

Sui test

Paola Bolgiani

Proveremo ad esaminare alcuni tipi di test attualmente considerati strumenti diagnostici. Ciò che mi preme mettere in luce è la logica in base alla quale tali test sono stati costruiti e vengono utilizzati.

Interessante notare che la parola “test” viene da “testa”, il vaso di coccio in cui gli alchimisti nel Medio Evo sperimentavano le reazioni di metalli con particolari sostanze al fine di produrre l’oro: quando la “testa” si rompeva rivelava se l’esperimento era o meno riuscito.

Naturalmente se prendiamo il concetto di “test” come prova, o esame, è ben chiaro che esso è antico come la civiltà: fin dalla Grecia antica, per esempio, esistevano prove alle quali un allievo doveva sottoporsi per valutare il suo apprendimento o la sua capacità fisica. E’ chiaro però che quello che invece caratterizza i test per come li intendiamo oggi è la dimensione di misurazione che essi introducono. Parlando dei test ci inoltriamo quindi nel campo di quella che viene definita la psicomètria.

Ciò vale anche nell’ambito più specifico della malattia mentale: se infatti per un verso fin dall’inizio dell’800 si sono messe in atto prove per individuare e anche classificare i malati mentali, questi non comportavano ancora l’uso della misura come fattore determinante.

I test attualmente utilizzati come strumenti diagnostici (ma, come vedremo, non solo diagnostici) si dividono in due grandi ambiti: quelli che misurano l’intelligenza e quelli di personalità, a loro volta suddivisi in questionari e proiettivi. Il primo campo in cui sorsero i test come misura si riferiscono tuttavia alla misura dell’intelligenza.

Test di intelligenza

Il primo ideatore di un test può essere considerato Esquirol, che nel 1838 propose di valutare il livello mentale di un individuo a partire da test di natura verbale. Nello stesso periodo nacquero anche prove di abilità considerate capaci di indicare il livello di intelligenza e consistenti nell’incastro di forme diverse in apposite strutture.

Fu nel 1890 che lo psicologo americano Cattell utilizzò per primo il termine “test mentale” per indicare una pratica di valutazione delle differenze psicologiche ma anche fisiologiche dei suoi studenti. Egli ipotizzò e studiò la correlazione fra alcuni caratteri somatici di un individuo e la sua capacità di reazione a stimoli esterni uditivi, visivi e tattili, inferendo da questi il grado di intelligenza, sulla base dell’idea che le capacità sensoriali di un individuo diano un’indicazione sulle sue capacità mentali – e quindi sulle capacità più complesse. Queste prove, che peraltro ebbero scarsi risultati nei termini di rispondenza fra le caratteristiche studiate e gli effettivi risultati scolastici, furono tuttavia i primi a mettere in rilievo uno studio delle differenze.

In quegli stessi anni nasceva in Europa la psicomètria, cioè la disciplina che si propone di misurare i fenomeni psichici attraverso procedimenti scientifici. Sia in Italia che in Germania questo diede impulso a nuove ricerche basate sugli stimoli precettivi e le corrispondenti reazioni, sempre a partire dall’idea di un collegamento fra queste e le facoltà mentali complesse. Essi si basavano

quindi sull'assunto della psicologia scientifica sperimentale, che suddivide in una somma di capacità semplici le attività umane più complesse.

Inoltre intorno alla metà dell'800 nasceva la scienza statistica, sebbene gli studi sulle probabilità avessero avuto inizio almeno a partire dal XVI secolo. La statistica, come vedremo, in breve si salderà con la psicomètria.

Nel 1895 Binet, in Francia criticò gli studi che lo avevano preceduto e i loro presupposti e propose un nuovo metodo valutativo delle capacità cognitive. Questo fu applicato pochi anni dopo quando Binet stesso fu chiamato a far parte di una commissione incaricata di valutare i bambini considerati ritardati per valutare la loro possibilità di apprendere, onde indirizzarli a classi speciali o eventualmente ad allontanarli dalle scuole. Binet costruì così la prima scala di valutazione dell'intelligenza, nota come scala Binet-Simon, che presentava una serie di quesiti in ordine crescente di difficoltà e che si riferivano a quelle che erano considerate le facoltà intellettive più importanti: capacità di esprimere giudizi, comprensione ed elaborazione di informazioni, capacità di definire vocaboli, memoria, ragionamento induttivo e deduttivo. Questa scelta oltre a denotare un'interpretazione specifica di cosa si intendesse per intelligenza, ipotizzava una correlazione fra i risultati alle prove e la capacità di apprendere dei bambini, correlazione nient'affatto scontata. Nella costruzione del test Binet valutò la difficoltà delle prove a partire da un gruppo di controllo, costituito da 50 bambini, di età diverse, oltre che da un piccolo gruppo di bambini "mentalmente ritardati".

La scala venne rivista e integrata diverse volte. Nel 1908, a partire dalla revisione della scala Binet Simon, fu introdotto il concetto di età mentale. L'età mentale è calcolata a partire dai risultati alle prove di un individuo confrontate con i risultati medi dei bambini della sua età appartenenti al gruppo di controllo. Da allora il concetto di età mentale a partire dalla sua semplicità e dalla sua applicabilità ebbe un grandissimo successo e una vasta diffusione.

Nel 1911 l'ultima scala Binet comportò l'estensione anche agli adulti. Nel 1917 negli Stati Uniti si cercò di applicare il test di Binet alle reclute per valutare la loro collocazione nell'esercito. L'insuccesso di questa operazione diede luogo alla formulazione dei primi test collettivi, costruiti secondo la medesima logica dei primi test di intelligenza (una serie di quesiti di diversa natura i cui risultati erano messi a confronto con quelli ottenuti dai gruppi di controllo), ma in grado di testare gruppi molto numerosi e quindi di essere utilizzati in contesti di massa, quale appunto il contesto militare.

Vediamo quindi che alla base dei test di intelligenza vi sono i seguenti presupposti: un'idea di intelligenza fondata sulla somma di diverse capacità mentali "semplici" rispetto all'intelligenza intesa come capacità tanto complessa da risultare difficilmente definibile agli stessi ricercatori. Inoltre l'idea, niente affatto scontata, che il risultato alle prove proposte sia in correlazione con l'intelligenza stessa di un individuo, idea costruita in maniera tautologica a partire dal raffronto con il gruppo di controllo e dalla media statistica dei risultati da questo ottenuto.

Questionari e test di personalità

I test di personalità sono considerati correlativi dei test di intelligenza, in quanto sarebbero finalizzati a rilevare le caratteristiche emotive, motivazionali, interpersonali e di atteggiamento di un individuo, distinte dalle sue capacità intellettive. Si tratta perlopiù di questionari finalizzati a

misurare in quale misura siano presenti determinate caratteristiche di personalità in un individuo. Ciò comporta evidentemente una categorizzazione previa di tali caratteristiche e, spesso, anche una valutazione positiva dell'adeguatezza di certe caratteristiche rispetto ad altre, considerate negative e/o inadeguate. I test di personalità vengono utilizzati non solo per la diagnosi in campo clinico (psicologico o psichiatrico), ma anche in altri ambiti, specialmente per la selezione del personale nelle aziende.

I questionari di personalità sono considerati meno attendibili rispetto ai test di intelligenza, considerati come maggiormente obiettivi a partire dall'idea che in questi ultimi un individuo non possa manifestare capacità che non possiede; nei questionari un individuo può infatti dare un'immagine di sé che considera più positiva o più socialmente accettata. Per questo motivo, oltre ad essere accompagnati dall'esortazione di risposte "sincere", questi test hanno al proprio interno una serie di quesiti "di controllo", ovvero miranti a rilevare eventuali incongruenze, contraddizioni o incompatibilità fra diverse risposte.

I questionari di personalità sono costituiti da domande molto strutturate (ovvero con possibilità di risposta limitata, ad esempio sì o no, oppure su una breve scala di valutazione, ad esempio da 1 a 4); la loro valutazione viene fatta secondo procedure standardizzate e confrontata con gruppi di controllo in base a criteri statistici.

Il più noto e utilizzato fra i questionari di personalità è l'MMPI (Minnesota Multiphasic Personality Inventory), che vide la luce nella sua prima versione nel 1942 negli Stati Uniti, a cura dell'Ospedale dell'Università del Minnesota. Nasce ad opera di due psichiatri la cui finalità era quella di poter pervenire a diagnosi psichiatriche in maniera facile e rapida e per valutare la gravità del disturbo psicopatologico e, conseguentemente, dislocare in modo opportuno i pazienti all'interno dell'ospedale. È quindi interessante notare che non nasce con finalità diagnostiche in senso stretto, ma piuttosto con finalità legate gestionali.

La costruzione del test avvenne con l'elaborazione di circa 1000 affermazioni che furono poi presentate ai pazienti dell'ospedale del Minnesota, già diagnosticati come ipocondriaci, depressi, paranoici, isterici, ecc. e parallelamente alla popolazione "normale" composta da parenti, visitatori e pazienti di altri reparti, scegliendo poi fra le affermazioni quelle che discriminavano i malati mentali dagli altri. Delle affermazioni (item) ne furono scelti 566, suddivise in dieci scale, a cui furono aggiunte 4 scale di controllo. Queste sono appunto quelle che servono per valutare la sincerità delle risposte date. Nel 1989, dopo numerose revisioni e aggiunte, è uscito l'MMPI-2, ancora rivisto nel 2001.

Per cogliere come funziona il test, ecco alcuni esempi delle scale di cui è composto l'MMPI fin dalla sua prima versione: Ipocondria (presenza di problemi fisici caratteristici degli ipocondriaci); Depressione (presenza di sintomi depressivi); Isteria di conversione (tendenza a somatizzare emozioni o disagi psichici); Deviazione psicopatica (carenza di controllo sulle risposte emotive e capacità di introiettare le regole sociali); Mascolinità-femminilità (tutti gli aspetti, interessi, atteggiamenti ecc. tipicamente maschili o femminili).

Successivamente sono state introdotte altre scale. Eccone alcune: Ansia (indica il livello di stress o di disagio provato dal candidato); Repressione (indica il livello di sottomissione e di convenzionalità); Forza dell'Io (indica la capacità del soggetto di trarre vantaggio dal trattamento psicoterapico); Ostilità ipercontrollata (indica la capacità di gestione della frustrazione); Leadership

(indica la tendenza ad assumere ruoli di controllo nelle relazioni sociali); Responsabilità sociale (indica la percezione del candidato di responsabilità sociale).

Nelle versioni più aggiornate, sono state inoltre introdotte le scale cosiddette “di contenuto”, che servono ad approfondire diverse variabili della personalità. Meritano di essere lette tutte:

Ansia: indica la presenza di sintomi ansiosi (somatizzazioni, difficoltà del sonno, problemi di concentrazione, etc.); Paure: indica la presenza di fobie; Ossessività: indica la presenza di comportamenti ossessivi; Depressione: indica la presenza di pensieri depressivi; Salute: indica la presenza di un'eccessiva preoccupazione per la propria salute fisica; Ideazione bizzarra: indica i processi di pensiero di tipo psicotico; Rabbia: indica problemi di controllo della rabbia; Cinismo: rileva le convinzioni misantropiche del candidato; Comportamento antisociale: indica soggetti che in passato hanno avuto comportamenti antisociali; Tipo A: indica i soggetti di tipo A (ipermotivati, centrati sul lavoro, irritabili); Bassa autostima: indica i soggetti con una bassa opinione di sé; Disagio sociale: misura il disagio a stare in gruppo; Problemi familiari: segnala la presenza di conflitti familiari; Problemi lavorativi: segnala la presenza di conflitti sul lavoro; Indicatori di difficoltà di trattamento: indica gli atteggiamenti negativi del soggetto verso i trattamenti di salute mentale.

Vediamo che, in particolare rispetto a queste scale di contenuto, i dati che si intendono ottenere vanno ben al di là di aspetti “diagnostici”, in quanto indagano molti aspetti della vita degli individui (come problemi familiari, passati comportamenti antisociali) in base ai quali considerare l'adeguatezza o meno sul piano dei comportamenti.

Come esempio di test proiettivi prenderemo il più noto, ovvero il test di Rorschach, così chiamato dal nome del loro creatore Hermann Rorschach (Zurigo 1884-1922). Il test è molto noto, e consiste nel sottoporre alla persona dieci tavole che presentano figure ottenute a partire da macchie di inchiostro di forma simmetrica. La persona è richiesta di dire che cosa vede in tali macchie.

L'interpretazione di forme ambigue, così come le illusioni ottiche, furono utilizzate fin dai tempi di Leonardo e di Botticelli, e d'altra parte pare che l'uso dell'interpretazione delle macchie di inchiostro fosse un gioco molto in voga nell'800, non solo conosciuto da Rorschach stesso, ma da lui tanto amato e praticato che pare che gli amici lo avessero soprannominato “Macchia”.

Dopo aver studiato botanica e biologia, Rorschach si iscrive alla facoltà di medicina e poi si specializza, nel 1909, in psichiatria, andando a lavorare in un ospedale psichiatrico. L'idea di far diventare il gioco allora in uso un test venne da un incontro casuale. Vicino all'ospedale in cui lavorava, un suo ex compagno insegnava in una scuola e intratteneva i suoi allievi con il gioco delle macchie. Il compagno gli disse di avere l'impressione che, con questo gioco, gli allievi migliorassero le loro prestazioni e la loro attenzione. Quello che Rorschach si chiese era se gli allievi più dotati potessero avere maggiore fantasia degli allievi meno dotati: così provò a “testare” questa sua ipotesi e successivamente a confrontare le risposte degli studenti con quelle dei suoi pazienti dell'ospedale. La prima osservazione, che presentò nella sua tesi di dottorato nel 1912 di cui era relatore Bleuler, fu che i pazienti schizofrenici producevano risposte molto diverse da quelle che davano gli individui “normali”, tesi questa però che non ebbe pressoché alcun seguito. Rorschach abbandonò quindi lo studio delle macchie fino a quando, alcuni anni dopo, nel 1917, un altro allievo di Bleuler, Hens presentò una ricerca sull'uso delle macchie di inchiostro come prova di

immaginazione. In questa ricerca egli aveva mostrato 8 macchie a più di mille soggetti (bambini, adulti, anziani, pazienti psichiatrici, ecc) e, rilevando delle differenze nel contenuto delle risposte, aveva ipotizzato che se utilizzate in modo corretto le macchie avrebbero potuto essere utili a fini diagnostici.

Da allora, e per i pochi anni di vita che gli rimanevano, Rorschach sviluppò il suo metodo con un intento sia di studio della percezione e della personalità, sia di diagnosi differenziale in psicopatologia, pubblicando nel 1921 il suo più importante manuale, *Psicodiagnostica. Metodologia e risultati di un esperimento diagnostico basato sulla percezione*.

I risultati che Rorschach descrive nel libro sono frutto di un metodo sperimentale, che comportava una metodologia per la somministrazione molto rigorosa, affinché la somministrazione delle macchie fosse paragonabile per soggetti diversi. Le macchie dovevano essere somministrate accompagnando sempre con la stessa frase la loro presentazione, e occorreva annotare fedelmente tutto quello che il soggetto diceva, e anche i suoi movimenti, le sue espressioni, i segni di "shock". L'insieme di queste annotazioni prende il nome di protocollo, e anche la sua valutazione fu rigorosamente delineata da Rorschach stesso: numero di risposte, i rifiuti, la localizzazione (interpretazione della figura nella globalità o in alcune parti), la determinante della risposta (forma, colore o sensazioni di movimento), il contenuto della risposta. Successivamente Rorschach inserì fra le determinanti anche il chiaroscuro (a seguito di una stampa delle tavole che aveva prodotto delle tavole sfumate e non in tinta unita) e la qualità del contenuto delle risposte (se banali o originali rispetto alla media delle risposte fino ad allora acquisite).

Ciò che è interessante rispetto al test di Rorschach è il fatto che la sua attendibilità sia riconosciuta a partire dalla standardizzazione delle procedure così come delle interpretazioni, mentre per alcuni non è corretto definirlo un test perché non è sufficientemente supportato da dati statistici significativi. Non è perciò considerato uno dei test psicometrici oggettivi, sebbene negli Stati Uniti a partire dagli anni '70 si siano prodotte ricerche di validazione (statistica) sempre più ampie.

Qualche considerazione

Nella misura in cui la psicologia si pone la questione di misurare le caratteristiche psicologiche e mentali di un individuo, deve anzitutto, sulla scorta della psicologia sperimentale, scegliere quali elementi prenderà in considerazione come rappresentativi di quella caratteristica. Questo passaggio, che è necessario in ogni ricerca scientifica, nel campo della diagnosi psicologica non può che risultare arbitrario e frutto di una concezione previa, che questa sia o meno esplicita. Ciò che colpisce non è tanto questo aspetto in sé, bensì il fatto che molto facilmente venga dimenticato, e dunque si consideri il risultato ad un test come una rappresentazione oggettiva di ciò che si va a valutare.

In secondo luogo, occorre tenere presente che ogni valutazione mediante test deve necessariamente, per la sua stessa impostazione, rendere irrilevante la presenza dell'esaminatore e le sue caratteristiche dalla valutazione che il test pretende di fare. Ciò viene attuato mediante la standardizzazione delle procedure, partendo dal presupposto che, se ciascun esaminatore agisce allo "stesso" modo, l'interferenza che egli comporta sarà annullata.

In terzo luogo, in generale non è tenuto in considerazione il fatto che la situazione stessa di esame rappresenta un'interferenza non eliminabile dalla valutazione pretesa oggettiva. La risposta a

questo ostacolo è ancora una volta cercata nella standardizzazione, inclusa la standardizzazione di ciò che occorrerebbe dire affinché la situazione di esame non infici i risultati.

Un test viene eseguito sempre con una certa finalità, come abbiamo visto anche nella storia del loro sviluppo. Che sia per indirizzare gli studenti in una scuola, che sia per collocare in modo più conveniente i soldati in un esercito, che sia per selezionare i candidati ad un lavoro, che sia infine per diagnosticare una patologia, si tratta in ogni caso di una finalità che ha delle precise connotazioni che prescindono dal campo strettamente clinico, e che sfocia nel campo politico, o biopolitico, di gestione dei corpi e delle popolazioni, e lo dimostra il fatto che in tutti questi casi ciò che conta è il raffronto con una popolazione considerata “nella norma”.