

Analisi Psicometrica del Multicontextual Empathy Test

Glossario delle Variabili

- **ID:** Codice identificativo univoco del partecipante.
- **Età:** Età del partecipante (in anni).
- **Residenza:** Regione di residenza del partecipante.
- **Genere:** Sesso biologico del partecipante (ad es. Donna o Uomo).
- **Educazione:** Titolo di studio del partecipante (colonna testuale) e relativo **codice** numerico (colonna adiacente) che ne indica il livello.
- **Occupazione:** Categoria occupazionale del partecipante (colonna testuale) e relativo **codice** numerico.
- **P.Prosociale (IRI):** Punteggio alla scala di preoccupazione empatica dell'*Interpersonal Reactivity Index* (IRI). Questa scala misura la tendenza del soggetto a provare empatia compassionevole e preoccupazione verso gli altri
- **E.Cognitiva (QCAE):** Punteggio di **Empatia Cognitiva** al **Questionnaire of Cognitive and Affective Empathy** (QCAE). È il punteggio totale delle sottoscale cognitive del QCAE (*Perspective Taking e Online Simulation*), indicando la capacità di comprendere i pensieri e le prospettive altrui.
- **E.Affettiva (QCAE):** Punteggio di **Empatia Affettiva** al QCAE. È il totale delle sottoscale affettive (*Emotional Contagion, Proximal Responsivity, Peripheral Responsivity*), indicando la tendenza a condividere e rispondere alle emozioni altrui.
- **PT (QCAE):** Sottoscala **Perspective Taking** del QCAE (abilità di assumere attivamente il punto di vista degli altri in contesti quotidiani).

- **OS (QCAE):** Sottoscala **Online Simulation** del QCAE (tendenza a immaginare le sensazioni e reazioni altrui anche in assenza di interazione diretta).
- **ECo (QCAE):** Sottoscala **Emotional Contagion** del QCAE (propensione a “contagiarsi” delle emozioni altrui).
- **PeR (QCAE):** Sottoscala **Peripheral Responsivity** del QCAE (responsività empatica verso situazioni remote o persone estranee, ad es. notizie o film).
- **PrR (QCAE):** Sottoscala **Proximal Responsivity** del QCAE (responsività empatica in situazioni vicine o che coinvolgono persone legate al soggetto).
- **QCAE tot:** Punteggio totale di empatia al QCAE (somma di tutte le sottoscale cognitive e affettive).
- **Alessitimia:** Punteggio totale alla **Toronto Alexithymia Scale (TAS-20)**, misura la difficoltà del soggetto nell’identificare e descrivere le proprie emozioni (maggiore punteggio = maggiore alessitimia).
- **E.Cognitiva MET:** Punteggio di **Empatia Cognitiva** al MET. È la somma dei punteggi ottenuti negli item cognitivi del MET (uno per ogni scenario); riflette l’abilità di mentalizzazione, cioè inferire lo stato mentale altrui in ciascun contesto.
- **E.Affettiva MET:** Punteggio di **Empatia Affettiva** al MET. Somma dei punteggi degli item affettivi del MET (uno per scenario); riflette la capacità di condivisione dell’esperienza emotiva altrui, ovvero quanto il soggetto risponde emotivamente in accordo con l’emozione altrui.
- **P.Prosociale MET:** Punteggio di **Preoccupazione Prosociale** al MET. Somma dei punteggi degli item prosociali del MET (uno per scenario); indica la tendenza a mostrare preoccupazione attiva e intenzione di aiutare l’altro.
- **MET Totale:** Punteggio totale al MET (somma di tutti i 18 item). Rappresenta la performance globale nel test situazionale di empatia.

- **Item 1.1 – 6.3 (MET):** Punteggi dei singoli item del MET. La notazione X.Y identifica l'item Y-esimo dello scenario X (ci sono 6 scenari numerati 1–6, ciascuno con 3 domande: .1 cognitiva, .2 affettiva, .3 prosociale). Ogni item è valutato da 0 a 2 punti (anche frazioni di 0,5 in caso di risposte parzialmente corrette). Ad esempio, **1.1** è l'item cognitivo dello scenario 1, **1.2** l'item affettivo dello scenario 1, **1.3** quello prosociale dello scenario 1, e così via fino a **6.3** (item prosociale dello scenario 6).

Introduzione e Ipotesi

Il Multicontextual Empathy Test (MET) è un test di giudizio situazionale progettato per valutare le abilità empatiche in contesti realistici. A differenza dei questionari self-report tradizionali, il MET presenta scenari sociali e richiede al partecipante di giudicare o reagire a ciascuna situazione, fornendo un punteggio di performance. L'idea alla base è che l'empatia, intesa come insieme di abilità cognitive, affettive e motivazionali, possa essere misurata tramite le risposte fornite in situazioni simulate, piuttosto che solo attraverso auto-valutazioni soggettive.

Le ipotesi guida dello studio sono le seguenti:

1. **Valutabilità dell'empatia tramite SJT:** Le abilità empatiche del soggetto sono valutabili mediante un test situazionale. Il MET dovrebbe fornire punteggi significativi che riflettono l'empatia del partecipante, analogamente ai questionari self-report esistenti, da un punto di vista differente.
2. **Validità convergente con misure esistenti (Ipotesi 1a):** I punteggi del MET devono mostrare una validità convergente con strumenti self-report di empatia già validati, in particolare il QCAE (Questionario di Empatia Cognitiva e Affettiva) e l'IRI (Indice di Reattività Interpersonale). Ci si attende correlazioni positive significative tra le sottoscale corrispondenti:

- La mentalizzazione misurata dal MET (E.Cognitiva MET) dovrebbe correlare con l'empatia cognitiva misurata dal QCAE (*Perspective Taking, Online Simulation*).
- La condivisione dell'esperienza del MET (E.Affettiva MET) dovrebbe correlare con l'empatia affettiva del QCAE (*Emotional Contagion, Proximal / Peripheral Responsivity*).
- La preoccupazione prosociale del MET (P.Prosociale MET) dovrebbe correlare con la preoccupazione empatica (*empathic concern*) misurata dall'IRI.
- Il punteggio totale MET dovrebbe correlare in modo positivo con il punteggio totale QCAE e (in misura minore) negativamente con l'alessitimia, data la relazione inversa attesa tra empatia e alessitimia.

3. Struttura fattoriale a tre componenti (Ipotesi 1b): Il MET è stato progettato per riflettere tre componenti distinte dell'empatia, la mentalizzazione (empatia cognitiva), la condivisione dell'esperienza (empatia affettiva) e la preoccupazione prosociale. Ci si attende quindi che gli item del test si raggruppino in tre fattori corrispondenti a queste dimensioni teoriche. In analisi fattoriale:

- Gli item cognitivi (.1) dovrebbero caricarsi su un fattore, separato dagli item affettivi (.2) e dagli item prosociali (.3).
- I tre fattori dovrebbero spiegare una porzione significativa della varianza totale e mostrare correlazioni moderate tra loro (poiché le componenti dell'empatia sono correlate ma distinguibili).
- Un'analisi fattoriale confermativa (CFA) dovrebbe evidenziare un buon fit per un modello a 3 fattori coerente col modello teorico, se il MET rispecchia chiaramente le tre componenti ipotizzate.

Analisi della Validità Convergente

Per verificare la validità convergente, sono state esaminate le correlazioni di Spearman fra le sottoscale del MET e le sottoscale corrispondenti dei questionari self-report (QCAE e IRI) data la distribuzione non normale di molti parametri. L'attenzione si è focalizzata sulle correlazioni tra:

- Empatia Cognitiva MET vs. Empatia Cognitiva QCAE (e sottoscale cognitive PT, OS),
- Empatia Affettiva MET vs. Empatia Affettiva QCAE (e sottoscale affettive ECo, PeR, PrR),
- Preoccupazione Prosociale MET vs. Preoccupazione Empatica IRI,
- nonché tra i punteggi totali MET e QCAE totali.

Risultati delle Correlazioni

- **Empatia Cognitiva (MET vs QCAE):** La correlazione tra E.Cognitiva MET e la corrispondente scala cognitiva del QCAE risulta molto bassa e non significativa ($r = 0.015$, $p = 0.45$). Anche rispetto alle due sottoscale cognitive del QCAE, *Perspective Taking* (PT) e *Online Simulation* (OS), la scala cognitiva del MET non mostra associazioni significative con la prima ma correlazione debole con la seconda (con PT $r = -0.16$, $p = 0.92$; con OS $r = 0.22$; $p = 0.03$). Questo dato indica che il punteggio di mentalizzazione ottenuto nel MET non tende a convergere con l'auto-valutazione di abilità cognitive empatiche (possibili cause sono discusse più avanti).
- **Empatia Affettiva (MET vs QCAE):** La sottoscala affettiva del MET mostra invece buona validità convergente. In particolare, E.Affettiva MET correla in maniera positiva e significativa con E.Affettiva QCAE ($r = 0.36$, $p < 0.001$). Ciò significa che chi ottiene un punteggio alto nelle risposte affettive del test situazionale tende ad auto-riportare maggior contagio emotivo e responsività empatica. Inoltre va notato che E.Affettiva MET presenta correlazioni moderate con tutte le sottoscale affettive specifiche del

QCAE, Emotional Contagion, Peripheral Responsivity e Proximal Responsivity (rispettivamente $r = 0.27$, $p = 0.03$; $r = 0.27$, $p = 0.008$; $r = 0.25$, $p = 0.015$), suggerendo che il MET catturi almeno in parte la componente di condivisione emotiva dell'empatia. Da notare, E.Affettiva MET correla significativamente anche con la scala IRI di preoccupazione empatica (*Empathic Concern*), $r = 0.32$ ($p = 0.003$), riflettendo il legame tra condividere l'emozione altrui e sentirsi motivati ad aiutare.

- **Preoccupazione Prosociale (MET vs IRI):** Come ipotizzato, il punteggio P.Prosociale del MET mostra una correlazione positiva significativa con la scala di Preoccupazione Prosociale dell'IRI ($r = 0.307$, $p = 0.003$) indicando che chi nei vari scenari MET manifesta maggior intenzione di aiutare o preoccupazione verso l'altro tende anche a dichiarare alti livelli di preoccupazione empatica nell'IRI.
- **Correlazioni tra punteggi totali:** Il punteggio MET Totale mostra una correlazione positiva ma modesta con il punteggio QCAE Totale ($r = 0.194$, $p = 0.047$). Questa correlazione, pur significativa, è di bassa entità, suggerendo che la performance complessiva al MET condivide soltanto in piccola parte la varianza col punteggio totale di empatia auto-riferita. In linea con le aspettative teoriche, il MET Totale correla negativamente (anche se debolmente) con l'Alessitimia ($r = -0.185$, $p = 0.94$) e positivamente con la Propensione Prosociale IRI ($r = 0.37$, $p < 0.001$) e con l'Empatia Affettiva QCAE ($r = 0.28$, $p = 0.008$), coerentemente con l'idea che elevata empatia si associa a minore difficoltà nell'identificare e nominare emozioni e con una maggiore prosocialità. Tuttavia, l'associazione tra i punteggi totali non è forte, a causa probabilmente della natura composita del punteggio MET e della moderata affidabilità interna discussa più avanti.

Interpretazione

- La componente affettiva del MET sembra avere la miglior validità convergente: nonostante problemi interni (discussi sotto), essa raggiunge buone correlazioni con QCAE affettivo ed Empathic Concern IRI. Ciò

implica che, pur essendo gli item affettivi MET poco coerenti tra loro, il punteggio complessivo che producono riesce comunque a intercettare qualcosa del tratto di empatia affettiva misurato dai *self-report*. È un risultato interessante: suggerisce che i partecipanti che “performano” bene nel MET sul piano emotivo sono effettivamente quelli che si auto-descrivono come empatici emotivamente, nonostante la variabilità delle risposte nei diversi scenari.

- La componente cognitiva del MET, al contrario, non mostra convergenza con l'empatia cognitiva auto-riferita. Diverse spiegazioni sono possibili: (a) il MET cognitivo potrebbe soffrire di *ceiling effect*, molti partecipanti ottengono il punteggio massimo o vicino al massimo negli item cognitivi (come indicato dalle medie elevate di E.Cognitiva MET = 9.3 su 12, con punteggio massimo 12 ottenuto da diversi soggetti); una varianza ridotta rende difficile ottenere correlazioni significative. (b) Il QCAE cognitivo misura soprattutto delle disposizioni (es. tendenza dichiarata a prendere prospettive), mentre il MET valuta *prestazioni* in compiti di teoria della mente; i due aspetti potrebbero non allinearsi perfettamente. (c) Possibili problemi negli item cognitivi stessi (ad es. alcuni scenari cognitivi possono essere troppo semplici o ambigui) riducono la validità. Questo punto evidenzia un'area problematica: la mentalizzazione situazionale misurata dal MET andrebbe migliorata perché attualmente non predice la capacità empatica cognitiva di tratto.
- La componente prosociale del MET ha mostrato correlazioni moderate con misure affini (IRI), ma anche con componenti affettive del QCAE. Questo indica che la tendenza ad agire in modo prosociale nei scenari del MET è collegata sia al tratto di preoccupazione empatica sia, in parte, alla risonanza emotiva. Ciò è coerente con la teoria: provare emozioni empatiche verso qualcuno spesso motiva l'azione di aiuto. Il MET cattura questa connessione, sebbene non in modo fortissimo ($r = 0.30$). Probabilmente la variabilità di contesto nei diversi scenari (es. il tipo di aiuto richiesto) incide sulle risposte, limitando la consistenza generale.

Il MET mostra evidenze di validità convergente soddisfacenti soprattutto sul versante affettivo e prosociale, mentre risulta carente sul versante cognitivo. Un aspetto cruciale, emerso dall'analisi incrociata, è il contrasto tra validità esterna e consistenza interna: ad esempio, la forte correlazione tra le scale affettive MET e QCAE contrasta con la debole coerenza interna degli item affettivi MET (bassa α di Cronbach), suggerendo che la dimensione di empatia affettiva è presente ma misurata in modo poco omogeneo nel test. Questo tema viene approfondito nell'analisi fattoriale e di affidabilità in seguito.

Analisi della Struttura Fattoriale

Per valutare se il MET rispecchia la struttura teorica a tre componenti (Ipotesi 1b), sono state condotte sia un'Analisi Fattoriale Esplorativa (EFA) sugli item, sia un'Analisi Fattoriale Confermativa (CFA) per testare il modello a 3 fattori.

Adeguatezza del campione e degli item

Come verifica preliminare, il **Test di Bartlett** sulla matrice di correlazione tra i 18 item è risultato altamente significativo ($\chi^2(153) = 269$, $p < 0.001$), indicando che le correlazioni osservate non sono dovute al caso e che gli item condividono un fattore comune almeno in parte. Tuttavia, l'indice MSA (*Measure Of Sampling Adequacy*) del test Kaiser–Meyer–Olkin (KMO) ha fornito un valore solo mediocre (MSA = 0.551). Un MSA intorno a 0.55 è sotto la soglia generalmente accettata di 0.70 per considerare l'analisi fattoriale pienamente attendibile.

Analizzando gli indici MSA per i singoli item, emergono differenze notevoli: molti item hanno MSA attorno a 0.5 - 0.6, ma alcuni scendono ben al di sotto. In particolare, l'item 4.1 (cognitivo, scenario 4) ha MSA = 0.406 e l'item 6.2 (affettivo, scenario 6) ha MSA = 0.364. Questi valori molto bassi indicano che tali item correlano debolmente con tutti gli altri, riducendo la coerenza fattoriale generale.

Analisi Fattoriale Esplorativa (EFA)

L'EFA ha evidenziato una struttura più complessa del previsto. In teoria ci aspetteremmo 3 fattori principali; tuttavia, l'esame degli autovalori e dello scree plot ha suggerito la possibile presenza di 4-5 fattori. Ciò è dovuto al fatto che alcuni item non si aggregano chiaramente nei fattori ipotizzati, generando fattori spuri.

Alcuni risultati salienti:

- Sono emersi fattori aggiuntivi con pochi item dai caricamenti significativi. Ad esempio, l'item 4.1 (già notato per bassa MSA) ha mostrato un caricamento elevato da solo su un fattore isolato (caricamento = 0.75) con alta unicità (= 0.41). Ciò indica un fattore specifico “ad hoc” quasi unicamente per quell'item, privo di altri item in comune, segno che 4.1 non condivide il costrutto con gli altri.
- Diversi item presentano unicità molto elevate (cioè bassa comunaltà). Ad esempio, l'item 2.2 (affettivo, scenario 2) ha unicità = 0.92, e l'item 2.3 (prosociale, scenario 2) unicità = 0.86, il che significa che oltre l'85 - 90% della loro varianza è indipendente dai fattori comuni. In pratica, questi item non mostrano caricamenti significativi su alcun fattore, suggerendo che non contribuiscono a misurare una dimensione condivisa con gli altri item (potrebbero essere mal posti o interpretati in modo idiosincratico dai partecipanti).
- Molti item hanno mostrato caricamenti incrociati su più fattori invece di un singolo fattore netto. Un caso emblematico è l'item 5.1 (cognitivo, scenario 5) che in analisi esplorativa carica moderatamente su tre fattori simultaneamente (caricamento = 0.57 su Fattore 1, = 0.45 su Fattore 3 e = 0.62 su Fattore 4) e una unicità praticamente nulla (= 0.005). Questo significa che 5.1 condivide varianza con diversi cluster di item differenti, rendendo difficile attribuirlo a una dimensione specifica (forse la sua interpretazione richiede capacità cognitive ma anche affettive/prosociali).

- Solo una piccola parte degli item si aggrega come atteso: ad esempio alcuni item cognitivi di certi scenari (es. 3.1, 3.3) tendono a caricarsi insieme su un fattore comune, e analogamente alcuni item prosociali (3.3, 6.3) condividono un fattore. In generale la mappa dei caricamenti non ha però mostrato la netta separazione 3 fattori = 3 tipi di item auspicata, ma piuttosto un pattern confuso con fattori spurii e *cross-loading* diffusi.

L'EFA quindi non conferma una chiara struttura a tre fattori. I dati suggeriscono che senza interventi sugli item, il MET potrebbe richiedere più fattori per spiegare le intercorrelazioni, ma questi fattori aggiuntivi non corrispondono a costrutti interpretabili teoricamente. Questo risultato riflette una struttura fattoriale “poco chiara” e non in linea con il modello teorico originario.

Analisi Fattoriale Confermativa (CFA)

La CFA è stata utilizzata per testare esplicitamente il modello teorico a tre fattori correlati (Mentalizzazione, Esperienza Condivisa, Preoccupazione Prosociale), ciascuno collegato ai rispettivi 6 item del MET. I risultati indicano che il modello teorico, applicato ai dati raccolti, presenta un adattamento insoddisfacente:

- Gli indici di fitness globali sono ben al di sotto dei criteri accettabili. Ad esempio, CFI (*Comparative Fit Index*) = 0.476 e TLI (*Tucker-Lewis Index*) = 0.393, valori drasticamente inferiori alla soglia convenzionale di 0.90. Ciò implica che la varianza spiegata dal modello a 3 fattori è meno della metà di quella spiegata da un modello nullo basato sulla media – segno di mancata corrispondenza tra modello teorico e dati osservati.
- L' RMSEA (*Root mean squared error of approximation*) = 0.088 (IC 90% 0.064 - 0.109) è al limite superiore accettabile (circa 0.08), indicando un misfit moderato. Considerando il campione ridotto, un RMSEA vicino a 0.09 suggerisce discrepanze significative tra la matrice di covarianza attesa e osservata.

- L' SRMR (*Standardized Root mean squared residual*) = 0.109 risulta anch'esso elevato (sopra 0.10), confermando residui significativi nelle correlazioni non spiegate dal modello.

Questi indici, nel complesso, evidenziano che il modello a tre fattori così com'è non riesce a spiegare adeguatamente i dati raccolti. In altre parole, le risposte ai 18 item non seguono le relazioni attese sotto un semplice schema tri-fattoriale.

Guardando ai dettagli del modello CFA, emergono problemi su specifici item/caricamenti:

- Gli item della dimensione condivisione dell'esperienza (affettiva) hanno mostrato caricamenti molto bassi e non significativi sul loro fattore. Nessuno dei 6 item affettivi ha un parametro di loading significativo ($p > 0.45$ per tutti). In particolare, item come 3.2 o 6.2 hanno pesi quasi nulli (< 0.05) sul fattore affettivo, indicando che questa dimensione nel modello CFA non è effettivamente rappresentata dai suoi indicatori. Di fatto, il fattore *Experience sharing* appare “vuoto” nel senso che non emergono indicatori validi a sostenerlo.
- Anche diversi item della dimensione mentalizzazione (cognitiva) non mostrano legami significativi col loro fattore. Ad esempio, per il fattore Mentalizzazione i primi tre item (1.1, 2.1, 3.1) hanno pesi moderati e significativi ($p < 0.01$), mentre gli item 4.1, 5.1 e 6.1 hanno loading trascurabili e non significativi ($p = 0.98, 0.11, 0.057$ rispettivamente). Ciò conferma quanto visto in EFA: gli scenari 4, 5, 6 per la parte cognitiva non contribuiscono alla dimensione come previsto. In pratica, metà degli item cognitivi non “caricano” sul fattore di mentalizzazione.
- Per la dimensione preoccupazione prosociale, la situazione è intermedia: alcuni item funzionano, altri no. Solo l'item 3.3 e 6.3 hanno coefficienti di loading significativi sul fattore ($p < 0.001$ per entrambi). Tutti gli altri item prosociali (1.3, 2.3, 4.3, 5.3) hanno loading non significativi (p tra 0.09 e 0.61), sebbene alcuni siano di entità moderata (0.10 - 0.17). Dunque il fattore

prosociale è sorretto principalmente da 2 item su 6, il che è chiaramente insufficiente e inaffidabile dal punto di vista della misura.

- Le covarianze tra i fattori stimate nel modello confermativo risultano poco significative. Ad esempio, la correlazione fattoriale tra Mentalizzazione ed Esperienza condivisa non è significativa ($p = 0.45$), il che potrebbe essere dovuto anche alla virtuale mancanza di varianza del fattore affettivo. In teoria ci si aspettava correlazioni moderate tra i fattori (essendo componenti dell'empatia), ma l'assenza di associazioni robuste nei dati riflette piuttosto la debole definizione dei fattori stessi con questi item.

La CFA dunque corrobora i risultati dell'EFA: il MET nella sua forma attuale non presenta la struttura trifattoriale prevista. La struttura fattoriale del test risulta quindi poco chiara e non rispetta il modello teorico di partenza. Ciò significa che, sebbene concettualmente il test fosse pensato per distinguere 3 componenti dell'empatia, empiricamente tali componenti non emergono nettamente dalle risposte dei partecipanti. Alcune possibili ragioni includono: qualità subottimale di alcuni item (che vanno rivisti o eliminati), numero insufficiente di item affidabili per ciascun fattore, oppure elementi degli scenari non ben accolti dai partecipanti stessi (negli scenari presentati le risposte cognitive/affettive/prosociali possono essere influenzate da un fattore generale di coinvolgimento empatico situazionale e sensibilità agli argomenti più che da tre costrutti separati).

Analisi dell'Affidabilità Interna

Si è valutata l'affidabilità interna del MET in termini di consistenza delle risposte agli item, utilizzando l' α di Cronbach e l' ω di McDonald, sia a livello di scala totale sia considerando l'impatto di ciascun item. Data la struttura multidimensionale prevista, si è calcolata inizialmente l'affidabilità dell'intera batteria di 18 item (trattandola come misura aggregata di empatia generale) e successivamente si è esplorato l'effetto di ciascun item su tale indice.

Consistenza interna della scala complessiva

Il MET completo (18 item) ha mostrato una consistenza interna insufficiente, con α di Cronbach = 0.557 e ω di McDonald = 0.621. Questi valori sono ben al di sotto del convenzionale 0.70 considerato accettabile per una scala psicometrica, indicando che gli item complessivamente non mostrano elevata inter-correlazione. In altri termini, le risposte ai diversi item del MET sono piuttosto eterogenee e il test fatica a comportarsi in modo coerente.

Contributo dei singoli item e possibile miglioramento

Un'analisi dettagliata "*if item dropped*" (ossia calcolando α e ω omettendo di volta in volta ciascun item) offre ulteriori indicazioni su quali item siano più problematici per la coesione interna:

- La rimozione dell'item 3.2 (affettivo, scenario 3) porterebbe α da 0.557 a 0.603 e ω da 0.621 a 0.645. Questo è il miglioramento più netto ottenuto rimuovendo un singolo item. Ciò suggerisce che 3.2, così com'è, indebolisce significativamente la consistenza interna e che correggerlo (o eliminarlo) gioverebbe alla scala.
- Anche eliminare l'item 4.1 (cognitivo, scenario 4) aumenterebbe α (da 0.557 a 0.580). Questo item, come visto nelle analisi fattoriali, è anomalo; qui notiamo che contribuisce poco alla coerenza interna e la sua assenza aumenterebbe leggermente la consistenza. È un candidato da rivedere o rimuovere.
- L'item 6.2 (affettivo, scenario 6) se rimosso alzerebbe α a 0.569, segnalando un piccolo beneficio. Anche 6.2 era risultato problematico in termini fattoriali (bassa MSA, nessun loading); il suo contenuto andrebbe rivalutato.
- La maggior parte degli altri item, se eliminati, farebbero diminuire l' α o la lascerebbero quasi invariata, suggerendo un apporto di informazioni utili e correlate alle scale. Ad esempio, rimuovendo item considerati "buoni" come 3.1 o 1.2 l' α scende (0.51), indicando che questi item condividono varianza significativa con gli altri e contribuiscono positivamente alla consistenza.

- Un caso particolare: rimuovendo l'item 2.2 (affettivo, scenario 2) l' α rimane identica a 0.557. Ciò implica che 2.2 è quasi del tutto scollegato dal resto della scala (correlazione item-totale ~ 0), ma non essendo invertito non disturba attivamente la consistenza, semplicemente non la aiuta. Questo risultato concorda con l'unicità molto alta osservata per 2.2 in EFA. Si potrebbe considerare di eliminare o riformulare anche questo item, poiché sembra apportare scarso valore aggiunto.

Dall'analisi dei singoli item emergono alcuni elementi critici già intravisti nelle analisi precedenti:

- Item da rivedere per scarso contributo: 4.1 (cognitivo) e 6.2 (affettivo) appaiono scarsamente correlati al resto del test e ne abbassano l'affidabilità. Potrebbero essere candidati per revisione o esclusione.
- Item con nessun contributo: 2.2 (affettivo) praticamente non condivide varianza col resto; pur non abbassando α , la sua utilità è dubbia ai fini della scala.
- Coerenza interna delle sottoscale: calcolando separatamente α per i 6 item cognitivi, affettivi e prosociali, si ottiene rispettivamente $\alpha_{\text{cog}} = 0.38$, $\alpha_{\text{aff}} = 0.01$, $\alpha_{\text{pro}} = 0.24$. Ciò conferma che, allo stato attuale, nessuna delle tre componenti possiede sufficiente consistenza interna.

L'affidabilità bassa suggerisce che le differenze inter-individuali osservate nel punteggio MET totale sono fortemente inficiate da errore di misura e specificità dei singoli item/scenario. In altri termini, un partecipante potrebbe fare bene in certi item e male in altri in modo poco coerente, rendendo il punteggio totale inattendibile.