



Validità e  
Attendibilità

---

# Che cosa si intende per *validità*?

La validità è la capacità di un indicatore di *rappresentare* empiricamente un concetto.

Il concetto di validità si riferisce alla capacità di un indicatore (o di più indicatori) di rappresentare empiricamente un dato concetto; una misura è valida se misura ciò che si intende misurare.

Esistono quattro approcci alla validazione degli strumenti: la validità di facciata, la validità di contenuto, la validità di criterio e la validità di costrutto.

- la validità di facciata
- la validità di contenuto
- la validità di criterio
- la validità di costrutto

# Validità di contenuto (content validity )

- La validità di contenuto è per certi versi molto simile alla validità di facciata e si concretizza nell'individuazione di un insieme di indicatori (item) in grado di garantire un'adeguata copertura semantica delle varie dimensioni rilevanti che compongono un dato costruito teorico (Neuman, 2000).
- La convalida dello strumento avviene su un piano logico e teorico piuttosto che empirico.

# La validità di criterio (criterion validity )

---

- La validità di criterio si traduce concretamente nella comparazione di un nuovo indicatore con una misura dello stesso costrutto (il criterio) di cui è stata già accertata o si accerterà la validità.
- Un test antidroga da eseguire su campioni di urina può rappresentare un criterio per validare le risposte a una o più domande sul consumo di droghe nell'ambito di un'inchiesta campionaria (Schutt, 2006).
- Ci sono due tipi di validità di criterio:
  - la validità *concorrente* è caratterizzata dal fatto che sia la nuova misura che il criterio sono empiricamente disponibili.
  - la validità *predittiva* è caratterizzata dal fatto che le informazioni sulla variabile criterio non sono ancora disponibili e verranno rilevate in un prossimo futuro.
- Un esempio di validità concorrente è quello di somministrare un nuovo test di intelligenza a un campione rappresentativo di studenti universitari e verificare il grado di correlazione del punteggio del test con la valutazione all'esame di maturità (criterio).
- Un esempio di validità predittiva è invece lo studio della correlazione tra il punteggio conseguito all'esame di maturità (test iniziale) e il punteggio che gli stessi studenti conseguiranno all'esame di laurea (test finale o criterio).

# Validità di facciata (face validity )

---

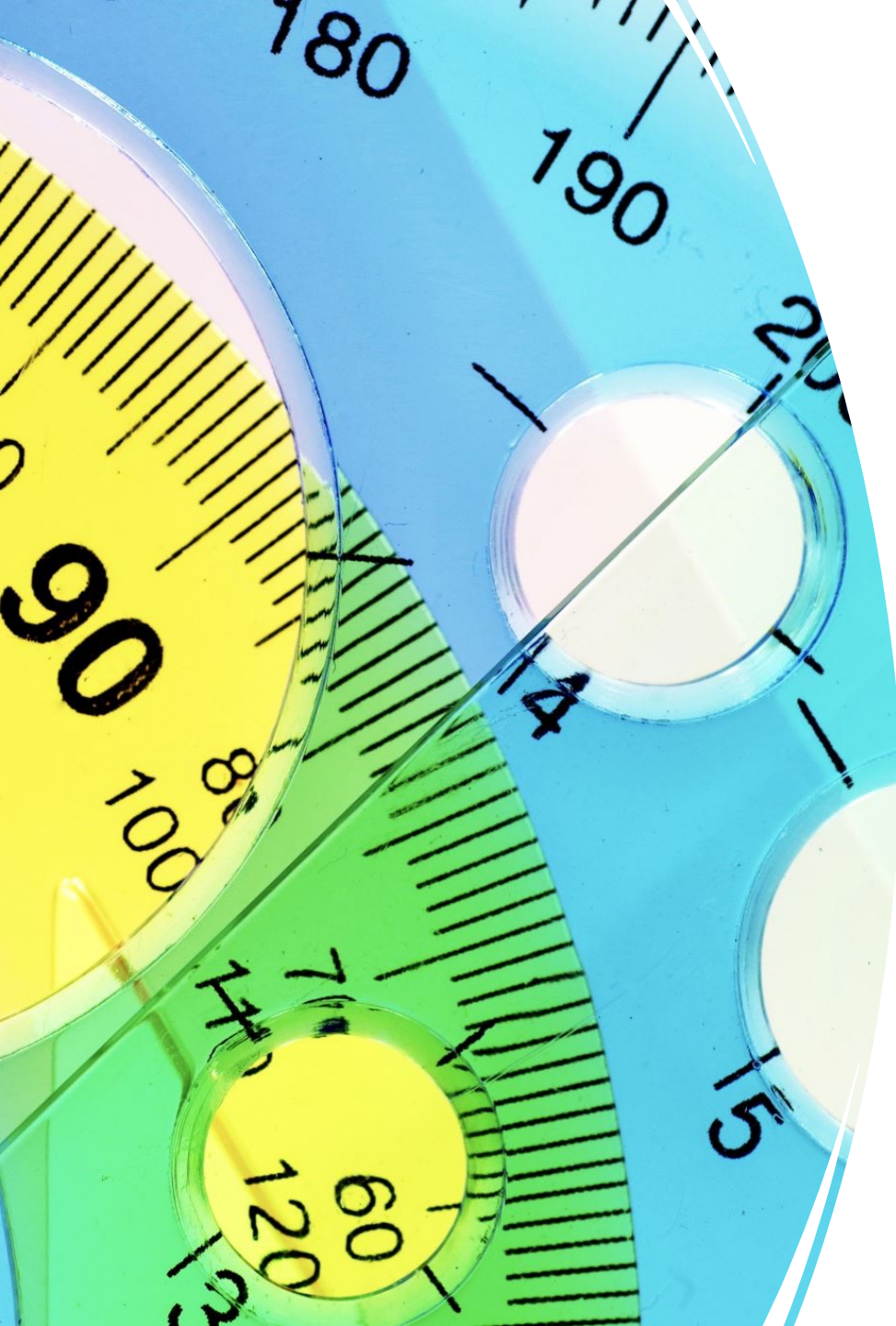
- La validità di facciata è essenzialmente un processo intuitivo, *prima facie*, di valutazione degli indicatori che rappresentano adeguatamente il contenuto del concetto.
- Tale valutazione di plausibilità e ragionevolezza nella scelta degli indicatori viene espressa da uno o più osservatori che hanno accumulato una certa esperienza entro un dato campo disciplinare
- Il numero di sigarette fumate da un soggetto nel corso dell'ultima settimana può rappresentare di primo acchito una misura valida del grado di dipendenza dei soggetti dalla nicotina. A uno sguardo più attento potremmo scoprire che i fumatori incalliti tendono a sottostimare il reale consumo di sigarette. Per tale ragione la validità di facciata non sempre esprime una misura convincente della validità (Schutt, 2006).

# La validità di costrutto (Construct Validity)

---

- La validità di costrutto valuta il grado di accordo tra indicatori che misurano uno stesso concetto teorico (validità convergente) e il grado di disaccordo tra indicatori che misurano costrutti differenti (validità divergente) (Campbell & Fiske, 1959).
- Mentre due o più indicatori di uno stesso costrutto dovrebbero mostrare una **forte** correlazione statistica per poter essere considerati misure valide del medesimo costrutto, due o più indicatori di costrutti diversi dovrebbero avere una bassa correlazione affinché i due costrutti possano essere considerati semanticamente diversi.
- Se vogliamo studiare la relazione tra “qualità del lavoro” e “benessere mentale” potremmo utilizzare come indicatori del primo costrutto la soddisfazione per la retribuzione e per la possibilità di intraprendere un percorso di carriera e come indicatori del secondo costrutto la soddisfazione verso la vita in generale e la percezione che la vita è ricca di opportunità



The image shows a close-up of a circular scale with multiple colored segments: yellow, green, blue, and pink. The scale has various numerical markings, including 180, 190, 200, 90, 100, 80, 14, 15, 60, and 120. A thin needle is visible, pointing towards the 14 mark on the blue segment.

# Che cosa si intende per *attendibilità*?

---

L'A. è sempre riferita all'affidabilità e alla precisione dello strumento di misurazione.

Si misura l'attendibilità con:

- Inter-rater
- Test-retest
- (Parallel forms)
- Internal consistency
- (Split-half)

# Attendibilità tra valutatori (inter-rater reliability )

---

- Il primo tipo di attendibilità consiste nel verificare la congruenza delle valutazioni espresse da *due o più* valutatori indipendenti (osservatori) in riferimento a uno o più indicatori di un tratto latente che si intende misurare.
- Per esempio, possiamo avere due sociologi che valutano il livello di attività di un gruppo di giovani oppure due ricercatori che valutano la bellezza di un campione di giovani donne.
- Il grado di corrispondenza o di associazione tra le risposte date dai due valutatori rappresenta una stima di inter-rater reliability.



# Attendibilità test-retest (test-retest reliability )

---

- Tale tipo di attendibilità esprime la *congruenza* tra due misurazioni dello stesso costrutto effettuate sugli stessi individui, nelle medesime circostanze e in momenti temporali diversi, nell'ipotesi che il fenomeno di studio mostri una certa stabilità.
- Per ottenere una stima della test-retest reliability basta calcolare la correlazione statistica tra due misurazioni.
- Per esemplificare, si potrebbe misurare la felicità chiedendo a un campione di soggetti quanto si sentono soddisfatti della loro vita su una scala a 10 posizioni, rimisurare a distanza di un mese e calcolare la correlazione statistica tra le due misurazioni.

# Metodo dello split-half (split-half reliability )

---

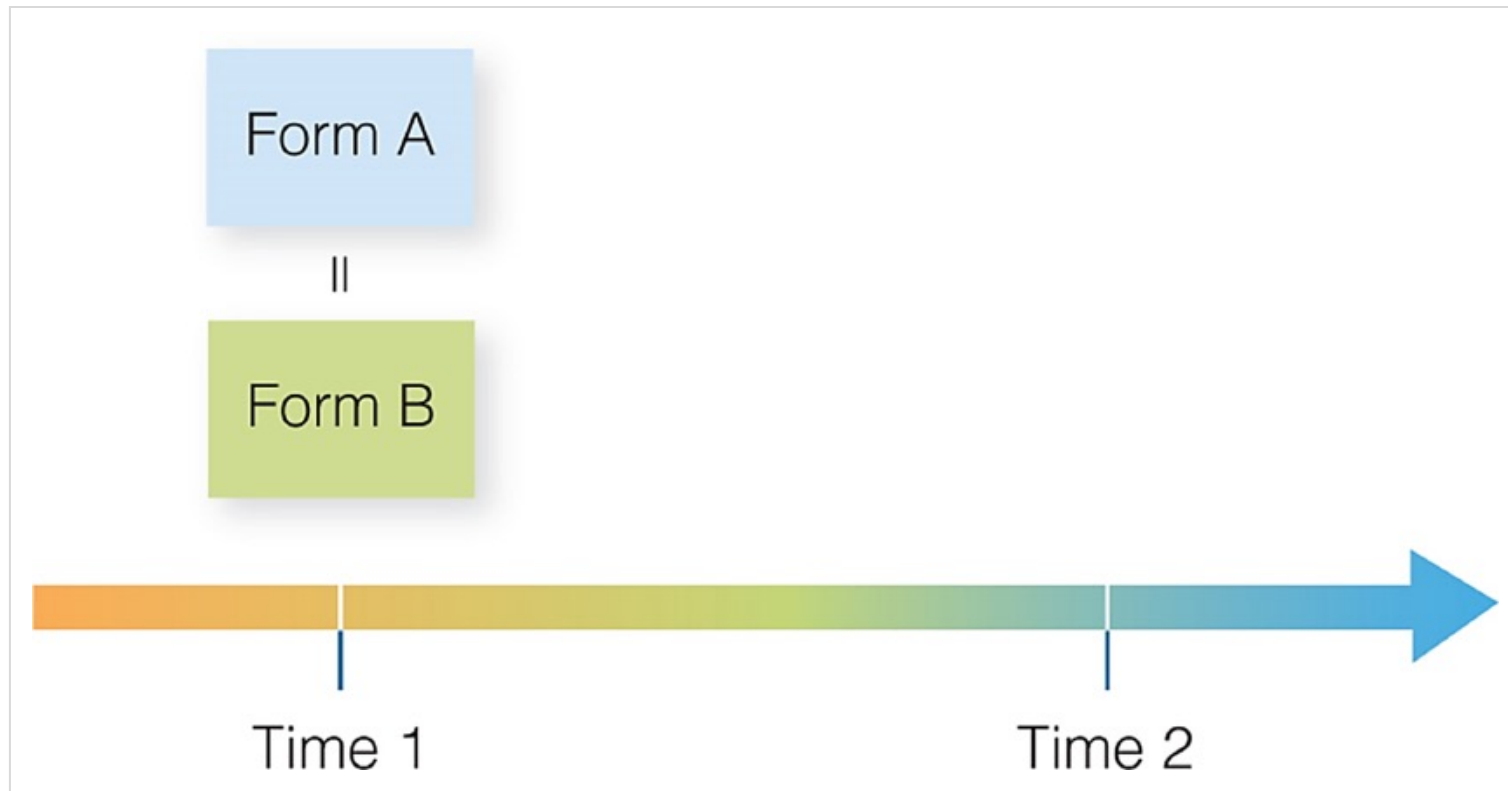
- Il metodo split-half verifica la coerenza tra due gruppi di indicatori che rappresentano *forme parallele* di un medesimo costrutto.
- In altre parole, si tratta di dividere lo strumento di misurazione in due metà e si considerano le due metà come strumenti tendenzialmente equivalenti.
- La correlazione tra i punteggi totali delle due misure rappresenta una stima della split-half reliability.
- Quanti più indicatori vengono utilizzati per misurare un dato costrutto, tanto più è probabile che le due metà dello strumento siano associate.
- Ne consegue che con strumenti di misurazione lunghi, connotati cioè da un elevato numero di indicatori, si finisce per sovrastimare questo tipo di attendibilità.

# Metodo della coerenza interna (internal consistency reliability )

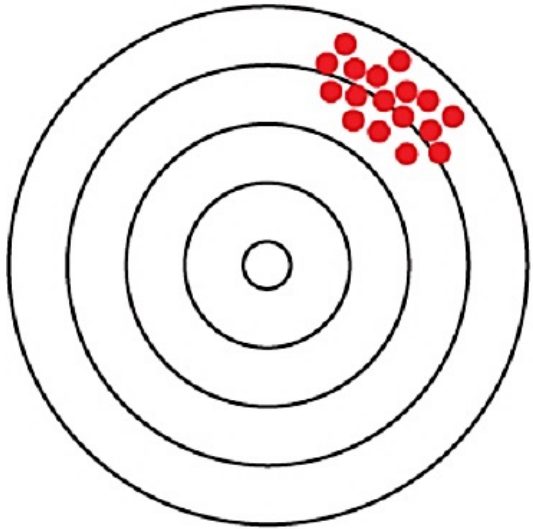
---

- Il metodo della coerenza interna si basa sulla valutazione di un certo numero di indicatori, che formano un dato strumento di misura e che vengono somministrati a un gruppo di soggetti in una singola occasione temporale.
- Gli items vengono giudicati attendibili se si muovono nella stessa direzione e se sono fortemente correlati uno all'altro.
- In altre parole, perché ci sia coerenza interna occorre che i soggetti dello studio esprimano valutazioni simili sugli indicatori prescelti.
- L'attendibilità di questo tipo può essere stimata facendo una media delle correlazioni bivariate tra gli indicatori prescelti oppure calcolando il coefficiente Alpha di Cronbach.

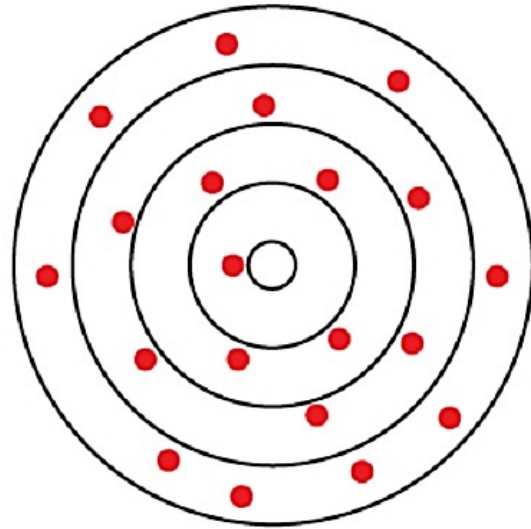
# Forme equivalenti (parallel forms)



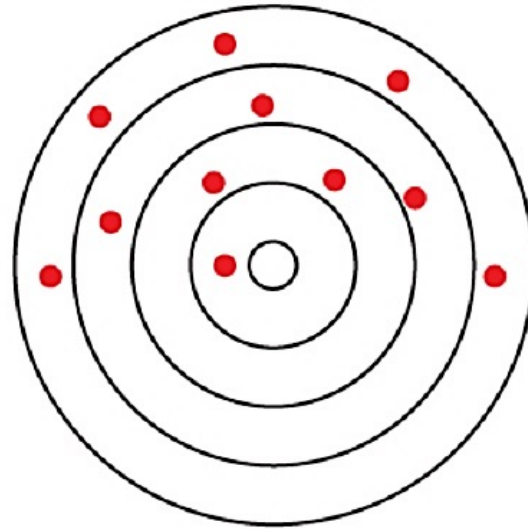
# Integrare Validità e Attendibilità



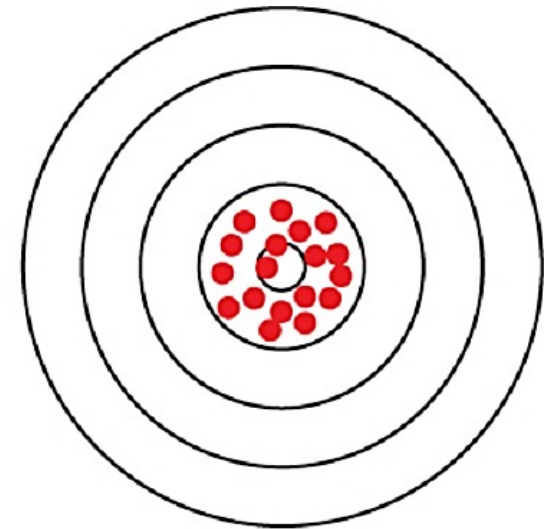
Reliable,  
not valid



Valid,  
not reliable



Neither reliable  
nor valid



Both reliable  
and valid

# Esercizio

- <https://www.psicologionline.net/test-psicologici-psicologia/test-dipendenza-gioco-azzardo>
- Raggruppare gli indicatori usati entro delle dimensioni considerate dal questionario.
- Indicare un possibile esempio di criterio per attendibilità e uno per validità che i redattori potrebbero usare per testare lo strumento.