

INTRODUZIONE

Negli ultimi decenni, si assiste alla proliferazione di ricerche che hanno l'obiettivo di misurare e comparare la performance degli Stati relativamente a determinati aspetti socio-economici. Alcune di queste ricerche vengono promosse da organismi internazionali quali l'OCSE e l'ONU, come nel caso del *Programme for International Student Assessment* (PISA), realizzata dall'OCSE per valutare la performance dei sistemi scolastici, o del *World Happiness Index*, sostenuta dal *Sustainable Development Solutions Network* dell'ONU con l'obiettivo di misurare il livello di benessere percepito. In molti casi, la produzione di indici di performance degli Stati ha stimolato la creazione di nuove organizzazioni internazionali come la *Transparency International*, un'organizzazione non governativa che nasce con l'obiettivo di misurare la corruzione percepita attraverso il *Corruption Perception Index*, la *Freedom House*, che ha creato un indice che misura il grado di libertà civili e diritti politici garantiti in ciascun paese, o la *Social Progress Imperative*, che produce il *Social Progress Index*, un indice che valuta le prestazioni sociali e ambientali dei diversi paesi. Altre volte, le unità di analisi non sono gli Stati ma singole istituzioni, come nel caso degli indici che misurano la performance delle università e dei principali istituti di educazione terziaria nel mondo.¹ Alcuni di questi indici sono costruiti aggregando dati che vengono raccolti a livello dei singoli cittadini di ogni Stato, spesso usando delle survey che rilevano percezioni, atteggiamenti e valutazioni.

-
1. Esistono decine di classifiche che valutano le università e i college. Tra queste, le più conosciute sono: l'*Academic Ranking of World Universities* (ARWU), redatta dall'Università Jiao Tong di Shanghai (<https://www.shanghairanking.com>, ultimo accesso: 12 agosto 2021); la *QS World University Rankings*, pubblicata ogni anno da Quacquarelli Symonds (QS), un'azienda britannica specializzata in educazione e studio all'estero (<https://www.qs.com>, ultimo accesso: 12 agosto 2021); la *Times Higher Education World University Rankings* (THE), curata da una rivista britannica che si occupa di notizie e questioni relative all'istruzione superiore (<https://www.timeshighereducation.com>, ultimo accesso: 12 agosto 2021).

Gli indici sono caratterizzati dal fatto che associano, per ogni unità di analisi, un valore numerico ad una particolare performance per produrre una classifica di vincenti e perdenti, come se stessero disputando una competizione sportiva. Ciò stimola i governi dei vari paesi coinvolti nella classifica a prestare attenzione a tali valutazioni e a cercare di migliorare il proprio punteggio, proprio come succede per i competitori sportivi.

Poiché tali indici sono costruiti attraverso un procedimento etichettato come “misurazione”, vengono percepiti come un valore oggettivi, dei «fatti sociali» (Durkheim, 1895; tr. it. 1969). Nonostante la vasta letteratura scientifica relativa ai limiti e alla natura contingente di tali procedure, e all’esistenza di indici alternativi che producono risultati molto diversi tra loro, i media che regolarmente pubblicano tali classifiche e i governi non si preoccupano di analizzare come vengono costruiti e quali distorsioni potrebbero creare, magari cercando anche di capire quali potrebbero essere gli effetti perversi prodotti dalla continua ricerca della migliore posizione nella classifica.

Misurare, nelle scienze sociali, non ha la stessa accezione che nelle scienze dure, come la fisica. È pressoché impossibile stabilire un’unità di misura standardizzata a livello mondiale, per cui i criteri di misurazione sono in genere stipulativi e contingenti. Ciò li rende poco stabili nel tempo e nello spazio; quando ci muoviamo attraverso contesti culturali differenti, è arduo mantenere l’invariabilità dei significati, mentre il processo di misurazione, così come è concepito nelle scienze dure, si basa su definizioni precise e ampiamente condivise.

Inoltre, spesso gli indici vengono calcolati usando i valori medi di dati ottenuti a livello disaggregato. Ci si chiede quindi se e quando la media di valori rilevati a livello individuale può rappresentare la performance nazionale. Sappiamo bene che la ricchezza di un paese non può essere ben rappresentata dai valori medi delle ricchezze dei suoi abitanti, a causa della sua variabilità. Ancora più complesso è il caso dagli indici che misurano il benessere e la felicità o qualche tipo di performance in cui le interazioni danno vita a fenomeni complessi che fanno emergere proprietà collettive che non sono semplicemente la somma delle parti.

Come mai, viste le ben note criticità della costruzione di questi indici, le classifiche vengono tenute in grande considerazione, in prima istanza dagli organismi internazionali e conseguentemente dai governi dei diversi paesi e dai media? Inoltre, perché, nonostante la pletera di procedure concorrenti, certi indici hanno più successo rispetto ad altri? Perché l'ONU, ad esempio, ha scelto di usare la survey della Gallup per redigere il suo report, nonostante l'abbondante disponibilità di metodi e dati che rilevano il benessere e la felicità?

L'obiettivo di questo volume è duplice. Innanzitutto intende analizzare gli aspetti epistemologici e metodologici che sono alla base della costruzione di queste classifiche. L'analisi approfondisce due casi particolari: lo studio OCSE-PISA che produce una graduatoria dei sistemi scolastici nazionali aggregando le valutazioni dei singoli studenti, e il *World Happiness Index*, che crea una graduatoria dei paesi sulla base del livello di felicità espresso dai cittadini.

Nella prima parte del volume, che comprende i capitoli 1, 2, 3 e 4, vengono esaminate le questioni associate alla misurazione e quantificazione nelle scienze sociali e al problematico passaggio dai dati individuali a quelli aggregati a livello nazionale. Nel capitolo 1 sono trattate le questioni relative al processo definitorio dei concetti che devono essere misurati. Vengono affrontati due problemi fondamentali: la difficoltà di individuare definizioni stabili e condivise, a causa del fatto che la società muta in relazione alle condizioni storiche e culturali, e l'impossibilità di tenere distinto il significato scientifico dei concetti da quello popolare, perché spesso gli enunciati scientifici si intrecciano con i discorsi di senso comune.

Il capitolo 2 analizza le modalità con cui vengono costruiti gli strumenti di misurazione usati per produrre gli indici e le conseguenti graduatorie. Nelle scienze sociali, a causa della difficoltà a individuare unità di misura standard, il concetto di misurazione viene "stirato", dando vita a metodi basati su procedure di assegnazione di giudizi che si realizzano in maniera tacita e implicita e che, di conseguenza, rendono particolarmente difficile il controllo intersoggettivo.

Molti indici usati per produrre le classifiche si basano su dati

raccolti a livello disaggregato attraverso procedure di campionamento. Il capitolo 3 esamina alcuni aspetti legati a queste procedure, che vengono attuate in contesti che presentano diversi livelli di eterogeneità interna e che non producono dei risultati puntuali ma delle stime che variano all'interno di intervalli di confidenza. A causa di queste peculiarità, i risultati delle ricerche non sono adatti ad essere strutturati sotto forma di graduatorie.

Il capitolo 4 illustra alcuni problemi relativi al passaggio dai dati individuali a quelli aggregati a livello nazionale, in particolare quello della fallacia individualistica (Richards *et al.*, 1990/1991) e dello slittamento semantico. Il trasferimento a unità di macro livello di proprietà rilevate a micro livello può risultare arbitraria e può comportare effetti indesiderati e risultati tra loro discordanti.

In definitiva, i metodi usati per costruire le classifiche, che dovrebbero monitorare e comparare la performance degli Stati rispetto a determinati aspetti socio-economici, non si basano su solide basi metodologiche ed epistemologiche. Non esistono criteri oggettivi per dichiarare che gli indicatori scelti per costruire gli indici descrivono esattamente ciò che intendono monitorare e che siano in grado di misurare un certo tipo di performance.

Le scelte relative alla costruzione degli indicatori, alla selezione dei dati e ai particolari tipi di analisi avvengono a volte in maniera arbitraria. Le classifiche sono il risultato dell'attivazione di processi molto complessi e difficili da ricostruire e quindi da comunicare in maniera esplicita. L'analisi critica delle procedure utilizzate richiederebbe molto tempo e un notevole impegno da parte di persone dotate di grandi e varie competenze, come quelle statistiche, economiche e sociologiche.

È legittimo quindi chiedersi come si crea e si sviluppa il consenso intorno a sistemi di valutazione e comparazione come quelli proposti dalle valutazioni comparative internazionali. Tra l'altro, la pubblicazione e la diffusione delle classifiche non hanno solamente il ruolo di descrivere i fenomeni sociali, ma contribuiscono a loro volta a cambiare la percezione pubblica di tali fenomeni. Non si limitano a fotografare la realtà, ma la modellano e la trasformano.

Le classifiche vengono costruite associando un valore numerico ad ogni stato di una particolare performance. Questi numeri

non vengono più percepiti come indizi o prove a sostegno di fatti scientifici, ma piuttosto come fatti scientifici veri e propri, entità imparziali e autonome rispetto alle procedure che li hanno determinati. «Il loro significato, ciò che si dice rappresentino, si configura come la risultante di scelte, accordi, imposizioni fatte da lunghe catene» (Neresini, 2015, p. 408). Più l'uso di questi numeri si diffonde più essi diventano oggetti indipendenti dal processo che li ha creati, accrescendone così il livello di accettazione sociale.

Il numero agisce di autorità. Secondo la visione di Latour (1987; tr. it. 1998), questa autorità emerge all'interno di una rete di attori che è in grado di costruire delle scatole nere. Se aumentano le scatole nere aumenta anche la complessità dei processi che hanno portato a quel numero e diventa difficile ricostruirne il percorso. Per cui, gli utilizzatori saranno costretti a usarli senza porsi tante domande. Gli attori che fanno parte del network diventano quindi dei mediatori, che garantiscono in qualche modo sulla affidabilità della scatola nera.

Come si costruiscono le scatole nere che consentono di accettare la rappresentazione di un mondo complesso come quello sociale in termini quantitativi? Come, e attraverso quali attori interessati, vengono istituzionalizzate certe pratiche di monitoraggio globale, come quelle del PISA o del *World Happiness Index*? Qual è l'impatto che questi studi hanno a livello politico, sociale e organizzativo?

Il secondo obiettivo del volume è cercare di capire quali sono le condizioni storiche, culturali e sociali che hanno portato ad accettare e promuovere la logica delle classifiche, nonostante la debolezza delle sue basi metodologiche, e i ruoli dei diversi agenti sociali, sia umani che non-umani, coinvolti in questo processo. Tali aspetti verranno approfonditi nella seconda parte. Non si tratta di uno studio sistematico ed esaustivo, ma di una panoramica degli aspetti più evidenti che contribuiscono alla diffusione e all'accettazione della logica delle classifiche nelle rilevazioni internazionali, il cui scopo è offrire spunti di riflessione sul fenomeno. Anche in questa seconda parte verranno analizzati, nello specifico, i casi OCSE-PISA e *World Happiness Index*.

In particolare, verranno presi in considerazione: il contesto

storico-culturale che ha favorito la diffusione delle classifiche mondiali, caratterizzato da una crescente istituzionalizzazione della valutazione standardizzata; il ruolo delle organizzazioni internazionali, che grazie a questi studi esercitano un certo tipo di soft power; l'accoglienza dei media e dei decisori politici locali, che adattano i risultati delle classifiche alle proprie esigenze; e infine l'*agency* dei numeri e il loro particolare confezionamento che concorrono a determinare l'immaginario collettivo e a riconfigurare la struttura delle relazioni sociali.

Nel capitolo 5, verrà esaminata l'evoluzione di quella che può essere identificata come la «società dell'audit» (Power, 1997), caratterizzata da una proliferazione di procedure di valutazione e di controllo di qualità e di performance, che dal mondo finanziario si espande a tutti i settori della vita economica e sociale. Questo tipo di controllo prevede la progettazione di indicatori quantitativi e test standardizzati, che ha stimolato l'emergere di una nuova etica di governance basata sulla competizione.

Il capitolo 6 si concentra sulle organizzazioni internazionali, che sono i principali promotori di studi i cui risultati vengono rappresentati attraverso tabelle che stilano graduatorie; in particolare verranno analizzati i ruoli dell'OCSE, della Banca Mondiale e dell'ONU. Attraverso questi studi le organizzazioni internazionali riescono a esercitare un tipo soft power che, pur non avendo una legalità formale, permette di governare attraverso i numeri e le comparazioni (Grek, 2009).

Il capitolo 7 si prefigge di analizzare le reazioni dei media e dei governi nazionali. La frequenza e le modalità con cui i risultati degli studi internazionali vengono presentati dai media contribuiscono a determinarne l'autorevolezza. L'informazione che viene trasmessa dai media deve essere semplice e accattivante per attrarre l'attenzione del pubblico; i risultati presentati sotto forma di graduatorie soddisfano questi requisiti, per cui questo è il tipo di dato che viene comunicato con maggiore frequenza. I messaggi vengono poi riadattati e reinterpretati all'interno di narrazioni locali già esistenti, dove i responsabili politici li utilizzano per giustificare decisioni predefinite.

Anche i numeri e le classifiche sono dotati di *agency* (Latour,

1987; tr. it. 1998), ossia della capacità di agire socialmente e in maniera autonoma e di contribuire alla configurazione delle interazioni che si sviluppano nel contesto in cui vengono collocati. Nel momento in cui sono introdotti all'interno delle relazioni sociali, sono in grado di prescrivere modelli e codici di comportamento, così come ogni altro attore sociale. Nel capitolo 8 verranno discussi gli effetti sociali della diffusione della logica delle classifiche e le modalità con cui queste ultime concorrono a costruire i significati del mondo e a trasmettere certi valori.

Infine, è doverosa una premessa relativa al significato di alcuni termini che vengono spesso usati come sinonimi quando si ha che fare con le graduatorie, in particolare i termini indice (*index*), indicatore (*indicator*) e punteggio (*rating*). Nel lessico delle scienze sociali, l'indicatore viene definito come un elemento che deve essere introdotto nel momento in cui è necessario rilevare un concetto complesso per il quale è difficile trovare una definizione operativa. L'indicatore è un concetto più semplice, semanticamente collegato al concetto di partenza, tale per cui diventa possibile individuare le operazioni necessarie per la sua rilevazione empirica. Se il concetto è molto astratto e complesso difficilmente potrà essere sostituito da un solo indicatore: in questo caso, occorre individuare più indicatori e poi combinarli in un unico indice. L'indice viene a volte definito anche "indicatore composito". L'indicatore, e di conseguenza anche l'indice, può assumere anche valori qualitativi, ma per costruire delle graduatorie è necessario usare indici e indicatori che sintetizzino un insieme variegato di informazioni in un numero. Il punteggio, o meglio il rating, esprime un giudizio rispetto ad una certa qualità attraverso un valore numerico. Quando si costruisce una classifica si definisce una gerarchia e ciò può essere fatto solo basandosi su un preliminare processo di valutazione. Nell'ambito della produzione di classifiche internazionali i tre termini diventano dunque sinonimi. Il punteggio assegnato alla valutazione è un numero che si basa sulla rilevazione di un indicatore semplice o composito, per cui quando si fa riferimento ai numeri che compaiono nelle graduatorie i tre termini possono assumere lo stesso significato ed essere quindi usati in maniera interscambiabile.



1. IL PROCESSO DEFINITORIO

La misurazione, nella scienza, consente di stabilire una serie di relazioni tra teoria e osservazione empirica. Possiamo dire che la pratica scientifica si basa su una continua e articolata interazione tra i costrutti della teoria e le osservazioni. La misurazione è vista come una rappresentazione della struttura empirica attraverso una struttura formale (Bruschi, 1999) basata sull'assunto «platonico che esiste una realtà vera (ιδέα) dell'attributo» (Lalla, 2015, p. 37). In sostanza, si fonda sulla costruzione di un isomorfismo tra le due strutture.

Secondo la definizione classica usata in fisica e in metrologia, misurare significa associare all'attributo di un oggetto un numero reale ottenuto dal rapporto con l'unità di misura. Quest'ultima è una grandezza definita in modo non ambiguo, per convenzione o per legge fisica, che diventa un riferimento condiviso.

Stevens (1951) ha rivisitato il significato del concetto di misurazione definendolo come il processo di assegnazione di numeri agli attributi di oggetti o eventi, sulla base di regole esplicite e convenzionali, stabilite in maniera chiara, univoca e intersoggettivamente verificabile. È evidente comunque che un presupposto cruciale per questo processo è che le definizioni semantiche degli attributi siano esplicite, non ambigue e ampiamente condivise dalla comunità scientifica.

1.1 Concetti e indicatori

Per misurare occorre che la concettualizzazione sia analiticamente definita e basata su un sistema teorico di riferimento. Nelle scienze cosiddette dure, in genere, esiste un'ampia condivisione sui significati che vengono attribuiti ai concetti scientifici: ognuno di loro ha una definizione specifica e univoca. Alcuni fanno parte esclusi-

vamente di un vocabolario specialistico usato dagli scienziati - ad esempio, i concetti di “forza gravitazionale” o di “energia cinetica” compaiono prevalentemente nei discorsi scientifici - altri, come quello di “calore” e “peso”, entrano frequentemente anche nel linguaggio comune. A volte, questi ultimi diventano ambigui o assumono un significato diverso rispetto a quello scientifico, come nel caso dei concetti di “temperatura” e “calore” che comunemente vengono percepiti come sinonimi anche se hanno delle definizioni scientifiche che li distinguono, o il concetto di “peso” che nel linguaggio comune corrisponde al concetto scientifico di massa, ma comunque ciò non influisce sulle procedure di misurazione. In fisica, la definizione dei concetti scientifici racchiude spesso anche le operazioni per rilevarla, coincide cioè con la sua definizione operativa”.¹

Cosa succede nelle scienze sociali? Esistono due problemi: il primo riguarda la difficoltà di definire i concetti in maniera oggettiva e condivisa da tutti gli scienziati, perché la società è mutevole e cambia rispetto al tempo e alle culture di riferimento. Il secondo, è che i concetti scientifici si intrecciano con i discorsi di senso comune: è impossibile spesso tenere distinti i due piani di uso dei concetti, quello strettamente scientifico e quello popolare. I concetti hanno a che fare con interpretazioni di interpretazioni per cui gli schemi teorici degli attori sociali possono influenzare quelli degli scienziati sociali e viceversa (Giddens, 1976; tr. it. 1979).

Le definizioni e le classificazioni sono preliminari indispensabili alla raccolta e all’analisi dei dati empirici. Secondo Madge (1962; tr. it. 1966), un problema nell’uso dei metodi sistematici nelle scienze sociali «è l’impossibilità di definire alcuni degli ambiziosi concetti coinvolti nell’analisi empirica» (*ivi*, p. 663). Il concetto, per diventare “scientifico” e “misurabile” deve essere esplicito e condiviso. Per procedere alla misurazione di un concetto, occorre innanzitutto partire da una definizione che sia evidente ed esente da ambiguità, dichiarando solo le

1. Secondo Bridgman (1927; tr. it. 1961), un concetto scientifico è tale solo se si identifica con il gruppo di operazioni necessarie per la sua traduzione empirica.