

Modelli di misurazione in psicometria

Carlo Chiorri, PhD

Principi di psicometria

Psicometria ingenua e scientifica

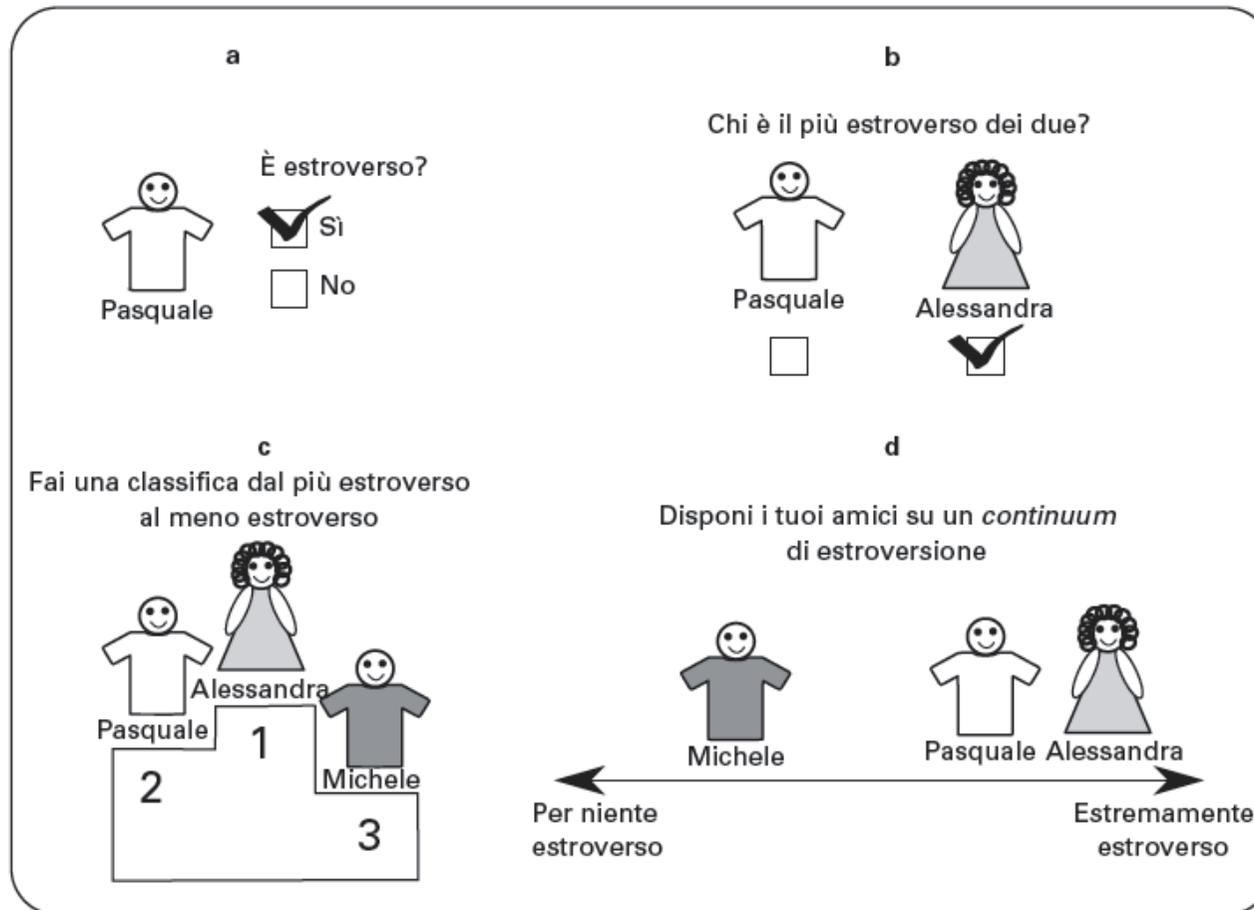
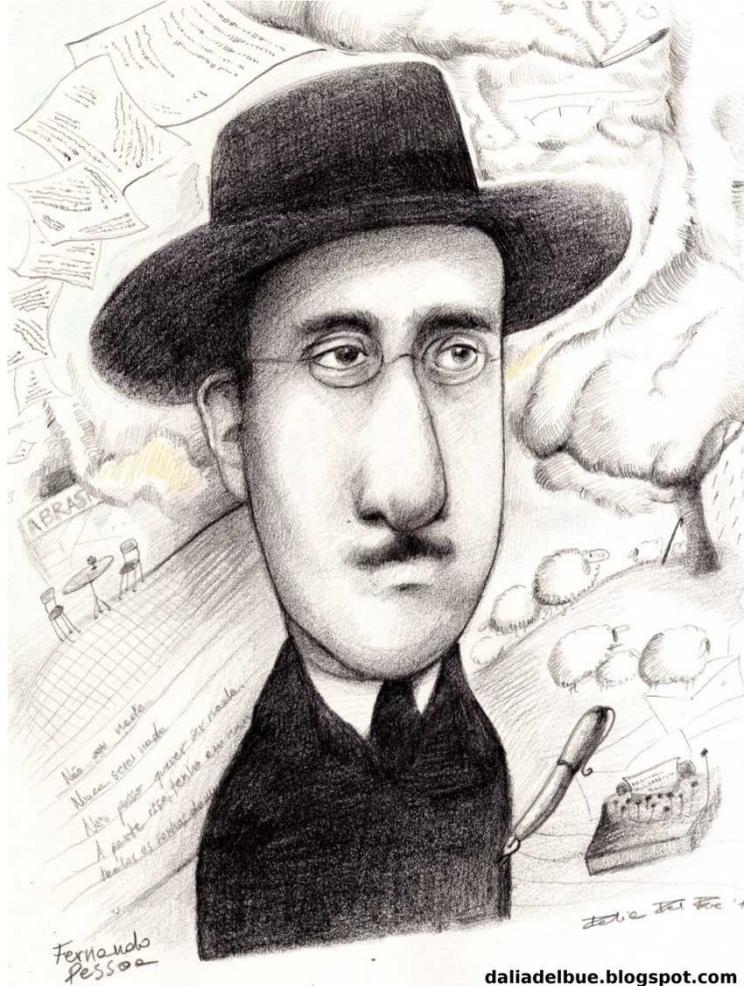


Figura 1.1 Stadi successivi della valutazione ingenua dell'estroversione.



Minha alma é uma orquestra oculta;
não sei que instrumentos tange e
range, cordas e harpas, tímboles e
tambores, dentro de mim. Só me
conheço como sinfonia.

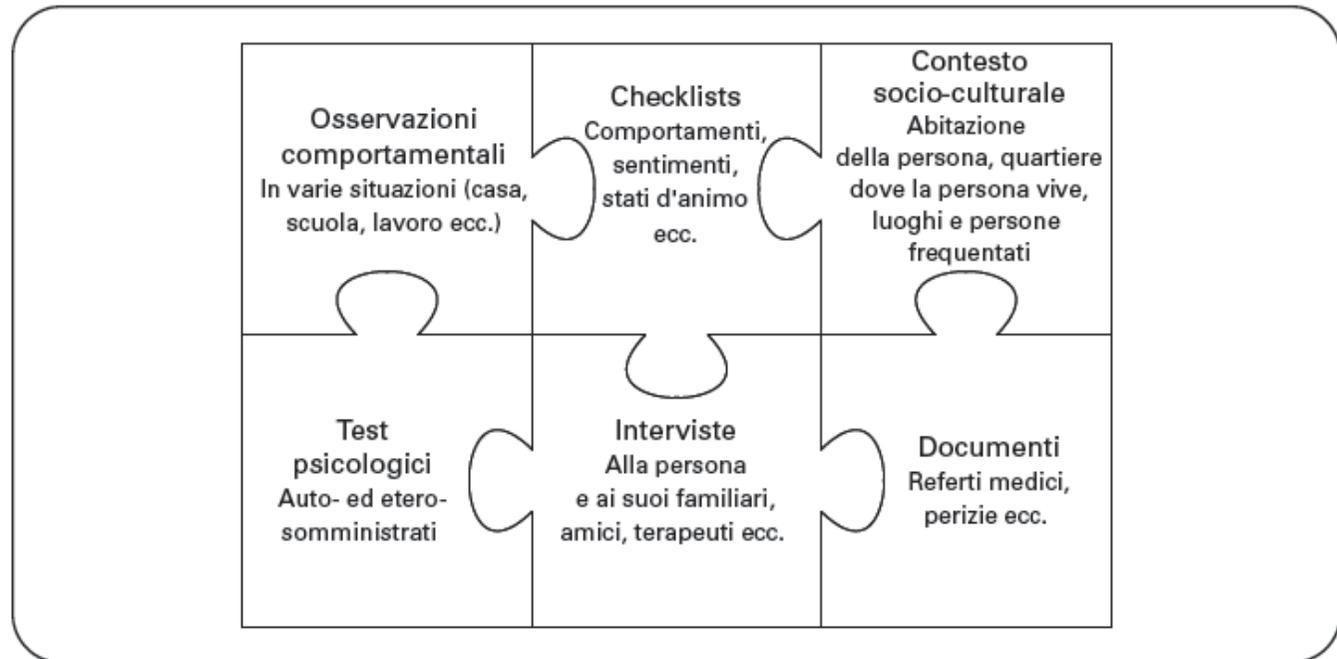
Fernando Pessoa
Livro do Desassossego

[La mia anima è una misteriosa
orchestra; non so quali strumenti suoni
e strida dentro di me: corde e arpe,
timpani e tamburi. Mi conosco come
una sinfonia]

Fernando Pessoa
Libro dell'inquietudine

Assessment psicologico

Figura 1.2 Il complesso mosaico di elementi che compongono l'assessment psicologico.



- Si misurano *attributi* e non persone!

Misurazione

- Fechner e la psicofisica
- *Candela* = quantifica la risposta dell'occhio umano alla luce ed è quindi basata su una risposta biologica alla radiazione ottica (Sistema Internazionale di Unità di Misura)

Misurabilità in psicologia

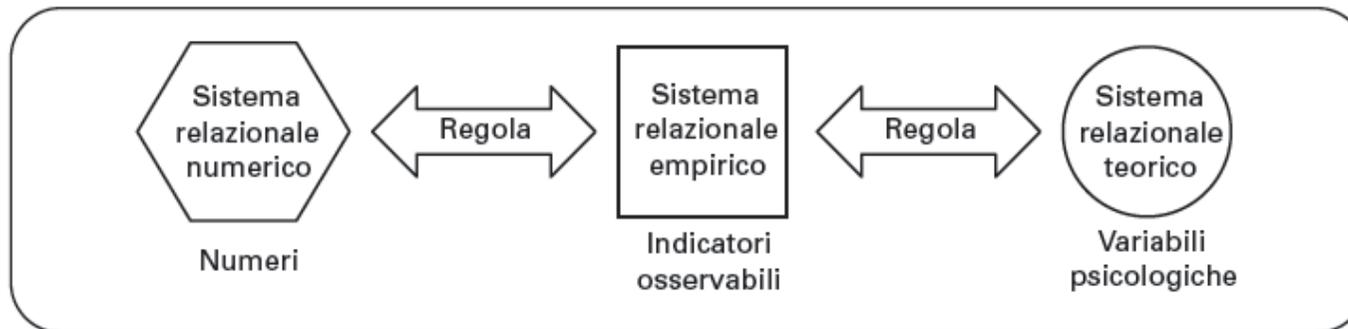
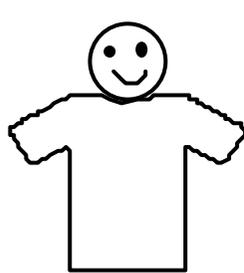


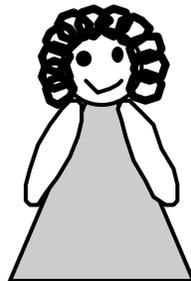
Figura 1.4 Connessioni fra i sistemi relazionali implicati nella misurazione.

Il problema della misurazione in psicologia

- Struttura additiva



QI = 116



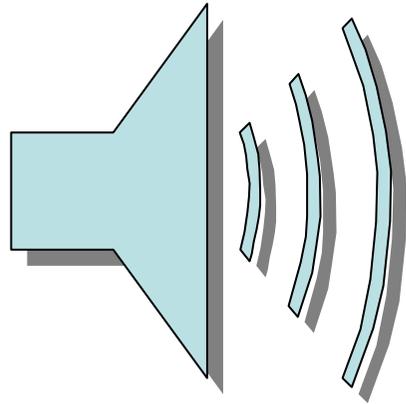
QI = 110

Quale significato ha la differenza di 6 punti di quoziente intellettivo?

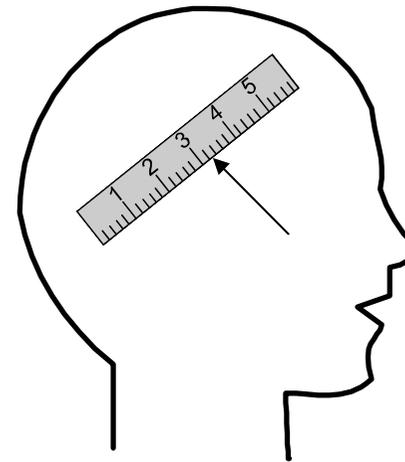
Per approfondimenti: http://www.ateneonline.it/chiorri/docenti2/approfondimenti/isbn6404-5_Approfondimento_1-1.pdf

Fechner (1860) e la psicofisica

Stimolo fisico di intensità b



Sensazione soggettiva γ



$$\gamma = k \ln b$$

Per approfondimenti: http://www.ateneonline.it/chiorri/docenti2/approfondimenti/isbn6404-5_Approfondimento_1-1.pdf

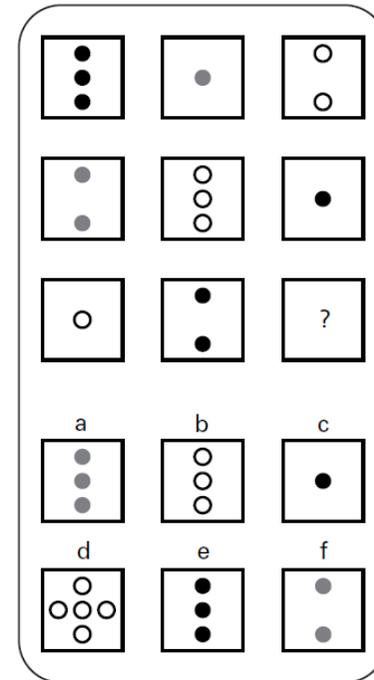
La misurazione in psicologia

- "The assignment of numerals to objects and events according to rules" (Stevens, 1946, p. 677)
- Le critiche (e.g., Michell, 1997)
- Le sensazioni hanno struttura additiva?
 - No (e.g., von Kries, 1882)
 - Sì (e.g. Fechner, 1887)

Per approfondimenti: http://www.ateneonline.it/chiorri/docenti2/approfondimenti/isbn6404-5_Approfondimento_1-1.pdf

Il modello di Spearman (1904)

- Regista : Film = Scrittore : ????
- 1 4 8 13 19 ???



$$x_i = \lambda g + s_i$$

La costruzione dei test psicologici: i primi passi e le tecniche di scaling

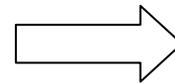
Il primo passo: cosa si vuole misurare?

- Definizione del costrutto
- E.g. Fobia Sociale (Mattick & Clarke, 1998)
 - **Fobia sociale:** ansia e paura che le persone provano di fronte alla prospettiva di essere osservati o guardati dagli altri, in particolare mentre stanno svolgendo alcune attività di routine come mangiare, bere, scrivere, firmare, usare le toilette pubbliche, lavorare, viaggiare su un mezzo pubblico di fronte agli altri o camminare in una stanza affollata. La preoccupazione principale della persona, in queste situazioni, è quella di apparire ansiosa, timida, con dei problemi, strana, e di cominciare a tremare, ad arrossire e/o a mostrare di sentirsi a disagio.
 - **Ansia da interazione sociale:** disagio nell'incontrare e parlare con gli altri, siano essi membri del sesso opposto, estranei o conoscenti. In particolare, la preoccupazione principale riguarda la paura di essere incapaci di esprimersi, di essere noiosi, di apparire stupidi, di non sapere cosa dire o come rispondere nelle interazioni sociali, e di essere ignorati.

Operazionalizzazione

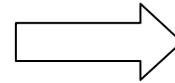


Intelligenza = capacità matematiche



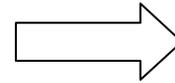
1 4 8 13 19 ???

Intelligenza = capacità logiche



Regista : Film = Scrittore : ????

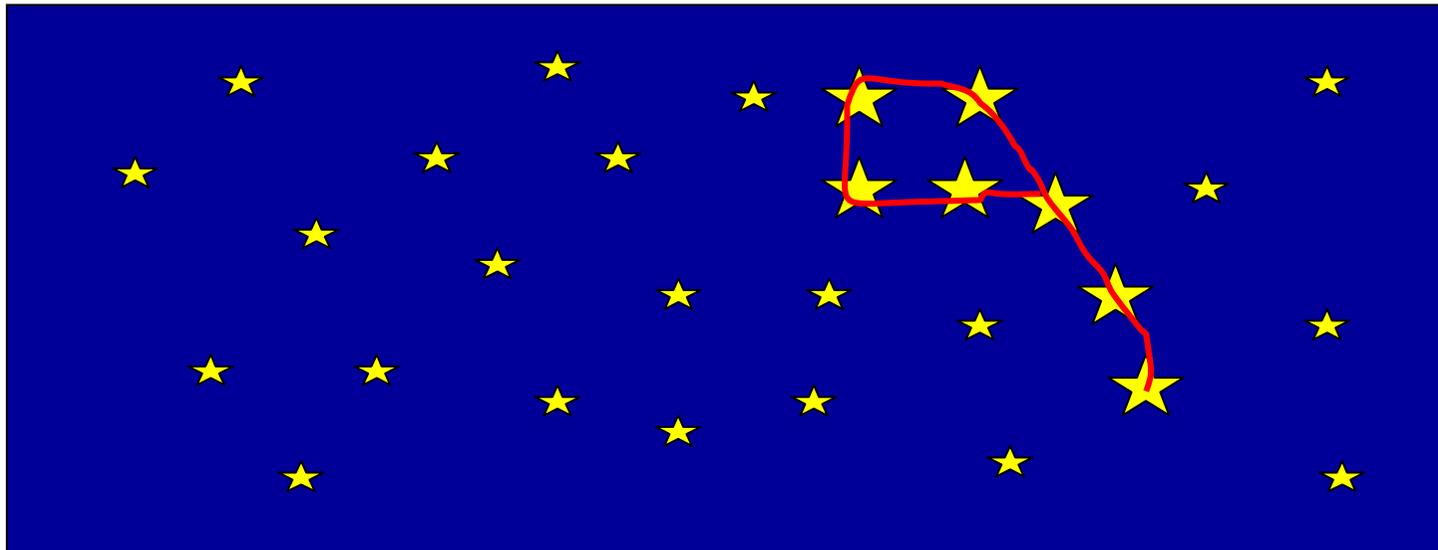
Intelligenza = capacità di adattamento



Quali operazionalizzazioni?

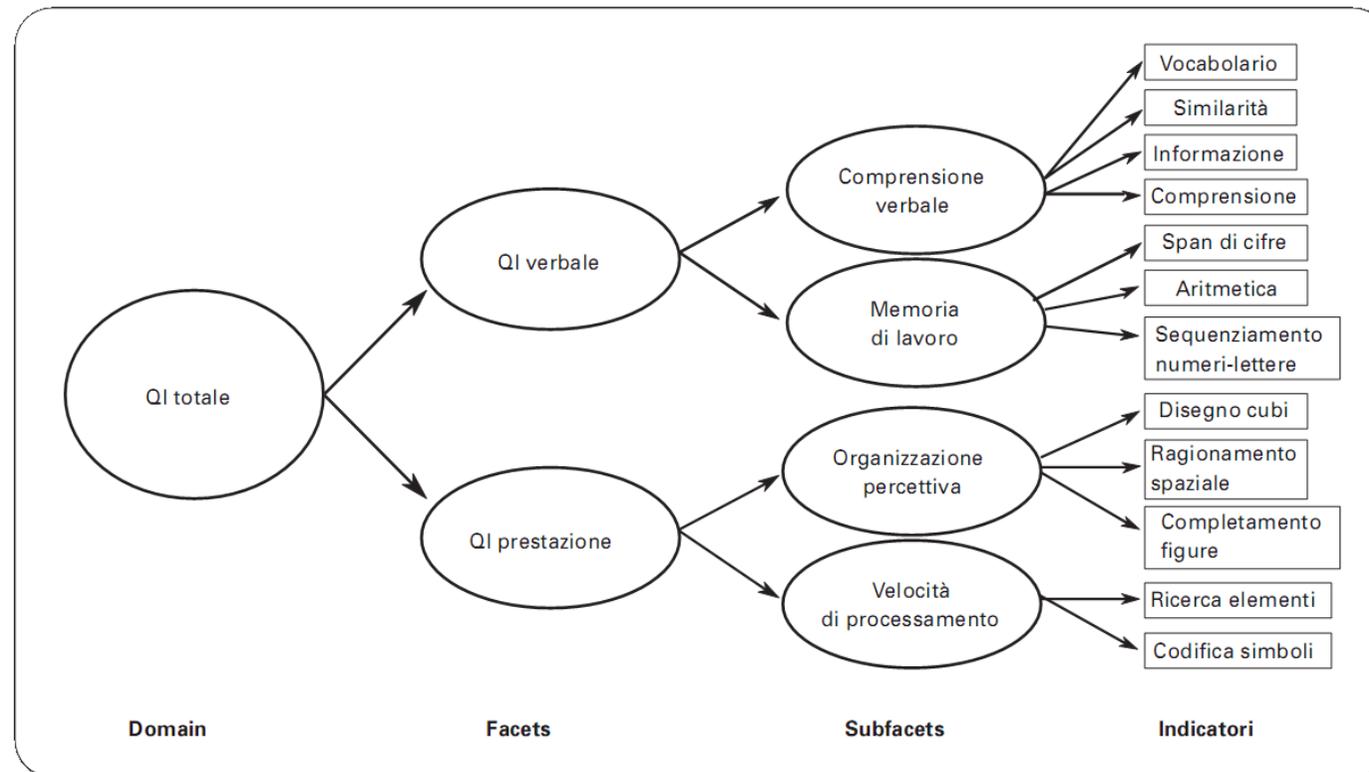
Dominio di contenuto

- Universo dei possibili comportamenti che, coerentemente con la definizione, possono rappresentare le operationalizzazioni del costrutto

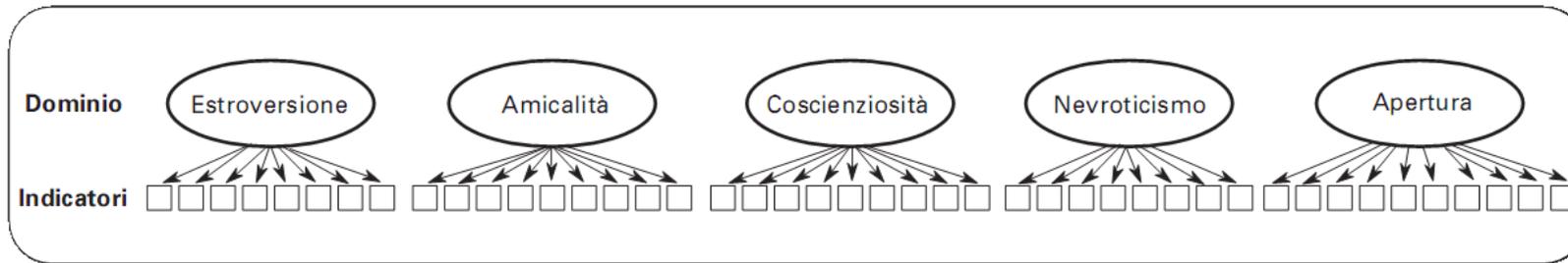


Definire il dominio di contenuto

- *Domain e facet*
- *Intelligenza e Five Factor Approach to Personality*

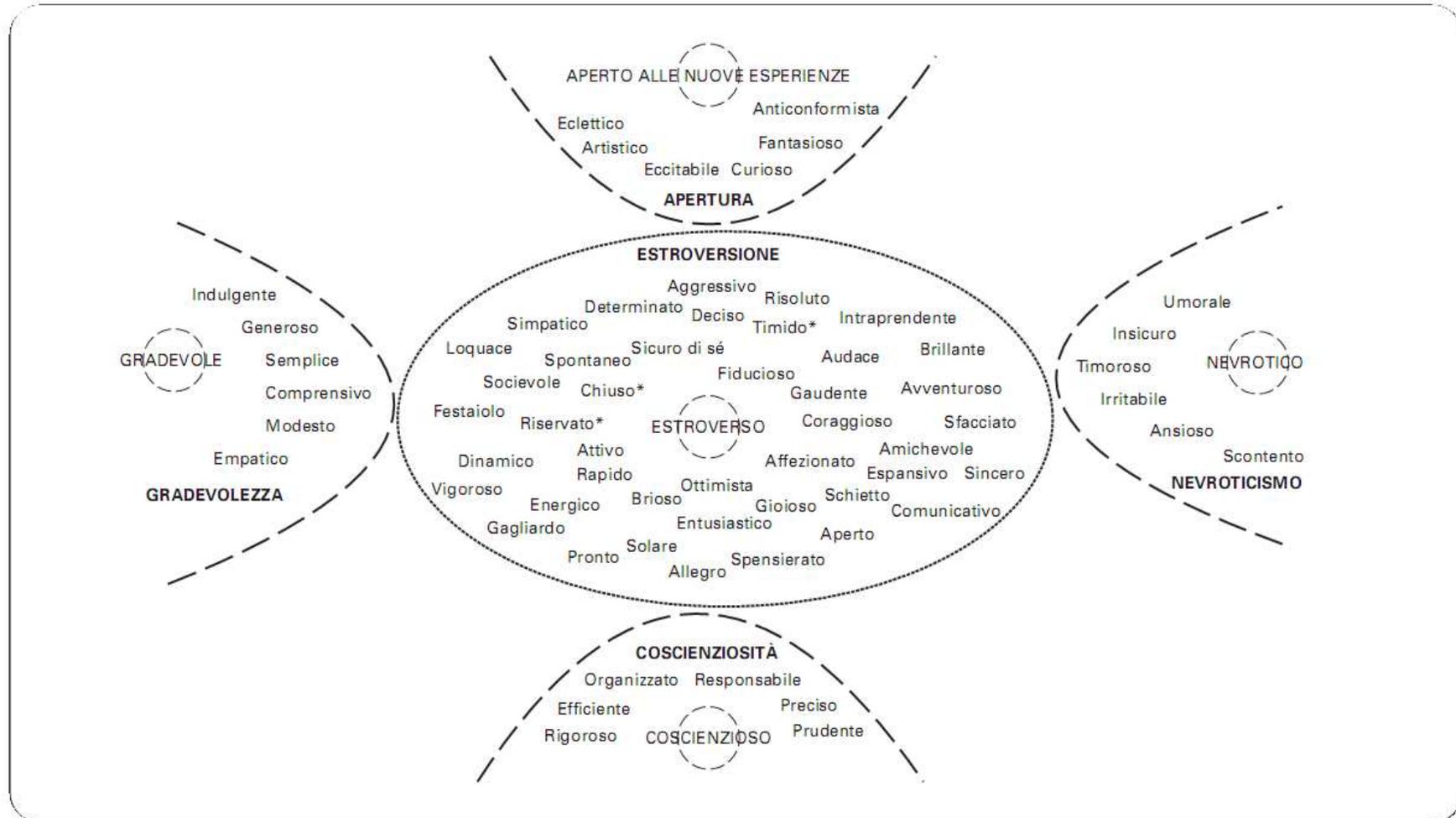


Caratteristiche di chi ha un punteggio basso	Scale dei tratti	Caratteristiche di chi ha un punteggio alto
Riservato, sobrio, non esuberante, distaccato, impegnato nel dovere, chiuso, taciturno	<p align="center">ESTROVERSIONE</p> Valuta la qualità e l'intensità dei rapporti interpersonali, il livello di attività, il bisogno di stimoli, la capacità di provare gioia	Socievole, attivo, loquace, interessato alle persone, ottimista, amante del divertimento, affettuoso
Cinico, rude, sospettoso, non collaborativo, vendicativo, crudele, irritabile, manipolatore	<p align="center">AMICALITA'</p> Valuta la qualità degli orientamenti interpersonali in una serie ininterrotta di pensieri, sentimenti e azioni che vanno dalla compassione all'antagonismo	Gentile, di animo buono, fiducioso, disponibile, indulgente, ingenuo, leale
Privo di scopi, inaffidabile, rigo, trascurato, molle, negligente, poca forza di volontà, edonista	<p align="center">COSCIENZIOSITA'</p> Valuta il grado di organizzazione degli individui, di perseveranza e impulso ad un comportamento che diritto allo scopo. Contrappone le persone sicure ed esigenti a quelle trasandate e indolenti	Organizzato, affidabile, lavoratore, autodisciplinato, puntuale, scrupoloso, ordinato, ambizioso, perseverante
Calmo, rilassato, non emotivo,duro, sicuro soddisfatto	<p align="center">NEVROTICISMO</p> Valuta l'adattamento in relazione all'instabilità emotiva. Identifica individui predisposti a stress psicologici, idee non realistiche, desideri o impulsi eccessivi e risposte di disadattamento	Preoccupato, nervoso, emotivo, insicuro, inadeguato, ipocondriaco
Conformista, con i piedi per terra interessi ristretti, non creativo, non analitico	<p align="center">APERTURA</p> Valuta la ricerca proattiva e apprezza l'esperienza spontanea, tolleranza e piacere di esplorare ciò che non è familiare	Curioso, di ampi interessi, creativo, originale, ricco di immaginazione, anticonformista

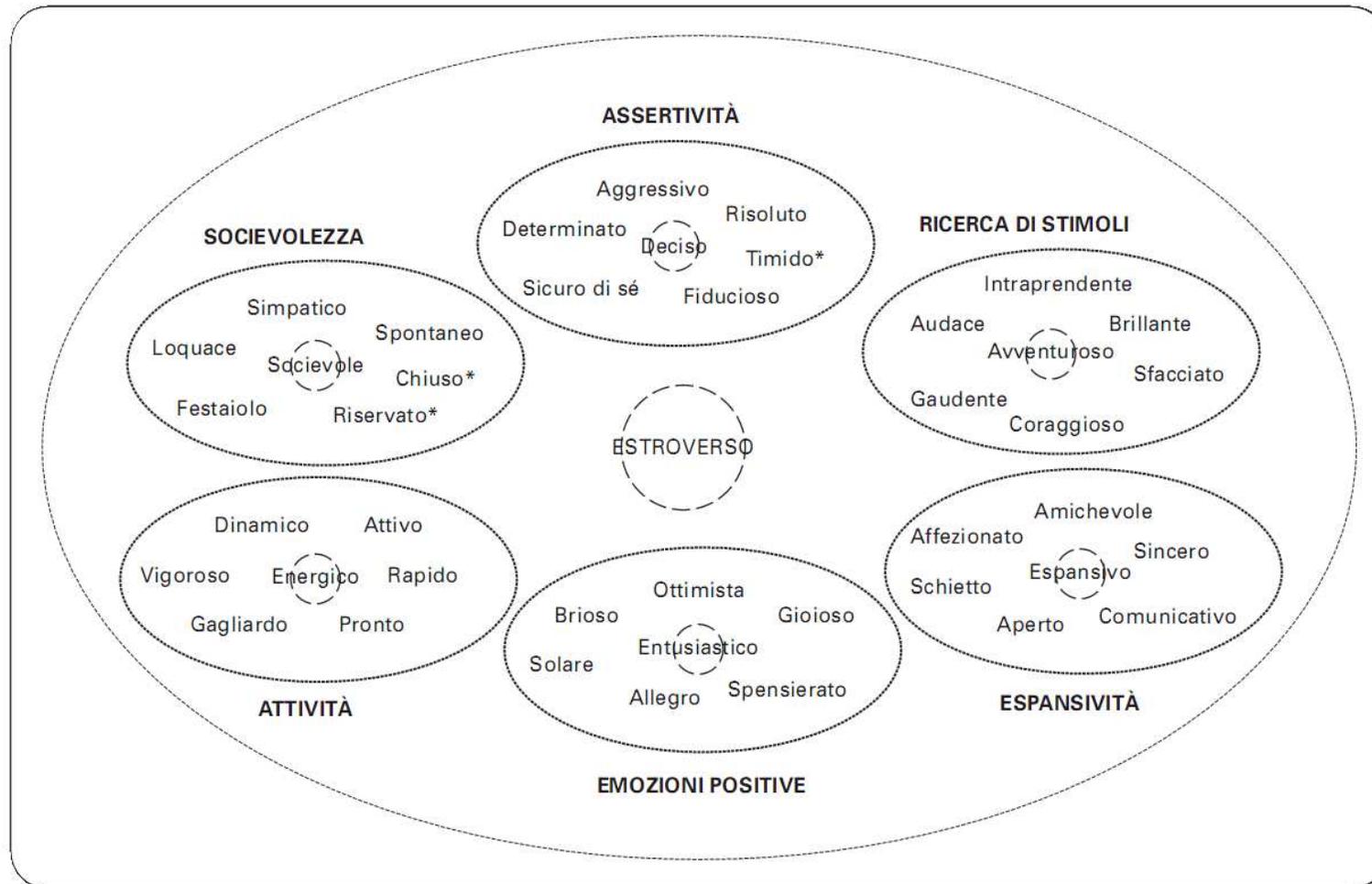


Dominio	Facets	Aggettivo prototipico
Estroversione	Socievolezza	Socievole
	Assertività	Deciso
	Attività	Energico
	Ricerca di stimoli	Avventuroso
	Emozioni positive	Entusiastico
	Espansività	Espansivo
Gradevolezza	Fiducia	Comprensivo
	Semplicità	Semplice
	Altruismo	Generoso
	Accondiscendenza	Indulgente
	Modestia	Modesto
	Solidarietà	Empatico
Coscienziosità	Competenza	Efficiente
	Ordine	Organizzato
	Responsabilità	Responsabile
	Impegno	Preciso
	Auto-disciplina	Rigoroso
	Prudenza	Prudente
Nevroticismo	Ansia	Ansioso
	Ostilità rabbiosa	Irritabile
	Depressione	Scontento
	Imbarazzo	Timoroso
	Lunaticità	Umorale
	Vulnerabilità	Insicuro
Apertura	Curiosità	Curioso
	Fantasia	Fantastico
	Estetica	Artistico
	Interessi	Eclettico
	Sensazioni	Eccitabile
	Anticonformismo	Anticonformista

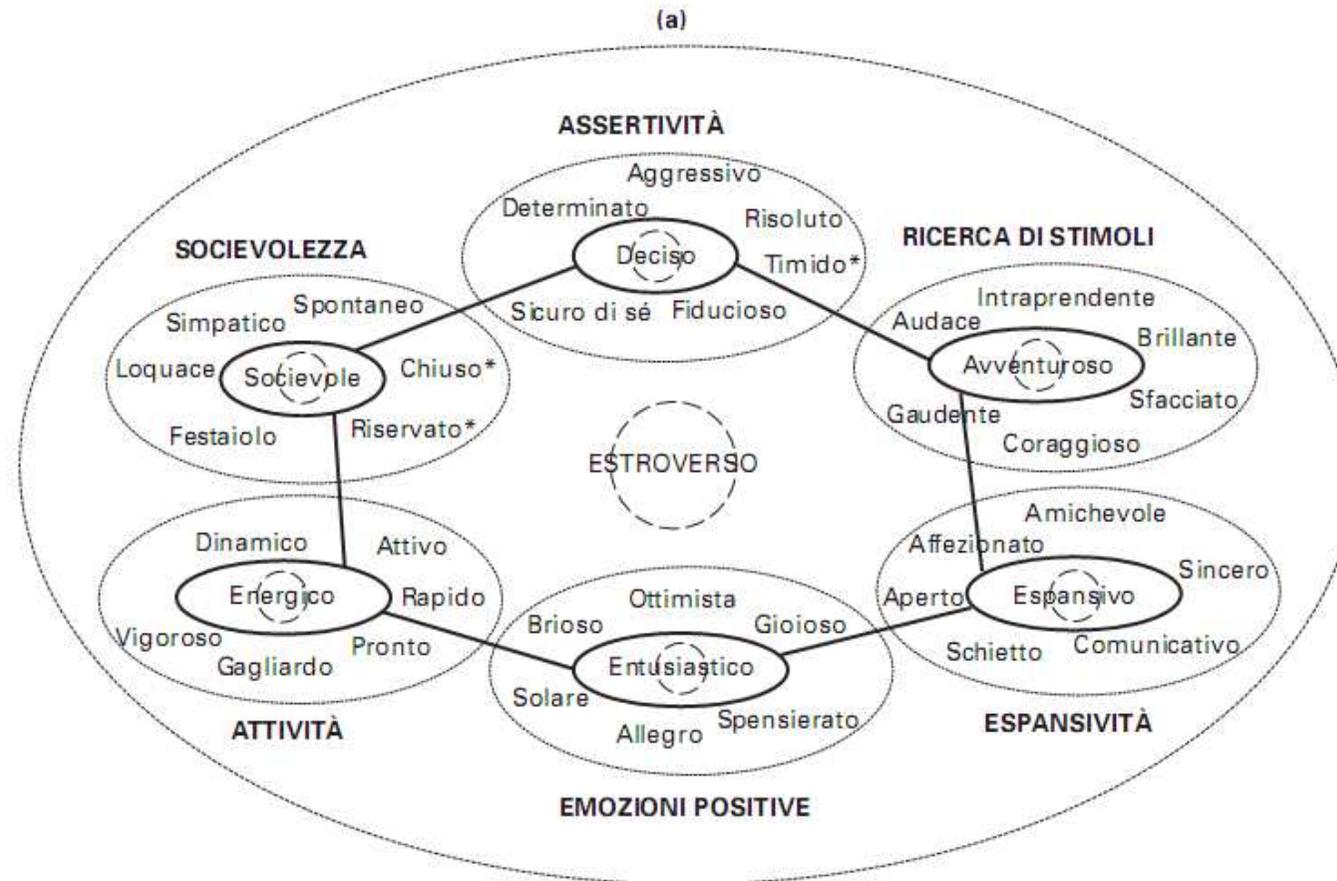
Domain



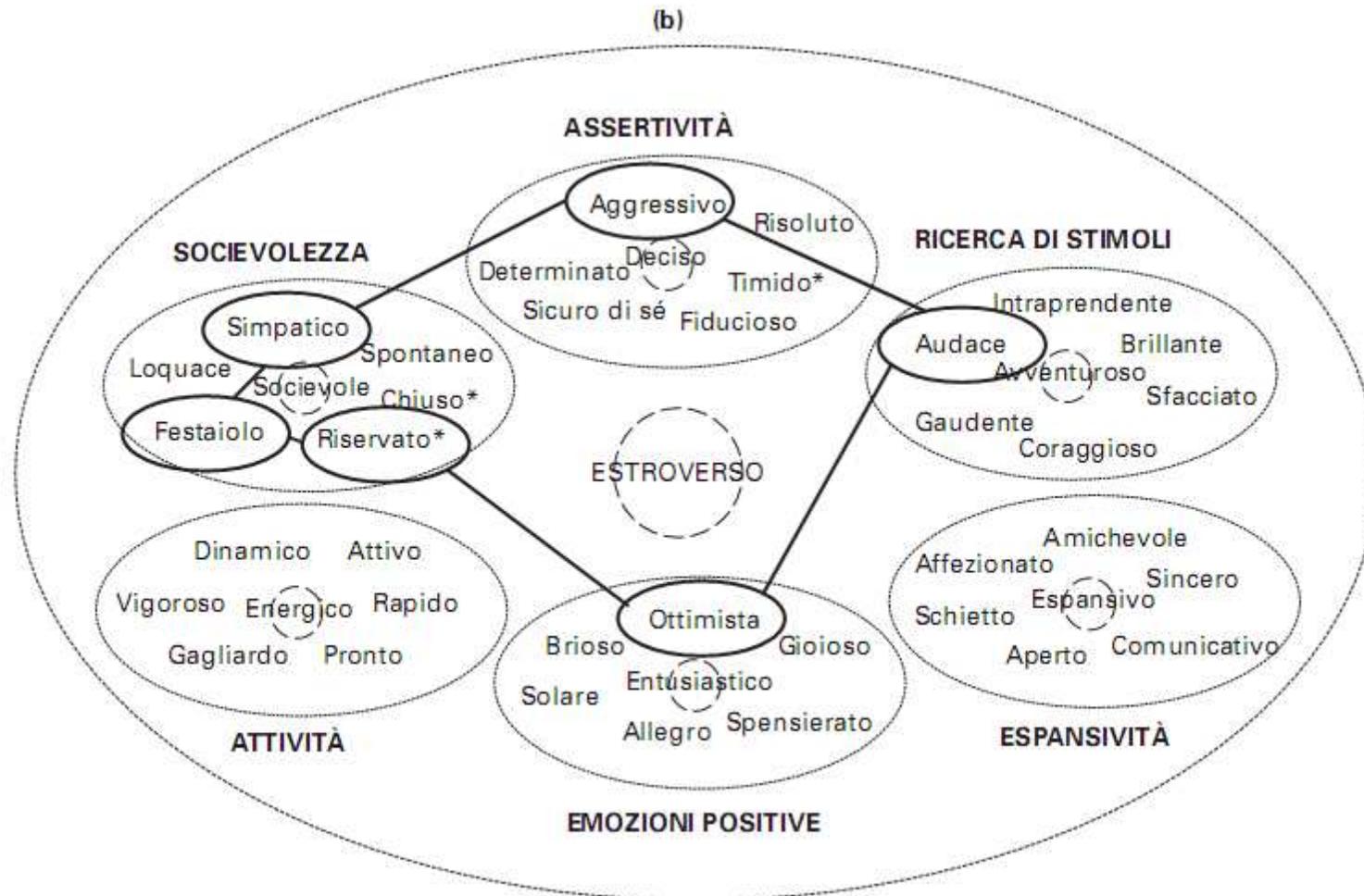
Facets



Campionamento del contenuto



Campionamento adeguato del contenuto



Campionamento impreciso del contenuto

(c)



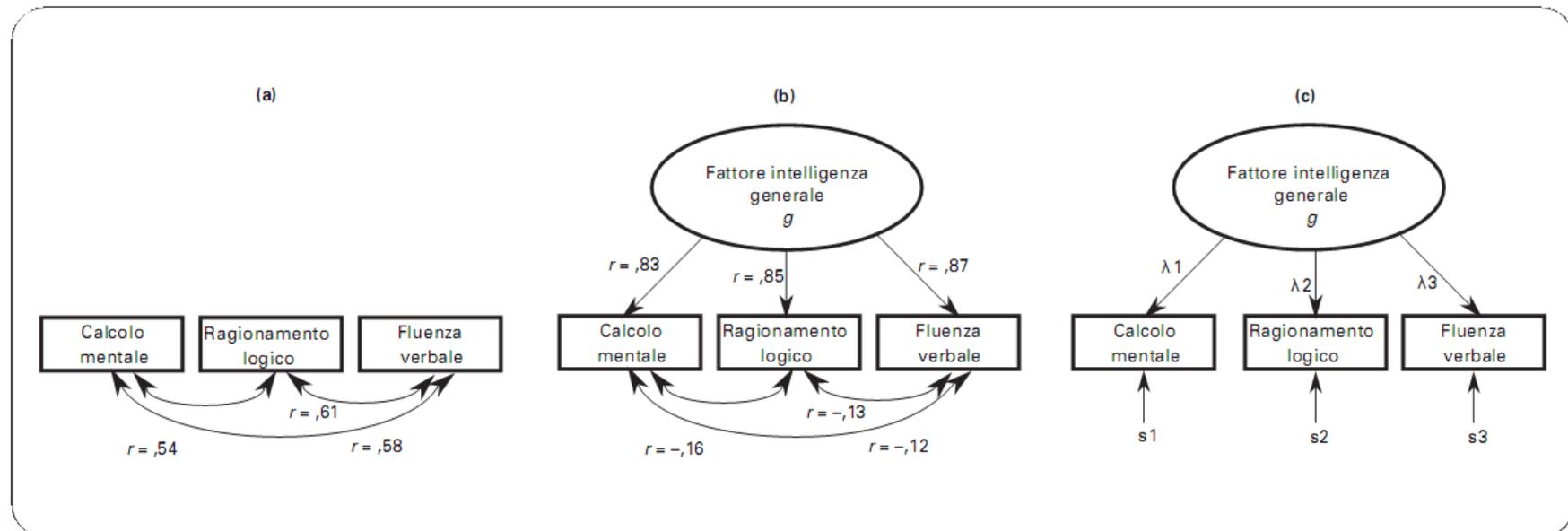
Campionamento inadeguato del contenuto

Scaling

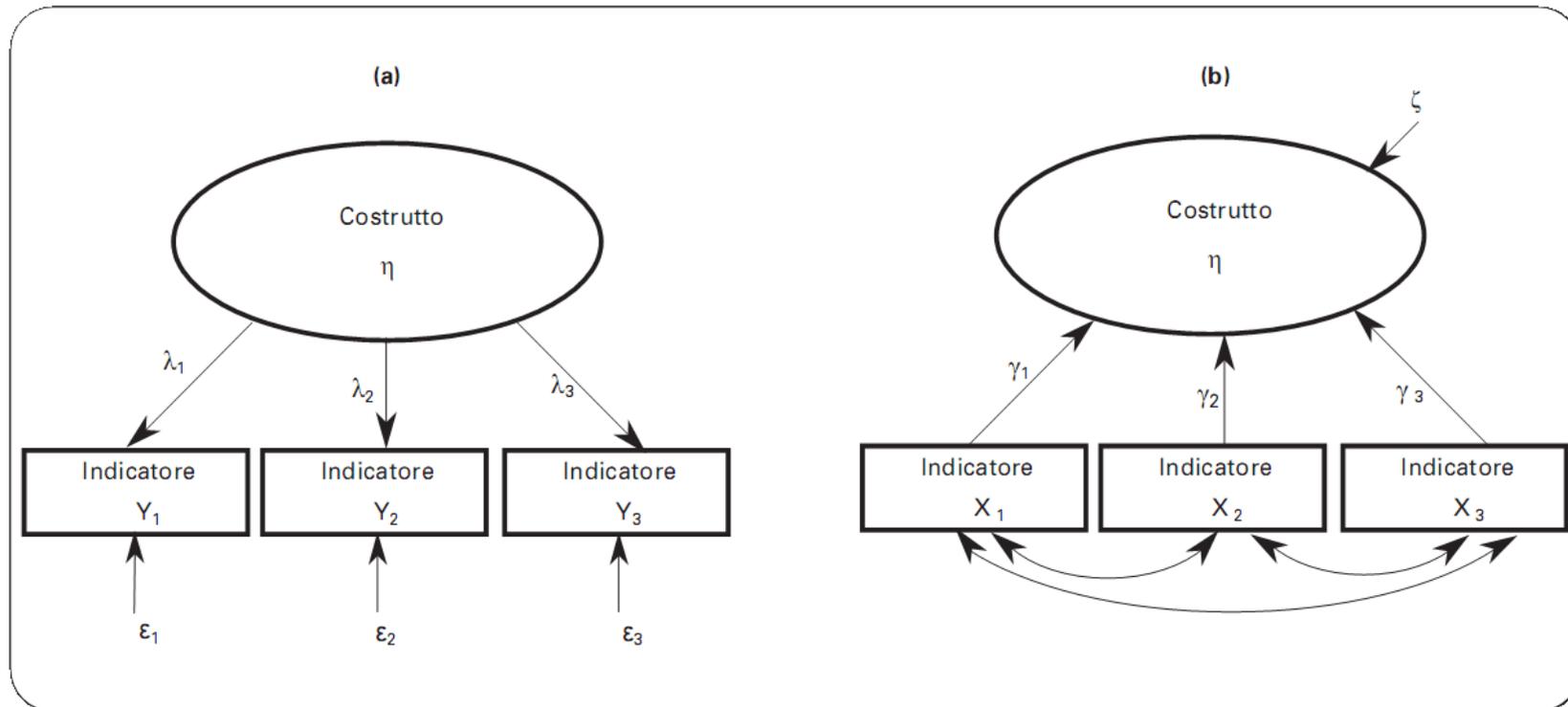
- Processo che permette di ottenere la misura quantitativa di una variabile psicologica, e quindi non osservabile direttamente (e.g., Giampaglia, 1990)
- Come trasformare le operazionalizzazioni in *item*?
- Come assegnare un punteggio alle risposte agli item?
- Come calcolare il punteggio nel test?

Scaling centrato sui soggetti

$$x_i = \lambda g + s_i \quad \Rightarrow \quad X = V + E$$



Indicatori riflessivi (a) e formativi (b)



$$y_i = \lambda_{i1} \eta_1 + \varepsilon_i$$

$$\eta_1 = \gamma_{11}x_1 + \gamma_{12}x_2 + \gamma_{13}x_3 + \zeta_1$$

Gli item di un test

Caratteristiche fondamentali

- Chiarezza
- Centralità rispetto alla definizione del costrutto e al contesto
- Non offensività

Item per i test di prestazione tipica

Item Sì/No e Vero/Falso

Mi piacciono i fiori V F

- Formulazione facile
- Risposta (in teoria) facile

Mi piacciono i fiori V F Non so

Mi piacciono i fiori V ? F

- Alternativa "non so" favorisce risposte difensive
- Problema dello scoring della risposta (0? 0,5?, etc.)

Cosa significa "non so"?

- Impossibilità di rispondere
- Non applicabilità al caso specifico
- Scarsa cooperatività o motivazione alla compilazione
- Difficoltà a comprendere la domanda
- Volontà di non esporsi o di rimanere sulla difensiva

Scelta multipla forzata

Sul lavoro sei una persona che (indica la risposta che ti caratterizza meglio):

- a) ama i periodi in cui ci si deve impegnare al massimo per rispettare una scadenza
- b) detesta lavorare sotto pressione;
- c) cerca di pianificare le cose da fare in anticipo così da non ritrovarsi in difficoltà successivamente

- Non sono né meglio né peggio di altre persone
- Penso di essere una persona speciale

Tom si sentiva ansioso, e divenne un po' stressato quando cominciò a pensare a tutto il lavoro che avrebbe dovuto fare. Quando il suo capo gli consegnò un'altra pratica da sbrigare, si sentì: (selezionare l'alternativa migliore):

- a. Sopraffatto b. Depresso c. Pieno di vergogna d. Imbarazzato e. Agitato

Problemi con scelta multipla forzata

- Le alternative giacciono sullo stesso continuum?
- Le alternative rappresentano gradi diversi di presenza del costrutto?
- Quale punteggio ottiene ogni alternativa?
- Auto-referenzialità → scale ipsative

Scale di valutazione (rating scales)

- Scale Likert \neq scale di risposta di tipo Likert
- Summated rating scales

Sono una persona organizzata

Per niente d'accordo

Completamente d'accordo

(a) Barrare valore numerico	Sono una persona organizzata	Per niente	1	2	3	4	5	Completamente
(b) Barrare casella	Sono una persona organizzata	Per niente	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	Completamente
(c) Scrivere direttamente il numero	Sono una persona organizzata	<u>4</u>						

Tipi di valutazioni

<i>Dominio</i>	<i>Esempio</i>
Frequenza	Con quale frequenza soffre di emicrania?
Intensità	Quanto influenza il suo umore il suo aspetto fisico?
Valutazione	Gli altri si approfittano di me

Ancoraggi

- Cosa significano i punteggi numerici?

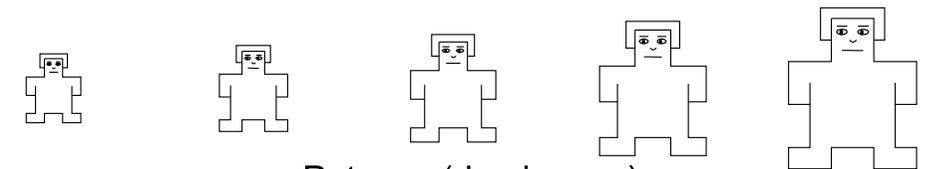
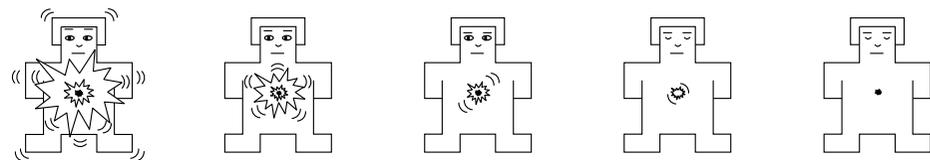
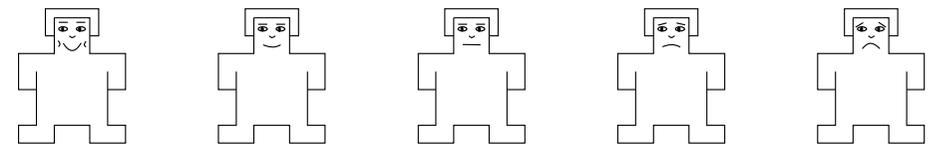
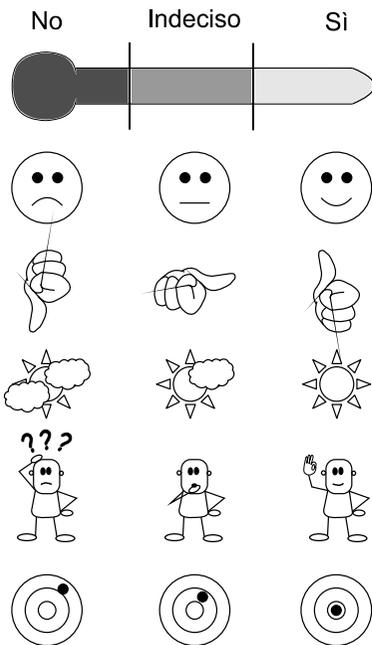
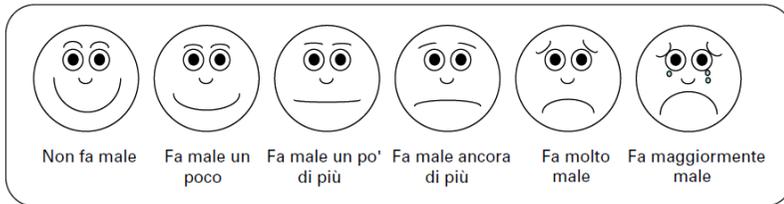
UMORE DEPRESSO (sentimento di tristezza, mancanza di speranza, sentimento di incapacità e di inutilità)

1. Assente
2. Manifesta questi sentimenti solo se interrogato
3. Esprime spontaneamente questi sentimenti
4. Comunica questi sentimenti con messaggi non verbali, cioè attraverso l'espressione del volto, la posizione del corpo, la voce e la tendenza al pianto
5. Il paziente manifesta quasi esclusivamente questi sentimenti mediante messaggi sia verbali sia non verbali

Quantificatori indeterminati

<i>Valore</i>	<i>Frequenza</i>	<i>Intensità</i>	<i>Valutazione</i>
1	Mai	Per niente	Per niente d'accordo
2	Raramente	Poco	Poco d'accordo
3	Talvolta	Mediamente	Né in accordo né in disaccordo
4	Spesso	Abbastanza	Abbastanza d'accordo
5	Sempre	Molto	Completamente d'accordo

Visual analogue scales





0: Hi. I am not experiencing any pain at all. I don't know why I'm even here.

1: I am completely unsure whether I am experiencing pain or itching or maybe I just have a bad taste in my mouth.

2: I probably just need a Band Aid.

3: This is distressing. I don't want this to be happening to me at all.

4: My pain is not fucking around.

5: *Why is this happening to me??*

6: *Ow*. Okay, my pain is *super* legit now.

7: I see Jesus coming for me and I'm scared.

8: I am experiencing a disturbing amount of pain. I might actually be dying. Please help.

9: I am almost definitely dying.

10: I am actively being mauled by a bear.

11: Blood is going to explode out of my face at any moment.

Too Serious For Numbers: You probably have ebola. It appears that you may also be suffering from Stigmata and/or pinkeye.

Quanti punti?

- Da 3 a 11
- Problema degli ancoraggi con molti punti
- Pari vs dispari

Scale di frequenza

Comportamento	Mai	Raramente	Qualche volta	Spesso	Sempre
(a)					
Scaricare la posta elettronica					
Guardare la TV					
Mandare un sms a un amico					
Andare al cinema					
(b)					
Fare un viaggio all'estero					
Comprare un computer					
Partecipare a una riunione di famiglia					
Andare al cinema					

Effetti distortenti

Oggetto della misurazione	Effetti distortenti
Personalità, atteggiamenti, opinioni ecc.	<p>Acquiescenza: essere sistematicamente d'accordo con le affermazioni</p> <p>Desiderabilità sociale: rispondere in maniera da mostrarsi nel modo più socialmente accettabile</p> <p>Estremismo: scegliere sempre le alternative più estreme</p> <p>Evasività o cautela: scegliere sempre i punti centrali della scala, o le alternative "non so"</p>
Esperienze attuali	<p>Inaccuratezza: rispondere a caso o in modo incoerente</p> <p>Devianza: scegliere alternative di risposta insolite</p> <p>Produttività: fornire più di una risposta, o produrre risposte prolisse se la domanda è aperta</p>
Comportamento presente o passato	<p>Effetto alone: farsi influenzare dalla valutazione generale su di sé nella valutazione di aspetti specifici</p> <p>Indulgenza/criticismo: valutarsi sistematicamente sopra o sotto la media</p>

Fattori di distorsione

- Lunghezza del test e numero di punti della scala di risposta
- Effetto attrattore del punto centrale della scala
 - "Non so", neutralità o atteggiamento difensivo?
 - Quando è necessario?

Item straight e reverse

Sono una persona affidabile

Polar opposite

Sono una persona inaffidabile

Negated regular

Non sono una persona affidabile

Negated polar opposite

Non sono una persona inaffidabile

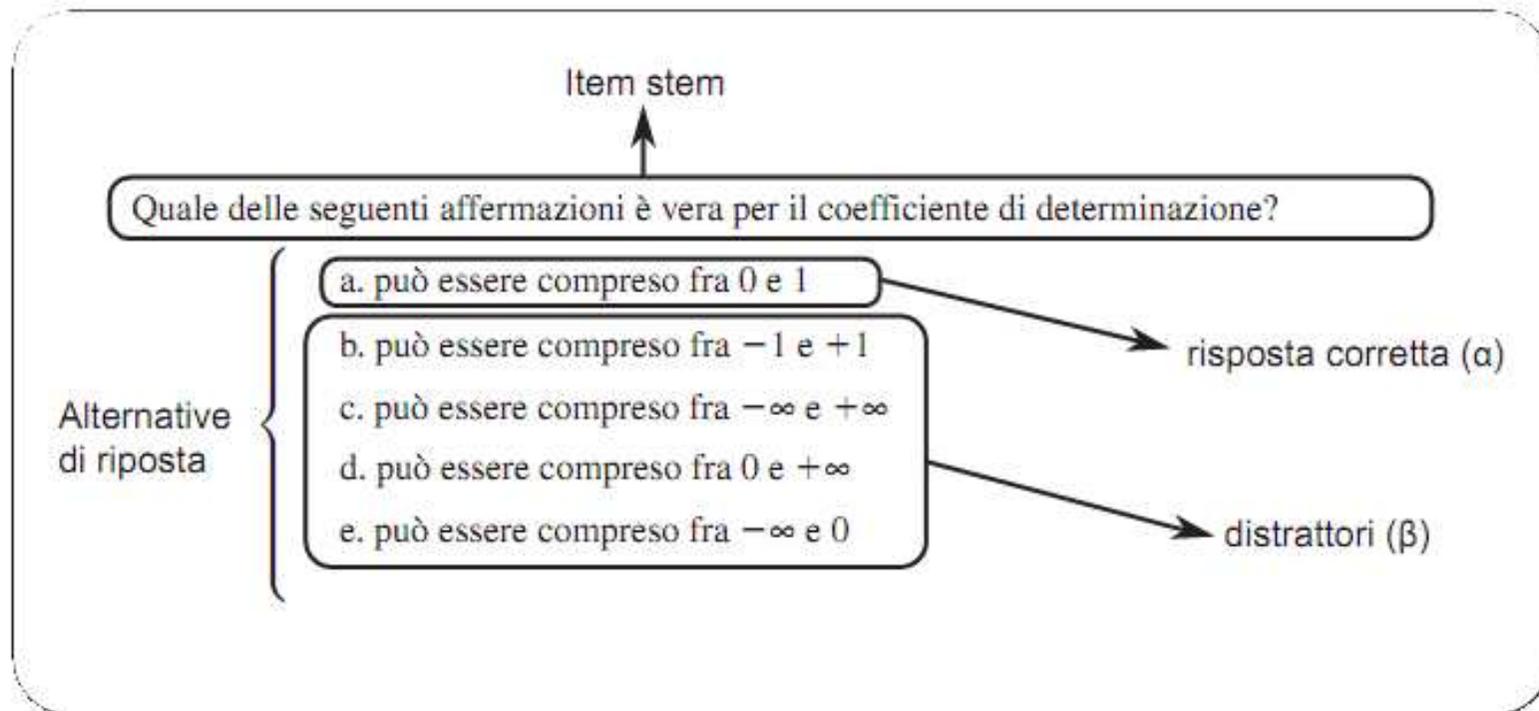
- Speed bump cognitivo
- Polar opposite scelta migliore ma l'opposto non esiste sempre
- Straight e reverse giacciono sullo stesso continuum?
- *Proximity effect*: straight devono stare lontani, reverse vicini (Wijters et al., 2009)

Item per i test di prestazione massima

Test di prestazione massima

- Domande aperte o chiuse?
- Rievocazione
 - Domande aperte → libera
 - Domande chiuse → guidata
- Quando è meglio aperta? → studi esplorativi
 - Partial credit scoring
 - Inter-rater agreement
- Metodo clinico
 - 8 10 12
 - 18 20 22
 - 100 102 104

Struttura dell'item di prestazione massima



Test di prestazione massima

- Item per i test di efficienza mentale (test di intelligenza, abilità, attitudine, informazione)
- Item per i test di profitto (prove di esame, test educativi standardizzati, etc.)
- Item di comprensione del testo

Test di intelligenza, abilità e attitudine

- Intelligenza fluida vs cristallizzata
- *Surgency*
- Abilità
- Attitudini

Analogie

SIAMESE: GATTO = ALANO: ???

a. Pesce b. Uccello c. Cane d. Rettile

- **Percezione empirica** → percezione e comprensione di ogni elemento dell'analogia
- **Deduzione di relazioni** → inferenza della relazione fra i primi due termini
- **Deduzione dei termini di correlazione** → capacità di applicare il rapporto inferito a un ambito diverso
- Intelligenza o cultura generale? La difficoltà deve stare nella relazione fra i termini, non nella conoscenza dei contenuti

TELEVISORE: MICROSCOPIO = RADIO: ???

a. Microprocessore b. Satellite c. Lente di ingrandimento d. Amplificatore

Numeri

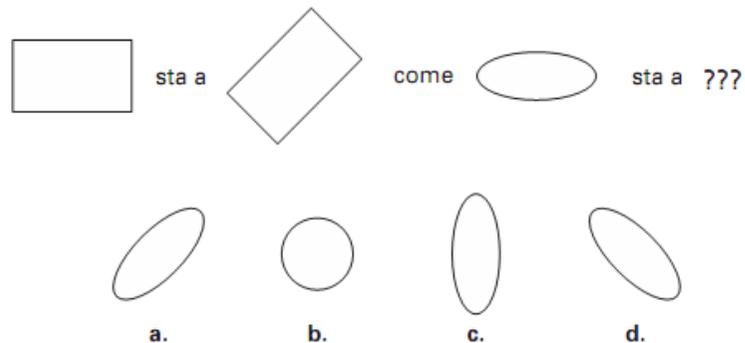
$13: 4 = 23: ???$
a. 11 b. 5 c. 12 d. 8

Lettere

$B: V = E: ???$
a. M b. P c. G d. S

- Relazione grandezza: Biglia : Pallone = Dado: ?????
- Relazioni fra parole: Bello : Brutto = Alto : ?????

Figure
(Fluida?)



Parola comune

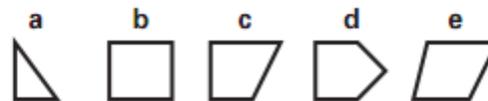
Casello, Linea, Scartamento, Biglietto, Tratta
a. Autostrada b. Ferrovia c. Teleferica d. Aviolinea e. Funivia

Item ad esclusione

- Individua l'intruso

Paola, Roberta, Marisa, Serena, Alice
a. Paola b. Roberta c. Marisa d. Serena e. Alice

16, 48, 32, 24, 20
a. 16 b. 48 c. 32 d. 24 e. 20



Item a sequenza

Numerico e alfabetico
(Alfabeto italiano o inglese?)

1, 4, 10, 19, 31, ???	A, B, D, G, M, ???
a) 40	a) T
b) 43	b) R
c) 46	c) S
d) 41	d) Q

Visivo
(varianti: carte da gioco, domino, etc.)

				???
a	b	c	d	

Matrice di Raven

			1	2	3	4
		???	5	6	7	8

Z	T	Q	N
A A D D G G L L			
Z	T	Q	N

I	M	I	M	O
O O O O N N N N I I				
I	M	I	M	O

a b c d e

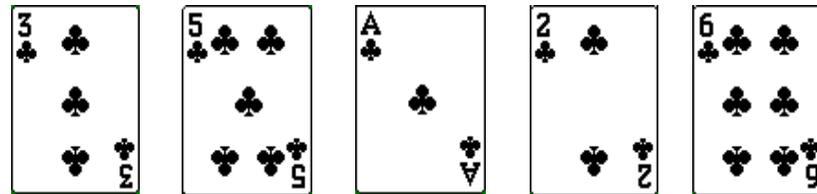
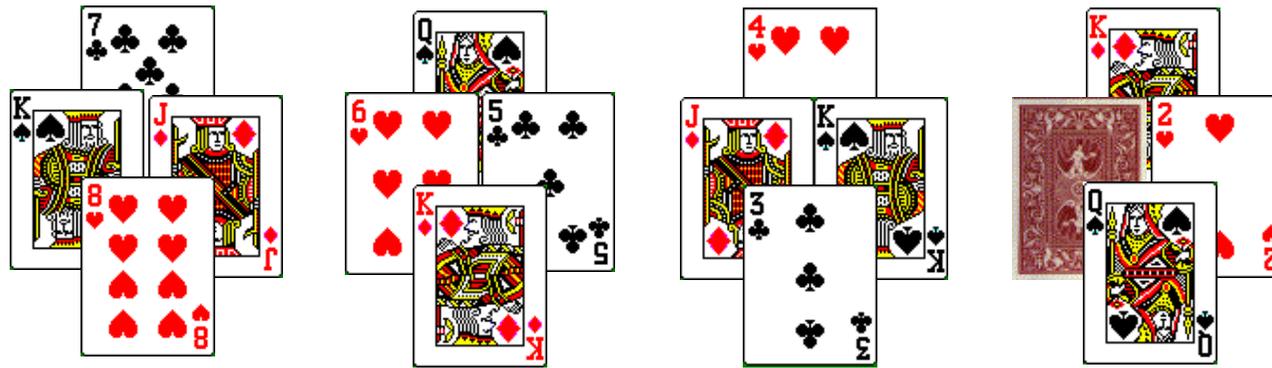
↑

●●●●	●●●		●●●●	?
●●	●		●●●●	?

	●●	●	●●	●
●●	●●	●●●		●

a b c d e

↑



a

b

c

d

e



Informazione (cultura) generale

- Argomenti vari

Quale autore ha scritto la poesia "Ed è subito sera"?	Chi fondò nel 1919 il Partito Popolare Italiano?	Qual è la capitale dell'Uruguay?
a) Carlo Emilio Gadda	a) Antonio Gramsci	a) La Paz
b) Giuseppe Ungaretti	b) Luigi Einaudi	b) Quito
c) Umberto Saba	→ c) Luigi Sturzo	c) Asunción
→ d) Salvatore Quasimodo	d) Paolo Boselli	→ d) Montevideo
e) Eugenio Montale	e) Pietro Badoglio	e) Caracas

Abilità verbale

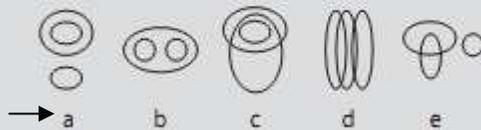
Il termine "idiosincrasia" indica:	In quale riga tutte le parole sono scritte con la corretta ortografia?	Quale termine tra i seguenti è sinonimo di "appropinquare"?
a) una passione sfrenata	a) esterefatto scientifico cielo	→ a) avvicinare
b) una malattia mentale	b) roccie camicie coscienza	b) distanziare
c) un culto pagano	→ c) familiare grigie geniale	c) rifornire
d) un intoppo burocratico	d) scialaquare ciliege ganasce	d) spiegare in modo adeguato
→ e) un'avversione per qualcosa	e) contrazione sciare dolce	e) rubare

- Scrivere parole che iniziano con "a" (tempo max 20 secondi)
- Selezionate fra loro le due parole più affini da un insieme di molte altre
- Trovare la parola che completa una frase
- Rimettere nell'ordine corretto le parole fornite in modo da formare una frase di senso compiuto
- Spiegare il significato di un proverbio

Item verbali di ragionamento

Indicare la rappresentazione che soddisfa la relazione insiemistica esistente fra i seguenti termini:

Ingegneria, Meccanica, Parapsicologia



Se:

Alcuni β sono γ .
Nessun α è β .

Quale fra le seguenti conclusioni è valida?

- a) Alcuni α non sono γ
- b) Alcuni γ non sono α
- c) Tutti i γ sono α
- d) Nessun γ è α
- e) Nessun α è γ

Per vincere alla lotteria, possedere almeno un biglietto è una condizione:

- a) Necessaria.
- b) Sufficient
- c) Necessaria e sufficiente
- d) Più necessaria che sufficiente
- e) Né necessaria né sufficiente

Test di profitto (achievement tests)

- Vantaggi vs svantaggi uso item scelta multipla
- PISA e TIMMS
- Invalsi

INVALSI grammatica

1. Nelle frasi seguenti il pronome “ne” sostituisce una o più parole. Per ogni frase indica quali parole sostituisce. Osserva bene l’esempio.

Nella frase:	“ne” sostituisce ...
Esempio: <i>Gli si avvicinò un venditore di fiori freschi e gliene offrì un mazzo.</i>	<i>di fiori freschi</i>
a. Il clima caldo e umido non fa bene alla salute, anzi ne derivano molte malattie.	
b. Ho avuto delle difficoltà ma preferisco non parlar ne .	
c. Ho scritto la relazione di storia: te ne ho portato una copia.	
d. Credo che tu abbia sbagliato ma non ne sono sicuro.	
e. Assumiti la responsabilità delle tue parole e di tutto quello che ne consegue.	

Risposta corretta

Nella frase:	“ne” sostituisce ...
Esempio: <i>Gli si avvicinò un venditore di fiori freschi e gliene offrì un mazzo</i>	<i>di fiori freschi</i>
a. Il clima caldo e umido non fa bene alla salute, anzi ne derivano molte malattie	(dal clima caldo e umido)
b. Ho avuto delle difficoltà ma preferisco non parlar ne	(delle difficoltà)
c. Ho scritto un articolo per un giornale tedesco: te ne ho portato una copia	(del giornale tedesco)
d. Credo che tu abbia sbagliato ma non ne sono sicuro	(del fatto che tu abbia sbagliato)
e. Assumiti la responsabilità delle tue parole e di tutto quello che ne consegue	(dalle tue parole)

Ambito grammaticale rilevato: 5

INVALSI matematica

1 Esempio per la II secondaria di II grado

In una città il costo di un biglietto dell'autobus è passato da 1 euro a 1,20 euro, se acquistato nelle biglietterie a terra, e 1,50 se acquistato a bordo. Qual è, in percentuale, il sovrapprezzo per l'acquisto a bordo rispetto all'acquisto in biglietteria?

- A. 20%
- B. 25%
- C. 30%
- D. 50%

Risposta corretta: B

I distrattori possono intercettare errori frequenti fra gli studenti:

- A. 20% (è quanto costa in meno il biglietto a terra rispetto a quello a bordo)
- C. 30% (è l'aumento in centesimi di euro, non in percentuale)
- D. 50% (è l'aumento rispetto al prezzo precedente e corrisponde a una lettura superficiale del testo)

Tipologia: Scelta multipla.

Ambito prevalente: Relazioni e funzioni.

Processo prevalente: Saper risolvere problemi utilizzando gli strumenti della matematica.

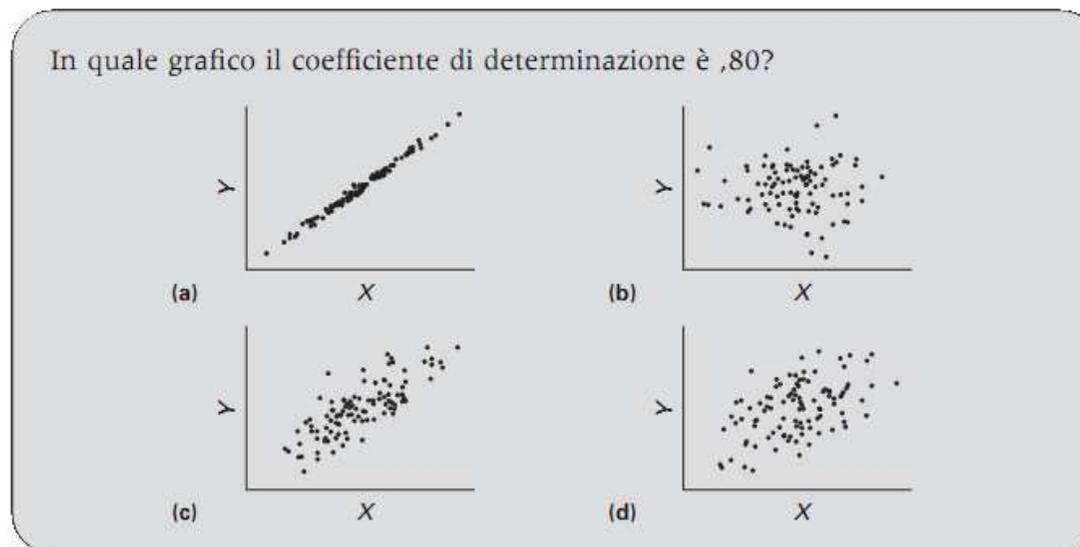
Nuovo Obbligo di Istruzione

- **Competenza** – Utilizzare le tecniche e le procedure del calcolo aritmetico ed algebrico, rappresentandole anche sotto forma grafica.
- **Abilità** – impostare uguaglianze di rapporti per risolvere problemi di proporzionalità e percentuale.

Item scelta multipla

Il coefficiente di determinazione:

- a. può essere compreso fra 0 e 1
- b. può essere compreso fra -1 e +1
- c. può essere compreso fra $-\infty$ e $+\infty$
- d. può essere compreso fra 0 e $+\infty$
- e. può essere compreso fra 0 e $-\infty$



Comprensione del testo

Senza l'evento o senza quella combinazione di eventi singolari e catastrofici di circa 65 milioni di anni fa, in nessun modo inevitabili e necessari, le regole di coesistenza fra mammiferi e dinosauri sarebbero potute restare inalterate ancora per un tempo indefinito. Certo, catastrofi analoghe sarebbero potute avvenire successivamente. Ma non vi è alcuna ragione stringente per cui una catastrofe sarebbe dovuta necessariamente avvenire nei milioni di anni che precedono il nostro tempo.⁶

Nel brano si afferma che gli eventi catastrofici avvenuti circa 65 milioni di anni fa sulla Terra:

- a) avrebbero anche potuto non accadere
- b) dovevano necessariamente accadere per una inevitabile e necessaria combinazione di fattori
- c) dovevano necessariamente accadere, e proprio in quel momento, per ragioni superiori alla comprensione umana
- d) sono in realtà accaduti più volte anche se la memoria umana non può ricordarlo
- e) si ripeteranno entro altri 65 milioni di anni circa

Valutazione degli item di un test

Valutazione preliminare

- Ruolo marginale della statistica
- Validità di contenuto e di facciata
- Analisi degli item e attendibilità
- Raffinamento del test

Validità di contenuto

Content validity is the degree to which elements of an assessment instrument are relevant to and representative of the targeted construct for a particular assessment purpose

(Haynes et al., 1995, p. 238)

- Rilevanza (*relevance*): appropriatezza degli elementi del test elementi per la valutazione del costrutto
- Rappresentatività (*representativeness*): grado in cui gli elementi del test sono rappresentativi delle varie sfaccettature del costrutto target e riescono a coprire in modo adeguato il dominio di contenuto

Esempio di metodo per la valutazione della validità di contenuto - esperti*

- Chiedere ad esperti (*Subject Matter Experts*) di valutare quanto è rilevante ogni item

<i>Item del test</i>	Quanto è rilevante l'item?		
	Per niente	Poco	Molto
<i>Item del test</i>	Per niente	Poco	Molto
<i>Item del test</i>	Per niente	Poco	Molto
<i>Item del test</i>	Per niente	Poco	Molto
[...]	[...]	[...]	[...]

- Regole di decisione per il mantenimento dell'item nell'item pool

* Per maggiori dettagli vedi Approfondimento 4.1

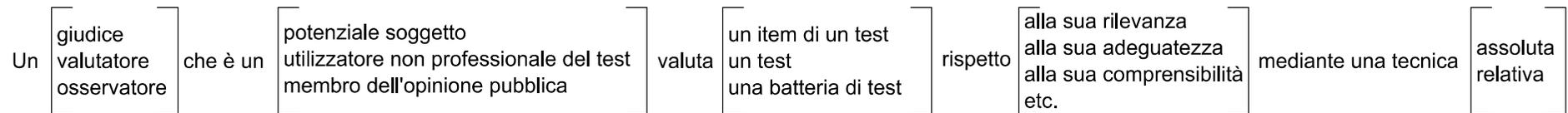
Validità di facciata

Content validity should not be confused with face validity. The latter is not validity in the technical sense; it refers not to what the test actually measures, but **to what it appears superficially to measure. Face validity pertains to whether the test 'looks valid'** to the examinees who take it, the administrative personnel who decide on its use, and other technically untrained observers

(Anastasi, 1988, p. 144)

- **Grado in cui gli item di un test *sembrano* misurare il costrutto che intendono misurare**
- Chi lo valuta?

Validità di facciata



- Deve essere eseguita per tutte le parti del test
- **Tecniche**
 - **riformulazione:** parafrasi
 - **thinking aloud:** pensare ad alta voce
 - **intervista cognitiva:** analisi dei processi di pensiero che portano alla risposta
 - **intervista in profondità:** compromesso fra il colloquio clinico e l'intervista direttiva → utile per analisi più approfondite

Validità di criterio

- Concorrente
- Postdittiva
- Predittiva

- Prima verifica empirica di validità
- Validità empirica
- Problemi nel trovare un buon criterio

Validità di costrutto

- Convergente: stesso costrutto, diverso metodo
- Discriminante: costrutto diverso, stesso metodo
- similar enough to other [concepts] to be recognizable, but different enough to be worth studying” (Mayer, 2000, p. 49)

Multi-Trait Multi-Method Matrix (Campbell & Fiske, 1959)

		Self-report			Osservatore			Esperto		
		DEP	ANX	VN	DEP	ANX	VN	DEP	ANX	VN
Self-report	DEP	(,90)								
	ANX	,50	(,91)							
	VN	,39	,40	(,92)						
Osservatore	DEP	,58	,24	,10	(,89)					
	ANX	,21	,59	,11	,61	(,88)				
	VN	,12	,13	,51	,60	,62	(,87)			
Esperto	DEP	,54	,23	,17	,68	,41	,37	(,86)		
	ANX	,25	,56	,13	,44	,65	,34	,65	(,84)	
	VN	,15	,16	,53	,31	,33	,61	,59	,62	(,85)

DEP = Depression; ANX = Ansia; VN = Vulnerabilità Narcisistica

Studio preliminare

Raccogliere i dati

- Decidere la popolazione target
- Rappresentatività
- Quanti soggetti?
- Criteri di inclusione ed esclusione

Analisi degli item

- Analisi statistica per valutare appropriatezza psicometrica degli item
- Non deve essere l'unico criterio per decidere se mantenere o meno l'item nell'item pool

Analisi degli item test di prestazione massima

Difficoltà

- Quanti soggetti NON rispondono correttamente ad ogni item?
- Evitare percentuali estreme ($< ,10$, $> ,90$)

Formato dell'item	Difficoltà media ideale
Due alternative	,75-,85
Tre alternative	,66-,77
Quattro alternative	,62-,74
Cinque alternative	,60-,70
Completamento/risposta breve	,45-,55

Item troppo facili

- Risposta a caso o basso livello nel costrutto?
- Ruolo motivatore

Item	P(corretta)
i01	,95
i02	,99
i03	,97
i04	,98
i05	,99
i06	,98
i07	,96
i08	,99
i09	,99
i10	,98
i11	,99
i12	,97
i13	,99

Punteggio	Frequenza
8	,01
9	,01
10	,01
11	,02
12	,14
13	,81

Item troppo difficili

- Soggetti con livello particolarmente alto o risposta a caso?
- Demotivatori?

Analisi della distribuzione di frequenza dei distrattori

Qual è la capitale del Tagikistan?					
(a)		(b)		(c)	
Alternativa	<i>P</i>	Alternativa	<i>P</i>	Alternativa	<i>P</i>
Mosca	,06	Mosca	,05	Astana	,12
Pechino	,00	Pechino	,00	Biškek	,13
Parigi	,00	Taškent	,47	Taškent	,12
Dušambe	,94	Dušambe	,48	Dušambe	,65

Discriminatività - Indice D

- Confronto fra gruppi precostituiti (test orientati al criterio)

$$D = P(\textit{Focus}) - P(\textit{Controllo})$$

- Confronto fra "livelli alti" e "livelli bassi" al test (test orientati al costrutto)

$$D = P(\textit{Alto}) - P(\textit{Basso})$$

Valore di D	Commento
$D \geq ,40$	Ottima, nessuna revisione
$,30 \leq D < ,40$	Buona, revisioni minime
$,20 \leq D < ,30$	Sufficiente, revisioni parziali
$D < ,20$	Insufficiente, riformulazione o eliminazione

Discriminatività - Effect size

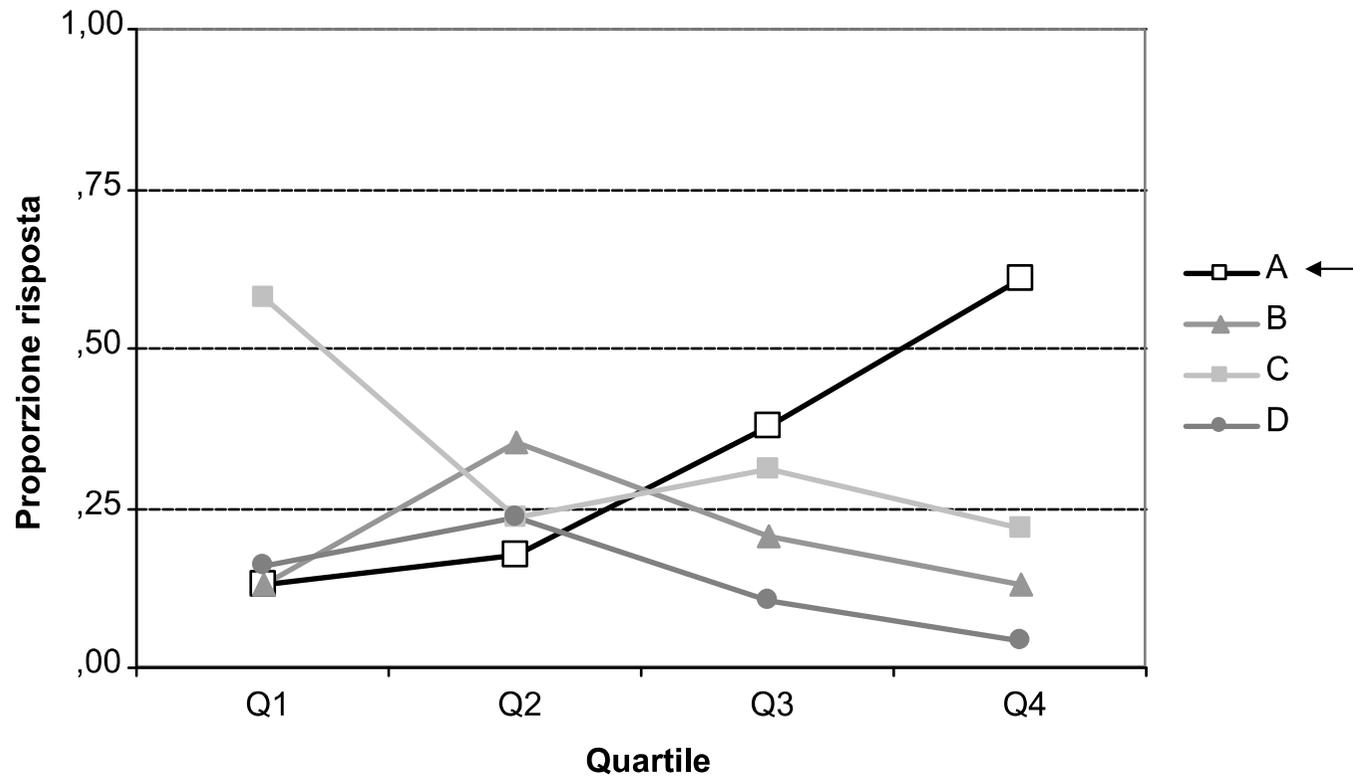
- Dimensione dell'effetto h

$$h = 2\arcsen\sqrt{P(\text{Focus})} - 2\arcsen\sqrt{P(\text{Controllo})}$$

$$h = 2\arcsen\sqrt{P(\text{Alto})} - 2\arcsen\sqrt{P(\text{Basso})}$$

Valore di h	Dimensione dell'effetto
$h < 0,20$	Trascurabile
$0,20 \leq h < 0,50$	Piccola
$0,50 \leq h < 0,80$	Moderata
$h \geq 0,80$	Grande

Discriminatività - Esame grafico



Correlazione item-totale corretta r_{it}

- Misura di quanto ogni item è rappresentativo dell'intera scala e di quanto ogni singolo item è in grado di rappresentare, da solo, il costrutto misurato dall'insieme degli altri item contenuti nel test
- Correlazione punto-biseriala
 - Il punteggio all'item è una vera dicotomia
 - Il punteggio al test è su scala a intervalli

$$r_{pbis} = \frac{(M_{corretta} - M_{errata})}{S_{TOT}} \sqrt{\frac{n_{corretta} n_{errata}}{n_{TOT} (n_{TOT} - 1)}}$$

Esempio

Item	0	0	0	0	0	0	0	1	1	1	1	1	1
Totale	2	8	5	3	3	7	9	13	7	14	7	10	14

- $M_{\text{corretta}} = 10,83$, $M_{\text{errata}} = 5,29$, $s_{TOT} = 3,92$, $n_{\text{corretta}} = 6$,
 $n_{\text{errata}} = 7$ e $n_{TOT} = 13$

$$r_{pbis} = \frac{(10,83 - 5,29)}{3,92} \sqrt{\frac{6 \times 7}{13(13 - 1)}} = ,73$$

Correlazione item-totale corretta r_{it}

- Perché "corretta"?

$$r_{it} = \frac{r_{pbis} S_{TOT} - S_i}{\sqrt{S_{TOT}^2 + S_i^2 - 2r_{pbis} S_{TOT} S_i}} \quad r_{it} = \frac{,73 \times 3,92 - ,52}{\sqrt{3,92^2 + ,52^2 - 2 \times ,73 \times 3,92 \times ,52}} = ,67$$

- Valore sufficiente: ,20
- Valore ottimale: ,30
- Non utilizzabile se utilizzata la correzione per guessing

Correzione per guessing

- Limitare le risposte a caso

$$X_{\text{corretto}} = C - \frac{E}{(k-1)}$$

- Valore atteso per chi risponde *completamente* a caso = 0
- Penalizza davvero i "gambler"?
- La correzione è realmente utile in:
 - test di velocità
 - test ad alto livello di difficoltà
 - test nei quali l'obiettivo è raggiungere un certo punteggio minimo

Analisi degli item test di prestazione tipica

Numero di casi mancanti

- Differenza di significato della risposta omessa test di prestazione massima vs test di prestazione tipica
- Cause
 - Il soggetto rifiuta di rispondere, per cui intenzionalmente non ha indicato la sua risposta
 - Il soggetto non sa cosa rispondere e, nel dubbio, non indica alcuna risposta
 - Il soggetto ha semplicemente saltato l'item, per sbaglio
 - Il dato non è stata inserito per errore nel database

Tipi di missing

- Missing per definizione della sottopopolazione
- Missing completely at random
 - La probabilità che un dato sia mancante è completamente indipendente sia dalla caratteristica misurata dal test di cui l'item fa parte, sia da qualsiasi altra variabile considerata nell'analisi
 - Little's MCAR Test
- Missing at random
 - La probabilità che un dato sia mancante non dipende dalla caratteristica misurata dal test di cui l'item fa parte una volta controllato l'effetto di un altre variabili, dato che queste altre variabili innescano il meccanismo che determina i dati mancanti
- Missing not at random

Gestire i missing

- **Pensarci prima**

- Ridurre la probabilità dei *missing* con un miglior disegno della ricerca, formulando meglio gli item, producendo un migliore *lay-out* del testo, etc.

- **Statisticamente**

- Solo se i missing sono casuali e si dispone di un campione sufficientemente ampio
 - E' assolutamente sconsigliabile utilizzare soluzioni semplicistiche come la sostituzione del missing con la media dei valori disponibili (purtroppo un'opzione in SPSS...)
 - Prorating → item test psicologici
 - FIML e FIWLS
 - Multiple Imputation
- } Qualunque tipo di dato

Prorating

- Utilizzabile solo per i missing relativi ad item di test psicologici di cui si conosca già la struttura fattoriale e/o la procedura di scoring
- Si calcola la media dei punteggi disponibili, e si sostituisce al valore mancante questo valore arrotondato all'intero
- NB: almeno il 50% dei punteggi degli item deve essere disponibile!
- Attenzione ad invertire i punteggi degli item reverse prima di fare questa operazione!

Esempio

Item	Ss01	Ss02	Ss03	Ss04	Ss05	Ss06
E' depressa, triste	1	2	1	3	miss	1
E' rilassata, gestisce bene lo stress (R)	5	miss	5	5	5	4
Può essere tesa	miss	3	3	1	miss	1
Si preoccupa molto	2	3	miss	3	1	3
È emotivamente stabile, non si turba facilmente (R)	5	5	miss	4	miss	4
Può essere lunatica	1	1	1	3	1	2
Rimane calma nelle situazioni di tensione (R)	5	4	5	5	miss	4
Diventa facilmente apprensiva	3	1	3	1	miss	2

NB: (R) = item reverse

Esempio

Item	Ss01	Ss02	Ss03	Ss04	Ss05	Ss06
E' depressa, triste	1	2	1	3	miss	1
E' rilassata, gestisce bene lo stress (R)	5	miss	5	5	5	4
Può essere tesa	miss	3	3	1	miss	1
Si preoccupa molto	2	3	miss	3	1	3
È emotivamente stabile, non si turba facilmente (R)	5	5	miss	4	miss	4
Può essere lunatica	1	1	1	3	1	2
Rimane calma nelle situazioni di tensione (R)	5	4	5	5	miss	4
Diventa facilmente apprensiva	3	1	3	1	miss	2

NB: (R) = item reverse

< 50% di missing
prorating ok

> 50% di missing
no prorating

Esempio

Item	Ss01	Ss02	Ss03	Ss04	Ss05	Ss06
E' depressa, triste	1	2	1	3	miss	1
E' rilassata, gestisce bene lo stress (R)	1	miss	1	1	1	2
Può essere tesa	miss	3	3	1	miss	1
Si preoccupa molto	2	3	miss	3	1	3
È emotivamente stabile, non si turba facilmente (R)	1	1	miss	2	miss	2
Può essere lunatica	1	1	1	3	1	2
Rimane calma nelle situazioni di tensione (R)	1	2	1	1	miss	2
Diventa facilmente apprensiva	3	1	3	1	miss	2

NB: (R) = item reverse

Inversione dei punteggi degli item reverse

In questo caso, poiché la scala è da 1 a 5 si avrà:

(1→5) (2→4) (3→3) (4→2) (5→1)

Esempio

Item	Ss01	Ss02	Ss03	Ss04	Ss05	Ss06
E' depressa, triste	1	2	1	3	miss	1
E' rilassata, gestisce bene lo stress (R)	1	miss	1	1	1	2
Può essere tesa	miss	3	3	1	miss	1
Si preoccupa molto	2	3	miss	3	1	3
È emotivamente stabile, non si turba facilmente (R)	1	1	miss	2	miss	2
Può essere lunatica	1	1	1	3	1	2
Rimane calma nelle situazioni di tensione (R)	1	2	1	1	miss	2
Diventa facilmente apprensiva	3	1	3	1	miss	2
Media valori disponibili	1.43	1.86	1.67	1.88		1.88
Punteggio prorated	1	2	2	2		2

NB: (R) = item reverse

FIML e MI

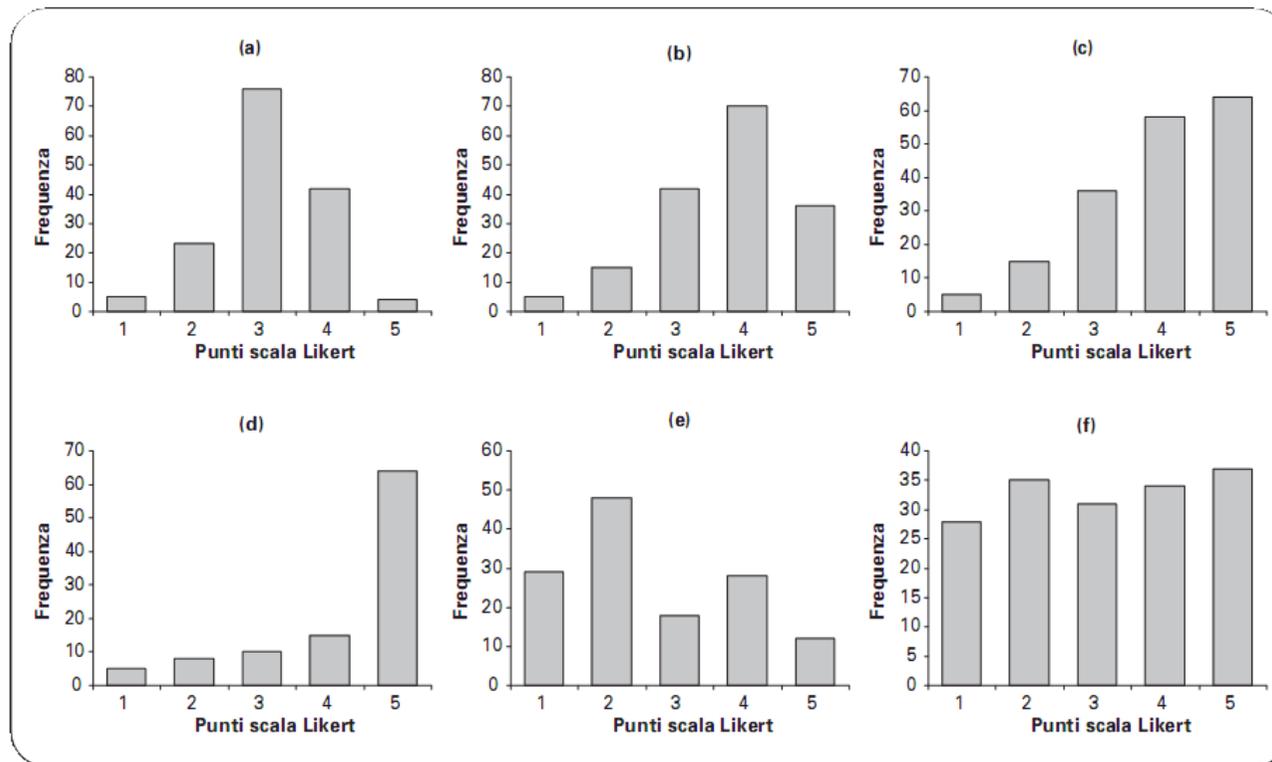
- Full Information Maximum Likelihood (FIML, dati metrici) e Full Information Weighted Least Squares (FIWLS, dati categoriali) usano tutti i dati osservati disponibili per produrre stime dei parametri (medie, coefficienti di regressione, saturazioni, etc.) che massimizzino la probabilità dei dati osservati di provenire dalla popolazione implicata da quelle stime
- Molto usata in analisi fattoriale e modelli di equazioni strutturali → alternativa migliore di prorating, ma necessita di software e competenze adeguate

Multiple Imputation

- Multiple Imputation (MI) è un metodo "data-based" per ottenere gli stessi risultati di FIML/FIWLS
- Questo metodo si basa sui dati disponibili e mediante particolari algoritmi genera un certo numero (di solito 5) di possibili valori per i dati mancanti, che costituiscono altrettanti dataset completi.
- L'analisi viene poi eseguita su ognuno dei dataset e i risultati di queste analisi vengono combinati (*pooled*) per ottenere un unico risultato

Forma della distribuzione

- Perché la distribuzione normale?



Punteggio minimo e punteggio massimo

- Adeguatezza scala di risposta di tipo Likert
- Cause
 - Item mal formulati
 - Scale di tipo Likert non adeguate per l'item (o viceversa)
 - Campioni troppo omogenei

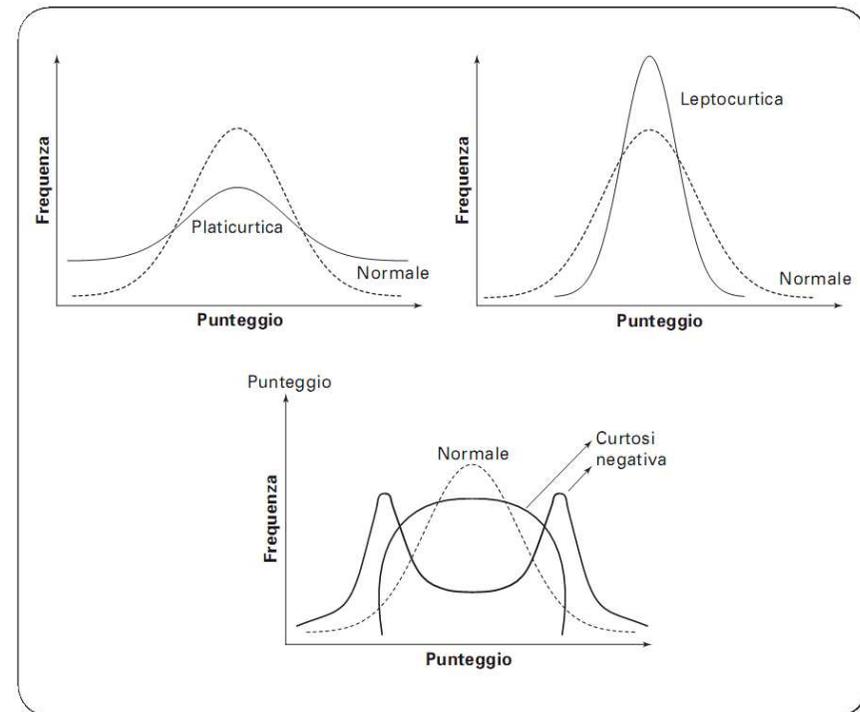
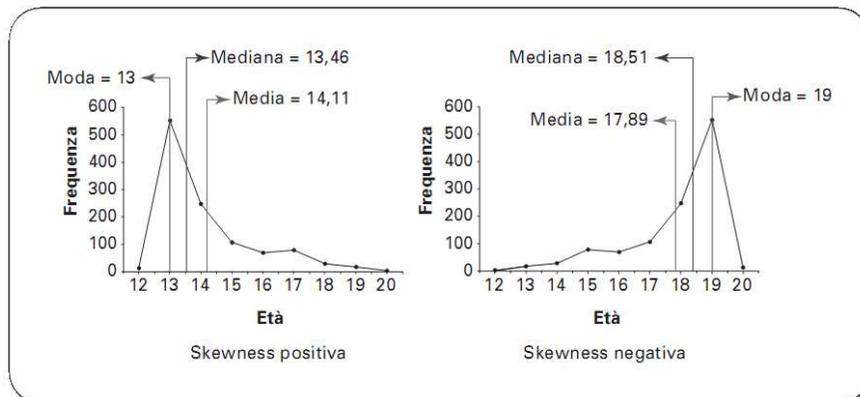
Indici di tendenza centrale e dispersione

- Valori indicativi (nessun supporto empirico)
- Possibile utilizzo indici ordinali (quartili)

Punti della scala	Range media	Punti della scala	Range media	DS
1-2-3-4	1,75-3,25	0-1-2-3	0,75-2,25	0,6
1-2-3-4-5	2,00-4,00	0-1-2-3-4	1,00-3,00	0,8
1-2-3-4-5-6	2,25-4,75	0-1-2-3-4-5	1,25-3,75	1,0
1-2-3-4-5-6-7	2,50-5,50	0-1-2-3-4-5-6	1,50-4,50	1,2
1-2-3-4-5-6-7-8	2,75-6,25	0-1-2-3-4-5-6-7	1,75-5,25	1,4
1-2-3-4-5-6-7-8-9	3,00-7,00	0-1-2-3-4-5-6-7-8	2,00-6,00	1,6

Indici di forma della distribuzione

- Skewness e curtosi
- Muthén e Kaplan (1985) e l'intervallo [-1;+1]



Discriminatività

- Confronto fra gruppi precostituiti (test orientati al criterio)
- Confronto fra "livelli alti" e "livelli bassi" al test (test orientati al costruito)
- Test t per campioni indipendenti (info metrica)
- Test di Mann-Whitney (info ordinale)
- Dimensione dell'effetto almeno "moderata"

d	r	Effect size
$< 0,20 $	$< 0,10 $	Trascurabile
$ 0,20 - 0,50 $	$ 0,10 - 0,30 $	Piccola
$ 0,50 - 0,80 $	$ 0,30 - 0,50 $	Moderata
$> 0,80 $	$> 0,50 $	Grande

Correlazione item-totale corretta r_{it}

$$r = \frac{n \sum_{j=1}^n i_j t_j - \sum_{j=1}^n i_j \sum_{j=1}^n t_j}{\sqrt{\left[n \sum_{j=1}^n i_j^2 - \left(\sum_{j=1}^n i_j \right)^2 \right] \left[n \sum_{j=1}^n t_j^2 - \left(\sum_{j=1}^n t_j \right)^2 \right]}}$$

$$r_{it} = \frac{rs_t - s_i}{\sqrt{s_t^2 + s_i^2 - 2rs_t s_i}}$$

Item	4	2	4	4	3	3	4	3	4	3	4	5	3	1	2
Totale	41	27	41	43	29	26	43	36	41	30	36	39	34	35	30

$r = ,64$, $s_i = 1,00$, e $s_t = 5,68$

$$r_{it} = \frac{,64 \times 5,68 - 1,00}{\sqrt{5,68^2 + 1,00^2 - 2 \times ,64 \times 5,68 \times 1,00}} = ,52$$

- Valore sufficiente: ,20
- Valore ottimale: ,30

Ridondanza

- Correlazioni maggiori di $|,70|$

Inter-Item Correlation Matrix

	ITEM1	ITEM2	ITEM3	ITEM4	ITEM5	ITEM6
ITEM1	1.000	.716	.713	.694	.691	.694
ITEM2	.716	1.000	.679	.653	.697	.698
ITEM3	.713	.679	1.000	.595	.685	.749
ITEM4	.694	.653	.595	1.000	.595	.665
ITEM5	.691	.697	.685	.595	1.000	.700
ITEM6	.694	.698	.749	.665	.700	1.000

The covariance matrix is calculated and used in the analysis.

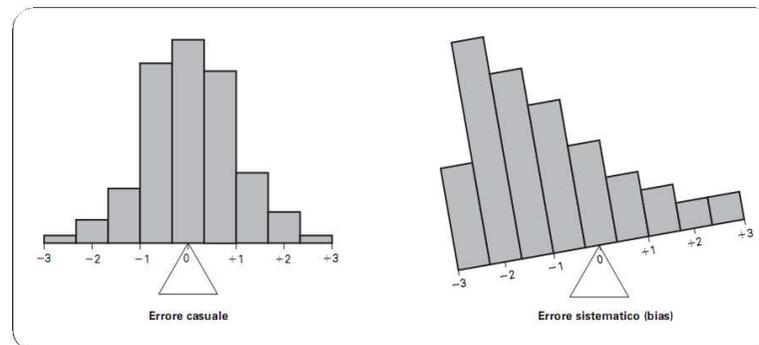
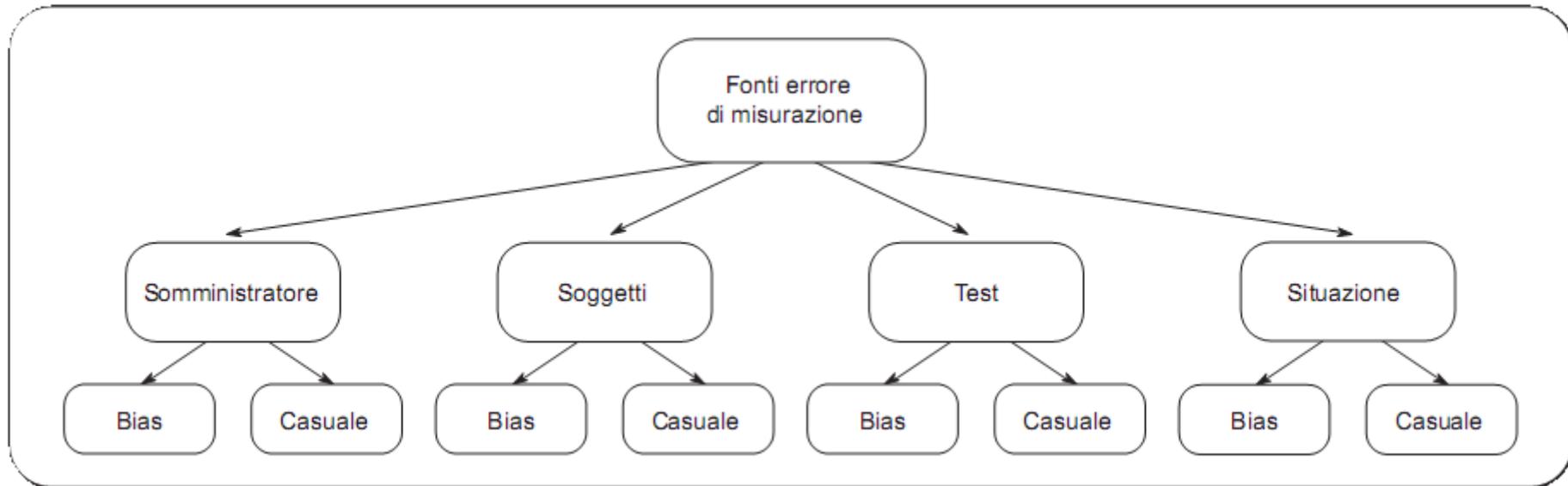
Attendibilità

Attendibilità (*reliability*)

- **Proprietà psicometrica relativa all'accuratezza con cui un test o una scala misura una certa variabile psicologica**
- Perché è diversa dalla validità?
- Proporzione del punteggio misurato che non riflette l'errore di misurazione

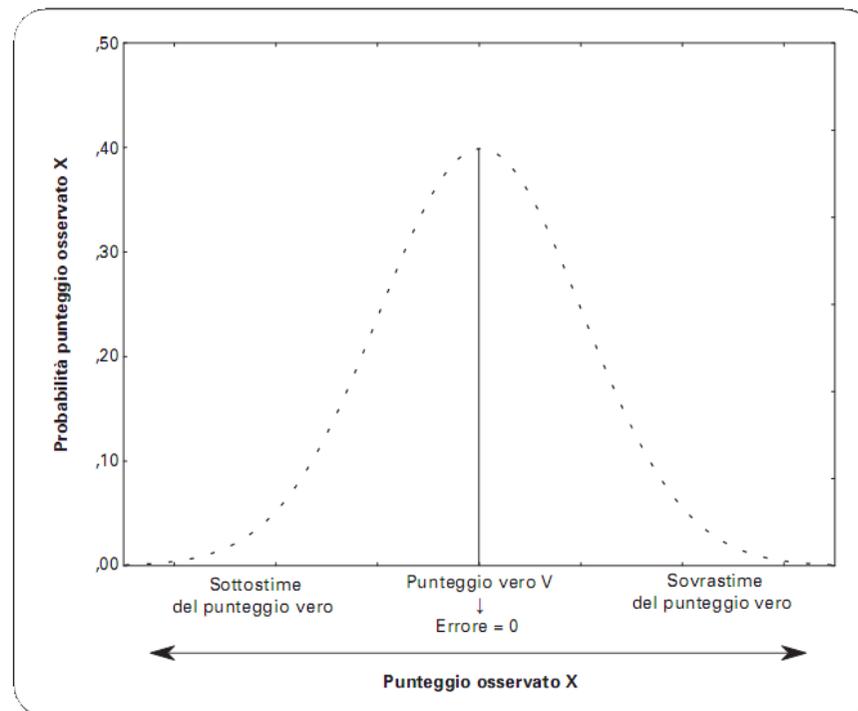
$$X = V + E$$

Fonti di errore nei test psicologici



Errore casuale nella TCT

- Variabile aleatoria o stocastica



Errore casuale nella TCT

- Punteggio osservato in un dato soggetto in una data situazione è solo uno dei possibili punteggi osservabili
- Il punteggio vero V , quindi, può essere concepito come **la media dei punteggi osservati di infinite misurazioni**
- $E(X) = V$

Assunzioni sull'errore di misurazione

1. L'errore di misurazione è casuale, è distribuito normalmente e ha valore atteso (media) uguale a zero: $E(E) = 0$.
 - Avremo quindi che $E(X) = E(V + E) = E(V) + E(E) = E(V) + 0 = E(V)$. Dato che V è costante, $E(X) = E(V) = V$

2. Punteggio vero ed errore di una misurazione sono indipendenti fra loro (= covarianza fra V ed E è uguale a zero)
 - L'entità dell'errore non dipende dalla quantità di punteggio vero, e quindi che l'errore di misurazione non dipende dalla quantità di costrutto posseduta dal soggetto

$$Cov(V,E) = \frac{\sum (V - M_V)(E - M_E)}{n}$$

Assunzioni sull'errore di misurazione

3. Gli errori di misurazione di due misurazioni sugli stessi soggetti sono indipendenti fra loro (= la covarianza tra due distribuzioni di errori in due misurazioni diverse è uguale a zero)
 - L'errore di misurazione che distorce il punteggio in un item non è in relazione con quello che distorce il punteggio in un altro item

Stimare l'attendibilità

$$r_{tt} = \frac{V}{V + E} = \frac{V}{X}$$

$$s_X^2 = \frac{\sum (X - M_X)^2}{n} = \frac{\sum [(V + E) - (M_V + M_E)]^2}{n} = \frac{\sum [(V - M_V) + (E - M_E)]^2}{n} =$$

$$= \frac{\sum (V - M_V)^2}{n} + \frac{\sum (E - M_E)^2}{n} + \frac{2 \sum [(V - M_V)(E - M_E)]}{n}$$

↓
↓
↓

Varianza
punteggio vero
 s_V^2
Varianza
errore
 s_E^2
Covarianza
errore-punteggio vero
(= 0!)

$$s_X^2 = s_V^2 + s_E^2$$

Stimare l'attendibilità

- **Attendibilità = rapporto tra varianza del punteggio vero e varianza del punteggio osservato (proporzione di punteggio vero che non riflette l'errore casuale di misurazione)**

$$r_{tt} = \frac{s_V^2}{s_X^2} = \frac{s_V^2}{s_V^2 + s_E^2}$$

$$r_{tt} = \frac{s_V^2}{s_X^2} = 1 - \frac{s_E^2}{s_V^2 + s_E^2}$$

Come stimare l'attendibilità?

- Due somministrazioni (stabilità)
 - Forme parallele e alternative
 - Test retest
- Una somministrazione (coerenza interna)
 - Split-half
 - Alpha di Cronbach

Parallelismo e tau-equivalenza

Tipo di misure	Relazione fra i punteggi veri τ	Relazione fra le varianze d'errore σ_E^2
Parallele	$\tau_1 = \tau_2$	$\sigma_{E1}^2 = \sigma_{E2}^2$
Tau equivalenti	$\tau_1 = \tau_2$	$\sigma_{E1}^2 \neq \sigma_{E2}^2$
Essenzialmente tau equivalenti	$\tau_1 = \tau_2 + c$	$\sigma_{E1}^2 \neq \sigma_{E2}^2$
Congeneriche	$\tau_1 = a\tau_2 + c$	$\sigma_{E1}^2 \neq \sigma_{E2}^2$

- Se parallelismo non soddisfatto → forme **alternative coefficiente di equivalenza** (valore ottimale: ,70)

Test-retest

The only meaningful estimate of reliability for a heterogeneous test is the retest variety

(Guilford, 1965, p. 450)

- Stabilità dei punteggi nel tempo (→solo tratti)
- Due o più somministrazioni dello stesso test agli stessi soggetti
- Valore accettabile: ,70
- Valore ottimale: ,90

Quanto tempo?

- **Coefficiente di replicabilità** (dependability) = poche ore/giorni
- **Coefficiente di stabilità** (varie settimane/mesi)
- Solitamente 2-4 settimane

Fonti di errore due somministrazioni

- Ricordo risposte fornite, strategie utilizzate, etc.
- Fluttuazioni nel livello del costrutto
- Variazione condizioni di somministrazione
- Interazione soggetto \times momento somministrazione

- Forme parallele
 - Interazione soggetto \times contenuto
- Forme parallele
 - Dipendenza seriale viola assunzione indipendenza errori di misurazione item diversi \rightarrow varianza di errore casuale è in realtà sistemica in quanto stabile nel tempo e "gonfia" l'attendibilità

Split-half

- Divisione a metà del test
 - Pari/dispari
 - Prima/Seconda metà
 - Etc.
- Correlazione fra i punteggi nelle due metà → non è attendibilità di un test con k item, ma di due test paralleli di $k/2$ item!
- Profezia di Spearman-Brown

$$r_{nt} = \frac{\left(\frac{k_{Modificata}}{k_{Esistente}} \right) r_{tt}}{1 + \left(\frac{k_{Modificata}}{k_{Esistente}} - 1 \right) r_{tt}} \quad \longrightarrow \quad r_{SH} = \frac{2r_{tt}}{1 + r_{tt}}$$

- Se due metà non parallele → coefficiente di Rulon

$$r_{tt} = 1 - \frac{s_D^2}{s_X^2}$$

s_D^2 è la varianza della differenza fra i punteggi nelle due metà

KR-20 e KR-21

- Problemi con split-half
- Soluzione: indice di attendibilità che sia la media di tutte le possibilità attendibilità split-half

$$KR_{20} = \frac{k}{k-1} \left[1 - \frac{\sum_{j=1}^k P_j(1-P_j)}{s_X^2} \right]$$

Se item hanno la stessa difficoltà

$$KR_{21} = \frac{k}{k-1} \left[1 - \frac{kM_P M_{(1-P)}}{s_X^2} \right]$$

- k = numero di item del test, P_j = proporzione di risposte corrette all'item j , s_X^2 = varianza del punteggio totale al test, M_P = media della proporzione di risposte corrette e $M_{(1-P)}$ = media della proporzione di risposte errate

Alpha (α) di Cronbach

- Estensione della KR-20 ad item non dicotomici

$$\alpha = \frac{k}{k-1} \left[1 - \frac{\sum_{j=1}^k s_j^2}{s_X^2} \right]$$

- k = numero di item, s_j^2 = varianza del punteggio all'item j , s_X^2 = varianza del punteggio totale al test

$$\alpha_z = \frac{kM_{rii}}{1 + M_{rii}(k-1)}$$

- k = numero di item, M_{rii} = correlazione media inter-item

Linee guida

<i>Valore di α o KR-20</i>		<i>Interpretazione</i>
$\alpha \geq ,90$		Ottimo
$,80 \leq \alpha < ,90$		Buono
$,70 \leq \alpha < ,80$		Discreto
$,60 \leq \alpha < ,70$		Sufficiente
$\alpha < ,60$		Inadeguato

Perché più item = più attendibilità?

- Ogni volta che vengono aggiunti item si modificano la proporzione di varianza totale, quella di varianza vera e quella di errore del test finale
- La varianza del punteggio vero aumenta in progressione quadratica, quella del punteggio osservato in progressione lineare

Coerenza interna \neq omogeneità

- Alpha di Crobach alto = alta coerenza interna ma non necessariamente unidimensionalità
- Costrutti ampi (domini) e specifici (facet)

<i>n</i>	Correlazione media inter-item						
	Costrutti ampi			Costrutti specifici			
	,10	,20	,30	,40	,50	,60	,70
5	,36	,56	,68	,77	,83	,88	,92
10	,53	,71	,81	,87	,91	,94	,96
15	,63	,79	,87	,91	,94	,96	,97
20	,69	,83	,90	,93	,95	,97	,98
25	,74	,86	,91	,94	,96	,97	,98
30	,77	,88	,93	,95	,97	,98	,99
35	,80	,90	,94	,96	,97	,98	,99
40	,82	,91	,94	,96	,98	,98	,99

Coerenza interna \neq omogeneità

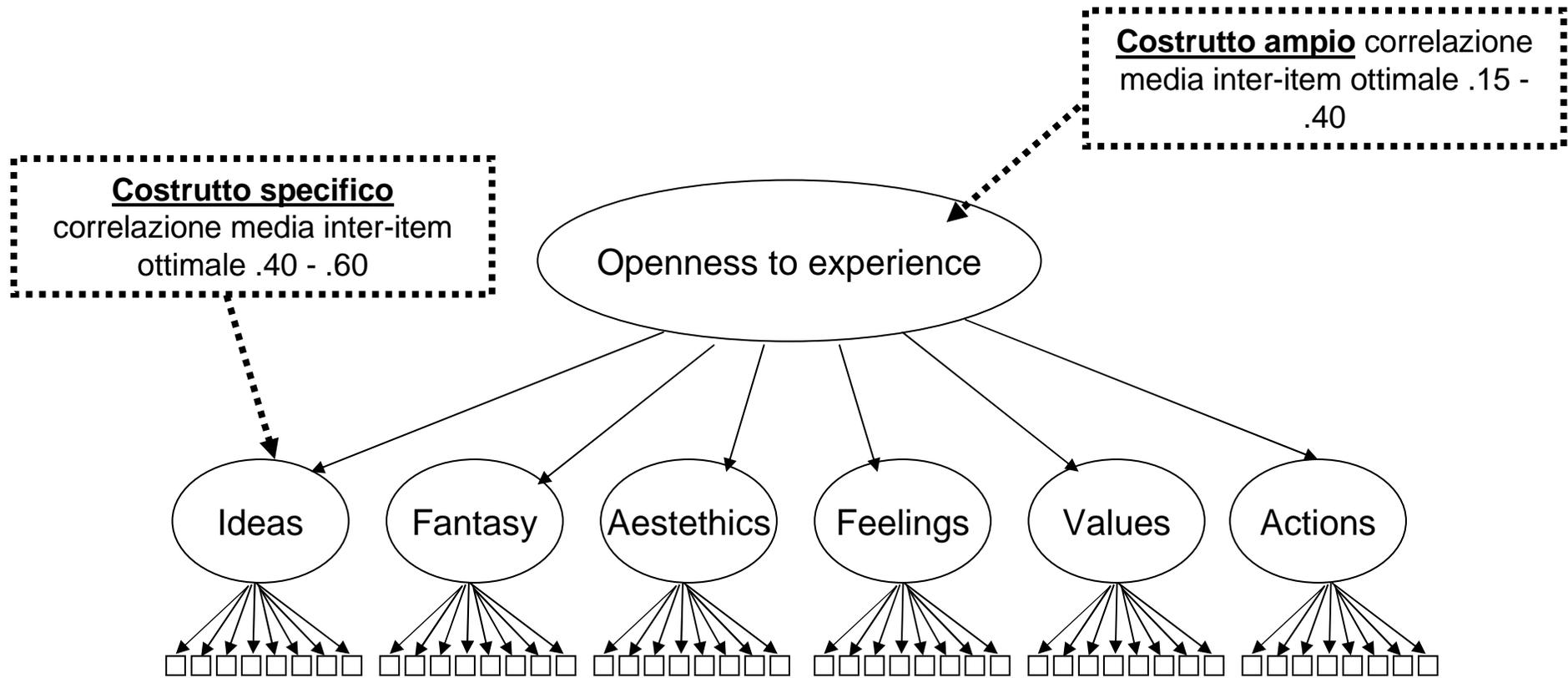
(a)							(b)						
	i01	i02	i03	i04	i05	i06		i01	i02	i03	i04	i05	i06
i01	-						i01	-					
i02	,77	-					i02	,48	-				
i03	,81	,82	-				i03	,50	,51	-			
i04	,25	,28	,30	-			i04	,52	,50	,50	-		
i05	,23	,32	,31	,78	-		i05	,47	,49	,50	,53	-	
i06	,35	,37	,29	,81	,81	-	i06	,53	,50	,47	,50	,50	-

Errore standard di alpha (Cortina, 1993)

$$\alpha_{SE} = \frac{s_{rij}}{\sqrt{0,5k(k-1)-1}}$$

,0661
,0047

s_{rij} = deviazione standard dei coefficienti di correlazione della matrice, k = numero di item



<i>Costrutto</i>	<i>α di Cronbach</i>	<i>r media inter-item</i>
Ampio (e.g., estroversione)	≥ .65	.15 - .40
Specifico (e.g., loquacità)	≥ .80	.40 - .60

Correlazione multipla al quadrato

- Non è un indice specifico delle analisi psicometriche
- Proporzione di varianza condivisa fra una variabile con un insieme di altre variabili
- Valore accettabile: ,10

$$R_{X.YZ}^2 = \frac{r_{XY}^2 + r_{XZ}^2 - 2r_{XY}r_{XZ}r_{YZ}}{1 - r_{YZ}^2}$$

Alpha senza l'item

- Quale sarebbe il valore di α se l'item venisse escluso dal test?
- Se item tutti necessari, eliminazione = peggioramento α
- Variazione minima sostanziale: +,01

Dimensionalità e validità del test

Unidimensionalità vs coerenza interna

- La scala (test) misura una cosa sola?
- Undimensionalità = omogeneità \neq coerenza interna
- Si possono ottenere valori di α alti anche in casi di non unidimensionalità

Stesso α (e M_{rij}), diversa dimensionalità

(a)

	<i>i01</i>	<i>i02</i>	<i>i03</i>	<i>i04</i>	<i>i05</i>	<i>i06</i>	<i>i07</i>	<i>i08</i>	<i>i09</i>
<i>i01</i>	1.00								
<i>i02</i>	.54	1.00							
<i>i03</i>	.50	.49	1.00						
<i>i04</i>	.21	.22	.20	1.00					
<i>i05</i>	.10	.24	.15	.61	1.00				
<i>i06</i>	.18	.17	.16	.58	.64	1.00			
<i>i07</i>	.13	.12	.11	.17	.10	.18	1.00		
<i>i08</i>	.21	.14	.17	.11	.14	.10	.55	1.00	
<i>i09</i>	.19	.15	.18	.12	.16	.10	.52	.50	1.00

(b)

	<i>i01</i>	<i>i02</i>	<i>i03</i>	<i>i04</i>	<i>i05</i>	<i>i06</i>	<i>i07</i>	<i>i08</i>	<i>i09</i>
<i>i01</i>	1.00								
<i>i02</i>	.21	1.00							
<i>i03</i>	.24	.27	1.00						
<i>i04</i>	.22	.24	.21	1.00					
<i>i05</i>	.28	.27	.23	.22	1.00				
<i>i06</i>	.24	.26	.26	.21	.28	1.00			
<i>i07</i>	.27	.27	.28	.29	.24	.24	1.00		
<i>i08</i>	.24	.22	.27	.25	.25	.27	.28	1.00	
<i>i09</i>	.29	.24	.25	.23	.28	.28	.29	.27	1.00

$$M_{rij} = .254, \alpha_Z = .754$$

Analisi fattoriale - aspetti teorici

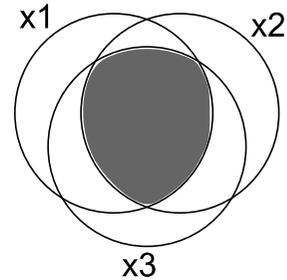
It is easy to see that correlation must be the consequence of the variations of the two organs being partly due to common causes. If they were wholly due to common causes, the correlation would be perfect, as is approximately the case with the symmetrically disposed parts of the body. If they were in no respect due to common causes, the correlation would be nil. Between these two extremes are an number of intermediate cases.

Galton (1888, p. 135)

Analisi fattoriale - aspetti teorici

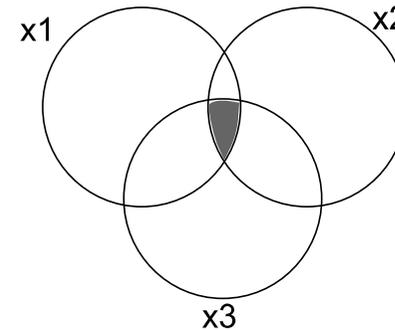
- Varianza comune

Molta variabilità in comune



	$x1$	$x2$	$x3$
$x1$	1,00		
$x2$,58	1,00	
$x3$,61	,57	1,00

Poca variabilità in comune



	$x1$	$x2$	$x3$
$x1$	1,00		
$x2$,08	1,00	
$x3$,14	,15	1,00

Cosa vuole trovare l'analisi fattoriale?

- Raggruppamenti di variabili (ossia, i fattori) derivati in base al grado di correlazione delle variabili stesse
- Quanti fattori sono necessari per riuscire a riassumere in modo adeguato le relazioni fra le variabili
- Quali variabili appartengono ad un fattore e qual è la forza del loro legame col fattore
- Un sistema di riferimento geometrico che permetta di descrivere le relazioni fra le variabili
- Il livello (punteggio fattoriale) di ogni soggetto in ogni fattore (punteggio del soggetto nel costrutto che si intende misurare) come somma pesata (combinazione lineare) dei punteggi nelle variabili osservate

Utilità dell'analisi fattoriale

- Sviluppo di test psicologici che misurino costrutti operazionalmente definiti
- Definizione operativa dei costrutti: in che misura gli item del test misurano la stessa cosa?
- Validità del test: gli item del test sono tutte misure valide del costrutto?
- Costruzione e verifica di teorie: in quale relazione stanno i costrutti misurati dal test?



Analisi fattoriale esplorativa e confermativa

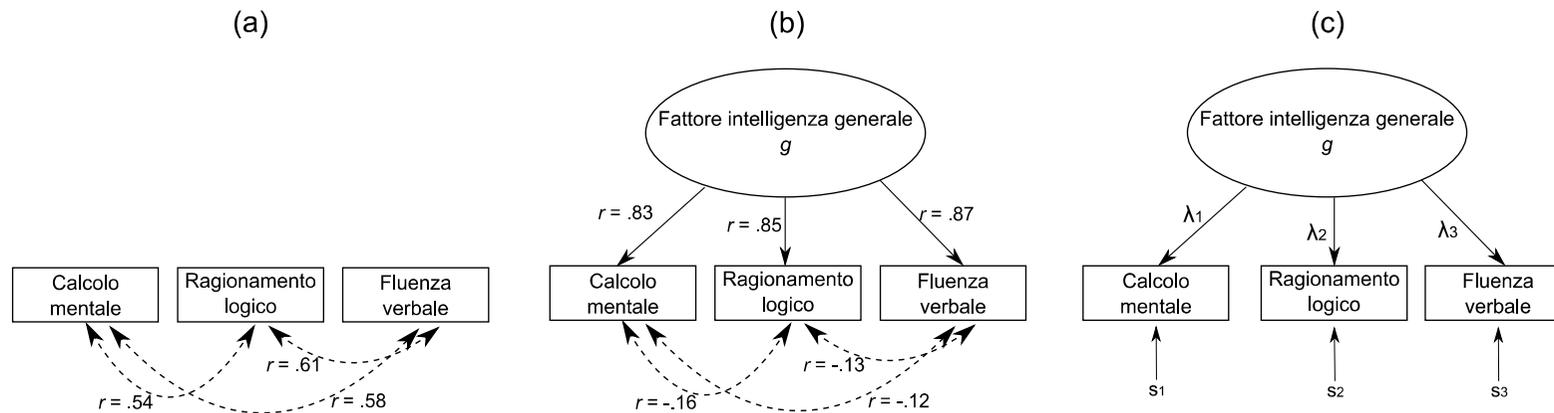
- L'analisi fattoriale esplorativa e i Big Five
- Indagine esplorativa vs verifica di ipotesi

Analisi fattoriale esplorativa

- Scomposizione matrice di correlazione
- Varianza comune e unica

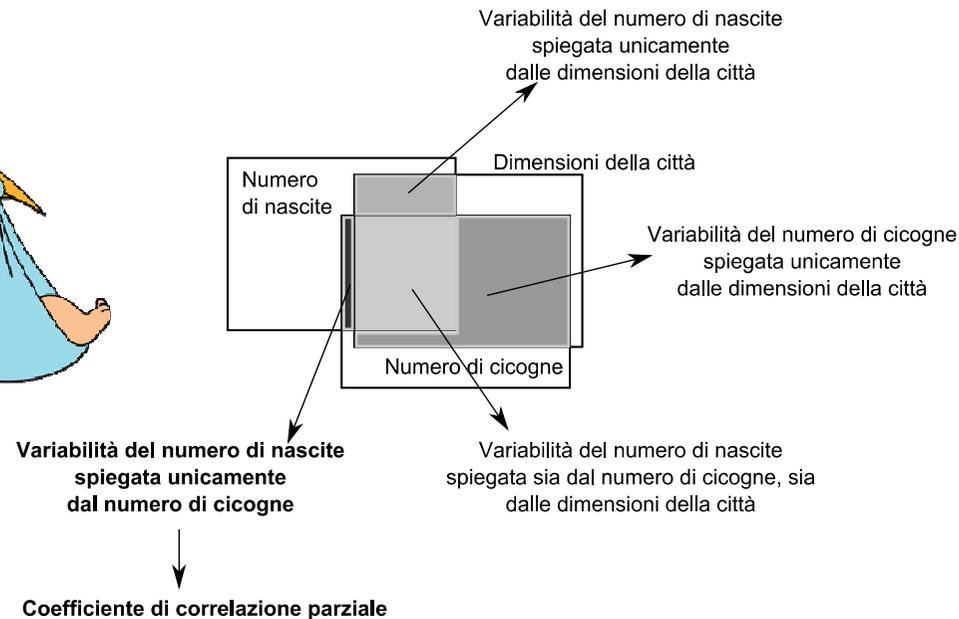
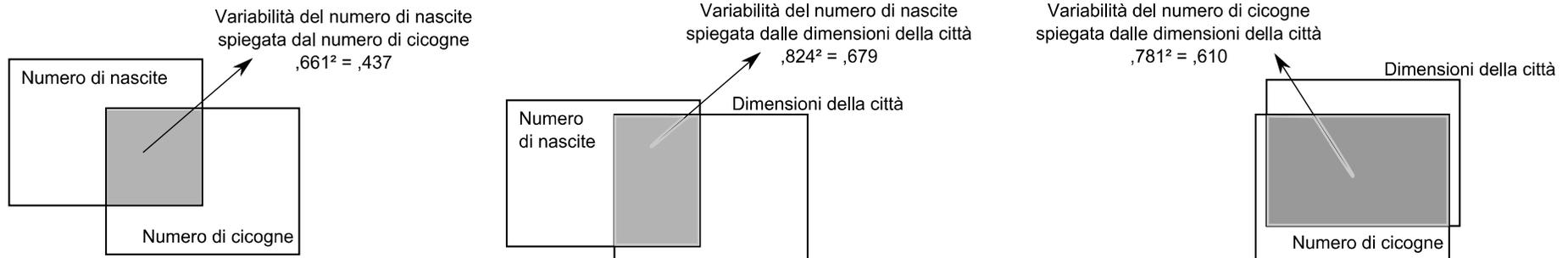
Principi fondamentali analisi fattoriale

1. Causazione



- Modello indicatori riflessivi
- Fattori come variabili di controllo (relazione spuria fra item)

Relazione spuria e correlazione parziale



Principi fondamentali analisi fattoriale

2. Parsimonia

- Spendere poco (pochi fattori) e fare bella figura (molta varianza comune spiegata!)
- “It is the faith of all science that an unlimited number of phenomena can be comprehended in terms of a limited number of concepts or ideal constructs” (Thurstone, L. L. (1935). *Multiple Factor Analysis*. Chicago: University of Chicago Press, p. 44)

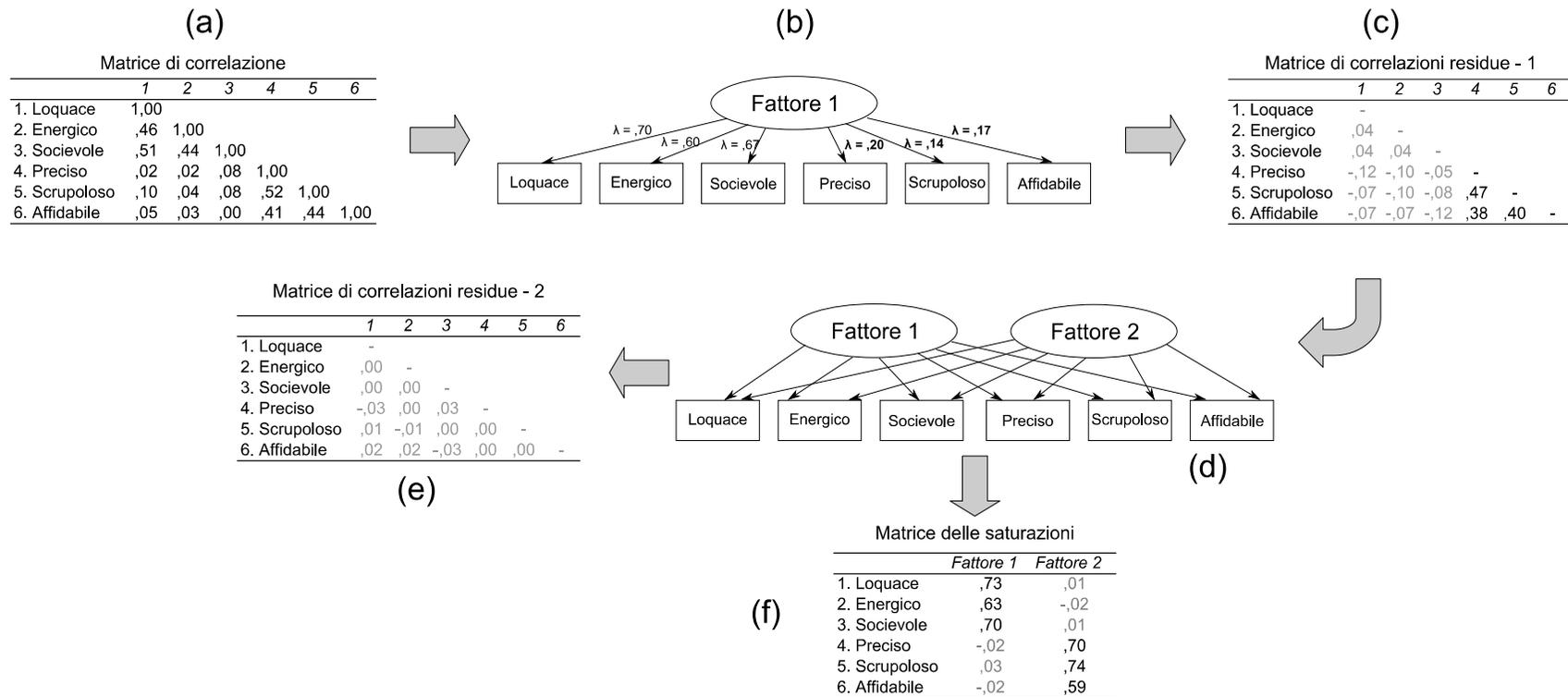
Principi fondamentali analisi fattoriale

3. Struttura semplice (Thurstone, 1947) e numero minimo di item per fattore

- Each row of the rotated matrix should contain at least one zero.
- In each factor the minimum number of zero loadings should be the number of factors in the rotation.
- For every pair of factors there should be variables with zero loadings on one and significant loadings on the other.
- For every pair of factors a large proportion of the loadings should be zero, at least in a matrix with a large number of factors.
- For every pair of factors there should be only a few variables with significant loadings on both factors.

Principi fondamentali analisi fattoriale

3. Struttura semplice e numero minimo di item per fattore



Saturazione (*factor loading*)

- Correlazione fra punteggio nell'item e nel fattore (costrutto)
- (Quasi) sempre compresa fra -1 e +1
- Se elevata al quadrato = varianza comune spiegata dal fattore per un item
- Somma saturazioni al quadrato = comunaltà (h^2)
- 1 - comunaltà = unicità (u^2)
- Sostanziale vs significativa

Struttura semplice (concezione moderna)

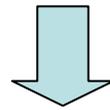
- Una sola saturazione sostanziale (uguale o maggiore di un valore minimo, in genere $|,30|$ o $|,40|$) su ogni riga (per ogni variabile osservata) → target loading
- Tutte le altre saturazioni il più basse possibile, ossia comprese fra $-,10$ e $+,10$, e/o con una differenza di almeno $,20$, in valore assoluto, rispetto alla saturazione principale → cross-loading

Struttura semplice - esempi

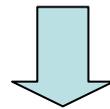
(a)				(b)				(c)			
	Fattore 1	Fattore 2	Fattore 3		Fattore 1	Fattore 2	Fattore 3		Fattore 1	Fattore 2	Fattore 3
i01	,37	,02	,08	i01	,11	,04	,07	i01	,01	,04	,07
i02	,45	,00	,04	i02	,51	,07	,03	i02	,51	,07	,03
i03	,46	,09	,05	i03	,31	,06	,09	i03	,03	,06	,09
i04	,12	,43	,05	i04	,49	,44	,51	i04	,49	,14	,08
i05	,02	,34	,12	i05	,05	,56	,17	i05	,05	,56	,17
i06	,03	,31	,15	i06	,11	,41	,11	i06	,11	,41	,11
i07	,15	,11	,44	i07	,04	,08	,50	i07	,44	,50	,35
i08	,05	,09	,42	i08	,05	,20	,35	i08	,05	,39	,35
i09	,01	,09	,33	i09	,16	,05	,45	i09	,16	,05	,42

Interpretazione dei fattori

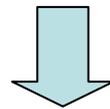
Perché gli item saturano sui fattori?



Perché sono correlati fra loro



Perché sono correlati fra loro?



Perché sono operationalizzazioni dello stesso costrutto (in teoria...)

Analisi fattoriale - aspetti statistici

- Equazione di specificazione (modello indicatori riflessivi)

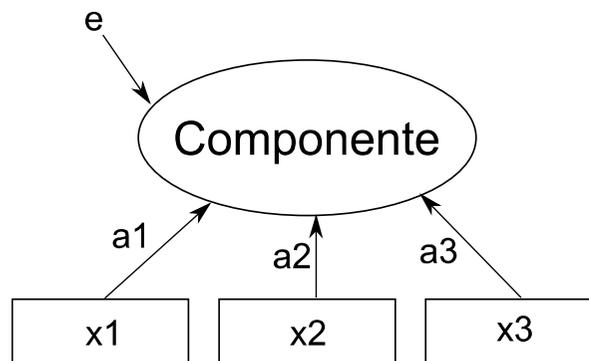
$$z_{ik} = \lambda_{i1}F_{1k} + \lambda_{i2}F_{2k} + \dots + \lambda_{ip}F_{pk} + \lambda_{is}S_{ik} + \lambda_{ie}E_{ik}$$

- Fattori comuni (F) = variabili che spiegano la variabilità *condivisa* dagli item
- Fattori specifici (S) = variabili che spiegano la variabilità *specifica e sistematica* degli item
- Fattori di errore (E) = parte di varianza dell'item dovuta all'errore di misurazione (*casuale*)

Analisi componenti principali

- Modello indicatori formativi

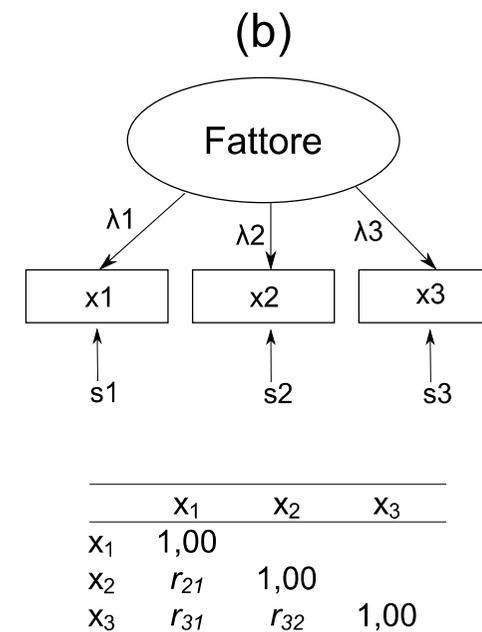
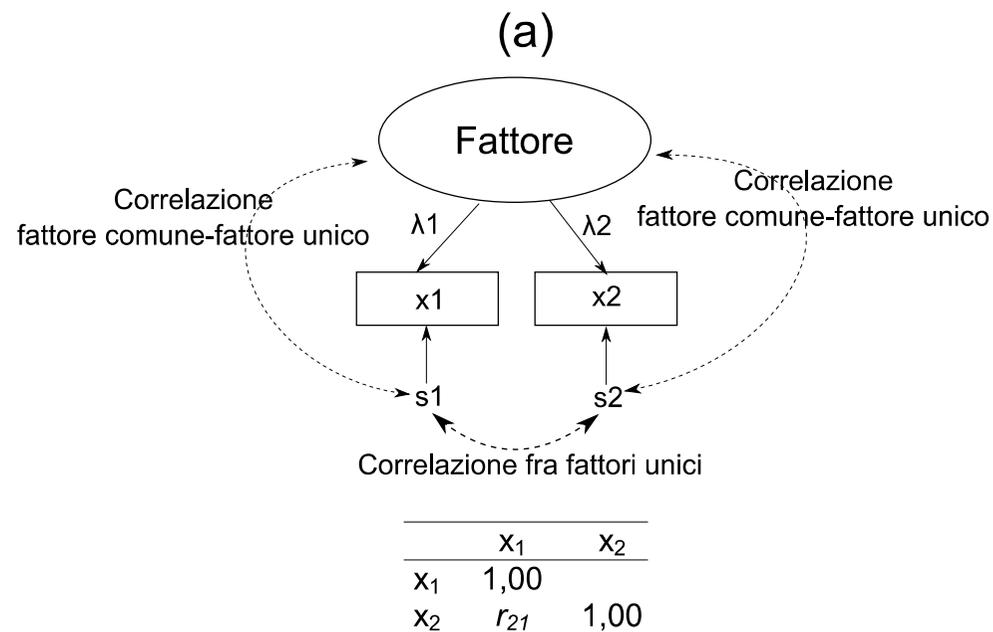
$$C_{jk} = \sum_{j=1}^p a_{ij} x_{ik} = a_{1j} x_{1k} + a_{2j} x_{2k} + a_{3j} x_{3k} + \dots + a_{pk} x_{nk}$$



	x ₁	x ₂	x ₃
x ₁	1,00		
x ₂	r ₂₁	1,00	
x ₃	r ₃₁	r ₃₂	1,00

Identificazione del modello

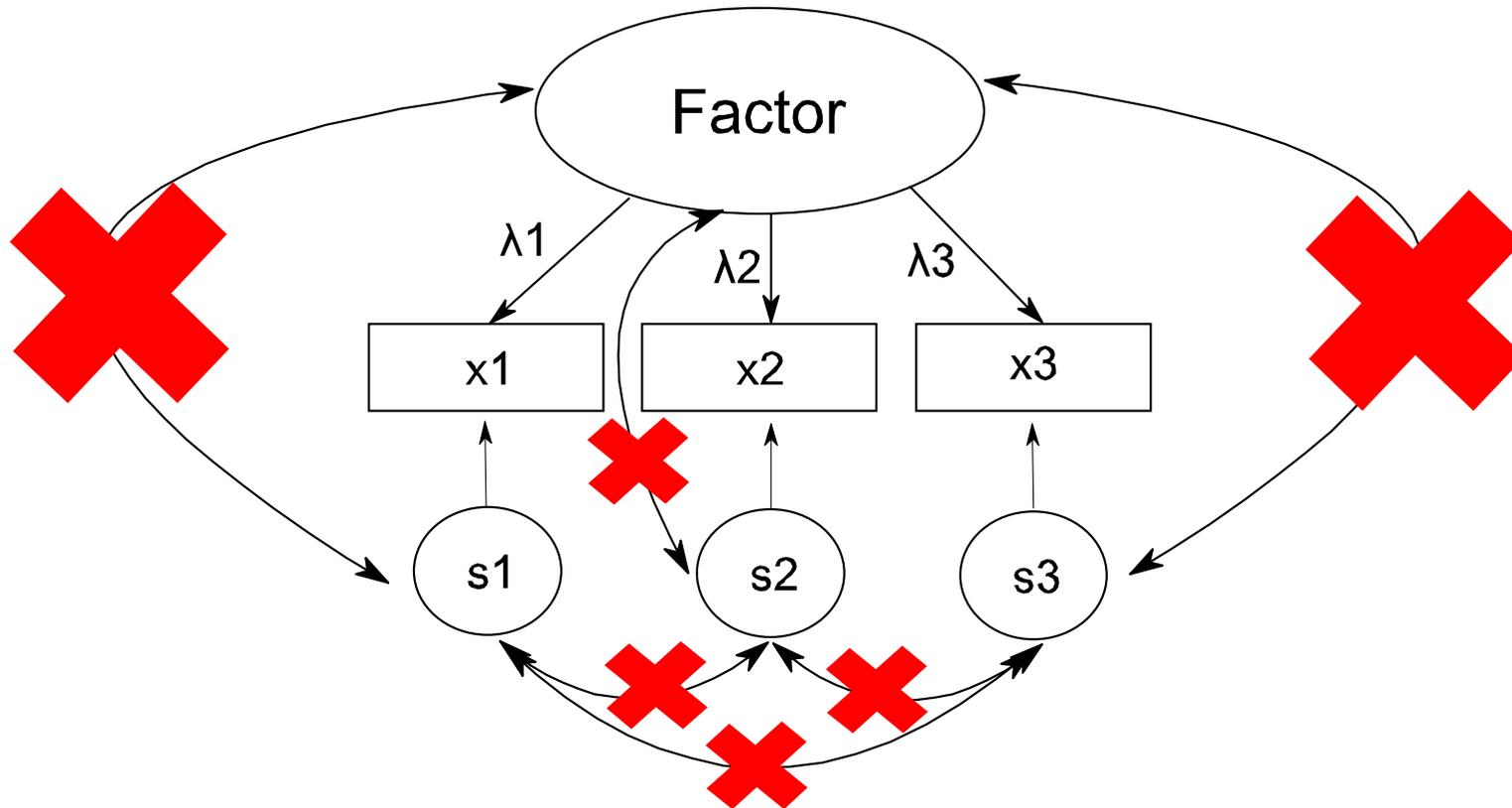
- Esprimere le relazioni ipotizzate fra le variabili latenti e manifeste del modello in funzione dell'informazione disponibile



Assunzioni statistiche analisi fattoriale

1. L'unica informazione nota è la varianza delle variabili osservate e le correlazioni fra queste
2. La correlazione fra i Fattori Comuni e quelli Unici è uguale a zero
3. La correlazione fra i Fattori Unici è uguale a zero

$$x_i = \lambda_i F + s_i$$



Fasi analisi fattoriale

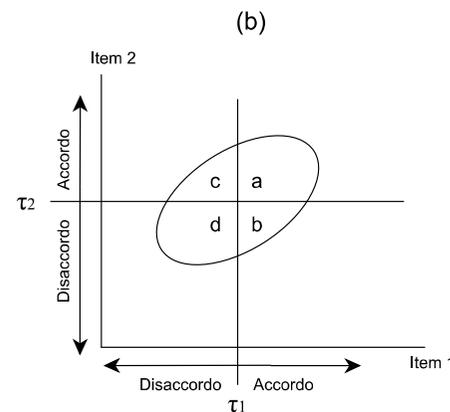
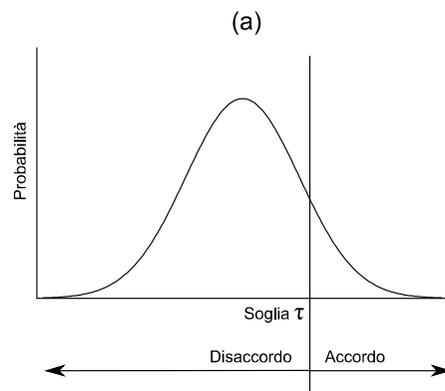
1. Pianificazione della ricerca e raccolta dei dati
2. Esame delle caratteristiche psicometriche delle variabili e degli outliers
3. Calcolo della matrice di correlazione ed esame della sua fattorizzabilità
4. Estrazione dei fattori
5. Rotazione dei fattori
6. Interpretazione dei fattori

Pianificazione della ricerca e raccolta dei dati

- Solida base teorica per la definizione del costrutto
- Aver seguito tutte le fasi precedenti di sviluppo del test
- Scala di risposta (dicotomica o Likert?)
- Numero di item: un costrutto ampio o varie facet?
- Numero di soggetti
- Raccolta dei dati nel modo adeguato

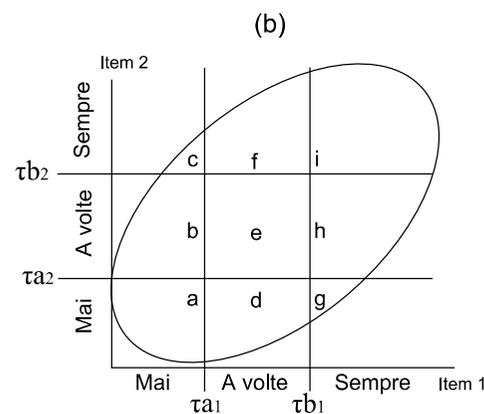
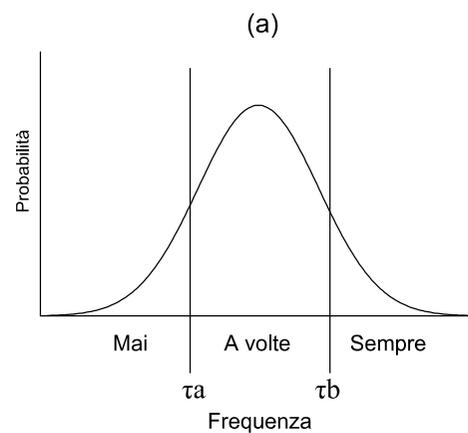
Correlazione tetracorica e policorica

- Underlying variable approach



(c)

		Item 2		
		Accordo	Disaccordo	
Item 1	Accordo	a	b	a+b
	Disaccordo	c	d	c+d
		a+c	b+d	1



(c)

		Item 2			
		Mai	A volte	Sempre	
Item 1	Mai	a	b	c	a+b+c
	A volte	d	e	f	d+e+f
	Sempre	g	h	i	g+h+i
		a+d+g	b+e+h	c+f+i	1

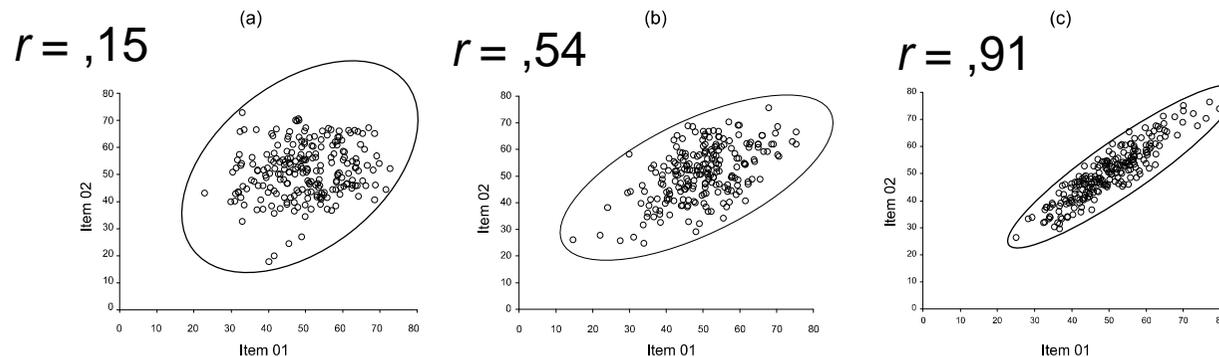
Quale matrice?

- item tutti dicotomici → matrice di correlazione tetracorica
- item Likert, molti con valori di skewness e curtosi *fuori* dal range $[-1;+1]$ → matrice di correlazione policorica;
- item Likert, molti con valori di skewness e curtosi *all'interno* del range $[-1;+1]$ → matrice di correlazione di Pearson

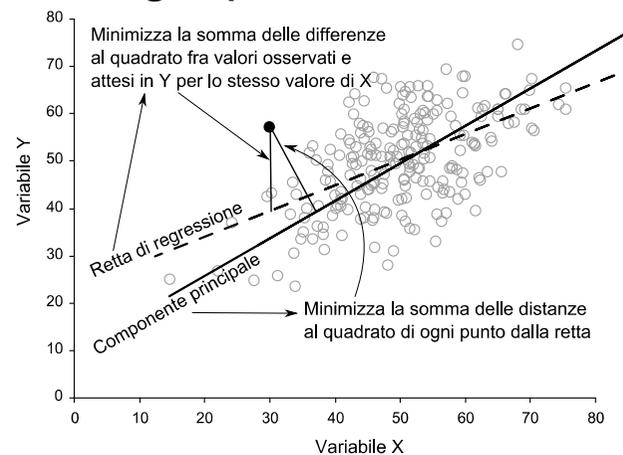
Estrazione dei fattori

Analisi componenti principali

- Partenza: matrice di correlazione



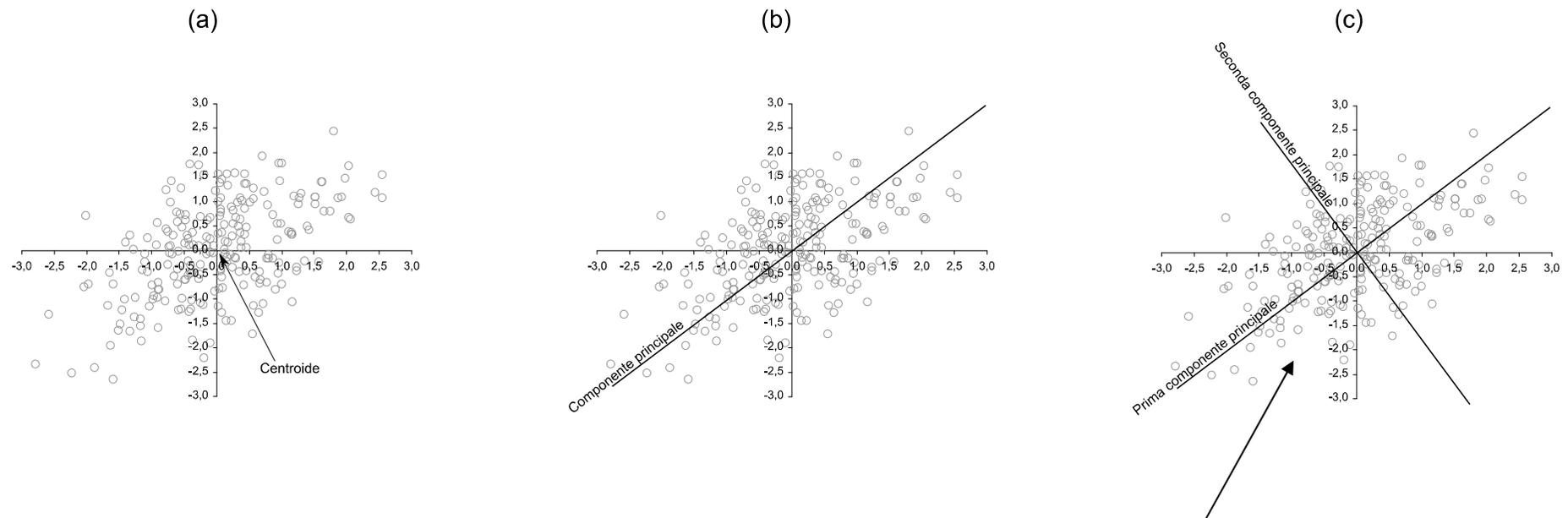
- Componente principale = asse che minimizza la somma delle distanze al quadrato di ogni punto



Estrazione dei fattori

Analisi componenti principali

- La correlazione standardizza ($M = 0 \pm 1$) le variabili \rightarrow centroide (0;0)



Le componenti principali creano un nuovo sistema di riferimento cartesiano

Estrazione dei fattori

Analisi componenti principali

- Saturazioni = coordinate cartesiane rispetto al nuovo sistema di riferimento cartesiano
- Matrice delle saturazioni = punto di arrivo dell'analisi

	Componente						h ²
	I	II	III	IV	V	IV	
Loquace	,64	-,51	,04	-,33	-,30	,35	1,00
Energico	,58	-,52	,26	,56	-,07	-,08	1,00
Socievole	,62	-,52	-,27	-,20	,41	-,25	1,00
Preciso	,55	,59	-,34	,25	,19	,36	1,00
Scrupoloso	,61	,56	-,20	-,05	-,39	-,35	1,00
Affidabile	,51	,57	,58	-,19	,22	-,02	1,00
Autovalore	2,07	1,79	,63	,56	,50	,45	
Proporzione di varianza spiegata	,34	,30	,10	,09	,08	,08	
Proporzione di varianza spiegata cumulata	,34	,64	,75	,84	,92	1,00	

Autovalori

Comunalità

	Componente						h ²
	I	II	III	IV	V	IV	
Loquace	,64	-,51	,04	-,33	-,30	,35	1,00
Energico	,58	-,52	,26	,56	-,07	-,08	1,00
Socievole	,62	-,52	-,27	-,20	,41	-,25	1,00
Preciso	,55	,59	-,34	,25	,19	,36	1,00
Scrupoloso	,61	,56	-,20	-,05	-,39	-,35	1,00
Affidabile	,51	,57	,58	-,19	,22	-,02	1,00
Autovalore	2,07	1,79	,63	,56	,50	,45	
Proporzione di varianza spiegata	,34	,30	,10	∞	∞	∞	
Proporzione di varianza spiegata cumulata	,34	,64	,75				

Somma dei quadrati

Gli autovalori decrescono

$$\text{Varianza spiegata} = \frac{\text{Autovalore}}{\text{Numero di item}}$$

Non siamo obbligati ad estrarre tutte le possibili componenti!

	Componente		h ²
	I	II	
Loquace	,64	-,51	,67
Energico	,58	-,52	,61
Socievole	,62	-,52	,66
Preciso	,55	,59	,66
Scrupoloso	,61	,56	,68
Affidabile	,51	,57	,58
Autovalore	2,07	1,79	
Proporzione di varianza spiegata	,34	,30	
Proporzione di varianza spiegata cumulata	,34	,64	

Varianza comune

- Matrice di correlazione "aggiustata" o "ridotta"

	1	2	3	4	5	6
1. Loquace	,34	,46	,51	,02	,10	,05
2. Energico	,46	,27	,44	,02	,04	,03
3. Socievole	,51	,44	,32	,06	,08	,00
4. Preciso	,02	,02	,06	,31	,52	,41
5. Scrupoloso	,10	,04	,08	,52	,34	,44
6. Affidabile	,05	,03	,00	,41	,44	,24

Stime di comunalità iniziale

Come si stimano comunalità iniziali?

- Metodo "classico"
 - valore assoluto della correlazione più alta della variabile con le altre → il quadrato della correlazione maggiore della variabile con tutte le altre è il valore minimo di quantità di varianza spiegabile da un fattore definito dalle due variabili in questione
- Metodo "comune"
 - Correlazione multipla al quadrato
- Metodo "iterativo"
 - Generazione casuale comunalità iniziali → estrazione fattori e calcolo comunalità di estrazione che vengono impiegate come comunalità iniziali al ciclo successivo

Principal Axis Factoring

	Fattore					h ²
	I	II	III	IV	V	
Loquace	,58	-,45	,07	-,14	-,01	,56
Energico	,50	-,42	,11	,17	-,06	,47
Socievole	,56	-,44	-,16	-,01	,07	,54
Preciso	,48	,52	-,14	,10	,01	,53
Scrupoloso	,54	,51	-,04	-,09	-,09	,57
Affidabile	,41	,46	,21	,00	,08	,43
Autovalore	1,59	1,31	,11	,07	,02	
Proporzione di varianza spiegata	,26	,22	,02	,01	,00	
Proporzione di varianza spiegata cumulata	,26	,48	,50	,51	,52	

	Componente						h ²
	I	II	III	IV	V	IV	
Loquace	,64	-,51	,04	-,33	-,30	,35	1,00
Energico	,58	-,52	,26	,56	-,07	-,08	1,00
Socievole	,62	-,52	-,27	-,20	,41	-,25	1,00
Preciso	,55	,59	-,34	,25	,19	,36	1,00
Scrupoloso	,61	,56	-,20	-,05	-,39	-,35	1,00
Affidabile	,51	,57	,58	-,19	,22	-,02	1,00
Autovalore	2,07	1,79	,63	,56	,50	,45	
Proporzione di varianza spiegata	,34	,30	,10	,09	,08	,08	
Proporzione di varianza spiegata cumulata	,34	,64	,75	,84	,92	1,00	

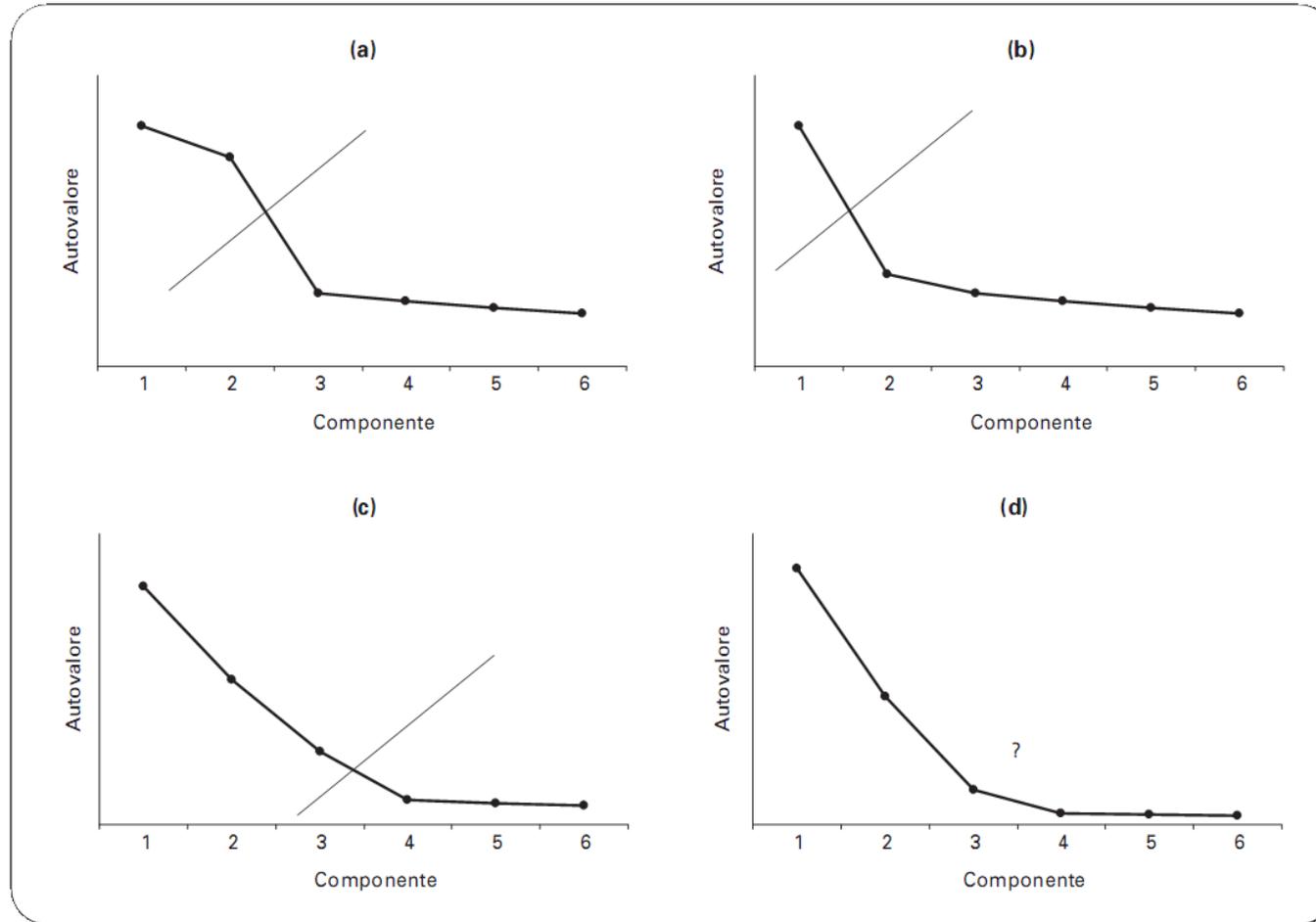
Metodi di estrazione dei fattori

- **Principal Axis Factoring (PAF)**
- **Maximum Likelihood (ML)**
 - Soluzione fattoriale che meglio riproduce la matrice di correlazione osservata stimando i valori delle saturazioni della popolazione che hanno la massima probabilità (= verosimiglianza) di produrre la matrice delle correlazioni campionaria
 - Indici di fit e test di significatività
 - Assunzione normalità multivariata
- **Unweighted Least Squares (ULS)**
 - *Distribution Free*
- **Weighted Least Squares Mean and Variance adjusted (WLSMV)**
 - Variabili dicotomiche / politomiche non normali

Quanti fattori estrarre?

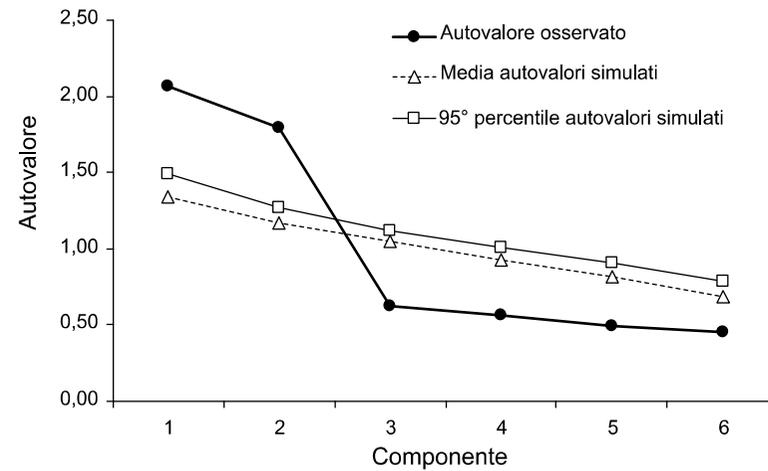
- Varianza spiegata minima (raro)
- Comunalità di estrazione minima (rarissimo)
- Kaiser-Guttman (obsoleto, porta ad estrarre troppi fattori)
- Scree-test (il migliore, ma soggettivo)
- Parallel Analysis (PA)
- Minimum Average Partial Correlation Statistic (MAP)
- Altri metodi (Hull)

Scree-plot

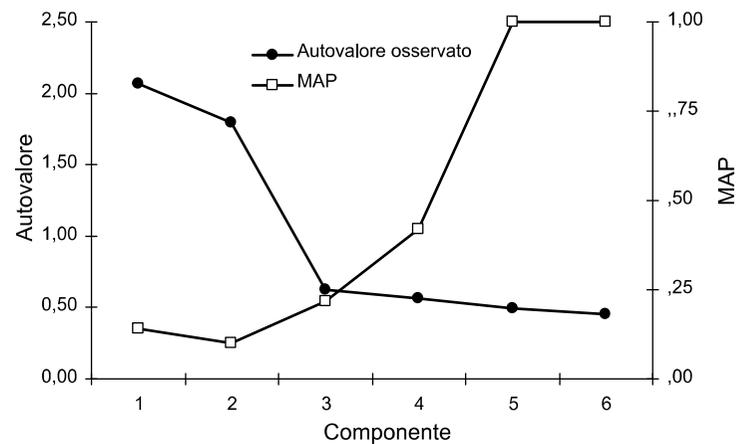


PA e MAP

Componente	Autovalore osservato	Media autovalori simulati	95° percentile autovalori simulati
1	2,07	1,34	1,49
2	1,79	1,17	1,27
3	0,63	1,04	1,12
4	0,56	0,93	1,01
5	0,50	0,82	0,90
6	0,45	0,69	0,79



Componente	Autovalore osservato	MAP
1	2,07	0,13957
2	1,79	0,10180
3	0,63	0,21619
4	0,56	0,41795
5	0,50	0,99999
6	0,45	1,00000



ACP vs AF

ACP

Non ha un modello statistico sottostante

per le variabili osservate, e si concentra sulla spiegazione della variazione totale nelle variabili osservate sulla base delle proprietà di massima varianza delle componenti principali

Modelli di misurazione formativi

Scompone la matrice di correlazione R , con 1.0 sulla diagonale della matrice

Cerca di spiegare la massima quota possibile di **varianza totale**

AF

Possiede un modello statistico sottostante

che ripartisce la varianza totale in varianza COMUNE + UNICA e si concentra sulla *varianza comune*

Modelli di misurazione riflessivi

Scompone la matrice di correlazione *adjusted* R' , con le stime della comunaltà sulla diagonale della matrice

Cerca di spiegare la massima quota possibile di **varianza comune**

ACP vs AF

ACP

Poiché non è altro che una riformulazione delle variabili mediante diversa combinazione lineare, **ha una soluzione unica**

Non implica l'ottenere variabili latenti, poiché le componenti sono trasformazioni lineari delle variabili originarie, ossia una sorta di *sommario empirico*

Permette di ottenere un punteggio nella variabile sintetica univocamente determinato

Saturazioni “gonfiate” e correlazioni fra i fattori “attenuate” dalla presenza della varianza d'errore

PAF

Non ha soluzione unica poiché cerca di spiegare la covarianza fra le variabili e l'indeterminatezza è legata ai diversi metodi di stima delle comunalità iniziali

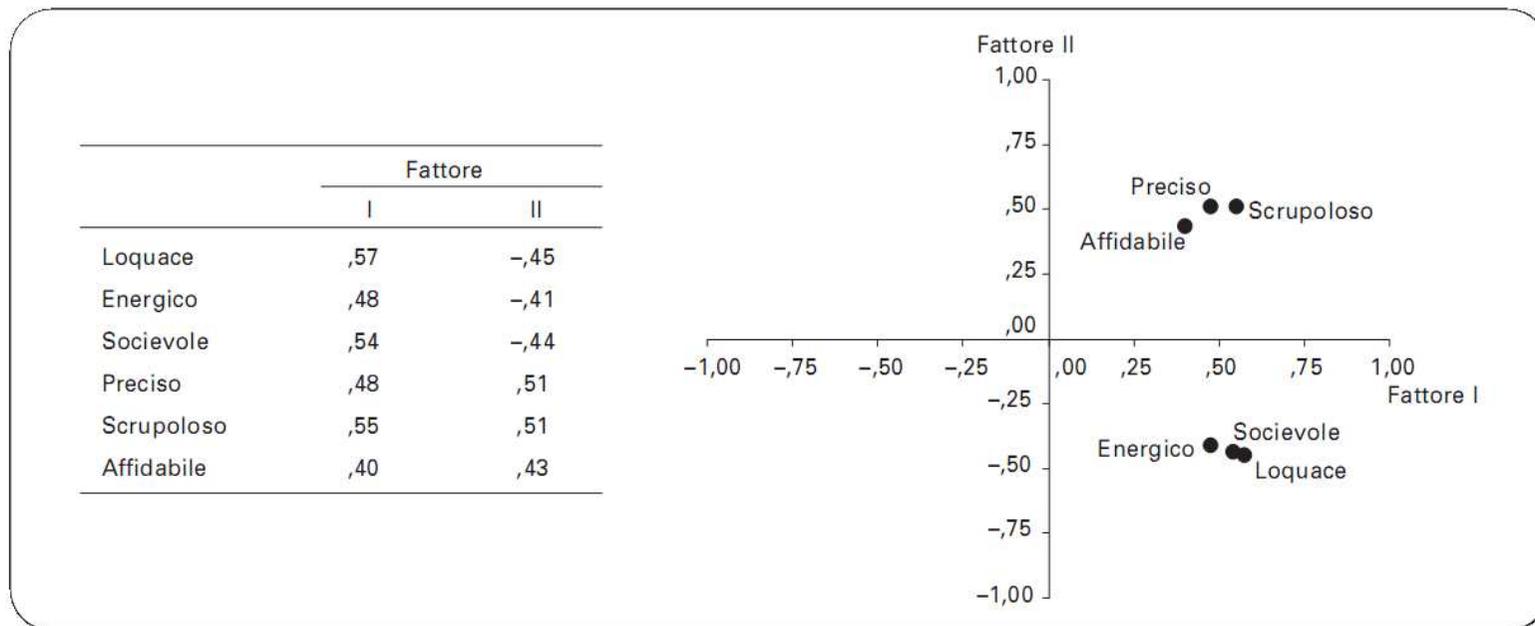
Implica l'ottenere variabili latenti, poiché i fattori sono variabili sintetiche che condensano la variabilità comune fra le variabili originarie

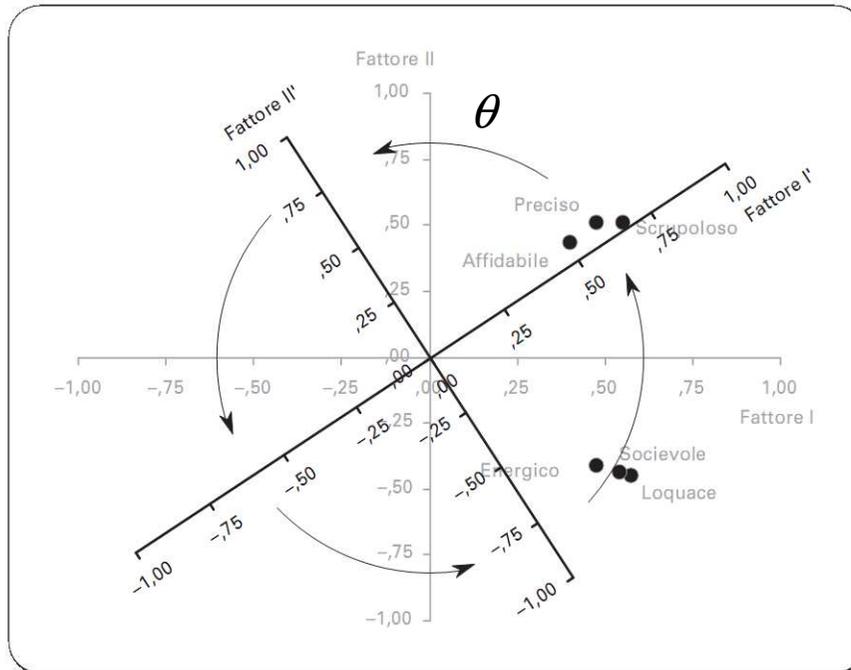
Indeterminatezza dei punteggi fattoriali

Saturazioni e correlazioni fra fattori non affette da varianza d'errore e quindi più vicine ai valori della popolazione

Rotazione dei fattori

- Problema di prospettiva



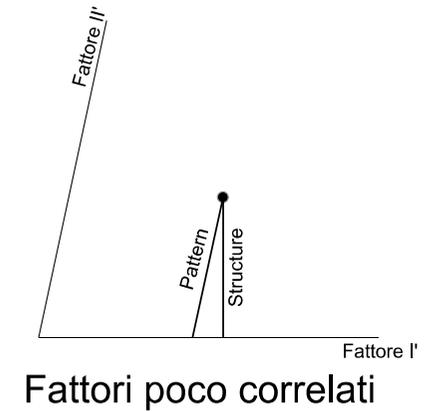
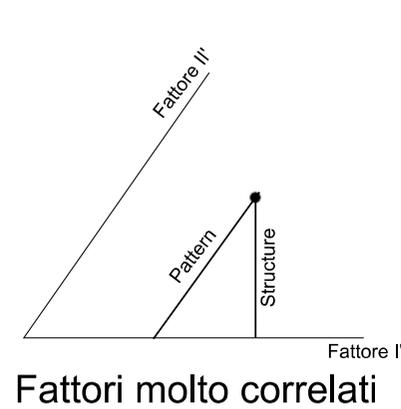
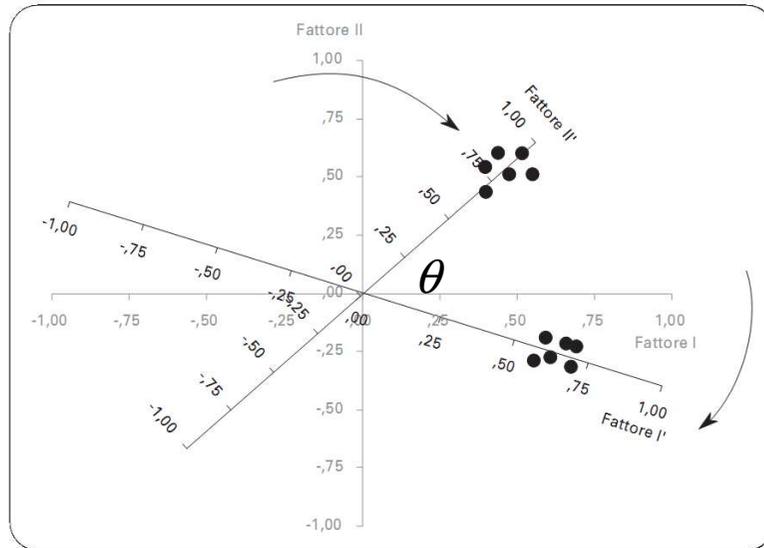


$$r_{jI'} = (\cos \theta)(r_{jI}) + (\sin \theta)(r_{jII})$$

$$r_{jII'} = (-\sin \theta)(r_{jI}) + (\cos \theta)(r_{jII})$$

	Soluzione non ruotata			Soluzione ruotata		
	Fattore			Fattore		
	I	II	h ²	I'	II'	h ²
Loquace	,57	-,45	,53	,12	-,72	,53
Energico	,48	-,41	,40	,08	-,62	,40
Socievole	,54	-,44	,49	,11	-,69	,49
Preciso	,48	,51	,48	,69	,06	,48
Scrupoloso	,55	,51	,56	,75	,01	,56
Affidabile	,40	,43	,35	,59	,05	,35
Autovalore	1,55	1,27		1,42	1,39	

- Ridistribuzione della varianza spiegata
- Fattori indipendenti o correlati?



	Pattern		Structure		h ²
	Fattore I	Fattore II	Fattore I'	Fattore II'	
Loquace	,73	,01	,73	,09	,53
Energico	,63	-,02	,63	,04	,40
Socievole	,70	,00	,70	,07	,49
Preciso	-,02	,70	,05	,69	,48
Scrupoloso	,03	,75	,11	,75	,56
Affidabile	-,02	,59	,04	,59	,35
<i>Autovalore</i>	<i>1,42</i>	<i>1,39</i>	<i>1,43</i>	<i>1,41</i>	

Correlazione fra i fattori = $\cos(\theta)$

Pattern vs structure matrix

- Coefficienti hanno significati diversi
- Pattern → coefficienti di regressione (possono essere $>|1.00|$)
- Structure → coefficienti di correlazione (non possono essere $>|1.00|$)
- Se rotazione ortogonale o correlazione fra i fattori = 0, le due matrici coincidono

Metodi analitici di rotazione

- Ortogonali
 - VARIMAX
 - QUARTIMAX
 - EQUAMAX

- Obliqui
 - PROMAX
 - DIRECT OBLIMIN
 - GEOMIN

Interpretazione dei fattori

- Almeno 3 item per fattore
- Replicabilità della soluzione fattoriale
- Gamma ridotta delle saturazioni (e.g., ,40-,60)
- Plausibilità teorica del fattore comune
- Evitare forzature

Modelli di strutture di covarianza

Analisi fattoriale confermativa
e modelli di equazioni strutturali

Per approfondimenti: http://www.ateneonline.it/chiorri/docenti2/approfondimenti/isbn6404-5_Approfondimento_5-4.pdf

Analisi fattoriale esplorativa vs confermativa

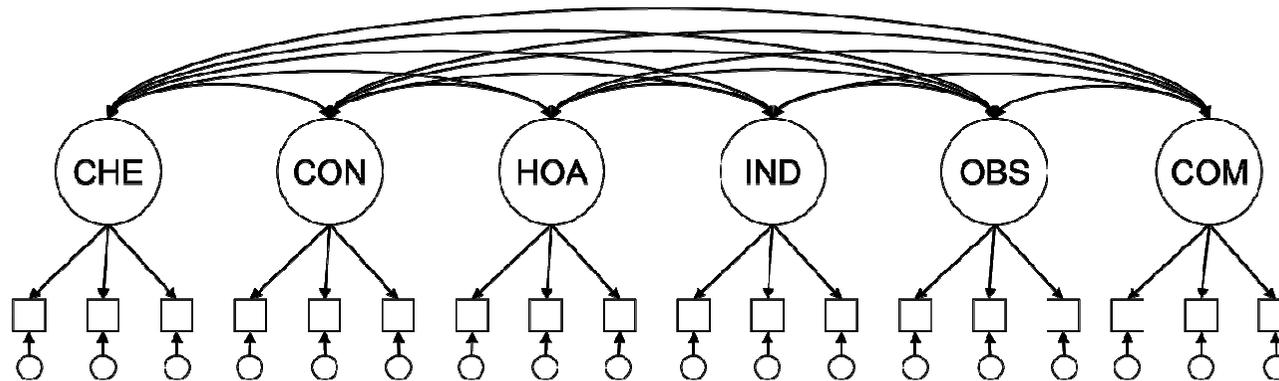
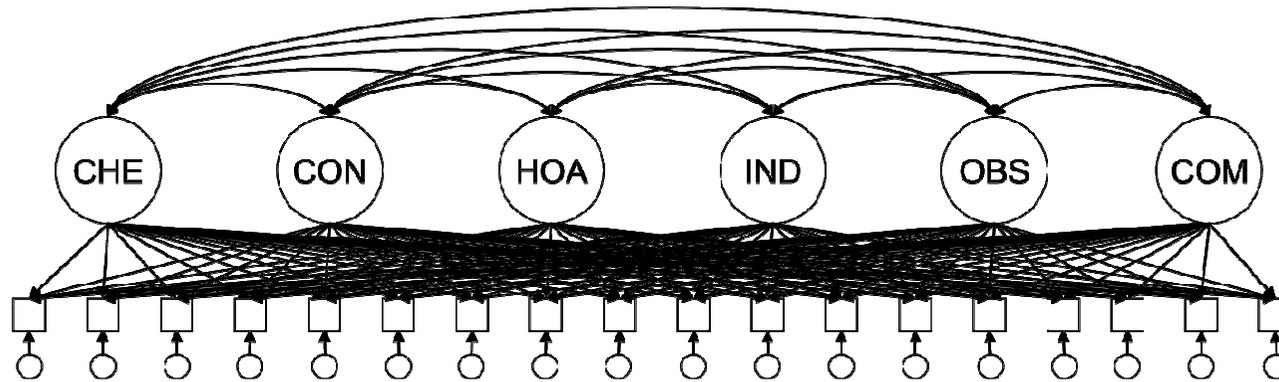
- In the role of an exploratory method, factor analysis has the peculiarity, among scientific investigative tools [...] that it can be profitably used with relatively little regard to prior formulation of a hypothesis

Cattell, R. B. (1952). *Factor Analysis*. New York: Harper & Bros, p.21

- The factorial study of human behavior might best be conceived as a program of studies rather than in terms of isolated, separate studies. Each study should build upon the knowledge gained from previous studies and add further to the verified fund of knowledge. Early studies in some domain, or class of behavior, will be more exploratory in nature and be made with less perfected batteries of measures. As knowledge increases concerning the interrelations of the various behaviors in such a domain, it should be possible to construct more satisfactory batteries for factorial analysis. Confirmatory studies should aid in firmly establishing the factorial structure.

Tucker, L. R. (1955). The objective definition of simple structure in linear factor analysis. *Psychometrika*, 20, 209-225, p. 210

EFA vs CFA



Note: CHE=Checking, CON=Contamination Fears, HOA=Hoarding, IND=Indecisiveness, OBS=Obsessions; COM=Compulsions

Analisi fattoriale confermativa (AFC)

- Cross-validation
- Metodo multi-gruppo per l'analisi fattoriale → deve essere noto in anticipo il numero dei fattori
- Rotazione verso matrice target (Procrustea) → deve essere nota in anticipo la struttura della matrice delle saturazioni
- Analisi strutture di covarianza

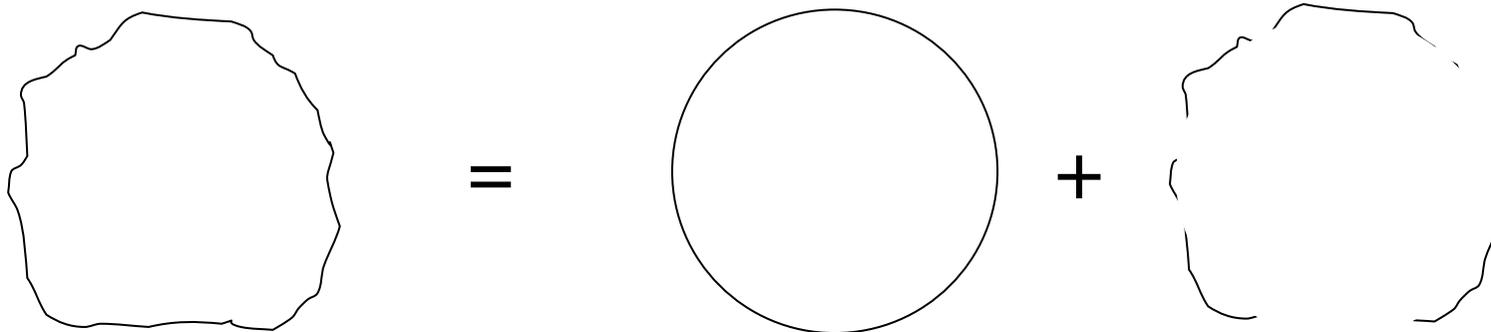
CFA oggi

- Bock, R. D., & Bargmann, R. (1966). Analysis of covariance structures. *Psychometrika* , 31, 443-449.
- Jöreskog, K. G. (1969). A general approach to confirmatory maximum likelihood factor analysis. *Psychometrika*, 34, 183-202.



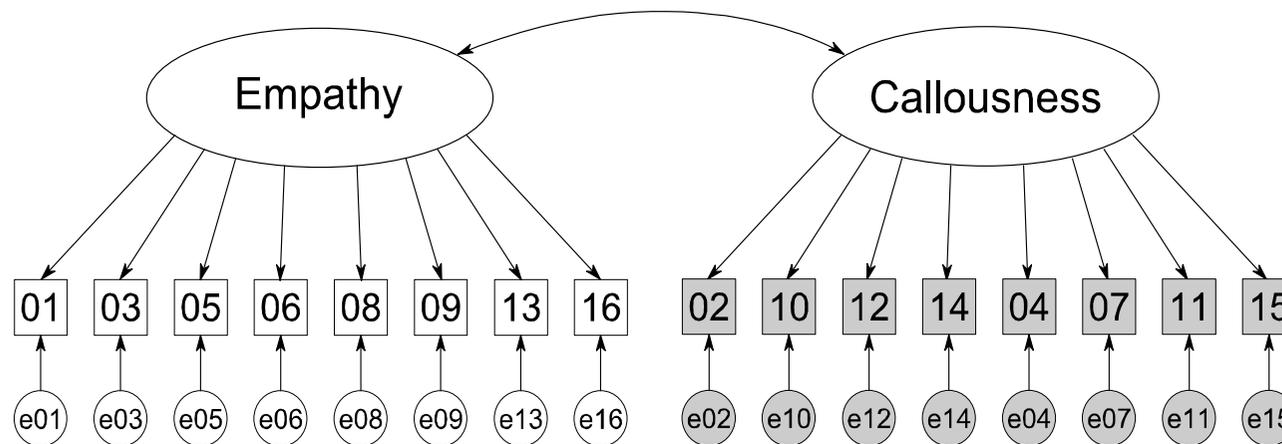
Logica della CFA

- Si parte da un modello teorico che definisce le relazioni fra tutte le variabili del modello (e.g., modello di misurazione)
- Si verifica il modello coi dati osservati (matrice di covarianza osservata)
- Si valuta l'adattamento (*fit*) del modello, ossia quanto bene la matrice di covarianza stimata coi dati a disposizione in base all'assunzione che il modello teorico da cui è stato specificato il modello sia vero riproduce la matrice di covarianza osservata



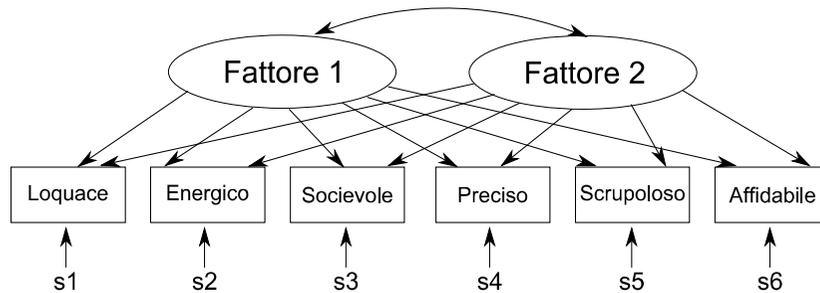
Specificazione del modello

- Su quale/i fattore/i saturano gli item?
- I fattori sono correlati? Se sì quali?
- Etc.



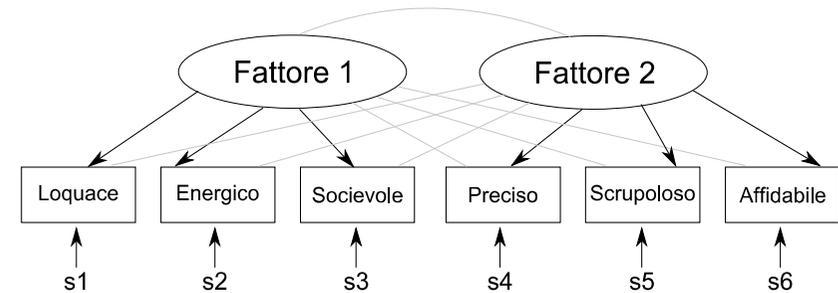
AFC vs AFE

Modello di
analisi fattoriale esplorativa



Item	I	II
Loquace	,73	,01
Energico	,63	-,02
Socievole	,70	,00
Preciso	-,02	,70
Scrupoloso	,03	,74
Affidabile	-,02	,59
Correlazione		,10

Modello di
analisi fattoriale confermativa



Item	I	II
Loquace	,73	,00
Energico	,63	,00
Socievole	,70	,00
Preciso	,00	,70
Scrupoloso	,00	,74
Affidabile	,00	,59
Correlazione		,00

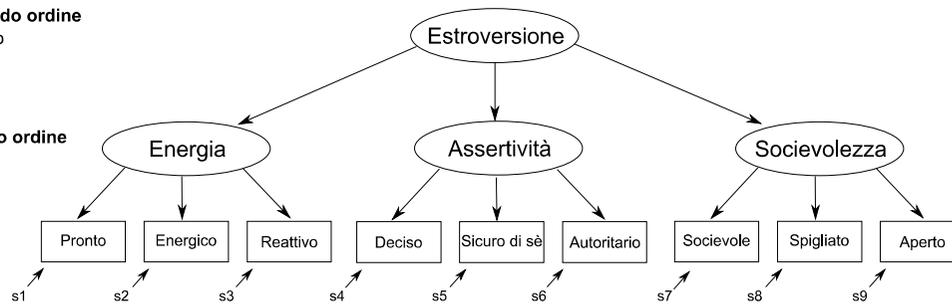
Flessibilità dei modelli struttura di covarianza

(a)

Modello di
analisi fattoriale di secondo ordine

Fattore di secondo ordine
Dominio

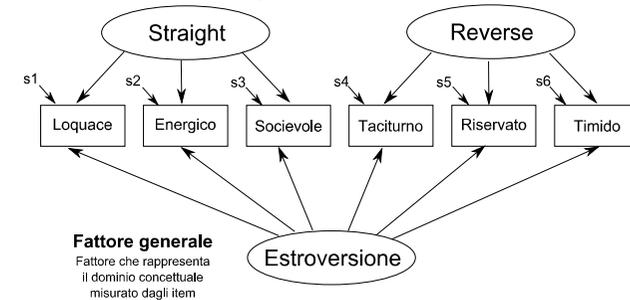
Fattori di primo ordine
Facets



(b)

Modello bifattoriale

Fattori di gruppo
Fattori specifici di alcuni item,
in questo caso legati al fatto
che gli item siano reverse o straight



Fattore generale
Fattore che rappresenta
il dominio concettuale
misurato dagli item

Identificazione del modello

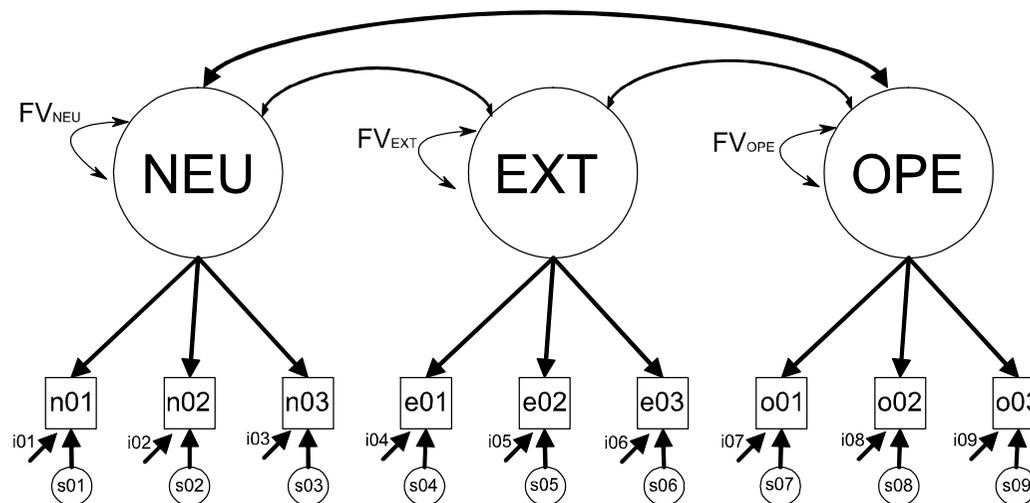
- Il modello di CFA si dice identificato se tutti i suoi parametri sono funzioni indipendenti degli elementi della matrice di covarianza
- L'informazione disponibile è nella matrice di covarianza che contiene $p = k(k+1)/2$ elementi, e $p - 1$ è il maggior numero di parametri che possono essere specificati

	Fearful	Worrying	Anxious	Talkative	Energetic	Sociable	Curious	Artistic	Imaginative
Fearful	0.651	0.217	0.168	-0.047	-0.035	0.107	0.051	0.039	0.042
Worrying	0.217	0.653	0.213	0.045	0.025	0.135	0.134	0.070	0.091
Anxious	0.168	0.213	0.500	-0.020	0.024	0.074	-0.060	-0.063	0.008
Talkative	-0.047	0.045	-0.020	0.257	0.064	0.179	0.141	0.099	0.071
Energetic	-0.035	0.025	0.024	0.064	0.251	0.095	0.013	0.052	0.007
Sociable	0.107	0.135	0.074	0.179	0.095	0.713	0.264	0.099	0.087
Curious	0.051	0.134	-0.060	0.141	0.013	0.264	0.745	0.321	0.408
Artistic	0.039	0.070	-0.063	0.099	0.052	0.099	0.321	0.527	0.199
Imaginative	0.042	0.091	0.008	0.071	0.007	0.087	0.408	0.199	0.633

$$\frac{9 \times (9 + 1)}{2} = 45$$

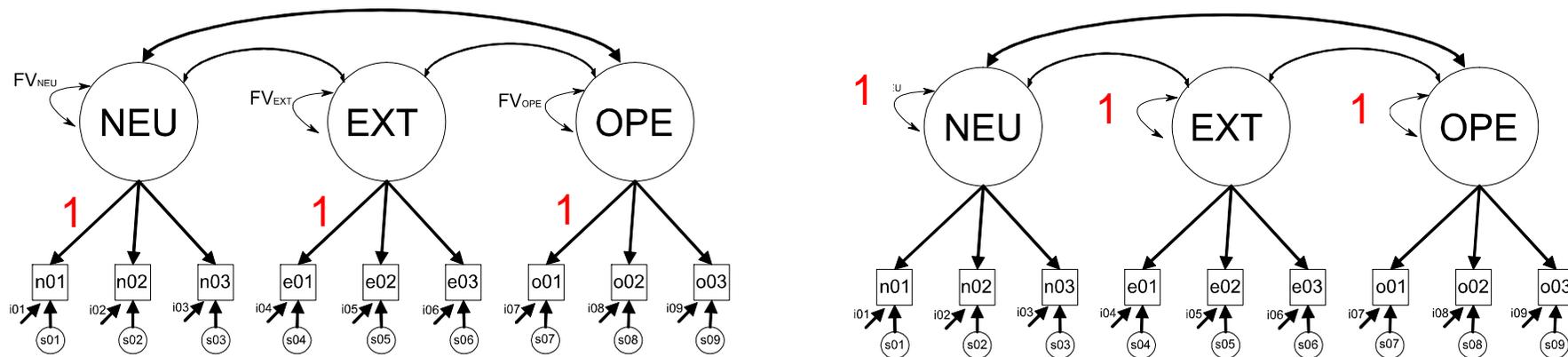
Identificazione del modello

- Quali sono i parametri del modello?
 - Saturazioni degli item (Factor loadings)
 - Varianze di errore degli item (Item error variances)
 - Intercette degli item (Item intercepts)
 - Correlazioni fra i fattori (Factor correlations)
 - Varianze dei fattori (Factor variances)



Identificazione del modello

- Le variabili latenti non hanno una metrica (unità di misura)!
- Fissare una saturazione @1 per ogni variabile latente, e stimare la varianza della variabile latente → la variabile latente prende la stessa metrica dell'item
- Fissare @1 la varianza della variabile latente e stimare tutte le saturazioni degli item → la variabile latente avrà media 0 e varianza 1



Identificazione del modello

- **Three Measure Rule:** un modello è identificato se ogni variabile latente ha almeno tre indicatori
- **Two Measure Rule:** un modello è identificato se ogni variabile latente ha almeno due indicatori e ogni costrutto è correlato con almeno un altro costrutto.
- Altri vincoli per l'identificazione

Stima del modello

- Il parametri del modello vengono stimati minimizzando la discrepanza fra la matrice di covarianza osservata (S) e quella implicata dal modello ($\hat{\Sigma}$)
- “Classiche” funzioni di discrepanza (ML, GLS, and ULS)
- Asymptotically Distribution Free Discrepancy Functions (ADF/WLS)

S

	Fearful	Worrying	Anxious	Talkative	Energetic	Sociable	Curious	Artistic	Imaginative
Fearful	0.651								
Worrying	0.217	0.653							
Anxious	0.168	0.213	0.500						
Talkative	-0.047	0.045	-0.020	0.257					
Energetic	-0.035	0.025	0.024	0.064	0.251				
Sociable	0.107	0.135	0.074	0.179	0.095	0.713			
Curious	0.051	0.134	-0.060	0.141	0.013	0.264	0.745		
Artistic	0.039	0.070	-0.063	0.099	0.052	0.099	0.321	0.527	
Imaginative	0.042	0.091	0.008	0.071	0.007	0.087	0.408	0.199	0.633

 $\hat{\Sigma}$

Fearful	0.646								
Worrying	0.227	0.649							
Anxious	0.138	0.215	0.497						
Talkative	0.031	0.048	0.029	0.255					
Energetic	0.015	0.024	0.014	0.045	0.249				
Sociable	0.063	0.099	0.060	0.185	0.092	0.708			
Curious	0.052	0.081	0.049	0.124	0.062	0.258	0.740		
Artistic	0.023	0.037	0.022	0.056	0.028	0.116	0.320	0.524	
Imaginative	0.030	0.046	0.028	0.071	0.035	0.147	0.405	0.183	0.629

Valutazione del fit del modello

- **Test del chi-quadrato**
- Verifica l'ipotesi nulla che i dati osservati siano stati generati dal modello teorico

$$F_{ML}(S, \hat{\Sigma}) = \text{tr}(S\hat{\Sigma}^{-1}) + \log |\hat{\Sigma}| - \log |S| - q$$

$$\chi^2 = (n-1)F_{ML}$$

gdl = numero di parametri del modello specificato – numero di parametri del modello nullo

- Troppo sensibile all'ampiezza campionaria

Valutazione del fit del modello

- **Fit assoluto**
- *Root Mean Square Error of Approximation (RMSEA)*
- Misura della discrepanza fra il modello specificato, con stime dei parametri ottimali, e la matrice di covarianza della popolazione
- Misura standardizzata non dipendente dalla scala di misura delle variabili osservate
- Tende ad essere distorta in campioni più piccoli, ma “premia” la parsimonia del modello
- Valori accettabili: $RMSEA < .08$, Valori ottimali: $RMSEA < .05$

$$RMSEA = \sqrt{\frac{\frac{X^2_{\text{model}}}{df_{\text{model}}} - 1}{n - 1}} = \sqrt{\frac{\frac{42.393}{24} - 1}{150 - 1}} = .072$$

Valutazione del fit del modello

- **Fit relativo** → confronto del fit modello in esame con il fit di uno modello “nullo” che assume che tra le variabili osservate non vi sia covarianza
- *Comparative Fit Index* (CFI, Bentler, 1990): esamina la discrepanza fra i dati e il modello ipotizzato
- Non risente dell'ampiezza campionaria, ma “paga una penalità” per i modelli con più parametri
- Valori accettabili: $CFI > .90$, Valori ottimali: $CFI > .95$

$$CFI = \frac{(X_{\text{null}}^2 - df_{\text{null}}) - (X_{\text{model}}^2 - df_{\text{model}})}{(X_{\text{null}}^2 - df_{\text{null}})}$$

$$CFI = \frac{(268.823 - 36) - (42.393 - 24)}{(268.823 - 36)} = .92$$

Valutazione del fit del modello

- *Tucker-Lewis Index* (TLI, Tucker & Lewis, 1973; or *Non Normed Fit Index*, NNFI; Bentler & Bonett, 1980): analizza la discrepanza fra il valore di chi-quadrato del modello ipotizzato e quello del modello nullo
- Non penalizza i modelli con più parametri
- Dipende dalla dimensione delle correlazioni, per cui set di dati poco correlati tendono a produrre TLI più bassi
- Valori accettabili: CFI > .90, Valori ottimali: CFI > .95

$$TLI = \frac{\frac{X_{null}^2}{gdl_{null}} - \frac{X_{model}^2}{gdl_{model}}}{\frac{X_{null}^2}{gdl_{null}} - 1}$$

$$TLI = \frac{\frac{268.823}{36} - \frac{42.393}{24}}{\frac{268.823}{36} - 1} = .88$$

Valutazione delle stime dei parametri

- Le stime dei parametri sono tutte statisticamente significative?
- I valori dei parametri sono nel range atteso? (e.g., le correlazioni fra i fattori sono $< |1.00|$)
- Indici di modifica: fonti di “misfit” → attenzione a specificare parametri campione-specifici

MODEL RESULTS

		Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
NEV	BY				
	N01	1.000	0.000	999.000	999.000
	N02	1.038	0.750	1.383	0.167
	N03	0.665	0.195	3.416	0.001
EST	BY				
	E01	1.000	0.000	999.000	999.000
	E02	0.524	0.086	6.061	0.000
	E03	0.907	0.130	6.987	0.000
APE	BY				
	O01	1.000	0.000	999.000	999.000
	O02	0.746	0.174	4.290	0.000
	O03	0.755	0.159	4.759	0.000
EST	WITH				
	NEV	0.051	0.035	1.462	0.144
APE	WITH				
	NEV	0.018	0.036	0.499	0.618
	EST	0.098	0.037	2.658	0.008
Intercepts					
	N01	2.817	0.045	62.574	0.000
	N02	2.871	0.046	61.946	0.000
	N03	3.106	0.046	67.016	0.000
	E01	3.194	0.046	68.706	0.000
	E02	3.131	0.044	71.655	0.000
	E03	3.071	0.046	66.770	0.000
	O01	2.949	0.046	64.365	0.000
	O02	2.837	0.045	62.588	0.000
	O03	2.914	0.044	65.919	0.000

Controllare che le correlazioni non siano maggiori di |1.00|!

Variances				
NEV	0.195	0.138	1.416	0.157
EST	0.487	0.084	5.783	0.000
APE	0.313	0.082	3.808	0.000
Residual Variances				
N01	0.514	0.138	3.717	0.000
N02	0.542	0.167	3.243	0.001
N03	0.665	0.074	8.945	0.000
E01	0.270	0.068	3.961	0.000
E02	0.535	0.045	11.970	0.000
E03	0.341	0.059	5.757	0.000
O01	0.421	0.076	5.568	0.000
O02	0.545	0.058	9.396	0.000
O03	0.505	0.053	9.539	0.000

Controllare che le varianze di errore degli item non siano maggiori di |1.00| o minori di zero!

Minimum M.I. value for printing the modification index 10.000

	M.I.	E.P.C.	Std E.P.C.	StdYX E.P.C.
--	------	--------	------------	--------------

BY Statements

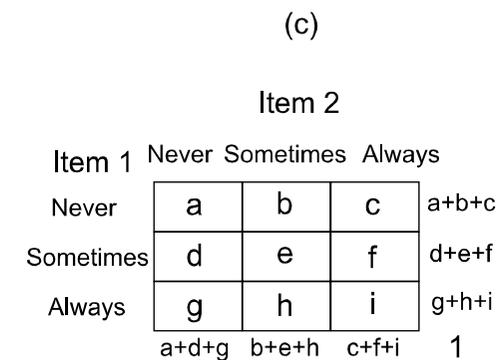
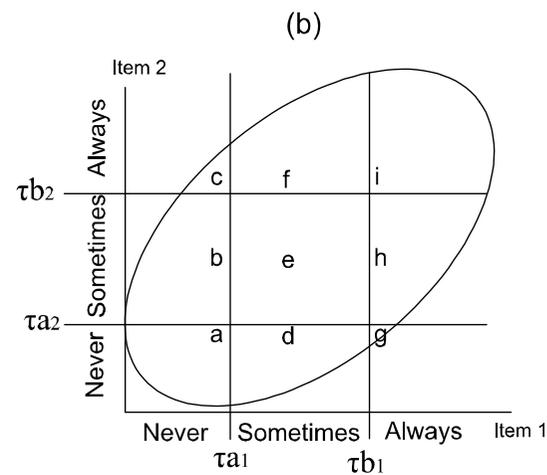
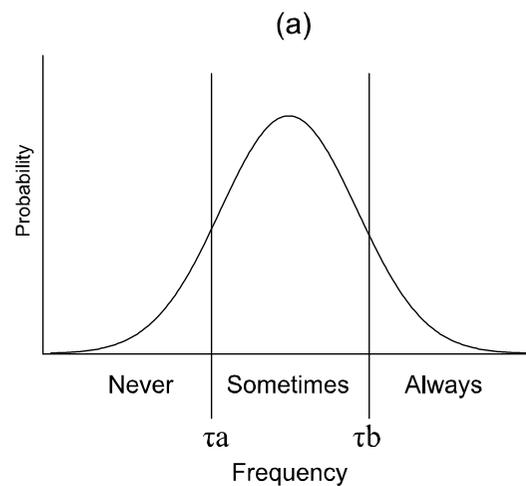
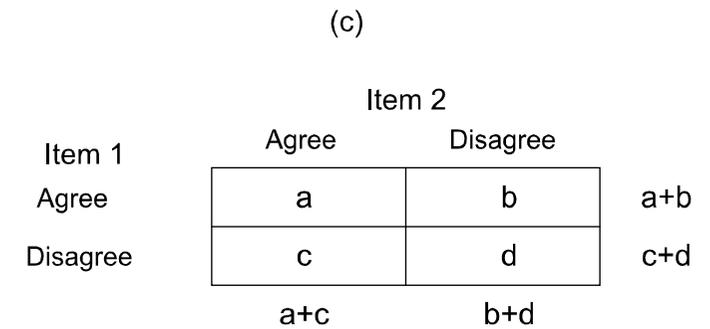
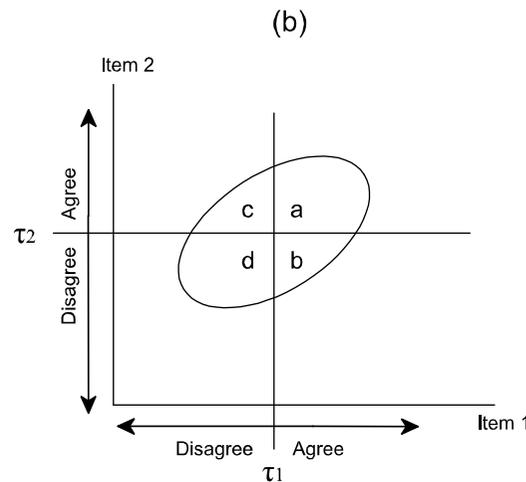
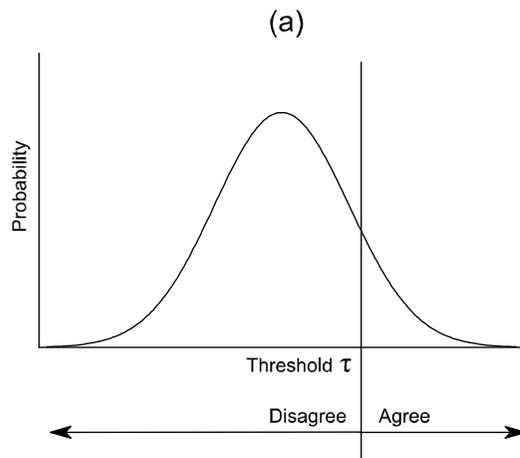
EST	BY N02	11.880	0.322	0.224	0.259
APE	BY N02	11.159	0.412	0.231	0.266
APE	BY N03	14.004	-0.412	-0.231	-0.266

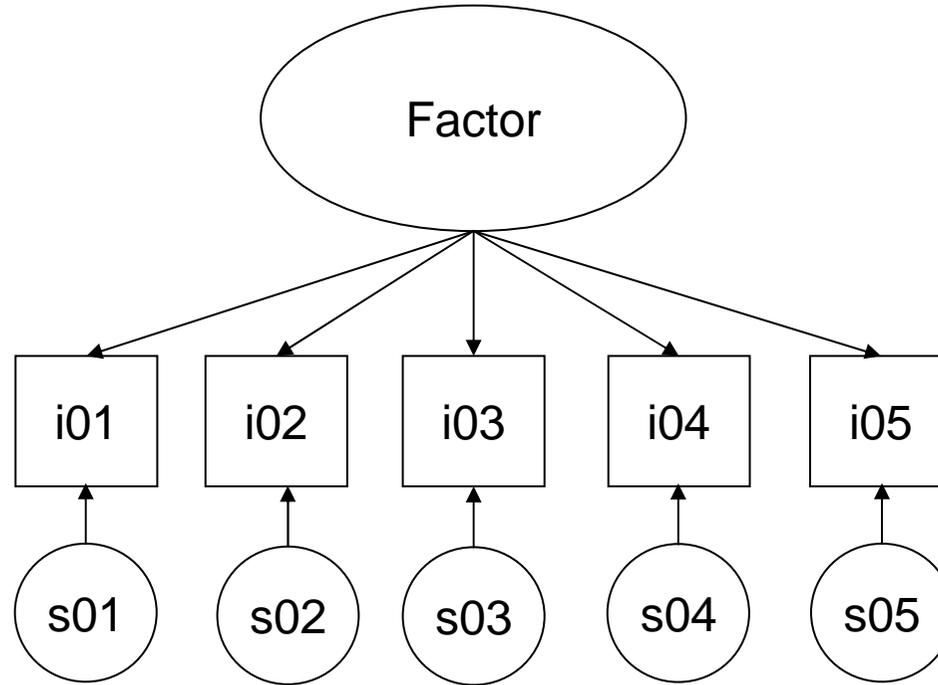
WITH Statements

N03	WITH N01	14.633	0.406	0.406	0.695
E01	WITH N03	10.144	-0.100	-0.100	-0.235
O03	WITH E03	10.278	-0.092	-0.092	-0.221

Variabili categoriali

- Underlying variable approach





CORRELATION MATRIX (WITH VARIANCES ON THE DIAGONAL)

Categorical
Polychoric correlation matrix

	I01	I02	I03	I04	I05
I01	1.000				
I02	0.440	1.000			
I03	0.503	0.264	1.000		
I04	0.494	0.722	0.353	1.000	
I05	0.464	0.682	0.324	0.561	1.000

Metric
Pearson correlation matrix

	I01	I02	I03	I04	I05
I01	1.000				
I02	0.299	1.000			
I03	0.319	0.141	1.000		
I04	0.345	0.585	0.154	1.000	
I05	0.339	0.537	0.181	0.435	1.000

Indici di fit con indicatori categoriali

MODEL FIT INFORMATION

Number of Free Parameters 23

Chi-Square Test of Model Fit

Value	30.461*
Degrees of Freedom	5
P-Value	0.0000

RMSEA (Root Mean Square Error Of Approximation)

Estimate	0.107
90 Percent C.I.	0.072
0.145	
Probability RMSEA <= .05	0.004

CFI/TLI

CFI	0.973
TLI	0.947

Chi-Square Test of Model Fit for the Baseline Model

Value	969.925
Degrees of Freedom	10
P-Value	0.0000

WRMR (Weighted Root Mean Square Residual)

Value	0.796
-------	-------

Indici di fit con indicatori metrici

MODEL FIT INFORMATION

Number of Free Parameters 15

Chi-Square Test of Model Fit

Value	24.609*
Degrees of Freedom	5
P-Value	0.0002
Scaling Correction Factor	1.661
for MLR	

RMSEA (Root Mean Square Error Of Approximation)

Estimate	0.094
90 Percent C.I.	0.059
0.132	
Probability RMSEA <= .05	0.021

CFI/TLI

CFI	0.918
TLI	0.836

Chi-Square Test of Model Fit for the Baseline Model

Value	249.866
Degrees of Freedom	10
P-Value	0.0000

SRMR (Standardized Root Mean Square Residual)

Value	0.052
-------	-------

Stime dei parametri

Indicatori categoriali

MODEL RESULTS

	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
F1				
BY				
I01	1.000	0.000	999.000	999.000
I02	1.448	0.126	11.532	0.000
I03	0.782	0.129	6.045	0.000
I04	1.354	0.113	11.996	0.000
I05	1.259	0.109	11.589	0.000
Thresholds				
I01\$1	0.600	0.063	9.458	0.000
I01\$2	1.355	0.084	16.099	0.000
I01\$3	1.860	0.117	15.906	0.000
I01\$4	2.366	0.184	12.840	0.000
I02\$1	0.184	0.060	3.080	0.002
I02\$2	1.238	0.079	15.608	0.000
I02\$3	1.829	0.114	16.015	0.000
I03\$1	1.288	0.081	15.843	0.000
I03\$2	1.892	0.120	15.778	0.000
I03\$3	2.366	0.184	12.840	0.000
I04\$1	0.384	0.061	6.289	0.000
I04\$2	1.370	0.085	16.144	0.000
I04\$3	1.927	0.123	15.629	0.000
I04\$4	2.612	0.241	10.833	0.000
I05\$1	0.634	0.064	9.917	0.000
I05\$2	1.745	0.107	16.253	0.000
I05\$3	2.282	0.169	13.475	0.000
I05\$4	2.612	0.241	10.833	0.000

Indicatori metrici

MODEL RESULTS

	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
F1				
BY				
I01	1.000	0.000	999.000	999.000
I02	1.729	0.333	5.196	0.000
I03	0.321	0.102	3.138	0.002
I04	1.518	0.239	6.361	0.000
I05	1.145	0.182	6.274	0.000
Intercepts				
I01	0.402	0.037	10.996	0.000
I02	0.569	0.037	15.537	0.000
I03	0.137	0.022	6.270	0.000
I04	0.467	0.035	13.241	0.000
I05	0.319	0.029	11.065	0.000

Indicatori categoriali

Variances				
F1	0.358	0.056	6.433	0.000
R-SQUARE				
Observed Variable	Estimate	Residual Variance		
I01	0.358	0.642		
I02	0.751	0.249		
I03	0.219	0.781		
I04	0.657	0.343		
I05	0.568	0.432		

Varianza residua della variabile latente continua sottostante la variabile categoriale osservata

Indicatori metrici

Variances				
F1	0.124	0.041	3.023	0.003
Residual Variances				
I01	0.472	0.073	6.430	0.000
I02	0.226	0.041	5.524	0.000
I03	0.200	0.040	4.955	0.000
I04	0.269	0.040	6.693	0.000
I05	0.208	0.028	7.355	0.000

Varianza residua della variabile continua osservata

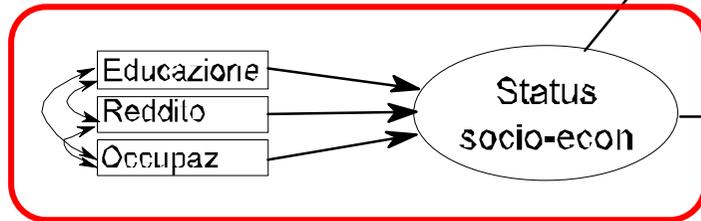
Modelli di equazioni strutturali

Estensione dei modelli di CFA con:

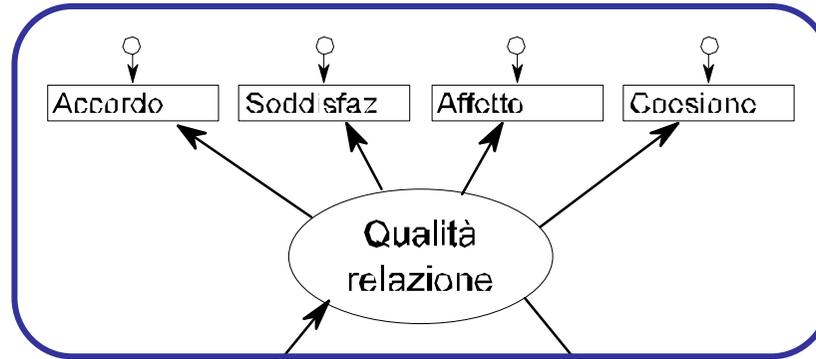
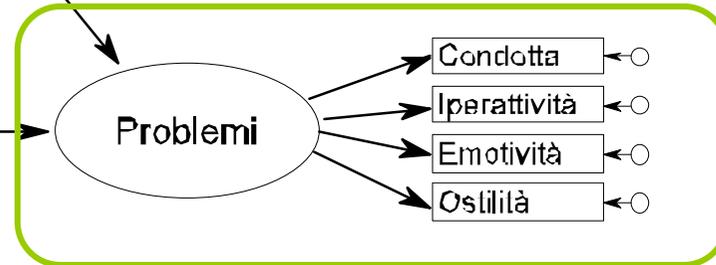
- Specificazione di relazioni causali fra variabili latenti e fra variabili latenti e osservate
- Modelli di misurazione e modelli strutturali
- Variabili esogene ed endogene
- Questione della causalità

Modello di misurazione per la qualità della relazione di coppia

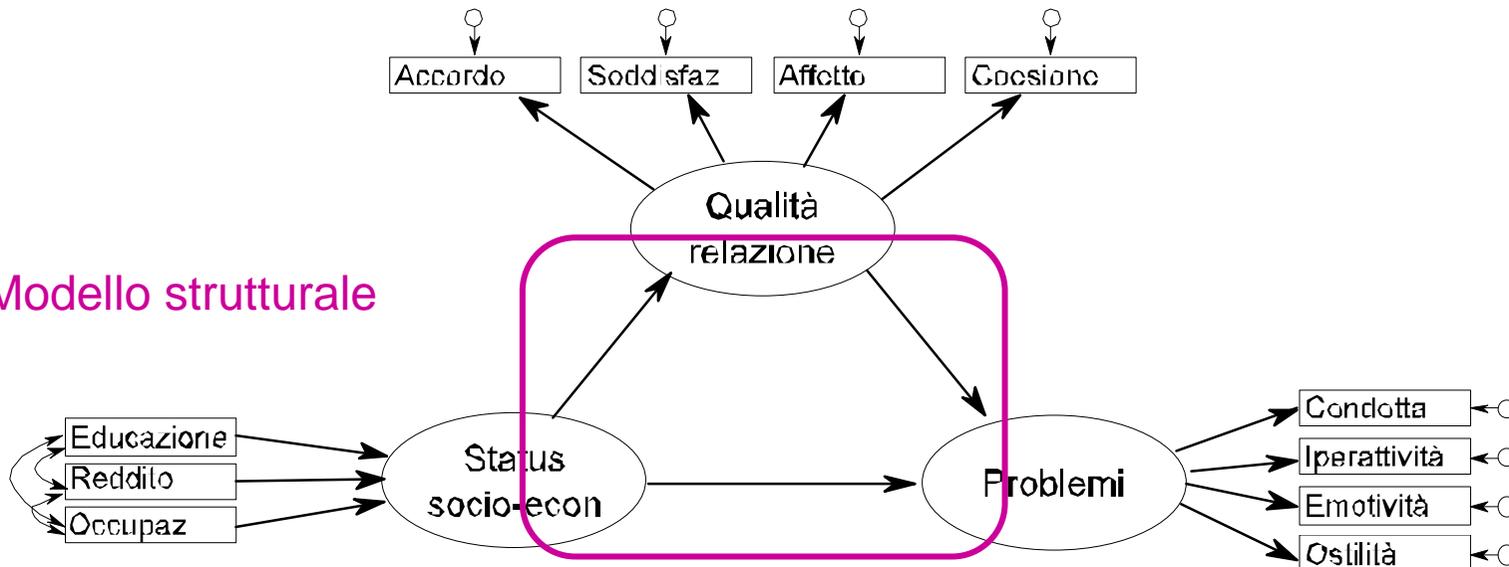
Modello di misurazione per il SES



Modello di misurazione per i problemi emotivo-comportamentali del bambino

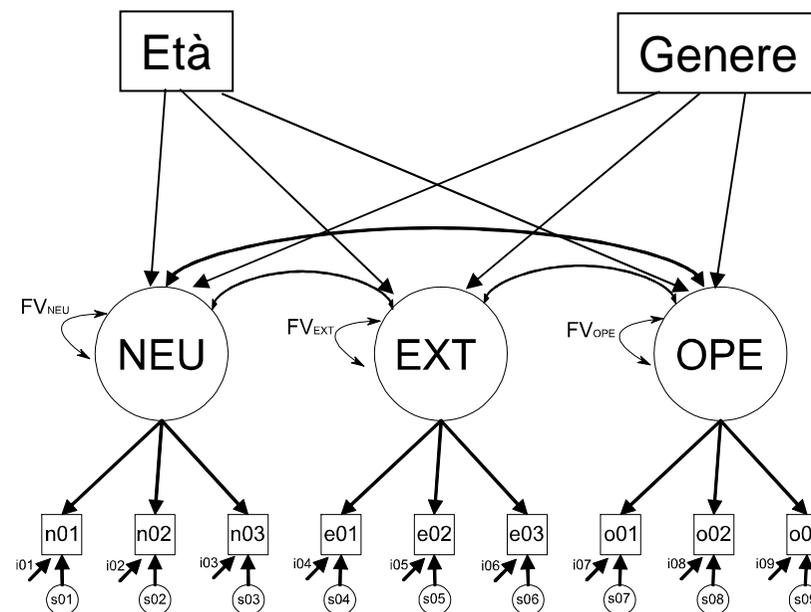


Modello strutturale



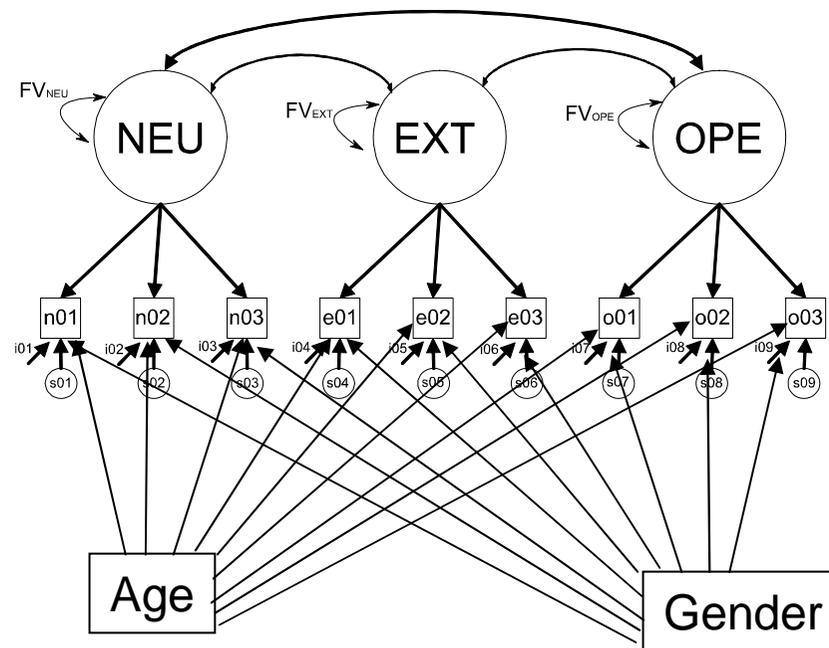
MIMIC Model

- I modelli Multiple Indicators Multiple Causes (MIMIC) prevedono relazioni fra variabili osservate e variabili latenti



Differential item functioning

- Le variabili osservate spiegano una quota di varianza degli indicatori osservati ulteriore rispetto a quella spiegata dai fattori



Bibliografia di riferimento

- Chiorri, C. (2011). *Teoria e tecnica psicometrica*. Milano: McGraw-Hill.
- Materiali supplementari:
<http://www.ateneonline.it/chiorri/studenti2/areastudenti1.asp>
- Chiorri, C. (2014). *Fondamenti di psicometria (2a edizione)*. Milano: McGraw-Hill.
- Materiali supplementari:
<http://www.ateneonline.it/chiorri2e/areastudenti.asp>