La fallacia della convalida soggettiva: una dimostrazione di credulità in aula

Bertram R. Forer

(1949)
The Journal of Abnormal Psychology
vol. 44, pp. 118-121
(Traduzione di Romolo Giovanni Capuano©)

Il presente articolo esamina alcuni errori metodologici che possono condizionare la stima della validità delle interpretazioni della personalità e degli strumenti di misurazione. Di primaria importanza è la natura delle interpretazioni stesse. Le valutazioni della personalità possono essere, e spesso sono, formulate in termini talmente generici da essere prive di significato in termini di comportamento denotabile. Oppure possono avere una "validità universale" e applicarsi a tutti. Bobertag (2) definisce *Universalscharakteristik* il tratto di personalità universalmente valido.

Avere due occhi è una caratteristica di tutti i vertebrati, quindi non è un fattore di distinzione valido tra essi. Il pollice opponibile non distingue un essere umano dall'altro. A livello psicologico, l'accettazione di alcuni tabù culturali sembra essere universale tra gli esseri umani che vivono all'interno di gruppi sociali. In pratica, tutti esibiscono qualche tratto psicologico in una certa misura. Se si intende caratterizzare un determinato individuo, è una procedura priva di senso mettere in evidenza i tratti che esibisce. Non è per la presenza o assenza di un tratto che gli individui differiscono. L'unicità di un individuo, come documenta ampiamente Allport (1), sta nell'importanza relativa dei vari elementi della personalità nel determinare il suo comportamento e nella consistenza relativa di questi tratti in confronto ad altri individui. L'individuo, dunque, è una configurazione unica di caratteristiche, ognuna delle quali può essere rinvenuta in tutti in misura variabile. Una affermazione universalmente valida si applica in maniera ugualmente soddisfacente alla maggioranza o alla totalità di una popolazione. Una affermazione universalmente valida è vera per l'individuo, ma è priva degli specifici elementi quantitativi ed esclusivi necessari ai fini di una diagnosi differenziale. In un certo senso, una affermazione universalmente valida descrive un gruppo culturale piuttosto che un dato psicologico personale.

È verosimile che un cliente accolga una descrizione universalmente valida della personalità come verità su di sé, giudicandola una descrizione che si adatta unicamente alla propria persona. Molti, se non la maggior parte, degli individui sono in grado di riconoscere determinate caratteristiche in se stessi, quando queste non vanno a loro discapito, ma ignorano le stesse caratteristiche quando emergono in altri individui. Un esempio è fornito dalla tendenza degli studenti a percepire i propri problemi secondo quanto riportato dai testi di psicologia anormale. In questi casi, l'individuo è privo della necessaria cornice quantitativa di riferimento per confrontare criticamente la descrizione riportata nel testo con la propria autovalutazione.

Talvolta, la conferma del cliente o di un terzo a conoscenza del suo caso è impiegata come criterio di convalida di inferenze e procedure diagnostiche (4). I risultati di un test possono suggerire problemi e modalità comportamentali caratteristiche che il terapeuta o lo stesso cliente possono confermare o smentire. Si può definire "convalida soggettiva" l'accertamento della correttezza delle inferenze relative a un cliente sulla base della valutazione da questi espressa sulle stesse. Quando le inferenze sono valide universalmente, come spesso accade, la conferma è inutile. I risultati positivi ottenuti tramite convalida soggettiva possono facilmente indurre un

falso senso di sicurezza in un terapeuta o in uno psicologo sperimentale; senso di sicurezza che rafforza la convinzione relativa alla bontà essenziale della propria filosofia della personalità o delle proprie capacità diagnostiche. Tale falsa convalida accresce la fiducia nell'uso di quello che altrimenti sarebbe da considerare uno strumento dubbio. Quando la conferma di una predizione è estesa in maniera acritica allo strumento, al sistema concettuale o alla persona che compie la predizione sorge un grande pericolo. Tali estensioni acritiche si verificano troppo frequentemente nel campo clinico.

La conferma di una predizione non dimostra necessariamente la validità delle proposizioni da cui scaturisce la predizione. Una predizione identica può scaturire da un insieme di proposizioni che contraddicono le proposizioni iniziali (3, p. 140). Taylor (12) ha dimostrato empiricamente che giudici diversi possono produrre predizioni identiche sulle medesime storie cliniche per motivi diversi. È necessaria la conferma di svariate predizioni in grado di discriminare tra un certo numero di clienti, se vogliamo accogliere una convalida con un qualche grado di fiducia.

È verosimile che gli indovini siano consapevoli di alcuni di questi meccanismi e che altri soggetti che producono pseudo-diagnosi, sebbene forse inconsapevoli delle fallacie inerenti alle loro procedure, ingannino i propri clienti facendo ricorso alla "validità universale" e alla "convalida soggettiva". Allport (1, p. 476) afferma che «uno dei modi con cui chi analizza il carattere si garantisce una reputazione di successo è tramite l'utilizzo di termini ambigui che possono applicarsi a chiunque». Uno sprovveduto a cui siano comunicate informazioni diagnostiche superficiali, soprattutto se il contesto sociale è connotato da prestigio, tende a fare proprie tali informazioni[1]. Egli rimane colpito dalle ovvie verità e ignora le discrepanze. Ma c'è dell'altro. Lo sprovveduto convalida anche lo strumento e chi produce la diagnosi. Gli studenti di Crider (4) giudicarono sorprendentemente accurate le analisi ricevute da un soggetto che produceva pseudo-diagnosi. Sembra che lo stesso Crider ne sia rimasto affascinato e che biasimi ogni rifiuto aprioristico delle pretese di questi individui. Se il ricorso a procedure di confronto dimostra che le inferenze derivate da test proiettivi eseguiti da clinici esperti posseggono una validità abbastanza alta (6, 7, 8, 9, 10), ciò non è vero per le asserzioni prodotte da chi utilizza tecniche grafologiche non standardizzate (11).

---

[1] D. G. Paterson, in una lettera personale all'autore, descrive e allega un profilo di personalità universalmente valido che egli utilizza nel corso delle sue conferenze. Lo riproduciamo di seguito con il suo permesso.

"Ha una intelligenza e una prontezza mentale superiori alla media. Anche la sua precisione è al di sopra della media; a volte è davvero meticoloso. È giustamente noto per la sua pulizia; non sopporta la sciatteria. Ha iniziativa, ossia, capacità di proporre suggerimenti e procurarsi nuove idee; è di mentalità aperta.

"A volte tendi a preoccuparti, ma non eccessivamente. Qualche volta ti senti depresso, ma non sei una persona di umore nero perché, in genere, sei allegro e piuttosto ottimista. Hai un buon carattere, anche se, in una fase precedente della tua vita, hai dovuto sforzarti per tenere a freno i tuoi impulsi e il tuo temperamento.

"Sei molto socievole; ti piace fare nuove conoscenze e soprattutto frequentare persone che conosci bene. Apprezzi l'arte, la pittura e la musica, ma non sfonderesti mai come artista, creativo o compositore musicale. Ti piacciono gli sport e gli eventi di atletica, ma dedichi più tempo alle pagine sportive dei quotidiani che alla partecipazione a eventi agonistici.

"Sei ambizioso ed è lodevole il fatto che tu voglia che la famiglia, i colleghi e gli amici abbiano una buona opinione di te. Le tue ambizioni si rivelano, per lo più, in una inclinazione ad abbandonarti a fantasie e sogni a occhi aperti, ma questo non significa che non sei capace di partecipare attivamente al gioco della vita.

"Continuerai ad avere buoni risultati lavorativi purché ti dedichi ad attività sociali ossia se ti impegni in occupazioni che ti portano a contatto con la gente. Non importa tanto quale lavoro sceglierai quanto il fatto che sei con altre persone. Il risvolto negativo è che non avresti successo se ti impegnassi in attività meramente teoriche o di pura ricerca in campi quali la fisica o la neurologia".

Recentemente, l'autore è stato avvicinato da un grafologo in un locale notturno che ha espresso il desiderio di "leggere" la sua scrittura. L'autore ha declinato la proposta e si è offerto di somministrare il test di Rorschach al grafologo. Ne è scaturita una gradevole discussione, nel corso della quale il grafologo ha sostenuto le basi scientifiche del suo lavoro sul fatto che i clienti affermano la correttezza delle sue interpretazioni. L'autore ha risposto che uno psicologo potrebbe eseguire una "lettura" a occhi chiusi e giungere allo stesso livello di convalida.

Esperimento

L'autore ha condotto il seguente esperimento con alcuni studenti di un corso introduttivo di psicologia per dimostrare la facilità con cui è possibile indurre i clienti a fornire una valutazione positiva, seppure ingiustificata, di uno strumento diagnostico attraverso una descrizione generica della personalità. In precedenza, l'autore aveva presentato agli studenti il *Diagnostic Interest Blank*[2] (5) ("Questionario diagnostico degli interessi", da qui in poi DIB), uno strumento elaborato per valutare il ruolo dei fattori motivazionali soggettivi nella selettività percettiva. Gli studenti hanno chiesto di essere sottoposti al test e di ricevere una valutazione della propria personalità. La richiesta è stata accolta dall'autore. Nel corso dell'incontro successivo, i 39 studenti sono stati sottoposti al DIB e informati che avrebbero ricevuto una breve profilo di personalità non appena l'autore avesse avuto il tempo di esaminare i risultati del test. Una settimana dopo, ogni studente ha ricevuto un profilo di personalità scritto a macchina e riportante il suo nome. L'autore ha condiviso il desiderio espresso dagli studenti di mantenere riservato il contenuto dei profili. Fortunatamente, essendo quel giorno in programma la somministrazione di un questionario, è stato possibile disporre ogni studente a distanza di due posti l'uno dall'altro senza suscitare sospetti. Dal punto di vista dello sperimentatore, era essenziale che nessuno studente vedesse i profili degli altri perché tutti i profili erano identici[3]. Gli studenti non hanno nutrito alcun sospetto.

Il profilo di personalità si sovrappone in parte a quello di Paterson, ma è composto da 13 affermazioni piuttosto che da una descrizione narrativa. Un'altra differenza sta nel fatto che il profilo è stato concepito per soddisfare il criterio della validità universale più di quanto non faccia quello di Paterson. Il profilo prevede i seguenti item:

1. Hai molto bisogno che gli altri ti apprezzino e ti ammirino.
2. Tendi a essere critico nei tuoi confronti.
3. Hai molte capacità inutilizzate che non hai sfruttato a tuo vantaggio.
4. Pur avendo alcune debolezze di carattere, sei generalmente in grado di porvi rimedio.
5. Il tuo adattamento sessuale ti ha posto alcuni problemi.
6. Anche se mostri disciplina e autocontrollo all'esterno, tendi a preoccuparti e a sentirti insicuro dentro di te.
7. Talvolta dubiti seriamente di aver preso la decisione giusta o di aver fatto la cosa giusta.
8. Preferisci una certa dose di cambiamento e variazione e ti senti insoddisfatto se sei impedito da limiti e restrizioni.
9. Ti vanti di saper pensare con la tua testa e di non accettare le opinioni degli

[2] Il DIB è composto da un elenco di passatempi, letture, caratteristiche personali, compiti lavorativi, speranze e ambizioni segrete relative a una persona ideale. Il test consente una interpretazione qualitativa e inferenze sulle dinamiche di personalità in modo simile a quanto accade con i test proiettivi

[3] Gli item provengono, in gran parte, da un testo di astrologia acquistato in una edicola. L'autore non era a conoscenza del profilo di Paterson quando l'esperimento è stato pianificato ed eseguito.

altri in mancanza di solide prove.

10. Hai imparato che è imprudente essere troppo sinceri nel rivelarsi agli altri.

11.Talvolta sei estroverso, affabile e socievole; altre volte sei introverso, diffidente e riservato.

12. Alcune tue aspirazioni sono piuttosto irrealistiche.

13. La sicurezza è uno degli obiettivi principali della tua vita.

Prima di ricevere i profili, gli studenti hanno avuto istruzioni di leggerne i contenuti e consegnare in seguito i fogli esprimendo le seguenti valutazioni:

A. Valuta su una scala da zero (nulla) a cinque (ottima) con quanta efficacia il DIB rivela la personalità.

B. Valuta su una scala da zero a cinque quanto la descrizione fornita riveli caratteristiche fondamentali della tua personalità.

C. Volta di nuovo il foglio e indica in corrispondenza di ogni item se esso esprime un giudizio vero o falso su di te o apponi un punto interrogativo quando sei incerto.

In risposta a una precisa richiesta, gli studenti sono stati informati che l'autore era in possesso di una copia del profilo che sarebbe stata consegnata loro dopo la raccolta dei dati. Una volta recuperati i fogli, l'autore ha chiesto agli studenti di alzare le mani nel caso avessero ritenuto soddisfacente il test. Quasi tutti gli studenti hanno alzato le mani, come essi stessi hanno osservato. A questo punto, è stato letto il primo item del profilo e gli studenti hanno ricevuto l'istruzione di indicare con la mano se esso assomigliava a uno degli item presenti nei loro profili. Essendosi alzate tutte le mani, la classe è scoppiata a ridere. Gli studenti hanno appreso, dunque, che l'esperimento era stato eseguito allo scopo di dimostrare la tendenza delle persone a lasciarsi eccessivamente colpire da affermazioni vaghe e ad attribuire a chi produce diagnosi una capacità di penetrazione immotivatamente alta. Sono state messe in evidenza le somiglianze tra l'esperimento e le attività dei ciarlatani. Il fatto che almeno un terzo degli studenti abbia chiesto di avere una copia del profilo in modo da sottoporlo per scherzo agli amici indica che si è trattato di una esperienza significativa per loro.

Risultati

I dati mostrano chiaramente che il gruppo si è lasciato raggirare. I giudizi relativi all'adeguatezza del DIB hanno evidenziato solo una valutazione inferiore a 4. Lo strumento, dunque, ha ricevuto un punteggio alto in termini di convalida soggettiva. La valutazione complessiva del profilo ha evidenziato cinque valutazioni inferiori a 4 (Tavola 1). Anche se alcuni studenti hanno espresso più giudizi critici nei confronti del profilo che del DIB, la maggior parte di essi ha dichiarato che il test rivelava senz'altro i tratti fondamentali della propria personalità.

Il numero di item specifici giudicati veri variava nel gruppo da 8 a 13, tranne per un soggetto che ha giudicato veri solo 5 item (Tavola 2). Il medesimo soggetto ha attribuito al test il punteggio di 4 e al profilo il punteggio di 2. In media, sono stati giudicati veri 10,2 item.

Non è emersa alcuna relazione significativa tra le valutazioni espresse dagli studenti e variabili come il sesso, l'età, l'occupazione e i punteggi conseguiti nella prova successiva.

A parte gli alti punteggi emersi dal DIB, che dimostrano un cospicuo livello di credulità o giudizio erroneo, un ulteriore indicatore di credulità scaturisce dal numero di valutazioni espresse in base a elementi non fattuali o contrari ai fatti. Se il soggetto

ritiene che tutti gli item si applichino ugualmente bene alla propria personalità, egli ha, in un certo senso, motivo di convalidare lo strumento; se esprime disaccordo rispetto a tutti gli item del profilo, egli ha motivo di respingere il DIB.

Il test del chi–quadro indica una certa associazione, significativa al livello dell'1%, tra le valutazioni espresse sulla bontà del profilo (B) e il numero di item contrassegnati come veri. Tuttavia, è evidente l'intervento di altri fattori che hanno condizionato il giudizio di inferenza dalla parte al tutto. Per alcuni soggetti, la presenza di 8 item contrassegnati come veri su 13 era di per sé sufficiente a giudicare ottimo il profilo. Per altri, un giudizio positivo, seppure non ancora ottimo, era esprimibile solo quando erano veri 12 item su 13. Possiamo, dunque, dire che, tra gli studenti partecipanti all'esperimento, i soggetti assegnavano pesi diversi alla verità o falsità degli item descrittivi per giungere a una valutazione complessiva del profilo.

Tavola 1
Distribuzione delle valutazioni

Valutazione	0	1	2	3	4	5	N
DIB	0	0	0	1	25	13	39
Profilo	0	0	1	4	18	16	39

Tavola 2
Distribuzione delle risposte "vero"

Numero di risposte "vero"	5	6	7	8	9	10	11	12	13	N
Frequenza	1	0	0	5	5	10	9	7	2	39

Non vi è alcuna relazione significativa tra le valutazioni espresse riguardo al DIB come strumento diagnostico (A) e il numero di item contrassegnati come veri (il valore di probabilità del chi–quadro è 0,4). Da un lato, la stima dell'adeguatezza del profilo di personalità è risultata dipendere, in parte, dal numero di conferme fattuali. Dall'altro, il grado di valutazione positiva del test è risultato non dipendere dal livello di concordanza tra risultati del test e autovalutazioni. Ciò significa che la convalida dello strumento del test rispondeva a una logica del tipo "tutto o niente" e dipendeva da un certo numero minimo di conferme. Il criterio relativo al numero di conferme necessarie per convalidare lo strumento adottato dagli studenti era variabile.

Tutti gli studenti hanno giudicato buono o ottimo il DIB come strumento di misurazione della personalità. La maggior parte di essi può essere accusata di aver commesso un errore logico, avendo giudicato positivamente lo strumento in mancanza di conferme adeguate. Coloro che hanno attribuito allo strumento il punteggio di 5, senza giudicare veri tutti i 13 item, hanno dimostrato di non prendere in considerazione i fatti rappresentati dal loro stesso atteggiamento critico. Si può dire lo stesso di quanti hanno attribuito allo strumento una valutazione più alta di quella attribuita al profilo di personalità. È interessante osservare che lo studente che ha espresso il giudizio più critico nei confronti del profilo di personalità, avendogli attribuito il punteggio di 2 e avendo giudicato veri solo 5 item, ha, però, attribuito al DIB il punteggio di 4.

La Tavola 3 riassume il numero di valutazioni positive attribuite ai 13 item da parte degli studenti. Nessun item ha raggiunto un livello completo di validità universale, sebbene più della metà di essi vi ci siano avvicinati.

Tavola 3
Valutazioni positive attribuite dagli studenti agli item del profilo

Risposta	Numero degli item												
	1	2	3	4	5	6	7	8	9	10	11	12	13
Vero	28	38	23	31	18	35	38	37	34	35	34	12	28
Falso	4	0	1	0	9	3	0	1	3	2	1	9	7
Incerto	7	1	15	8	12	1	1	1	2	2	4	18	4

Ricordo delle valutazioni

Dal momento che molti studenti avevano provato imbarazzo per essersi lasciati "abbindolare", l'autore ha ipotizzato che la dinamica dei processi mnestici avrebbe favorito una attenuazione dei risultati di questa aggressione alla autostima. La classe era stata edotta della distribuzione delle valutazioni. Tre settimane dopo, gli studenti furono informati che, come promesso, l'autore aveva cancellato i loro nomi dai fogli riportanti le valutazioni, ma che gli sarebbe piaciuto confrontare le valutazioni con i voti ottenuti dalle risposte al questionario. Fu loro chiesto se erano disponibili ad annotare, facendo leva sulla memoria, le valutazioni espresse a proposito del DIB e del profilo. Le scale di valutazione furono riportate sulla lavagna. Gli studenti si mostrarono inizialmente scettici, come comprensibile, ma alla fine furono collaborativi. Erano presenti solo 32 studenti sui 39 che erano stati sottoposti al DIB e al profilo.

I risultati hanno più o meno confermato le attese. Per quanto riguarda il punto A (relativo al DIB), non è emersa alcuna tendenza generale: due studenti hanno alzato le loro valutazioni da 4 a 5 e tre le hanno abbassate da 5 a 4. Tuttavia, le valutazioni riguardanti il punto B (relativo al profilo) sono state tendenzialmente abbassate. Sette valutazioni di 5 sono state abbassate a 4 e una di 5 a 3. Nessuna è stata alzata. La Tavola 4 evidenzia le due distribuzioni delle valutazioni relative al profilo.

Tavola 4
Valutazioni originarie e rievocate relative al profilo di personalità

Valutazione	2	3	4	5	N
Originaria	1	3	12	16	32
Rievocata	1	4	19	8	32

Il test t, che misura gli scostamenti dalle medie, mostra significatività al livello dell'1%. Ciò conferma che, tra i più creduloni, si è verificato una significativa riduzione del consenso prestato alla bontà del profilo.

Conclusioni

1. Quanto riferiscono i soggetti che producono pseudo-diagnosi a proposito della validità dei loro metodi e risultati può essere riprodotto o superato in laboratorio senza far ricorso a uno strumento diagnostico. È possibile dimostrare la validità delle stime di personalità alla cieca quando si utilizza il metodo della convalida soggettiva (la conferma del cliente) per item descrittivi contenenti affermazioni di validità praticamente universale.

2. La convalida di uno strumento o di un profilo di personalità tramite convalida soggettiva è una procedura fallace che presuppone l'oggettività dell'autovalutazione e la comprensione degli altri da parte del cliente.

3. Facendo ricorso al metodo della convalida soggettiva, un profilo di personalità fittizio può facilmente indurre a valutare positivamente uno strumento diagnostico, anche quando lo stesso profilo non è valutato in modo interamente positivo. È sufficiente un livello minimo di corrispondenza tra il profilo e l'autovalutazione per suscitare un atteggiamento complessivamente favorevole al profilo; atteggiamento che si estende, in maniera acritica, allo strumento.

4. È probabile che la procedura della convalida soggettiva generi risultati più fallaci nel caso di valutazioni complessive di un profilo di personalità che quando sono valutate individualmente affermazioni specifiche.

5. Quando l'autostima è minacciata, le funzioni della memoria operano in modo tale da allontanare la minaccia e aumentare l'autostima. Queste alterazioni mnestiche sono distorsioni difensive del ricordo, non mere dimenticanze.

6. Gli psicologi clinici e gli altri soggetti che compiono inferenze relative alle caratteristiche della personalità possono essere spinti ad attribuire un livello di significatività eccessivamente alto a queste inferenze. È assolutamente fondamentale che i clinici sottopongano le loro procedure, presupposizioni e, forse, proiezioni a un approfondito esame sperimentale.

Testi consultati

(1) ALLPORT, G. W. (1937). *Personality, a psychological interpretation*. New York: Holt.

(2) BOBERTAG, O. (1934). "Bemerkungen zum Verifikationsproblem", *Zeitschrift für angewandte Psychologie*, vol. 46, pp. 246-249.

(3) COHEN, M. R., NAGEL, E. (1934). *Introduction to logic and the scientific method*. New York: Harcourt, Brace.

(4) CRIDER, B. (1944). "A study of a character analyst", *The Journal of Social Psychology*, vol. 20, pp. 315-318.

(5) FORER, B. R. *A diagnostic interest blank* (in corso di stampa).

(6) HARRISON, R. (1940). "Studies in the use and validity of the Thematic Apperception Test with mentally disordered patients. II. A quantitative validity study. III. Validation by the method of 'blind analysis'", *Character & Personality*, vol. 9, pp. 122-133; 134-138.

(7) HARRISON, R. (1943). "The Thematic Apperception and Rorschach methods of personality analysis in clinical practice", *Journal of Psychology*, vol. 15, pp. 49-74.

(8) HARRISON, R., ROTTER, J. B. (1945). "A note on the reliability of the Thematic Apperception Test", *The Journal of Abnormal Psychology*, vol. 40, pp. 97-99.

(9) HERTZ, M. R., RUBINSTEIN, B. B. (1939). "A comparison of three 'blind' Rorschach analyses", *American Journal of Orthopsychiatry*, vol. 9, pp. 295-314.

(10) MURRAY, H. A., STEIN, M. (1943). "Note on the selection of combat officers", *Psychosomatic Medicine*, vol. 4, pp. 386-391.

(11) PASCAL, G. R., SUTTELL, B. (1947) "Testing the claims of a graphologist", *Journal of Personality*, vol. 16, pp. 192-197.

(12) TAYLOR, D. W. (1947). "An analysis of predictions of delinquency based on case studies", *The Journal of Abnormal Psychology*, vol. 42, pp. 45-56.