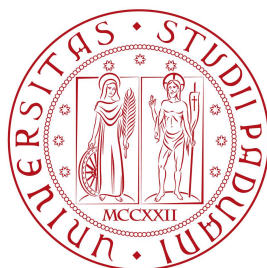


Università degli Studi di Padova
Dipartimento di Scienze Statistiche
Corso di Laurea Triennale in

Statistica per le Tecnologie e le Scienze



**Analisi della struttura fattoriale di un questionario
self-report: un esempio di applicazione di modelli di
analisi fattoriale confermativa**

Relatore: Prof.ssa Laura Ventura
Dipartimento di Scienze Statistiche

Correlatore: Dott.ssa Tatiana Marci
Dipartimento di Psicologia dello Sviluppo e della Socializzazione

Laureando: Pietro Renna

Matricola n. 1224873

Anno Accademico 2021/2022

*Alla mia famiglia, in particolare
a nonna Carmela, a nonna
Angela e a zia Tonia.*

Indice

1	Introduzione	1
1.1	Problema della misurazione in psicologia	1
1.2	Validazione del costrutto	2
1.3	Obiettivo e presentazione della tesi	3
2	Structural Equation Modelling (SEM)	6
2.1	Introduzione	6
2.2	Path Analysis	7
2.2.1	Cenni storici	9
2.3	Analisi Fattoriale Confermativa (CFA)	9
2.3.1	Definizione	9
2.3.2	Specificazione del modello	10
2.3.3	Identificazione del modello	12
2.3.4	Stima del modello	15
2.4	Bontà di adattamento	17
2.4.1	Statistica test χ^2	17
2.4.2	Goodness of Fit Index (GFI)	19
2.4.3	Standardized Root-Mean-Square Residual Index (SRMR)	19
2.5	Confronto tra modelli	20
2.5.1	Statistica test χ^2	20
2.5.2	Tucker–Lewis Index (TLI)	20
2.5.3	Comparative Fit Index (CFI)	21
2.5.4	Root Mean Square Error of Approximation (RMSEA)	22
2.6	Valori di trade-off	22
3	Caso di studio	24
3.1	Introduzione	24
3.2	Procedura e metodo	25
3.2.1	Partecipanti	25

3.3	Strumenti	28
3.3.1	Caratteristiche socio-demografiche	28
3.4	Analisi dei dati	29
3.5	Risultati	30
3.5.1	Analisi descrittive	30
3.5.2	Struttura fattoriale	43
3.5.3	Indici di bontà di adattamento dei modelli CFA di primo ordine	55
4	Conclusioni e discussione	56
A	Questionario EATQ-R	58
B	Specificazione modelli CFA in R	62

Capitolo 1

Introduzione

1.1 Problema della misurazione in psicologia

La misurazione è l'insieme delle operazioni che descrivono ogni tipo di approccio utilizzato dai ricercatori per quantificare e rappresentare una variabile di interesse (Flake e Fried, 2020).

In psicologia, nella maggior parte dei casi, l'oggetto di misura non può essere direttamente osservato. Il processo di individuazione dei possibili comportamenti (i.e., dominio di contenuto) che permette di rilevare la presenza di una caratteristica psicologica non direttamente osservabile (variabile latente) e trasformarla in operazioni che chiunque può osservare o eseguire è detto operazionalizzazione del costrutto. Il valore della ricerca scientifica dipende completamente dal grado di successo e di validità delle operazionalizzazioni.

Di recente, in psicologia è stata messa in discussione l'affidabilità dei risultati (Open Science Collaboration, 2015) e sono stati individuati numerosi aspetti potenzialmente problematici riconducibili alla mancata replica dei risultati, dovuti da una parte all'applicazione di metodologie e procedure statistiche non adeguate, dall'altra all'utilizzo di misure di dubbia validità e affidabilità (Flake *et al.*, 2022). Questo rappresenta il primo ostacolo per condurre studi di replica e suggerisce la necessità di migliorare le pratiche di misurazione.

1.2 Validazione del costrutto

La validità delle conclusioni di uno studio è un concetto ampio che può essere valutato in base a molti aspetti tra cui, ma non solo, la validità del costrutto (Flake *et al.*, 2022). La validazione del costrutto è il processo con il quale gli scienziati generano evidenza a sostegno dell'affermazione che un punteggio (score) mantenga uno specifico significato (Cronbach e Meehl, 1955), dove un costrutto è inteso come un concetto teorico. Ci si riferisce quindi alla misura in cui le operazioni delle variabili indipendenti e delle variabili dipendenti in uno studio corrispondano alle corrispettive variabili concettuali (costrutti) (Fabrigar *et al.*, 2020), e che per esse ne venga generata evidenza. Come precedentemente anticipato, in psicologia i costrutti corrispondono perlopiù a fenomeni che non sono direttamente osservabili e si tratta quindi di concetti latenti che vengono definiti a partire da un insieme di fenomeni misurabili. È dunque fondamentale studiare la validità dei costrutti a partire dalla loro identificazione, dalla scelta delle misurazioni osservabili necessarie alla loro definizione e dalla teoria che sostiene questa struttura, fino a giungere alla verifica dell'evidenza a supporto del costrutto per le misure che vengono utilizzate. L'utilizzo di misure con scarsa validità porta inevitabilmente a misurazioni non attendibili, con la conseguente incapacità di giungere a conclusioni affidabili. Sebbene all'interno della letteratura vengano raccomandate e suggerite specifiche procedure, necessarie per lo sviluppo di misure valide e affidabili, non sempre tali raccomandazioni vengono seguite in ogni fase dello sviluppo e valutazione degli strumenti (i.e., dalla fase di costruzione alla fase di validazione). In Flake *et al.* (2017) vengono suggeriti alcuni step fondamentali per lo sviluppo degli strumenti in psicologia e per la verifica della loro validità, e vengono individuate tre fasi. La fase sostanziale si occupa della definizione dei costrutti a livello teorico ed è svolta consultando la letteratura. Nella seconda fase, detta strutturale, si effettua un'analisi descrittiva per le risposte degli item e si definiscono i modelli per i costrutti di interesse. Segue infine la fase esterna, che prevede una serie di analisi e processi di verifica della validità esterna dello strumento e che studia le differenze tra i diversi costrutti.

Fase	Evidenza della validità	Descrizione
Sostanziale	Revisione della letteratura e concettualizzazione del costrutto	Identificazione di profondità e ampiezza del costrutto (Gelbach & Brinkworth, 2011)
	Sviluppo dell'item e selezione della scala Rilevanza del contenuto e rappresentatività	Revisione di un esperto (Gelbach & Brinkworth, 2011) Mappatura degli item (Dawis, 1987), gruppi di attenzione, e colloquio cognitivo (i.e., think aloud; Willis, 2004), indagare il costrutto sotto rappresentazione o irrilevanza (i.e., content validity, Sireci, 1998)
Strutturale	Analisi degli item Analisi dei fattori	Distribuzione delle risposte, correlazioni totali degli item, e difficoltà Analisi esplorativa e confermativa inclusi i modelli di equazioni strutturali e la item response theory
	Affidabilità	Coefficienti α e ω (McDonald, 1999); correlazioni interitem, test-retest (McCrae, Kurtz, Yamagata, & Terracciano, 2011), affidabilità (Chmielewski & Watson, 2009)
	Test sull'invarianza di misurazione (i.e. funzionamento differenziale degli item)	Analisi fattoriale per gruppi multipli, item response theory, e test sul funzionamento differenziale degli item (Millsap, 2011)
Esterna	Convergente e discriminante	Correlazioni tra altre scale volte a catturare costrutti simili e diversi, analisi sulla matrice multitrait-multimethod (Campbell & Fiske, 1959)
	Previsione/criterio Gruppi noti	Regressioni sui criteri delle variabili di importanza Individuare le differenze tra i gruppi noti per differire su un costrutto

Figura 1.1: Fasi di validazione di un costrutto (Flake *et al.* (2017), traduzione della Tabella 1, *Examples of Validity Evidence and Resources for Each Phase of Construct Validation*)

1.3 Obiettivo e presentazione della tesi

L'obiettivo generale di questo lavoro è presentare alcune delle procedure analitiche utilizzate nella fase di verifica della validità strutturale degli strumenti (Fase strutturale, Figura 1.1). A tal fine è stato considerato uno strumento sviluppato per la valutazione del temperamento durante l'adolescenza riconducibile all'approccio teorico di Rothbart e Derryberry (1981). Sebbene non ci si soffermerà su questo punto, è importante tenere conto che attualmente esistono più approcci allo studio del temperamento, in quanto i gruppi di ricerca interpretano il concetto in modi diversi, e per le diverse concettualizzazioni vengono utilizzati approcci metodologici differenti (Zentner e Shiner, 2012) producendo importanti ripercussioni su come il costrutto viene operazionalizzato.

In Rothbart e Derryberry (1981) si definisce il temperamento come una serie di differenze nella reattività (facilità con cui nell'individuo si possono suscitare delle risposte a livello neuroendocrino, autonomico e affettivo) e nell'autoregolazione (processi che contribuiscono a modulare la reattività e che si manifestano in comportamenti quali l'inibizione e la regolazione attentiva). Tale modello assume che la reattività e i comportamenti regolatori siano legati a processi neurobiologici. Le misure sviluppate all'interno di questo approccio includono sia misure comportamentali (es. compiti strutturati in laboratorio) che questionari speci-

fici che permettono di valutare il costrutto dall'infanzia all'età adulta. I questionari sono gli strumenti più utilizzati per valutare il temperamento in bambini e adolescenti, e hanno il vantaggio di poter essere impiegati per raccogliere informazioni da diverse fonti (genitori, insegnanti), nonché dai diretti interessati (self-report). L'*Early Adolescent Temperament Questionnaire* (EATQ-R; Capaldi e Rothbart, 1992) è uno dei pochi questionari disponibili per valutare il temperamento durante l'adolescenza. Nonostante l'ampio utilizzo a livello internazionale, pochi studi ne hanno eseguito un'approfondita valutazione delle proprietà psicometriche, suggerendo spesso risultati non sovrapponibili e rendendo potenzialmente discutibili le conclusioni che potrebbero emergere dalla sua applicazione. Inoltre, non esistono ad oggi studi che ne hanno indagato le proprietà psicometriche nella popolazione italiana.

Questo rappresenta una grande lacuna, considerato che lo studio dei tratti temperamentali costituisce un argomento privilegiato di indagine nel campo della ricerca psicologica e di come le caratteristiche individuali riconducibili a tali aspetti siano state spesso indicate come possibili fattori di rischio per lo sviluppo di problemi emotivi e comportamentali in adolescenza (Sanson *et al.*, 2004). Di conseguenza, la verifica dell'utilizzo di uno strumento valido e affidabile è indispensabile sia nella ricerca che in ambito applicativo.

Tenuto conto dell'attuale utilizzo e dei precedenti studi in letteratura, il seguente lavoro ha focalizzato l'attenzione sull'analisi descrittiva degli item e l'analisi della struttura fattoriale che rappresentano parte del processo di verifica della validità strutturale del questionario. Nello specifico, a livello descrittivo l'obiettivo è stato esplorare le distribuzioni delle risposte, l'analisi dei principali indici di tendenza centrale e di dispersione, mentre per valutare la struttura fattoriale è stata applicata l'analisi fattoriale confermativa. Nel Capitolo 2 vengono presentati i modelli di equazioni strutturali (SEM), in particolare definendo i modelli di analisi fattoriale confermativa e gli indici di bontà e di confronto per essi. Nel Capitolo 3 viene presentata la metodologia di analisi e i risultati ottenuti, i quali vengono discussi in conclusione nel Capitolo 4. Sebbene questo lavoro possa essere ritenuto un primo step nella valutazione dello strumento, i risultati evidenziano l'importanza dell'analisi descrittiva come prima fase nel processo di valutazione degli strumenti e la necessità di verificare in prima istanza la buona operazionalizzazione delle singole sottoscale in quei casi in cui il costrutto generale sia rappresentato da più dimensioni (costrutto multidimensionale). Infatti, qualora alcune delle dimensioni contenessero degli indicatori con bassa

dimensionalità (i.e., bassi loading) o presentassero problemi di multicollinearità (i.e., item molto simili), questo condurrebbe a ripercussioni nella bontà di adattamento di modelli più complessi e in particolare nei modelli di ordine superiore costruiti, la cui definizione necessita dei fattori di primo livello.

Capitolo 2

Structural Equation Modelling (SEM)

2.1 Introduzione

I modelli di equazioni strutturali sono un insieme di tecniche statistiche multivariate che consentono di modellare in modo complesso dati multivariati correlati, al fine di individuare le interrelazioni tra variabili osservate e variabili latenti (si veda, ad esempio, Ramlall, 2016).

L'obiettivo dell'analisi attraverso i SEM è determinare in quale misura il modello teorico è supportato dai dati. Se i dati campionari supportano il modello teorico, si possono ipotizzare modelli teorici più complessi, mentre se i dati del campione non supportano il modello teorico, è possibile modificare e testare il modello originale oppure sviluppare e testare altri modelli teorici (Schumacker e Lomax, 2010). Altri nomi con cui ci si riferisce ai SEM in letteratura sono:

- *Covariance Structure Analysis*: perché con i SEM si analizza la matrice di covarianza e non i valori osservati delle variabili;
- *Analysis of Moment Structure*: perché, oltre all'analisi della matrice di covarianza, le applicazioni moderne dei SEM si occupano anche dello studio della media e dei momenti di ordine superiore;
- *LISREL (Linear Structural Relations) Model*: nome dato da un noto software per i SEM;
- *Causal Modelling*: perché, sotto determinate assunzioni, un SEM supporta l'inferenza causale come un modello strutturale causale.

Ci sono diverse definizioni di SEM, ma in termini generali possono essere visti come un'estensione della *path analysis* a cui si aggiunge l'uso di variabili latenti (Figura 2.1).

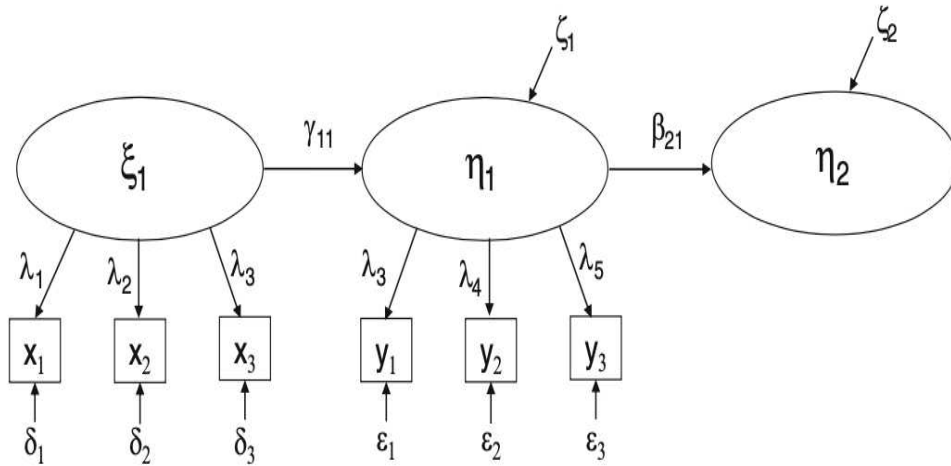


Figura 2.1: Esempio di rappresentazione grafica di un modello SEM (Stein *et al.*, 2012, Figura 1)

2.2 Path Analysis

La *path analysis* è un'estensione del modello di regressione, utilizzata per testare l'adattamento della matrice di correlazione rispetto a due o più modelli causali che vengono confrontati dal ricercatore (Garson, 2013). Il modello che si vuole adattare ai dati viene rappresentato tramite diagrammi (*path diagram*) invece che in forma di equazioni matematiche (Figura 2.3), utilizzando una notazione standardizzata (Figura 2.2) per evidenziare i sistemi di relazioni tra le variabili ed esaminare gli effetti causali. Le frecce unidirezionali rappresentano percorsi causali, in cui una variabile influenza direttamente un'altra, mentre le frecce a doppia direzione rappresentano correlazioni tra variabili (Stein *et al.*, 2012).

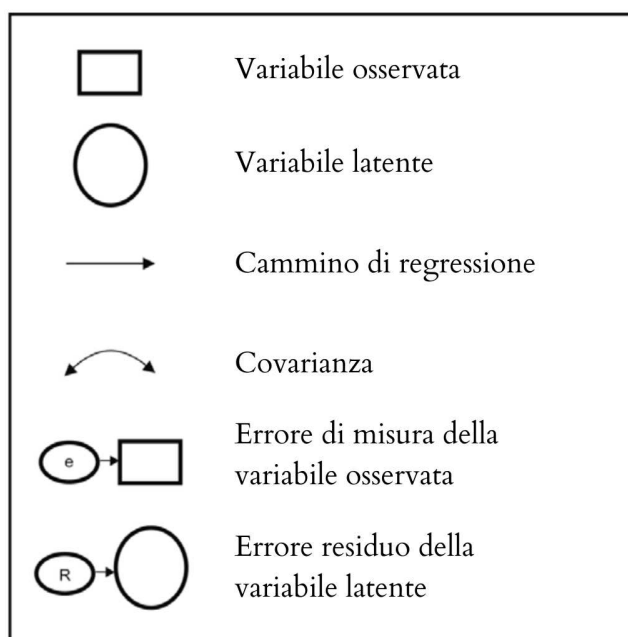


Figura 2.2: Notazione nei modelli di equazioni strutturali (Civelek, 2018, traduzione della Figura 3)

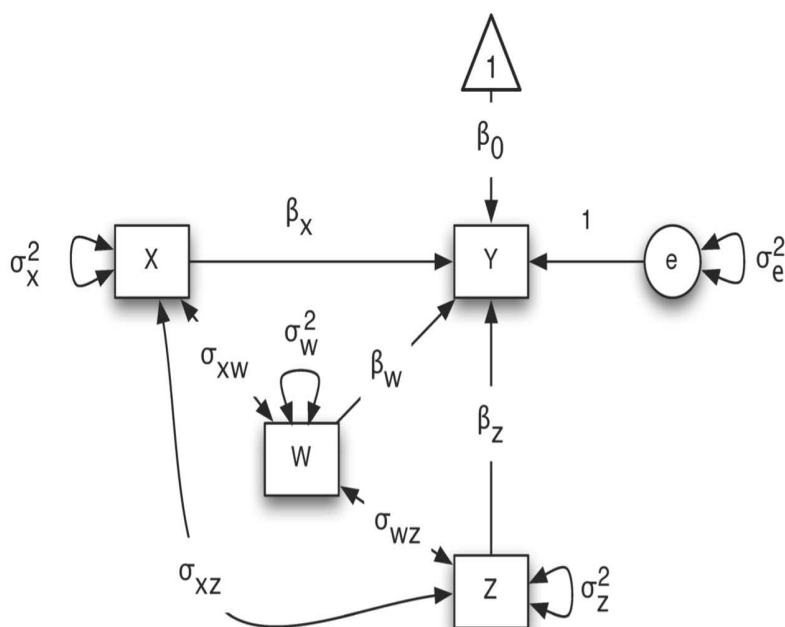


Figura 2.3: Rappresentazione del modello di regressione multipla $Y = \beta_0 + \beta_X X + \beta_W W + \beta_Z Z + \varepsilon$, (Zhang e Wang, 2017, Capitolo 13)

2.2.1 Cenni storici

Seppur non si conosca con esattezza l'origine storica, è noto che nel 1920 il genetista Sewall Wright fornì un primo sviluppo della *path analysis* per degli studi di filogenetica (Hoyle, 2012).

Negli anni 70, la *path analysis* è diventata più popolare e sono stati pubblicati numerosi articoli in sociologia, psicologia, scienze politiche, ecologia e altri campi. A partire dagli anni 80, la *path analysis* si è evoluta in una serie di programmi e pacchetti informatici di modellazione delle equazioni causali o strutturali (SEM), che hanno superato i problemi dei metodi ai minimi quadrati con sviluppi dei modelli lineari generalizzati (GLM). Il vantaggio principale dei SEM è che forniscono misure migliori dei costrutti teorici sottostanti le variabili e possono stimare modelli sia ricorsivi che non, modelli con errori di misura, modelli che includono variabili non osservate (latenti). Si ricorda che:

- **Modello ricorsivo:** tutti gli effetti causali vanno nella stessa direzione (unidirezionali) e i termini di disturbo sono incorrelati tra loro;
- **Modello non ricorsivo:** contiene uno o più effetti causali che si puntano a vicenda, o vi è la presenza di fattori di disturbo correlati.

Inoltre, i SEM forniscono assunzioni più flessibili tra cui la modellazione con termini di errore correlati, interazioni e non linearità (Lleras, 2005).

2.3 Analisi Fattoriale Confermativa (CFA)

2.3.1 Definizione

La maggior parte dei concetti sociali scientifici non è misurabile direttamente (intelligenza, capacità produttiva di un individuo, felicità,...) e nasce quindi la necessità di considerare questi costrutti come variabili latenti: seppur non siano direttamente misurabili, possono essere ottenute tramite le variabili osservate dalle quali sono causate.

L'*analisi fattoriale* (FA) è una delle tecniche di riduzione della dimensionalità più utilizzate nelle scienze sociali e comportamentali, e può essere vista come un paradigma di modellazione di variabili latenti in cui un insieme di variabili osservate rappresenta gli indicatori di una variabile latente (Finch e French, 2015).

Viene effettuata tipicamente tramite due tipi di approcci:

- *Analisi Fattoriale Esplorativa (EFA)*: nota anche con il nome di FA non restrittiva perché non pone restrizioni sulla matrice di covarianza. Si tratta di un approccio induttivo, preferibile quando non si fanno assunzioni a priori sulla struttura fattoriale del fenomeno di interesse o quando, in generale, il ricercatore non impone una struttura latente ma fissa il numero ottimale di fattori basandosi sull'interpretazione dei risultati ottenuti a seguito dell'analisi (ad esempio, considerando la quantità di varianza spiegata).
- *Analisi Fattoriale Confermativa (CFA)*: si valutano delle ipotesi formulate a priori e richiede di fissare il numero di fattori e la struttura latente da ricercare prima dell'analisi. Il ricercatore deve esplicitare gli indicatori di riferimento per ogni fattore di interesse secondo l'approccio teorico che si vuole seguire. È utile anche per comparare diversi modelli e più ipotesi teoriche ma richiede una forte evidenza empirica che supporti, almeno parzialmente, gli assunti teorici della struttura latente (Finch e French, 2015).

Quest'ultima sarà proprio l'approccio che si andrà ad approfondire e che viene trattato nelle analisi del caso applicativo.

2.3.2 Specificazione del modello

Terminologia

1. **Variabile endogena (dipendente)**: variabile la cui variazione è spiegata da una o più variabili nel modello (Lleras, 2005). Graficamente presenta almeno una freccia in entrata;
2. **Variabile esogena (indipendente)**: variabile la cui variazione è spiegata da fattori esterni al modello, e inoltre spiega altre variabili presenti nel modello (Lleras, 2005). Non presenta frecce in entrata, ad eccezione del termine di errore;
3. **Variabile di mediazione**: variabile sia esogena che endogena. Ha il ruolo di condurre l'effetto di una variabile verso un'altra.

Il modello fa parte della famiglia più ampia e generale dei modelli a variabili latenti. Si consideri come riferimento il modello 2.1 LISREL nella notazione modificata di (Jöreskog e Sörbom, 1978), dato da:

$$\boldsymbol{\eta} = \boldsymbol{\alpha} + \mathbf{B}\boldsymbol{\eta} + \boldsymbol{\Gamma}\boldsymbol{\xi} + \boldsymbol{\zeta}, \quad (2.1)$$

dove:

- $\boldsymbol{\eta}$: vettore $m \times 1$ di variabili latenti endogene;
- $\boldsymbol{\alpha}$: vettore $m \times 1$ dei termini di intercetta;
- \mathbf{B} : matrice $m \times m$ dei coefficienti dato l'impatto che le variabili endogene hanno fra di loro;
- $\boldsymbol{\xi}$: vettore $n \times 1$ di variabili latenti esogene;
- $\boldsymbol{\Gamma}$: matrice $m \times n$ dei coefficienti dato l'effetto delle variabili esogene su quelle endogene;
- $\boldsymbol{\zeta}$: vettore $m \times 1$ dei termini di errore di disturbo.

Le assunzioni del modello 2.1 sono:

- $E(\boldsymbol{\zeta}) = 0$;
- $COV(\boldsymbol{\xi}', \boldsymbol{\zeta}) = 0$;
- $(\mathbf{I} - \mathbf{B})$ invertibile.

Il caso specifico del modello di misura CFA ha equazioni

$$\mathbf{y} = \boldsymbol{\alpha}_y + \boldsymbol{\Lambda}_y \boldsymbol{\eta} + \boldsymbol{\epsilon} \quad (2.2)$$

$$\mathbf{x} = \boldsymbol{\alpha}_x + \boldsymbol{\Lambda}_x \boldsymbol{\xi} + \boldsymbol{\delta}, \quad (2.3)$$

che esplicita il legame tra il vettore di variabili latenti e le risposte osservate (indicatori) \mathbf{y} e \mathbf{x} , dove $\boldsymbol{\alpha}_y$ e $\boldsymbol{\alpha}_x$ rappresentano i termini di intercetta, $\boldsymbol{\Lambda}_y$ e $\boldsymbol{\Lambda}_x$ sono le matrici dei loadings, ossia i coefficienti di regressione dato l'impatto delle variabili latenti $\boldsymbol{\eta}$ e $\boldsymbol{\xi}$ su \mathbf{y} e \mathbf{x} , mentre $\boldsymbol{\epsilon}$ e $\boldsymbol{\delta}$ sono i fattori specifici di \mathbf{y} e \mathbf{x} . Per questi ultimi, si assume valore atteso nullo, matrici di covarianza rispettivamente $\boldsymbol{\Theta}_\epsilon$ e $\boldsymbol{\Theta}_\delta$, incorrelazione. Inoltre, se tutte le variabili sono lontane dalla loro media, in modo tale che le variabili osservate e le variabili latenti hanno media pari a zero, allora il termine di intercetta non è necessario e si ha, utilizzando l'equazione relativa a \mathbf{x} ,

$$\mathbf{x} = \boldsymbol{\Lambda}_x \boldsymbol{\xi} + \boldsymbol{\delta}, \quad (2.4)$$

che è la forma più comune utilizzata dai ricercatori (Bollen *et al.*, 2010). A partire da definizioni generali applicabili su ogni tipo di SEM, il vettore media implicito della popolazione risulta:

$$\boldsymbol{\mu}(\boldsymbol{\theta}) = \begin{bmatrix} \boldsymbol{\alpha}_y + \boldsymbol{\Lambda}_y(\mathbf{I} - \mathbf{B})^{-1}(\boldsymbol{\alpha}_n + \boldsymbol{\Gamma}\boldsymbol{\mu}_\xi) \\ \boldsymbol{\alpha}_x + \boldsymbol{\Lambda}_x\boldsymbol{\mu}_\xi \end{bmatrix}, \quad (2.5)$$

con $\boldsymbol{\theta}$ vettore dei parametri contenente i coefficienti di regressione, le varianze e le covarianze del modello.

La matrice di varianza/covarianza implicita è definita come:

$$\boldsymbol{\Sigma}(\boldsymbol{\theta}) = \begin{bmatrix} \boldsymbol{\Sigma}_{yy}(\boldsymbol{\theta}) & \boldsymbol{\Sigma}_{yx}(\boldsymbol{\theta}) \\ \boldsymbol{\Sigma}_{xy}(\boldsymbol{\theta}) & \boldsymbol{\Sigma}_{xx}(\boldsymbol{\theta}) \end{bmatrix}, \quad (2.6)$$

che risulta essere partizionata e composta dalle matrici di covarianza:

$$\boldsymbol{\Sigma}_{xx}(\boldsymbol{\theta}) = \boldsymbol{\Lambda}_x\boldsymbol{\Phi}\boldsymbol{\Lambda}'_x + \boldsymbol{\Theta}_\delta \quad (2.7)$$

$$\boldsymbol{\Sigma}_{xy}(\boldsymbol{\theta}) = \boldsymbol{\Lambda}_x\boldsymbol{\Phi}\boldsymbol{\Gamma}'(\mathbf{I} - \mathbf{B})^{-1}\boldsymbol{\Lambda}'_y \quad (2.8)$$

$$\boldsymbol{\Sigma}_{yy}(\boldsymbol{\theta}) = \boldsymbol{\Lambda}_y(\mathbf{I} - \mathbf{B})^{-1}(\boldsymbol{\Gamma}\boldsymbol{\Phi}\boldsymbol{\Gamma}' + \boldsymbol{\Psi})(\mathbf{I} - \mathbf{B})^{-1'}\boldsymbol{\Lambda}'_y + \boldsymbol{\Theta}_\epsilon, \quad (2.9)$$

dove $\boldsymbol{\Phi}$ è la matrice $n \times n$ di covarianza delle variabili latenti esogene $\boldsymbol{\xi}$, $\boldsymbol{\Psi}$ è la matrice $m \times m$ di covarianza dei termini di errore di disturbo $\boldsymbol{\zeta}$, ed entrambe sono quadrate e simmetriche. Il modello è valido se le seguenti ipotesi risultano vere:

$$\boldsymbol{\mu} = \boldsymbol{\mu}(\boldsymbol{\theta}) \quad (2.10)$$

$$\boldsymbol{\Sigma} = \boldsymbol{\Sigma}(\boldsymbol{\theta}), \quad (2.11)$$

ossia se è possibile riprodurre esattamente le medie, le varianze e le covarianze implicite utilizzando i valori osservati dei parametri della popolazione del modello (Bollen *et al.*, 2010).

2.3.3 Identificazione del modello

Per i modelli CFA le condizioni di sufficienza per l'identificazione del modello sono date dalle *regole dei tre indicatori* e dalle *regole dei due indicatori* (Bollen, 1989).

Regola dei tre indicatori (Three-indicator rule)

1. $n \geq 1$;
2. è presente un solo elemento non nullo per riga nella matrice Λ_x ;
3. ci sono tre o più indicatori per fattore;
4. Θ_δ è diagonale.

Regole dei due indicatori (Two-indicator rules)

1. $n \geq 1$;
2. tutti gli elementi di Φ sono non nulli;
3. è presente un solo elemento non nullo per riga nella matrice Λ_x ;
4. ci sono due o più indicatori per fattore;
5. Θ_δ è diagonale.

Tuttavia, in Bollen (1989) si elimina la seconda condizione assumendo che almeno un elemento non diagonale di Φ sia non nullo, per ogni riga. La limitazione più grande di queste regole è l'assunzione di Θ_δ diagonale, ossia nessun errore è correlato. A causa delle limitazioni di queste regole, molti ricercatori effettuano test empirici di identificazione locale forniti dai principali pacchetti software invece di identificare il modello globalmente tramite l'algebra sulla covarianza (Davis, 1993). Un vettore dei parametri θ è globalmente identificato se non ci sono dei vettori θ_1 e θ_2 tali per cui $\Sigma(\theta_1) = \Sigma(\theta_2)$, tranne che nel caso $\theta_1 = \theta_2$. L'identificazione locale è invece un concetto più debole di unicità: un vettore dei parametri θ è localmente identificato in un punto θ_1 se nelle vicinanze di θ_1 non è presente un vettore θ_2 tale per cui $\Sigma(\theta_1) = \Sigma(\theta_2)$. Seppur i due tipi di identificazione possano sembrare simili, vi sono delle differenze notevoli: l'identificazione locale valuta l'identificazione per θ_1 , specifico punto di θ , e determina se la matrice di covarianza nei parametri cambia per piccoli cambiamenti di θ_1 , invece l'identificazione globale non è così limitata. Per quest'ultima, infatti, si considera ogni coppia di vettori di possibili valori di θ . L'identificazione locale è necessaria, ma non è sufficiente, per l'identificazione globale, mentre l'identificazione globale implica quella locale (Bollen, 1989).

Vengono mostrati in sintesi nella Figura 2.4 tutti i metodi presentati da Bollen (1989).

Regole di identificazione	Identifica:	Riassunto delle regole ^a	Condizione necessaria	Condizione sufficiente
<i>t</i> -Rule	modello	$t \leq \frac{1}{2}(q)(q + 1)$	si	no
Regola dei tre indicatori	modello	$n > 1$ un elemento non nullo per riga di Λ_x tre o più indicatori per fattore	no	si
Regole dei due indicatori		Θ_δ diagonale		
Regola 1	modello	$n > 1$ $\phi_{i.} \neq 0$ per ogni i e j un elemento non nullo per riga di Λ_x due o più indicatori per fattore	no	si
Regola 2	modello	Θ_δ diagonale $n > 1$ $\phi_{ij} \neq 0$ per almeno una coppia di i e j , dove $i \neq j$ un elemento non nullo per riga di Λ_x due o più indicatori per fattore	no	si
		Θ_δ diagonale		

^a Si assume che a tutte le variabili latenti venga assegnata una scala

Figura 2.4: Metodi di identificazione (Bollen, 1989, traduzione della Tabella 7.3, *Summary of Identification Rules for Confirmatory Factor Analysis*)

Valutazione degli stati di identificazione

Nei SEM l'informazione nota è costituita dalla matrice di varianza e covarianza, dalla media e dai momenti di ordine superiore. Rimangono invece ignoti i parametri che si vogliono stimare per il modello, e a seconda del numero di variabili osservate p , del numero di parametri ignoti t e del numero di parametri non ridondanti $\frac{1}{2}p(p + 1)$. Si presentano tre casi (Sturgis, 2022):

- **Modello sotto-identificato (under-identified)**. La quantità di informazione nota è minore dell'informazione ignota:

$$t > \frac{1}{2}p(p + 1).$$

- **Modello identificato (just-identified)**. La stessa quantità di informazione nota e ignota (il numero di parametri non ridondanti e il numero di parametri da stimare coincide):

$$t = \frac{1}{2}p(p + 1).$$

- **Modello sovra-identificato (over-identified)**. La quantità di informazione nota è maggiore dell'informazione ignota:

$$t < \frac{1}{2}p(p + 1).$$

Nei SEM, qualora si fosse nella situazione di modello sotto-identificato o identificato, serve ricondursi a una forma di modello sovra-identificato: questo è possibile grazie a operazioni di aggiunta di variabili osservate o di diminuzione del numero di parametri (ad esempio, assumendo l'uguaglianza fra due parametri).

2.3.4 Stima del modello

Massima Verosimiglianza

Gli stimatori a informazione piena sono quelli più utilizzati, e infatti nella maggior parte dei software di analisi è impostato di *default* lo stimatore di massima verosimiglianza a informazione piena. La funzione di verosimiglianza implementata è:

$$F_{ML} = \ln|\Sigma(\boldsymbol{\theta})| - \ln|\mathbf{S}| + \text{tr}[\Sigma^{-1}(\boldsymbol{\theta})\mathbf{S}] - P_z + (\bar{\mathbf{z}} - \boldsymbol{\mu}(\boldsymbol{\theta}))'\Sigma^{-1}(\boldsymbol{\theta})(\bar{\mathbf{z}} - \boldsymbol{\mu}(\boldsymbol{\theta})), \quad (2.12)$$

dove \mathbf{S} è la matrice di covarianza campionaria, $\bar{\mathbf{z}}$ è il vettore delle medie campionarie delle variabili osservate, P_z è il numero di variabili osservate (Bollen *et al.*, 2010). Lo stimatore $\hat{\boldsymbol{\theta}}$ è scelto per minimizzare la funzione F_{ML} e essendo uno stimatore di massima verosimiglianza, gode delle proprietà di consistenza, non-distorsione asintotica e matrice di covarianza asintotica pari alla matrice di informazione attesa. Inizialmente

si assumeva che tutte le variabili osservate provenissero da una distribuzione normale multivariata, ma in seguito Browne (1984) ha dimostrato che, se le variabili non presentano un eccesso di curtosi multivariata, le proprietà di massima verosimiglianza continuano a essere valide anche in presenza di asimmetria multivariata. Altri studi hanno evidenziato le condizioni in cui lo stimatore ML rimane robusto anche in presenza di un eccesso di curtosi multivariata (Bollen *et al.*, 2010).

Two Stage Least Squared (2SLS)

Bollen (1996) sviluppò uno stimatore basato sulla famiglia dei *Two Stage Least Squares* (2SLS) applicabile per l'analisi fattoriale e ai SEM. L'approccio si basa sull'utilizzo di un indicatore di scala per ogni $\boldsymbol{\eta}_i$, tale per cui $\boldsymbol{y}_{1i} = \boldsymbol{\eta}_i + \boldsymbol{\epsilon}_i$, con \boldsymbol{y}_{1i} vettore di indicatori di scala, e allo stesso tempo $\boldsymbol{x}_{1i} = \boldsymbol{\xi}_i - \boldsymbol{\delta}_i$ porta a ottenere $\boldsymbol{\xi}_i = \boldsymbol{x}_{1i} - \boldsymbol{\delta}_i$. Sostituendo $\boldsymbol{\eta}_i$ con $(\boldsymbol{y}_{1i} - \boldsymbol{\xi}_i)$ e ponendo $\boldsymbol{\xi}_i = \boldsymbol{x}_{1i} - \boldsymbol{\delta}_i$ nel modello a variabili latenti e nel modello di misura, il SEM diventa

$$\boldsymbol{y}_{1i} = \boldsymbol{\alpha}_n + \boldsymbol{B}\boldsymbol{y}_{1i} + \boldsymbol{\Gamma}\boldsymbol{x}_{1i} + \boldsymbol{\epsilon}_{1i} - \boldsymbol{B}\boldsymbol{\epsilon}_{1i} - \boldsymbol{\Gamma}\boldsymbol{\delta}_{1i} + \boldsymbol{\zeta}_i, \quad (2.13)$$

$$\boldsymbol{y}_{2i} = \boldsymbol{\alpha}_{y2} + \boldsymbol{\Lambda}_{y2}\boldsymbol{y}_{1i} - \boldsymbol{\Lambda}_{y2}\boldsymbol{\epsilon}_{1i} + \boldsymbol{\epsilon}_{2i} \quad (2.14)$$

$$\boldsymbol{x}_{2i} = \boldsymbol{\alpha}_{x2} + \boldsymbol{\Lambda}_{x2}\boldsymbol{x}_{1i} - \boldsymbol{\Lambda}_{x2}\boldsymbol{\delta}_{1i} + \boldsymbol{\delta}_{2i}, \quad (2.15)$$

dove \boldsymbol{y}_{2i} e \boldsymbol{x}_{2i} sono le variabili (non scalari) osservate.

Le singole equazioni di questo modello appaiono come equazioni di regressione. Tuttavia, l'uso della regressione con i *minimi quadrati ordinari* (OLS) non è considerata in generale, poiché alcuni termini di disturbo sono correlati con una o più variabili esplicative, motivo per cui l'OLS diventa, in questo caso, uno stimatore inconsistente. Lo stimatore 2SLS fa uso di variabili strumentali (*Instrumental Variables*; IV) che vengono selezionate per ogni equazione. Le variabili strumentali sono variabili osservate del modello che non sono correlate con i termini di disturbo dell'equazione, e che introducono la variabilità esogena in una variabile endogena. In Bollen e Bauer (2004) viene proposto un algoritmo che trova le variabili strumentali per una data struttura del modello.

I metodi sinora proposti valgono per variabili continue, mentre di seguito si presenta un metodo per dati ordinali.

Weighted Least Squares (WLS)

Il metodo è utilizzato quando sono presenti uno o più indicatori di tipo categoriale e, in tal caso, non sarebbe corretto utilizzare il metodo della

massima verosimiglianza trattando le variabili categoriali come continue perché si otterrebbero delle stime attenuate delle correlazioni tra gli indicatori, statistiche test e standard error scorretti, ed inoltre il metodo potrebbe ad avere "pseudofattori" che sono artefatti della difficoltà o dell'estremità degli indicatori (Brown, 2015).

L'obiettivo è minimizzare la funzione

$$F_{WLS}(\theta) = [\mathbf{s} - \boldsymbol{\sigma}(\theta)]' \mathbf{W}^{-1} [\mathbf{s} - \boldsymbol{\sigma}(\theta)], \quad (2.16)$$

dove \mathbf{s} è il vettore di elementi non ridondanti nella matrice di covarianza empirica, $\boldsymbol{\sigma}(\theta)$ è il vettore di elementi non ridondanti nella matrice di covarianza del modello $\boldsymbol{\Sigma}(\theta)$, θ vettore dei parametri, \mathbf{W}^{-1} è una matrice dei pesi definita positiva ($k \times k$), dove $k = p(p+1)/2$, con p numero di variabili osservate (Schermelleh-Engel *et al.*, 2003). Tuttavia, la stima WLS richiede campioni di numerosità elevata per stimare accuratamente la matrice dei momenti di ordine quarto, ossia un campione di numerosità superiore ad $a + p$ (dove a è il numero di elementi di \mathbf{S}) per evitare che \mathbf{W} sia singolare (Brown, 2015).

2.4 Bontà di adattamento

Nei modelli di equazioni strutturali la valutazione della bontà di adattamento non è semplice e immediata, e negli ultimi anni vi sono diversi dibattiti in letteratura sui metodi da utilizzare: pertanto, è necessario considerare più criteri per affermare se il modello è consistente rispetto ai dati empirici (Schermelleh-Engel *et al.*, 2003). L'adattamento del modello è quindi determinato dalla capacità di adattamento della matrice di varianza/covarianza del campione al modello di equazione strutturale (Schumacker e Lomax, 2010). Si illustrano di seguito i criteri utilizzati durante le analisi del caso applicativo, che sono basati sulle differenze tra la matrice \mathbf{S} di covarianza campionaria e la matrice $\boldsymbol{\Sigma}(\theta)$ di covarianza implicita del modello.

2.4.1 Statistica test χ^2

Tra i criteri presentati, il χ^2 è l'unico test di significatività per i modelli teorici, in quanto gli altri criteri si limitano a fornire solo statistiche descrittive per i modelli.

La statistica test è definita come (Schermelleh-Engel *et al.*, 2003):

$$\chi_{df}^2 = (n - 1)F[\mathbf{S}, \boldsymbol{\Sigma}(\hat{\theta})], \quad (2.17)$$

con $df = s - t$ gradi di libertà, dove:

- n : numerosità campionaria;
- s : numero di elementi non ridondanti in \mathbf{S} , matrice di covarianza empirica;
- t : numero totale di parametri da stimare;
- $F[\mathbf{S}, \Sigma(\hat{\theta})]$: funzione del metodo di stima in uso.

La statistica (2.17), in presenza di n elevato e assumendo normalità multivariata, risulta asintoticamente distribuita come una chi-quadro con df gradi di libertà: per questo motivo, la statistica è nota con il nome di *funzione di adattamento minimo chi-quadro* o *rapporto di log-verosimiglianza chi-quadro* (Kline, 2015). Per valori alti della statistica rispetto ai gradi di libertà si conclude che la matrice di covarianza implicita $\Sigma(\hat{\theta})$ e la matrice di covarianza empirica \mathbf{S} sono significamente diverse, e quindi i residui $\mathbf{S} - \Sigma(\hat{\theta})$ devono essere prossimi a zero per avere un buon adattamento del modello (Schermelleh-Engel *et al.*, 2003). È utile avere come riferimento due casi notevoli:

- $\chi_{df}^2 = 0$: il modello si adatta perfettamente ai dati (modello saturo);
- χ_{df}^2 massimo: modello di indipendenza, non ci sono percorsi nel modello (Schumacker e Lomax, 2010).

Il criterio, tuttavia, è sensibile rispetto a:

- **Variazioni della numerosità campionaria:** al crescere di n (generalmente oltre i 200), la statistica test assume valori alti (Schumacker e Lomax, 2010);
- **Complessità del modello:** il valore della statistica diminuisce all'aumentare del numero dei parametri nel modello;
- **Violazione delle assunzioni:** la statistica test si basa su assunzioni di distribuzione multivariata asintotica delle variabili osservate che non sono necessariamente rispettate.

2.4.2 Goodness of Fit Index (GFI)

Il GFI misura la quantità di varianza e covarianza nella matrice di covarianza empirica \mathbf{S} che è predetta dalla matrice implicita di covarianza $\Sigma(\boldsymbol{\theta})$ (Schermelleh-Engel *et al.*, 2003).

L'indice è definito da:

$$GFI = 1 - \frac{\chi_t^2}{\chi_n^2}, \quad (2.18)$$

dove

- χ_n^2 è la statistica test del modello nullo (*baseline*), che è il modello di indipendenza;
- χ_t^2 è la statistica test del modello di interesse.

Solitamente la quantità (2.18) assume valori tra zero e uno (in alcuni casi, anche negativi), e si ha un buon adattamento per valori alti.

2.4.3 Standardized Root-Mean-Square Residual Index (SRMR)

Si definisce inizialmente l'indice RMR, basato sui residui stimati (Schermelleh-Engel *et al.*, 2003), che è dato da

$$RMR = \sqrt{\frac{\sum_{i=1}^p \sum_{j=1}^i (s_{ij} - \hat{\sigma}_{ij})^2}{p(p+1)/2}}, \quad (2.19)$$

con:

- s_{ij} : elemento della matrice \mathbf{S} ;
- $\hat{\sigma}_{ij}$: elemento della matrice $\Sigma(\hat{\boldsymbol{\theta}})$;
- p : numero di variabili osservate.

Valori dell'indice prossimi a zero indicano un buon adattamento, seppur sia doveroso notare che i residui stimati siano dipendenti dalla scala e che quindi il valore del RMR dipende dalla grandezza delle varianze e delle covarianze delle variabili osservate che non sono standardizzate (Kline, 2015). Il problema della scala rende il valore del RMR di difficile interpretazione. Dunque si è reso necessario definirne una versione standardizzata in cui i residui vengono divisi per le deviazioni standard, che è proprio l'indice SRMR (Schermelleh-Engel *et al.*, 2003).

2.5 Confronto tra modelli

L'idea degli indici di comparazione è confrontare l'adattamento di un modello di interesse con un modello *baseline* di riferimento, che solitamente coincide con il modello di indipendenza: quest'ultimo assume che le variabili osservate siano misurate senza errore e che siano tra loro incorrelate. Un modello più restrittivo che può essere utilizzato come riferimento è il modello nullo, dove i parametri vengono corretti a zero (Schermelleh-Engel *et al.*, 2003). L'obiettivo è quantificare l'eventuale miglioramento del modello di interesse rispetto al modello *baseline* che si è scelto, e come nel caso di bontà di adattamento ci si limita a illustrare i soli criteri utilizzati nelle analisi del caso applicativo.

2.5.1 Statistica test χ^2

La statistica (2.17) può essere utilizzata anche per effettuare confronti tra modelli annidati (Schermelleh-Engel *et al.*, 2003), nella forma:

$$\chi_{diff}^2 = \chi_{df_A}^2 - \chi_{df_B}^2, \quad (2.20)$$

dove:

- $\chi_{df_A}^2$ indica il valore della statistica per un primo modello A (che è un modello ristretto rispetto a B), con df_A gradi di libertà;
- $\chi_{df_B}^2$ indica il valore della statistica per un secondo modello B avente gradi di libertà (df_B) minori rispetto ai gradi di libertà del modello A ;
- $diff = df_A - df_B$.

Trattandosi di una differenza tra due statistiche chi-quadro, la statistica χ_{diff}^2 è a sua volta distribuita come una chi-quadro con $diff$ gradi di libertà. Se la differenza è significativa, si preferisce il modello B , altrimenti è preferibile il modello A in quanto presenta un numero inferiore di parametri. Anche nella situazione di confronto tra modelli annidati si deve tener conto delle limitazioni della statistica test presentate in precedenza.

2.5.2 Tucker–Lewis Index (TLI)

Sviluppato inizialmente da Tucker e Lewis (1973) per l'analisi fattoriale, è stato successivamente esteso per i modelli di equazioni strutturali per

confrontare un modello di interesse con il modello nullo (Schumacker e Lomax, 2010) ed è noto anche con il nome di *Nonnormed Fit Index* (NNFI)

È costruito a partire dalle statistiche χ^2 (Schumacker e Lomax, 2010), come:

$$TLI = NNFI = \frac{[(\frac{\chi_{null}^2}{df_{null}}) - (\frac{\chi_{proposed}^2}{df_{proposed}})]}{[(\frac{\chi_{null}^2}{df_{null}}) - 1]}, \quad (2.21)$$

composto da:

1. χ_{null}^2 : statistica test del modello nullo, con df_{null} gradi di libertà;
2. $\chi_{proposed}^2$: statistica test del modello di interesse, con $df_{proposed}$ gradi di libertà;

Solitamente assume valori tra zero e uno, anche se talvolta può andare oltre questo range, e valori alti indicano un buon adattamento: si ha un adattamento accettabile per valori maggiori di 0.95, e un adattamento ottimo per valori maggiori di 0.97. Inoltre, è uno degli indici meno affetti dalla numerosità del campione (Schermelleh-Engel *et al.*, 2003).

2.5.3 Comparative Fit Index (CFI)

È una versione modificata dell'indice RNI (*Relative Noncentrality Index*) definito come (Schermelleh-Engel *et al.*, 2003):

$$CFI = 1 - \frac{\max[(\chi_t^2 - df_t), 0]}{\max[(\chi_t^2 - df_t), (\chi_i^2 - df_i), 0]}, \quad (2.22)$$

con:

- χ_i^2 : statistica test del modello di indipendenza (*baseline*);
- χ_t^2 : statistica test del modello di interesse;
- df : numero di gradi di libertà.

L'indice varia tra zero e uno, e a valori alti corrisponde un buon adattamento del modello: solitamente, si considera un buon adattamento con valori maggiori o uguali a 0.97, e si ha un adattamento accettabile per valori maggiori o uguali a 0.95.

Come per il TLI, è uno degli indici meno affetti dalla numerosità campionaria

2.5.4 Root Mean Square Error of Approximation (RMSEA)

È un indice di adattamento assoluto che misura quanto la stima sia sbagliata, dove un valore di zero indica una stima perfetta (Kline, 2015). È definito come la radice quadrata della discrepanza stimata dovuta all'approssimazione per grado di libertà (Schermelleh-Engel *et al.*, 2003):

$$\hat{\epsilon}_a = \sqrt{\max\left\{\left(\frac{F(\mathbf{S}, \boldsymbol{\Sigma}(\hat{\theta}))}{df} - \frac{1}{N-1}\right), 0\right\}}, \quad (2.23)$$

dove:

- $F(\mathbf{S}, \boldsymbol{\Sigma}(\hat{\theta}))$: minimo della funzione di stima in uso;
- $df = s - t$: gradi di libertà, con s numero di elementi non ridondanti della matrice \mathbf{S} e t numero totale di parametri stimati.

2.6 Valori di trade-off

A conclusione del capitolo, si presenta la Tabella 2.5 con i valori di trade-off per alcuni dei criteri utilizzabili, tratta da Schermelleh-Engel *et al.* (2003), i cui scopi principali sono:

1. penalizzare il modello per il numero di parametri;
2. quantificare l'adattamento del modello ai dati in analisi.

Misura di bontà	Buon adattamento	Adattamento accettabile
χ^2	$0 \leq \chi^2 \leq 2df$	$2df < \chi^2 \leq 3df$
p value	$.05 < p \leq 1.00$	$.01 \leq p \leq .05$
χ^2/df	$0 \leq \chi^2/df \leq 2$	$2 < \chi^2/df \leq 3$
$RMSEA$	$0 \leq RMSEA \leq .05$	$.05 < RMSEA \leq .08$
p value per test di adattamento ($RMSEA < .05$)	$.10 < p \leq 1.00$	$.05 \leq p \leq .10$
Intervallo di confidenza (IC)	vicino al $RMSEA$, estremo sinistro del IC = .00	vicino al $RMSEA$
$SRMR$	$0 \leq SRMR \leq .05$	$.05 < SRMR \leq .10$
NFI	$.95 \leq NFI \leq 1.00^a$	$.90 \leq NFI < .95$
$NNFI$	$.97 \leq NNFI \leq 1.00^b$	$.95 \leq NNFI < .97^c$
CFI	$.97 \leq CFI \leq 1.00$	$.95 \leq CFI < .97^c$
GFI	$.95 \leq GFI \leq 1.00$	$.90 \leq GFI < .95$
$AGFI$	$.90 \leq AGFI \leq 1.00$, vicino al GFI	$.85 \leq AGFI < .90$, vicino al GFI
AIC	più piccolo del AIC per il modello di confronto	
$CAIC$	più piccolo del $CAIC$ per il modello di confronto	
$ECVI$	più piccolo del $ECVI$ per il modello di confronto	

Figura 2.5: Criteri di valutazione e di confronto dei modelli CFA, (Schermelleh-Engel *et al.*, 2003, traduzione della Tabella 7.1, *Recommendations for Model Evaluation: Some Rules of Thumb*)

Il prossimo capitolo tratterà il caso di studio, introducendo inizialmente il questionario che è stato somministrato e gli strumenti utilizzati, e in seguito presentando il campione e le analisi descrittive per le distribuzioni degli item, fino a giungere ai modelli CFA di primo ordine e alla valutazione della bontà di adattamento di questi ultimi.

Capitolo 3

Caso di studio

3.1 Introduzione

Come introdotto nel capitolo precedente, una delle tecniche statistiche maggiormente utilizzate per valutare la validità di costruito di uno strumento sono i modelli di equazioni strutturali. A titolo illustrativo, in questo lavoro è stato considerato un questionario volto a rilevare il temperamento in adolescenti di età compresa tra i 9 e i 16 anni cioè l'*Early Adolescent Temperament Questionnaire-Revised* (Capaldi e Rothbart, 1992), che è uno dei pochi questionari disponibili per valutare tale costruito durante l'adolescenza. Nonostante l'ampio utilizzo di questa scala, pochi studi ne hanno eseguito un'approfondita valutazione delle proprietà psicometriche, suggerendo spesso risultati non sovrapponibili e rendendo potenzialmente discutibili le conclusioni che potrebbero emergere dalla sua applicazione. Inoltre, non esistono a oggi studi che ne hanno indagato le proprietà psicometriche nella popolazione italiana. Di conseguenza, l'obiettivo principale di questo lavoro è contribuire allo studio delle proprietà psicometriche dell'EATQ-R in un campione di adolescenti italiani. In particolare, tenendo conto dei risultati emersi dalla letteratura, gli obiettivi principali sono stati:

- Eseguire un'analisi approfondita degli item (*item analysis*) attraverso l'analisi delle principali statistiche descrittive e della distribuzione delle risposte agli item;
- Studio preliminare della struttura fattoriale dello strumento attraverso l'applicazione dell'analisi fattoriale confermativa su modelli di primo ordine, al fine di valutare le scale latenti misurate dai relativi item osservati e di discuterne la validità.

3.2 Procedura e metodo

3.2.1 Partecipanti

I dati utilizzati all'interno di questa tesi sono stati raccolti all'interno del progetto "Sviluppo socio-emotivo in bambini di età scolare" (Marci *et al.*, 2020). Il progetto è stato approvato dal Comitato Etico della Scuola di Psicologia dell'Università degli Studi di Padova. Tutti i partecipanti hanno fornito il consenso informato firmato da entrambi i genitori e il consenso verbale prima della somministrazione. Prima del coinvolgimento, ogni partecipante è stato messo a conoscenza della possibilità di ritirarsi dalla ricerca in qualsiasi momento e senza alcuna penalizzazione. I partecipanti hanno compilato durante l'orario scolastico una serie di questionari tra cui una scheda socio-demografica, il *Family Affluence Scale* (FAS; Currie *et al.*, 2008) per valutare lo status socio-economico della famiglia di provenienza e l'*Early Adolescent Temperament Questionnaire-Revised* (Capaldi e Rothbart, 1992) che rappresenta lo strumento target di questo lavoro.

L'analisi è stata eseguita su 397 studenti (51% femmine, 49% maschi, Figura 3.1) di età compresa tra gli 11 e 15 anni (Media = 13.193; Deviazione Standard = 0.661, Figura 3.2), di cui 204 (52%) frequentanti il secondo anno e 193 (48%) il terzo anno (Figura 3.3), provenienti da 4 scuole secondarie di primo grado della provincia di Padova. Due partecipanti non hanno riportato informazioni sul genere e l'età. Dei 397 soggetti che hanno compilato il questionario, la maggior parte ha riportato un livello socio-economico alto (58%) o medio (41%) rilevato attraverso la compilazione del *Family Affluence Scale* (FAS; Currie *et al.*, 2008).

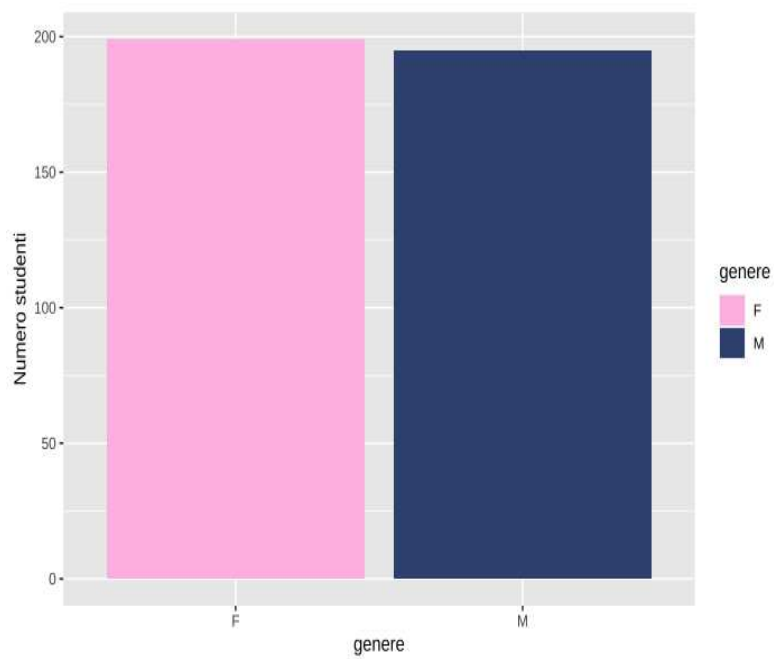


Figura 3.1: Diagramma a barre per le modalità della variabile genere.

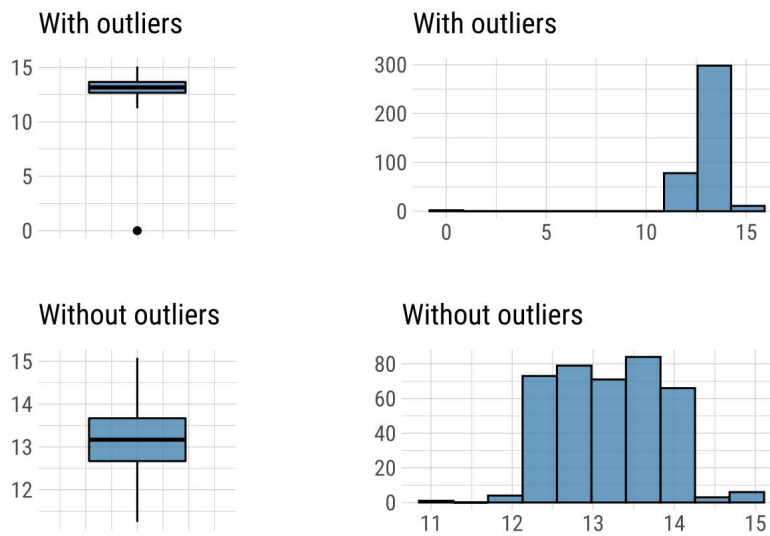


Figura 3.2: Boxplot e istogramma dell'età nei casi di assenza e presenza di un valore anomalo.

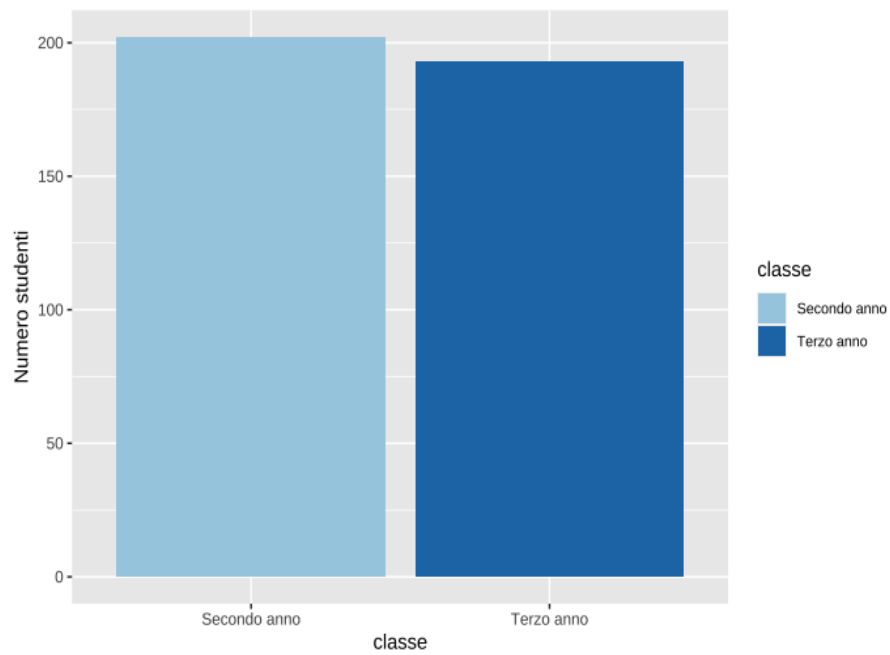


Figura 3.3: Diagramma a barre per la variabile classe (trattata come fattore).

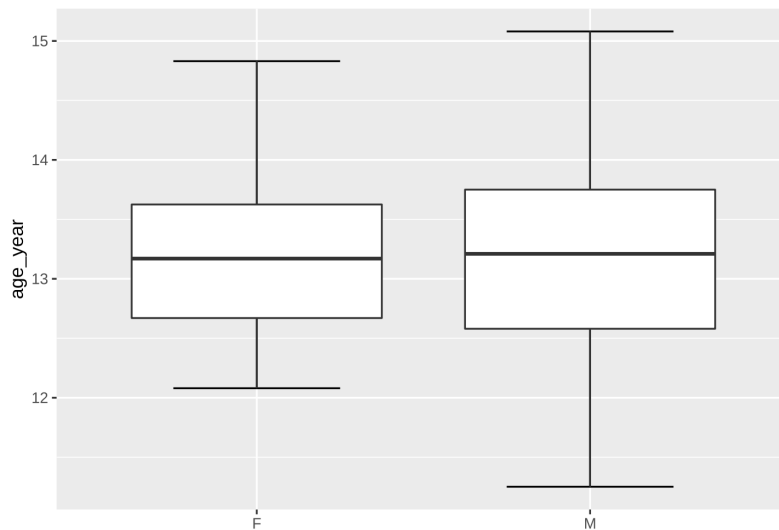


Figura 3.4: Boxplot di età distinta per genere.

3.3 Strumenti

3.3.1 Caratteristiche socio-demografiche

Attraverso la compilazione di una scheda socio-demografica è stato chiesto ai partecipanti di indicare la propria data di nascita e il genere, oltre che fornire risposta alle domande del *Family Affluence Scale* (FAS; Currie *et al.*, 2008). Il FAS è un questionario self-report costituito da quattro domande volto a fornire un indice di benessere economico (*Socioeconomic Status*; SES). Il punteggio ottenuto in ciascun item viene sommato al fine di ottenere un punteggio totale compreso tra 0 e 9. Punteggi compresi fra 0 e 2 sono indicativi di un SES basso, un punteggio compreso fra 3 e 5 si riferisce ad un SES medio e un punteggio da 6 a 9 indica un SES alto. Lo strumento è ampiamente utilizzato a livello internazionale e ha dimostrato buona validità e affidabilità in numerose culture compreso il contesto italiano (Vieno *et al.*, 2009). EATQ è un questionario *self-report* costituito da 64 item che valutano il temperamento in adolescenti di età compresa tra i 9 e i 16 anni. È costituito da 12 scale (fattori di primo livello) che a loro volta costituiscono 4 ampie dimensioni di secondo ordine (macro-fattori): *Effortful Control*, *Surgency*, *Negative Affect*, *Affiliativeness*. In dettaglio:

1. *Effortful Control*: costituito da *Attention* (capacità di focalizzare l'attenzione e di spostarla quando lo si desidera), *Inhibitory Control* (capacità di pianificare e di sopprimere le risposte inappropriate) e *Activation Control* (capacità di compiere un'azione quando c'è una forte tendenza a evitarla);
2. *Surgency*: include *High Intensity Pleasure* (il piacere che deriva da attività ad alta intensità o novità), *Fear* (impatto negativo legato all'interruzione dei compiti in corso o al blocco degli obiettivi.), *Shyness* (inibizione comportamentale alle novità e alle sfide, soprattutto sociali);
3. *Negative Affect*: costituito dai fattori di primo ordine *Frustration* (affetto negativo legato all'interruzione dei compiti in corso o al blocco degli obiettivi), *Depressive Mood* (affetto sgradevole e abbassamento dell'umore, perdita di piacere e di interesse per le attività) e *Aggression* (azioni ostili e aggressive, compresa la violenza fisica diretta a persone e oggetti, l'aggressione verbale diretta e indiretta e la reattività ostile);

4. *Affiliativeness*: include i fattori di primo ordine *Affiliation* (il desiderio di calore e vicinanza con gli altri, indipendentemente dalla timidezza o dall'estroversione), *Perceptual Sensitivity* (rilevamento o consapevolezza percettiva di stimoli lievi e di bassa intensità nell'ambiente), *Pleasure Sensitivity* (quantità di piacere legata ad attività o stimoli che comportano bassa intensità, velocità, complessità, novità e incongruenza).

A loro volta, i fattori di secondo ordine confluiscono in un fattore generale, ossia il temperamento. Ogni item è formulato su una scala Likert a 5 punti che va da 1 (per niente vero per te/per tuo figlio) a 5 (molto vero per me). Alcuni item prevedono una ricodifica (*item reversed*) prima dello scoring (si veda Appendice A). Per ciascuna sottoscala (fattori di primo ordine) e macroscale (fattori di secondo livello) è stato calcolato un punteggio dato dalla media del punteggio ottenuto negli item, e viene inoltre calcolato un punteggio generale di temperamento.

3.4 Analisi dei dati

L'analisi dei dati è stata effettuata con il software R (<https://www.r-project.org/>). Sono state condotte delle analisi descrittive preliminari che includono il calcolo della media, deviazione standard, range, simmetria e curtosi di ciascun item. Successivamente, sono stati definiti i modelli CFA separatamente per ogni sottoscala (modelli di primo ordine). Tutti i modelli sono stati stimati utilizzando lo stimatore WLSMV (*Weighted Least Square Mean and Variance Adjusted*), versione modificata dello stimatore (2.16) raccomandato per i dati su scala Likert (Flora e Curran, 2004). I dati mancanti (<.1%) sono stati gestiti utilizzando il metodo di *pairwise* (Katsikatsou *et al.*, 2012), che permette di considerare anche le unità statistiche le quali presentano valori mancanti solo per alcune variabili. I modelli e gli item sono stati attentamente valutati per verificare i loading nel fattore corrispondente. Data l'ampia dimensione del campione, ci si aspettava che, anche un item con bassa dimensionalità (loading basso), potesse essere significativo: pertanto, al fine di fornirne una valutazione descrittiva, si è considerato come accettabile un loading pari o superiore a 0.30 (Kline, 2015). L'adattamento dei modelli è stato valutato tenendo conto degli indici di fit presentati nel Capitolo 2 e con i trade-off della Tabella 2.5.

3.5 Risultati

3.5.1 Analisi descrittive

Per le risposte date agli item di ogni scala, si fornisce una tabella riassuntiva delle frequenze relative e degli indici descrittivi ottenuti dalla funzione *describe*, insieme a un diagramma a barre per le frequenze relative. La valutazione descrittiva delle scale è effettuata considerando l'asimmetria e la media, e si vuole verificare se gli item presentano distribuzione Normale: i valori che hanno probabilità minore di essere scelti sono quelli estremi. La scala Likert funziona bene se tutti i valori vengono scelti almeno una volta e, nel caso in cui la distribuzione non fosse Normale, sarà necessario indagare eventuali fattori di distorsione. L'asimmetria assume valore zero quando la distribuzione è perfettamente Normale, quindi valori lontani da zero indicano una asimmetria a destra (maggiori di zero) o a sinistra (minori di zero).

Considerando la media come un indice della difficoltà/affettività, è possibile stabilire un intervallo intorno al valore medio teorico che definisce un livello di difficoltà/affettività adeguato tale da includere item che non siano nè troppo difficili (nessuno tende a essere in accordo con il contenuto) e nè troppo facili (tutti tendono a essere in accordo con il contenuto). Nel caso di una scala Likert con punteggi da 1 a 5, la media degli item deve variare in un intervallo che va da 2 a 4.

Activation Control

Le medie riportano valori compresi tra 2.58 e 3.54. Per 4 item su 5 si ha una distribuzione delle risposte leggermente asimmetrica (ac_1, ac_3, ac_4, ac_5), mentre per l'item ac_2 l'asimmetria è più evidente. Nel complesso la scala può ritenersi accettabile.

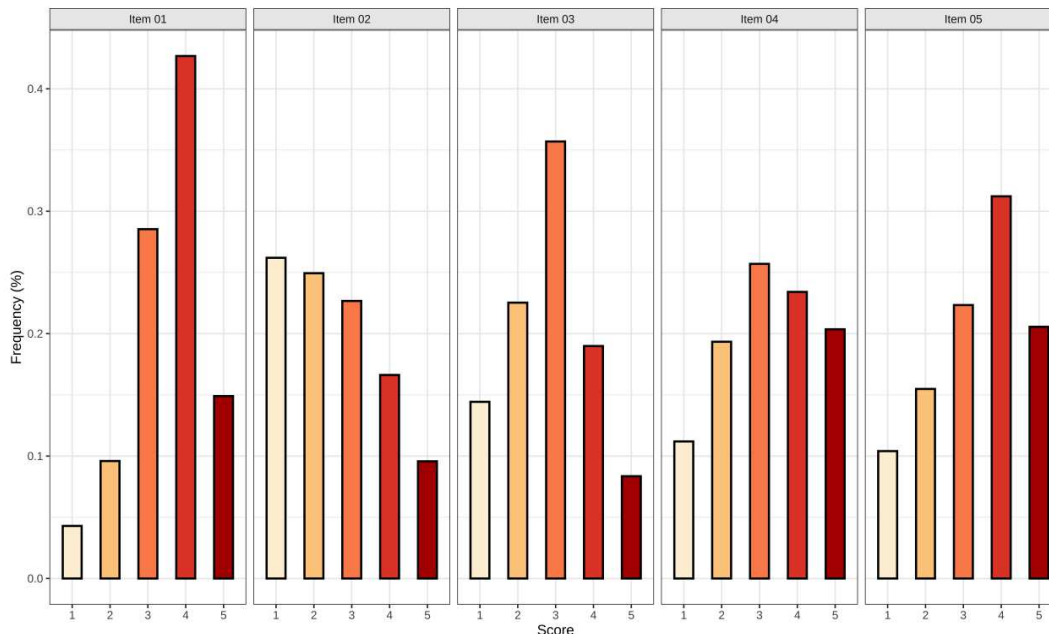


Figura 3.5: Diagramma a barre della distribuzione degli item di Activation Control.

	Mod. 1	Mod. 2	Mod. 3	Mod. 4	Mod. 5	n	mean	sd	median	min	max	skew	kurtosis
ac_1	0.04	0.10	0.29	0.43	0.15	396	3.54	1	4	1	5	-0.60	0.07
ac_2	0.26	0.25	0.23	0.17	0.10	397	2.58	1.30	2	1	5	0.34	-1.01
ac_3	0.14	0.23	0.36	0.19	0.08	395	2.84	1.14	3	1	5	0.06	-0.69
ac_4	0.11	0.19	0.26	0.23	0.20	393	3.22	1.28	3	1	5	-0.16	-1.04
ac_5	0.10	0.15	0.22	0.31	0.21	394	3.36	1.26	4	1	5	-0.39	-0.88

Attention

Le medie riportano valori compresi tra 2.93 e 3.86. Per 4 item su 6 si ha una distribuzione delle risposte leggermente asimmetrica ($att_1, att_3, att_4, att_5$), mentre per gli item att_2 e att_6 l'asimmetria è più evidente. Nel complesso la scala può ritenersi accettabile.

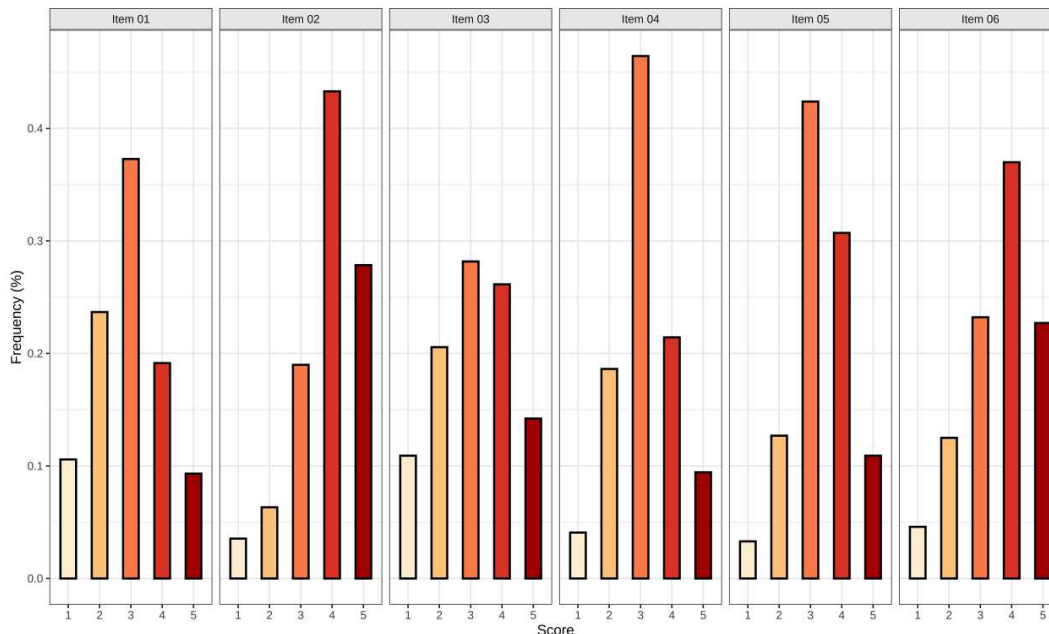


Figura 3.6: Diagramma a barre della distribuzione degli item di Attention.

	Mod. 1	Mod. 2	Mod. 3	Mod. 4	Mod. 5	n	mean	sd	median	min	max	skew	kurtosis
att_1	0.11	0.24	0.37	0.19	0.09	397	2.93	1.11	3	1	5	0.08	-0.59
att_2	0.04	0.06	0.19	0.43	0.28	395	3.86	1.01	4	1	5	-0.90	0.53
att_3	0.11	0.21	0.28	0.26	0.14	394	3.12	1.21	3	1	5	-0.12	-0.91
att_4	0.04	0.19	0.46	0.21	0.09	392	3.14	0.96	3	1	5	0.09	-0.18
att_5	0.03	0.13	0.42	0.31	0.11	394	3.33	0.95	3	1	5	-0.16	-0.14
att_6	0.05	0.12	0.23	0.37	0.23	392	3.61	1.11	4	1	5	-0.55	-0.42

Inhibitory Control

Le medie riportano valori compresi tra 0.94 e 1.29. Per 2 item su 5 si ha una distribuzione delle risposte leggermente asimmetrica (ic_3, ic_5), mentre per i rimanenti item l'asimmetria è più evidente. Nel complesso la scala non è soddisfacente.

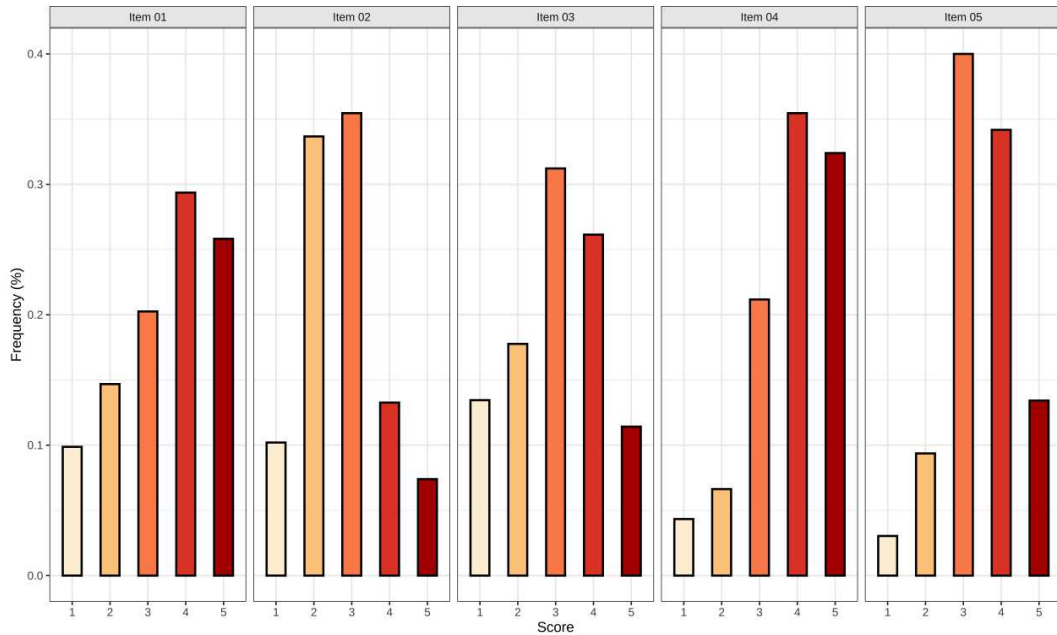


Figura 3.7: Diagramma a barre della distribuzione degli item di Inhibitory Control.

	Mod. 1	Mod. 2	Mod. 3	Mod. 4	Mod. 5	n	mean	sd	median	min	max	skew	kurtosis
ic_1	0.10	0.15	0.20	0.29	0.26	395	3.47	1.29	4	1	5	-0.46	-0.89
ic_2	0.10	0.34	0.35	0.13	0.07	392	2.74	1.05	3	1	5	0.39	-0.31
ic_3	0.13	0.18	0.31	0.26	0.11	394	3.04	1.20	3	1	5	-0.15	-0.84
ic_4	0.04	0.07	0.21	0.35	0.32	392	3.85	1.08	4	1	5	-0.83	0.14
ic_5	0.03	0.09	0.40	0.34	0.13	395	3.46	0.94	3	1	5	-0.27	-0.04

Affiliation

La media riporta un valore compreso tra 3.30 e 4.21, e quindi va oltre l'estremo superiore fissato di 4. Per 2 item su 5 si ha una distribuzione delle risposte leggermente asimmetrica (afl_1, afl_5), tuttavia per i rimanenti item l'asimmetria è più evidente. Dunque, la scala non è ritenuta soddisfacente.

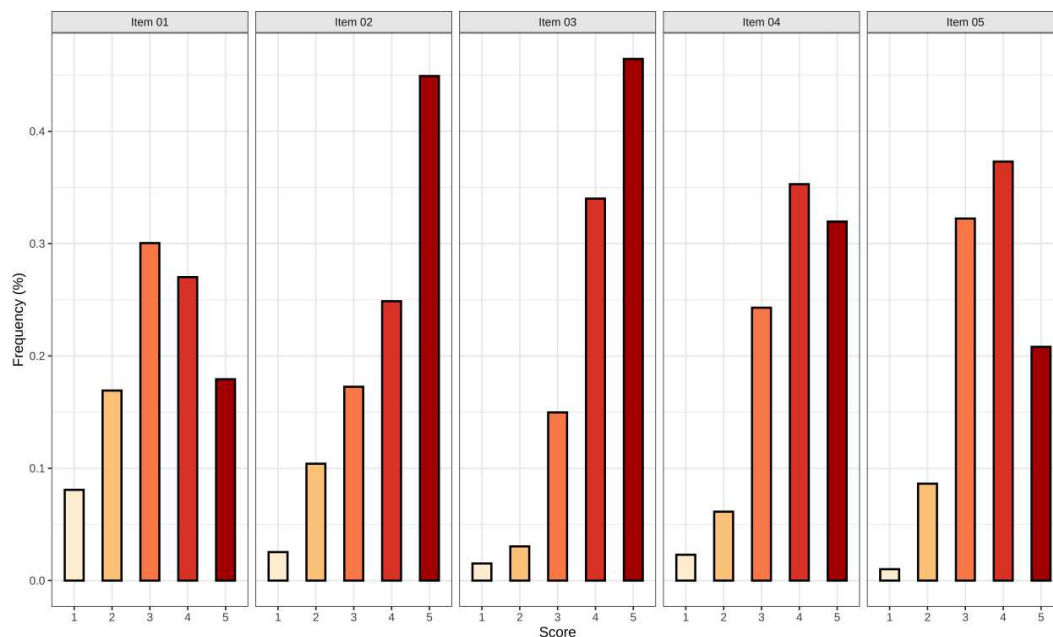


Figura 3.8: Diagramma a barre della distribuzione degli item di Affiliation.

	Mod. 1	Mod. 2	Mod. 3	Mod. 4	Mod. 5	n	mean	sd	median	min	max	skew	kurtosis
af1_1	0.08	0.17	0.30	0.27	0.18	396	3.30	1.18	3	1	5	-0.23	-0.79
af1_2	0.03	0.10	0.17	0.25	0.45	394	3.99	1.13	4	1	5	-0.85	-0.33
af1_3	0.02	0.03	0.15	0.34	0.46	394	4.21	0.91	4	1	5	-1.14	1.10
af1_4	0.02	0.06	0.24	0.35	0.32	391	3.88	1	4	1	5	-0.68	-0.02
af1_5	0.01	0.09	0.32	0.37	0.21	394	3.68	0.93	4	1	5	-0.27	-0.46

Fear

Le medie riportano valori compresi tra 2 e 3.34. Per 4 item su 6 si ha una distribuzione delle risposte leggermente asimmetrica ($fear_2$, $fear_3$, $fear_4$, $fear_5$), e la scala può essere ritenuta accettabile.

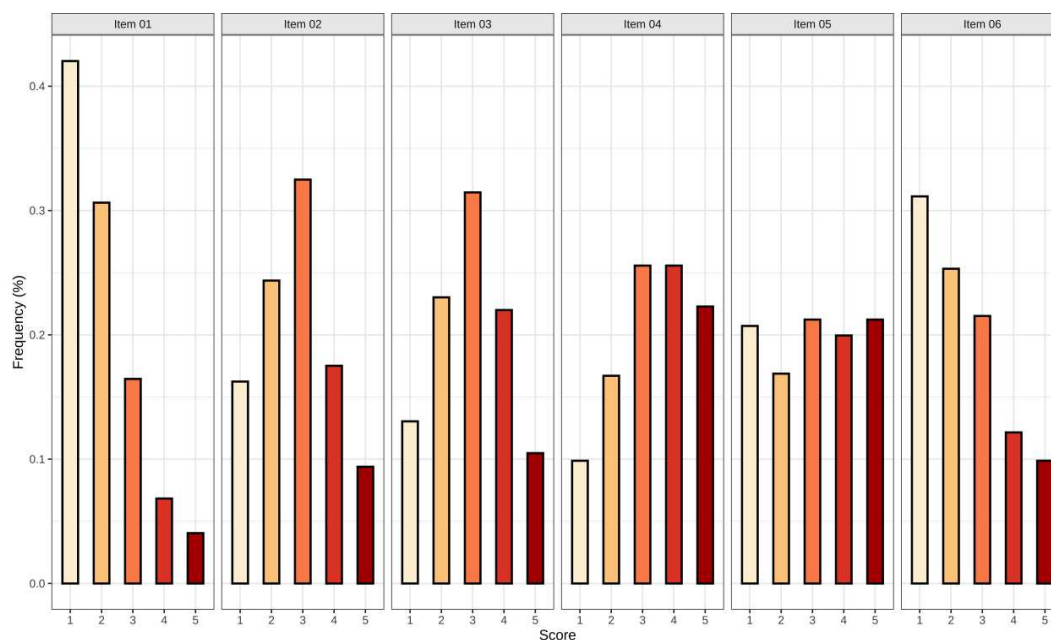


Figura 3.9: Diagramma a barre della distribuzione degli item di Fear.

	Mod. 1	Mod. 2	Mod. 3	Mod. 4	Mod. 5	n	mean	sd	median	min	max	skew	kurtosis
fear_1	0.42	0.31	0.16	0.07	0.04	395	2	1.11	2	1	5	1.01	0.29
fear_2	0.16	0.24	0.32	0.18	0.09	394	2.79	1.19	3	1	5	0.15	-0.79
fear_3	0.13	0.23	0.31	0.22	0.10	391	2.94	1.18	3	1	5	0.02	-0.83
fear_4	0.10	0.17	0.26	0.26	0.22	395	3.34	1.26	3	1	5	-0.28	-0.95
fear_5	0.21	0.17	0.21	0.20	0.21	391	3.04	1.43	3	1	5	-0.06	-1.31
fear_6	0.31	0.25	0.22	0.12	0.10	395	2.44	1.31	2	1	5	0.53	-0.84

Frustration

Le medie riportano valori compresi tra 2.32 e 3.78. Per 5 item su 7 si ha una distribuzione delle risposte leggermente asimmetrica ($frus_1, frus_2, frus_3, frus_4, frus_7$), e la scala è dunque accettabile.

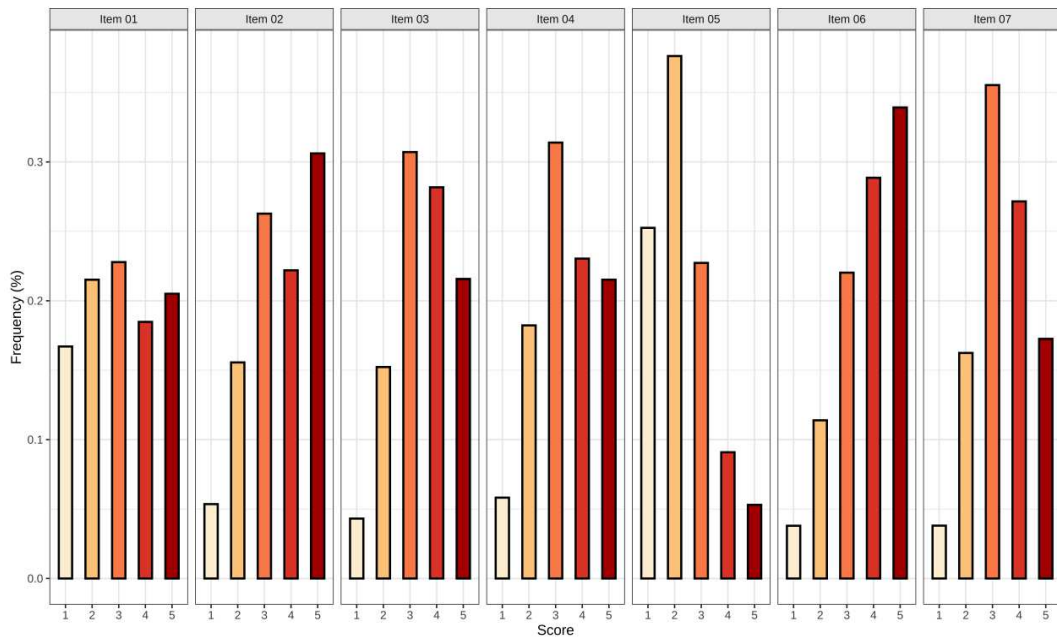


Figura 3.10: Diagramma a barre della distribuzione degli item di Frustration.

	Mod. 1	Mod. 2	Mod. 3	Mod. 4	Mod. 5	n	mean	sd	median	min	max	skew	kurtosis
frus_1	0.17	0.22	0.23	0.18	0.21	395	3.05	1.38	3	1	5	0.01	-1.23
frus_2	0.05	0.16	0.26	0.22	0.31	392	3.57	1.22	4	1	5	-0.36	-0.94
frus_3	0.04	0.15	0.31	0.28	0.22	394	3.47	1.12	3	1	5	-0.26	-0.72
frus_4	0.06	0.18	0.31	0.23	0.22	395	3.36	1.17	3	1	5	-0.15	-0.88
frus_5	0.25	0.38	0.23	0.09	0.05	396	2.32	1.11	2	1	5	0.70	-0.13
frus_6	0.04	0.11	0.22	0.29	0.34	395	3.78	1.14	4	1	5	-0.62	-0.53
frus_7	0.04	0.16	0.36	0.27	0.17	394	3.38	1.07	3	1	5	-0.13	-0.64

High Intensity Pleasure/Surgency

Le medie riportano valori compresi tra 2.30 e 3.70. Per 4 item su 6 si ha una distribuzione delle risposte leggermente asimmetrica ($Sur_2, Sur_4, Sur_5, Sur_6$), e quindi la scala risulta essere accettabile.

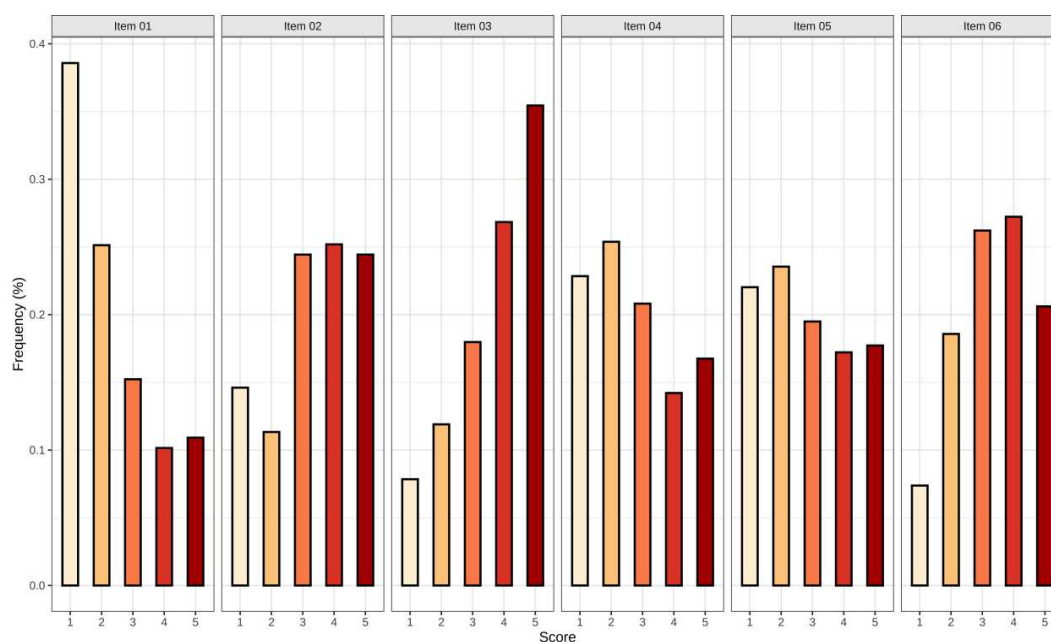


Figura 3.11: Diagramma a barre della distribuzione degli item di High Intensity Pleasure/Surgency.

	Mod. 1	Mod. 2	Mod. 3	Mod. 4	Mod. 5	n	mean	sd	median	min	max	skew	kurtosis
Sur_1	0.39	0.25	0.15	0.10	0.11	394	2.30	1.36	2	1	5	0.74	-0.70
Sur_2	0.15	0.11	0.24	0.25	0.24	397	3.34	1.35	3	1	5	-0.38	-1
Sur_3	0.08	0.12	0.18	0.27	0.35	395	3.70	1.28	4	1	5	-0.68	-0.66
Sur_4	0.23	0.25	0.21	0.14	0.17	394	2.77	1.39	3	1	5	0.28	-1.17
Sur_5	0.22	0.24	0.19	0.17	0.18	395	2.85	1.41	3	1	5	0.17	-1.27
Sur_6	0.07	0.19	0.26	0.27	0.21	393	3.35	1.21	3	1	5	-0.24	-0.91

Pleasure Sensitivity

Le medie riportano valori compresi tra 2.85 e 3.19. Tutti gli item presentano una leggera asimmetria (valori molto vicini allo 0). La scala risulta essere ottima.

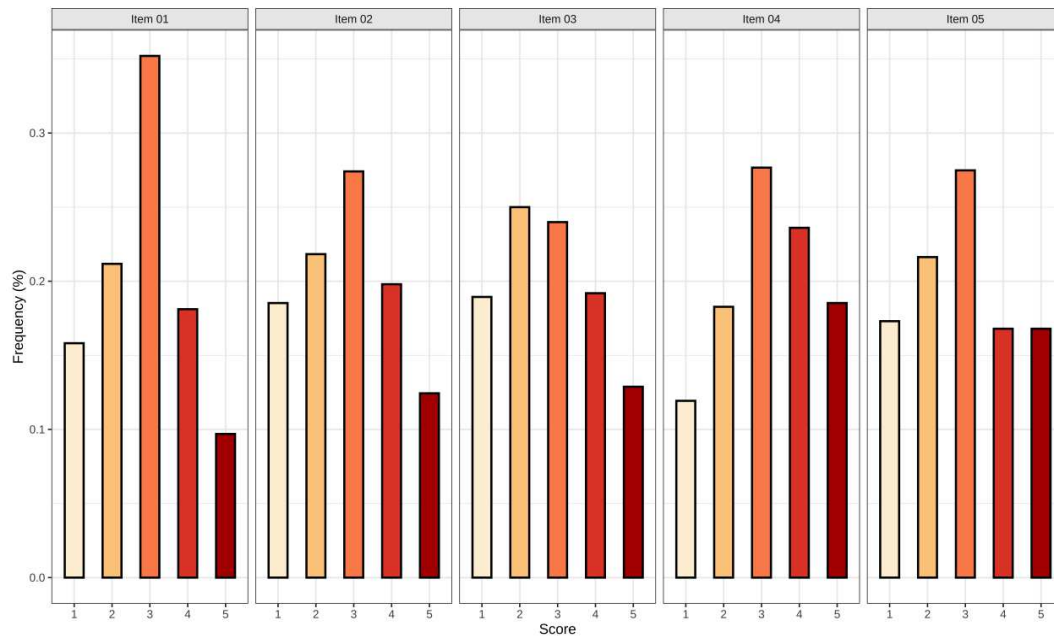


Figura 3.12: Diagramma a barre della distribuzione degli item di Pleasure Sensitivity.

	Mod. 1	Mod. 2	Mod. 3	Mod. 4	Mod. 5	n	mean	sd	median	min	max	skew	kurtosis
pls_1	0.16	0.21	0.35	0.18	0.10	392	2.85	1.18	3	1	5	0.07	-0.76
pls_2	0.19	0.22	0.27	0.20	0.12	394	2.86	1.28	3	1	5	0.09	-1.03
pls_3	0.19	0.25	0.24	0.19	0.13	396	2.82	1.30	3	1	5	0.17	-1.08
pls_4	0.12	0.18	0.28	0.24	0.19	394	3.19	1.27	3	1	5	-0.16	-0.99
pls_5	0.17	0.22	0.27	0.17	0.17	393	2.94	1.32	3	1	5	0.09	-1.09

Perceptual Sensitivity

Le medie riportano valori compresi tra 2.32 e 3.80. Per ps_1, ps_2 si ha una distribuzione delle risposte con asimmetria leggera, e per ps_3, ps_4 si ha invece un'asimmetria più forte ma non troppo alta. Nel complesso, la scala si può ritenere accettabile.

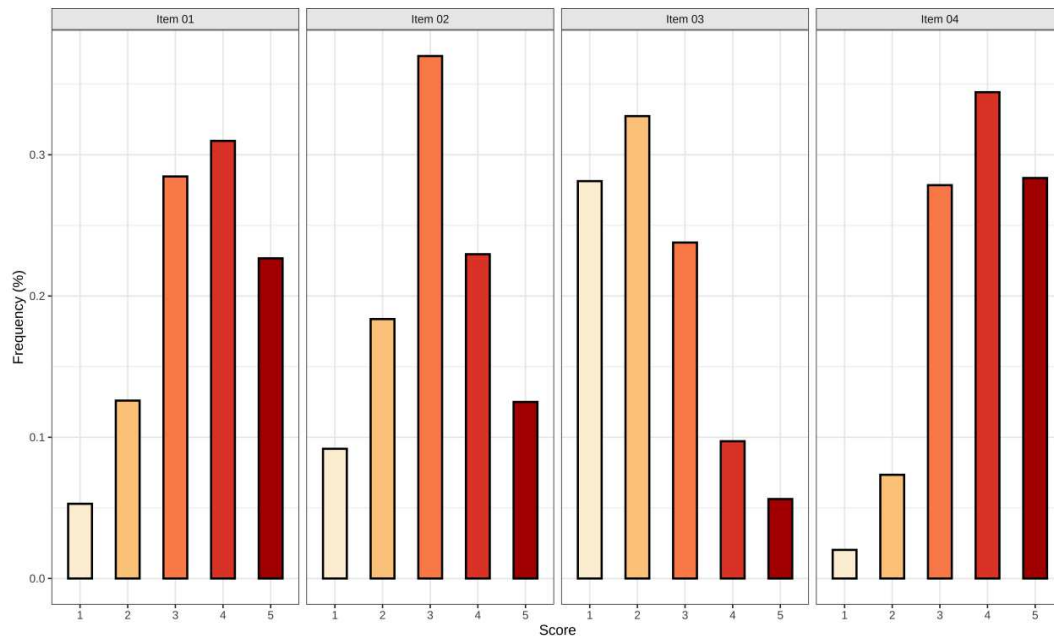


Figura 3.13: Diagramma a barre della distribuzione degli item di Perceptual Sensitivity.

	Mod. 1	Mod. 2	Mod. 3	Mod. 4	Mod. 5	n	mean	sd	median	min	max	skew	kurtosis
ps_1	0.05	0.13	0.28	0.31	0.23	397	3.53	1.13	4	1	5	-0.42	-0.57
ps_2	0.09	0.18	0.37	0.23	0.12	392	3.11	1.13	3	1	5	-0.08	-0.62
ps_3	0.28	0.33	0.24	0.10	0.06	391	2.32	1.15	2	1	5	0.64	-0.35
ps_4	0.02	0.07	0.28	0.34	0.28	395	3.80	1	4	1	5	-0.52	-0.28

Shyness

Le medie riportano valori compresi tra 2.16 e 3.22. Solo per l'item Shy_1 l'asimmetria è più evidente, dunque la scala può essere ritenuta accettabile.

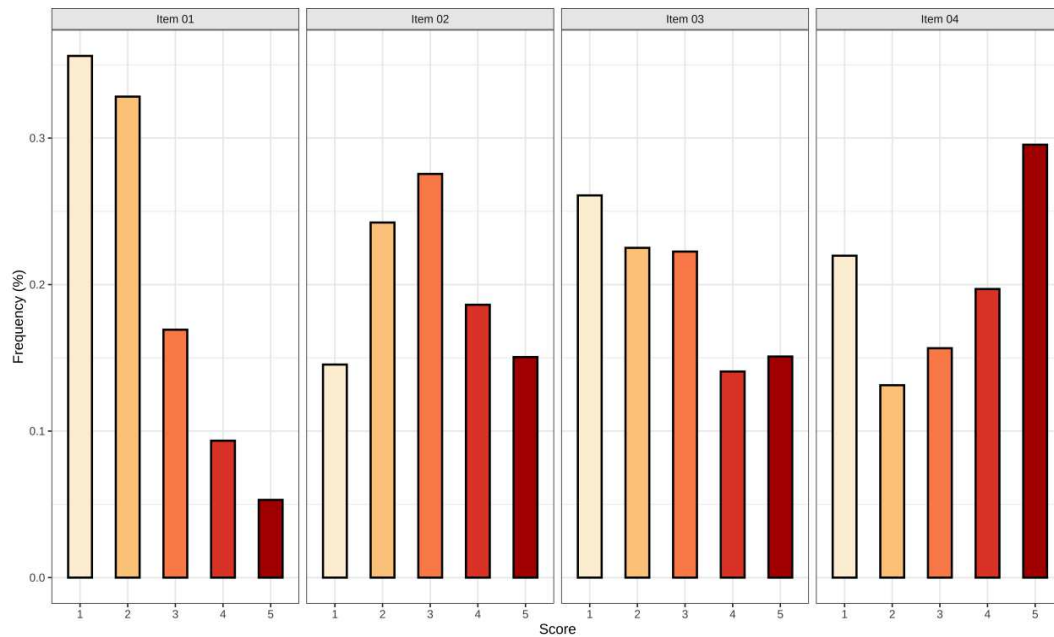


Figura 3.14: Diagramma a barre della distribuzione degli item di Shyness.

	Mod. 1	Mod. 2	Mod. 3	Mod. 4	Mod. 5	n	mean	sd	median	min	max	skew	kurtosis
Shy_1	0.36	0.33	0.17	0.09	0.05	396	2.16	1.16	2	1	5	0.85	-0.14
Shy_2	0.15	0.24	0.28	0.19	0.15	392	2.95	1.27	3	1	5	0.10	-1.01
Shy_3	0.26	0.23	0.22	0.14	0.15	391	2.70	1.39	3	1	5	0.31	-1.15
Shy_4	0.22	0.13	0.16	0.20	0.30	396	3.22	1.53	3	1	5	-0.24	-1.43

Aggression

Le medie riportano valori compresi tra 1.06 e 1.46, e non sono compresi nell'intervallo di range fissato da 2 a 4. Per 4 item su 6 si ha una distribuzione delle risposte leggermente asimmetrica (ag_1, ag_3, ag_4, ag_5), mentre vi è un'asimmetria più evidente per ag_2 e ag_6 . La scala non può essere ritenuta accettabile.

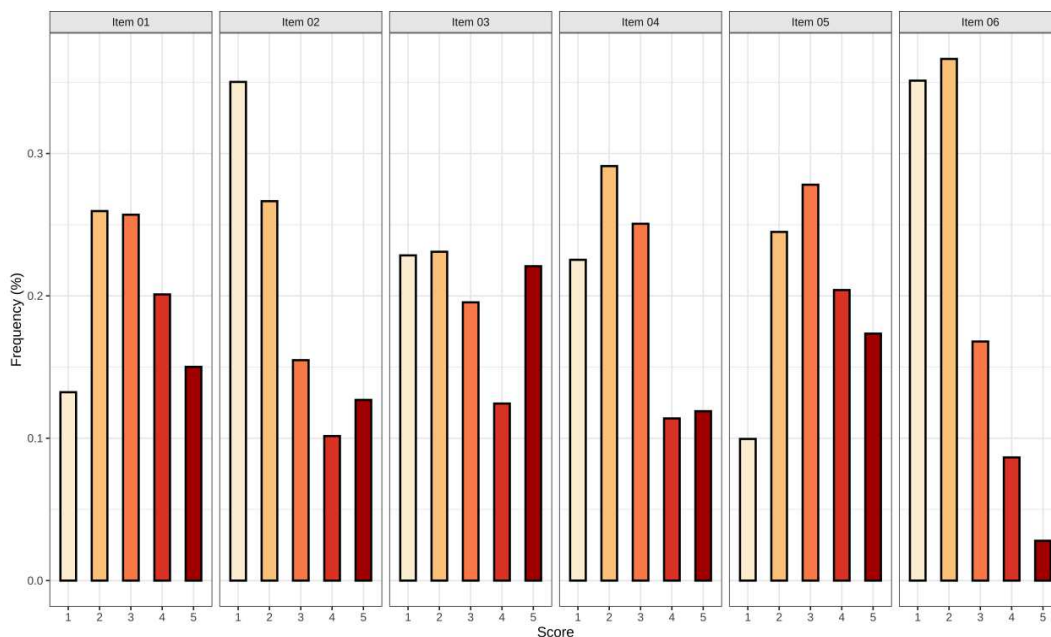


Figura 3.15: Diagramma a barre della distribuzione degli item di Aggression.

	Mod. 1	Mod. 2	Mod. 3	Mod. 4	Mod. 5	n	mean	sd	median	min	max	skew	kurtosis
ag_1	0.13	0.26	0.26	0.20	0.15	393	2.98	1.26	3	1	5	0.10	-1.03
ag_2	0.35	0.27	0.15	0.10	0.13	394	2.39	1.38	2	1	5	0.67	-0.82
ag_3	0.23	0.23	0.20	0.12	0.22	394	2.88	1.46	3	1	5	0.20	-1.33
ag_4	0.23	0.29	0.25	0.11	0.12	395	2.61	1.28	2	1	5	0.45	-0.80
ag_5	0.10	0.24	0.28	0.20	0.17	392	3.11	1.24	3	1	5	0.03	-1.01
ag_6	0.35	0.37	0.17	0.09	0.03	393	2.07	1.06	2	1	5	0.87	0.09

Depression Mood

La media riporta un valore compreso tra 1.88 e 3.59, e va oltre l'estremo inferiore fissato. Per 4 item su 6 si ha una distribuzione delle risposte leggermente asimmetrica ($dep_1, dep_3, dep_4, dep_5$), mentre è molto evidente per l'item dep_6 . Nel complesso la scala può essere ritenuta accettabile, seppur potrebbe essere necessario riformularla.

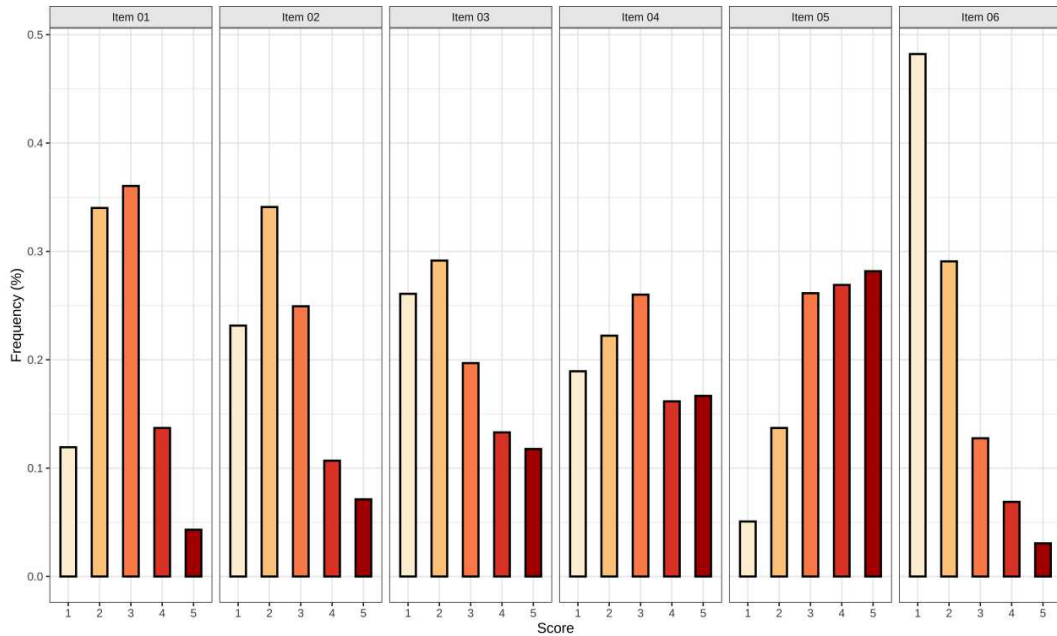


Figura 3.16: Diagramma a barre della distribuzione degli item di Depressive Mood.

	Mod. 1	Mod. 2	Mod. 3	Mod. 4	Mod. 5	n	mean	sd	median	min	max	skew	kurtosis
dep_1	0.12	0.34	0.36	0.14	0.04	394	2.64	1	3	1	5	0.30	-0.29
dep_2	0.23	0.34	0.25	0.11	0.07	393	2.45	1.16	2	1	5	0.57	-0.44
dep_3	0.26	0.29	0.20	0.13	0.12	391	2.55	1.32	2	1	5	0.48	-0.92
dep_4	0.19	0.22	0.26	0.16	0.17	396	2.89	1.34	3	1	5	0.14	-1.12
dep_5	0.05	0.14	0.26	0.27	0.28	394	3.59	1.18	4	1	5	-0.43	-0.75
dep_6	0.48	0.29	0.13	0.07	0.03	392	1.88	1.07	2	1	5	1.18	0.65

3.5.2 Struttura fattoriale

Vengono di seguito presentati i modelli CFA di primo ordine, di cui è possibile trovare la specificazione in R in Appendice B.

Per ognuno dei modelli, viene fornita la rappresentazione grafica nella quale vengono evidenziati i loadings e una tabella che presenta le stime dei loadings, gli standard error, lo z-value e il p-value. Si ricorda che le stime dei loadings verranno considerate soddisfacenti se risultano maggiori della soglia fissata di 0.3, mentre gli indici di fit robusti, presentati successivamente nella Figura 3.5.3, saranno interpretati secondo i trade-off di Schermelleh-Engel *et al.* (2003), definiti nella Figura 2.5.

Activation Control

Tutti i loadings riportano delle stime soddisfacenti. Gli indici di fit che mostrano un buon adattamento sono solo GFI (0.994) e SRMR (0.050).

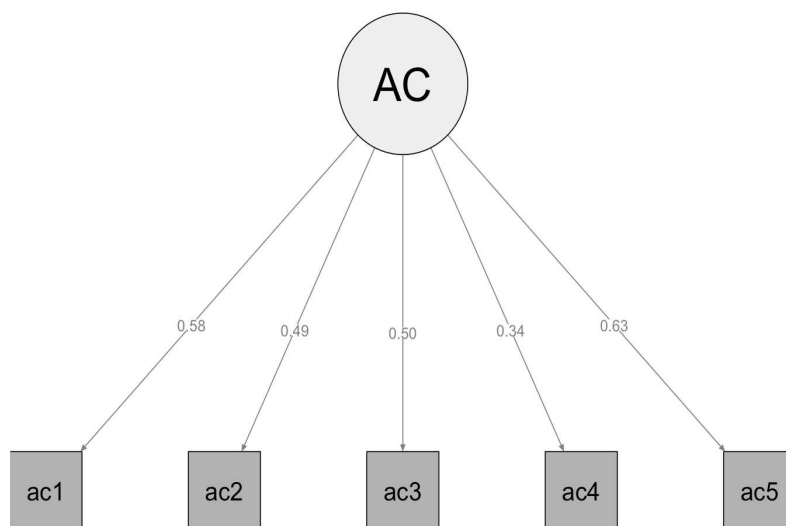


Figura 3.17: Modello CFA di primo ordine di Activation Control.

Variabili	loadings	Std. Error	z-value	p-value
ac_1	0.585	0.055	10.623	0.000
ac_2	0.488	0.054	9.103	0.000
ac_3	0.499	0.054	9.231	0.000
ac_4	0.336	0.058	5.748	0.000
ac_5	0.631	0.055	11.578	0.000

Tabella 3.1: Risultati per il modello della scala AC

Affiliation

Tutte le stime dei loadings sono soddisfacenti. Gli indici di fit indicano un buon adattamento per CFI (0.976), GFI (0.997) e SRMR (0.038), mentre si ha un adattamento accettabile per TLI (0.952) e RMSEA (0.079).

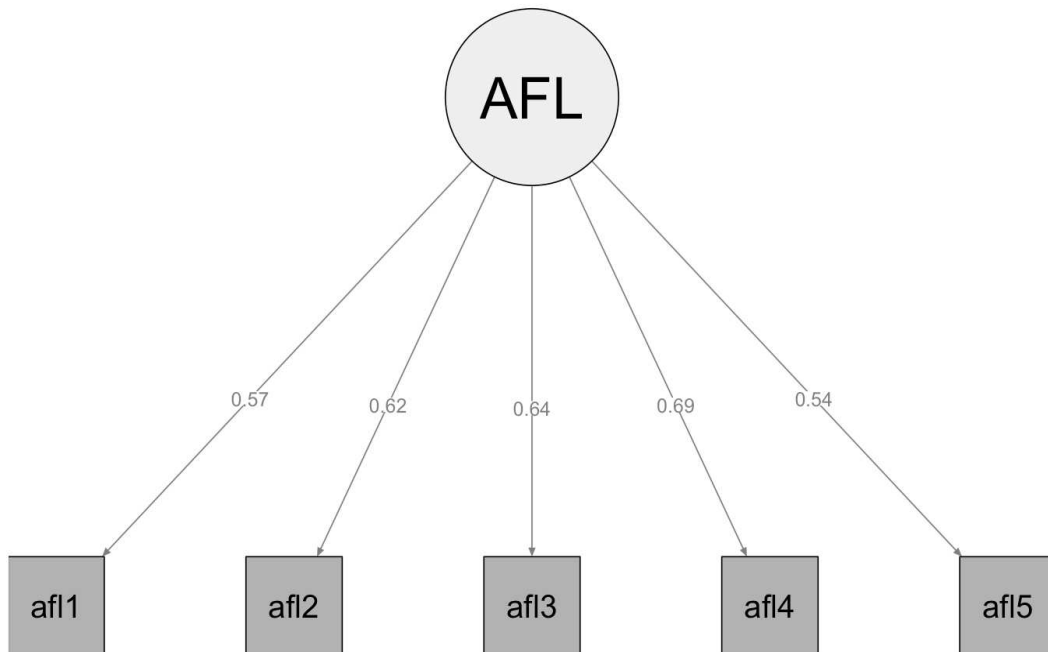


Figura 3.18: Modello CFA di primo ordine di Affiliation.

Variabili	loadings	Std. Error	z-value	p-value
<i>afl</i> ₁	0.566	0.044	12.797	0.000
<i>afl</i> ₂	0.619	0.046	13.321	0.000
<i>afl</i> ₃	0.637	0.042	15.003	0.000
<i>afl</i> ₄	0.688	0.044	15.770	0.000
<i>afl</i> ₅	0.540	0.047	11.386	0.000

Tabella 3.2: Risultati per il modello della scala AFL

Attention

Tutte le stime dei loadings sono soddisfacenti. L'unico indice che mostra un buon adattamento è il GFI (0.989), mentre SRMR (0.066) evidenzia un adattamento accettabile.

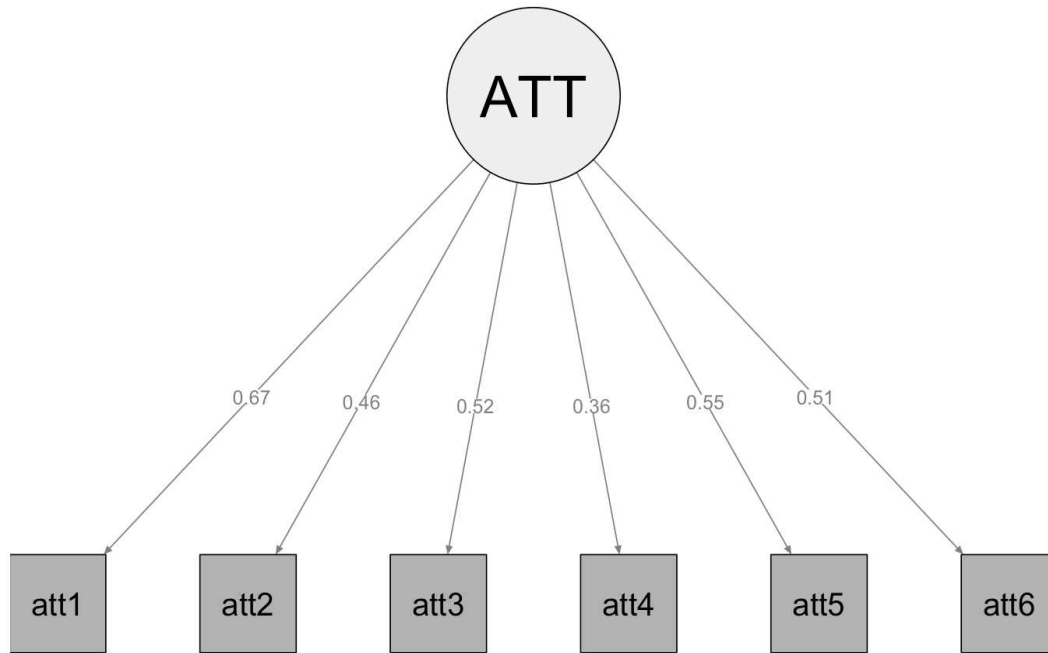


Figura 3.19: Modello CFA di primo ordine di Attention.

Variabili	loadings	Std. Error	z-value	p-value
<i>att</i> ₁	0.669	0.045	14.703	0.000
<i>att</i> ₂	0.459	0.049	9.348	0.000
<i>att</i> ₃	0.517	0.047	11.085	0.000
<i>att</i> ₄	0.364	0.050	7.337	0.000
<i>att</i> ₅	0.552	0.047	11.668	0.000
<i>att</i> ₆	0.508	0.047	10.800	0.000

Tabella 3.3: Risultati per il modello della scala ATT

Fear

Tutte le stime dei loadings sono soddisfacenti. L'unico indice che mostra un buon adattamento è il GFI (0.884), mentre SRMR (0.062) evidenzia un adattamento accettabile.

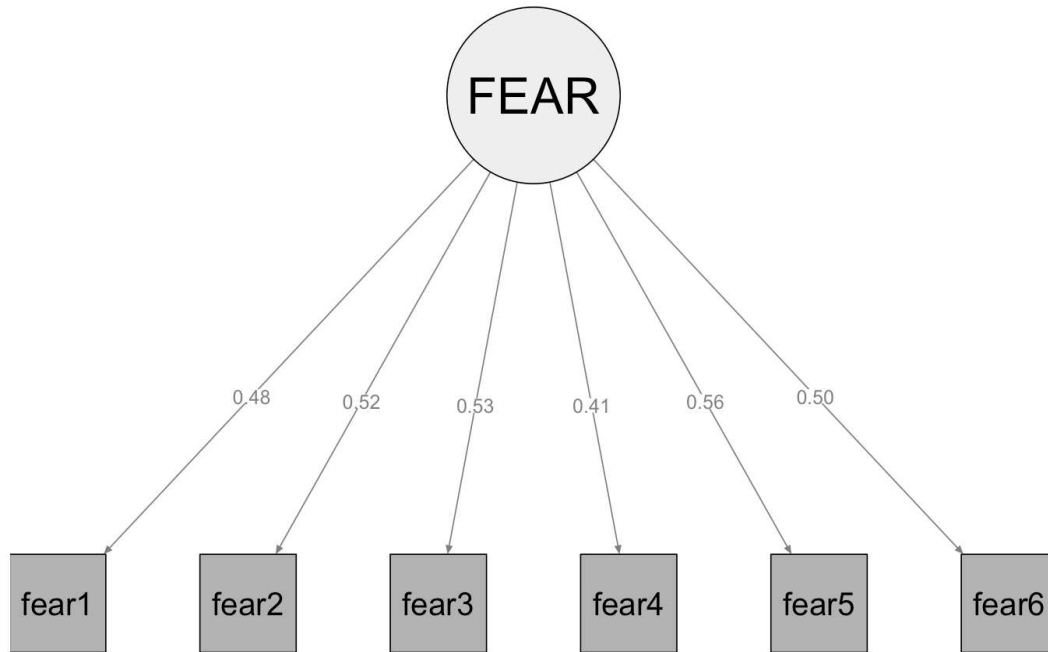


Figura 3.20: Modello CFA di primo ordine di Fear.

Variabili	loadings	Std. Error	z-value	p-value
<i>fear</i> ₁	0.475	0.055	8.715	0.000
<i>fear</i> ₂	0.517	0.054	9.646	0.000
<i>fear</i> ₃	0.534	0.052	10.194	0.000
<i>fear</i> ₄	0.414	0.061	6.814	0.000
<i>fear</i> ₅	0.556	0.054	10.362	0.000
<i>fear</i> ₆	0.499	0.050	9.921	0.000

Tabella 3.4: Risultati per il modello della scala FEAR

Frustration

Le stime dei loadings sono soddisfacenti, ad eccezione di $frus_5$ (0.264). Si ha un buon adattamento per GFI (0.983) e per SRMR (0.007).

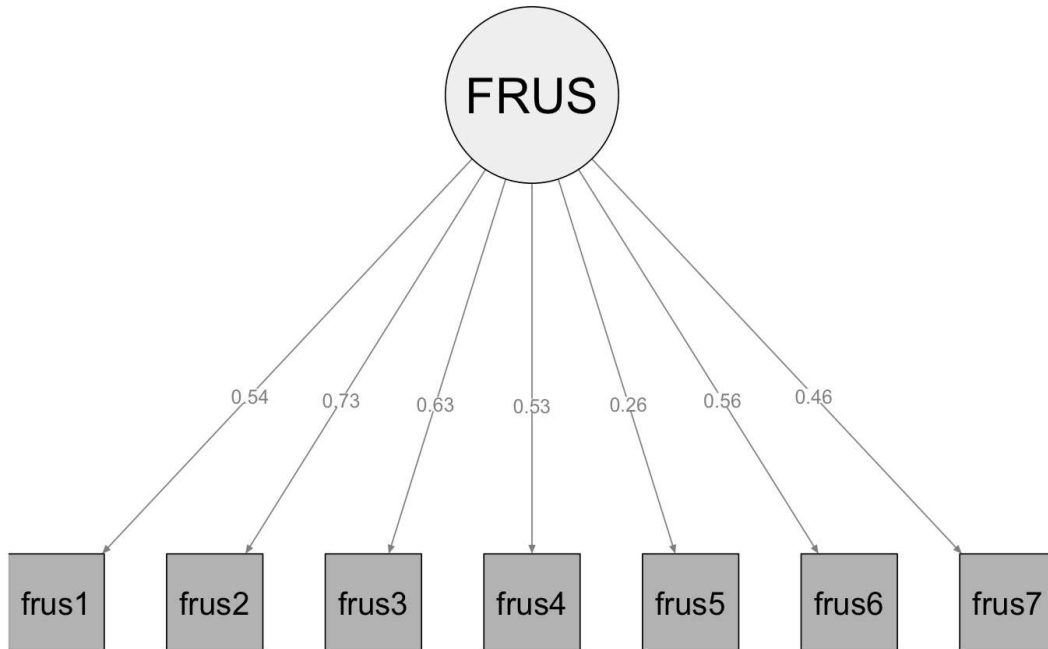


Figura 3.21: Modello CFA di primo ordine di Frustration.

Variabili	loadings	Std. Error	z-value	p-value
$frus_1$	0.545	0.043	12.630	0.000
$frus_2$	0.735	0.036	20.135	0.000
$frus_3$	0.629	0.037	17.182	0.000
$frus_4$	0.527	0.043	12.158	0.000
$frus_5$	0.264	0.051	5.168	0.000
$frus_6$	0.556	0.043	13.060	0.000
$frus_7$	0.463	0.045	10.380	0.000

Tabella 3.5: Risultati per il modello della scala FRUS

High Intensity Pleasure/Surgency

Le stime dei loadings non sono soddisfacenti per gli item Sur_1 (0.230), Sur_2 (0.187). Si ha un buon adattamento per GFI (0.983) e per SRMR (0.008).

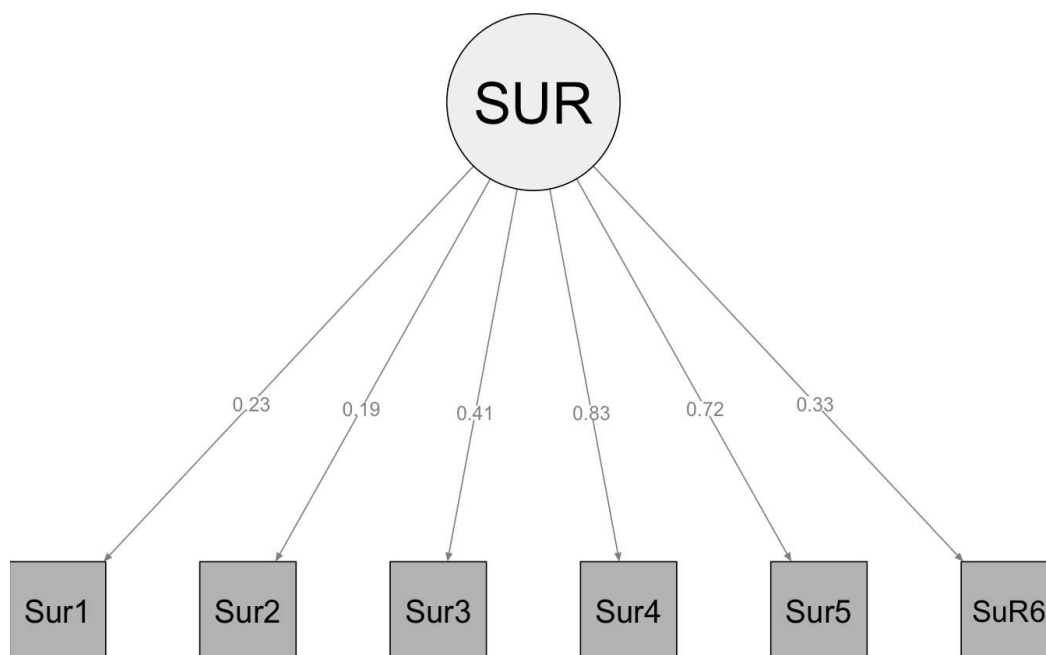


Figura 3.22: Modello CFA di primo ordine di High Intensity Pleasure/Surgency.

Variabili	loadings	Std. Error	z-value	p-value
Sur_1	0.230	0.059	3.883	0.000
Sur_2	0.187	0.056	3.359	0.000
Sur_3	0.406	0.052	7.794	0.000
Sur_4	0.827	0.051	16.319	0.000
Sur_5	0.724	0.050	14.600	0.000
Sur_6	0.333	0.052	6.416	0.000

Tabella 3.6: Risultati per il modello della scala SUR

Inhibitory Control

Le stime dei loadings sono soddisfacenti. L'unico indice che mostra un buon adattamento è il GFI (0.981), mentre SRMR (0.096) presenta un valore borderline.

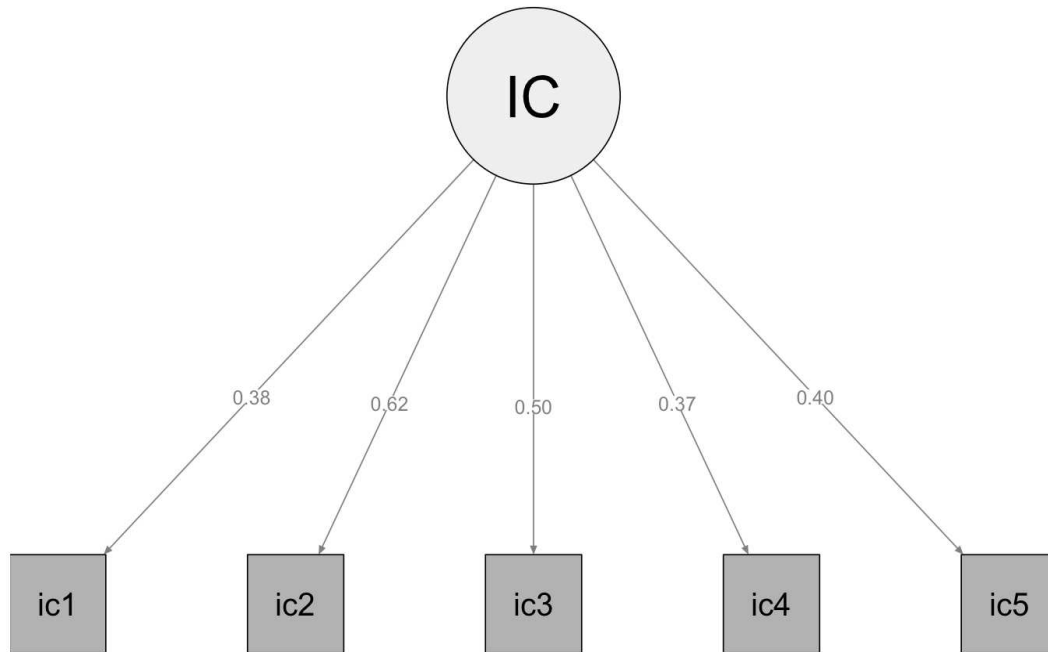


Figura 3.23: Modello CFA di primo ordine di Inhibitory Control.

Variabili	loadings	Std. Error	z-value	p-value
<i>ic</i> ₁	0.385	0.059	6.504	0.000
<i>ic</i> ₂	0.624	0.063	9.975	0.000
<i>ic</i> ₃	0.499	0.061	8.132	0.000
<i>ic</i> ₄	0.371	0.057	6.476	0.000
<i>ic</i> ₅	0.401	0.058	6.859	0.000

Tabella 3.7: Risultati per il modello della scala IC

Pleasure Sensitivity

Le stime dei loadings sono soddisfacenti. Tutti gli indici di fit mostrano un buon adattamento.

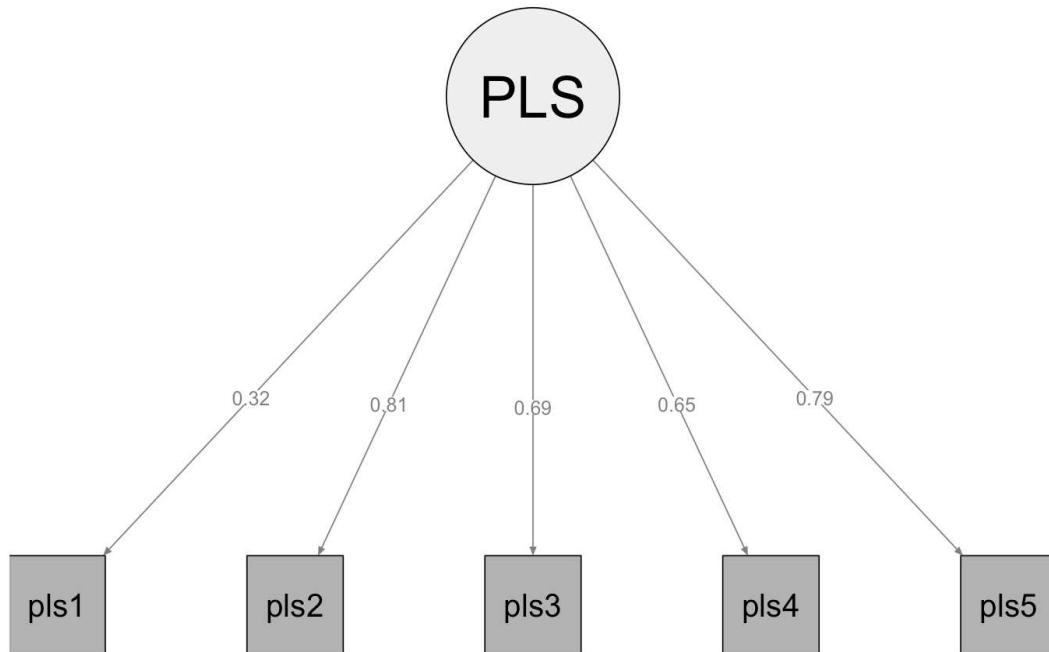


Figura 3.24: Modello CFA di primo ordine di Pleasure Sensitivity.

Variabili	loadings	Std. Error	z-value	p-value
pls_1	0.320	0.050	6.360	0.000
pls_2	0.809	0.025	32.941	0.000
pls_3	0.692	0.030	22.689	0.000
pls_4	0.652	0.035	18.696	0.000
pls_5	0.785	0.028	28.139	0.000

Tabella 3.8: Risultati per il modello della scala PLS

Perceptual Sensitivity

Le stime dei loadings, rispetto alla soglia fissata, sono inferiori solo per ps_3 . Si ha un buon adattamento per tutti gli indici fatta eccezione di chi-quadro (5.952) e RMSEA (0.071), che evidenziano un adattamento accettabile.

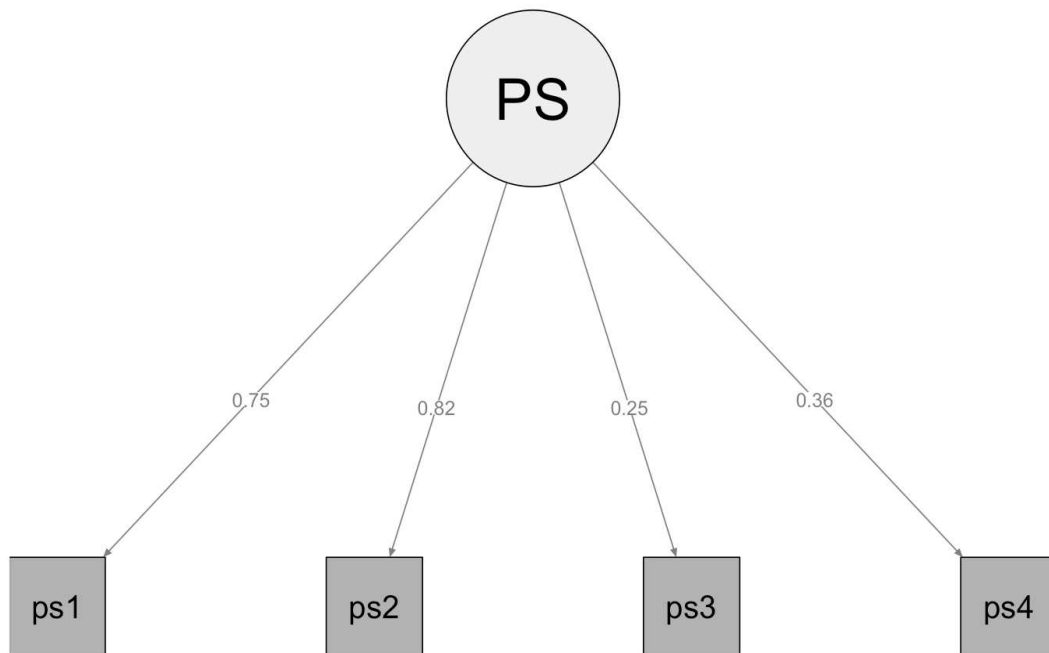


Figura 3.25: Modello CFA di primo ordine di Perceptual Sensitivity.

Variabili	loadings	Std. Error	z-value	p-value
ps_1	0.746	0.054	13.697	0.000
ps_2	0.818	0.054	15.162	0.000
ps_3	0.246	0.056	4.418	0.000
ps_4	0.360	0.049	7.407	0.000

Tabella 3.9: Risultati per il modello della scala PS

Shyness

Tutte le stime dei loadings sono soddisfacenti e tutti gli indici di fit mostrano un buon adattamento.

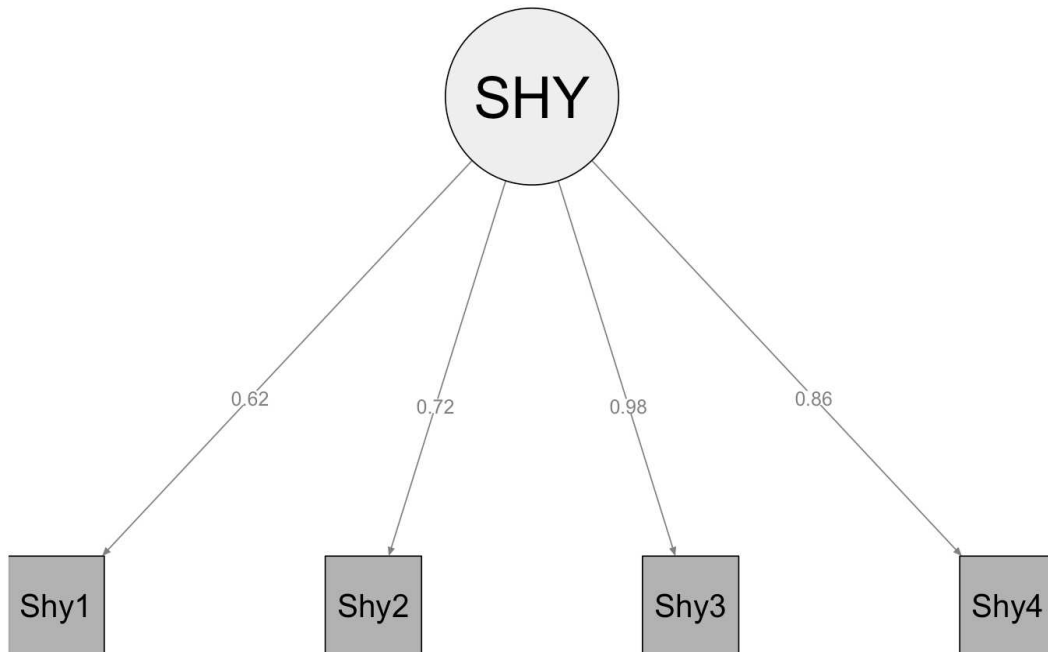


Figura 3.26: Modello CFA di primo ordine di Shyness.

Variabili	loadings	Std. Error	z-value	p-value
<i>Shy₁</i>	0.620	0.035	17.968	0.000
<i>Shy₂</i>	0.719	0.026	27.183	0.000
<i>Shy₃</i>	0.976	0.015	64.137	0.000
<i>Shy₄</i>	0.862	0.020	43.081	0.000

Tabella 3.10: Risultati per il modello della scala SHY

Aggression

Tutte le stime dei loadings risultano essere soddisfacenti. Si ha un buon adattamento per CFI (0.985), TLI (0.974), GFI (0.996) e SRMR (0.035) e un adattamento accettabile per chi-quadro (23.243) e RMSEA (0.063).

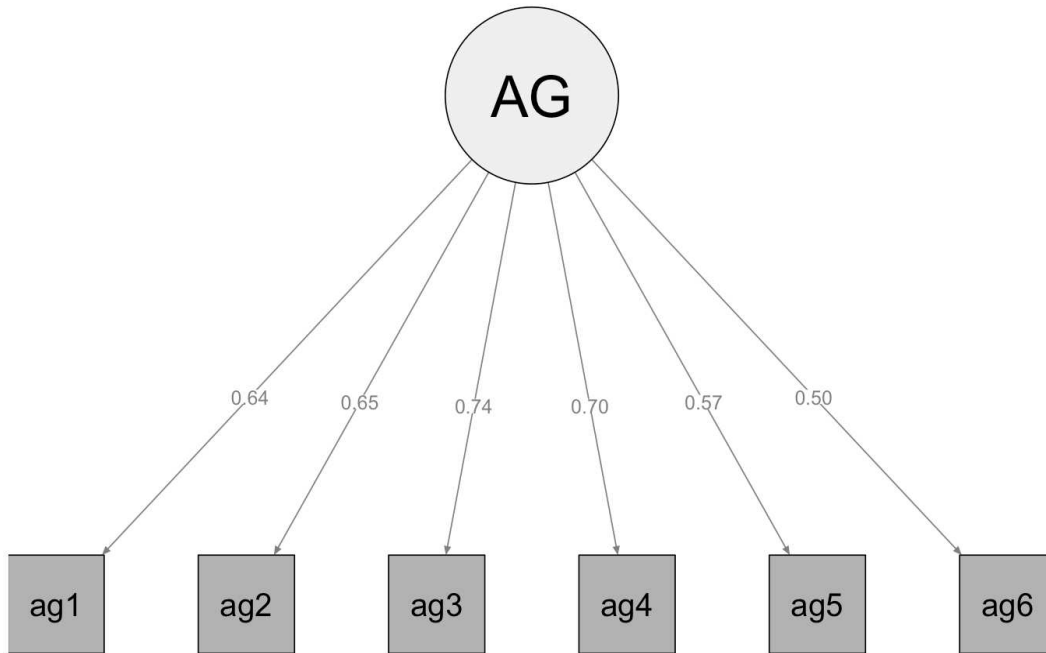


Figura 3.27: Modello CFA di primo ordine di Aggression.

Variabili	loadings	Std. Error	z-value	p-value
<i>ag</i> ₁	0.637	0.039	16.281	0.000
<i>ag</i> ₂	0.646	0.038	17.041	0.000
<i>ag</i> ₃	0.740	0.033	22.369	0.000
<i>ag</i> ₄	0.705	0.033	21.442	0.000
<i>ag</i> ₅	0.572	0.041	13.977	0.000
<i>ag</i> ₆	0.495	0.045	10.891	0.000

Tabella 3.11: Risultati per il modello della scala AG

Depressive Mood

Tutte le stime dei loadings sono soddisfacenti. Negli indici di fit si evidenzia un buon adattamento solo per GFI (0.992), e un adattamento accettabile per CFI(0.957) e SRMR (0.056).

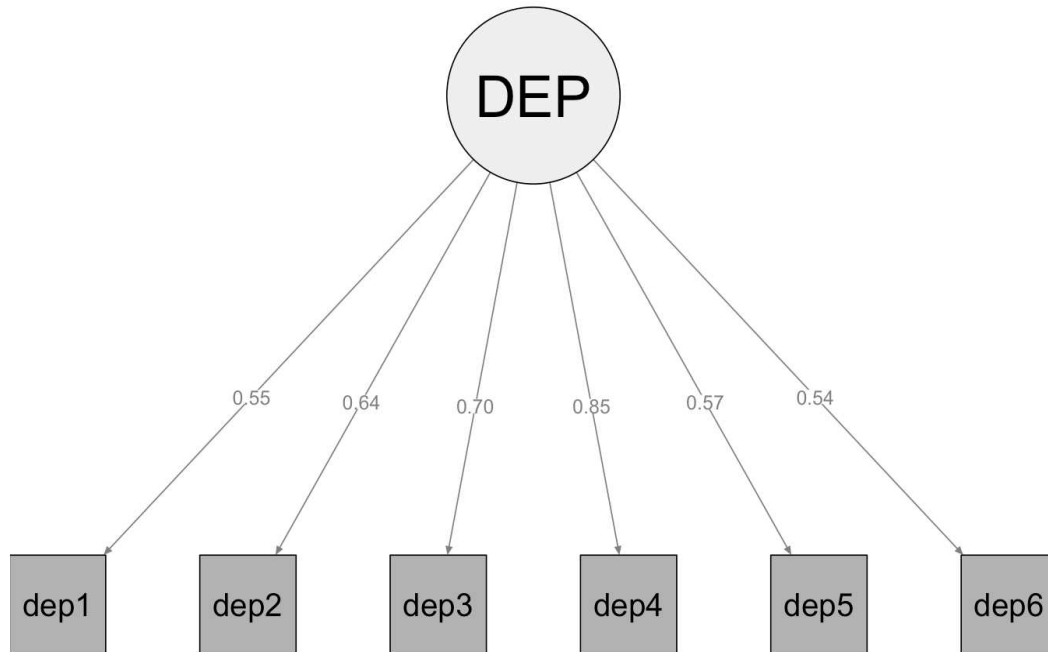


Figura 3.28: Modello CFA di primo ordine di Depressive Mood.

Variabili	loadings	Std. Error	z-value	p-value
<i>dep₁</i>	0.553	0.043	12.730	0.000
<i>dep₂</i>	0.644	0.034	19.176	0.000
<i>dep₃</i>	0.696	0.034	20.412	0.000
<i>dep₄</i>	0.855	0.027	31.992	0.000
<i>dep₅</i>	0.574	0.040	14.521	0.000
<i>dep₆</i>	0.540	0.042	10.800	0.000

Tabella 3.12: Risultati per il modello della scala DEP

3.5.3 Indici di bontà di adattamento dei modelli CFA di primo ordine

In conclusione, si riportano gli indici per tutte le scale valutate, i cui valori sono già stati discussi durante la presentazione dei modelli.

	χ^2	<i>df</i>	<i>p-value</i>	χ^2/df	<i>CFI</i>	<i>TLI</i>	<i>GFI</i>	<i>RMSEA</i>	<i>SRMR</i>
<i>ac</i>	24.472	5	0	4.894	0.932	0.864	0.994	0.092	0.050
<i>afl</i>	17.589	5	0.003	3.518	0.976	0.952	0.997	0.079	0.038
<i>att</i>	62.541	9	0	6.949	0.885	0.809	0.989	0.123	0.066
<i>fear</i>	48.230	9	0	5.359	0.884	0.807	0.989	0.105	0.062
<i>frus</i>	106.028	14	0	7.573	0.879	0.819	0.983	0.129	0.007
<i>Sur</i>	66.194	9	0	7.355	0.891	0.818	0.983	0.127	0.008
<i>ic</i>	66.491	5	0	13.298	0.709	0.418	0.981	0.176	0.096
<i>pls</i>	1.722	5	0.886	0.344	1	1.006	0.999	0	0.010
<i>ps</i>	5.952	2	0.051	2.976	0.991	0.972	0.998	0.071	0.032
<i>Shy</i>	0.275	2	0.872	0.137	1	1.002	0.999	0	0.003
<i>ag</i>	23.243	9	0.006	2.582	0.985	0.974	0.996	0.063	0.035
<i>dep</i>	58.056	9	0	6.450	0.957	0.928	0.992	0.117	0.056

Figura 3.29: Indici robusti di bontà di adattamento.

Capitolo 4

Conclusioni e discussione

Lo studio proposto aveva l'obiettivo di sostenere il ruolo considerevole che la *item analysis* ha come prima fase del processo di definizione dei fattori latenti, ossia la necessità di testare il funzionamento degli indicatori all'interno dei SEM partendo dai modelli più semplici (di primo ordine) prima di passare allo studio dei costrutti di livello più alto: qualora alcune delle scale non fossero valide, non sarà possibile procedere alla definizione di modelli di ordine superiore per cui è richiesto l'uso di queste scale, in quanto essi sono definiti a partire dai costrutti di primo livello.

Per maggiore chiarezza espositiva sono stati dapprima presentati e argomentati i problemi della misurazione in psicologia e di quanto la procedura di validità dei costrutti sia fondamentale, risaltando il ruolo degli strumenti statistici che devono essere scelti al fine di raggiungere una validazione corretta delle scale del questionario EATQ-R.

Al fine di giustificare l'uso dei modelli di equazioni strutturali, nel Capitolo 2 ne è stata illustrata la teoria sottostante, ponendo particolare attenzione sui modelli CFA e sugli indici di valutazione e di confronto che sono stati utilizzati nelle analisi (Capitolo 3), evidenziandone l'adeguatezza per il contesto trattato nel caso di studio.

Prima di procedere con la definizione dei modelli CFA di primo ordine, si è ritenuto doveroso presentare le caratteristiche del campione in analisi e delle distribuzioni delle risposte per ogni item. Il campione è apparso sostanzialmente omogeneo nel genere e nella classe frequentata, mentre le risposte non sempre sono state bilanciate e hanno riportato, nella maggior parte dei casi, una leggera asimmetria.

Dallo studio dei modelli di primo ordine e a seguito dei valori assunti dagli indici di fit, per alcune delle 12 scale non è stata confermata la struttura fattoriale originariamente proposta dal questionario. Menzio-

ni notevoli sono la scala *Inhibitory Control*, per la quale si è ottenuto un solo indice a sostegno del buon adattamento, seguita da *Attention* e *Fear*, per le quali si è avuto un solo indice a sostegno del buon adattamento e uno per adattamento accettabile. Al contrario, le scale *Pleasure Sensitivity* e *Shyness* sono soddisfacenti per tutti gli indici.

I loadings hanno indicato un buon adattamento per tutte le scale, ad eccezione di alcuni item (*frus₅*, *Sur₁*, *Sur₂*).

Alcuni dei risultati negli indici di fit potrebbero essere stati influenzati da differenze culturali e linguistiche, ricordando inoltre che lo strumento EATQ-R non era mai stato testato sulla popolazione italiana in precedenza.

Sviluppi futuri includono lo studio dei costrutti di livello superiore e la rivalutazione per alcuni item la cui bontà non è stata soddisfacente, riconsiderandone la presenza all'interno del questionario. Non si esclude l'utilizzo di altri strumenti statistici: come già discusso nel Capitolo 1, vi sono tuttora dei dibattiti in letteratura su quali strumenti utilizzare per evitare il problema della replicabilità.

Appendice A

Questionario EATQ-R

Vengono indicati gli item per ogni scala, per i quali si specifica se si tratta di *item reversed*¹, la posizione della domanda all'interno del questionario e la domanda corrispondente all'item.

item	
Activation Control, N=5	
ac1	R 7) Faccio fatica a finire le cose in tempo.
ac2	R 18) Faccio qualcosa di divertente per un po' prima di cominciare i compiti, anche se non dovrei.
ac3	30) Se mi assegnano un compito difficile comincio subito a farlo.
ac4	39) Di solito finisco i compiti prima della scadenza.
ac5	R 49) Quando devo fare una ricerca per la scuola rimando fino all'ultimo momento.
Affiliation, N=5	
aff1	17) Voglio avere la possibilità di condividere i miei pensieri con qualcun altro.
aff2	27) Mi piace abbracciare le persone a cui voglio bene.
aff3	31) Farei qualsiasi cosa per aiutare le persone a cui tengo.
aff4	44) Per me è importante avere dei rapporti stretti con gli altri.
aff5	54) Sono una persona piuttosto calorosa e amichevole.
Aggression, N=6	
ag1	5) Se sono arrabbiato con qualcuno tendo a dire cose che so benissimo feriranno i suoi sentimenti.
ag2	9) Quando mi arrabbio lancio in giro le cose o le rompo.
ag3	13) Quando mi arrabbio sul serio con qualcuno potrei anche picchiarlo.

¹Con la lettera R si fa riferimento agli item reversed, ossia domande con formulazione opposta rispetto al resto del test. Prima di effettuare le analisi, gli item reversed vengono ricodificati ponendo le modalità come segue: 5 diventa 1, 4 diventa 2, 3 rimane 3, 1 diventa 5

ag4	22)	Tendo ad essere maleducato con le persone che non mi piacciono.
ag5	50)	Quando mi arrabbio veramente con un amico, tendo ad esplodere.
ag6	58)	Faccio i dispetti alle persone senza un vero motivo.

Attention, N=6

att1	1)	Per me è facile concentrarmi a fondo quando faccio i compiti.
att2	R 34)	A scuola faccio fatica a ingranare quando passo da una materia all'altra.
att3	R 38)	Quando cerco di studiare faccio fatica a isolarmi dai rumori intorno a me e a concentrarmi.
att4	41)	Sono bravo a star dietro a tutte le varie cose che mi succedono intorno.
att5	59)	Ascolto molto attentamente quando qualcuno mi spiega come fare una cosa.
att6	R 61)	Ho la tendenza a lasciare le cose a metà per passare a fare qualcos'altro.

Depressive Mood, N=6

dep1	R 2)	Mi sento felice per la maggior parte della giornata.
dep2	11)	I miei amici sembrano divertirsi più di me.
dep3	20)	Spesso basta poco per farmi piangere.
dep4	29)	Mi rattristo più di quanto gli altri si immaginano.
dep5	37)	Mi rattristo quando tutto va storto.
dep6	55)	Mi sento triste anche quando dovrei divertirmi, come a Natale o durante un viaggio.

Fear, N=6

fear1	32)	Ho paura ad andare in macchina con persone a cui piace correre veloce.
fear2	35)	Mi preoccupa per la mia famiglia quando non siamo insieme.
fear3	40)	Mi preoccupa l'idea di mettermi nei guai.
fear4	46)	Mi innervoliscono alcuni ragazzi a scuola che spintonano i compagni e buttano in giro i libri.
fear5	51)	L'idea che i miei genitori possano abbandonarmi o non esserci più mi spaventa.
fear6	57)	Mi spavento quando a casa entro in una stanza buia.

Frustration, N=7

frus1	25)	Mi infastidisco quando cerco di telefonare e la linea è occupata.
frus2	36)	Mi arrabbio molto quando voglio fare una cosa e i miei genitori non mi lasciano.

frus3		47)	Mi irrito quando devo smettere di fare qualcosa che mi piace.
frus4		56)	Aspettare in coda è una cosa che mi infastidisce molto.
frus5		60)	Quando faccio un errore nei compiti mi scoraggio molto.
frus6		62)	Mi infastidisco quando le persone mi interrompono mentre sto parlando.
frus7		64)	Mi arrabbio molto quando non riesco a far bene una cosa.

Inhibitory Control, N=5

ic1	R	10)	Faccio fatica a non aprire i regali prima del dovuto.
ic2		14)	Quando mi dicono di smettere di fare qualcosa, per me è facile fermarmi.
ic3	R	26)	Più mi trattengo dal fare una cosa che non dovrei, più è probabile che io la faccia.
ic4		43)	Per me è facile mantenere un segreto.
ic5		63)	In genere riesco a seguire i miei piani e i miei obiettivi.

Pleasure Sensitivity, N=5

pls1		4)	Mi piace sentire una brezza calda sul viso.
pls2		16)	Mi piace ascoltare gli uccellini che cantano.
pls3		23)	Mi piace osservare la forma delle nuvole nel cielo.
pls4		33)	Mi piace guardare gli alberi e camminare nei boschi.
pls5		65)	Mi piace ascoltare il il fruscio delle foglie in autunno.

Perceptual Sensitivity, N=4

ps1		6)	Riesco a notare anche piccoli cambiamenti intorno a me, ad esempio le luci che diventano più luminose in una stanza.
ps2		12)	Tendo a notare piccoli cambiamenti che gli altri non notano.
ps3		21)	Sono molto sensibile ai rumori.
ps4		24)	Riesco a capire se una persona è arrabbiata dalla sua espressione del viso.

Shyness, N=4

Shy1		8)	Sono timido con ragazzi dell'altro sesso.
Shy2		15)	Mi sento timido quando incontro persone nuove.
Shy3		45)	Io sono timido.
Shy4	R	53)	Non sono timido.

Surgency, N=6

Sur1		3)	Penso che sarebbe entusiasmante trasferirsi in un'altra città.
Sur2	R	19)	Non mi piacerebbe vivere in una città molto grande,

- anche se fosse sicura.
- Sur3 R 28) L'idea di fare una discesa ripida con gli sci mi spaventa.
- Sur4 42) Non avrei paura a provare qualche sport rischioso,
come le immersioni in mare aperto.
- Sur5 48) Non avrei paura a provare cose tipo
l'arrampicata in montagna.
- Sur6 52) Mi piace andare nei posti in cui c'è molta gente
e tanto divertimento.

Appendice B

Specificazione modelli CFA in R

A titolo di esempio, viene mostrato il codice relativo alla scala AC. Per le altre scale la procedura è la medesima e può essere facilmente generalizzata.

Definizione del modello

```
library(lavaan)
m_ac = 'ac=~ac_1 + ac_2 + ac_3 + ac_4 + ac_5'

fit_ac<-cfa(m_ac,data=data1,ordered=(colnames(data1[,12:104])),
missing="pairwise",estimator="WLSMV",std.lv=T)
```

Stime e risultati

```
summary(fit_ac,fit.measures=TRUE,standardized=T)
```

Rappresentazione grafica

```
library(semPlot)
semPaths(fit_ac,
         layout="tree2",
         mar = c(5, 1, 5, 1),

         style="LISREL",
         #intercepts = T,
         sizeMan=10.0,
```

```
sizeLat=18,  
edge.label.cex = 0.9,  
edge.label.position=0.6,  
asize=2,  
residuals = FALSE,  
nodeLabels = c("ac1","ac2",  
"ac3","ac4","ac5", "AC"),  
color = list(lat = rgb(239,  
239, 239, maxColorValue =255),  
man = rgb(178, 178, 178,  
maxColorValue = 255)),  
thresholds=FALSE,  
intercept=F,  
exoVar=F,  
whatLabels="est",  
optimizeLatRes = TRUE,  
exoCov = FALSE)
```

Bibliografia

- Bollen K. A. (1989). *Structural equations with latent variables*. John Wiley & Sons.
- Bollen K. A. (1996). An alternative two stage least squares (2sls) estimator for latent variable equations. *Psychometrika*, **61**(1), 109–121.
- Bollen K. A.; Bauer D. J. (2004). Automating the selection of model-implied instrumental variables. *Sociological Methods & Research*, **32**(4), 425–452.
- Bollen K. A.; Bauer D. J.; Christ S. L.; Edwards M. C. (2010). Overview of structural equation models and recent extensions. *Statistics in the social sciences: Current methodological developments*.
- Brown T. A. (2015). *Confirmatory factor analysis for applied research*. Guilford publications.
- Browne M. W. (1984). Asymptotically distribution-free methods for the analysis of covariance structures. *British journal of mathematical and statistical psychology*, **37**(1), 62–83.
- Capaldi D. M.; Rothbart M. K. (1992). Development and validation of an early adolescent temperament measure. *The Journal of Early Adolescence*, **12**(2), 153–173.
- Civelek M. E. (2018). Essentials of structural equation modeling. Available at SSRN: <https://ssrn.com/abstract=3338325>.
- Cronbach L. J.; Meehl P. E. (1955). Construct validity in psychological tests. *Psychological bulletin*, **52**(4), 281.
- Currie C.; Molcho M.; Boyce W.; Holstein B.; Torsheim T.; Richter M. (2008). Researching health inequalities in adolescents: the development of the health behaviour in school-aged children (hbsc) family affluence scale. *Social science & medicine*, **66**(6), 1429–1436.

- Davis W. R. (1993). The fc1 rule of identification for confirmatory factor analysis: A general sufficient condition. *Sociological Methods & Research*, **21**(4), 403–437.
- Fabrigar L. R.; Wegener D. T.; Petty R. E. (2020). A validity-based framework for understanding replication in psychology. *Personality and Social Psychology Review*, **24**(4), 316–344.
- Finch W. H.; French B. F. (2015). *Latent variable modeling with R*. Routledge.
- Flake J. K.; Fried E. I. (2020). Measurement schmeasurement: Questionable measurement practices and how to avoid them. *Advances in Methods and Practices in Psychological Science*, **3**(4), 456–465.
- Flake J. K.; Pek J.; Hehman E. (2017). Construct validation in social and personality research: Current practice and recommendations. *Social Psychological and Personality Science*, **8**(4), 370–378.
- Flake J. K.; Davidson I. J.; Wong O.; Pek J. (2022). Construct validity and the validity of replication studies: A systematic review. *American Psychologist*.
- Flora D. B.; Curran P. J. (2004). An empirical evaluation of alternative methods of estimation for confirmatory factor analysis with ordinal data. *Psychological methods*, **9**(4), 466.
- Garson G. D. (2013). *Path analysis*. Statistical Associates Publishing Asheboro, NC.
- Hoyle R. H. (2012). *Handbook of structural equation modeling*. Guilford Press.
- Jöreskog K. G.; Sörbom D. (1978). *LISREL IV: Analysis of Linear Structural Relationship by the Method of Maximum Likelihood*. Internat. Educational Services.
- Katsikatsou M.; Moustaki I.; Yang-Wallentin F.; Jöreskog K. G. (2012). Pairwise likelihood estimation for factor analysis models with ordinal data. *Computational Statistics & Data Analysis*, **56**(12), 4243–4258.
- Kline R. B. (2015). *Principles and practice of structural equation modeling*. Guilford publications.
- Lleras C. (2005). Path analysis. *Encyclopedia of social measurement*, **3**(1), 25–30.

- Marci T.; Moscardino U.; Lionetti F.; Santona A.; Altoé G. (2020). Using harter and likert response formats in middle childhood: A comparison of attachment measures. *Assessment*, **27**(8), 1821–1835.
- Open Science Collaboration (2015). Estimating the reproducibility of psychological science. *Science*, **349**(6251), aac4716.
- Ramlall I. (2016). *Applied structural equation modelling for researchers and practitioners: Using R and Stata for behavioural research*. Emerald Group Publishing.
- Rothbart M. K.; Derryberry D. (1981). Theoretical issues in temperament. In *Developmental disabilities*, pp. 383–400. Springer.
- Sanson A.; Hemphill S. A.; Smart D. (2004). Connections between temperament and social development: A review. *Social Development*, **13**(1), 142–170.
- Schermelleh-Engel K.; Moosbrugger H.; Müller H. *et al.* (2003). Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of-fit measures. *Methods of psychological research online*, **8**(2), 23–74.
- Schumacker R.; Lomax R. (2010). A beginners guide to structural equation modeling: Sem. *New Jersey: Lawrence Erlbaum Associates*.
- Stein C.; Morris N.; Nock N. (2012). Structural equation modeling. *Methods in molecular biology (Clifton, N.J.)*, **850**, 495–512.
- Sturgis P. (2022). Structural equation modelling (sem): What it is and what it isn't. national centre for research methods online learning resource. available at <https://www.ncrm.ac.uk/resources/online/all/?mainid=10416>.
- Tucker L. R.; Lewis C. (1973). A reliability coefficient for maximum likelihood factor analysis. *Psychometrika*, **38**(1), 1–10.
- Vieno A.; Santinello M.; Lenzi M.; Baldassari D.; Mirandola M. (2009). Health status in immigrants and native early adolescents in italy. *Journal of community health*, **34**(3), 181–187.
- Zentner M.; Shiner R. L. (2012). *Fifty years of progress in temperament research: A synthesis of major themes, findings, and challenges and a look forward*. The Guilford Press.
- Zhang Z.; Wang L. (2017). Advanced statistics using r. *Granger, IN: ISDSA Press*, **3**.