo federale statunitense di vincolare l'allocazione delle risorse alla dimostrazione dell'efficacia degli interventi educativi tramite ricerche e valutazioni con disegni sperimentali o quasi-sperimentali (Julnes & Debra, 2007). Tale dibattito si è sempre più sviluppato a livello internazionale e può essere identificato con l'espressione *Evidence Based Evaluation* (Clarke, 2006). Si tratta di un aspetto da considerare all'interno di un movimento più generale che sembra voler sostituire l'utopia della «società sperimentale» (Campbell, 1991) con quella della «società globale basata sull'evidenza» (Donaldson, 2009).

La scintilla che ha riaperto il dibattito è stata fatta scoccare nel 2003 dal neo costituito Istituto di Scienze dell'Educazione afferente al Dipartimento dell'Educazione statunitense. È di quell'anno, infatti, la pubblicazione del documento che traduceva operativamente il *Reform Act* del 2002 di George Bush (U.S. Department of Education, 2003) individuando nell'esperimento classico e in alcuni disegni quasi-sperimentali i disegni di ricerca e di valutazione privilegiati per l'erogazione dei fondi pubblici. Nelle disposizioni federali del Dipartimento dell'Educazione degli Stati Uniti seguite alla pubblicazione della legge *No child left behind Act* del 2001 (U.S. Department of Education, 2002) viene richiesto, per esempio, che le iniziative innovative in ambito scolastico mostrino chiaramente forti prove di efficacia. Nelle disposizioni del dicembre 2003, predisposte al fine di aiutare a identificare e mettere in atto pratiche educative corroborate da prove rigorose (U.S. Department of Education, 2003, p. III), si legge:

I responsabili a livello statale e locale, e gli stessi educatori, devono scegliere tra una miriade di proposte per decidere quali interventi meritino considerazione per le loro scuole e per le loro classi. Molti di questi hanno visto forme di intervento, presentate con grande enfasi come capaci di produrre grandi vantaggi, venire e andarsene negli anni, lasciando dietro di sé poco di positivo e di cambiamento permanente.

Come criterio di scelta ufficiale di efficacia (*The gold standard*) è stato indicato l'uso di esperimenti con gruppo di controllo e campioni scelti casualmente (*Randomized Controlled Trials* – RCT), in grado di misurare la forza e l'ampiezza degli effetti riconducibili alle iniziative innovative e di assicurare l'affidabilità, validità, pertinenza e significatività dei risultati. Si tratta di una scelta rispetto alla quale ha manifestato non poche perplessità anche la comunità dei ricercatori in ambito educativo, in particolare quelli dell'American Educational Research Association (Schneider *et al.*, 2007).

Un mese prima, nello stesso anno (il 4 novembre 2003), in occasione del suo congresso annuale, l'Associazione americana di valutazione (AEA) aveva predisposto un documento (*AEA statement*) intitolato *Scientifically based evaluation methods*, in cui si prendeva una posizione netta contro le scelte del Dipartimento dell'Educazione. I contenuti principali di questo documento sono riassumibili nei seguenti punti:

- non solo gli esperimenti sono capaci di determinare conoscenze sui legami di causalità;
- non solo gli esperimenti, ma anche studi diversi e con metodi misti, sono capaci di dimostrare rigore scientifico;
- le decisioni politiche solide traggono beneficio da dati non solo sulla causalità ma anche sui fattori condizionali.

Una posizione però non unanimemente condivisa dal momento che, un mese dopo, alcuni membri senior dell'AEA elaborarono un documento di reazione (denominato *The not AEA statement*), nel quale esprimevano la loro posizione favorevole agli esperimenti e, nei casi dove questi non fossero possibili, ad alcuni disegni quasi-sperimentali.

Negli anni successivi il dibattito tra gli esperti sul concetto di evidenza nella valutazione è rimasto vivo ed è documentato da varie pubblicazioni (per es. Donaldson & Christie, 2005; Julnes & Debra, 2007; Schwandt, 2007; Chatterji, 2009; Donaldson *et al.*, 2009; Green, 2013; Hall, 2013; Mertens & Hesse-Biber, 2013), nelle quali si può individuare la proposta dei metodi misti e della credibilità dell'evidenza nella valutazione.

Per comprendere le linee del dibattito negli Stati Uniti, occorre considerare che il contesto normativo e quello organizzativo della valutazione sono a tal punto strutturati da giustificare le reazioni menzionate. Infatti, da un lato, si trova un governo che prende scelte che vanno ben oltre le linee di indirizzo entrando anche nel merito di scelte teoriche e metodologiche specifiche e rese vincolanti; dall'altro lato, il sistema della valutazione ha assunto forme di istituzionalizzazione e di professionalizzazione che lo rendono più facilmente vulnerabile alle scelte governative.

### 3. VOCI DEL DIBATTITO SULL'«EVIDENCE BASED EVALUATION» IN EUROPA

Per quanto riguarda il contesto europeo, ci limitiamo a prendere in esame solo i principali elementi che possono essere ricavati da alcuni provvedimenti europei che, per la loro influenza sulle possibili allocazioni di risorse economiche, possono essere assunti come un riferimento fondamentale.

Considerando i documenti europei in materia di politiche di ricerca in ambito educativo, che rivelano un interessante e consistente richiamo al movimento dell'*Evidence Based Education*, non sembra rintracciabile un'omogeneità di posizione circa le possibili scelte metodologiche da proporre. La Commissione Europea (Direzione Generale - DG Bilancio, 2004), per esempio, ha fatto proprie le indicazioni dell'OCSE non facendo alcun riferimento, nelle sue linee guida sulla valutazione, ai disegni sperimentali. La DG Occupazione, invece, ha riconosciuto che i disegni quasi-sperimentali sono il modo più importante per valutare gli effetti di un intervento (European Commission, 1999, p. 14). Un riferimento esplicito alla ricerca *evidence based* è presente nel rapporto della DG Istruzione e Cultura del 2007, *Towards more knowledge-based policy and practice in education and*

*training* (Commission of the European Communities, 2007), in cui emerge la proposta di creare agenzie e centri specifici di ricerca sulla base di priorità nazionali e di diffondere i risultati della ricerca con nuove modalità per informare e orientare le pratiche e le politiche educative. In egual modo troviamo riferimenti alle decisioni politiche in ambito educativo e formativo basate su evidenze anche nella Conclusione del Consiglio d'Europa del 12 maggio 2009 (Consiglio Europeo, 2009), ove si afferma che per contribuire all'elaborazione di politiche «fondate su elementi concreti» è necessario effettuare il «monitoraggio periodico dei progressi» realizzati per conseguire uno degli obiettivi strategici della nuova programmazione europea. Anche nel Rapporto finale del *Progetto EIPEE – Evidence Informed Policy in Education in Europe* è presente un riferimento al rapporto tra valutazione e decisioni basate su evidenze laddove si raccomanda di prevedere una valutazione di processo e di esito in tutte le nuove attività progettate al fine di accrescere l'uso dei risultati della ricerca nelle decisioni politiche (Gough *et al.*, 2011).

In ogni caso, sembra emergere una tendenza, soprattutto a partire dal nuovo secolo, a non vincolare in modo rigoroso le attività valutative a specifici modelli e/o metodi pur individuando nell'attività valutativa una funzione rilevante per promuovere una cultura delle decisioni basate su evidenze.

Anche tra i ricercatori e gli studiosi europei dell'ambito educativo è vivo il dibattito tra i sostenitori dell'approccio *evidence based* e coloro che esprimono posizioni critiche; questi ultimi non sempre rinnegano l'importanza delle «prove» nello sviluppo di nuove politiche educative o di nuove modalità di intervento, ma esprimono l'esigenza di salvaguardare modalità rigorose per la loro individuazione (Pawson, 2002; Thomas, 2004) e democratiche (per es. Biesta, 2007; Saussez & Lessard, 2009). La discussione è tuttavia incentrata soprattutto sull'*Evidence Based Education* e non verte in modo specifico sull'*Evidence Based Evaluation*.

#### 4. ASPETTI DEL DIBATTITO SULL'«EVIDENCE BASED EVALUATION» IN ITALIA

In Italia, il dibattito sull'*Evidence Based Education* sembra essersi avviato in modo strutturato a partire dal 2007, quando sulla rivista *Je-LKS: Journal of E-Learning and Knowledge Society* appaiono due articoli che mettono a fuoco questo argomento (Calvani, 2007; Ranieri, 2007). A partire da quella data i contributi cominciano ad intensificarsi (Calvani, 2011; Federighi, 2011; Margiotta, 2012; Cottini, 2013); su tale argomento nel 2012 esce anche una

monografia specifica (Calvani, 2012), mentre nel 2013 *Form@re* pubblica due numeri monografici<sup>1</sup>.

In relazione al tema specifico del rapporto tra valutazione ed *evidence based*, va segnalato il XV Congresso dell'Associazione Italiana di Valutazione (AIV) dal titolo *Politiche basate sull'evidenza (Evidence Based Policy) e valutazione. Dal mito alla realtà?* che si è svolto a Bari nei giorni 19 e 20 aprile 2012<sup>2</sup>, di cui però non sono ancora stati pubblicati gli Atti. Secondo gli organizzatori del congresso, in un contesto di profonda crisi economica, le scelte pubbliche devono essere effettuate con una maggiore oculatezza. Questo richiede ai decisori politici di avvalersi di «dati» e «informazioni valutative sempre più accurate e solide» per effettuare le proprie scelte. In questa direzione si muove anche l'Unione Europea, la cui politica di coesione per il periodo 2014-2020 vincola gli investimenti di risorse alla «disponibilità di dati ed informazioni valutative derivanti dalla applicazione di modelli e metodi solidi e difendibili»<sup>3</sup>. L'interrogativo proposto come fondamentale è il seguente: in che modo la valutazione è in grado di produrre tali evidenze «sia che si adotti una definizione ristretta di *evidence based evaluation*» (valutazione basata su approcci sperimentali) sia che se ne adotti «una accezione ampia» («valutazione degli effetti» – basata su «solide e rigorose evidenze empiriche» – effettuata anche attraverso «metodi non-sperimentali» e «non-standard»)»? Connessi a questo, sono alcuni degli altri interrogativi che emergono dall'analisi della situazione: «Esiste o meno in Italia un sistema decisionale orientato ai risultati? Come stimolare la domanda di valutazioni basate su solide modalità di costruzione di evidenze empiriche? Quali sono i prerequisiti di *governance* necessari alla EBP (*Evidence Based Policy*)? Quali metodi utilizzare e in quale modo al fine di offrire rigorose basi di evidenza alle valutazioni? Quali condizioni organizzative e di contesto sono necessarie per attivare domande di valutazione basate su solide evidenze empiriche? Come possono i committenti delle valutazioni controllare la qualità delle evidenze empiriche sulle quali si

---

<sup>1</sup> Nell'ambito dei servizi sociali, ricordiamo, tra l'altro, che nel 2012 appare un approfondimento monografico sul tema delle «Prove di efficacia nel lavoro a diretto contatto con le persone» nel volume 13(2) della rivista *Studi Zancan*. Si tratta dei contributi di un seminario internazionale che affrontano il tema dell'*evidence based practice* da diverse prospettive riportando anche esperienze realizzate in diverse nazioni.

<sup>2</sup> Il congresso è stato organizzato e realizzato in collaborazione con l'Università degli Studi di Bari (Dipartimento di Psicologia e Scienze Pedagogiche e Didattiche e Dipartimento di Agraria) e il Nucleo di Valutazione e Verifica degli Investimenti Pubblici della Regione Puglia.

<sup>3</sup> Le nuove norme e leggi che governano il prossimo ciclo di investimenti della politica di coesione europea per il periodo 2014-2020 sono state pubblicate sulla *Gazzetta Ufficiale dell'Unione Europea*, L 347, 20 Dicembre 2013, <http://eur-lex.europa.eu/JOHtml.do?uri=OJ:L:2013:347:SOM:IT:HTML>.

basano i giudizi che vengono loro consegnati? Quali esperienze di valutazione [...] possono essere indicate come esempi di valutazioni effettivamente *evidence based*?». Alcuni interrogativi sono relativi in modo specifico ai sistemi educativi, quali per esempio: «In che misura le esperienze più recenti di valutazione di sistema attive in Italia (Servizio Nazionale di Valutazione, indagini OCSE-PISA, indagini OCSE-PIACC, CPV Trento) sono utilizzabili in chiave di supporto a *Evidence Based Policy* nel campo dell'*education*? Quali elementi di metodo risultano di particolare interesse e quali invece sono carenti e da migliorare? A quali condizioni i casi 'di successo' si sono dimostrati tali? Come la 'politica' usa (o dovrebbe usare) i risultati di queste valutazioni? Come sono utilizzati dalle scuole? Quali ambiti analitici e valutativi andrebbero approfonditi e sviluppati per una valutazione puntuale ed approfondita del sistema educativo italiano nelle sue diverse articolazioni e rispetto ai suoi fenomeni caratterizzanti (diseguaglianza di opportunità, di accesso, di risultato; dispersione ed insuccesso)? In che senso la 'valutazione delle competenze' è – se lo è – *evidence based*?» (Bastianelli *et al.*, 2012).

Si tratta di interrogativi molto interessanti che possono sollecitare anche in Italia un dibattito già affrontato in altri contesti.

## 5. I TEMI DEL DIBATTITO SULL'«EVIDENCE BASED» NELL'AMBITO DELLA VALUTAZIONE

Come abbiamo visto in precedenza, il dibattito all'interno dell'*Evidence Based Evaluation* sembra essere focalizzato soprattutto attorno a questioni che rientrano prevalentemente nell'ambito metodologico. Tuttavia, come afferma Clark (2008, p. 307), quando i ricercatori devono affrontare il tema delle evidenze «devono affrontare numerosi temi chiave» ad esse relativi. Infatti, il concetto di evidenza è

legato da vicino alla nozione epistemologica di verità e di conoscenza. Intrinsecamente, vi è una profonda componente ad alto contenuto filosofico e valoriale quando si prende in considerazione l'evidenza e come e se l'evidenza è vista come tale in relazione a una visione del mondo, teoria o paradigma sottostante. Sebbene per i ricercatori sia importante essere esperti nei metodi, essi devono essere consapevoli di queste dimensioni epistemologiche sottostanti e riflettervi criticamente.

Nel contesto specifico della valutazione, Schwandt (2009, p. 197) amplia ulteriormente lo spettro delle dimensioni che toccano il dibattito sulle evidenze. Egli infatti afferma che

il dibattito originario spesso ha generato più fervore che chiarezza perché è stato troppo frequentemente focalizzato su questioni relative alle differenze circa i metodi piuttosto che sui temi più complessi sottostanti l'intersezione con l'epistemologia e la teoria sociale e politica della valutazione. Allo stesso modo, la riproposizione attuale del dibattito è sbagliata perché, ancora una volta, si concentra su questioni di metodo mentre ignora problemi più importanti che riguardano ciò che costituisce una evidenza e il suo uso responsabile.

Ecco perché egli avanza la proposta di una «teoria dell'evidenza per la valutazione» che sia in grado di trasformare il dibattito su ciò che costituisce una «evidenza credibile» in un dibattito «più educativo e chiarificatore» (*ivi*, p. 199).

Gli ambiti toccati dallo studioso statunitense per sviluppare la sua teoria sono i seguenti: (1) il carattere dell'evidenza; (2) i riferimenti valoriali dell'evidenza; (3) il contesto di applicazione dell'evidenza; (4) la natura della razionalità e dell'argomentazione. Noi riteniamo importante aggiungere un'altra dimensione, strettamente connessa a quelle sopra richiamate, che è il sistema organizzativo deputato a generare evidenze, ossia l'insieme delle organizzazioni (pubbliche o private) che hanno come obiettivo quello di costruire evidenze per i decisori.

Svilupperemo di seguito questi aspetti, avvalendoci sia della riflessione di Schwandt, sia di altri autori che hanno toccato gli stessi temi, anche se in modo non così sistematico.

### 5.1. *La distinzione tra dati ed evidenze: non solo questioni terminologiche*

Un aspetto importante, quando si cerca di definire il concetto di evidenza sia in relazione all'attività di ricerca in ambito educativo sia a proposito di valutazione, è la distinzione tra il concetto di *dato* e quello di *evidenza*.

Nell'ambito della riflessione sulla metodologia della ricerca in campo sociale esistono diverse posizioni circa il concetto di *dato*. Corbetta (1999, p. 83), per esempio, distingue tra *dati* e *informazioni*: «In genere si dà il nome di *informazioni* ai materiali empirici grezzi non ancora sistematizzati, e quello di *dati* agli stessi materiali organizzati in modo tale da poter essere analizzati». In Kerlinger e Lee (2000, p. 192), invece, si afferma che con il termine «dato», così come viene usato di solito nella ricerca sul comportamento umano, si intendono i «risultati di ricerca da cui sono ricavate inferenze». Di solito si tratta di «risultati numerici, come i punteggi di un test o statistiche come le medie, le percentuali e i coefficienti di correlazione». Tuttavia, i dati possono essere più di questo, cioè «informazioni tratte dai giornali o dagli articoli di riviste, da materiali biografici, diari e così via – insomma, materiale verbale in generale. In altre parole, il termine 'dati' è un termine generale con mol-

teplici significati». Per Marradi (2002, p. 10) il dato è «il valore simbolico assegnato allo stato di uno specifico caso su una specifica proprietà sulla base delle convenzioni e regole stabilite dalla definizione operativa». Si tratta di valori simbolici espressi solitamente con delle cifre numeriche o, talvolta, con lettere. Miles e Huberman (1994, p. 9) affermano che «in qualche senso tutti i dati sono qualitativi: essi si riferiscono alle essenze di persone, oggetti e situazioni [...]. Noi abbiamo un'esperienza «grezza» che è in seguito convertita in parole [...] o in numeri [...]. Noi ci focalizziamo sui dati nella forma di parole, cioè di linguaggio nella forma di testo esteso [...]. Le parole di basano su osservazioni, interviste o documenti [...] tali dati di solito non sono immediatamente accessibili per l'analisi, ma richiedono una qualche elaborazione (*processing*)».

Se passiamo ad analizzare il concetto di *evidenza*, ci troviamo in qualche misura nella medesima situazione: anche in questo caso esso non viene definito in modo univoco (Calvani, 2011). Tuttavia, possiamo individuare due tendenze interpretative del concetto.

La prima che identifica l'evidenza con «il presentarsi o manifestarsi di un oggetto qualsiasi come tale» (Abbagnano, 1993, pp. 369-370). Si tratta di una concezione «oggettivistica» dell'evidenza che, dopo la parentesi soggettivistica cartesiana, è stata riaffermata nella cultura contemporanea dalla fenomenologia di Husserl secondo cui l'evidenza è tornata a designare il suo significato originario del presentarsi o manifestarsi di un oggetto come tale, qualunque sia l'oggetto e quali che siano i metodi con cui s'intende certificare o garantire la sua presenza o manifestazione. È questo anche il significato che, nel linguaggio comune, viene dato al termine. In questa accezione, quindi, il termine evidenza sembra coincidere con quello di «dato» inteso come «materiale empirico» primario o grezzo.

Il secondo filone, invece, individua l'evidenza come un elemento che supporta (comprova) la validità di una determinata affermazione o decisione. In questo caso, evidenza sembra sinonimo di «prova» oppure «argomento a sostegno di». Troviamo questa accezione del termine, per esempio, in Clark (2008, p. 307), il quale afferma inoltre che i «concetti di evidenza sono inseriti attraverso tutta la ricerca qualitativa: dalla base per l'indagine, attraverso i processi di ricerca, ai prodotti finali della ricerca qualitativa [...]. La determinazione di ciò che conta come evidenza, chi o cosa determina questo, e su quali basi filosofiche questo viene fatto rimane oggetto di dibattito».

Anche Schwandt (2009, p. 199) ritiene che, da un punto di vista epistemologico, il termine evidenza significhi «un'informazione che riguarda se una convinzione o proposizione sia vera o falsa, valida o non valida, giustificata o non dimostrata». Tuttavia egli ritiene che tale informazione non vada limitata al concetto di «efficacia causale» ma ampliata per essere adeguata e

utilizzabile all'interno di una teoria dell'evidenza per la valutazione. Qui egli introduce un nuovo ambito di riflessione sul concetto di evidenza: quello della *spendibilità* o utilizzabilità che è stato ripreso anche da Calvani (2013, p. 92) quando parla di evidenza come di una «conoscenza affidabile e spendibile nelle politiche e nelle pratiche operative»<sup>4</sup>.

Al fine di una chiarezza concettuale, riteniamo che sia più opportuno distinguere tra i «dati», intesi come materiale empirico iniziale «grezzo», raccolto attraverso adeguati strumenti, e i «risultati» delle analisi ed elaborazioni dei dati, così come le «informazioni» che si possono ricavare dalle analisi dei dati, e gli «esiti» della ricerca ossia il prodotto finale delle inferenze (di tipo logico) che possono essere ricavate.

Calvani (2011, p. 78), invece, utilizza l'espressione «conoscenza affidabile» come sinonimo di conoscenza fondata o *evidence based*, espressione intesa nell'accezione lata (basata su evidenze empiriche rigorosamente raccolte) senza entrare nel dettaglio delle metodologie e dei criteri impiegati.

Anche a proposito dell'espressione «basata su», che spesso accompagna il termine evidenza, esiste un problema terminologico e concettuale. Rinvia infatti a un «qualcosa» che «si basa» sulle evidenze. Che cos'è questo qualcosa? Nelle espressioni più ricorrenti appaiono i termini di educazione (*Evidence Based Education*), politica (*Evidence Based Policy*), pratica (*Evidence Based Practice*) o valutazione (*Evidence Based Evaluation*).

Limitandoci all'ambito della valutazione, chi ha contribuito a elaborare una logica della valutazione mette in evidenza come siano molteplici gli elementi che entrano nel processo complessivo della valutazione (rilevazione dell'oggetto da valutare o evaluando e formulazione del giudizio di valore). Pertanto, utilizzare l'espressione «basata su» a proposito della valutazione sembra attribuire un ruolo unico o prevalente nel processo valutativo alle evidenze. «Basarsi su» qualcosa può infatti veicolare la convinzione che quel «qualcosa» su cui ci si basa sia l'elemento fondamentale. Ma i modelli valutativi ci dicono che non è così. Pertanto ecco l'esigenza di introdurre una modifica nell'espressione da cui siamo partiti e «ammorbidirla» (Calvani, 2011) con espressioni che intendono significare la molteplicità di fattori che entrano in gioco nel processo valutativo: valutazione «informata» dalle o valutazione «consapevole» delle evidenze (*evidence informed* o *evidence aware*).

Se si accetta il carattere «costruttivo» e «contestuale» dell'evidenza, appare chiaro che l'espressione «basata sull'evidenza» o «base di evidenza» «debba essere interpretata con cautela: affermare che l'evidenza possa apparire

---

<sup>4</sup> Nei contesti francofoni, il termine inglese *evidence* viene tradotto con *donnée probante* e l'espressione *evidence based* con *fondée/basée sur des données probantes*. Si tratta di un'espressione il cui significato si avvicina a quello di «prova».

come importante nella nostra decisione di agire è una cosa; affermare che essa sia il fondamento della nostra azione è un'altra» (Schwandt, 2009, p. 203).

Pensiamo che tali riferimenti siano utili a farci riflettere sul significato concettuale dei diversi termini.

## *5.2. L'apparente «evidenza» delle evidenze: il carattere costruttivo dell'evidenza*

Se escludiamo la concezione dell'evidenza come «dato empirico originario» – così come concepito per esempio da Dewey (1938) – tutte le altre concezioni sottolineano come le evidenze siano elementi finali di un processo di costruzione. In tale accezione, l'evidenza perde – e deve perdere – quell'alone di immediatezza, di elemento che appare per folgorazione che è sempre presente nell'immaginario collettivo.

L'evidenza deve essere intesa come un elemento che, costruito attraverso un processo rigoroso, viene riconosciuto da una comunità scientifica come un dato «accettabile», quindi attendibile e valido. A questo proposito va fatta una distinzione tra quattro elementi: (1) il processo che porta alla costruzione dell'«enunciato» (attraverso l'esperienza, la pratica o la ricerca); (2) il processo entro il quale viene utilizzato l'«enunciato»; (3) il contenuto dell'enunciato; (4) il numero di conferme del medesimo enunciato.

Ci chiediamo: quale di questi quattro aspetti fa diventare quell'enunciato una «evidenza»? Possiamo parlare di evidenza solo nel caso di un esito di ricerca metodologicamente rigoroso? Solo quando viene utilizzato, per esempio, un disegno sperimentale<sup>5</sup>? Solo quando quel dato viene assunto all'interno di un processo decisionale in risposta a specifici requisiti di spendibilità? Possiamo parlare di evidenza solo quando l'enunciato descrive un rapporto causale? Oppure quando quell'enunciato si ripresenta in modo costante in più di una «esperienza» o in più di una ricerca, come per esempio nel caso delle meta-analisi o delle rassegne sistematiche?

Il processo più importante che deve subire un enunciato per essere trasformato in evidenza è di certo quello che Schwandt (2009, p. 201) chiama «interpretazione» e che non è altro che un momento importante del processo di «costruzione» dell'evidenza. Infatti, un elemento di per sé non determina la verità o falsità di una proposizione. È necessario che questo elemento – in relazione alla proposizione di cui si deve stabilire la verità o meno, ma anche

---

<sup>5</sup> Circa gli aspetti metodologici, Scriven (2009, pp. 137-138) si chiede anche: la valutazione deve focalizzarsi solo sull'individuazione di legami causali tra gli *outcome* (impatto) e l'intervento realizzato (variabile indipendente)? L'individuazione dei legami causali è possibile solo attraverso l'esperimento con gruppo di controllo?

alle circostanze e alla pratica rispetto alla quale è presa in esame – risponda ai seguenti criteri: (1) rilevanza (l'informazione riguarda direttamente l'ipotesi o l'enunciato in questione?); (2) credibilità (possiamo credere all'informazione?); (3) forza probativa o inferenziale (con quale forza l'informazione porta all'enunciato o all'ipotesi considerata?). La rilevanza è collegata al tipo di interrogativo a cui desideriamo rispondere. La credibilità, invece, all'accordo sulle regole metodologiche per generare l'informazione. La forza probativa all'argomentazione o al consenso razionale.

A questo punto, però, è necessario fare una distinzione tra *evaluative evidence* o *evaluation evidence* (Gargani & Donaldson, 2011) ed *evidence based evaluation* (Clarke, 2006; Stake, 2009) che a nostro avviso non sempre viene fatta in letteratura. Essa scaturisce dal considerare la valutazione come un processo all'interno del quale occorre distinguere il momento della «rilevazione dei dati» per la determinazione dell'oggetto da valutare da quello della valutazione dell'«oggetto» con il ricorso a criteri e standard. Possiamo utilizzare l'espressione *evidence based evaluation* nel momento in cui il processo di formulazione del giudizio di valore prevede una preliminare fase di raccolta di «evidenze empiriche» – costruite seguendo i criteri metodologici della ricerca così come definiti dalla comunità scientifica di riferimento – e l'utilizzo di questo tipo di «evidenze» per la costruzione del giudizio di valore. Parliamo invece di «evidenza valutativa» a proposito del prodotto finale del processo di formulazione del giudizio di valore, ossia del giudizio stesso. Tuttavia, per riuscire a specificare e identificare meglio questi aspetti, riteniamo che – riprendendo alcune sollecitazioni presenti in Hurteau *et al.*, 2012 – possa essere più utile utilizzare l'espressione «valutazione basata su evidenze e sulla credibilità del giudizio». Con tale espressione, infatti, si vuole sottolineare come la fase di rilevazione/misurazione sia effettuata seguendo i criteri metodologici della ricerca empirica e la fase di valutazione sia effettuata secondo i criteri sopra richiamati (Schwandt, 2009) a cui aggiungiamo quelli della trasparenza e della partecipazione democratica (Green, 2009).

Una simile concettualizzazione viene espressa anche da Schwandt (2009, p. 209), laddove fa una distinzione tra una «evidenza credibile» all'interno del processo valutativo e una «valutazione credibile». Nel processo di formulazione del giudizio di valore (di un programma, progetto o politica) alcuni degli enunciati sono enunciati fattuali; altri invece sono giudizi di valore, altri definizioni, principi, spiegazioni causali, raccomandazioni operative. Ciò che costituisce una «valutazione credibile» dipende anche da come ciascuno dei diversi tipi di enunciato viene spiegato e difeso e pure da come tutti i diversi tipi di enunciato vengono organizzati all'interno di una argomentazione. Lo sviluppo di «evidenze credibili» in una valutazione, quindi, non è sufficiente per stabilire la credibilità di una valutazione.

### 5.3. I riferimenti valoriali dell'evidenza

La dimensione valoriale contraddistingue trasversalmente tutti i vari momenti che attraversa l'evidenza valutativa (il giudizio di valore credibile e basato su evidenze): produzione, interpretazione, disseminazione e uso (Schwandt, 2009).

Nella fase di *raccolta, produzione e interpretazione* degli elementi utili a formulare il giudizio di valore, il soggetto che valuta deve attentamente «valutare» se essi siano elementi fraudolenti, privi di significato oppure fuorvianti sia che si tratti di «elementi empirici», costruiti attraverso procedure di ricerca, sia che si tratti di criteri valutativi e standard, individuati attraverso processi di tipo democratico. Inoltre, il valutatore deve attentamente conoscere e rispettare le regole dell'argomentazione nel costruire il giudizio di valore per evitare di cadere in argomentazioni fallaci.

Nella fase di *disseminazione* delle evidenze valutative, la dimensione valoriale emerge nel momento in cui si prendono in esame soprattutto le caratteristiche dei destinatari del giudizio e il particolare uso per cui è stato elaborato il giudizio di valore. In ragione delle caratteristiche dei destinatari sarà necessario ponderare in che modo rendere conto della complessità dell'evidenza valutativa e del grado di credibilità della logica con cui è stato costruito il giudizio di valore al fine di renderlo comprensibile nelle sue varie dimensioni. Inoltre il giudizio di valore, una volta formulato, è in sé carico di aspetti valoriali che aumentano se si considera che esso ha la possibilità di convincere qualcuno.

Nella fase di uso delle evidenze valutative da parte di coloro che non sono la fonte delle evidenze utilizzate, emerge il problema di quali siano le responsabilità del valutatore circa il supporto per far comprendere la portata qualitativa delle evidenze prese in esame all'interno del processo decisionale.

Esiste, tuttavia, anche un piano *meta-processuale* (ossia che si pone su un piano preliminare rispetto al processo complessivo di valutazione) che è anch'esso intriso di riferimenti valoriali e che non può non essere preso in considerazione riflettendo sul piano di una logica della valutazione. Tutti e due i principali processi dell'intero percorso valutativo (rilevazione dei dati e formulazione del giudizio di valore) sono caratterizzati da una serie di decisioni che vengono prese sulla base di presupposti di vario tipo, sia teorici sia valoriali.

### 5.4. Il contesto di utilizzo dell'evidenza

Un terzo elemento che non dobbiamo assolutamente dimenticare è che la produzione e l'utilizzo dell'evidenza valutativa sono processi che avvengono

all'interno di una precisa cornice teorico-valoriale e con un particolare sistema di procedure e strumenti, come abbiamo visto in precedenza, ma sempre collocati in un ben preciso contesto sociale caratterizzato in primo luogo dalle esigenze, dai bisogni, dagli interessi, dagli interrogativi e dagli obiettivi dei diversi soggetti coinvolti. In una prospettiva democratica della valutazione questo implica che, oltre a prendere in considerazione i diversi punti di vista dei soggetti coinvolti nel processo valutativo e decisionale, dovrà essere prestata attenzione al coinvolgimento e alla partecipazione dei soggetti più deboli anche nei diversi momenti deliberativi così da garantire un equo equilibrio delle forze (Green, 2012). Possiamo quindi affermare che il giudizio di valore, elaborato con il contributo di evidenze empiriche e attraverso procedure logico-argomentative tali da renderlo credibile, è un prodotto complesso frutto di elementi teorici, valoriali, metodologici ma anche contestuali.

### *5.5. Il processo di formulazione del giudizio di valore e il rischio di riduzionismi*

Un altro aspetto che è opportuno prendere in considerazione nel momento in cui si cerca di delineare una cornice teorica che possa fare da sfondo al fenomeno delle evidenze in valutazione è quello relativo ai processi di ragionamento (razionalità) e di argomentazione che vengono utilizzati sia durante il processo di formulazione del giudizio di valore sia nei diversi contesti di discorso in cui le evidenze vengono utilizzate.

A questo proposito è necessario fare una precisazione terminologica relativa al concetto di «decisione». Se facciamo nostra la seguente definizione di presa di decisione: «confrontare due o più opzioni rispetto a un certo numero di caratteristiche ed esprimere una preferenza riguardo all'opzione che, più delle altre, sembra soddisfare i criteri di accettabilità stabiliti dal decisore» (Rumiati, 2006), ci accorgiamo che il processo di decisione deliberativo si differenzia in parte da quello di formulazione del giudizio di valore proprio del percorso di valutazione.

Rifacendoci a Scriven (1991, p. 120), nel contesto della formulazione del giudizio di valore le evidenze – intese come l'insieme dei dati empirici raccolti – rivestono un ruolo fondamentale, poiché vengono a costituire l'oggetto che deve essere valutato. Tuttavia esse non sono i soli elementi che vengono considerati durante il processo di formulazione del giudizio. Infatti una funzione fondamentale viene svolta dai criteri valutativi e dagli standard.

Nel momento della decisione deliberativa (ossia un processo che porta alla decisione circa la scelta di un'azione da compiere nei confronti di un X già valutato), invece, le evidenze che entrano in gioco per supportare la decisione possono essere anche i giudizi di valore già formulati, ossia le

evidenze valutative. Anche in questo caso, tuttavia, le evidenze valutative non sono i soli elementi che entrano nell'elaborazione della decisione. Uno spazio importante hanno anche altri fattori. In primo luogo il contesto della decisione. Un elemento contestuale, per esempio, è lo scopo per cui si vuole prendere una decisione. I processi decisionali sono sempre interconnessi con uno scopo. Per esempio, in ambito educativo una situazione è quella di decidere se attivare o meno una sperimentazione di tipo innovativo, un'altra, invece, è quella in cui si vuole «applicare» un determinato dispositivo didattico per diffonderlo a fronte degli esiti positivi riscontrati in altre situazioni.

Il rischio che si corre è quello di fare propria un'idea riduttiva del processo decisionale, ossia di sminuire la complessità degli elementi e delle loro relazioni. Altro rischio, inoltre, è quello di proporre una concezione del processo decisionale di tipo razionale. Oggi le ricerche in questo ambito ci dicono che il processo decisionale è sottoposto a una complessità di elementi e che la «razionalità assoluta» deve essere sostituita da una «razionalità limitata» che meglio sembra descrivere i reali processi decisionali anche nell'ambito dell'educazione.

### *5.6. Gli aspetti organizzativo-culturali della costruzione delle evidenze*

Il movimento culturale che ruota attorno all'*evidence based* non è semplicemente relegato a un dibattito che si sviluppa sulle riviste o attraverso pubblicazioni di vario genere. Vi è anche un risvolto di tipo organizzativo che va attentamente preso in esame per le implicazioni di tipo culturale che può avere sul sistema educativo e scolastico.

Possiamo infatti individuare specifiche istituzioni organizzative (centri, associazioni ecc.) – con relative risorse umane, economiche e strumentali – che si dedicano a perseguire l'obiettivo di fornire le conoscenze concrete necessarie a sviluppare processi decisionali efficaci ai diversi livelli cercando in tal modo di individuare ambiti di connessione tra ricerca e pratiche educative (Calvani, 2007; Bruni & Vivanet, 2013).

Ciascuna di queste organizzazioni ha elaborato documenti di tipo metodologico per individuare i criteri da utilizzare al fine di predisporre rassegne di studi.

A livello internazionale ricordiamo, per esempio, la Campbell Collaboration Education Coordinating Group, uno dei sei gruppi coordinati che operano all'interno della Campbell Collaboration, un *network* internazionale di ricerca che dal 2000 ha avviato un'intensa attività di rassegne sistematiche e sintesi quantitative circa gli effetti di interventi sociali in diversi ambiti

(educazione, criminologia, giustizia, sviluppo internazionale, servizi sociali)<sup>6</sup>. Nelle diverse edizioni delle linee guida per la preparazione di protocolli di rassegne si afferma che le evidenze possono derivare da studi che utilizzano vari metodi, ma l'orizzonte di riferimento resta sempre quello della ricerca quantitativa (The Campbell Collaboration, 2001; Shadish & Myers, 2004). In questi documenti non viene mai esplicitamente fatto riferimento a studi di tipo valutativo.

Negli Stati Uniti, nel 2002, ha iniziato ad operare il What Works Clearinghouse (WWC)<sup>7</sup> che intende costituire una risorsa per i processi decisionali supportati da informazioni in ambito educativo. Per raggiungere questo obiettivo, il WWC identifica studi che forniscono evidenze credibili e affidabili dell'efficacia di una determinata pratica, programma o politica e ne diffonde rassegne (sistematiche, brevi e di singoli casi) e relazioni sul proprio sito web insieme anche a guide pratiche rivolte agli educatori. Nel manuale di procedure e standard del marzo 2014 (What Works Clearinghouse, 2014) viene precisato che gli studi che possono essere inclusi nei processi di sintesi sono solo quelli che presentano un disegno sperimentale (*Randomized Controlled Trials* – RCT) o quasi sperimentale. In questo documento tra i diversi tipi di disegni degli studi che possono essere selezionati sono inseriti esplicitamente anche i disegni valutativi.

In Europa, l'istituzione più nota è denominata Evidence for Policy and Practice Information and Co-ordinating Centre (EPPI)<sup>8</sup>, la cui attività è stata avviata fin dal 1993 in ambito sanitario; solo nel 2000 è stata aperta un'area specifica sull'educazione (Torgerson *et al.*, 2001; Oakley *et al.*, 2005; Ranieri, 2007). Due sono gli obiettivi del centro: da un lato, realizzare rassegne sistematiche e sintesi di ricerche; dall'altro, condurre studi sull'uso delle

---

<sup>6</sup> Cfr. <http://www.campbellcollaboration.org/ECG/Education/index.php>. Sempre a livello internazionale, ricordiamo che nel 2003 il Centre for Educational Research and Innovation (CERI) dell'OCSE ha realizzato il Progetto «Evidence-based policy research in education» che si è sviluppato attraverso una serie di quattro *workshop* tenutisi dal 2004 al 2007 a cui hanno partecipato esperti nella ricerca e nelle politiche da vari paesi. L'esito di questo lavoro è stato pubblicato nel volume *Evidence in education: Linking research and policy* (OECD, 2007). In questo rapporto da un punto di vista metodologico si è optato per il superamento della contrapposizione tra metodi quantitativi e qualitativi e per considerare gli esperimenti controllati randomizzati (RCT) «come una delle metodologie appropriate per la ricerca educativa» (p. 23).

<sup>7</sup> È sorta come iniziativa amministrata dal National Center for Education Evaluation and Regional Assistance (NCEE) nell'Institute for Education Sciences (IES) all'interno del Department of Education (Gersten & Hitchcock, 2009). Per approfondimenti cfr. <http://ies.ed.gov/ncee/wwc/default.aspx>.

<sup>8</sup> L'EPPI è parte dell'Unità di Ricerca delle Scienze Sociali dell'Istituto di Educazione dell'Università di Londra. Cfr. <http://eppi.ioe.ac.uk/cms>.

ricerche sia a livello personale sia a livello di politiche. Le rassegne sistematiche realizzate da EPPI seguono un approccio che, a differenza di quelli adottati dalle altre organizzazioni, fa propria una concezione ampia del concetto di evidenza (elementi quantitativi e qualitativi), degli approcci (predefinito o emergente/iterativo) e dei metodi per realizzare le sintesi di ricerche (sintesi concettuale narrativa, sintesi meta-etnografica, meta-analisi statistica) (EPPI, 2010)<sup>9</sup>. Tra i diversi tipi di ricerca che vengono inseriti nell'elaborazione di ricerche sistematiche si fa esplicito riferimento anche alle ricerche valutative, ma solo di tipo sperimentale.

Al di là di questa carrellata di centri, occorre guardare queste esperienze con un'ottica critica. In primo luogo è necessario chiedersi: dal punto di vista istituzionale quali sono le caratteristiche del soggetto che gestisce queste organizzazioni? chi concretamente opera queste rassegne e studi? quali bisogni/interessi vengono presi in considerazione nella fase di elaborazione degli studi di sintesi? quale concezione di evidenza è fatta propria da chi realizza le rassegne? qual è la consapevolezza del dibattito sulla distinzione tra ricerca e valutazione? qual è la riflessione sulla complessità del processo che conduce alla formulazione del giudizio di valore? in che misura, nonostante le aperture presenti nell'approccio dell'EPPI, si tiene conto anche del dibattito metodologico sulle strategie di ricerca diverse da quelle dell'esperimento controllato randomizzato per individuare i nessi causali (il *What works*) (Scriven, 2009)?

Dai dati raccolti, l'impressione che si ricava è che nei documenti metodologici predisposti dalle diverse organizzazioni, tutta la complessità che circonda la riflessione sul rapporto tra evidenze e valutazione non sia stata adeguatamente valorizzata. E questo a nostro avviso non agevola la diffusione di un'adeguata cultura della valutazione in considerazione del fatto che questi centri, pur non avendo un diretto controllo sulla produzione di ricerche e di valutazioni di qualità, di fatto svolgono un'influenza indiretta circa le scelte metodologiche dei ricercatori e dei valutatori mediante la diffusione degli standard metodologici privilegiati nel selezionare e analizzare le ricerche utilizzate per le rassegne (OECD, 2007).

### *5.7. Quale rilevanza di questo dibattito per il contesto italiano?*

Nell'analisi del rapporto tra ricerche, pratiche, politiche e valutazioni educative nel contesto internazionale, abbiamo visto che questo dibattito si è sviluppato in modo particolare, anche se non esclusivo, negli Stati Uniti. In

---

<sup>9</sup> Per gli aspetti metodologici che riguardano le rassegne sistematiche, l'EPPI si rifà anche a Gough *et al.*, 2011.

quel contesto, negli ultimi decenni, si è assistito alla nascita, allo sviluppo e alla progressiva istituzionalizzazione del movimento culturale dell'*evidence based education* ed *evaluation* che ha fatto dello slogan *What works* il *Leitmotiv* del rapporto tra politiche educative e mondo della scuola, della ricerca e della valutazione.

Il contesto in cui si sta sviluppando questa riflessione in Italia è molto diverso da quello statunitense. In comune, di certo, vi sono tendenze che caratterizzano la gestione delle pubbliche amministrazioni e delle risorse economiche che ormai sono diventate «globali». Basti pensare al bisogno di migliorare la qualità dei servizi erogati e di dimostrare un utilizzo efficace delle risorse utilizzate (responsabilità sociale e rendicontativa). Ma non troviamo nel nostro paese – neanche di riflesso dall'Europa – scelte di politica educativa così vincolanti come quelle del Dipartimento dell'Educazione statunitense, anche se va ricordata l'influenza sul sistema scolastico e formativo delle scelte alla base dei Fondi strutturali europei.

Tuttavia pensiamo che accogliere nelle nostre riflessioni sulla realtà della valutazione i diversi aspetti del dibattito internazionale, che abbiamo cercato di delineare nelle sue linee generali nelle pagine precedenti, possa favorire lo sviluppo della nostra cultura della valutazione in varie direzioni. Da un lato, *prevenendo* scelte che potrebbero indirizzare – anche attraverso il filtro dei documenti a livello europeo – la valutazione verso forme riduttive (ossia non in grado di tenere in considerazione la complessa serie di elementi strettamente interconnessi, come ad es. i valori di riferimento a livello sociale e teorico) e non pienamente democratiche. Dall'altro, *rilanciando* il dibattito sull'importanza che «l'impianto di valutazione sia internamente coerente, costruito in base a una profonda conoscenza della cultura valutativa per rispondere alle esigenze della situazione specifica e che i valutatori esplicitino e argomentino sempre i presupposti delle proprie scelte e del proprio agire» (Giovannini & Marcuccio, p. 23). Infine, *richiamando* i decisori politici a una complessità del momento decisionale che, sebbene successivo al processo valutativo (Gattullo, 1986), tenga conto non solo di scelte ideologiche o di suggestioni personali ma anche di «elementi» valutativi che – inseriti nella cornice del «paradigma della ricerca sociale» (Christie & Fleischer, 2009) – siano riconoscibili dagli altri (credibili) e, in quanto tali, anche apprezzabili per il loro valore.

## RIFERIMENTI BIBLIOGRAFICI

- Abbagnano, N. (1993). *Dizionario di filosofia*. Milano: TEA.
- Bastianelli, M., Mineo, S., & Roma, F. (2012). *Policy e valutazione delle competenze: l'indagine ISFOL-OCSE «PLAAC»*. Presentazione al XV Congresso dell'AIV – Associazione Italiana di Valutazione *Politiche basate sull'evidenza (Evidence Based Policy) e valutazione. Dal mito alla realtà?* Prelevato il 27 Febbraio 2014 da <http://isfoloia.isfol.it/handle/123456789/224>.
- Becchi, E., & Vertecchi, B. (a cura di). (1986). *Manuale critico della sperimentazione e della ricerca educativa*. Milano: Franco Angeli.
- Biesta, G. (2007). Why «What works» won't work: Evidence-based practice and the democratic deficit in educational research. *Educational Theory*, 57(1), 1-22.
- Bruni, F., & Vivant, G. (2013). Evidence based education: centri di ricerca e risorse in Rete. *Form@re*, 13(2), 102-106.
- Calvani, A. (2007). Evidence-Based Education: ma «funziona» il «che cosa funziona»? *Journal of e-Learning and Knowledge Society*, 3(3), 139-146.
- Calvani, A. (2011). «Decision making» nell'istruzione. «Evidence based education» e conoscenze sfidanti. *Journal of Educational, Cultural and Psychological Studies*, 3, 77-99.
- Calvani, A. (2012). *Per un'istruzione evidence based. Analisi teorico-metodologica internazionale sulle didattiche efficacy e inclusive*. Trento: Erickson.
- Calvani, A. (2013). Evidence Based (Informed?) Education: neopositivismo ingenuo o opportunità epistemologica? *Form@re*, 13(2), 91-101.
- Campbell, D. T. (1969). Reforms as experiments. *American Psychologist*, 24, 409-429.
- Campbell, D. T. (1991). Methods for the experimenting society. *American Journal of Evaluation*, 12(3), 223-260.
- Chatterji, M. (2009). Enhancing impact evidence on how global education work. Theory, epistemological foundations, and principles for applying multiphase, mixed method designs. In K. E. Ryan & J. B. Cousins (Eds.), *The Sage international handbook of educational evaluation* (pp. 97-111). Thousand Oaks, CA: Sage.
- Christie, C. A., & Fleischer, D. (2009). Social inquiry paradigms as a frame for the debate on credible evidence. In S. Donaldson, C. Christie, & M. Mark (Eds.), *What counts as credible evidence in evaluation and evidence based practice?* (pp. 19-30). Thousand Oaks, CA: Sage.
- Clark, A. M. (2008). Evidence. In L. M. Given (Ed.), *The Sage encyclopedia of qualitative research methods*, Vol. 1 (pp. 307-308). Thousand Oaks, CA: Sage.
- Clarke, A. M. (2006). Evidence-based evaluation in different professional domains: similarities, differences and challenges. In I. F. Shaw, J. C. Greene, &

- M. M. Mark (Eds.), *The Sage handbook of evaluation, policies, programs and practices* (pp. 559-581). London: Sage.
- Commission of the European Communities (2007). *Towards more knowledge-based policy and practice in education and training*. Commission staff working document SEC2007.1098. Bruxelles: European Commission. Prelevato il 3 Febbraio 2014 da <http://bookshop.europa.eu/en/towards-more-knowledge-based-policy-and-practice-in-education-and-training-pbNC7807296>.
- Consiglio Europeo (2009). *Conclusioni del Consiglio del 12 maggio 2009 su un quadro strategico per la cooperazione europea nel settore dell'istruzione e della formazione (ET 2020)*. Prelevato il 18 Febbraio 2014 da <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:52009XG0528%2801%29:EN:NOT>.
- Corbetta, P. (1999). *Metodologia e tecniche della ricerca sociale*. Bologna: Il Mulino.
- Cottini, L., & Morganti, A. (2013). Evidence based education and special education: A possible dialogue. *Italian Journal of Special Education for Inclusion*, 1, 65-82.
- Cronbach, L. J., Ambron, S. R., Dornbusch, S. M., Hess, R. D., Hornik, R. C., Phillips, D. C., Walzer, D. F., & Weiner, S. S. (1980). *Toward reform of program evaluation: Aims, methods, and institutional arrangements*. San Francisco, CA: Jossey-Bass.
- Damiano, E. (2011). Il «senso» della valutazione. Fenomenologia sociale e opzioni epistemologiche. *Education Sciences & Society*, 2(2), 10-39.
- Datta, L. E. (1994). Paradigm wars. A basis for peaceful coexistence and beyond. *New Directions for Program Evaluation*, 61, 53-70.
- Dewey, J. (1938). *Logic, the theory of inquiry*. New York: Henry Holt and Co. (trad. it., *Logica, teoria dell'indagine*. Torino: Einaudi, 1974).
- Domenici, G. (1993). *Manuale della valutazione scolastica*. Roma - Bari: Laterza.
- Donaldson, S. I. (2009). In search of the blueprint for an evidence-based global society. In S. Donaldson, C. Christie, & M. Mark (Eds.), *What counts as credible evidence in evaluation and evidence based practice?* (pp. 2-18). Thousand Oaks, CA: Sage.
- Donaldson, S. I., & Christie, C. A. (2005). The 2004 Claremont Debate: Lipsey versus Scriven. Determining causality in program evaluation and applied research: Should experimental evidence be the gold standard? *Journal of Multidisciplinary Evaluation*, 2(3), 60-77.
- Donaldson, S. I., Christie, C. A., & Mark M. (Eds.) (2009). *What counts as credible evidence in evaluation and evidence based practice?* Thousand Oaks, CA: Sage.
- EPPI (2010). *EPPI-Centre methods for conducting systematic reviews*. London: EPPI-Centre, Social Science Research Unit, Institute of Education, University of London. Prelevato il 22 Febbraio da <http://eppi.ioe.ac.uk/cms/LinkClick.aspx?fileticket=hQBu8y4uVw1%3d&tabid=88>.

- European Commission (1999). *Guidelines for systems of monitoring and evaluation of ESF assistance in the period 2000-2006*. Prelevato il 12 Gennaio 2014 da [http://www.igfse.pt/upload/docs/aval\\_LP\\_orientacoes\\_processo\\_aval\\_inter\\_DGEmprego.pdf](http://www.igfse.pt/upload/docs/aval_LP_orientacoes_processo_aval_inter_DGEmprego.pdf).
- Federighi, P. (2011). La ricerca evidence based in educazione degli adulti. *Pedagogia Oggi, 1-2*, 112-120.
- Gargani, J., & Donaldson, S. I. (2011). What works for whom, where, why, for what, and when? Using evaluation evidence to take action in local contexts. *New Directions for Evaluation, 130*, 17-30.
- Gattullo, M. (1967). *Didattica e docimologia. Misurazione e valutazione nella scuola*. Roma: Armando.
- Gattullo, M. (1986). Misurare, valutare, decidere. Un modello e tre adattamenti. *La Scuola Se, 17-18*, 55-58.
- Gersten, R., & Hitchcock, J. (2009). What is credible evidence in education? The role of the What Work Clearinghouse in informing the process. In S. Donaldson, C. Christie, & M. Mark (Eds.), *What counts as credible evidence in evaluation and evidence based practice?* (pp. 78-95), Thousand Oaks, CA: Sage.
- Giovannini, M. L. (a cura di). (1988). *La valutazione delle innovazioni nella scuola*. Bologna: Cappelli.
- Giovannini, M. L., & Marcuccio, M. (2012). *Ricerca sulla valutazione di un progetto scolastico innovativo. Scelte e interrogativi*. Milano: Franco Angeli.
- Gough, D., Tripney, J., Kenny, C., & Buk-Berge, E. (2011). *Evidence informed policy in education in Europe: EIPEE final project Report*. London: EPPI-Centre, Social Science Research Unit, Institute of Education, University of London. Prelevato il 12 Febbraio 2014 da <http://www.eipee.eu/Default.aspx?tabid=2510&language=en-US>.
- Greene, J. C. (2009). Evidence as «proof» or evidence as «inkling». In S. Donaldson, C. Christie, & M. Mark (Eds.), *What counts as credible evidence in evaluation and evidence based practice?* (pp. 153-167). Thousand Oaks, CA: Sage.
- Green, J. C. (2012). La contribution des données probantes au processus de crédibilisation d'une évaluation. In M. Hurteau, S. Houle, & F. Guillemette, *L'évaluation de programme axée sur le jugement credible* (pp. 57-74). Québec: Press de l'Université du Québec.
- Green, J. C. (2013). Reflections and ruminations. *New Directions for Evaluation, 138*, 109-119.
- Guba, E. G. (1990). The alternative paradigm dialog. In Id. (Ed.), *The paradigm dialog*. Thousand Oaks, CA: Sage.
- Hall, J. N. (2013). Pragmatism, evidence, and mixed methods evaluation. *New Directions for Evaluation, 138*, 15-26.

- House, E. R. (2003). Introduction. In T. Kellaghan, D. L. Stufflebeam, & L. A. Wingate, *International handbook of educational evaluation* (pp. 9-14). Dordrecht - Boston - London: Kluwer Academic Publishers.
- Hurteau, M., Houle, S., & Guillemette, F. (2012). *L'évaluation de programme axée sur le jugement credible*. Québec: Press de l'Université du Québec.
- Julnes, G., & Debra, J. R. (2007). Current federal policies and controversies over methodology in evaluation. *New Directions for Evaluation*, 113, 1-12.
- Kerlinger, F. N., & Lee, H. B. (2000). *Foundations of behavioural research* (4th ed.). Orlando, FL: Harcourt College Publishers.
- Kushner, S., & Norris, N. (2007). The new public management and evaluation. *Advances in Program Evaluation*, 10, 1-16.
- Margiotta, U. (2012). The nature of evidence: Improving educational research in Italy. *Pedagogia Oggi*, 2, 37-56.
- Marradi, A. (2002). *Linee guida per l'analisi bivariata nelle scienze sociali* (4<sup>a</sup> ed.). Milano: Franco Angeli.
- Mertens, D. M., & Hesse-Biber, S. (2013). Mixed method and credibility of evidence in evaluation. *New Directions for Evaluation*, 138, 5-13.
- Miles, A. B., & Huberman, A. M. (1994). *Qualitative data analysis. An expanded sourcebook* (2nd ed.). Thousand Oaks, CA: Sage.
- Oakley, A., Gough, D., Oliver, S., & Thomas, J. (2005). The politics of evidence and methodology: Lessons from the EPPI-Centre. *Evidence & Policy*, 1(1), 5-31.
- OECD (2007). *Evidence in education. Linking research and policy*. Paris: OECD.
- Pawson, R. (2002). Una prospettiva realista. Politiche basate sull'evidenza empirica. *Sociologia e Ricerca Sociale*, 68-69, 11-57.
- Ranieri, M. (2007). Evidence based education: un dibattito in corso. *Journal of e-Learning and Knowledge Society*, 3(3), 147-152.
- Rumiati, R. (2006). Decisione, psicologia della. In *Enciclopedia italiana di scienze, lettere ed arti. Appendice 7*. Prelevato il 15 Gennaio 2014 da [http://www.treccani.it/enciclopedia/psicologia-della-decisione\\_%28Enciclopedia\\_Italiana%29](http://www.treccani.it/enciclopedia/psicologia-della-decisione_%28Enciclopedia_Italiana%29).
- Saussez, F., & Lessard, C. (2009). Entre orthodoxie et pluralisme, les enjeux de l'éducation basée sur la preuve. *Revue Française de Pédagogie*, 168, 111-136.
- Schneider, B., Carnoy, M., Kilpatrick, J., Schmidt, W. H., & Shavelson, R. J. (2007). *Estimating causal effects. Using experimental and observational designs*. Washington, DC: American Educational Research Association.
- Schwandt, T. A. (2007). Thoughts on using the notion of «evidence» in the controversy over methods choice. *New Directions for Evaluation*, 113, 111-127.
- Schwandt, T. A. (2009). Toward a practical theory of evidence for evaluation. In S. Donaldson, C. Christie, & M. Mark (Eds.), *What counts as credible evidence in evaluation and evidence based practice?* (pp. 197-212). Thousand Oaks, CA: Sage.

- Scriven, M. (1991). *Evaluation thesaurus* (4th ed.). Newbury Park, CA: Sage.
- Scriven, M. (Ed.). (1993). Hard-won lessons in program evaluation. *New Directions for Evaluation*, 58, 1-107.
- Stake, R. E. (2009). The incredible lightness of evidence: Problems of synthesis in educational evaluation. *Studies in Educational Evaluation*, 35(1), 3-6.
- The Campbell Collaboration (2001). *Guidelines for preparation of review protocols*. Prelevato il 16 Febbraio 2014 da [http://www.campbellcollaboration.org/resources/research/the\\_production.php](http://www.campbellcollaboration.org/resources/research/the_production.php).
- Thomas, G. (2004). Introduction: Evidence and practice. In G. Thomas & R. Pring (Eds.), *Evidence-based practice in education* (pp. 1-18). Maidenhead: Open University Press.
- Torgerson, C., Roberts, B., Thomas, J., Dyson, A., & Elbourne, D. (2002). *Developing protocols for systematic reviews in education: Early experiences from EPPI-Centre review*. Paper presented at Third International, Inter-Disciplinary Evidence-Based Policies and Indicator Systems Conference. Prelevato il 22 Febbraio 2014 da <http://www.cem.org/attachments/ebe/P271-278%20Carole%20Torgerson%20et%20al.pdf>.
- U.S. Department of Education (2002). *No child left behind Act of 2001. Public law 107-110 – Jan. 8, 2002*. Prelevato il 22 Febbraio 2014 da <http://www2.ed.gov/policy/elsec/leg/esea02/index.html>.
- U.S. Department of Education (2003). *Identifying and implementing educational practices supported by rigorous evidence: A user friendly guide*. Washington, DC: Institute of Education Sciences, U.S. Department of Education. Prelevato il 22 Febbraio 2014 da [http://ies.ed.gov/ncee/pdf/evidence\\_based.pdf](http://ies.ed.gov/ncee/pdf/evidence_based.pdf).
- Visalberghi, A. (1955). *Misurazione e valutazione nel processo educativo*. Milano: Edizioni di Comunità.
- What Works Clearinghouse (2014). *Procedures and standards handbook. Version 3.0*. Washington, DC: U.S. Department of Education. Prelevato il 7 Marzo 2014 da [http://ies.ed.gov/ncee/wwc/pdf/reference\\_resources/wwc\\_procedures\\_v3\\_0\\_standards\\_handbook.pdf](http://ies.ed.gov/ncee/wwc/pdf/reference_resources/wwc_procedures_v3_0_standards_handbook.pdf).

## RIASSUNTO

*Il contributo intende presentare i principali aspetti del vivace dibattito nazionale e internazionale sulle evidenze nell'ambito della valutazione, con uno sguardo particolare al contesto statunitense in cui tale dibattito ha avuto particolare rilevanza. In primo luogo vengono considerate alcune tappe fondamentali della riflessione e del confronto sul tema, nella più ampia cornice del movimento dell'«Evidence Based Education». Viene poi presa in esame la specificità delle evidenze nella valutazione rispetto alla ricerca e al processo de-*

*cisionale. In particolare, l'attenzione è stata focalizzata su alcuni degli aspetti rilevanti per l'elaborazione di una teoria dell'evidenza nella valutazione: il concetto di evidenza nelle sue molteplici accezioni, i risvolti valoriali dell'evidenza, il contesto di produzione e di utilizzo dell'evidenza, il processo di formulazione del giudizio di valore e l'uso di tale giudizio nei processi decisionali. Nella conclusione vengono infine illustrati alcuni dei motivi che rendono rilevante questo dibattito anche per il contesto italiano nella direzione di produrre evidenze in un processo valutativo credibile e trasparente.*

*Parole chiave:* Evidence based education, Evidenze, Valutazione, Valutazione basata sulle evidenze, Valutazione credibile.