



# TECNICHE DI ANALISI DEI DATI

**AA 2019/2020**

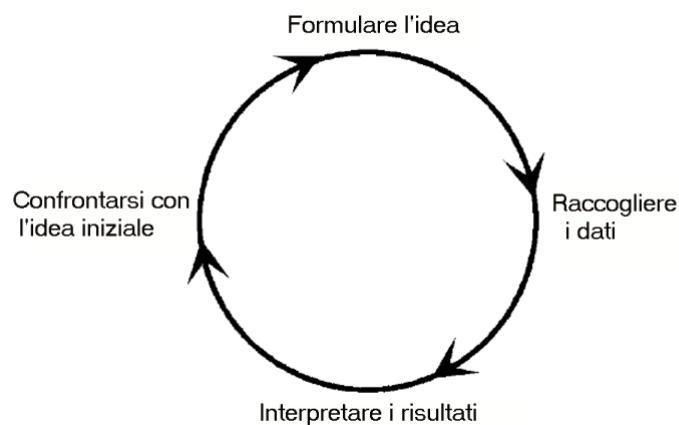
**PROF. V.P. SENESE**

Questi materiali sono disponibili per tutti gli studenti al seguente indirizzo:

[http://psiclab.altervista.org/TecnAnDat2020/2019\\_2020.html](http://psiclab.altervista.org/TecnAnDat2020/2019_2020.html)

Università della Campania «Luigi Vanvitelli» – Dipartimento di Psicologia – TECNICHE DI ANALISI DEI DATI – © Prof. V.P. Senese

## LA RICERCA



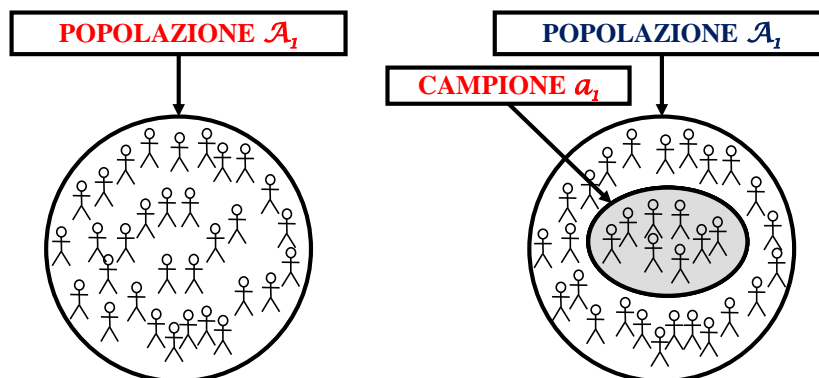
# FASI DELLA RICERCA

CONOSCENZE ◀▶ TEORIA

Analisi della  
letteratura

- identificazione del problema
  - **pianificazione dello studio**
  - formulazione delle ipotesi di ricerca
  - raccolta dei dati
- analisi dei dati
- interpretazione dei risultati

## CAMPIONAMENTO



Insieme degli **elementi** ( $N$ ) cui il ricercatore rivolge la propria attenzione per l'indagine.

Sottoinsieme di elementi dell'insieme popolazione ( $n$ ), ottenuto mediante un'ideale tecnica di **estrazione campionaria** (**campionamento**).

Si dice **representativo** un campione in possesso di tutte le più importanti caratteristiche della popolazione da cui è estratto.

# CAMPIONE

In statistica, la parola **campione** si utilizza in tre sensi distinti, fra loro interrelati:

**(1) insieme delle unità d'analisi** che sono state scelte per esse trattate ed esaminate nel corso di un esperimento (partecipanti)

**(2) insieme dei numeri-misure** che sono stati rilevati e registrati in una particolare esecuzione di un esperimento (valori)

**(3) insieme di variabili casuali**, e quindi esso stesso una variabile casuale (valori)

## ALCUNE DOMANDE DI RICERCA

### Descrizione

**Come è fatto "x"? Quali sono le sue caratteristiche? Quanto sono frequenti?**

Es., Qual è l'aspetto terapeutico nei gruppi di auto-aiuto? Quali sono le risposte verbali utilizzate più di frequente dai terapeuti cognitivi? Quanto frequente è il disturbo borderline di personalità?

### Descrizione - confronto

**Il gruppo "x" è diverso dal gruppo "y"?**

Es., Gli uomini e le donne differiscono nella predisposizione alla genitorialità? Che tipo di interazioni avvengono nelle famiglie con ragazzi aggressivi, rispetto a quelle con ragazzi non aggressivi?

### Correlazione

**C'è una relazione tra la variabile "x" e la variabile "y"? Questa relazione è influenzata da una terza variabile ("z")?**

Es., Il grado di sostegno del marito è associato al superamento della depressione? Le caratteristiche socio-economiche influenzano la relazione tra successo scolastico e autostima?

### Causalità

**La variabile "x" influenza il cambiamento osservato in "y"? La variabile "x" ha un effetto sulla variabile "y" maggiore della variabile "z"?**

Es., Il training genitoriale determina una maggiore *sensitivity* nella relazione adulti-bambini? Nella manifestazione delle condotte devianti adolescenziali il comportamento di quale dei due genitori ha un peso maggiore?

### Misurazione

**Quanto è adeguata la scala "M" nel misurare la variabile "x"?**

Es., La scala PARQ è in grado di misurare l'accettazione-rifiuto genitoriale nella popolazione Italiana? È possibile misurare e distinguere i diversi tipi di conflitto matrimoniale?



# TECNICHE DI ANALISI DEI DATI

AA 2019/2020

PROF. V.P. SENESE

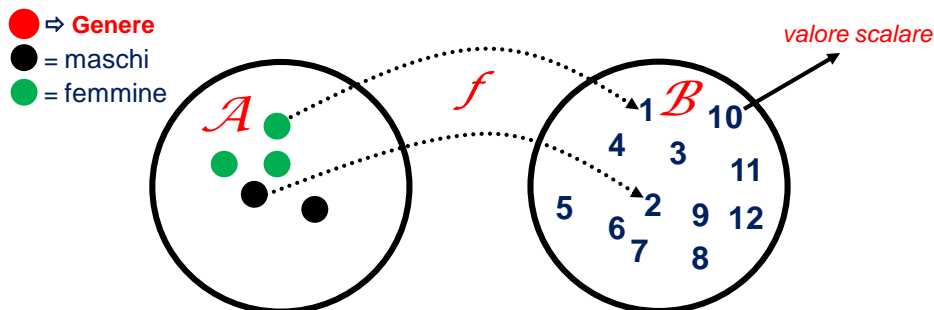
Questi materiali sono disponibili per tutti gli studenti al seguente indirizzo:

[http://psiclab.altervista.org/TecnAnDat2020/2019\\_2020.html](http://psiclab.altervista.org/TecnAnDat2020/2019_2020.html)

Università della Campania «Luigi Vanvitelli» – Dipartimento di Psicologia – TECNICHE DI ANALISI DEI DATI – © Prof. V.P. Senese

## MISURAZIONE

Possiamo definire il concetto di **misura** come una procedura di **CLASSIFICAZIONE** che consenta di attribuire un **oggetto** ( $x \in \mathcal{A}$  {caratteristiche, persone, ecc.}) ad una determinata **classe** ( $\mathcal{B}$ ) stabilendo all'interno della classe una relazione **nota e determinata dall'oggetto**.



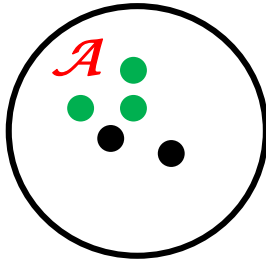
$\mathcal{A}$  = sistema relazionale **empirico** (dominio)  
 $\mathcal{B}$  = sistema relazionale **numerico** (codominio)  
 $f$  = **funzione** o scala di misura

# MISURAZIONE

Quali sono gli oggetti della **Psicologia**?



**LE VARIABILI**



Tutti i comportamenti o tutte le caratteristiche (fisiche o psichiche), gli eventi, gli oggetti, ecc. che **VARIANO** da un caso all'altro o nel tempo e che si ritiene siano in relazione al fenomeno studiato.

**PER LA MISURAZIONE DI UNA STESSA CARATTERISTICA È POSSIBILE DISPORRE DI *DIFFERENTI CRITERI* O PROCEDIMENTI TRA LORO EQUIVALENTI**

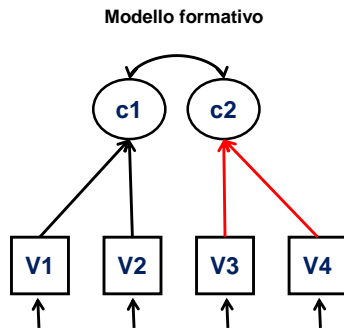
## OPERAZIONALIZZAZIONE

Le **variabili psicologiche** sono dimensioni teoriche (**costrutti**) non osservabili direttamente, ma misurabili mediante l'individuazione di indicatori osservabili di cui sono l'espressione.

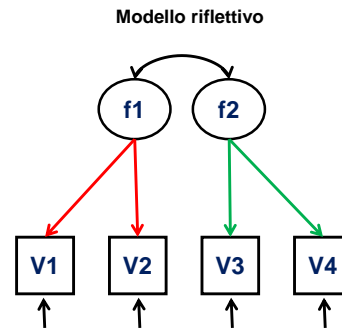
L'**indicatore** è una variabile osservabile che si ipotizza cogliere il costrutto o parte di esso.

La scelta degli **indicatori** non è ovvia. C'è sempre una **componente arbitraria** poiché dipendono da un lato dalle **teorie** dei ricercatori e dall'altro dagli **strumenti** adottati per misurarli.

## MODELLI DI MISURA

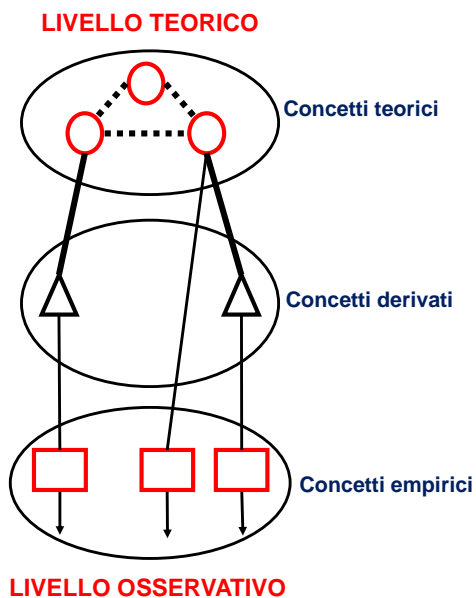


Gli indicatori **formano** il costrutto (*indice*; es., “**stress**” ← trasloco, divorzio, nascita di un figlio, lutto, ecc.).



Gli indicatori **riflettono** (manifestazione osservabile) il costrutto (*scala*; es., “**sensibilità al rumore**” → sono svegliato dal rumore, mi abituo al rumore, mi irrita se i vicini sono rumorosi, ecc.).

## MISURAZIONE



- Definizione del problema (livello di astrazione teorico)
- Definizione teorica (individuazione costrutti)
- Definizione empirica (individuazione degli indicatori)

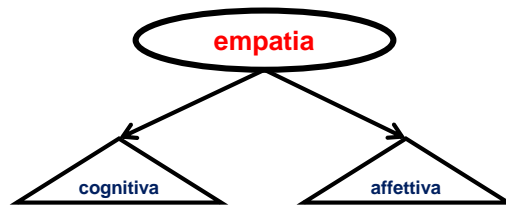
..... PROPOSIZIONE NON OSSERVABILE  
 ——— DEFINIZIONE TEORICA  
 ——— REGOLA DI CORRISPONDENZA  
 → DEFINIZIONE EMPIRICA (METODO)

Bagozzi, 1994

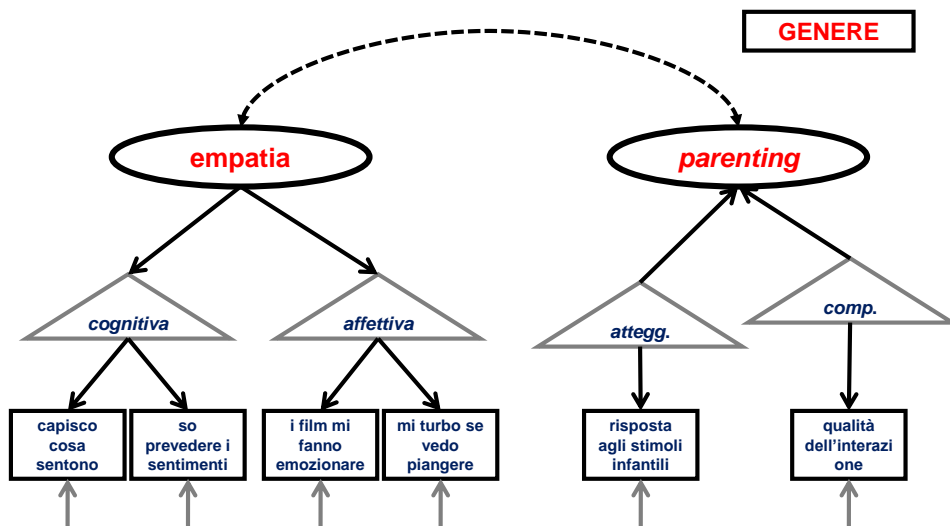
## Es. EMPATIA

L'empatia è la capacità di adottare il punto di vista psicologico di un'altra persona per comprenderne i pensieri, i sentimenti e le emozioni (Baron-Cohen, 2011).

L'empatia è caratterizzata da due dimensioni: affettiva (emozioni) e cognitiva (mettersi nei panni dell'altro; Baron-Cohen & Wheelwright, 2004; Davis, 1983).



## Es. EMPATIA & PARENTING



## Es. VARIABILI #1

SESSO  
 ETÀ  
 ISTRUZIONE  
 OCCUPAZIONE  
 REDDITO  
 CONDIZ. SPERIM.  
 N° ERRORI  
 VELOCITÀ DELLA R  
 ESPLOAZIONE SPAZ.  
 ATTACCAMENTO  
 FREQ. CARDIACA  
 CARATTERISTICHE GENETICHE  
 ∞

SESSO	N° ERRORI
ETÀ	VELOCITÀ R
ISTRUZIONE	ESPLORAZ. S.
OCCUPAZIONE	FOCUS ATT.
REDDITO	CARA. GENET.
FREQ. CARD.	MEMORIA
COND. SPERIM.	ATTACCAM.

## Es. VARIABILI PSICOLOGICHE

### ATTITUDINI E ABILITÀ

Abilità intellettive  
 Abilità amministrative  
 Capacità motorie  
 Capacità linguistiche  
 Capacità sociale  
 Successo scolastico  
 Creatività

### ATTEGGIAMENTI E OPINIONI

Atteggiamenti politici  
 Atteggiamenti sociali  
 Giudizi  
 Preferenze estetiche  
 Morale  
 Valutazioni  
 Umore

### CARATTERISTICHE DI PERSONALITÀ

Motivazioni  
 Temperamento  
 Carattere  
 Valori  
 Interessi

### CARATTERISTICHE FISICHE

Dimensioni del fisico  
 Variabili biologiche  
 Risposte fisiologiche  
 Fenomeni percettivi  
 Sintomi fisici

Comrey, 1995



# MISURE PSICOLOGICHE

## #Performance



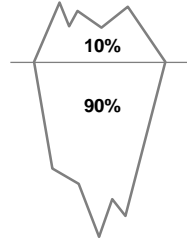
Risposte comportamentali  
Correttezza (C/E)  
Tempo di reazione o Latenza  
(reaction time, RT)

## #Self-report



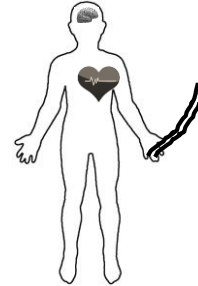
Risposte verbali e  
comportamentali a test,  
questionari o interviste.

## #Implicite



Risposte comportamentali  
Automatiche o Implicite  
raccolte mediante procedure  
sperimentali (es. IAT).

## #SN



Risposte psicofisiologiche  
raccolte mediante specifiche  
apparecchiature (es., fMRI,  
EEG, ERP, Respiro, Battito  
cardiaco, Conduttanza  
cutanea, Ecc.).

## LE VARIABILI #1

*LA CLASSIFICAZIONE DIPENDE DAL  
FENOMENO INDAGATO E DALLE IPOTESI*

### INDIPENDENTI

la cui variazione o manipolazione determina o influenza  
altre variabili

### DIPENDENTI

la cui variazione è determinata dal variare della variabile  
indipendente

### DI DISTURBO

la cui variazione non influenza il fenomeno indagato  
(confondente) o che lo influenza in modo indiretto  
(confusa).

## LE VARIABILI #3

### **MODO X**

quando viene controllata facendole assumere solo determinati valori noti al ricercatore (es., variabile indipendente);

### **MODO Y**

quando viene controllata facendola variare liberamente e registrandone i valori che assume (es., variabile dipendente);

### **MODO K**

quando viene controllata facendole assumere solo uno specifico valore (costante; es., variabile di disturbo);

### **MODO M**

quando viene controllata creando dei sotto-gruppi nei quali la variabile si manifesti allo stesso modo (es., variabile di disturbo);

### **MODO R**

quando non viene controllata ma fatta liberamente variare, assumendo che si distribuisca casualmente (es., variabile di disturbo);

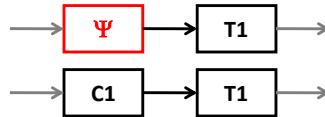
## I TIPI DI RICERCA

Nella psicologia esistono tre tipologie principali di disegni di ricerca che si differenziano in base al **controllo del ricercatore** sulle variabili indipendenti, di disturbo o confondenti:

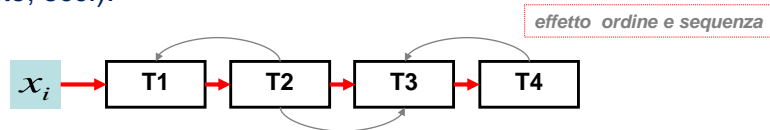
- (1) i **veri esperimenti**
- (2) i **quasi esperimenti**
- (3) la **ricerca osservazionale o correlazionale**

## I DISEGNI DI RICERCA

Quando si utilizzano delle **misurazioni indipendenti**, ovvero relative a gruppi distinti il disegno viene detto tra i soggetti o a misure indipendenti (*between subjects*).

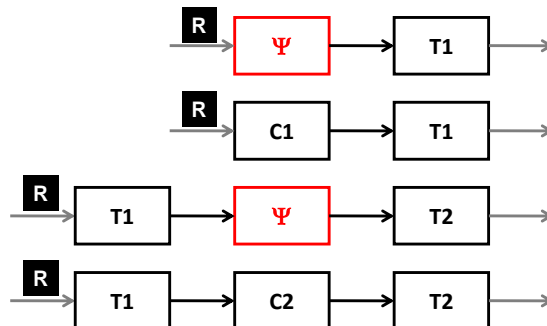


Quando si utilizzano delle **misurazioni successive** relative agli stessi soggetti il disegno viene detto entro i soggetti, a misure ripetute o dipendenti (*within subjects*). In questo caso non si parla di gruppi diversi ma di  $k$  condizioni diverse o momenti diversi (es. prima e dopo il trattamento, ecc.).



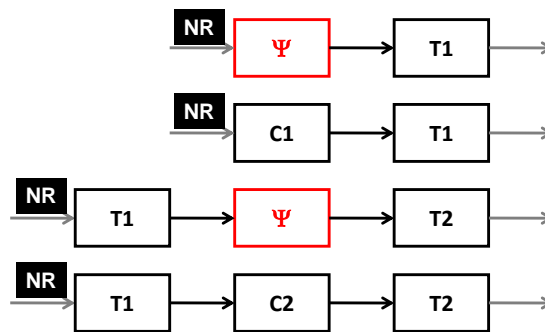
## I VERI ESPERIMENTI

Studi di laboratorio che permettono il **massimo controllo sulla variabile indipendente** o sulla manipolazione di interesse (variabili). Per questo consentono la **base più consistente per un'interpretazione causale**. In questa tipologia di ricerche il ricercatore è in grado di **assegnare i partecipanti** alle diverse condizioni sperimentali in **modo casuale** (*randomized designs oppure randomized controlled trials [RCTs]*).



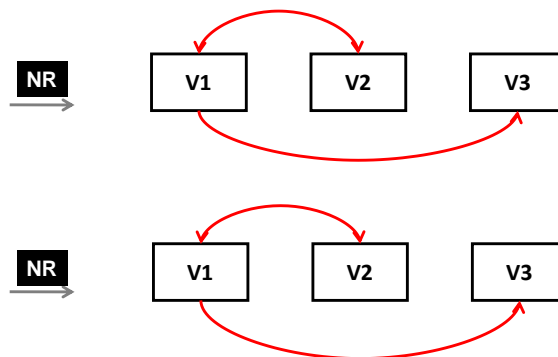
## I QUASI ESPERIMENTI

Disegni sperimentali in cui non è possibile tenere completamente sotto controllo tutti gli aspetti implicati (variabili) nello studio.

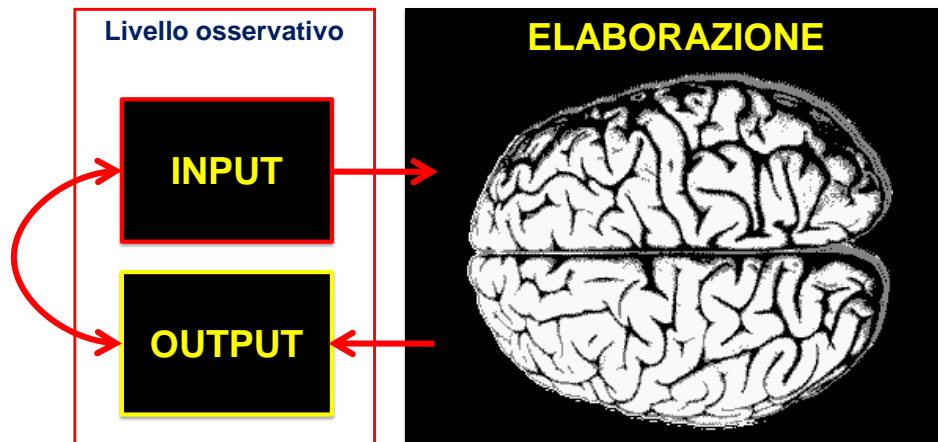


## RIC. OSSER./CORREL.

Le ricerche **passivo-osservazionali** o **correlazionali** consentono di osservare le relazioni tra variabili, ma non implicano alcuna manipolazione delle variabili. Possono essere valutate sia relazioni **associative** sia **causali**.



## SCHEMATIZZANDO



## TECNICHE DI ANALISI DEI DATI

**AA 2019/2020**

**PROF. V.P. SENESE**

Questi materiali sono disponibili per tutti gli studenti al seguente indirizzo:

[http://psiclab.altervista.org/TecnAnDat2020/2019\\_2020.html](http://psiclab.altervista.org/TecnAnDat2020/2019_2020.html)

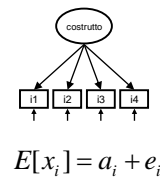
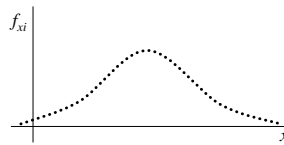
## LA TCT

Secondo la Teoria Classica del Testing (TCT) in ogni misura sono riconoscibili due componenti: una vera e una d'errore:

$$x_i = v_i + e_i$$

La componente d'errore non è conoscibile, ma viene definita come **casuale (distribuzione Gaussiana)**. Da ciò ne conseguono alcune proprietà:

$$\begin{aligned} \bar{e}_i &= 0 \\ r_{e_i x_i} &= 0 \\ r_{e_1 e_2} &= 0 \end{aligned}$$



$$E[x_i] = a_i + e_i$$

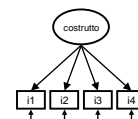
40

## ATTENDIBILITÀ

Le tecniche classicamente utilizzate per il calcolo dell'attendibilità assumono:

- (1) che la misura di interesse rappresenti un *costrutto unico* (**unidimensionalità**);
- (2) che la misura sia **valida**, ovvero che la variabilità sistematica dei punteggi sia dovuta *esclusivamente al costrutto misurato* e non ad altri **costrutti intervenienti**.

Il punto centrale dei metodi di calcolo dell'attendibilità è avere delle **forme multiple** o misurazioni multiple della stessa misura (es., item, scale, osservatori, ecc.). In questo modo maggiore è la **correlazione** tra le forme maggiore è la **variabilità sistematica**.



## LA TCT

L'attendibilità è dunque l'**accuratezza** con cui un test misura un certo attributo psicologico ed è **inversamente proporzionale** all'**errore casuale** presente nella misurazione stessa.

$$\text{Varianza totale} = \text{Var. vera} + \text{Var. errore}$$



$$r_{tt} = \frac{\text{Var. vera}}{\text{Var. totale}} = \frac{\sigma_v^2}{\sigma_x^2}$$

L'indice di attendibilità può essere usato per calcolare il **coefficiente di correlazione** tra i punteggi **osservati** ( $X$ ) e i punteggi **veri** ( $V$ ) che viene detto **indice di fedeltà** e si ottiene facendo la radice quadrata dell'indice di attendibilità.

$$r_{VX} = \sqrt{r_{tt}}$$

42

## LA TCT

Da un punto di vista empirico, esistono differenti metodi per la stima dell'**indice di attendibilità** o di **precisione** (su **scala percentuale**) della misura.



Regola Pratica (Nunnally, Bernstein, 1994):

- [.90 - 1] Ottima;
- [.80 - .89] Buona;
- [.70 - .79] Discreta;
- [.60 - .69] Sufficiente;
- [ 0 - .59] **Inadeguati**.

Come sottolineato da Cronbach, **non esistono "differenti procedimenti"** per il calcolo dell'attendibilità, ma **esistono aspetti differenti** dell'attendibilità.

Ogni metodo misura qualcosa di diverso e fornisce delle indicazioni specifiche.

**Metodi:**

- *forme parallele*
- *test-retest*
- *split-half*
- *alfa di Cronbach o KR<sub>20</sub>*
- *coefficiente di Rulon*
- *coefficiente L<sub>2</sub> di Guttman*
- ...

43

# ALFA DI CRONBACH

L'**alfa di Cronbach** è la misura di attendibilità (**coerenza interna**) maggiormente diffusa in letteratura e utilizzabile quando la scala è **unidimensionale** e le **forme parallele** (item) hanno la **stessa varianza vera** ( $v$  [*upsilon*] oppure  $\tau$  [*tau*]).

$$\alpha_r = K \cdot \frac{\bar{r}}{1 + (K-1) \cdot \bar{r}}$$

**Alfa standardizzato.** Nel caso di forme **perfettamente parallele** (stessa varianza vera e stessa varianza di errore).

$$\alpha_r = K \cdot \frac{c\bar{v}}{\bar{s}^2 + (K-1) \cdot c\bar{v}}$$

**Alfa non standardizzato.** Nel caso di forme **tau-equivalenti** (stessa varianza vera [ $\tau$ ], ma diversa varianza di errore).

$$\rho_{tt} = \frac{\sum(\lambda_i)^2}{\sum(\lambda_i)^2 + \sum\delta^2}$$

**Item congenerici** [*rho*]. Nel caso di forme con **diversa** varianza vera e **diversa** varianza di errore analizzate con tecniche di **CFA** (Raykov, 1997).

## LA VALIDITÀ

Scopo della **ricerca** è giungere a delle conclusioni ben fondate, ovvero **valide** (Kazdin, 1996), sul **nesso causale** tra alcune **variabili oggetto di indagine** (*preferibilmente manipolate sperimentalmente*, **VI**) e il **fenomeno di interesse** (**VD**).

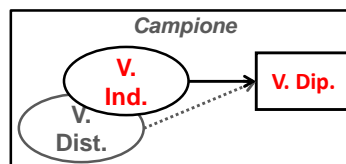
Sono stati identificati **4 principali tipi di validità** (Cook & Campbell, 1979):

- (1) **interna**
- (2) **esterna**
- (3) **di costruito**
- (4) **statistica**



## VALIDITÀ INTERNA

Il grado in cui l'**intervento** (v. ind.), **piuttosto che altre cause**, è **responsabile dei risultati**, delle variazioni osservate o delle differenze tra i gruppi.

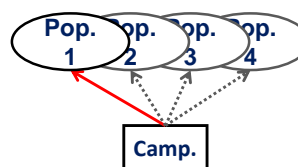


In ciascuna ricerca diversi sono i fattori o **minacce** che possono agire **limitando la validità dello studio**:

- (1) la storia
- (2) la maturazione
- (3) le prove o misure ripetute
- (4) la strumentazione
- (5) la regressione statistica
- (6) il campionamento
- (7) l'abbandono o la mortalità sperimentale
- (...) ...

## VALIDITÀ ESTERNA

Il grado in cui è possibile **estendere i risultati alla popolazione**. Vale a dire **la generalizzazione dei risultati** ad altre persone, ad altri ambienti, ad altri momenti, ad altre misure, ecc.

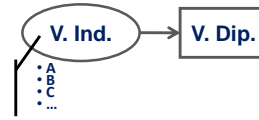


Le **minacce** sono tutti quei fattori specifici che possono limitare **la generalizzabilità e quindi la validità esterna**:

- (1) il campionamento
- (2) lo sperimentatore
- (3) le caratteristiche dell'ambiente
- (4) l'interferenza dovuta al trattamento multiplo
- (5) l'effetto novità e valutazione
- (6) l'intrusività delle misure
- (7) la sensibilizzazione dovuta alla misura
- (...) ...

# VALIDITÀ DI COSTRUTTO

Il grado in cui è possibile **definire le componenti responsabili dell'effetto osservato**. Vale a dire **la capacità di individuare le cause della relazione causale**.

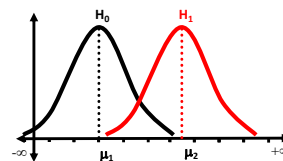


Le **minacce** che possono limitare **la validità di costrutto**:

- (1) effetto osservazione (attenzione) o placebo
- (2) l'operationalizzazione della VI e dei suoi livelli
- (3) le aspettative dello sperimentatore
- (4) le informazioni date ai partecipanti
- (...) ...

# VALIDITÀ STATISTICA

Il grado in cui è possibile **verificare una relazione tra le variabili**. Vale a dire **la capacità di creare le condizioni adatte alla verifica delle ipotesi**.



Le **minacce** che possono limitare **la validità statistica**:

- (1) bassa potenza statistica ( $1 - \beta$ )
- (2) la variabilità nelle procedure di raccolta dati
- (3) l'eterogeneità dei partecipanti
- (4) la bassa attendibilità (affidabilità) delle misure
- (5) la necessità di confronti multipli (aumento dell'errore  $\alpha$ )
- (...) ...

# MISURA

Aspetto	Definizione
<b>ATTENDIBILITÀ</b>	grado in cui una misura rimane <b>stabile</b> nel corso di molteplici misurazioni e grado di <b>consistenza interna</b> della misura
<b>VALIDITÀ</b>	
<b>CONTENUTO</b>	<b>rappresentatività</b> degli item come misura del costrutto
<b>FACCIATA</b>	capacità degli item di misurare il <b>costrutto</b>
<b>COSTRUTTO</b>	grado in cui la misura <b>riflette</b> il costrutto accuratamente
<i>convergente</i>	associazione della misura considerata con <b>altre misure</b> dello stesso costrutto
<i>discriminante</i>	<b>mancanza</b> di correlazione tra la misura e misure di costrutti diversi
<b>CRITERIO</b>	associazione della misura considerata con altre <b>misure esterne</b> in relazione con essa
<i>predittiva</i>	se la misura viene eseguita <b>prima</b> del criterio
<i>concorrente</i>	se la misura e il criterio sono eseguite <b>contemporaneamente</b>