



UNIVERSITÀ DEGLI STUDI DI PADOVA

Dipartimento di Psicologia dello Sviluppo e della Socializzazione (DPSS)

Corso di laurea triennale in Scienze psicologiche dello sviluppo, della personalità
e delle relazioni interpersonali

Elaborato finale

Proprietà psicometriche dell'Affective Reactivity Index (ARI): Una revisione sistematica della letteratura.

*Psychometric properties of the Affective Reactivity Index (ARI):
A systematic review.*

Relatrice

Prof.ssa Tatiana Marci

Laureanda:
Vanessa Almici

Matricola:
2049472

Anno Accademico 2023/2024

INDICE

INTRODUZIONE	3
CAPITOLO I – L’IRRITABILITÀ	4
1.1 Il dibattito nosologico sul disturbo bipolare in età evolutiva	4
1.2 Studi longitudinali	6
1.3 Uno studio sull’ereditabilità familiare	7
1.4 Patofisiologia dell’irritabilità	8
1.5 Il disturbo da disregolazione dell’umore dirompente (DMDD)	9
CAPITOLO II – L’AFFECTIVE REACTIVITY INDEX (ARI)	11
2.1 Sviluppo ed evoluzione dello strumento	12
2.2 Proprietà psicometriche dell’ARI	13
2.2.1 ARI: versione self- e parent-report	14
2.2.2 ARI: versione teacher-report	17
2.3 Sintesi e conclusioni	19
CAPITOLO III – PROPRIETÀ PSICOMETRICHE DELL’ARI	21
3.1 Struttura fattoriale	22
3.2 Affidabilità	24
3.3 Validità	27
CAPITOLO IV – CONCLUSIONI E DISCUSSIONE	34
4.1 Prospettive future	36
BIBLIOGRAFIA	38
APPENDICE A	42
APPENDICE B	44

INTRODUZIONE

Negli ultimi anni, l'irritabilità rappresenta un fenomeno psicologico complesso che ha attirato un'attenzione crescente nella ricerca clinica e ha suscitato un dibattito nosologico su quale posizione debba avere nella psicopatologia dell'età evolutiva.

Questa tesi si propone di condurre una revisione sistematica della letteratura scientifica relativa all'Affective Reactivity Index ARI; Stringaris et al., 2012), uno strumento concepito per valutare l'irritabilità da una prospettiva dimensionale, in bambini e adolescenti di età compresa tra i 6 e i 17 anni. Il questionario è disponibile in diverse modalità: self-report (somministrabile a partire dagli 11 anni) e parent-/teacher-report.

Il primo capitolo della tesi è dedicato a un'analisi teorica e clinica dell'irritabilità, esaminando il dibattito nosologico nato tra la metà degli anni '90 e i primi 2000, lo sviluppo della sindrome della disregolazione severa dell'umore e il ruolo che ha oggi l'irritabilità dal punto di vista clinico e nosologico. Questa prima parte, sebbene concisa, fornisce una base essenziale per comprendere l'importanza della misurazione dell'irritabilità e la prospettiva dimensionale dell'Affective Reactivity Index (ARI – Stringaris et al., 2012).

Il secondo e il terzo capitolo presentano la scala e compiono una revisione sistematica rispetto agli studi esistenti che hanno esaminato le proprietà psicometriche dell'ARI, dal punto di vista della struttura fattoriale, dell'affidabilità, della validità di contenuto, costruito e di criterio.

In conclusione, la tesi sottolinea l'importanza di una valutazione accurata delle proprietà psicometriche dell'ARI e propone direzioni future per la ricerca, evidenziando la necessità di ulteriori studi per approfondire la comprensione di tale costrutto psicologico.

CAPITOLO I

L'IRRITABILITÀ

L'irritabilità è definita dall'APA Dictionary of Psychology come uno stato eccessivo e facilmente provocato di frustrazione, fastidio o impazienza. È uno degli stati d'animo principali di cui l'uomo fa esperienza e a cui è in grado di dare un nome. Essa è caratterizzata dalla componente affettiva della rabbia e da manifestazioni comportamentali aggressive come scatti collerici (Stringaris, 2011).

Da ormai diversi anni si ritiene che alti livelli di irritabilità rappresentino un fattore di rischio per la psicopatologia in bambini e adolescenti. Uno studio epidemiologico (Brotman et al., 2006), condotto su 1420 bambini e adolescenti tra i 9-19 anni, ha evidenziato come il 3,3% di essi soffrisse di disregolazione severa dell'umore (Severe Mood Dysregulation – SMD), coincidente con alti livelli di irritabilità. Inoltre, lo studio sottolinea come la disregolazione dell'umore nella tarda infanzia ($10,6 \pm 1.4$ anni di età) aumenti il rischio di disturbi depressivi nella prima età adulta. Pertanto, alti livelli di irritabilità possono avere un'influenza clinicamente significativa nel definire le traiettorie di sviluppo dell'individuo.

Nonostante l'interesse crescente della comunità scientifica, la ricerca e la letteratura riguardanti l'irritabilità sono ancora relativamente limitate.

Il primo capitolo di questa tesi si propone di offrire una panoramica sul costrutto di irritabilità nella letteratura scientifica, analizzandone l'evoluzione nel tempo e le implicazioni cliniche ad esso associate.

1.1 Il dibattito nosologico sul disturbo bipolare in età evolutiva

La necessità di intraprendere uno studio approfondito sul costrutto dell'irritabilità è emersa a seguito dell'osservazione di un incremento esponenziale della diagnosi di disturbo bipolare in età infantile e adolescenziale, registrato tra la metà degli anni '90 e l'inizio degli anni 2000 negli Stati Uniti. Invero, secondo i dati annuali registrati dal National Hospital Discharge Survey (NHDS), nel 1996 il tasso di diagnosi infantile di disturbo bipolare era di 1,3 ogni 10.000 bambini, numero salito a 7,3 ogni 10.000 nel

2004. Allo stesso tempo, venne rilevato un aumento di quattro volte delle dimissioni ospedaliere correlate al disturbo bipolare tra gli adolescenti (Blader et al., 2007).

Stando a una revisione della letteratura condotta da Brotman et al. (2017), l'aumento del tasso di diagnosi di disturbo bipolare in giovane età alimentò un dibattito nosologico, specie considerando come molti di questi giovani non presentassero conclamati episodi di mania o ipomania, bensì una sintomatologia caratterizzata da alti livelli di irritabilità e iper arousal. Da una parte, questa revisione suggerisce l'ipotesi che – prima della metà degli anni '90 – il disturbo bipolare fosse sotto-diagnosticato in età giovanile; tuttavia, la spiegazione più plausibile secondo Brotman et al. è un allargamento dei criteri diagnostici utilizzati nella pratica clinica, incorporando ai classici episodi maniacali o ipomaniacali, anche sintomi non episodici di irritabilità e iper arousal. La ricerca arrivò a dimostrare come questa incorporazione fosse inesatta.

Ellen Leibenluft, psichiatra e ricercatrice americana, ha dedicato molto del suo lavoro allo studio del disturbo bipolare in età evolutiva e al dibattito circa i limiti diagnostici (Leibenluft et al., 2003); oltre che allo studio dell'irritabilità grave (Leibenluft, 2011), arrivando alla creazione di una sindrome chiamata “disregolazione severa dell'umore” (Severe Mood Dysregulation – SMD). Lo scopo del definire la SMD non era la creazione di una nuova vera e propria categoria diagnostica, bensì di una categoria differenziata dal disturbo bipolare e utile a testare se esistesse una relazione tra l'irritabilità grave e la mania/ipomania. In altre parole, l'obiettivo era verificare se un individuo di età evolutiva con una sintomatologia di alta irritabilità potesse essere considerato bipolare. I criteri diagnostici stabiliti da Leibenluft et al. sono forniti nell'Appendice A.

A partire dal 2002 (Leibenluft, 2011), presso il National Institute of Mental Health (NIMH) sono stati studiati 146 bambini con grave disregolazione dell'umore (SMD) di età media 11,7 anni, con i genitori che in media collocavano a 7 anni prima l'insorgenza dei sintomi. La diagnosi di SMD è stata fatta utilizzando un modulo aggiuntivo all'interno della Schedule for Affective Disorders and Schizophrenia-Present and Lifetime Version (K-SADS-PL). All'interno dello studio è stato esaminato anche un gruppo di 107 giovani individui con disturbo bipolare. Entrambi i gruppi (SMD e disturbo bipolare) sono stati sottoposti alla Children's Global Assessment Scale (CGAS). Il gruppo di bambini diagnosticati con severa disregolazione dell'umore ha ottenuto all'interno della scala un

punteggio medio di 45.8 (DS= 6.9); mentre il gruppo di bambini con disturbo bipolare ha ottenuto un punteggio medio di 46.5 (DS=12.4). Sebbene il gruppo del disturbo bipolare dimostri una maggiore variabilità nella distribuzione dei punteggi, il confronto tra i punteggi medi dei due gruppi suggerisce come sia il disturbo bipolare che la disregolazione severa dell'umore siano molto compromettenti per i partecipanti

La revisione della ricerca condotta da Leibenluft (2011) ha permesso di identificare delle differenze significative tra la categoria diagnostica del disturbo bipolare e la sindrome della disregolazione severa dell'umore, riscontrabili attraverso studi longitudinali, l'approfondimento dell'ereditabilità familiare e approfondimenti patofisiologici.

1.2 Gli studi longitudinali

Gli studi longitudinali rappresentano un metodo empiricamente efficace per valutare la validità nosologica dell'incorporare una sintomatologia di forte irritabilità non episodica e iper arousal all'interno dei criteri per il disturbo bipolare in età evolutiva.

Due sindromi pediatriche, infatti, sono considerate patofisiologicamente simili quando presentano un decorso simile e prevedono rischi simili per il fenotipo adulto. Ciò significa che – nel valutare se la disregolazione severa dell'umore sia una manifestazione evolutiva della mania – questione cruciale è stabilire se gli individui con questa sindrome, sviluppino durante l'adolescenza e l'età adulta veri e propri episodi maniacali o ipomaniacali, o un disturbo bipolare non meglio specificato (Leibenluft, 2011).

Uno studio longitudinale condotto da Stringaris et al. (2010) ha seguito 93 giovani partecipanti affetti da disturbo bipolare di tipo I (76%) o tipo II (24%), diagnosticati secondo i criteri del DSM-IV e 84 giovani partecipanti con sindrome di disregolazione severa dell'umore (SMD). La presenza/assenza di episodi maniacali o ipomaniacali o misti è stata valutata tramite alcuni moduli della K-SADS, con un intervallo di follow-up di 6 mesi. I risultati hanno evidenziato come solo un partecipante appartenente al gruppo della sindrome SMD ha esibito un episodio maniacale, ipomaniacale o misto durante lo studio (mediana di follow-up: 28.4 mesi). Per quanto riguarda il gruppo di partecipanti con disturbo bipolare, 58 di loro (il 62,4%) durante i follow-up hanno presentato almeno un episodio maniacale, ipomaniacale o misto. Pertanto, nel campione di soggetti con

SMD, la probabilità di sviluppare un episodio maniacale, ipomaniacale o misto era 50 volte minore che nei soggetti diagnosticati con disturbo bipolare.

In un altro studio epidemiologico, già citato nell'introduzione a questo primo capitolo (Brotman et al., 2006), è stato constatato come soffrire di disregolazione severa dell'umore a $10,6 \pm 1.4$ anni di età aumenti di 7 volte il rischio di soddisfare i criteri diagnostici per la depressione unipolare all'età di 18,3 anni. All'interno del grande campione utilizzato per questo studio (1420 individui tra i 9-19 anni), il 3,3% soddisfaceva i criteri per la SMD, mentre solo lo 0,1% rispettava i criteri diagnostici del disturbo bipolare.

Il disturbo bipolare in età evolutiva risulta essere più raro della sindrome SMD e il campione di Brotman et al. – a differenza dello studio di Stringaris et al. (2010) – era stato reclutato senza ricercare esplicitamente individui diagnosticati con disturbo bipolare. Di conseguenza, lo studio di Brotman et al. non opera un confronto tra SMD e disturbo bipolare. Questo studio, tuttavia, offre uno sguardo interessante rispetto ai potenziali esiti psicopatologici della SMD: un aumentato rischio per la depressione unipolare in età adulta, ma non per il disturbo bipolare.

1.3 Uno studio sull'ereditabilità familiare

In uno studio creato ad hoc (Brotman et al., 2007), è stata condotta un'indagine approfondita riguardante la familiarità della disregolazione severa dell'umore, con l'obiettivo di metterla a confronto con quella del disturbo bipolare. Il presupposto fondamentale di questo studio era che, se la sindrome della disregolazione severa dell'umore (SMD) è davvero un fenotipo del disturbo bipolare, allora ci si aspetta che la familiarità delle due condizioni sia molto simile, ossia che la probabilità di un bambino con SMD di avere un genitore con disturbo bipolare sia la stessa di un bambino che presenta il disturbo bipolare (Leibenluft, 2011). Lo studio ha riscontrato una differenza significativa circa la prevalenza del disturbo bipolare nei genitori del campione di bambini preso in esame: 33,3% nel campione di bambini con disturbo bipolare pediatrico e 2,7% nel campione di bambini con disregolazione severa dell'umore.

Per quanto questo studio presenti dei limiti significativi – come la dimensione ridotta del campione e il bias di accertamento – esso rappresenta un ulteriore passo con

cui la comunità scientifica ha riscontrato la necessità di approfondire lo studio dei sintomi della disregolazione severa del tono dell'umore, in particolare l'irritabilità, e di separarli dall'inquadramento nosologico del disturbo bipolare.

1.4 Patofisiologia dell'irritabilità

Secondo un'analisi della letteratura operata da Rao (2014), Leibenluft e collaboratori (Leibenluft, 2011) condussero una serie di studi pilota osservando i potenziali evento-correlati e con strumenti di misurazione quali la risonanza magnetica funzionale (fMRI) e la magnetoencefalografia (MEG) per confrontare soggetti sani (gruppo di controllo), soggetti affetti da disturbo bipolare e soggetti che soddisfacevano criteri per la SMD. I risultati comportamentali di queste ricerche hanno evidenziato come entrambi i gruppi sperimentali presentassero difficoltà nel riconoscere le emozioni facciali, nella tolleranza alla frustrazione e nella modifica delle risposte comportamentali apprese.

In uno studio sulla risposta alla frustrazione (Rich et al., 2007) entrambi i gruppi sperimentali (bipolare e SMD) hanno svolto un compito attentivo (affective Posner Task) e sono stati confrontati con un gruppo di controllo, rispetto al quale hanno mostrato un arousal maggiore quando sottoposti a frustrazione. I due gruppi sperimentali, tuttavia, si differenziavano tra loro per quanto riguarda la performance comportamentale e psicofisiologica. I soggetti giovani con disturbo bipolare – quando sottoposti alla condizione di frustrazione – mostravano una riduzione dell'attenzione esecutiva (rappresentata da onde P3 parietali diminuite). I soggetti con SMD – in qualunque condizione, frustrante e non – evidenziavano difficoltà nei processi attentivi precoci (con onde N1 e P1 parietali, temporali e centrali diminuite).

In uno studio successivo (Rich et al., 2011), il task attentivo di Posner è stato sottoposto a 60 individui tra gli 8-17 anni (20 con disturbo bipolare, 20 con SMD e 20 come gruppo di controllo). Lo studio prevedeva tre condizioni in cui – dopo ogni stimolo a cui il partecipante rispondeva – veniva manipolato il feedback. La condizione 1 prevedeva un feedback non-emozionale sull'accuratezza della risposta (“Good Job!” o “Incorrect!”). La condizione 2 introduceva delle contingenze in cui il soggetto vinceva o perdeva denaro in base alla performance (“Great Job! Win 25 Cents” o “Wrong! Lose 25

Cents”). La condizione 3 manipolava il feedback delle risposte corrette fingendo di calcolare la velocità di risposta: (“Too Slow! Lose 25 Cents” per il 56% delle risposte corrette; “You’re Quick! Win 25 Cents” dopo il 44% di risposte corrette), le risposte sbagliate risultavano sempre in una perdita in denaro. Lo studio ha analizzato l’attività neurale a seguito dei feedback tramite la magnetoencefalografia (MEG); sono stati, inoltre, analizzati il tempo di reazione, l’accuratezza delle risposte e la risposta affettiva del partecipante dopo il completamento del compito (misurazione self-report).

I risultati hanno riscontrato nei partecipanti con SMD un aumento dell’arousal a seguito dei feedback negativi maggiore che negli altri due gruppi. Inoltre, in risposta ai feedback negativi, nel gruppo SMD è stata rilevata una maggiore attivazione della corteccia cingolata anteriore e nella circonvoluzione frontale mediale. Il gruppo del disturbo bipolare, diversamente, a seguito dei feedback negativi subiva una maggiore attivazione nella circonvoluzione frontale superiore e una minore attivazione del lobo dell’insula.

Un altro studio (Brotman et al., 2010) ha riscontrato una differenza tra soggetti giovani con disturbo bipolare e soggetti giovani con SMD in un compito di elaborazione delle emozioni facciali: i soggetti con disregolazione severa del tono dell’umore, infatti, presentavano un’attivazione minore dell’amigdala sia rispetto al gruppo sperimentale del disturbo bipolare, che rispetto al gruppo di controllo.

1.5 Il disturbo da disregolazione dell’umore dirompente (DMDD)

Sono numerosi i disordini mentali caratterizzati da alti livelli di irritabilità: il disturbo oppositivo-provocatorio, l’ADHD, i disturbi d’ansia, il disturbo di depressione maggiore e il disturbo bipolare (Brotman et al., 2017).

Il disturbo da disregolazione dell’umore dirompente (Disruptive Mood Dysregulation Disorder – DMDD) si contraddistingue da questi disturbi, poiché ha come sintomo primario un livello elevato di irritabilità. Esso si colloca nella sezione dei disturbi depressivi del DSM-5 – considerati in tutto l’arco della vita – nonostante il suo esordio è riscontrabile nell’infanzia (Rao, 2014).

Questo disturbo è nato a seguito degli studi sulla disregolazione severa dell'umore (SMD), come tentativo di creare una vera e propria categoria diagnostica ad hoc per i bambini che soffrono di un eccessivo livello di irritabilità.

I criteri diagnostici del DMMD (forniti nell'Appendice B) sono basati sulla SMD, tuttavia, il DMDD si differenzia dalla disregolazione severa dell'umore (SMD), poiché non presenta tra i propri criteri i sintomi di iper arousal. L'inserimento del disturbo dell'umore dirompente nel DSM-5 ha scatenato un dibattito nella comunità scientifica (Rao, 2014), poiché molta della evidenza empirica a suo sostegno proviene da studi sulla disregolazione severa dell'umore e in cui non sono stati utilizzati i nuovi criteri del DSM-5. Pertanto, si può dire che il dibattito nosologico sull'irritabilità rimane più vivo che mai.

Stringaris – precedentemente alla nascita del DMDD – si è espresso su quella che per lui dovrebbe essere la posizione dell'irritabilità nella nosologia psichiatrica, sottolineando come l'irritabilità sembri possedere le proprietà di una dimensione nella psicopatologia: essa rientra in un'ampia gamma di patologie psichiatriche e sembra poter essere misurata attraverso un continuum.

Usare un approccio dimensionale piuttosto che categoriale per l'irritabilità, secondo Stringaris, potrebbe aiutare a semplificare la diagnosi delle psicopatologie dove essa è presente come sintomo, togliendola dai criteri diagnostici ed evitando così le comorbidità artificiali. Inoltre, incorporando all'interno del DSM-5 una scala specifica per misurare l'irritabilità, si fornirebbe ai clinici uno strumento utile sia in fase di assessment che per valutare l'efficacia dei trattamenti (Stringaris, 2011). Questa prospettiva portò Stringaris et al. (2012) alla creazione dell'Affective Reactivity Index (ARI).

CAPITOLO II

L’AFFECTIVE REACTIVITY INDEX (ARI)

Come già menzionato nel capitolo precedente, l’idea di sviluppare uno strumento di misurazione specifico per l’irritabilità è nata dal dibattito nosologico circa la posizione dell’irritabilità all’interno della psicopatologia. Nella sua revisione (Stringaris, 2011), Stringaris ha sottolineato l’importanza di adottare un approccio dimensionale, soprattutto considerando che – al momento della pubblicazione del suo articolo – il DMS-5 (APA, 2013) non era ancora disponibile e il DSM-IV (APA, 2000) si basava prevalentemente su un modello categoriale per la diagnosi. Inoltre, il DSM-IV incorporava l’irritabilità nella sintomatologia di numerosi disordini mentali, senza tuttavia mai fornire una definizione specifica.

L’Affective Reactivity Index (ARI – Stringaris et al., 2012) si basa su una definizione molto semplice e ampia di irritabilità, intesa come uno stato d’animo di facile fastidio e suscettibilità, caratterizzato da scoppi di rabbia e collera (“*a mood of easy annoyance and touchiness characterized by anger and temper outbursts*” – Stringaris, 2011; Stringaris et al., 2012).

Nella Sezione III del DSM-5 (APA, 2013), è presente un adattamento dell’ARI proposto dall’American Psychological Association (APA), creato utilizzando i medesimi items della versione originale (ARI; Stringaris et al., 2012). Questa nuova versione dello strumento rientra nella categoria “Scale di valutazione dei sintomi trasversali”, nate con lo scopo di introdurre un’indagine dimensionale dei sintomi e criteri diagnostici. L’irritabilità è, invero, un sintomo trasversale a molteplici quadri clinici, sia esternalizzanti che internalizzanti, tra cui i disturbi comportamentali, l’ansia e la depressione. Sebbene l’adattamento dell’ARI proposto dall’APA non sia stato validato dagli autori della scala originale, esso ha dimostrato una buona affidabilità a livello applicativo in conformità con i criteri riportati dal DSM-5 (Narrow et al., 2013).

L’adattamento dell’ARI è presente anche nella Sezione III del DSM-5-TR (APA, 2022), sia in forma self-report che in forma parent/guardian report:

- LEVEL 2—Irritability—Child Age 11–17 (Affective Reactivity Index [ARI])
- LEVEL 2—Irritability—Parent/Guardian of Child Age 6–17 (Affective Reactivity Index [ARI]).

Questo capitolo è dedicato alla presentazione della struttura, dello sviluppo e dei primissimi studi sulle proprietà psicometriche dell’Affective Reactivity Index (ARI – Stringaris et al., 2012).

2.1 Struttura e sviluppo dell’ARI

L’Affective Reactivity Index (ARI – Stringaris et al., 2012) è uno strumento di misurazione dell’irritabilità di bambini e adolescenti tra i 6 e i 17 anni. Esso è disponibile in due versioni: *self-report* – somministrabile a partire dagli 11 anni di età – e *parent/teacher report*.

Le istruzioni di compilazione sono molto simili nelle due versioni: viene chiesto al genitore/insegnante/bambino/adolescente il grado di rispecchiamento negli ultimi 6 mesi rispetto alla descrizione dei sentimenti e comportamenti indicati dagli items. Ai genitori o insegnanti vengono presentate le seguenti istruzioni: *“In the last 6 months and compared to others of the same age, how well does each of the following statements describe the behavior/feelings of your child? Please try to answer all questions”*. Le istruzioni della versione *self-report* sono quasi identiche: varia, naturalmente, la parte finale della domanda che diventa: *“your behaviours/feelings”*.

La scala è stata progettata per essere semplice e concisa; infatti, è composta da soli sette items. I primi sei indagano la sintomatologia legata all’irritabilità, sia sul piano emotivo che comportamentale: la facilità con cui il bambino/adolescente si sente infastidito dagli altri, se perde spesso la calma, se rimane arrabbiato a lungo, se è arrabbiato la maggior parte del tempo, se perde la calma frequentemente, se si arrabbia facilmente. L’ultimo item valuta il grado percepito di compromissione del funzionamento del bambino/adolescente (*“overall, irritability causes him/her – o “me” nella versione self report – problems”*).

Ai compilanti è richiesto di indicare il loro livello di accordo su una scala Likert a 3 punti con tre opzioni di risposta: “not true”, “somewhat true”, “certainly true” che ricevono, rispettivamente, un punteggio di 0, 1 e 2 punti. Il punteggio grezzo totale è dato dalla somma dei punteggi dei primi sei items e può variare tra 0 e 12 punti – viene escluso dal conteggio l’item finale della compromissione del funzionamento. Punteggi alti corrispondono a livelli di irritabilità più severi.

Lo studio pilota dell’Affective Reactivity Index è stato effettuato su un campione di 80 soggetti statunitensi, divisi tra gruppo clinico e gruppo di controllo. La prima versione del questionario era più estesa e composta da 21 items; da questa prima versione è stata fatta una selezione tenendo conto della ridondanza di alcuni items e assicurandosi che gli items rimanenti, per quanto concisi, coprissero adeguatamente vari aspetti dell’irritabilità: durata, frequenza e soglia. Sono stati eliminati gli items non fondamentali alla coerenza interna o non in grado di discriminare adeguatamente tra soggetti clinici e soggetti del gruppo di controllo.

Nella versione della scala contenuta all’interno del DSM-5 (APA, 2013) e DSM-5-TR (APA, 2022) viene richiesto al bambino/adolescente o genitore di valutare gli items in base agli ultimi sette giorni, poiché si tratta di uno strumento pensato per essere utilizzato all’interno di una batteria – insieme ad altri strumenti di Livello 2 – con cui valutare la risposta del paziente al trattamento. È fondamentale, in aggiunta al punteggio grezzo, il calcolo punteggio medio totale (*average total score*), utilizzabile sia nella ricerca che nella pratica clinica. Esso si ottiene dividendo il punteggio grezzo per 6 (corrispondente al numero degli items considerati). Il valore ottenuto (*average total score*) può variare tra 0 e 3 e permette al clinico o al ricercatore di valutare il livello di irritabilità del soggetto indagato da nullo (0) a moderato (1) fino a grave (2 in su).

L’utilizzo dell’*average total score* si è rivelato essere non solo affidabile, ma anche pratico e rilevante sia nel lavoro di ricerca che nella pratica clinica, in accordo con i criteri riportati dal DSM-5 (Narrow et al., 2013).

2.2 Proprietà psicometriche dell’ARI

Le caratteristiche psicometriche dell’Affective Reactivity Index sono state oggetto di studio in diversi paesi del mondo, tra cui gli Stati Uniti, il Regno Unito, la Cina, il Brasile, l’Australia e la Spagna. Gli studi, per quanto ancora non numerosi, hanno riportato risultati promettenti sia in termini di validità che di affidabilità dello strumento.

Il terzo capitolo si concentrerà in maniera più approfondita su una revisione sistematica degli studi attualmente disponibili; tuttavia, in questo secondo capitolo verranno già presentati i primi studi riguardanti le proprietà psicometriche delle versioni

self-report e parent-report (Stringaris et al., 2012) e della versione teacher-report. (Ezpeleta et al., 2020).

2.2.1 ARI: versioni self- e parent-report

Il primissimo studio sulle proprietà psicometriche dell’Affective Reactivity Index (Stringaris et al., 2012) ha preso in considerazione un campione di bambini statunitensi e britannici di età compresa tra i 5-6 e i 18 anni.

Il campione statunitense ($n=218$) è stato scelto direttamente tra i partecipanti di un altro studio sulla disregolazione severa dell’umore (SMD), condotto presso il National Institute of Mental Health (NIMH) e già menzionato nel capitolo I (Brotman et al., 2007; Stringaris et al., 2010). Il campione era composto da 67 bambini con SMD, 39 bambini con disturbo bipolare di tipo I o II (reclutati dallo studio del NIMH), 35 bambini appartenenti a famiglie a rischio per il disturbo bipolare e 77 volontari sani (entrambi reclutati tramite pubblicità). A partire dal campione di 218 bambini, sono stati ottenuti i dati sulla versione parent-report per 214 soggetti (98%), i dati self-report di 194 soggetti (89%) e i dati di entrambe le versioni del questionario per 192 soggetti (88%). L’età media era 12.90 anni ($DS = 2.70$; range 6–17) con il 60% di maschi e il 40% di femmine.

Il campione britannico è stato diviso in gruppo clinico e gruppo di controllo. Il gruppo clinico ($n=34$) era composto da bambini e adolescenti tra i 5 e i 17 anni appartenenti ai Servizi di Salute Mentale Infantile e Adolescenziale della South West London & St Georges Mental Health NHS Trust. Le diagnosi variavano dal disturbo oppositivo provocatorio (ODD, 15%, $n=5$), all’ADHD (15%, $n=5$), al disturbo dello spettro autistico (ASD, 9%, $n=3$). Inoltre, il 9% dei soggetti ($n=3$) presentava comportamenti auto-lesivi senza diagnosi specifica. Il gruppo di controllo ($n=54$) era composto da bambini e adolescenti tra i 6 e i 18 anni, reclutati in una scuola primaria e in tre scuole secondarie appartenenti alla stessa area geografica del gruppo clinico. L’età media dell’intero campione britannico era di 11.70 anni ($DS = 3.46$, range 5–18) con il 59% composto da maschi e il 41% da femmine.

Dal punto di vista descrittivo, innanzitutto, i risultati hanno rilevato che l’item “*easily annoyed by others*” ha ottenuto il punteggio medio più elevato, mentre “*angry most of the time*” ha registrato il punteggio medio più basso.

In entrambe le versioni (*self-* e *parent-report*) e in entrambi i campioni (statunitense e britannico) è stata valutata la presenza di un effetto del genere; tuttavia, non sono risultate differenze significative nei livelli di irritabilità di maschi e femmine.

Struttura fattoriale – L'*Affective Reactivity Index* (ARI - Stringaris et al. (2012)) è stato analizzato dagli autori mediante un'analisi fattoriale confermativa (CFA) allo scopo di verificare l'adeguatezza dell'impiego di una struttura a un fattore, per la quale sono stati impiegati diversi indici di fit:

- Il Comparative Fit Index (CFI), in cui un valore uguale o superiore a 0,95 indica un buon adattamento; Il Tucker-Lewis Index (TLI), in cui valori vicino all'1 indicano un buon adattamento; Il Root Mean Square Error of Approximation (RMSEA), in cui valori $> 0,05$ indicano un buon adattamento; Il Weighted Root Mean Square Residual (WRMR), soglia raccomandata a 0,6 (Yu, 2002).

I risultati degli indici CFI, TLI, WRMR hanno evidenziato come la soluzione a un fattore sia adeguata; tuttavia, l'indice RMSEA delle scale *self-report* risulta superiore ai valori raccomandati (RMSEA= 0.09 nel campione statunitense e RMSEA= 0.21 in quello britannico).

Affidabilità – Dallo studio di Stringaris et al. (2012) è emersa un'elevata consistenza interna: nel campione statunitense, infatti, l'alpha di Cronbach era di 0.88 nella versione *self-report* e 0.92 nella *parent-report*. Per le versioni *self-* e *parent-report* del campione britannico, l'alpha di Cronbach era rispettivamente di 0.90 e 0.89. I punteggi totali di entrambe le versioni erano maggiori nel campione statunitense rispetto a quello britannico. L'alta consistenza interna è significativa, poiché indica come gli items siano in grado di valutare in modo adeguato le diverse sfaccettature del costrutto d'interesse (in questo caso l'irritabilità).

È emersa, inoltre, una correlazione alta e significativa ($p < .001$) tra i punteggi di genitori e figli, con valore di $r = 0.58$ per il campione statunitense e $r = 0.73$ per il campione britannico. Questo risultato è importantissimo, poiché indica nell'ARI una buona capacità nel fornire una misurazione affidabile dell'irritabilità in una prospettiva *multi-informant*.

Il coefficiente di correlazione per la stabilità longitudinale (misurata su un periodo medio di circa un anno) è risultato elevato e significativo per l'ARI *parent-report* ($r = 0.88$, $p < 0.001$), ma non per la versione *self-report* ($r = 0.29$, $p = 0.28$).

Validità – Un punto di forza dello studio di Stringaris et al. è che i risultati hanno fatto emergere in entrambe le versioni dell'ARI una buona capacità di discriminazione tra soggetti sani e soggetti affetti da psicopatologia o con disregolazione severa dell'umore, i quali riportavano punteggi di irritabilità più elevati. Inoltre, nella versione *parent-report*, i bambini e adolescenti affetti da disturbo bipolare hanno riportato punteggi di irritabilità inferiori rispetto ai bambini e adolescenti che soffrivano di disregolazione severa dell'umore.

In aggiunta a ciò, sono state condotte delle analisi esplorative limitate al campione britannico. Ai partecipanti con età uguale o maggiore degli 11 anni e ai genitori è stato sottoposto – oltre all'ARI – anche *lo Strengths and Difficulties Questionnaire* (SDQ, Goodman, 2001): un questionario di 25 items che valuta i comportamenti internalizzanti, esternalizzanti e prosociali in bambini e adolescenti di età compresa tra 4 e 16 anni. Il questionario è suddiviso in 5 sottoscale:

1. Difficoltà emotive;
2. Problemi di condotta;
3. Iperattività e disattenzione;
4. Problemi con i pari;
5. Comportamenti prosociali.

L'obiettivo era testare la validità convergente tra lo SDQ e l'ARI. L'ipotesi di partenza era che alti livelli di irritabilità fossero associati a problemi emotivi, piuttosto che a problemi comportamentali o antisociali: Stringaris et al. sostengono, invero, che uno stato d'animo irritabile non motiverà necessariamente il bambino a compiere un'azione aggressiva. L'aggressività e l'ostilità possono manifestarsi in bambini con alta irritabilità, ma non obbligatoriamente. (Stringaris et al., 2012).

Oltre all'utilizzo di correlazioni per fini esplorativi, l'ipotesi è stata testata tramite diversi modelli di regressione. In un caso, la variabile di outcome era la sottoscala relativa alle difficoltà emotive dello SDQ, mentre tutte le altre sottoscale venivano utilizzate come variabili predittive insieme al punteggio totale dell'ARI; nell'altro caso, la variabile outcome era la sottoscala dei problemi di condotta e tutti gli altri punteggi precedentemente menzionati (compresa la sottoscala delle difficoltà emotive) sono stati utilizzati come predittori. Le variabili outcomes e predittive delle versioni *self-report*, inoltre, venivano distinte e separate da quelle delle versioni *parent-report*.

I risultati delle analisi correlazionali univariate indicano che l'ARI correla positivamente con tutte le sottoscale dello SDQ, ad eccezione della sottoscala dei problemi con i pari nella versione self-report. In queste prime analisi non è emersa una differenza rilevante nella correlazione dell'ARI con la sottoscala delle difficoltà emotive o con quella dei problemi di condotta.

Tuttavia, i modelli di regressione hanno dimostrato come – secondo i punteggi dell'ARI e dello SDQ *parent-report* – l'irritabilità emerge come unica variabile predittiva delle difficoltà emotive; allo stesso tempo, la sottoscala dell'iperattività e disattenzione si configura come l'unica variabile predittiva nei problemi di condotta.

2.2.2 ARI: versione teacher-report

In uno studio molto recente (Ezpeleta et al, 2020), è stata condotta un'analisi approfondita delle proprietà psicometriche dell'Affective Reactivity Index (ARI) in contesto scolastico, più specificatamente utilizzando gli insegnanti come *informant* circa i livelli di irritabilità dei loro studenti.

Questo studio assume una particolare importanza dal punto di vista applicativo, poiché rende possibile una valutazione *multi-informant* del bambino o adolescente. Dal punto di vista clinico, il confronto dei punteggi dei diversi informatori (genitori, insegnanti e autovalutazione) permette di ottenere una visione globale del funzionamento clinico del paziente. Dal punto di vista della ricerca, lo studio dell'ARI nelle sue diverse versioni consente di delineare in modo preciso le caratteristiche e le proprietà dello strumento.

Il campione era composto da bambini di 7 e 11 anni, facenti parte di uno studio longitudinale spagnolo più ampio, il cui scopo era monitorare il benessere psicologico e i problemi comportamentali dei partecipanti attraverso la somministrazione di diversi strumenti a famiglie e insegnanti, tra i quali: l'ARI, lo SDQ, la *Children's Aggression Scale* (CAS), la *Diagnostic Interview for children and adolescents for Parents of Preschool and Young Children* (DICA-PPYC) e la *Children's Global Assessment Scale* (CGAS).

Ezpeleta et al. hanno analizzato nello specifico un totale di 471 risposte teacher-report per i bambini di 7 anni e 454 risposte per i bambini di 11 anni. Ciascun insegnante

ha valutato da 1 a 11 bambini di 7 anni; oppure da 1 a 15 bambini di 11 anni. Al momento della compilazione, l'insegnante aveva una conoscenza del bambino pari a una media di 11.2 mesi (DS = 4.8) nel caso degli alunni di 7 anni; mentre il tempo medio di conoscenza per i bambini di 11 anni era di 9.4 mesi (DS = 4.7).

I risultati delle analisi descrittive hanno evidenziato che i punteggi di irritabilità nei bambini maschi di 7 anni risultavano leggermente superiori rispetto alle femmine della stessa età; tuttavia, a 11 anni non sono state registrate differenze rilevanti tra maschi e femmine. In aggiunta a questo, i punteggi totali del campione di 7 anni sono risultati mediamente più alti rispetto a quelli del campione di 11 anni.

Struttura fattoriale – Ezpeleta et al. (2020) hanno utilizzato, nella loro analisi fattoriale confermativa (CFA), il metodo WLSMV per esaminare un modello a 6 items unidimensionale con bambini di 7 e 11 anni. I risultati hanno mostrato un adattamento soddisfacente a 7 anni, ma insufficiente a 11 anni (RMSEA = .121). Nel campione di 11 anni, il modello è risultato soddisfacente solo dopo aver preso in considerazione l'elevata correlazione tra gli items 2 (i.e., “*Often loses their temper*”) e 6 (i.e., “*Loses their temper easily*”).

Affidabilità – La consistenza interna dell'ARI si è dimostrata soddisfacente, con un alpha di Cronbach pari a 0.85 per i bambini di 7 anni e 0.88 per i bambini di 11 anni (Ezpeleta et al., 2020). Il punteggio grezzo di entrambe le fasce d'età mostrava una correlazione positiva e significativa con l'item 7 (riguardante la compromissione del funzionamento del bambino).

Validità – La validità convergente è stata testata confrontando i punteggi dell'ARI con altri strumenti compilati dagli insegnanti: sono state prese in considerazione due sottoscale dello SDQ (Difficoltà emotive e Problemi di condotta); i punteggi dell'aggressività totale, dell'aggressività verso i pari e il punteggio totale della *Children's Aggression Scale* (CAS); e, infine, il livello di rabbia rilevato tramite la *PROMIS – Anger—Parent/Guardian of Child Age 6–17* (PROMIS, 2016). In generale, i risultati hanno rilevato delle elevate correlazioni tra il punteggio dell'ARI e i punteggi degli altri strumenti considerati. Sono state valutate, inoltre, le associazioni dei punteggi ARI *teacher-report* con i dati *self-report* e *parent-report* di altri strumenti: nel caso della *Child Behaviour Checklist* (CBCL) compilata dai genitori, la correlazione nella valutazione dell'irritabilità era bassa nel campione di bambini di 7 anni (0.17) e leggermente più alta

con i bambini di 11 anni (0.32). È stata rilevata una discreta associazione per l'*Emotion Regulation Checklist* (ERC). Per quanto riguarda, invece, gli strumenti compilati dai bambini stessi – come lo *Youth Self-Report* (YSR) o l'*Anger Questionnaire* (AQ) – l'associazione è risultata molto bassa.

L'associazione dei punteggi dell'ARI *teacher-report* con altri strumenti *teacher-report* è maggiore rispetto a quella con misure ottenute da altri informatori, risultando tanto più forte quanto più il costrutto misurato si avvicina all'irritabilità. Complessivamente, il grado di accordo tra informatori è da considerarsi moderato (Ezpeleta et al., 2020).

Lo studio dimostra, infine, la capacità dell'ARI di discriminare in modo longitudinale e trasversale tra individui con ADHD e ODD e soggetti sani, in base al punteggio totale di irritabilità ottenuto; tuttavia, lo strumento non è sensibile alle differenze tra soggetti che presentano sintomi di ansia e soggetti del gruppo di controllo.

2.3 Sintesi e conclusioni

Entrambi gli studi presentati questo capitolo (Stringaris et al., 2012 e Ezpeleta et al., 2020) mostrano al loro interno diversi punti di forza, così come dei limiti. Innanzitutto, le analisi della struttura fattoriale compiute dai due studi supportano l'unidimensionalità del costrutto dell'irritabilità così come è stata teorizzata da Stringaris (Stringaris, 2011; Stringaris et al., 2012).

Il grande merito dello studio di Stringaris et al. (2012) può essere trovato nell'utilizzo di campioni provenienti da due Stati diversi; nell'impiego di numerosi metodi di analisi e nell'inclusione di partecipanti con diverse diagnosi cliniche.

I limiti dello studio includono principalmente la dimensione ridotta del campione britannico; la natura trasversale dei risultati che limita la possibilità di poter fare inferenze significative e, infine, il fatto che l'ARI sia uno strumento relativamente conciso, che non valuta in profondità il fenomeno dell'irritabilità. Proprio per questo, gli autori sottolineano la necessità di compiere ulteriori studi per esplorare in modo più approfondito il rapporto tra irritabilità e discontrollo emotivo/comportamentale in età evolutiva (Stringaris et al., 2012).

Il punto di forza dello studio di Ezpeleta et al. (2020) è l'introduzione del contesto scolastico – estremamente centrale nella vita del bambino – all'interno della valutazione dell'irritabilità. La validità convergente è stata testata confrontando i punteggi dell'ARI *teacher-report* con altri strumenti che misuravano aspetti affini all'irritabilità (rabbia e difficoltà emotive e/o comportamentali in età evolutiva). L'associazione era soddisfacente quando gli strumenti venivano somministrati allo stesso soggetto e diventava discreta o bassa se il punteggio dell'ARI *teacher-report* veniva confrontato con altre misurazioni *self-* o *parent-report* (Ezpeleta et al., 2020). Questo risultato potrebbe suggerire una capacità limitata dell'ARI nella valutazione dell'irritabilità in ottica *multi-informant*; tuttavia, è importante considerare che le difficoltà del bambino sono molto spesso specifiche a un determinato contesto. Ciò potrebbe spiegare le discrepanze dei punteggi forniti da vari informatori.

Complessivamente, questi due studi rappresentano una prima dimostrazione della potenzialità dell'ARI: uno strumento dalle proprietà psicometriche promettenti in termini di affidabilità e validità e adeguato alla misurazione dell'irritabilità in età evolutiva, sia nel contesto clinico che nel contesto di ricerca.

Il terzo capitolo sarà dedicato a una revisione più approfondita dei vari studi effettuati in merito alle proprietà psicometriche dell'*Affective Reactivity Index* (ARI – Stringaris et al., 2012).

CAPITOLO III

PROPRIETÀ PSICOMETRICHE DELL'ARI

Il terzo capitolo di questo elaborato è dedicato alla revisione sistematica della letteratura esistente riguardante le proprietà psicometriche dell'*Affective Reactivity Index* (ARI – Stringaris et al., 2012).

1. Gli obiettivi specifici di questa revisione sistematica possono essere articolati nei seguenti punti: Analizzare i risultati dei diversi studi relativi alla struttura fattoriale dell'ARI, allo scopo di verificare la validità interna della scala e accertare se essa misuri adeguatamente l'irritabilità come variabile latente;
2. Valutare l'affidabilità e la consistenza interna della scala;
3. Indagare la validità dell'ARI, ponendo particolare attenzione alla presenza di studi riguardanti la validità di contenuto.

Gli articoli accademici sono stati ricercati tramite i database di Scopus e PsychInfo, utilizzando le seguenti parole chiave:

Scopus	(TITLE-ABS-KEY (affective reactivity index) AND TITLE-ABS-KEY (factor AND analysis OR factor AND structure OR psychometric OR propert OR validat OR validit OR reliability))
PsychInfo	TX (affective reactivity index) AND TX (factor analysis OR factor structure OR psychometric OR validity OR reliability)

Tabella 1 – Parole chiave della ricerca bibliografica condotta sui database Scopus e PsychInfo

Dalla ricerca bibliografica condotta, sono emersi 100 articoli, di cui 50 provenienti dal database Scopus e 50 dal database PsychInfo; successivamente, è stato effettuato un confronto dei risultati e sono stati rimossi 16 duplicati identificati su PsychInfo. Il risultato finale ha portato a un totale di 84 articoli: 50 appartenenti al database Scopus e 34 appartenenti a PsychInfo.

Degli 84 articoli identificati, sono stati considerati solo quelli che analizzavano l'ARI in relazione alle sue proprietà psicometriche, a condizione che rispettassero almeno uno degli obiettivi della revisione: pertanto, sono stati selezionati 16 studi che prendevano

in analisi la struttura fattoriale, l'affidabilità e la consistenza interna, la validità e/o la validità di contenuto dell'ARI.

3.1 Struttura fattoriale

Come già anticipato nel capitolo precedente, l'*Affective Reactivity Index* (ARI – Stringaris et al. (2012)) è stato analizzato dagli autori mediante un'analisi fattoriale confermativa (CFA) allo scopo di verificare l'adeguatezza dell'impiego di una struttura a un fattore, per la quale sono stati impiegati diversi indici di fit: CFI, TLI, RMSEA e WRMR.

I risultati degli indici CFI, TLI, WRMR hanno evidenziato come la soluzione a un fattore sia adeguata; tuttavia, l'indice RMSEA delle scale *self-report* risulta superiore ai valori raccomandati (0.09 nel campione statunitense e 0.21 nel campione britannico).

Uno studio successivo (DeSousa et al., 2013) ha utilizzato una metodologia simile per valutare la struttura a un fattore dell'ARI nel contesto brasiliano, impiegando il metodo di stima WLSMV (Weighted Least Squares Mean and Variance adjusted) e utilizzando il Chi-square (χ^2 : $p > 0.05$) e gli stessi indici di Stringaris et al., 2012. I risultati della CFA hanno mostrato un buon adattamento del modello, con χ^2 (gdl)=9.753 (9); $p = .371$, CFI = .999, TLI = .998, RMSEA = .025 e WRMR = .398.

Mulraney et al. (2014) nel loro articolo sulla validazione della versione australiana dell'ARI ("*Psychometric Properties of the Affective Reactivity Index in Australian Adults and Adolescents*") hanno confermato la struttura a un fattore nella versione *parent-report* dell'ARI (CFI = 1.00, RMSEA = .00); tuttavia, la struttura non si è rivelata adeguata alla versione *self-report* adolescenziale (CFI = .95, RMSEA = .12) e *self-report* per adulti (CFI = .91, RMSEA = .13). È stata ipotizzata una ridondanza nel modo in cui gli Items 2 e 6 venivano interpretati nelle versioni *self-report* (tra i due, infatti, risultava un'elevata covarianza). Sulla base di ciò, Mulraney et al. dimostrarono che rimuovendo l'Item 6 alla versione *self-report* adolescenziale, il grado di adattamento della scala alla struttura a un fattore migliorava (CFI = .98, RMSEA = .08). Allo stesso modo, la rimozione dell'Item 2 forniva il miglior fit per la versione *self-report* degli adulti (CFI = 1.00, RMSEA = .03).

Ezpeleta et al. (2020) è un altro studio già menzionato nel secondo capitolo e che indaga la struttura fattoriale dell'ARI *teacher-report* con il metodo WLSMV per

esaminare un modello a 6 items unidimensionale con bambini di 7 e 11 anni. I risultati hanno mostrato un adattamento soddisfacente a 7 anni, ma insufficiente a 11 anni (RMSEA = .121). L'adattamento è stato considerato accettabile dopo aver correlato l'unicità degli Items 2 e 6 che – come nello studio di Mulraney et al. (2014) – venivano interpretati in maniera molto simile dai partecipanti.

Anche Evans et al. (2021) hanno utilizzato il metodo WLSMV come metodo di stima per implementare le CFA nel loro studio di validazione dell'ARI *self-* e *parent-report* su un campione clinico. I risultati hanno evidenziato degli indici di fit CFI e TLI superiori a .95, indicando un buon adattamento della scala; tuttavia, l'indice RMSEA per la versione *parent-report* era discutibile. La successiva analisi di un modello *multi-informant* ha confermato che il modello a un fattore si adattava bene (CFI/TLI > .98, RMSEA < .07).

Dougherty et al. (2021) hanno utilizzato un'analisi fattoriale esplorativa (EFA) seguita da una CFA per testare l'unidimensionalità dell'ARI. I risultati delle EFA e delle CFA per la versione *parent-report* (compilata dalle madri) e *self-report* hanno supportato l'unidimensionalità con indici di fit soddisfacenti (CFI > .95, TLI > .95, RMSEA < .10).

Anche nello studio di Sugaya et al. (2022) su un campione brasiliano di bambini di età prescolare, gli indici di fit sono stati impiegati in modo simile a quanto fatto da Stringaris et al. I risultati hanno mostrato che una struttura unidimensionale a un fattore descriveva adeguatamente l'ARI *parent-report* in entrambi i campioni analizzati: campione scolastico, CFI = .99, TLI = .98, RMSEA = .07 (90% CI 0.05–0.10) e campione clinico, CFI = 1.00, TLI = .99, RMSEA = .060 (90% CI .00–.12).

Wilson et al. (2022) hanno impiegato una CFA per valutare l'ARI-P (*Parent-Report*) in un campione clinico di bambini di età compresa tra 3 e 8 anni tramite il metodo di stima WLSMV. Gli indici di fit includevano il Chi-quadrato (χ^2), il CFI, l'RMSEA e il WRMR. I risultati hanno mostrato che la soluzione unidimensionale dell'ARI-P forniva un fit nel complesso accettabile, con due dei tre indici considerati buoni (CFI = .993 e WRMR = .781).

Mallidi et al. (2023) hanno riscontrato nel loro studio una discrepanza di accordo tra informatori (*self-* e *parent-report*) e a causa di ciò hanno esplorato un approccio bifattoriale con, da un lato, un fattore di irriducibilità condiviso tra gli informatori e, dall'altro lato, un fattore specifico che catturasse la varianza nel singolo informatore. Il

modello bifattoriale è stato scelto proprio perché consente di separare la varianza specifica dell'informatore dalla varianza condivisa. Il modello bifattoriale è stato testato utilizzando la CFA sulle versioni *self-* e *parent-report* dell'Affective Reactivity Index: *ARI-youth* e *ARI-parent*. Gli indici di fit impiegati includevano RMSEA, CFI e lo Standardized Root Mean Square Residual (SRMR). I risultati hanno mostrato un eccellente fit sia per il campione del National Institute of Mental Health – NIMH (CFI = .99; RMSEA = .05; SRMR = .02) che per il campione del Child Mind Institute – CMI (CFI = .99; RMSEA = .04; SRMR = .02).

Lo studio più recente (Lu et al., 2024) ha esaminato la struttura fattoriale unidimensionale dell'ARI-S (*Self-Report*) e ARI-P (*Parent-Report*) in un campione di bambini e adolescenti cinesi, tra i 9 e 17 anni. Gli indici di fit evidenziano un buon adattamento del modello sia per la ARI-P (CFI = .990, TLI = .984, RMSEA = .074 con 90%CI=[.045, .105], WRMR=0.753) che per la ARI-S (CFI = .995, TLI = .992, RMSEA = .056 con 90% CI = [.022, .089], WRMR = 0.632).

3.2 Affidabilità

Come già riscontrato nel secondo capitolo, gli autori dell'ARI (Stringaris et al., 2012) hanno evidenziato nel loro studio un'elevata consistenza interna (con un alpha di Cronbach di $\alpha = 0.88$ per la versione *self-report* e $\alpha = 0.92$ per la *parent-report* nel campione statunitense, e valori di $\alpha = 0.90$ e $\alpha = 0.89$ nel campione britannico. I punteggi totali erano superiori nel campione statunitense. È stata riscontrata anche una buona correlazione tra i punteggi di genitori e figli: $r = 0.58$ (CI 0.47–0.66) nel campione US, $r = 0.73$ (CI 0.56–0.85) nel campione UK. Ciò indica che l'ARI misura efficacemente l'irritabilità da una prospettiva *multi-informant*. Il coefficiente di correlazione per la stabilità longitudinale è risultato elevato e significativo per l'ARI *parent-report* ($r = 0.88$, $p < 0.001$), ma non per la versione *self-report* ($r = 0.29$, $p = 0.28$).

DeSousa et al. (2013) hanno confermato la buona consistenza interna dell'ARI nel campione brasiliano, con un coefficiente di Cronbach di $\alpha = 0.843$.

Mulraney et al. (2014), nel loro studio sulle proprietà psicometriche dell'ARI in un campione australiano, hanno riportato a loro volta una consistenza interna soddisfacente (coefficiente alpha .80 nel *self-report adulto*, .85 nel *self-report*

adolescenziale e .80 nel *parent-report*). L'affidabilità test-retest è stata indagata tramite il coefficiente di concordanza intraclassa (ICC) in un sotto-campione di adulti nell'intervallo di tempo di una settimana. I risultati erano buoni, con un ICC=.80 (95% CI [.59,.90]).

Lo studio di Tseng et al. (2017) ha indagato l'affidabilità test-retest dell'ARI su un periodo di 2–4 settimane, utilizzando il coefficiente di correlazione di Pearson (r) e il coefficiente di concordanza intraclassa (ICC). I valori di ICC sono stati interpretati secondo la classificazione di Cicchetti e Sparrow (1981) come segue: $< .40$ = poor; $.40$ – $.59$ = fair; $.60$ – $.74$ = good; $> .74$ = excellent. I risultati hanno mostrato un'ottima affidabilità test-retest, sia per l'ARI *self-report* ($r=.79$, ICC=.88) che per il *parent-report* ($r=.82$, ICC=.90), con valori significativi ($p < .001$). Tuttavia, l'accordo tra ARI tra le versioni *self-* e *parent-report* era da basso a moderato. ($r = .37$ e $.22$, $p < .05$ – come riportato nel lavoro).

Lo studio di Lee et al., 2017 ha rilevato nuovamente una buona consistenza interna nell'ARI *self-report* (alpha di Cronbach $\alpha = 0.820$).

Pan et al. (2019) hanno indagato la consistenza interna e l'affidabilità test-retest (valutata nell'intervallo di due settimane) nella versione cinese dell'ARI sia per un campione di controllo che per un campione clinico, all'interno del quale erano presenti diversi bambini affetti da disturbo dello spettro autistico (ASD, $n=21$). In generale, la elevata consistenza interna è stata nuovamente confermata (alpha di Cronbach $\alpha = .91$ per il *parent-report* e $\alpha = .86$ per il *self-report*). L'affidabilità test-retest – valutata in un intervallo di due settimane – era più elevata nell'ARI *parent-report* (ICC = $.82$, 95% CI $.71$ – $.88$) rispetto alla versione *self-report* (ICC = $.66$, 95% CI $.50$ – $.78$). Tuttavia, nel campione clinico affetto da ASD, l'ARI *self-report* aveva un'alta consistenza interna ($\alpha = .87$) ma un'affidabilità test-retest modesta (ICC = $.51$, 95% CI $.06$ – $.79$).

Anche lo studio di Ezpeleta et al. (2020) – già citato nel capitolo II – ha evidenziato una consistenza interna soddisfacente nell'ARI *teacher-report* ($\alpha = .85$ per i bambini di 7 anni e $\alpha = .88$ per i bambini di 11 anni. Inoltre, era presente una correlazione positiva tra il punteggio grezzo di entrambe le fasce d'età e l'Item 7 di Impairment. L'accordo tra informatori è stato ritenuto moderato.

Evans et al. (2021) hanno valutato la consistenza interna attraverso i coefficienti alpha di Cronbach e omega di McDonald's. La consistenza interna si è rivelata maggiore

nel *parent-report* ($\alpha = .88$ e $\omega = .90$) rispetto al *self-report* ($\alpha = .79$ e $\omega = .80$) con, in entrambi i coefficienti un CI [95% CI]. Inoltre, la correlazione tra i punteggi delle due versioni *self-* e *parent-report* non è risultata molto forte ($r = .32$), suggerendo una possibile discrepanza nel modo in cui genitori e figli interpretano la scala.

Lo studio di Kalvin et al. (2021) ha analizzato le proprietà psicometriche dell'ARI *parent-report* in un campione clinico di bambini con spettro autistico (ASD). I risultati hanno evidenziato una consistenza interna eccellente ($\alpha = .90$).

Wilson et al. (2022) hanno valutato la consistenza interna attraverso il calcolo del coefficiente di Cronbach nei primi 6 Items dell'ARI-P (*parent-report*) in un campione clinico. I risultati hanno mostrato un coefficiente $\alpha = .87$, indicando una buona consistenza interna. Inoltre, non sono emerse evidenze di ridondanza tra gli items (nessun coefficiente $\alpha > .90$), suggerendo che non vi è necessità di accorciare la scala.

Sugaya et al. (2022) hanno valutato la consistenza interna della versione brasiliana dell'ARI *parent-report* in un campione clinico (suddiviso in 3 gruppi sottoposti a diversi programmi di intervento) e un campione scolastico (quest'ultimo suddiviso a sua volta in 3 gruppi, due sottoposti a programmi di intervento e uno di controllo) utilizzando il coefficiente di Cronbach e il Composite Reliability Index (CRI). I risultati hanno mostrato una buona consistenza interna: nel campione scolastico, il coefficiente α di Cronbach era $\alpha = .78$ e il CRI era $.88$; nel campione clinico, il coefficiente di Cronbach era $\alpha = .86$ e il CRI era $.94$. L'affidabilità test-retest è stata valutata attraverso l'ICC in un intervallo di tempo di 15 settimane, ma solo nel gruppo di controllo del campione scolastico (poiché era l'unico all'interno dello studio a non essere sottoposto ad alcun programma di intervento). Il livello di test-retest reliability è risultato moderato: ICC = $.55$ (95% CI $.42 - .66$).

Lo studio di Lee et al. (2022) non si occupava di analizzare nel dettaglio le proprietà psicometriche della scala, bensì l'associazione con diversi fattori natali e perinatali in un campione del Taiwan; nonostante ciò, gli autori hanno comunque analizzato e valutato come alta la consistenza interna dell'ARI *parent-report* ($\alpha = .89$).

La consistenza interna dell'ARI *self-* e *parent-report* è stata valutata anche da Mallidi et al. (2023), sempre tramite l'alpha di Cronbach: l'ARI *parent-report* ha mostrato un'eccellente consistenza interna ($\alpha = .92$ nel campione NIMH e $\alpha = .90$ nel campione CMI); l'ARI *self-report* ha mostrato una buona consistenza interna ($\alpha = .88$ nel campione

NIMH e $\alpha = .82$ nel campione CMI). Per quanto riguarda l'affidabilità test-retest, 177 coppie genitore-figlio nel campione NIMH hanno completato l'ARI due volte, con un intervallo medio di 45,3 giorni tra le due valutazioni. È stato utilizzato il coefficiente ICC per valutare l'affidabilità test-retest, con risultati che indicano una buona affidabilità sia nel *parent-report* (ICC = .85) che nel *self-report* (ICC = .78).

Lo studio più recente (Lu et al., 2024) ha valutato l'ARI-S (*self-report*) e ARI-P (*parent-report*) sia in termini di consistenza interna (utilizzando l'alpha di Cronbach e l'omega di McDonald's) che di affidabilità test-retest (utilizzando l'ICC).

I risultati hanno riportato dei valori buoni di consistenza interna: nell'ARI-P un coefficiente omega di $\omega = .83$ (95% CI = [.795, .853]) e nell'ARI-S un coefficiente di omega $\omega = .82$ (95% CI = [.790, .850]). I valori dell'alpha di Cronbach per l'ARI-P e l'ARI-S sono stati entrambi di 0.81, confermando la buona consistenza interna.

Per quanto riguarda l'affidabilità test-retest, è stata misurata su un intervallo di quattro settimane utilizzando il coefficiente ICC con un modello casuale a due vie. I risultati hanno mostrato una buona affidabilità test-retest, con un ICC di .77 (95% CI = [.666 – .841]) per l'ARI-P e di 0.81 (95% CI = [.725 – .872]) per l'ARI-S.

3.3 Validità

Validità di costrutto e validità di criterio – Per quanto riguarda la validità dell'ARI, sono stati molti gli studi che hanno indagato la validità di costrutto (in particolare la validità convergente) e la validità di criterio. Di seguito i risultati, in ordine, dal meno recente al più recente.

Gli autori dell'ARI (Stringaris et al., 2012), come già evidenziato nel primo capitolo, hanno rilevato una correlazione positiva dell'ARI con tutte le sottoscale dello *Strengths and Difficulties Questionnaire* (SDQ – Goodman, 2001), ad eccezione della sottoscala dei problemi con i pari nella versione *self-report*. Inoltre, i modelli di regressione hanno evidenziato che – nei punteggi dell'ARI e dello SDQ *parent-report* – l'irritabilità emerge come unica variabile predittiva della sottoscala delle difficoltà emotive, ma non dei problemi di condotta.

DeSousa et al. (2013) hanno studiato la validità di costrutto tramite il coefficiente di Pearson per analizzare l'associazione tra l'ARI e la *Spence Children's Anxiety Scale*

(SCAS), una misurazione *self-report* dei sintomi ansiosi in bambini e adolescenti. L'ipotesi era che la correlazione fosse più significativa con la sottoscala SCAS-GAD (ansia generalizzata e ansia da distress), in accordo con le evidenze della letteratura, che collegano alti livelli di irritabilità in infanzia allo sviluppo successivo di disturbi quali il disturbo d'ansia generalizzata, la depressione e la distimia. I risultati hanno mostrato correlazioni da piccole a moderate tra i punteggi ARI e tutte le sottoscale della SCAS. I successivi Z test hanno evidenziato che la correlazione con la SCAS-GAD (ansia generalizzata) è risultata significativamente più forte (con un punteggio di $p = 0.56$) rispetto alle altre sottoscale. Questo risultato conferma l'ipotesi degli autori.

Mulraney et al. (2014), nel loro studio intitolato "*Psychometric Properties of the Affective Reactivity Index in Australian Adults and Adolescents*", hanno analizzato la validità convergente sia nell'ARI *self-report adolescenziale* che nel *self-report adulto*.

Nel *self-report adolescenziale*, l'obiettivo era indagare la validità convergente e predittiva dell'ARI con alcune sottoscale dello SDQ (difficoltà emotive, problemi di condotta e iperattività/disattenzione) e con il *Reynolds Adolescent Depression Scale* (RADS-2 – Reynolds, 2002). Le analisi sono state condotte utilizzando il coefficiente di Pearson, correlazioni crossinformant e diversi modelli di regressione in cui le sottoscale dello SDQ sopracitate venivano utilizzate una a una come outcome e il punteggio dell'ARI diventava predittore, insieme a tutte le altre sottoscale dello SDQ.

I risultati hanno evidenziato relazioni da forti a moderate per i punteggi dell'ARI *self-* e *parent-report* e delle sottoscale SDQ, ad eccezione della sottoscala del rapporto con i pari (nel *parent-report*) e della sottoscala del comportamento prosociale (nel *self-report*). Le correlazioni crossinformant sono risultate moderate tra i punteggi dell'ARI *parent-report* e quelli dello SDQ *self-report*; tuttavia, solo le sottoscale delle problematiche emotive e dell'iperattività/disattenzione erano correlate all'ARI *self-report*. Infine, dai modelli di regressione analizzati, l'ARI *parent-report* risultava in grado di predire tre sottoscale dello SDQ: difficoltà emotive, problemi di condotta e iperattività. L'ARI *self-report* prediceva, a sua volta, le difficoltà emotive e i problemi di condotta, ma non l'iperattività.

Inoltre, i risultati indicano una correlazione positiva e significativa tra ARI *self-report* e la RADS-2 ($r=.49$ con $p < 0.01$).

Nel *self-report* adulto, l'obiettivo era indagare la validità convergente dell'ARI con la CES-D (Radloff, 1977), la GAD-7 (Spitzer, Kroenke, Williams, & Lowe, 2006) e la *Liebowitz Social Anxiety Scale* (LSAS – Liebowitz, 1987); tre scale che misurano, rispettivamente, la depressione, l'ansia generalizzata e l'ansia sociale. L'associazione è stata misurata sia con il coefficiente di correlazione di Pearson, che con dei modelli di regressione in cui ognuna delle tre misure rappresentava l'outcome e il punteggio ARI totale era il predittore. I risultati evidenziano correlazioni moderate tra l'ARI e queste tre misure. Nella CES-D: $r = .42$ ($p < .01$); nella GAD-7: $r = .44$ ($p < .01$); e nella LSAS: $r = .31$ ($p < .01$) nella LSAS.

Mulraney et al. hanno svolto nello stesso anno (2014) un altro studio intitolato “*Brief report: Can irritability act as a marker of psychopathology?*” in cui è stata testata la validità di criterio tra l'ARI *self-report* e le sottoscale dello SDQ con un'analisi di regressione. I risultati hanno mostrato che l'ARI fosse un forte predittore per tre delle cinque sottoscale dello SDQ:

- Problemi di condotta, con $r = .54$ ($p < .01$) e $\beta = .54$ ($p < .05$);
- Difficoltà emotive, con $r = .39$ ($p < .01$) e $\beta = 0.39$ ($p < .05$);
- Iperattività, con $r = 0.50$ ($p < .01$) e $\beta = .50$ ($p < .05$);

Nello studio di Tseng et al., (2017) – sebbene il focus principale fosse la validazione dell'*Affective Posner 2 Task* – vengono evidenziate correlazioni significative tra l'ARI *self-report* e le valutazioni di affetto negativo durante le condizioni di frustrazione del compito. L'ARI è stato somministrato ai partecipanti due volte, in due visite. I risultati delle correlazioni tra l'ARI e le valutazioni di infelicità ($r = .32$ e $.33$, $p < .001$) e frustrazione ($r = .23$ e $.26$, $p < .05$) suggeriscono che l'ARI *self-report* possa essere un indicatore valido delle emozioni negative evocate in situazioni di stress.

Nello studio di Pan et al. (2019), la validità concorrente è stata testata con il coefficiente di Pearson. La versione cinese dell'ARI *parent-report* ha mostrato una correlazione significativa con diverse sottoscale della *Child Behaviour Checklist* (CBCL – Achenbach & Dumenci, 2001; Achenbach & Rescorla, 2000): la correlazione era buona per la scala del comportamento aggressivo ($r = .667$, $p < .001$) e significativa per la scala di ansia/depressione ($r = .434$, $p < .001$) e per la scala dei problemi sociali ($r = .453$, $p < .001$). Le correlazioni dell'ARI *self-report* erano basse con tutte le sottoscale della CBCL.

Ezpeleta et al. (2020), come già detto nel primo capitolo, hanno rilevato convergenze particolarmente elevate tra l'ARI *teacher-report* con altri strumenti compilati dagli insegnanti: due sottoscale dello SDQ (Difficoltà emotive e Problemi di condotta); i punteggi dell'aggressività totale, dell'aggressività verso i pari e il punteggio totale della *Children's Aggression Scale* (CAS); e, infine, il livello di rabbia rilevato tramite la *PROMIS—Anger—Parent/Guardian of Child Age 6–17* (PROMIS, 2016).

Dal confronto dell'ARI *teacher-report* con diverse misurazioni *self-report* (YSR e AQ) e *parent-report* (CBCL ed ERC), non sono emerse convergenze significative, tranne una discreta associazione con l'ERC. L'associazione dei punteggi dell'ARI *teacher-report* con altri strumenti *teacher-report* era maggiore, risultando tanto più forte quanto più il costrutto misurato si avvicinava all'irritabilità. Complessivamente, il grado di accordo tra informatori era da considerarsi moderato.

Evans et al. (2021) hanno valutato la validità convergente in relazione a diversi strumenti di valutazione. Innanzitutto, l'analisi delle correlazioni zero-order ha mostrato che l'ARI *parent-report* presenta correlazioni positive significative con la *Conners 3rd Edition, Parent Short Form* (Conners, 2008), utilizzata per valutare l'associazione tra l'ARI e i comportamenti esternalizzanti ($r = .26-.66, p < .01$). L'analisi della regressione multipla ha confermato come l'ARI *parent-report* sia predittivo dell'aggressione reattiva ($\beta = .55$), misurata attraverso la scala *Proactive/Reactive Aggression Rating Scale* (PRA; Dodge & Coie, 1987). Inoltre, sono presenti correlazioni moderate con la *Revised Children's Manifest Anxiety Scale 2nd Edition, Short Form* (RCMAS; Reynolds & Richmond, 2008) ($r = .23-.42, p < .01$) e il *Children's Depression Inventory 2nd Edition: Self-Report, Short Form* (CDI; Kovacs, 2010) ($r = .23-.42, p < .01$), suggerendo una validità convergente moderata in relazione ad ansia e depressione.

Lo studio di Sugaya et al. (2022) ha analizzato la validità di criterio dell'ARI *parent-report* con diverse scale:

- *Children's Behavior Checklist* (CBQ), con un focus sulla CBQ rabbia/frustrazione
- *Swanson, Nolan, and Pelham-IV* (SNAP-IV), una scala di valutazione dell'ADHD e dell'ODD (disturbo oppositivo provocatorio). Sono stati, inoltre, utilizzati 3 degli items dedicati all'ODD per costruire una scala SNAP-irritability;

- *Child behavior checklist 1.5–5* (CBCL), in particolare le scale dell'ansia/depressione, dei problemi attentivi e la CBCL-irritability (una scala a 3 items sviluppata ad hoc per misurare l'irritabilità).
- *Multidimensional assessment profile of disruptive behavior* (MAP-DB), un questionario multidimensionale che analizza quattro dimensioni del comportamento dirompente (i.e. temper loss, aggression, noncompliance and low concern to others).

L'ARI ha mostrato correlazioni moderate con lo SNAP-irritability e la CBCL-irritability, con coefficienti di correlazione che vanno da .55 a .68. Sono emerse, inoltre, correlazioni moderate con costrutti legati all'irritabilità, come ansia/depressione, aggressività e comportamento oppositivo.

Wilson et al. (2022), nel loro campione clinico di bambini tra i 3 e gli 8 anni, hanno analizzato la validità concorrente dell'ARI-P (*parent-report*) con diverse scale sindromiche della CBCL: la scala dei problemi d'ansia, la scala dell'ansia/depressione, la scala del comportamento aggressivo e quella dei problemi di attenzione. I risultati riscontrano correlazioni parziali significative ma moderate (controllate per età) tra il punteggio totale dell'ARI-P e le varie sottoscale della CBCL: problemi d'ansia ($r=.24, p < .05$); ansia/depressione ($r=.27, p < .01$); comportamento aggressivo ($r= .63, p < .001$) e problemi di attenzione ($r=.24, p < .05$). Gli autori, inoltre, hanno valutato la validità convergente dell'ARI-P (*parent-report*) con una scala a 3 items che misura l'irritabilità, tratta dalla CBCL. I risultati evidenziano una correlazione parziale forte ($\geq .5$) tra l'ARI-P e la CBCL-irritability ($r = .55, p < .001$).

Lo studio più recente di Lu et al. (2024) ha analizzato la validità concorrente dell'ARI (ARI-S e ARI-P) con le varie sottoscale dello SDQ; il metodo impiegato è stato quello delle correlazioni parziali, controllando le variabili del genere e dell'età. I risultati hanno riscontrato, in generale, delle correlazioni da deboli a moderate. Per quanto riguarda la validità convergente, per l'ARI-P (*parent-report*) è stata riscontrata una correlazione alta con la CBCL-irritability ($r = .73, p < .02$, così riportato nel lavoro). Per l'ARI-S (*self-report*) è stata valutata la correlazione con la sottoscala dell'irritabilità dello Youth Self-Report (YSR) che è risultata essere moderata ($r = .69, p < .01$).

Validità di contenuto – Nella presente revisione sistematica, non è emerso alcun riferimento alla valutazione della validità di contenuto nello studio degli autori originali

(Stringaris et al., 2012); tuttavia, l'adattamento dell'ARI nella versione brasiliana della scala (DeSousa et al., 2013) è stato valutato da un gruppo di esperti, i quali hanno concordato sulla rilevanza e adeguatezza dei fondamenti teorici dell'ARI originale rispetto al contesto brasiliano. L'ARI – nelle sue versioni self- e parent-report – è stato tradotto due volte da due traduttori indipendenti, mentre un terzo traduttore ha creato la traduzione definitiva sintetizzando le prime due. Gli esperti hanno espresso un giudizio positivo sulla pertinenza che avevano gli items tradotti nel rispecchiare il costrutto di irritabilità, ad eccezione dell'item 1 (“*Easily annoyed by others*”), in cui “annoyed” era stato tradotto come “perturbado” e, quindi, interpretato dagli esperti come “preoccupato” al posto di “infastidito”; ciò ha convinto gli esperti a ritenere “incomodado” come un termine più appropriato. Le istruzioni per la compilazione sono state giudicate come appropriate.

Validità esterna – Gli studi presenti – per quanto il loro numero sia ancora limitato – rilevano come l'Affective Reactivity Index (ARI – Stringaris et al., 2012) dimostri di possedere delle buone proprietà psicometriche in diversi contesti culturali: Stati Uniti, Regno Unito, Spagna, Brasile, Cina, Australia.

Inoltre, diversi studi hanno dimostrato nell'ARI una capacità nel discriminare la psicopatologia: Stringaris et al. (2012) hanno riscontrato punteggi di irritabilità più elevati nel loro campione clinico rispetto al gruppo di controllo (in particolare, per i bambini con disregolazione severa dell'umore, subito seguiti dai bambini con disturbo bipolare). Sugaya et al. (2022) hanno riscontrato punteggi elevati nei bambini con ADHD in comorbidità con un disturbo oppositivo provocatorio, rispetto all'ADHD con altre comorbidità o senza comorbidità. Lo studio di Kalvin et al. (2021) ha riscontrato punteggi elevati di irritabilità nei bambini con lo spettro autistico (ASD), specie quando questo era in comorbidità con un disturbo d'ansia o un disturbo dell'umore dirompente. Lo studio di Pan (2019) riscontrava punteggi elevati nel campione clinico di bambini con disturbo dell'umore disregolato e dirompente (DMDD). Altri studi già citati in precedenza (Evans et al., 2021; Pan et al., 2019; Sugaya et al., 2022) hanno riscontrato un'associazione moderata con diverse scale di misurazione di comportamenti internalizzanti (ansia e depressione) o esternalizzanti (comportamento dirompente, aggressività).

Le psicopatologie considerate da questi studi (ADHD, ODD, ASD, DMDD, ansia, disturbo bipolare e depressione unipolare) sono caratterizzate da una sintomatologia che

include al suo interno livelli elevati di irritabilità. Il fatto che bambini e adolescenti con questi disturbi ottengano punteggi più elevati nell'ARI rispetto a un gruppo di controllo fornisce un'ulteriore evidenza della capacità dell'ARI di misurare l'irritabilità in modo dimensionale.

CAPITOLO IV

DISCUSSIONE E CONCLUSIONI

Il primo capitolo di questo elaborato è stato dedicato a una prospettiva teorica su come il costrutto di irritabilità si è evoluto nel corso degli anni, suscitando un dibattito nosologico circa la posizione dell'irritabilità all'interno della psicopatologia dell'età evolutiva.

Stringaris (2011) ha evidenziato le proprietà dimensionali dell'irritabilità, la quale si colloca nella sintomatologia di diverse psicopatologie. *L'Affective Reactivity Index* (ARI – Stringaris et al., 2012) è stato creato sulla base di questo approccio dimensionale.

In generale, la struttura fattoriale e l'affidabilità risultano soddisfacenti in entrambe le versioni della scala, ma con alcuni limiti nella versione *self-report*.

Dal punto di vista della struttura fattoriale, diversi autori hanno studiato la soluzione a un fattore della scala attraverso analisi fattoriali confermate (CFA). I risultati mostrano che la struttura a un fattore sembra essere adeguata alla versione *parent-report* dell'ARI, ma dimostra alcune criticità nella versione *self-report* (Stringaris et al., 2012; Mulraney et al., 2014, “*Psychometric Properties of the Affective Reactivity Index in Australian Adults and Adolescents*”) e *teacher-report* (Ezpeleta et al., 2020). Lo studio di Mulraney et al. ha considerato la covarianza tra gli Items 2 e 6 dell'ARI (“*often loses their temper*”) e (“*loses their temper easily*”) e ha ipotizzato che questi due items venissero interpretati in modo ridondante dai soggetti che compilavano la scala. L'ipotesi è stata confermata, poiché gli indici di fit miglioravano omettendo l'Item 6 dall'ARI *self-report* adolescenziale e l'Item 2 dall'ARI *self-report* adulto. Anche nello studio di Ezpeleta et al. (2020), gli indici di fit nel campione di 11 anni sono migliorati dopo aver correlato l'unicità dell'interpretazione degli Items 2 e 6.

L'evidenza di alcuni studi sulle discrepanze di accordo nei punteggi tra genitori e figli (in questa revisione, per esempio: Mulraney et al., 2014; Tseng et al., 2017), ha portato Mallidi et al. (2023) a considerare per la prima volta un modello bifattoriale, composto da un fattore di irritabilità condiviso tra gli informatori e un fattore specifico che catturasse la varianza nel singolo informatore. I risultati hanno mostrato un eccellente fit sia per entrambi i campioni clinici considerati (NIMH e CMI). Nonostante questo risultato eccellente, il contributo di Mallidi et al. è molto recente e – dalla revisione

condotta – non emergono altri contributi che abbiano replicato l’efficacia dell’approccio bifattoriale.

I risultati per quanto riguarda l’affidabilità dell’ARI, evidenziano una consistenza interna molto alta in tutti gli studi, anche se in alcuni studi era maggiore nella versione *parent-report* rispetto alla *self-report* (Evans et al., 2021; Mallidi et al., 2023).

Anche l’affidabilità test-retest dell’ARI è risultata in generale elevata, anche se nello studio di Pan et al. (2019) è risultata modesta per quanto riguarda un campione di 21 bambini con disturbo dello spettro autistico (ASD).

Nella valutazione della validità convergente, sono emerse alcune criticità date dai risultati contrastanti per quanto riguarda l’associazione tra l’ARI e lo *Strengths and Difficulties Questionnaire* (SDQ – Goodman, 2001). Stringaris et al., 2012 hanno rilevato una correlazione positiva dell’ARI con tutte le sottoscale dello SDQ, ad eccezione della sottoscala dei problemi con i pari nella versione *self-report*. Mulraney et al. (2014, “*Psychometric Properties of the Affective Reactivity Index in Australian Adults and Adolescents*”), al contrario, hanno ottenuto una correlazione positiva tra l’ARI e tutte le sottoscale dello SDQ, ad eccezione della sottoscala dei problemi con i pari nel *parent-report*.

Allo stesso modo, i risultati di Stringaris et al. (2012) sostengono che l’irritabilità è predittiva solo dei problemi internalizzanti (sottoscala delle difficoltà emotive) ma non dei problemi esternalizzanti (sottoscala dei problemi di condotta). Un altro studio di Mulraney et al. (2014) intitolato “*Brief report: Can irritability act as a marker of psychopathology?*”, ha invece riscontrato il valore predittivo dell’irritabilità anche nella sottoscala dei problemi di condotta e ha, dunque, riscontrato un valore predittivo dell’ARI sia a livello di comportamenti internalizzanti che a livello di comportamenti esternalizzanti.

Altri studi hanno evidenziato una correlazione da moderata ad alta tra l’ARI e la CBCL-irritability (Sugaya et al., 2022; Wilson et al., 2022; Lu et al., 2024), dando un ulteriore contributo positivo alla validità di criterio dell’ARI. La validità di criterio dell’ARI viene rispecchiata anche da correlazioni moderate con altri strumenti di misurazione di comportamenti internalizzanti ed esternalizzanti: depressione, ansia, aggressività, rabbia, frustrazione e comportamento dirompente.

La validità di contenuto è stata valutata da un solo studio (DeSousa et al., 2013) che ha adattato la scala originale al contesto brasiliano. Stringaris et al. (2012), infatti, non riportano alcun riferimento alla valutazione della scala da parte di un gruppo di esperti. La mancanza di valutazione della validità di contenuto nello studio originale potrebbe spiegare il motivo per cui in alcuni studi successivi (Mulraney et al., 2014 “*Psychometric Properties of the Affective Reactivity Index in Australian Adults and Adolescents*”; Ezpeleta et al., 2020) è stata riscontrata una ridondanza nel modo in cui i partecipanti interpretavano gli Items 2 e 6 dell’ARI *self- e teacher-report*.

4.1 Prospettive future

Un punto di forza emerso dalla revisione sistematica è che gli studi sull’*Affective Reactivity Index*, per quanto limitati, hanno dimostrato delle proprietà psicometriche soddisfacenti in diversi contesti culturali (Stati Uniti, Regno Unito, Spagna, Brasile, Cina, Australia) e un’efficacia nel discriminare i gruppi di controllo da diversi gruppi clinici che presentavano varie psicopatologie (ASD, ADHD, DMDD, ODD, ansia, disturbo bipolare, depressione unipolare). Ulteriori studi potrebbero dare un contributo alla validazione dell’ARI sia per quanto riguarda la replicabilità dei risultati già presenti, sia per quanto concerne l’esplorazione di nuovi contesti clinici o culturali.

I risultati delle analisi sulla struttura fattoriale dell’ARI sollevano delle criticità; sia sull’efficacia del modello a un fattore (Mallidi et al., 2023) che sul possibile ruolo della ridondanza nell’interpretazione degli Items 2 e 6 (Mulraney et al., 2014 “*Psychometric Properties of the Affective Reactivity Index in Australian Adults and Adolescents*”; Ezpeleta et al., 2020). Alla luce di questo, è necessario testare in studi futuri la replicabilità del modello bifattoriale proposto da Melvin et al. (2023); inoltre, sarebbe utile compiere uno studio sulla validità di contenuto della versione originale dell’ARI (Stringaris et al., 2012), per valutare la misura in cui ciascun item riflette il costrutto dell’irritabilità e correggere eventuali ridondanze o ambiguità nell’interpretazione dei vari items.

Infine, visti i risultati contrastanti, sarebbe utile valutare ulteriormente la validità convergente e predittiva dell’ARI con le varie sottoscale dello SDQ, così come con altri

strumenti di misurazione, allo scopo di approfondire il legame esistente tra irritabilità, comportamenti internalizzanti ed esternalizzanti.

BIBLIOGRAFIA

- American Psychological Association. (n.d.) Irritability. APA Dictionary of Psychology [Online]. Available from: <https://dictionary.apa.org/irritability>.
- Blader J.C., Carlson G.A. (2007). Increased rates of bipolar disorder diagnoses among U.S. child, adolescent, and adult inpatients, 1996-2004. *Biological Psychiatry*; 62:107-14, DOI: [10.1016/j.biopsych.2006.11.006](https://doi.org/10.1016/j.biopsych.2006.11.006)
- Brotman M.A., Kassem L., Reising M.M., Guyer A.E., Dickstein D.P., Rich B.A., Towbin K.E., Pine D.S., McMahon F.J., Leibenluft E. (2007). Parental diagnoses in youth with narrow phenotype bipolar disorder or severe mood dysregulation. *American Journal of Psychiatry*; 164:1238–1241, DOI: <https://doi.org/10.1176/appi.ajp.2007.06101619>
- Brotman M.A., Kirkanski K., Leibenluft E. (2017). Irritability in Children and Adolescents. *Annual Review of Clinical Psychology*, Vol. 13, 317-341, DOI: <https://doi.org/10.1146/annurev-clinpsy-032816-044941>
- Brotman M.A., Rich B.A., Guyer A.E., Lunsford J.R., Horsey S.E., Reising M.M., Thomas L.A., Fromm S.J., Towbin K., Pine D.S., Leibenluft E. (2010). Amygdala activation during emotion processing of neutral faces in children with severe mood dysregulation versus ADHD or bipolar disorder. *The American Journal of Psychiatry*, 167, 61-69, DOI: <https://doi.org/10.1176/appi.ajp.2009.09010043>
- Brotman M.A., Schmajuk M., Rich B.A., Dickstein D.P., Guyer A.E., Costello E.J., Egger H.L., Angold A., Pine D.S., Leibenluft E. (2006). Prevalence, Clinical Correlates, and Longitudinal Course of Severe Mood Dysregulation in Children. *Biological Psychiatry*, Vol. 60, Issue 9, 991-997, DOI: <https://doi.org/10.1016/j.biopsych.2006.08.042>
- De Sousa D.A., Stringaris A., Leibenluft E., Koller S.H., Manfro G.G., Salum G.A. (2013). Cross-cultural adaptation and preliminary psychometric properties of the Affective Reactivity Index in Brazilian youth: Implications for DSM-5 measured irritability. *Trends Psychiatry and Psychotherapy*, 35(3), DOI: <https://doi.org/10.1590/S2237-60892013000300004>

- Dougherty L.R., Galano M.M., Chad-Friedman E., Olinio T.M., Bufferd S.J., Klein D.N. (2021). Using Item Response Theory to Compare Irritability Measures in Early Adolescent and Childhood Samples. *Assessment* (3):918-927, DOI: <https://doi.org/10.1177/1073191120936363>
- Evans S.C., Abel M.R., Doyle R.L., Skov H., Harmon S.L. (2021). Measurement and correlates of irritability in clinically referred youth: Further examination of the Affective Reactivity Index. *Journal of Affective Disorders*, 283: 420-429, DOI: <https://doi.org/10.1016/j.jad.2020.11.002>
- Evans S.C., Shaughnessy S., Karlovich A.R. (2023). Future Directions in Youth Irritability Research. *Journal of Clinical Child and Adolescent Psychology*, 52:5, 716-734, DOI: <https://doi.org/10.1080/15374416.2023.2209180>
- Ezpeleta, L., Penelo, E., de la Osa, N., Navarro, J.B., Trepato, E. (2020). How the Affective Reactivity Index (ARI) works for teachers as informants. *Journal of Affective Disorders*, 261:40-48, DOI: <https://doi.org/10.1016/j.jad.2019.09.080>
- Kalvin C.B., Gladstone T.R., Jordan R., Rowley S., Marsh C.L., Ibrahim K., Sukhodolsky D.G. (2021). Assessing irritability in children with autism spectrum disorder using the Affective Reactivity Index. *Journal of Autism and Developmental Disorders*, 51(5):1496–1507. DOI: [10.1007/s10803-020-04627-9](https://doi.org/10.1007/s10803-020-04627-9)
- Lee K.S., Xiao J., Liew Z., Gau S.S.-F., Tseng W.-L. (2022). Perinatal and birth correlates of childhood irritability in Taiwan's national epidemiological study. *Journal of Affective Disorders*, 299:273-280, DOI: <https://doi.org/10.1016/j.jad.2021.12.016>
- Lee M., Aggen S.H., Carney D.M., Hahn S., Moroney E., Machlin L., Brotman M.A., Towbin K.E., Leibenluft E., Pine D.S., Roberson-Nay R., Hettema J.M. (2017). Latent structure of negative valence measures in childhood. *Depression and Anxiety* 34(8), 742–751, DOI: <https://doi.org/10.1002/da.22656>
- Leibenluft E. (2011). Severe mood dysregulation, irritability, and the diagnostic boundaries of bipolar disorder in youths. *The American Journal of Psychiatry*, 168:129–42, DOI: <https://doi.org/10.1176/appi.ajp.2010.10050766>
- Leibenluft, E., Charney, D.S., Towbin, K.E., Bhangoo, R.K., Pine, D.S. (2003). Defining clinical phenotypes of Juvenile Mania. *American Journal of Psychiatry*, 160(3):430–437. DOI: <https://doi.org/10.1176/appi.ajp.160.3.430>

- Lu B., Fang Y., Cai J., Chen Z. (2024). Psychometric Evaluation of the Affective Reactivity Index Among Children and Adolescents in China: A Multi-Method Assessment Approach. *Assessment*, (5):1020-1037, doi: <https://doi.org/10.1177/10731911231199424>
- Mallidi A., Meza-Cervera T., Kircanski K., Stringaris A., Brotman M.A., Pine D.S., Leibenluft E., Linke J.O. (2023). Robust caregiver-youth discrepancies in irritability ratings on the Affective Reactivity Index: An investigation of its origins. *Journal of Affective Disorders*, 332:185-193, DOI: <https://doi.org/10.1016/j.jad.2023.03.091>
- Mulraney M., Melvin G., Tonge B. (2014). Brief report: Can irritability act as a marker of psychopathology? *Journal of Adolescence*, 37(2014):419-423, DOI: <https://doi.org/10.1016/j.adolescence.2014.03.005>
- Mulraney M.A., Melvin G.A., Tonge B.J. (2014). Psychometric Properties of the Affective Reactivity Index in Australian Adults and Adolescents, 26(1):148-155, DOI: <https://doi.org/10.1037/a0034891>
- Pan P-Y, Yeh C-B. (2019). Irritability and Maladaptation Among Children: The Utility of Chinese Versions of the Affective Reactivity Index and Aberrant Behavior Checklist-Irritability Subscale. *Journal of Child & Adolescent Psychopharmacology*, (3):213-219. DOI: <https://doi.org/10.1089/cap.2018.0070>
- Rao U. (2014) DSM-5: disruptive mood dysregulation disorder. *Asian Journal of Psychiatry*, 11:119-23, DOI: <https://doi.org/10.1016/j.ajp.2014.03.002>
- Rich B.A., Carver F.W., Holroyd T., Rosen H.R., Mendoza J.K., Cornwell B.R., Fox N.A., Pine D.S., Coppola R., Leibenluft E. (2011). Different neural pathways to negative affect in youth with pediatric bipolar disorder and severe mood dysregulation, *Journal of Psychiatric Research*, 45(10): 1283-1294, DOI: <https://doi.org/10.1016/j.jpsychires.2011.04.006>
- Rich B.A., Schmajuk M., Perez-Edgar K.E., Fox N.A., Pine D.S., Leibenluft E. (2007). Different psychophysiological and behavioral responses elicited by frustration in pediatric bipolar disorder and severe mood dysregulation. *American Journal of Psychiatry*, 164(2):309-17. DOI: <https://doi.org/10.1176/ajp.2007.164.2.309>

- Stringaris A. (2011). Irritability in children and adolescents: a challenge for DSM-5. *Eur Child Adolesc Psychiatry* **20**, 61–66, DOI: <https://doi.org/10.1007/s00787-010-0150-4>
- Stringaris A., Baroni A., Haimm C., Brotman M., Lowe C.H., Myers F., Rustgi E., Wheeler W., Kayser R., Towbin K., Leibenluft E. (2010). Pediatric bipolar disorder versus severe mood dysregulation: risk for manic episodes on follow-up. *Journal of the American Academy of Child Adolescent Psychiatry*; 49:397–405, DOI: <https://doi.org/10.1097/00004583-201004000-00014>
- Stringaris, A., Goodman, R., Ferdinando, S., Razdan, V., Muhrer, E., Leibenluft, E., Brotman, M.A. (2012). The Affective Reactivity Index: a concise irritability scale for clinical and research settings. *Journal of Child Psychology and Psychiatry*, 53(11):1109-17, DOI: <https://doi.org/10.1111/j.1469-7610.2012.02561.x>
- Sugaya L.S., Kircanski K., Stringaris A., Polanczyk G.V., Leibenluft E. (2022). Validation of an irritability measure in preschoolers in school-based and clinical Brazilian samples. *European Child and Adolescent Psychiatry*, (4):577-587, DOI: <https://doi.org/10.1007/s00787-020-01701-6>
- Tseng W-L, Moroney E., Machlin L., Roberson-Nay R., Hetttema J.M., Carney D., Stoddard J., Towbin K.A., Pine D.S., Leibenluft E., Brotman M.A. (2017). Test-retest reliability and validity of a frustration paradigm and irritability measures. *Journal of Affective Disorders*, 212:38-45, DOI: <https://doi.org/10.1016/j.jad.2017.01.024>
- Wilson M.K., Cornacchio D., Brotman M.A., Comer J.S. (2022). Measuring Irritability in Early Childhood: A Psychometric Evaluation of the Affective Reactivity Index in a Clinical Sample of 3- to 8-Year-Old Children. *Assessment*, 29(7), 1473–1481, DOI: <https://doi.org/10.1177/10731911211020078>

APPENDICE A

Criteria di inclusione della sindrome da disregolazione severa dell'umore (SMD)

1. Età attuale 7-17 anni, con esordio della sindrome prima dei 12 anni;
2. Umore anomalo (in particolare rabbia o tristezza), presente per almeno metà giornata nella maggior parte dei giorni e di gravità sufficiente da essere notato dalle persone facenti parte l'ambiente del bambino (es: genitori, insegnanti, pari);
3. Iper arousal, definito da almeno tre dei seguenti sintomi: insonnia, agitazione, distraibilità, pensiero accelerato o fuga delle idee, eloquio pressante, intrusività;
4. Rispetto ai pari, il bambino mostra una reattività marcatamente aumentata agli stimoli emotivi negativi, la quale si manifesta modo verbale o comportamentale. Per esempio, il bambino risponde alla frustrazione con capricci prolungati (inadeguati per età e/o evento scatenante), ira verbale e/o aggressione nei confronti di persone o beni. Tali eventi si verificano, in media, almeno tre volte alla settimana;
5. I sintomi dei punti 2, 3 e 4 sono attualmente presenti e lo sono stati per almeno 12 mesi, senza alcun periodo privo di sintomi superiore a 2 mesi;
6. I sintomi risultano gravemente compromettenti in almeno un contesto di vita (casa, scuola, o con i pari) e sono almeno leggermente compromettenti in un secondo contesto.

Criteria di esclusione della sindrome da disregolazione severa dell'umore (SMD)

1. Mostra uno qualsiasi di questi sintomi maniacali cardinali:
 - Umore elevato o espansivo
 - Grandiosità o autostima gonfiata
 - Ridotto bisogno di sonno in modo episodico
2. I sintomi si presentano in periodi distinti che durano più di un giorno;

3. Soddisfa i criteri diagnostici per la schizofrenia, il disturbo schizoaffettivo, il disturbo pervasivo dello sviluppo o il disturbo da stress post-traumatico;
4. Soddisfa i criteri per il disturbo da uso di sostanze negli ultimi 3 mesi;
5. $QI > 70$
6. I sintomi sono dovuti agli effetti fisiologici diretti di una sostanza d'abuso o a una condizione medica o neurologica generale.

(tradotto da Leibenluft, 2011 – adattamento di Leibenluft et al., 2003)

APPENDICE B

Criteri di inclusione del disturbo da disregolazione dell'umore dirompente (DMDD)

1. Scoppi di collera severi e ricorrenti che si verificano almeno 3 volte alla settimana (sia rabbia verbale che comportamentale), sproporzionati nell'intensità o nella durata alla situazione e non coerenti con il livello di sviluppo dell'individuo;
2. Tra uno scoppio di collera e l'altro, l'umore è persistentemente irritabile o arrabbiato per la maggior parte della giornata, quasi tutti i giorni ed è osservabile anche dalle altre persone;
3. I sintomi devono essere stati presenti per almeno 12 mesi, con un periodo privo di sintomi non superiore a 3 mesi;
4. I sintomi sono presenti in almeno tre contesti di vita (es: casa, scuola, con i pari) e devono avere raggiunto un livello grave in almeno un contesto;
5. La diagnosi del disturbo non dovrebbe essere fatta per la prima volta prima dei 6 anni o dopo i 18 anni;
6. Dall'anamnesi o dall'osservazione, l'esordio dei sintomi è prima dei 10 anni.

Criteri di esclusione del disturbo da disregolazione dell'umore dirompente (DMDD)

1. I comportamenti non si verificano esclusivamente durante un episodio di disturbo di depressione maggiore e non sono spiegabili da altre psicopatologie (es: disturbo dello spettro autistico, disturbo depressivo persistente/distimia, disturbo post traumatico da stress o disturbo d'ansia da separazione).
2. Il disturbo dell'umore dirompente non può coesistere con il disturbo bipolare, il disturbo esplosivo intermittente o il disturbo oppositivo provocatorio (individui i cui sintomi soddisfano sia i criteri del DMDD che del

disturbo oppositivo provocatorio – Opposite Defiant Disorder, ODD –
dovrebbero ricevere solo la diagnosi di ODD.

(tradotto da Rao, 2014)