ARTICOLO BERSAGLIO

LA DEBOLE RELAZIONE FRA TEORIE DI INTELLIGENZA E TEST DI INTELLIGENZA

CESARE CORNOLDI E ENRICO TOFFALINI

Università di Padova

Riassunto. I test di intelligenza costituiscono probabilmente lo strumento psicologico maggiormente utilizzato. Tuttavia a questo grande uso non corrisponde una riflessione teorica sufficientemente ampia e soprattutto la relazione fra i test e la teoria appare debole, anche se alcuni proponenti dei test giustificano implicitamente questa debolezza lasciando margine ai singoli utilizzatori per una interpretazione legata ad una propria posizione teorica (ma poi non sempre mettono a disposizione degli utilizzatori le informazioni necessarie). Negli ultimi anni si è cercato di superare questa debolezza anche con il richiamo a precise posizioni teoriche (come per esempio la teoria CHC) che assumono la coesistenza di un fattore *g* unitario e di forme multiple dell'intelligenza. Il modo più diffuso e popolare con cui è stato fatto questo da parte degli utilizzatori dei test si è basato su una serie di assunzioni che però meritano una discussione teorica. Il presente articolo discute criticamente soprattutto le seguenti assunzioni: l'esistenza di un'abilità generale legata al fattore *g*, l'esistenza di abilità ampie ben definite, la loro indipendenza e equidistanza, la collocazione delle abilità ampie ad uno stesso livello della gerarchia, la stabilità della struttura dell'intelligenza nell'arco di vita.

1. INTRODUZIONE

Nel sentire comune c'è un ovvio rapporto fra l'idea di cosa sia l'intelligenza e la sua misurazione, ma questo rapporto è solo parzialmente testimoniato dalla ricerca psicologica di base e dalla ricerca applicata, perché molto spesso lo studio dell'intelligenza si è svolto indipendentemente dalla ricerca sugli strumenti di valutazione e questa, a sua volta, ha seguito percorsi che tenevano conto solo in parte della ricerca di base. Il presente intervento vuole contribuire a colmare questa lacuna con una discussione generale sul rapporto fra teorie e test e con un focus sui concetti fondamentali cui sembrano ispirarsi molti test contemporanei, soffermandosi sulla teoria cui si richiamano attualmente quasi tutti i test di intelligenza, cioè la teoria CHC.

Quando Boring (1923) scriveva che «l'intelligenza è ciò che è misurato dai test di intelligenza», non intendeva negare il costrutto né criticare i test, ma solo problematizzare l'uso dell'etichetta «intelligenza», che nel linguaggio ordinario ha un significato vago. E infatti commentava che sarebbe stato meglio se gli psicologi avessero usato un termine più tecnico, così da differenziare il costrutto teorico indagato dalla psicologia dalla «intelligenza» intesa nel linguaggio comune, ma ormai la distinzione non era stata fatta e avrebbe aggiunto ulteriori complicazioni lasciar convivere due termini.

Il rischio di una definizione circolare, che mette il carro dei test davanti ai buoi della teoria è comunque connaturato al fatto che i test di intelligenza costituiscono uno dei più significativi risultati pratici e applicativi della ricerca psicologica, per alcuni addirittura il maggiore di tutti i tempi (Herrnstein & Murray, 1994), testimoniato del resto dal larghissimo uso che se ne fa (sono gli strumenti maggiormente usati in Psicologia, v. per es. Evers et al., 2013). Se ne fa infatti quasi inevitabilmente riferimento nella clinica dello sviluppo, nelle decisioni sulle disabilità e sul deterioramento cognitivo, in questioni giudiziarie e nelle procedure di selezione scolastica o professionale (v. per es. Ones et al., 2005) anche perché la stima del fattore generale dell'abilità cognitiva, il Quoziente di Intelligenza (QI), generalmente assunto come vera sintesi dell'intelligenza, è un indice abbastanza rapido da misurare, attendibile, stabile, e capace di predire con buona approssimazione una vasta gamma di esiti nella vita reale, anche al netto di più ovvi aspetti come lo stato socio-economico di partenza.

La storia dei test di intelligenza dimostra che essi sono nati proprio per rispondere a esigenze pratiche. Il primo test famoso fu ideato da Binet e Simon (1905) per rispondere ad una richiesta del sistema scolastico francese volta a identificare il percorso scolastico più idoneo per ciascuno studente. I test utilizzati dall'esercito americano già all'epoca della Prima guerra mondiale servivano ad assegnare i soldati a diverse mansioni diversificate per complessità cognitiva (v. Yerkes, 1921). Ai tempi di Binet la precocità intellettiva, definita dall'età mentale, era considerata un predittore potente di capacità intellettiva, ben superiore al puro successo scolastico osservato in un dato momento, e questa ipotesi è stata poi confermata da numerosissime ricerche che hanno documentato che il QI (la versione moderna dell'età mentale), è predittore a lungo termine di successo scolastico e professionale, di capacità di affrontare i problemi della vita, di longevità e dello stesso QI rilevato di nuovo a 70 anni di distanza (v. Deary et al., 2004).

L'idea di intelligenza sottostante al concetto di *età mentale* era di tipo globale-maturativo e faceva riferimento ad un unico indice quan-

titativo rappresentato dal QI. Binet includeva infatti prove cognitive di svariata natura (verbali, visivi, mnestici, perfino motori) nella propria batteria, ma solo per avere una stima complessivamente più robusta dell'abilità generale, non perché credesse a tipi diversi di intelligenza. Questa idea è stata poi fatta propria dai principali test di intelligenza come quelli legati al nome di Wechsler. Wechsler (1958) infatti concepiva l'intelligenza come un'entità generale o globale (la cui misura era il QI di deviazione), che includeva diversi aspetti ma non implicava abilità multiple che meglio definissero l'intelligenza. Di fatto Wechsler proponeva l'uso di una batteria multicomponenziale solo per garantirne una maggiore attendibilità, e per evitare che particolari deficit o competenze legate all'uso di specifici tipi di materiale viziassero i risultati, non perché credesse in tipi diversi di intelligenza.

3. CENTRALITÀ DEL QI (E DI UN'ABILITÀ GENERALE)

Praticamente tutti i test di intelligenza forniscono una stima del QI come stima dell'intelligenza complessiva. In taluni casi si ritiene che tale stima sia sufficiente a riassumere le informazioni rilevanti sulle capacità intellettive di una persona, in altri, una misura fondamentale, ma non esaustiva.

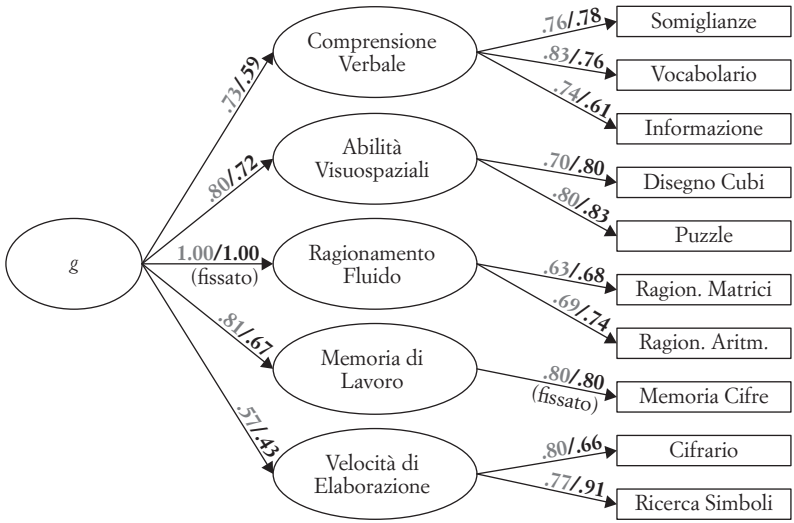
Il riferimento ad un singolo costrutto di intelligenza, di solito legato ad un fattore generale *g* che sta sotto i vari aspetti dell'intelligenza, e a un indice fondamentale, il QI, è basato sulle teorie unitarie dell'intelligenza. Le teorie unitarie traggono origine dalla considerazione che tipicamente distinte misure intellettive relative ad abilità apparentemente molto diverse risultano in realtà altamente correlate nella popolazione tipica e un singolo fattore (o una singola componente principale) spiega una percentuale di varianza molto superiore a qualsiasi ulteriore fattore. La teoria unitaria storicamente più importante è quella di Spearman (1904) che, non a caso, si valse ampiamente della statistica per sostenere questa posizione, ma essa è ampiamente presente anche oggi, sulla base della constatazione che la maggior parte della varianza comune tra diverse prove cognitive è attribuibile a un unico fattore generale anziché a fattori dominio-specifici (es. verbale, visivo, mnestico) (es. Watkins, 2006).

Tuttavia, le teorie unitarie sono anche state oggetto di critiche. Alcune critiche hanno osservato che la unitaria singola misura, non per niente spesso chiamata Quoziente di Intelligenza Totale QIT (*Full Scale Intelligence Quotient*, FSIQ) e non semplicemente QI, riguarda una somma di punteggi che riflettono funzioni della mente molto diverse, talvolta con riscontri psicometrici differenti. Dal punto di vista degli utilizzatori dei test, una critica fondamentale è rappresentata dal fatto

che il singolo valore di QI è insufficientemente informativo e, oltre a non giustificare i costi economici e di lavoro richiesti per la somministrazione di un'intera batteria di test, può non rendere interamente conto della complessità del profilo cognitivo di una persona, mascherando i suoi eventuali punti di forza e di debolezza, questione ampiamente discussa fra sostenitori di una teoria unitaria e psicologi che sostengono l'utilità di un profilo articolato (per una discussione che ci ha riguardato direttamente, v. Beaujean, 2017; Toffalini et al., 2017a, 2017b). A questo proposito, un uso fondamentale dei test di intelligenza riguarda popolazioni cliniche ove si evidenziano specifici disturbi incompatibili con una teoria unitaria e con il riferimento ad un singolo QI. Per esempio, numerosi studi hanno suggerito che varie popolazioni caratterizzate da disturbi del neurosviluppo, inclusi bambini e adulti con DSA, e bambini con ADHD, mostrano differenze rilevanti rispetto ai neurotipici sia nel profilo medio che nella struttura dell'intelligenza (Giofrè & Cornoldi, 2015; Giofrè et al., 2019; Pizzigallo et al., 2023; Toffalini et al., 2022; v. anche Figura 1, con i risultati ottenuti da Pizzigallo et al., 2023). Le differenze maggiori riguardano gli aspetti di elaborazione cognitiva di base (memoria di lavoro verbale e velocità di elaborazione visiva) che non solo dimostrano prestazioni medie deficitarie a dispetto di aspetti centrali dell'intelligenza in genere preservati, ma anche una minore saturazione rispetto al fattore *g*. Questo suggerisce che in tali popolazioni vi siano compromissioni specifiche relative al disturbo, che sono abbastanza scollegate dal funzionamento cognitivo generale. Va sottolineato che si tratta di popolazioni specifiche altamente selezionate per avere condizioni particolari (quali dislessia, difficoltà in matematica, o disattenzione e iperattività), e dunque per definizione «minoritarie» (o «neurodivergenti»). Ciononostante, esse dimostrano che almeno in alcuni casi semplificare lo studio dell'intelligenza a «tutto è *g*» non è un'operazione interamente appropriata.

Profili differenziati possono essere trovati talvolta anche in gruppi della popolazione altrimenti «neurotipica», per esempio comparando letterati, tipicamente con alta intelligenza verbale, e architetti, mediamente molto competenti sulla componente spaziale (Cavallini et al., 2009). E infatti gli stessi proponenti di test di intelligenza che offrono un QIT raccomandano di non considerare il QIT come attendibile nei casi in cui la dispersione fra diversi punteggi sia alta (v. Kaufman et al., 2015). A questo si potrebbe obiettare che la presenza di casi eccezionali e discrepanti è sempre prevista da qualsiasi modello statistico, ciò che conta essendo semmai la loro (limitata) frequenza e il loro (modesto) impatto nelle previsioni a livello di popolazione nell'insieme. Ma, per l'intelligenza, frequenza e impatto sono rilevanti in molti casi, e di fronte al caso singolo è sempre l'individuo che conta.

Figura 1
Modello a cinque fattori dell'intelligenza per subtest della WAIS-IV con loading standardizzati in un campione neurotipico versus con DSA della popolazione adulta italiana.



Legenda loadings: **Neurotipici / DSA**

Note. Loading standardizzati per la popolazione adulta generale ($N = 1424$, campione di standardizzazione della WAIS-IV) e per un campione adulto con DSA ($N = 298$, ottenuto via studio multicentrico) da un modello fattoriale confermativo di tipo gerarchico con fattore *g* al vertice, basato su una ricomposizione in 5 fattori analoga alla WISC-V. Sono stati fissati: il loading del fattore Memoria di Lavoro sul punteggio di Memoria di Cifre a un valore plausibile di .80, in quanto il fattore include un unico punteggio; e quello di *g* su Ragionamento Fluido a 1.00, in quanto quest'ultimo fattore avrebbe altrimenti presentato varianza leggermente negativa, suggerendo che il Ragionamento Fluido è sostanzialmente isomorfo a *g* (sebbene meno attendibile, perché composto da due soli *subtest*). RMSEA = 0.06, CFI = 0.96. Adattato da Pizzigallo et al. (2023).

In questi casi, mettere insieme tutti i punteggi di prove intellettive per valutare la capacità intellettiva generale può essere fuorviante. Si potrebbe pensare che questo ragionamento non valga per i soggetti con alto QIT perché, per ottenere un punteggio così alto, dovrebbero aver fatto bene un po' dappertutto. In realtà, va osservato che non è corretto aspettarsi un profilo mediamente omogeneo nei casi con elevato QI, dato che le diverse aree del funzionamento riflettono in misura diversa il fattore generale (es. Cornoldi et al., 2023). E inoltre il profilo intellettivo eterogeneo riscontrato per i DSA di intelligenza

media è stato riscontrato anche nei DSA di intelligenza superiore (Toffalini, Pezzuti & Cornoldi, 2017) e più in generale in tutta la popolazione con alto QI (Morrone et al., 2019).

Il rapporto tra teorie sull'intelligenza e test di intelligenza è complicato anche dal fatto che le teorie empiriche si basano necessariamente su un certo insieme di prove di intelligenza, e la scelta delle particolari prove finisce per influenzare i risultati e quindi il modello derivato. L'insistenza su una unica abilità generale potrebbe quindi essere in parte dovuta alla scelta delle prove utilizzate per valutarla, soprattutto se la scelta definitiva di prove e item si è basata sulla loro capacità di discriminare fra soggetti con alta e bassa abilità generale. Per esempio, è stato osservato che l'inclusione, poco frequente ma legittima, fra le prove intellettive di numerosi test sensoriali (Stankov, 2005) produce il risultato che un fattore generale spiega il 25% delle differenze fra gli individui, ma un secondo fattore legato ad abilità sensoriali ne spiega una proporzione di poco inferiore (il 20%). Ma lo stesso problema vale anche per quelle ricerche, ispirate ad altre teorie, in cui la scelta delle prove è stata influenzata dalla teoria adottata, con una conseguente forma più sottile di circolarità (la struttura dell'intelligenza che risulta dalle prove proposte è coerente con la teoria che ha ispirato la scelta delle prove) rispetto a quella classicamente denunciata (l'intelligenza è ciò che è misurato dai test).

4. TEORIE NON UNITARIE E LA TEORIA CHC

Nella storia della intelligenza sono state proposte molte teorie non-unitarie e anzi esse hanno raggiunto particolare fama ed evidenza, come è testimoniato dalla risonanza di nomi come quelli di Thurstone, Vernon, Gardner e Sternberg. Molte di queste teorie non hanno però avuto implicazioni dirette e di successo per la misurazione dell'intelligenza. Il caso è diverso per la cosiddetta teoria CHC proposta da Carroll (1993), su influenza di Horn e Cattell, e poi più volte sistematizzata. Per quanto non trovi un posto particolarmente rilevante nei Manuali di Psicologia Generale, la teoria CHC è assolutamente dominante nel campo della Diagnostica e della Psicometria, e si rifà soprattutto al contributo di John Carroll che ha sostenitori assolutamente entusiasti, come emerge da questa citazione di un suo erede: «I believe Carroll's (1993) publication represents to the field of applied psychometrics a work similar in stature to other principia publications in other scientific fields (e.g., Newton's three volumes, and Whitehead & Russell's, Principia Mathematica)» (McGrew, 2009).

La teoria a 3 strati di John Carroll (1993) costituisce una rielaborazione relativamente recente e più rappresentativa della posizione

psicometrica classica di tipo gerarchico. Attraverso la costituzione di una imponente, soprattutto per l'epoca, banca dati formata da più di 400 dataset relativi a 130.000 individui, Carroll pervenne a definire una struttura dell'intelligenza basata su tre strati organizzati gerarchicamente: uno strato III più elevato (fattore *g*), uno strato II con 8 o 9 abilità ampie e uno strato I con molti fattori specifici. Le abilità ampie sono state oggetto nel corso degli anni di descrizioni lievemente diverse, ma – riprendendole da una celebre sintesi (Flanagan et al., 2007) – possono essere ricondotte alle seguenti:

– intelligenza fluida, come capacità di risolvere problemi nuovi indipendentemente da competenze già acquisite;

– conoscenza quantitativa, come l'insieme delle abilità matematiche;

– intelligenza cristallizzata, come insieme di competenze e strategie principalmente codificate verbalmente, acquisite attraverso l'esperienza;

– visualizzazione, come capacità di percepire e manipolare mentalmente pattern visivi e rappresentazioni spaziali;

– memoria di lavoro, come capacità di mantenere temporaneamente informazioni nella memoria a breve termine;

– velocità d'elaborazione, come rapidità e accuratezza nell'eseguire compiti semplici ma che richiedono un controllo attentivo;

– immagazzinamento e recupero dalla memoria a lungo termine;

– percezione uditiva, come capacità di percepire e analizzare pattern sonori;

– capacità di lettura e scrittura.

La teorizzazione di Carroll presentava vari elementi in comune con altre importanti formulazioni proposte da John Horn e da Raymond Cattell celebri anche per il riconoscimento di due fondamentali fattori (Horn & Cattell, 1966), ovvero intelligenza fluida, che riguarda la capacità della mente di operare (in particolare classificare, seriare, ritrovare analogie) senza ricorrere a conoscenze acquisite, intelligenza «cristallizzata», che si riferisce invece alla attività della mente sostenute dalle conoscenze pregresse. C'erano somiglianze fra la teorizzazione di Carroll e quella di Horn e Cattell, ma ci furono anche scambi diretti fra questi autori e questo diede origine a una categoria di teorizzazioni siglate CHC dalle iniziali dei cognomi dei tre studiosi. In particolare, McGrew (2023) ricorda un incontro avvenuto nel Marzo 1986 a Dallas, in cui, oltre a Carroll e Horn, c'era Richard Woodcock, che diede spunto alla creazione di un test di intelligenza (legato ai nomi di Woodcok e Johnson) direttamente originato dall'approccio CHC.

La teoria CHC, oltre a diventare riferimento per quasi tutti i test di intelligenza, ha trovato sostegno in numerose ricerche soprattutto di ispirazione psicometrica (v. per es. Keith & Reynolds, 2010; McGrew, 2023; Schneider & McGrew, 2018) e sostenitori della loro

stretta relazione con i test, per es. con la WISC-V, come Wilson et al. (2023) che tuttavia non paiono del tutto disinteressati rappresentando in parte la casa editrice che ne detiene i diritti commerciali (un problema che si aggrava là dove la casa editrice non mette a disposizione i dati sullo strumento).

Nella divulgazione e nell'uso clinico la teoria CHC e altre teorie sono venute a caratterizzarsi tuttavia per una serie di punti che costituiscono una sistematizzazione semplificata, ma anche una forzatura del pensiero originario di Carroll (1993). Si è infatti schematizzata l'intelligenza come un insieme di «abilità» che si collocano a tre diversi livelli. Al livello più alto viene posta un'abilità generale *g* (di cui il QI totale è l'approssimazione numerica). Il fattore *g* viene dunque qui inteso come una particolare «abilità» e non, come sarebbe tecnicamente più corretto (e teoricamente meno impegnativo), come un «fattore latente», ovvero una causa o un insieme di concause comuni non direttamente osservate e la cui natura non è necessariamente specificata. A un livello intermedio viene posto un numero ristretto di abilità fondamentali ben identificate, del tutto distinte tra loro (ovvero ortogonali al netto di *g*), e di pari importanza gerarchica, su cui soprattutto si deve focalizzare la valutazione psicometrica. Infine, a un livello più basso, ci sono abilità specifiche che possono essere oggetto di valutazioni mirate, ma non sono di solito discusse nel report finale della valutazione intellettiva.

Questi punti richiedono una discussione approfondita che ne mostra, a nostro avviso, l'utilità euristica e pratica, ma anche evidenti limiti che devono essere presi in seria considerazione (come del resto usualmente avviene) dalla ricerca di base sull'intelligenza ma devono anche allertare (cosa che invece di solito non avviene) lo psicologo che usa il test di intelligenza. Nella parte che segue analizziamo e discutiamo alcune tipiche assunzioni fatte proprie da chi usa i test di intelligenza e che ci paiono particolarmente importanti e critiche.

5. IL FATTORE *G* COME ELEMENTO CENTRALE MA (FORSE) NON ESAUSTIVO

Come già anticipato, uno dei fenomeni empirici più forti e universalmente replicati in psicologia è la *positive manifold*, cioè la correlazione positiva tra prestazioni in compiti cognitivi anche completamente diversi. Chi è bravo in un compito di conoscenza del vocabolario, ad esempio, tende a esserlo anche in un compito di ragionamento non-verbale o in uno *span* di memoria di lavoro, a parità di altri fattori noti. Di questo si era già accorto Spearman (1904) a inizio del secolo scorso, che propose il costrutto di fattore *g* attribuendo alla positive manifold un valore centrale per capire la natura dell'intelligenza.

L'ipotesi e l'interesse per il fattore *g* sono sempre rimasti *mainstream* tra gli studiosi dell'intelligenza nel Novecento e lo sono ancora oggi, nonostante i ripetuti tentativi di proporre ipotesi multiple o multicomponenziali. La «potenza» di *g* deriva da un lato dalla sua capacità di spiegare la grande maggioranza della variabilità individuale interpretabile nei compiti cognitivi (ossia al netto dell'errore di misura), dall'altro dalla sua capacità di predire (come del resto abbiamo visto essere emerso per i test di intelligenza) esiti rilevanti nella vita reale, dal successo scolastico a quello lavorativo, dallo status socioeconomico raggiunto (al netto di quello di partenza) alla longevità (es. Warne, 2020).

La superiorità di *g* nello spiegare le differenze individuali in svariate prestazioni cognitive non dipende, si noti, dal fatto che queste ultime correlano sempre in alto grado (in parte le correlazioni sono addirittura smorzate dall'imperfetta attendibilità dei singoli test), ma dal fatto che il fattore generale spiega da solo la gran parte della varianza comune, lasciando ai fattori specifici (es. l'abilità verbale) quote modeste di varianza comune da spiegare. Ad esempio, nel campione di standardizzazione della WISC-IV americana oltre il 70% della varianza comune ai diversi subtest risulta attribuibile al solo fattore *g* (Watkins, 2006), indipendentemente dall'età, e risultati analoghi vengono ottenuti sistematicamente da diverse batterie di intelligenza in diversi Paesi.

La preponderanza di *g* è tale che i ricercatori mainstream che si occupano di intelligenza tendono a prediligere l'interesse per i fenomeni che riguardano il fattore generale (es. Beaujean, 2017; Deary et al., 2007; Jensen, 1998; Warne, 2020), mettendo i fattori specifici in secondo piano, se non relegandoli a elementi marginali da tenere in conto solo per garantire un buon *fit* del modello fattoriale. Un corollario importante per la discussione su teorie dell'intelligenza gerarchiche (come la CHC) è che tanto più è preponderante *g*, tanto meno importante e più difficile è stabilire con precisione un catalogo di abilità distinte e specifiche, potendo queste essere sempre definite e ridefinite in base al particolare insieme di compiti cognitivi che si è deciso di utilizzare, correndo il rischio di identificare «troppe intelligenze» (es. Austin & Saklofske, 2005). Spearman (1927) aveva postulato il principio dell'indifferenza dell'indicatore secondo cui il particolare compito cognitivo usato è irrilevante, perché qualsiasi scelta si faccia si misura soprattutto *g*, il rimanente «residuo specifico» potendo essere sostanzialmente derubricato a errore di misura. Se questa posizione è estrema, è anche vero che l'onere di dimostrare l'utilità del «residuo specifico» al netto della varianza comune è in capo a chi propugna l'importanza di misure di abilità distinte, cosa che infatti è stata fatta in alcuni casi. In particolare la sottolineatura delle abilità specifiche ha acquisito importanza nello studio e nella valutazione co-

gnitiva nelle popolazioni cliniche, ad esempio quelle con disturbi del neurosviluppo, come già spiegato sopra, caso comunque non irrilevante dal punto di vista scientifico e sociale.

Se la positive manifold è un fenomeno empirico indubitabile, un altro problema è che la sua natura rimane in parte misteriosa. Spearman, e con lui gran parte dei teorici dell'intelligenza da sempre, la attribuiscono a un singolo fattore latente chiamato fattore *g*, causa o insieme di concause unitarie che influiscono simultaneamente su tutti i tipi di abilità cognitiva di una persona. La stessa natura di «fattore latente» fa però apparire *g* un costrutto sfuggente, tanto più che nessun correlato cognitivo, genetico o neurobiologico riesce a renderne conto in modo pienamente soddisfacente. Secondo una buona metafora (Haier, 2014) il fattore *g* è un po' come la «materia oscura» del modello cosmologico standard, la sostanza che secondo i modelli matematici costituisce la gran parte della massa dell'universo, tiene insieme le galassie, è necessaria a spiegare numerosi fenomeni gravitazionali, eppure rimane così elusiva all'osservazione diretta che qualcuno dubita della sua esistenza e soprattutto della possibilità di poterla considerare una vera e propria entità o, nell'analogia fatta col fattore *g*, una vera e propria dimensione psichica, quella «abilità generale» che invece compare in certe applicazioni psicometriche delle teorie gerarchiche.

Una delle alternative al fattore *g* oggi presa in considerazione è l'ipotesi mutualista (van der Maas et al., 2006), secondo cui i diversi domini delle abilità cognitive nascono indipendenti e finiscono per correlare durante la maturazione del bambino, perché le interazioni reciproche tra i diversi domini di abilità sono mutuamente benefiche per il loro sviluppo. Questo genererebbe la positive manifold senza invocare un fattore *g* sovraordinato. A livello neurale, una maggiore o minore correlazione tra diverse aree di abilità potrebbe riflettere la struttura della rete di connessioni nel cervello, con il compromesso tra il costo energetico di collegare parti del cervello più o meno distanti e il vantaggio di connettere determinate aree che lavorano insieme (Bullmore & Sporns, 2012). Nonostante la plausibilità biologica di un'intelligenza pensata come rete di connessioni, l'ipotesi mutualista rimane minoritaria, anche a causa della sua limitata capacità a spiegare fenomeni quali la presenza di una forte positive manifold già in età molto precoci (cioè già dalle primissime età in cui la prestazione cognitiva diventa misurabile), la maggiore parsimonia dell'ipotesi di *g* (Gignac, 2014; Gignac & Kretzschmar, 2017), e l'esistenza dell'*effetto Jensen* (secondo cui un'abilità cognitiva è tanto più predittiva di un *outcome* o legata a un suo correlato quanto più fortemente riflette *g*, suggerendo che in effetti sia proprio *g* la cosa che conta veramente; Jensen, 1998).

Un'altra alternativa recente è la *process overlap theory* (Kovacs & Conway, 2016). Questa rielabora l'ipotesi cognitivista classica per cui compiti apparentemente diversi attingerebbero in realtà a fattori e risorse cognitive comuni, quali il controllo attentivo, la memoria di lavoro o la flessibilità cognitiva (Cornoldi, 2006). Anche questa teoria è rimasta finora minoritaria, forse anche per la difficoltà a isolare con chiarezza empirica inequivocabile queste presunte risorse cognitive comuni. Va anche detto che tali risorse «comuni», qualora venissero identificate, rappresenterebbero esse stesse *g* (inteso come causa comune), facendo della *process overlap theory* più un tentativo di esplicitare la natura del fattore *g* che una reale alternativa teorica.

Di fatto i test oggi più popolari si pubblicizzano proprio per la capacità di offrire una stima del fattore *g* (di cui il QIT può essere considerato un'approssimazione), ma anche di andare oltre esso e di offrire punteggi differenziati che mettano in luce punti di forza e di debolezza dell'individuo valutato. Sorprendentemente, però, in luogo di fare questo in base all'intero corpo di discussioni e teorie che hanno riguardato gli studi sull'intelligenza, tutte le batterie principali di test di intelligenza evocano, quasi si trattasse di una parola magica, la capacità del test di valutare la struttura dell'intelligenza prevista da una particolare teoria e cioè la teoria CHC. Lo fa ovviamente la batteria Woodcock-Johnson (che ha proprio avuto un legame storico con i proponenti del modello CHC; v. Schrank & Wendling, 2018), ma lo fanno anche le scale DAS (Elliott et al., 2007) e, con riferimento agli strumenti utilizzati anche in Italia, le scale Leiter (Roid & Koch, 2017), Wechsler (v. in particolare Wechsler, 2014), KABC (v. Kaufman & Kaufman, 2004) e IDS (Grob et al., 2018). Questo riferimento è spesso forzato perché la corrispondenza appare lieve o empiricamente solo parzialmente fondata (v. Dombrowski et al., 2021), addirittura comico nel caso delle scale Leiter (puramente non-verbali) che, per loro specifica missione, intendono prescindere da molte delle componenti fondamentali previste dal modello CHC. Da questo punto di vista lo strumento più credibile appare eventualmente l'IDS-2 (Intelligence and Developmental Scales for Children and Adolescents; Scale di intelligenza e di sviluppo per bambini e adolescenti) di Alexander Grob e Priska Hagmann-von Arx (Grob et al., 2018) (adattamento italiano di Ferri, Rea e Casagrande, 2022) che misura direttamente molte delle abilità ampie previste dal modello CHC (Elaborazione visiva, Velocita di elaborazione, Memoria fonologica a breve termine, Memoria visuospaziale a breve termine, Ragionamento concettuale, Ragionamento verbale e Memoria a lungo termine).

Come abbiamo visto dalla presentazione dei test di intelligenza un loro punto di forza viene attualmente considerato il fatto che riescono a offrire, oltre ad una misura del QIT, anche una misura delle abilità fondamentali, con un richiamo alla teoria CHC che, come viene tipicamente riportata, fornisce un elenco sistematizzato delle abilità umane, una specie di mappa articolata dell'intelligenza. Il punto più apprezzato dai clinici della teoria CHC è tuttavia quello per cui essa prevede un numero ben definito e limitato di abilità ampie, supportate da robusta ricerca, e quindi potenziale riferimento per ragionamenti e strumenti che vogliono offrire una riflessione sulla articolazione dell'intelligenza umana. Le abilità ampie assumono quindi una centralità che finisce per mettere in ombra tanto il fattore *g* quanto le abilità specifiche.

L'importanza data alle abilità ampie apparenta dunque, più di quanto possa sembrare, le teorie gerarchiche come la CHC alle teorie multiple dell'intelligenza che costituirono l'ovvia alternativa alle teorie unitarie nel riconoscimento di abilità intellettive distinte. Anche in questo caso la posizione ha illustri rappresentanti nella storia della Psicologia. In particolare Thurstone (1938) propose differenti componenti basilari dell'intelligenza che associò a un test (Primary Mental Abilities, PMA) che per l'appunto valutava queste abilità primarie (v. Thurstone & Thurstone, 1941). Per arrivare alla identificazione delle abilità primarie, Thurstone somministrò un'ampissima gamma di prove e lo fece, piuttosto che sui bambini, su giovani adulti, in cui il grado di differenziazione fra le abilità sarebbe maggiore. Queste scelte metodologiche crearono le premesse per la identificazione di abilità distinte. Le principali abilità primarie identificate da Thurstone erano le seguenti (fra parentesi è indicata una prova da lui utilizzata per riconoscere l'abilità; v. Cornoldi, 2007):

1) Comprensione verbale (es. capire un proverbio).

2) Fluenza verbale (es. produrre il maggior numero di parole che cominciano per una certa lettera).

3) Abilità numerica (es. compiere calcoli complessi, per es. qual è il 30% di 60?).

4) Memoria (es. apprendimento associativo, cioè memorizzare coppie di termini e quindi ricordare il secondo sulla base del primo).

5) Velocità percettiva (es. stimare se due figure sono identiche procedendo più velocemente possibile)

6) Ragionamento: induttivo (es. scoprire una regola dati gli esemplari, come nel caso della serie da completare 0, 11, 222, 3333, ?); deduttivo (es. valutare un sillogismo).

7) Visualizzazione spaziale (es. giudicare l'identità di due figure differentemente ruotate).

Un problema della teoria di Thurstone e di molte altre teorie che propongono forme diverse di intelligenza è che ciascuna fa riferimento a un insieme distinto di abilità. Per esempio, Sternberg ha proposto di distinguere tre forme fondamentali di intelligenza (v. Sternberg & Spear-Swerling, 1997) riferibili rispettivamente ad aspetti logici, pratici e creativi.

In una versione estrema dell'approccio multicomponenziale, l'ancor oggi popolare teoria di Gardner (2011), nella sua versione più famosa, propone sette forme distinte di intelligenza: *linguistica*, *musicale*, *logico-matematica*, *spaziale*, *corporeo-cinestetica*, *intrapersonale*, *interpersonale.* Il messaggio forte che essa fornisce è che ogni intelligenza non solo è ugualmente «nobile», ma anche ugualmente indipendente dalle altre, al punto che si può pensare a forme di intelligenza che sono ugualmente distanti fra loro.

La critica relativa allo scarso accordo delle teorie multiple su quali possano essere considerate le abilità fondamentali si applica anche alle teorie gerarchiche che abbiano l'ambizione di fornirne un elenco, compresa quindi la teoria CHC (come ammise lo stesso Carroll). Infatti anche studiosi che si richiamano al modello CHC ammettono che l'elenco delle abilità ampie può essere modificato. Per esempio Keith e Reynolds (2010) ritengono vadano ridefinite le abilità ampie relative alla intelligenza cristallizzata e alla memoria; McGrew (2009) propone di aggiungerne diverse legate alle varie modalità sensoriali.

Due problemi metodologici stanno sotto alla disputa di quali siano le abilità ampie:

1) Le abilità ampie emergono dal raggruppamento dei punteggi alle prove proposte e quindi la loro natura è influenzata dalla scelta di tali prove. Per quanto Carroll e altri studiosi si siano basati su una gamma ampia di prove rappresentative di diverse tradizioni e diversi ricercatori, è possibile che tali ricercatori fossero ispirati da idee comuni o semplicemente tenessero conto solo degli ambiti in cui era più facile trovare misure robuste, di modo che le prove non erano rappresentative di tutte le reali abilità che compongono l'intelligenza. Se, per esempio, fossero state incluse prove che misuravano le intelligenze molto diverse di cui parla Gardner (come l'abilità musicale o l'abilità corporea) forse sarebbero emerse altre abilità ampie.

2) La identificazione delle abilità ampie è anche legata al rapporto col fattore *g*. In cosa consiste una «abilità ampia»? in un modulo del funzionamento neurocognitivo separato e ben identificabile ma influenzato da *g*? in un residuo specifico di variabilità individuale di cui *g* non riesce a rendere conto all'interno di un modello psicometrico? Se la relazione con *g* è troppo forte, l'abilità diventa psicometricamente indistinguibile da *g* (e dunque inutile ai fini predittivi), mentre se la relazione è troppo debole (come nel caso di costrutti quali l'in-

telligenza «emotiva») se ne può invocare l'eliminazione da un modello cognitivo dell'intelligenza. Più in generale, dal punto di vista empirico, il proponente di un modello gerarchico potrebbe anche dare tutta la centralità a *g*, e mantenere le abilità ampie a mera garanzia della tenuta del fit statistico del modello fattoriale: paradossalmente, la preponderanza del fattore *g* sia nello spiegare la varianza comune tra diversi test, sia nel predire gli outcome rilevanti nella vita reale, potrebbe essere la ragione dietro al proliferare di svariate proposte riguardo le «intelligenze» specifiche: queste ultime, proprio per il fatto di avere un ruolo minore (al netto di *g*), sono sempre difficilmente confermabili o confutabili.

Data la pluralità di abilità ampie, un suggerimento potrebbe essere dunque quello di fissarsi su quelle *broad abilities* che servono effettivamente al professionista per la casistica che sta esaminando (probabilmente diverse in casistiche differenti), così che un modello «statico» che ha l'ambizione di abbracciare la totalità delle differenze individuali cognitive nella popolazione, del tipo CHC, risulterebbe essere al tempo stesso troppo ampio e poco utile. Ma anche questa posizione, di fatto agnostica relativa al quesito relativo alle abilità ampie, non risolverebbe il problema per cui non c'è accordo su quali esse siano.

7. GERARCHIA COME LIVELLI O COME UN CONTINUUM

Un problema delle teorie multiple è che danno l'impressione di accordare uguale importanza a tutte le forme di intelligenza, senza tenere conto che alcune espressioni dell'intelligenza sono più necessarie nella vita e nell'adattamento di una persona (per esempio la capacità verbale) e altre sono a tal punto indicative del funzionamento psichico generale (per esempio la logica) che un loro deficit generalmente si accompagna a difficoltà ad avere una vita autonoma. Queste e altre osservazioni hanno portato a ipotizzare una struttura gerarchica dell'intelligenza, in cui certe forme di intelligenza sono di gerarchia più elevata e altre di gerarchia più bassa.

Tuttavia il problema sembra proporsi anche per alcune teorie gerarchiche o almeno per la concettualizzazione e utilizzazione che di esse viene fatta. Se non altro per praticità descrittiva, alcune teorie gerarchiche, come la teoria CHC, vengono associate alla identificazione di strati in cui collocare le abilità, suggerendo l'idea che le abilità che si pongono nello stesso strato (I e II rispettivamente, nella teoria CHC) si pongano a uno stesso livello gerarchico. Possiamo convenire che l'idea di strati di abilità abbia un valore euristico, ma riteniamo che essa è ben lontana da una fotografia fedele del sistema di abi-

lità. In primo luogo, come lo stesso Carroll (1993) suggeriva, ci sono probabilmente abilità intermedie inesplorate, che si pongono fra uno strato e l'altro. Inoltre, non c'è motivo di pensare che le abilità di uno stesso strato debbano per forza avere uguale importanza nella gerarchia. Per esempio, le abilità ampie non sono ugualmente relate al fattore *g*, come del resto emergeva, sia pur con una differenziazione di valori non elevata (dovuta anche al fatto che in questo caso si tratta di medie calcolate su valori più variabili), nel lavoro di Carroll. Ecco nell'ordine le saturazioni medie di fattore *g* ritrovate, per una serie di abilità in gran parte riconducibili alle sue abilità ampie, trovate da Carroll (1993, p. 597) basandosi su una serie di studi (fra parentesi il numero di studi interessato): Induzione .57 (6), Visualizzazione .57 (12), Ragionamento quantitativo .51 (7), Abilità verbale .49 (13), Memoria associativa .46 (9), Relazioni spaziali .40 (12), Facilità numerica .39 (9), Ragionamento sequenziale .39 (13), Span di memoria .38 (10), Velocità percettiva .34 (10). Questi loading sono in genere più elevati e ancor più differenziati tra loro in altre indagini (per esempio a quelli generalmente trovati per analoghi indici di abilità ampie nei campioni di standardizzazione delle Wechsler). Infatti molte altre evidenze suggeriscono che le abilità ampie abbiano saturazioni di fattore *g* più diversificate, e che queste saturazioni configurino una specie di continuum al cui vertice c'è un'intelligenza fluida a stento distinguibile dal fattore *g*. Fra queste ricerche scegliamo come esempio quella svolta da noi recentemente (Pizzigallo et al., 2023, v. Figura 1) risultato dell'analisi dei punteggi ottenuti alla WAIS-IV da un campione di popolazione adulta italiana ($N = 1424$, da campione di standardizzazione della WAIS-IV, che si può ritenere complessivamente neurotipico) e da un campione con DSA (~~Disturbo Specifico dell'Apprendimento;~~ $N = 298$) raccolto da uno studio multicentrico. Come si può vedere, abbiamo isolato cinque abilità ampie, ma queste hanno saturazioni di *g* estremamente diversificate, a conferma di una ipotesi del continuum piuttosto che di un livello gerarchico unico. Inoltre, come già ricordavamo in precedenza, queste saturazioni sono minori negli adulti con DSA (e il calo è ancor più ampio nella popolazione evolutiva, v. Giofrè & Cornoldi, 2015) a riprova che nella popolazione clinica un approccio unitario non solo è meno informativo quando si scrive un report, ma anche meno giustificato empiricamente.

L'idea del continuum era del resto implicitamente presente nella teoria di Vernon (1950) e nell'analisi dell'intelligenza basata sul modello radex utilizzato da Marshalek, Lohman e Snow (1983), descrivendo l'intelligenza come un cerchio che ha al suo centro un ipotetico punto centrale rispetto al quale le varie abilità si discostano più o meno e si dispongono lungo diversi raggi che rappresentano le loro specificità. Il cerchio può essere inoltre diviso in spicchi; uno spicchio

potrebbe riferirsi alle abilità linguistiche, uno alle abilità spaziali, uno alle abilità matematiche, e includere prove più o meno centrali.

Tradizionalmente la gerarchia è stata definita su base psicometrica individuando un fattore più centrale che spiega la varianza comune, e poi fattori a diversi livelli che spiegano varianze specifiche, e solo a posteriori si è cercato di dare un senso psicologico al fattore *g* o alla gerarchia. Tuttavia, vi sono anche teorie che sono partite da una riflessione teorica, come nel caso della *teoria a cono* proposta da uno di noi (v. per es. Cornoldi, 2007) per esempio assumendo che un'abilità è tanto più centrale quanto più: 1) riesce ad essere esplicativa di ontogenesi e filogenesi, 2) *è essenziale al funzionamento psichico e adattivo,* 3) sorregge le altre abilità, 4) mette in gioco risorse controllate di memoria di lavoro. In questo caso, poiché prescinde dal riscontro psicometrico, il fattore gerarchicamente più elevato non corrisponde necessariamente al fattore *g* reperito statisticamente e può essere collocato più in alto, ma non necessariamente ad un vertice estremo della gerarchia. Per esempio, si può assumere che l'abilità di ragionamento logico sia gerarchicamente elevata, perché estremamente necessaria per il funzionamento mentale e adattivo, elemento di differenziazione ontogenetica e filogenetica, tale da richiedere grandi risorse di memoria di lavoro controllata, ma non necessariamente tale da spiegare una grande parte di varianza comune.

8. SOLO PARZIALE INDIPENDENZA (CONTINUITÀ?) A LIVELLO ORIZZONTALE

Come già abbiamo visto, nella breve analisi critica delle teorie multiple, la narrativa enfatica sul ritrovamento di un certo numero di abilità ampie ne assume una loro piena distinguibilità e ortogonalità al netto del sovraordinato fattore *g*, tale per cui sembrerebbero non esserci rapporti diretti fra le varie abilità o, se rapporti ci sono, sono tali da non privilegiare alcune abilità rispetto ad altre, secondo un principio di equidistanza. Questa posizione, per quanto non presente in forma esplicita a livello teorico, viene considerata caratteristica in molte descrizioni delle teorie fornite e viene fatta propria nell'uso dei test di intelligenza e nell'analisi clinica.

Le evidenze dimostrano tuttavia che le abilità ampie sono correlate fra loro, con vari gradi di relazione (discorso che vale anche allo strato inferiore relativamente a abilità specifiche riferibili alla stessa abilità ampia o a diverse abilità ampie). Ovviamente, la correlazione tra abilità ampie è prevista dal fatto che esiste un sovrastante fattore *g*, e anzi è ciò che ne giustifica l'esistenza. A noi però sembra che ci possa essere minore o maggiore vicinanza anche al netto di *g*, ovvero, in termini tecnici, che ci possano essere dei residui correlati. Ci sono

numerosissime evidenze, per esempio, del fatto che le abilità visive e le abilità spaziali sono fra loro più prossime, e invece più distanti dalle abilità verbali, anche al netto di quanto riflettano *g*. In un nostro studio abbiamo documentato, per esempio, l'esistenza di un *continuum* per cui l'abilità spaziale sequenziale si pone ad un polo e ha, a livello più prossimo (ovvero un maggiore grado di somiglianza e condivisione di risorse), l'abilità spaziale simultanea, cui segue l'abilità visiva, mentre l'abilità verbale si pone più distante, con una minima condivisione di risorse con l'abilità spaziale (Pazzaglia & Cornoldi, 1999). Inoltre certe abilità, per esempio la memoria di lavoro o la velocità di elaborazione, sono più prossime all'uno o all'altro versante a seconda del materiale implicato (per esempio, visivo o verbale). Questo diverso grado di apparentamento potrebbe essere associato alla scarsa «purezza» delle misure utilizzate, come accade nel caso della valutazione di memoria di lavoro o velocità con l'uso di materiali che mettono inevitabilmente in gioco altre abilità. Secondo questa obiezione, togliendo il peso del materiale, si otterrebbe una misura più pulita e indipendente, ma c'è da domandarsi se questo ragionamento abbia senso, dal momento che tali misure non solo richiedono necessariamente l'uso di un qualche materiale, ma assumono anche significato solo nell'inscindibile rapporto fra processo e materiale.

Si può quindi pensare che le abilità fondamentali non siano equidistanti e del tutto indipendenti, ma si pongano in una certa collocazione del sistema gerarchico non solo con riferimento alla gerarchia, ma anche con riferimento alla loro maggiore o minore prossimità. Se condividiamo l'ipotesi di una pluralità di abilità, il grado diverso di relazioni fra le diverse abilità può dunque essere rappresentato come una specie di continuum orizzontale, o forse meglio come più continua che tengano conto di più dimensioni implicate dalle varie abilità, come per l'appunto previsto dal modello a cono (Cornoldi, 2007). Va da sé che, se si pensa a dei *continua* piuttosto che a dei fattori ben distinti in modo discreto, anche il problema della identificazione del numero e delle caratteristiche delle abilità viene meno perché ogni punto dell'insieme dei continua definisce potenzialmente un'abilità.

9. VARIAZIONI DELLA STRUTTURA IN GRUPPI DIVERSI

Un problema con le teorie dell'intelligenza che propongono strutture molto dettagliate, fra cui la CHC, è che esse si propongono per valere per tutta la popolazione, indipendentemente dalle caratteristiche dei rispondenti (e invece abbiamo visto che l'organizzazione può essere diversa in presenza di determinati disturbi, e anche in funzione dell'età).

Ci sono fenomeni e aspetti che suggeriscono invece la necessità di avere una teoria flessibile in grado di tenere conto in modo puntuale di certe variazioni fra gruppi e in particolare di fasce d'età diverse. Un caso celebre è quello della differenziazione/dedifferenziazione, per cui a certe età (per esempio nel giovane adulto) le abilità intellettive risultano ampiamente differenziate e ad altre età (per esempio nell'anziano) essere molto meno differenziate (v. per es. Koen et al., 2020). Ma vi sono numerosissimi altri esempi che riguardano abilità più specifiche come, per citare casi a noi meglio noti, la memoria di lavoro dei bambini piccoli (v. Carretti et al., 2022) o il controllo esecutivo (v. Feraco & Cona, 2022).

Si potrebbe obiettare che le teorie, le quali in genere non sono costruite per prendere in considerazione particolari caratteristiche di certi gruppi clinici o di età, propongono solo una «configurazione» generale di fattori, che dovrebbe valere all'incirca sempre. Le eventuali specificità emerse in certi gruppi clinici o anagrafici potrebbero così intaccare aspetti particolari dell'invarianza di misura (es. loadings, varianze dei residui) e non minare per forza la validità «configurale» della teoria. Quand'è così, tuttavia, si deve prendere atto non solo di una certa genericità della teoria, ma anche dei limiti di generalizzabilità e di sensibilità dello strumento psicometrico derivatone per valutare l'intelligenza (e proposto in genere per ampie gamme di età e di condizioni cliniche).

10. PER UN TEST DI BASE

Come già si anticipava, la fortuna (meritata) dei test di intelligenza non nasce dal fatto che riflettono una particolare teoria dell'intelligenza, ma dal fatto che sono capaci di fornire al professionista informazioni estremamente importanti sul funzionamento della persona e su quello che potrà essere e fare. Molti clinici dichiarano infatti che le informazioni più importanti e affidabili ricavate dai loro assessment riguardano i risultati dei test di intelligenza. Che questo elemento sia cruciale è provato anche dalla storia di David Wechsler, legato ai test più famosi e utilizzati di intelligenza, che aveva un'idea vaga e «aperta» di intelligenza e si preoccupava piuttosto che il test funzionasse bene. Wechsler, infatti, si limitò a assemblare una batteria che raccogliesse prove con buone proprietà psicometriche e che rappresentasse aspetti diversi del funzionamento mentale. Le prove scelte non erano ideate da lui ma facevano parte della tradizione psicometrica del tempo, dunque non nascevano da una vera e propria teoria. E tuttavia le prove scelte da Wechsler e poi dai suoi successori, per la loro solidità misurativa e la loro flessibilità, hanno portato non solo

a infiniti usi che ne hanno rivelato la loro utilità, ma anche alla loro utilizzabilità a fini di ricerca, sia relativa a profili specifici, sia relativa alla teorizzazione sull'intelligenza. Questa utilizzabilità, almeno con riferimento allo strumento più utilizzato, e cioè la WISC, viene però oggi minata dal fatto che le informazioni psicometriche sullo strumento, almeno in Italia, non sono più rese disponibili al ricercatore.

Volendo seguire la logica di Wechsler per costruire uno strumento di libero accesso, come da più parti viene proposto (con un ottimismo forse eccessivo sulla fattibilità della operazione), una batteria dovrebbe quindi avere diverse componenti. Suggeriamo che includa almeno una prova di ragionamento logico (preferibilmente del tipo «Matrici»), una prova di ragionamento verbale di tipo induttivo (del tipo «reperimento di elemento comune fra due elementi» o di «individuazione dell'intruso in una serie di elementi»), una prova di ragionamento spaziale (combinazione di elementi spaziali? Rotazione mentale?), una prova di intelligenza cristallizzata verbale (vocabolario? Comprensione del testo?) e infine qualche prova specifica di velocità/attenzione e di memoria di lavoro da includere nel calcolo del QI totale solo a condizione che i punteggi emersi non siano troppo divergenti. Una batteria del genere dovrebbe poi essere soggetta a una sperimentazione volta a verificare le proprietà psicometriche e a una ricerca di validazione. Nel rispetto della tradizione scientifica, bisognerebbe poi intendere sempre gli strumenti e i risultati ottenuti come una base, e non come il punto d'arrivo definitivo di una teorizzazione sulla intelligenza. Questo sarebbe tanto più possibile quanto più i materiali fossero oggetto di una sperimentazione continua e i dati fossero disponibili alla comunità scientifica, cosa che invece oggi – come menzionavamo – succede in parte o di fatto per nulla.

BIBLIOGRAFIA

Austin, E. J., & Saklofske, D. H. (2005). Far too many intelligences? On the communalities and differences between social, practical, and emotional intelligences. In R. Schulze & R. D. Roberts (Eds.), *Emotional intelligence: An international handbook* (pp. 107-128). Hogrefe & Huber Publishers.

Beaujean, A. A. (2017). Commentary on «strengths and weaknesses in the intellectual profile of different subtypes of specific learning disorder» (Toffalini et al., 2017). *Clinical Psychological Science*, *5*(5), 874-877. https://psycnet.apa.org/doi/10.1177/2167702617708629

Binet, A., & Simon, T. (1905). New methods for the diagnosis of the intellectual level of subnormals. *L'Année Psychologique*, *12*(1), 191-244.

Boring, E. G. (1923) Intelligence as the tests test it. *New Republic*, *36*, 35-37.

Bullmore, E., & Sporns, O. (2012). The economy of brain network organization. *Nature Reviews Neuroscience*, *13*(5), 336-349. https://doi.org/10.1038/nrn3214

Carretti, B., Giofrè, D., Toffalini, E., Cornoldi, C., Pastore, M., & Lanfranchi, S. (2022). Structure of working memory in children from 3 to 8 years old. *Developmental Psychology*, *58*(9), 1687-1701. https://doi.org/10.1037/dev0001385

Carroll, J., B., (1993). *Human cognitive abilities: A survey of factor-analytic studies.* Cambridge University Press.

Cavallini, E., Cornoldi, C., & Vecchi, T. (2009). The effects of age and professional expertise on working memory performance. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition*, *23*(3), 382-395. https://psycnet.apa.org/doi/10.1002/acp.1467

Cornoldi, C. (2006). The contribution of cognitive psychology to the study of human intelligence. *European Journal of Cognitive Psychology*, *18*(1), 1-17. http://dx.doi.org/10.1080/09541440500215889

Cornoldi, C. (2007). *L'intelligenza*. Il Mulino.

Cornoldi, C., Giofrè, D., & Toffalini, E. (2023). Cognitive characteristics of intellectually gifted children with a diagnosis of ADHD. *Intelligence*, *97*, articolo 101736. https://doi.org/10.1016/j.intell.2023.101736

Deary, I. J., Strand, S., Smith, P., & Fernandes, C. (2007). Intelligence and educational achievement. *Intelligence*, *35*(1), 13-21. https://psycnet.apa.org/doi/10.1016/j.intell.2006.02.001

Deary, I. J., Whiteman, M. C., Starr, J. M., Whalley, L. J., & Fox, H. C. (2004). The impact of childhood intelligence on later life: following up the scottish mental surveys of 1932 and 1947. *Journal of Personality and Social Psychology*, *86*(1), 130-147. https://doi.org/10.1037/0022-3514.86.1.130

Dombrowski, S. C., McGill, R. J., & Morgan, G. B. (2021). Monte Carlo modeling of contemporary intelligence test (IQ) factor structure: Implications for IQ assessment, interpretation, and theory. *Assessment*, *28*(3), 977-993. https://psycnet.apa.org/doi/10.1177/1073191119869828

Elliott, C. D., Salerno, J. D., Dumont, R., & Willis, J. O. (2018). The Differential Ability Scales – Second Edition. In D. P. Flanagan & E. M. McDonough (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (4th ed., pp. 360-382). The Guilford Press.

Evers, A., Muñiz, J., Bartram, D., Boben, D., Egeland, J., Fernández-Hermida, J. R., Frans, Ö., Gintiliené, G., Hagemeister, C., Halama, P., Iliescu, D., Jaworowska, A., Jiménez, P., Manthouli, M., Matesic, K., Schittekatte, M., Sümer, H. C., & Urbánek, T. (2012). Testing practices in the 21st century: Developments and European psychologists' opinions. *European Psychologist*, *17*(4), 300-319. https://doi.org/10.1027/1016-9040/a000102

Feraco, T., & Cona, G. (2022). Differentiation of general and specific abilities in intelligence. A bifactor study of age and gender differentiation in 8-to 19-year-olds. *Intelligence*, *94*, articolo 101669. https://doi.org/10.1016/j.intell.2022.101669

Flanagan, D. P., Ortiz, S. O., & Alfonso, V. C. (2007). Essentials of cross-battery assessment. (2nd ed.). John Wiley & Sons Inc.

Gardner, H. (2011). *Frames of mind: The theory of multiple intelligences*. Basic Books.

Gignac, G. E. (2014). Dynamic mutualism versus *g* factor theory: An empirical test. *Intelligence*, *42*, 89-97. https://psycnet.apa.org/doi/10.1016/j.intell.2013.11.004

Gignac, G. E., & Kretzschmar, A. (2017). Evaluating dimensional distinctness with correlated-factor models: Limitations and suggestions. *Intelligence*, *62*, 138-147. https://psycnet.apa.org/doi/10.1016/j.intell.2017.04.001

Giofrè, D., & Cornoldi, C. (2015). The structure of intelligence in children with specific learning disabilities is different as compared to typically developing children. *Intelligence*, *52*, 36-43. https://doi.org/10.1016/j.intell.2015.07.002

Giofrè, D., Pastore, M., Cornoldi, C., & Toffalini, E. (2019). Lumpers vs. splitters: Intelligence in children with specific learning disorders. *Intelligence*, *76*, articolo 101380. https://doi.org/10.1016/j.intell.2019.101380

Grob, A., Hagmann-von Arx, P., Jaworowska, A., Matczak, A., & Fecenec, D. (2018). *Intelligence and Development Scales-2 (IDS-2)*. Hogrefe.

Haier, R. J. (2014). The universe, dark matter, and streaming intelligence. *Intelligence*, *42*, 152-155. https://doi.org/10.1016/j.intell.2013.07.022

Herrnstein, R. J., & Murray, C. A. (1994). *The bell curve: Intelligence and class structure in American life.* Free Press.

Horn, J. L., & Cattell, R. B. (1966). Refinement and test of the theory of fluid and crystallized general intelligences. *Journal of Educational Psychology*, *57*(5), 253.

Jensen, A. R. (1998). *The g factor: The science of mental ability*. Praeger Publishers/Greenwood Publishing Group.

Kaufman, A. S., & Kaufman, N. L. (2004). *Kaufman brief intelligence test* (2nd ed.). American Guidance Service Inc.

Kaufman, A. S., Raiford, S. E., & Coalson, D. L. (2015). *Intelligent testing with the WISC-V*. John Wiley & Sons Inc.

Keith, T. Z., & Reynolds, M. R. (2010). Cattell-Horn-Carroll abilities and cognitive tests: What we've learned from 20 years of research. *Psychology in the Schools*, *47*(7), 635-650. https://doi.org/10.1002/pits.20496

Koen, J. D., Srokova, S., & Rugg, M. D. (2020). Age-related neural dedifferentiation and cognition. *Current Opinion in Behavioral Sciences*, *32*, 7-14. https://doi.org/10.1016%2Fj.cobeha.2020.01.006

Kovacs, K., & Conway, A. R. (2016). Process overlap theory: A unified account of the general factor of intelligence. *Psychological Inquiry, 27*(3), 151-177. https://awspntest.apa.org/doi/10.1080/1047840X.2016.1153946

Marshalek, B., Lohman, D. F., & Snow, R. E. (1983). The complexity continuum in the radex and hierarchical models of intelligence. *Intelligence*, *7*(2), 107-127. https://doi.org/10.1016/0160-2896(83)90023-5

McGrew, K. S. (2009). CHC theory and the human cognitive abilities project: Standing on the shoulders of the giants of psychometric intelligence research. *Intelligence*, *37*(1), 1-10. https://doi.org/10.1016/j.intell.2008.08.004

McGrew, K. S. (2023). Carroll's three-stratum (3S) cognitive ability theory at 30 years: Impact, 3S-CHC theory clarification, structural replication, and cognitive-achievement psychometric network analysis extension. *Journal of Intelligence*, *11*(2), articolo 32. https://doi.org/10.3390/jintelligence11020032

Morrone, C., Pezzuti, L., Lang, M., & Zanetti, M. A. (2019). Analisi del profilo WISC-IV in un campione italiano di bambini e adolescenti intellettualmente gifted. *Psicologia Clinica dello Sviluppo*, *23*(1), 71-96. http://dx.doi.org/10.1449/93134

Ones, D. S., Viswesvaran, C., & Dilchert, S. (2005). Cognitive Ability in Selection Decisions. In O. Wilhelm & R. W. Engle (Eds.), *Handbook of understanding and measuring intelligence* (pp. 431-468). Sage Publications. https://doi.org/10.4135/9781452233529.n24

Pazzaglia, F., & Cornoldi, C. (1999). The role of distinct components of visuo-spatial working memory in the processing of texts. *Memory*, *7*, 19-41.

Pizzigallo, E., Cornoldi, C., Buono, S., Città, S., Viola, F., & Toffalini, E. (2023). The intellectual profile of adults with specific learning disabilities. *Journal of Intelligence*, *11*(12), articolo 223. https://doi.org/10.3390/jintelligence11120223

Roid, G. H., & Koch, C. (2017). Leiter-3: Nonverbal cognitive and neuropsychological assessment. In R. S. McCallum (Ed.), *Handbook of nonverbal assessment* (2nd ed., pp. 127-150). Springer International Publishing/ Springer Nature. https://doi.org/10.1007/978-3-319-50604-3_8

Schrank, F. A., & Wendling, B. J. (2018). The Woodcock-Johnson IV: Tests of cognitive abilities, tests of oral language, tests of achievement. In D. P. Flanagan & E. M. McDonough (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (4th ed., pp. 383-451). The Guilford Press.

Schneider, W. J., & McGrew, K. S. (2018). The Cattell-Horn-Carroll theory of cognitive abilities. In D. P. Flanagan & E. M. McDonough (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (4th ed., pp. 73-163). The Guilford Press.

Spearman, C. (1904). General intelligence, objectively determined and measured. *The American Journal of Psychology*, *15*, 201-292. https://doi.org/10.2307/1412107

Spearman, C. (1927). *The Abilities of Man*. Macmillan.

Stankov, L. (2005). g-Factor: Issue of design and interpretation, In O. Wilhelm e R. W. Engle (Eds.), *Handbook of Understanding and Measuring Intelligence*. Sage Publications.

Sternberg R. J., & Spear-Swerling, L. (1997), *Le tre intelligenze. Come potenziare le capacità analitiche, creative e pratiche*. Erickson.

Thurstone, L. L. (1938). *Primary mental abilities*. University of Chicago Press.

Toffalini, E., Buono, S., & Cornoldi, C. (2022). The structure, profile, and diagnostic significance of intelligence in children with ADHD are impressively similar to those of children with a specific learning disorder. *Research in Developmental Disabilities*, *129*, articolo 104306. https://doi.org/10.1016/j.ridd.2022.104306

Toffalini, E., Giofrè, D., & Cornoldi, C. (2017a). Strengths and weaknesses in the intellectual profile of different subtypes of specific learning disorder: A study on 1,049 diagnosed children. *Clinical Psychological Science*, *5*(2), 402-409. http://dx.doi.org/10.1177/2167702616672038

Toffalini, E., Giofrè, D., & Cornoldi, C. (2017b). Pros and cons of using intelligence batteries for the study of clinical populations: A response to Beaujean (2017). *Clinical Psychological Science*, *5*(5), 878-879. http://dx.doi.org/10.1177/2167702617720716

Toffalini, E., Pezzuti, L., & Cornoldi, C. (2017). Einstein and dyslexia: Is giftedness more frequent in children with a specific learning disorder than in typically developing children? *Intelligence*, *62*, 175-179. https://psycnet.apa.org/doi/10.1016/j.intell.2017.04.006

Van Der Maas, H. L., Dolan, C. V., Grasman, R. P., Wicherts, J. M., Huizenga, H. M., & Raijmakers, M. E. (2006). A dynamical model of general intelligence: The positive manifold of intelligence by mutualism. *Psychological Review*, *113*(4), 842-861. https://doi.org/10.1037/0033-295x.113.4.842

Vernon, P.E. (1950). *The structure of human abilities.* Methuen.

Warne, R. T. (2020). *In the know: Debunking 35 myths about human intelligence*. Cambridge University Press.

Wechsler, D. (1958). *The measurement and appraisal of adult intelligence (4th ed.).* Williams & Wilkins Co. https://doi.org/10.1037/11167-000

Wechsler, D. (2014). *Wechsler Intelligence Scale for Children (fifth edition)*. Pearson.
Yerkes, R. M. (Ed.) (1921). *Psychological examining in the US army: Memoirs of the National Academy of Sciences (Vol. 15)*. US Government Printing Office.

The weak relationship between models of intelligence and tests of intelligence

Summary. Psychological tests probably represent the most used psychological tool. However this use is not associated with a corresponding development of theoretical reflection. Furthermore this reflection is only weakly related with the creation and use of the tests. Some tests implicitly consider this weakness as a strength as it leaves the user and the researcher free of interpreting the data on the basis of their data and particular theoretical orientation, but an increasing tendency to not make available the basic psychometric information actually limits this possibility. In the last years test producers have also tried to reduce this weakness making reference to intelligence models (in particular the CHC model) assuming that the main indexes of intelligence are related with general intelligence and/or broad intellectual abilities. The most popular and simplified view followed by practioners has consequently adopted a series of assumptions that require a critical analysis. The present paper discusses some of these assumptions and in particular the existence of a general ability (related with *g* factor, the existence of a well defined series of broad abilities, their independence and equidistance, their location at a same level of the hierarchical structure of intelligence, the stability of the intelligence structure across the life span.

Keywords: Structure of intelligence, intelligence tests, CHC model, general ability, broad abilities.

La corrispondenza va inviata a Cesare Cornoldi, Dipartimento di Psicologia Generale, Università di Padova, Via Venezia 8, 35131 Padova.
E-mail: cesare.cornoldi@unipd.it