

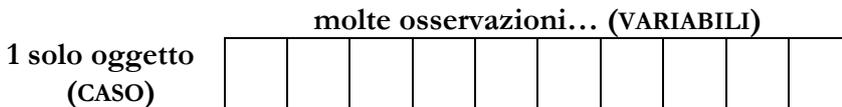
ATTENDIBILITÀ E FEDELTA'

Da: Marradi A. (1990), *Fedeltà di un dato, affidabilità di una definizione operativa* in Rassegna italiana di sociologia, a. XXXI, n. 1, gennaio-marzo 1990, pagg. 55- 96.

ATTENDIBILITÀ'

Il concetto di **attendibilità** nasce in astronomia, in connessione con i concetti di errore di osservazione e di precisione/accuratezza. "L'accuratezza è concepita come una proprietà della coppia costituita da uno strumento e dall'osservatore che lo utilizza. Essa può essere considerata l'inverso della varianza degli esiti di tutte le osservazioni relative allo stesso fenomeno" (Marradi 1990, 55).

In astronomia, in chimica e nelle altre scienze fisiche risulta abituale analizzare l'accuratezza delle proprie osservazioni, misurando più volte lo stesso oggetto. In questi termini il concetto di attendibilità può essere considerato di natura vettoriale:



Nel 1947 Cronbach esplicita due degli assunti su cui si fonda questa concezione:

- a) lo stato dell'oggetto non cambia spontaneamente durante l'intervallo fra la prima e l'ultima osservazione (o quantomeno possono essere considerati trascurabili gli effetti del cambiamento spontaneo rispetto a quelli della coppia strumento-osservatore);
- b) lo stato dell'oggetto non è alterato dal processo di osservazione.

Recentemente Marradi ha esplicitato altri due assunti necessari ad estendere la portata del concetto di attendibilità alle scienze umane, i cui oggetti, anche quando sono dello stesso tipo non sono fungibili. Parlando di 'non fungibilità' non ci riferiamo unicamente alle ovvie diversità biologiche tra individui, ma principalmente alle differenze culturali e sociali che caratterizzano profondamente ognuno di noi.

Questi assunti (pacifici per le scienze fisiche) sono:

- c) esiste uno stato effettivo dell'oggetto sulla proprietà in questione;
- d) le differenze fra oggetti dello stesso tipo sono, se esistono, trascurabili al fine di stabilire l'accuratezza di una coppia strumento-operatore.

Ma nelle scienze umane l'assunto d) è insostenibile. Questa è infatti la differenza epistemologica determinante fra le scienze fisiche e le scienze umane: le unità non possono essere considerate fungibili. D'altra parte anche gli assunti a) e b) hanno suscitato diverse perplessità: il primo perché è altamente improbabile che il pensiero e il comportamento umano non subiscano cambiamenti di rilievo tra una rilevazione e l'altra; il secondo perché l'interazione tra ricercatore e soggetto della ricerca finisce inevitabilmente per influenzare i risultati della rilevazione.

Concentrandosi sulla sostenibilità degli assunti a) - assenza di cambiamento spontaneo tra la prima e l'ultima osservazione e b) - lo stato dell'oggetto non è alterabile dal processo di osservazione - gli psicometrici, a partire dalla fine dell'800 hanno messo a punto alcune tecniche per valutare l'attendibilità. La più famosa è il **test-retest reliability**.

Questa tecnica consiste nel somministrare due volte nel tempo (t1 e t2) un test ad una pluralità di soggetti¹ – cercando di mantenere inalterato il luogo e il modo in cui viene sottoposto il test - e nel confrontare le rilevazioni.

Se nelle scienze fisiche l'accuratezza di una coppia strumento-osservatore è rilevata sulla base di numerose osservazioni dello stesso oggetto in momenti diversi e viene misurata con una grandezza scalare (l'inverso della varianza dei risultati delle osservazioni); nelle scienze umane il concetto di attendibilità – rilevato attraverso il test-retest – assume una natura bivettoriale e dovrebbe per questo produrre una grandezza vettoriale (una cifra per ciascun soggetto):

due sole osservazioni... (VARIABILI)

	t1	t2
molti		
oggetti		
...		
(CASI)		

Ma è qui che l'armamentario statistico è venuto in soccorso degli psicologi. Poiché ogni rilevazione produce un vettore, nel caso in cui gli elementi dei vettori siano (o siano considerati) numeri cardinali, è stato possibile calcolare una "misura" di attendibilità attraverso il coefficiente di correlazione delle due rilevazioni — in questo caso denominato **test-retest reliability coefficient**.

Gli psicometrici hanno abitualmente considerato il coefficiente come un attributo permanente dello strumento impiegato; essi cioè sono tranquillamente passati sopra al fatto che tale coefficiente è stato calcolato su uno specifico campione, e che quindi non dovrebbe essere riferito ad altre popolazioni e ad altre situazioni di rilevazione.

Com'è ampiamente acquisito nella letteratura psicometrica, la tecnica test-retest presenta un rilevante svantaggio: la ripetizione del test produce dei cambiamenti nelle risposte dei soggetti, il che non necessariamente provoca una riduzione del valore del coefficiente di attendibilità test-retest. Infatti si può produrre un innalzamento artificiale dell'attendibilità nel caso in cui i soggetti ricordano nella seconda rilevazione le risposte fornite durante la prima, e le ripetono per il desiderio di apparire coerenti.

Per far fronte a questo inconveniente negli anni venti è stata ideata la tecnica denominata **parallel forms o equivalent forms** con la quale si correlano i punteggi ottenuti su due tests differenti, ma definiti "equivalenti", invece che sullo stesso test somministrato due volte. Il relativo coefficiente è stato ribattezzato **coefficient of equivalence**. Il problema che però insorge in questo caso riguarda l'effettiva equivalenza dei tests. La valutazione da parte del ricercatore dell'equivalenza tra i due stimoli somministrati è arbitraria, e altrettanto arbitraria rischia di essere la valutazione dell'attendibilità.

Negli anni trenta si introduce un'ulteriore semplificazione: si è diviso il test in due parti di uguale lunghezza e si correlano i punteggi ottenuti dai soggetti nelle due metà. Questo coefficiente è stato denominato **split-half reliability coefficient**. Il limite principale di questa tecnica sta nella non-univocità della stima dell'attendibilità. I risultati di tale stima dipendono infatti fortemente dal criterio seguito per la suddivisione degli *items* in gruppi.

A cavallo degli anni trenta il concetto di attendibilità non è, quindi, più centrato sulla ripetizione delle osservazioni nel tempo, ma sulla correlazione di due vettori di punteggi.

¹ Il ricorso a più soggetti è reso necessario dall'insostenibilità dell'assunto d).

La decisiva trasformazione del concetto di attendibilità nella manipolazione matematica di una matrice di cifre è compiuta da **Kuder e Richardson** che, in un celebre saggio del 1937, propongono una formula per calcolare un coefficiente di attendibilità sulla base di una matrice delle correlazioni fra vettori costituiti dai punteggi dei soggetti su ciascuno degli *items* di un solo test. Sulla strada tracciata da Kuder e Richardson – il cui coefficiente presupponeva solo *items* dicotomici - si inseriscono Cronbach con il suo **coefficiente alfa** (basato sull'elaborazione di matrici di punteggi su *items* di qualsiasi formato), Raju con il **coefficiente beta** (che applica i principi di alfa ad una situazione in cui un test è diviso in varie parti con un numero diseguale di *items*), Heise e Bohrnstedt con il **coefficiente omega** (fondato sull'analisi fattoriale) ed Armor con il **coefficiente theta** (fondato sull'analisi delle componenti principali).

La riduzione del concetto di attendibilità ad una manipolazione matematica di due o più vettori ha prodotto due conseguenze:

- 1) Il progressivo allontanamento del concetto di attendibilità usato dagli psicometrici dal corrispondente concetto del senso comune. Viene, in altri termini, definitivamente perso di vista il problema della corrispondenza delle cifre ottenute con gli stati effettivi dei soggetti sulle proprietà;
- 2) L'attenuazione della percezione dei ricercatori delle radicali diversità esistenti fra il concetto di attendibilità (rapporto tra linguaggio e realtà - confronto tra M3 e M1) e quello di validità (confronto tra due concetti che ha in mente il ricercatore – M2), dovuta alla sostanziale corrispondenza delle tecniche per calcolare i coefficienti di attendibilità (trasformati in test di "congruenza tra variabili") e di validità.

FEDELTA'

Il concetto di fedeltà può essere definito come la corrispondenza dello stato effettivo del soggetto nella proprietà al singolo dato che esso rappresenta nella matrice. Sottolineando l'importanza del confronto delle cifre della matrice con gli stati effettivi dei soggetti sulle proprietà considerate, Marradi impiega il termine 'fedeltà' per distinguere il concetto riferito al livello del singolo atto di rilevazione dal concetto di attendibilità degli psicometrici relativo all'intero strumento.

Nel proporre il concetto di "fedeltà" Marradi revisiona quegli assunti ricordati in precedenza:

- a) lo stato di un soggetto può cambiare spontaneamente fra una rilevazione e l'altra;
- b) la rilevazione può alterare tale stato, o registrarne uno anche laddove non ne esiste alcuno;
- c) anche supponendo che, in ogni dato istante, ci sia uno stato effettivo di ciascun oggetto sulla proprietà che ci interessa, nella maggior parte dei casi non si può esser sicuri che lo si conosce, e non si può postularlo uguale alla media o al risultato di qualunque altra elaborazione logico-matematica degli stati osservati.
- d) oltre alle differenze fra individui sono rilevanti anche le differenze fra le situazioni di rilevazione (che includono le differenze fra i rilevatori).

Da questo complesso di assunti consegue che occorre prestare una particolare attenzione ad ogni singolo atto di rilevazione per poter affermare qualcosa riguardo alla corrispondenza del dato prodotto con lo stato del soggetto su una data proprietà. Non ha alcun senso quindi attribuire un qualsiasi grado di attendibilità ad uno strumento.

La letteratura metodologica evidenzia infatti la molteplicità di aspetti, eventi e condizioni che possono influenzare l'esito di una rilevazione: le condizioni fisiche e mentali del soggetto al momento del test, l'atteggiamento di chi sottopone il test, il modo in cui è strutturato il questionario, il linguaggio impiegato, la fretta dell'intervistato o dell'intervistatore... per non parlare degli errori materiali compiuti dagli intervistatori o codificatori.

Il controllo della fedeltà di una rilevazione, soprattutto se non sono disponibili documenti ufficiali con cui confrontare gli stati dichiarati dai soggetti, impone un meticoloso lavoro sul campo. I principali strumenti impiegati per rilevare la fedeltà sono: l'osservazione diretta da parte dei ricercatori, le informazioni fornite dai vicini e conoscenti, ma soprattutto l'intervista di controllo in profondità (**l'intervista sull'intervista**). Quest'ultimo strumento - elaborato e impiegato da alcuni metodologi polacchi - permette di raccogliere informazioni sui processi cognitivi ed emotivi dell'intervistato nell'atto di rispondere a una domanda chiusa.

La fedeltà è dunque una proprietà di natura cellulare (si pone a livello di cella della matrice) che ha l'obiettivo di correggere i propri dati dopo aver appurato la veridicità delle dichiarazioni dei soggetti.