



LJMU Research Online

Giofré, D, Toffalini, E and Provazza, S

Are the WISC-IV Italian norms similar to the UK norms? A comparison between the two standardizations

<http://researchonline.ljmu.ac.uk/id/eprint/6698/>

Article

Citation (please note it is advisable to refer to the publisher's version if you intend to cite from this work)

Giofré, D, Toffalini, E and Provazza, S (2017) Are the WISC-IV Italian norms similar to the UK norms? A comparison between the two standardizations. *Psicologia Clinica dello Sviluppo*, 21 (1). pp. 143-154. ISSN 1824-078X

LJMU has developed **LJMU Research Online** for users to access the research output of the University more effectively. Copyright © and Moral Rights for the papers on this site are retained by the individual authors and/or other copyright owners. Users may download and/or print one copy of any article(s) in LJMU Research Online to facilitate their private study or for non-commercial research. You may not engage in further distribution of the material or use it for any profit-making activities or any commercial gain.

The version presented here may differ from the published version or from the version of the record. Please see the repository URL above for details on accessing the published version and note that access may require a subscription.

For more information please contact researchonline@ljmu.ac.uk

<http://researchonline.ljmu.ac.uk/>

The Version of Scholarly Record of this Article is published in *Psicologia Clinica dello Sviluppo*, available online at: <http://dx.doi.org/>. Note that this article may not exactly replicate the final version published in *Psicologia Clinica dello Sviluppo*.

Giofrè, D., Toffalini, E., & Provazza, S. (2017). Are the WISC-IV Italian norms similar to the UK norms? A comparison between the two standardizations. *Psicologia Clinica Dello Sviluppo*, (1), 143–154.

La WISC-IV sovrastima le competenze dei ragazzi italiani? Discrepanze tra la standardizzazione UK e quella Italiana della scala.

David Giofrè¹, Enrico Toffalini², & Serena Provazza²

¹Natural Sciences and Psychology, Liverpool John Moores University, UK

²Dipartimento di Psicologia Generale, Univesità di Padova, Italy

Manuscript accepted: 23/09/2016

Per corrispondenza,
Dr. David Giofrè
Liverpool John Moores University
Natural Sciences and Psychology
Tom Reilly Building Byrom Street, Liverpool, L3 3AF
Email. david.giofre@gmail.com
Tel. +44 151 904 6336
Fax. +44 151 904 6302

Abstract

La WISC-IV è la batteria psicologica più utilizzata al mondo; tuttavia non tutti i subtest rimangono gli stessi nei vari adattamenti della scala. Rimangono praticamente invariati, infatti, solo quelli relativi all'indice di ragionamento visuoperceptivo e di velocità di elaborazione. In questo lavoro abbiamo confrontato le norme italiane con quelle del Regno Unito in questi indici: i risultati mettono in luce come le differenze nell'indice di velocità di elaborazione siano tutto sommato modeste, mentre vi sono delle differenze più marcate per quanto riguarda l'indice di ragionamento visuoperceptivo. Le disuguaglianze tra le standardizzazioni, inoltre, non sono omogenee ma tendono a concentrarsi in alcune specifiche fasce di età. Questo risultato ha importanti ripercussioni per il calcolo dell'indice di ragionamento percettivo e del QI e sembra essere dovuto ad assunzioni diverse sulle curve di crescita di alcuni subtest. Nel presente contributo verranno esaminate le implicazioni cliniche e teoriche dello studio.

Are the WISC-IV Italian norms similar to the UK norms? A comparison between the two standardizations.

Abstract. The WISC-IV is the most popular psychological battery in the world. However, not all subtests remain the same in the various adaptations of the scale. In fact, only the subtests included in the perceptual reasoning and in the processing speed indexes remain virtually the same across various adaptations of the scale. In this paper we compared Italian with United Kingdom norms in these two indices. Results showed that differences in the processing speed index were small and that they were larger in the perceptual reasoning index. Further, differences between the two standardizations, were not homogeneous, but concentrated in specific age groups. The clinical and theoretical implications of the study will be discussed.

Keywords: QI, intelligenza; WISC-IV; bambini.

**La WISC-IV sovrastima le competenze dei ragazzi italiani?
Discrepanze tra la standardizzazione UK e quella Italiana della scala.**

Nel panorama italiano e internazionale esistono numerosi strumenti per la valutazione dell'intelligenza; tra di essi, le scale Wechsler sono sempre state tra le più utilizzate per l'assessment in ambito psicologico. Di fatto la WISC-IV (*Wechsler Intelligence Scale for Children-fourth edition*) è la batteria più utilizzata nella valutazione psicologica in ambito evolutivo a livello mondiale (Evers et al., 2012) grazie alla sua versatilità, multicomponenzialità e praticità d'uso. Non sorprende quindi che le scale ed i loro adattamenti siano disponibili in molti Paesi tra cui, ad esempio, Stati Uniti (USA), Regno Unito (UK) ed Italia.

Nei paesi come il Regno Unito, l'adattamento della scala al contesto locale è un'operazione necessaria; le scale Wechsler, infatti, pur essendo prodotte negli Stati Uniti, vengono comunque adattate anche nei restanti Paesi anglofoni, poiché vi sono alcuni stimoli che risultano familiari ad un bambino americano e che lo sono poco o per nulla per un bambino inglese, e viceversa. Anche il linguaggio è lievemente diverso e richiede quindi piccoli adattamenti. La realtà italiana, invece, è un po' più complessa e necessita di adattamenti maggiori: gli stimoli verbali ad esempio, devono essere tradotti e adattati dall'inglese all'italiano. Per questa ragione vi sono prove che cambiano poco o per nulla nei vari Paesi (ad esempio disegno con cubi, ragionamento con matrici, concetti illustrati e ricerca di simboli), e prove che, essendo di natura verbale (ad esempio vocabolario, somiglianze, e comprensione), subiscono notevoli cambiamenti nell'adattamento ad una lingua diversa.

Un discorso a parte va fatto per le prove di "memoria di lavoro" che richiedono il mantenimento, l'elaborazione ed il ricordo di stimoli presentati oralmente. Infatti, la memoria di lavoro, almeno per quanto riguarda la WISC-IV, è di tipo verbale e possono pertanto emergere delle differenze negli adattamenti ad una lingua diversa. Nel caso della memoria di cifre, che consiste nella ripetizione in avanti o all'indietro di numeri, gli stimoli presentati non cambiano tra una cultura e un'altra, ma necessitano di essere tradotti, processo che può portare ad alcune discrepanze.

Un altro esempio è il riordinamento di lettere e numeri, che richiede di riordinare prima le lettere in ordine alfabetico e poi i numeri in ordine crescente, in cui lettere di uso comune nelle lingue anglofone (ad esempio J, K, X, Y e W) sono state sostituite con altre.

Il fatto che alcune prove rimangano sostanzialmente invariate mentre altre cambino sensibilmente ha delle conseguenze notevoli se si vogliono confrontare i risultati ottenuti in standardizzazioni diverse. Infatti, il confronto tra culture anche molto diverse tra di loro (Inglese, Statunitense, e Italiana) è possibile soltanto nelle prove che per loro natura cambiano molto poco; al contrario, il confronto è molto difficoltoso nelle prove che richiedono un “accomodamento”. Per quanto riguarda nello specifico la WISC-IV, che è l’ultima versione della batteria in Italia, le prove che rimangono pressappoco inalterate sono disegno con cubi, concetti illustrati, ragionamento con matrici (che fanno tutte parte dell’indice di ragionamento visuoperceptivo o IRP), cifrario e ricerca di simboli (che invece fanno parte dell’indice di velocità di elaborazione o IVE).

I confronti tra edizioni ed adattamenti diversi sono molto comuni nei paesi anglofoni; se consideriamo la versione UK della WISC-IV (Wechsler, 2003) possiamo notare che le standardizzazioni USA e UK sono avvenute nello stesso periodo ed i risultati della versione UK sono stati confrontati con quelli USA (Tabelle C.4 e C5; Wechsler, 2003). I punteggi tra le due culture si sono dimostrati molto simili, ed infatti le piccole differenze emerse sono state attribuite alla campionatura e “corrette” tramite metodi di *smoothing* (ossia aggiustamenti ottenuti attraverso particolari analisi statistiche) o tramite il cosiddetto *continuous norming* (si veda Zachary & Gorsuch, 1985 per una trattazione approfondita). Pertanto, le due versioni della scala WISC-IV – seppur diverse soprattutto nella parte verbale, la quale è stata adattata al contesto UK – sono sostanzialmente sovrapponibili. Questo vuol dire che un ipotetico bambino inglese con un determinato punteggio in uno dei subtest della WISC-IV, otterrà un punteggio ponderato molto simile sia con le norme statunitensi che con quelle britanniche. Sorge allora spontanea la domanda: “che succederebbe se confrontassimo le norme britanniche o statunitensi con quelle italiane?”.

Nella taratura italiana della scala WISC-IV (Orsini, Pezzuti, & Picone, 2012) non sono stati effettuati, o almeno non sono stati riportati, confronti tra adattamenti diversi della scala. Non sappiamo quindi cosa succederebbe se confrontassimo norme UK con norme Italiane. Questo confronto, tuttavia, appare essere molto interessante perché permetterebbe di osservare se la scala funziona in modo simile in culture diverse.

Lo scopo di questo articolo è quello di confrontare l'adattamento italiano della scala WISC-IV (Orsini et al., 2012) con quello britannico (Wechsler, 2003). Pur essendo i dati UK e USA simili, abbiamo preferito quelli britannici, dato che il Regno Unito è un paese europeo con una cultura ed una popolosità più simili alla nostra. Il confronto è stato effettuato solo tra prove la cui procedura di somministrazione, gli stimoli, il materiale e le istruzioni fossero perfettamente sovrapponibili tra le due culture. Inoltre, abbiamo limitato il confronto ai subtest principali perché sono quelli che vengono comunemente utilizzati nella pratica clinica per la stima del quoziente intellettivo (QI).

Metodo

Procedura

Per il confronto abbiamo preso in considerazione i manuali della standardizzazione britannica (Wechsler, 2003) e italiana (Orsini et al., 2012) e, non avendo a disposizione i dati grezzi, ci siamo basati sulle tabelle pubblicate sui rispettivi manuali. In un primo momento abbiamo confrontato le medie dei punteggi ponderati tra l'adattamento UK e quello italiano, successivamente abbiamo diviso le differenze medie per le deviazioni standard Italiane. La deviazione standard, per ciascuna fascia d'età, è stata calcolata sulla distribuzione dei punteggi grezzi ricostruita tramite la relativa tabella di conversione grezzo-ponderato. Sono state generate un milione di misure per ciascuna fascia d'età ed è stata assunta la perfetta normalità delle distribuzioni dei punteggi ponderati. Laddove a un unico punteggio ponderato corrispondevano più punteggi grezzi, è stato considerato il punteggio grezzo mediano, ad eccezione dei ponderati di 1 e 19; in questi ultimi casi è stata applicata una correzione finale sulla deviazione standard per aggiungere la variabilità dovuta

alle code poste oltre le tre deviazioni standard dalla media nella distribuzione gaussiana. Le analisi sono state svolte con “R” (R Core Team, 2014).

Materiali

Sono stati presi in esame cinque subtest della WISC-IV: disegno con cubi, concetti illustrati, ragionamento con matrici, cifrario e ricerca di simboli. Per quanto riguarda gli indici abbiamo considerato QI, IRP, e IVE. Per il calcolo degli indici ci siamo basati sulle norme di conversione italiane ottenute dal manuale (Orsini et al., 2012; Orsini & Pezzuti, 2014).

Risultati

In tabella 1 sono presentati i risultati del confronto. Si noti che per il calcolo del QI, non essendo possibile il confronto tra le due popolazioni, abbiamo assunto che le due versioni della scala fossero identiche nelle prove verbali (quindi aggiungendo una costante di 50 per il calcolo del QI).

I punteggi riportati in Tab. 1 per i singoli subtest ci offrono una indicazione del punteggio ponderato che otterrebbe con le norme italiane un bambino che con le norme inglesi ottiene un punteggio di 10. Per essere più precisi abbiamo però operato sulle differenze di punteggio (i punteggi sono stati ottenuti sottraendo la media italiana da quella UK e dividendo il risultato per la deviazione standard italiana). A titolo di esempio, riportiamo il caso di un ipotetico bambino inglese di sei anni e un giorno che ottiene un punteggio grezzo di 11 punti nel subtest ragionamento con matrici. Il bambino in questione otterrebbe un punteggio ponderato di 10 con entrambe le tarature (UK e ITA). Tuttavia, solo avendo un punteggio di 11 il bambino UK otterrebbe un punteggio ponderato di 10, invece, con punteggi grezzi di 10 o 9 il bambino otterrebbe un punteggio ponderato di 9. Con la standardizzazione Italiana, infatti, punteggi grezzi di 10 o 9 corrispondono sempre ad un ponderato di 10 punti (ottenibile con punteggi da 9 a 11). È evidente quindi la presenza di una differenza tra i punteggi medi delle due tarature, e per calcolare questa differenza si è proceduto come segue.

- 1) Il punteggio centrale UK è stato confrontato con il punteggio centrale italiano (ossia 11 vs. 10) sottraendo i due punteggi ($11 - 10 = 1$).
- 2) La differenza di 1 è stata divisa per la deviazione standard dei punteggi ponderati italiani, facilmente inferibile dalle norme (vedi la procedura per maggiori informazioni), dividendo 1 per la DS (in questo caso 4.10), otteniamo un valore di 0.244. Questo valore è analogo ad un punto z (basato sulla DS della standardizzazione italiana).
- 3) A questo punto è facile trasformare il punto z (con media 0 e $DS = 1$) in un punteggio ponderato standardizzato (con media 10 e $DS = 3$) ($[0.244 \times 3] + 10 = 10.7$).

L'esempio sopra descritto mette in luce come con questa procedura sia più precisa rispetto ad un confronto "classico" (trasformando il punteggio grezzo medio di un ipotetico individuo UK e attribuendo il punteggio ponderato in base alle norme italiane). Si noti però che i risultati sono molto simili usando i due metodi: ad esempio nell'indice di Ragionamento Visuoperceptivo lo scarto tra i due metodi non eccede mai ± 2 punti, con uno scarto medio inferiore ad un punto standardizzato.

La tabella 1 dimostra come i subtest in cui le differenze sono più marcate siano quelli dell'indice di ragionamento visuoperceptivo ed in particolare nel ragionamento con matrici.

Inserire qui tabella 1

Dato che le differenze erano maggiori nelle prove di ragionamento visuoperceptivo (disegno con cubi, concetti per immagini e ragionamento con matrici) abbiamo svolto delle analisi ulteriori su di esse. In figura 1, in alto, vengono riportati i valori medi delle due popolazioni come riportati nei rispettivi manuali di standardizzazione; come si può vedere, e in analogia con i dati in tabella, emergono alcune differenze: i punteggi della popolazione UK sono infatti più alti di quelli italiani e questo dato emerge soprattutto in alcune fasce di età ed in alcuni subtest.

Per comprendere meglio la natura di queste differenze abbiamo svolto delle regressioni lineari aventi come variabili dipendenti le prove dell'indice di ragionamento visuoperceptivo e come predittori il gruppo (italiano e UK) e l'età. Per quanto riguarda il disegno con cubi, l'interazione

gruppo per età è significativa ($t = 3.36, p = .001$), così come l'effetto dell'età ($t = 34.79, p < .001$), mentre non emergono differenze significative tra gruppi ($t = -0.90, p = .369$). Il modello, nel suo complesso, è significativo $F(3, 62) = 949.5, p < .001, R^2 = .98$. Per quanto riguarda i concetti illustrati, invece, l'interazione gruppo per età non è significativa ($t = 1.47, p = .146$), così come l'effetto del gruppo ($t = -0.93, p = .927$), mentre è significativo l'effetto dell'età ($t = 18.87, p < .001$). Anche questo modello risulta, nel complesso, significativo $F(3, 62) = 273.1, p < .001, R^2 = .93$. Rispetto al subtest ragionamento con matrici, infine, l'interazione gruppo per età è significativa ($t = -2.02, p = .048$), così come l'effetto dell'età ($t = 22.80, p < .001$), e del gruppo ($t = 4.86, p < .001$). Anche il terzo modello, nel complesso, è significativo, $F(3, 62) = 344.7, p < .001, R^2 = .94$.

I valori di adattamento del modello lineare sembrano essere buoni (i valori di R^2 sono molto alti); tuttavia, la figura 1 dimostra come gli andamenti siano lineari soprattutto per le norme italiane mentre appaiono curvilinei (o quadratici) per quelle UK. Per questa ragione abbiamo svolto delle ulteriori regressioni aggiungendo come predittore l'andamento quadratico: il confronto tra i modelli lineari e quadratici, non riportato, è sempre significativo ($p < .001$), dimostrando un migliore adattamento di questi ultimi ai dati.

Nello specifico, per quanto riguarda il disegno con cubi, il modello è significativo, $F(5, 60) = 1117, p < .001, R^2 = .99$ e tutte le interazioni e gli effetti principali risultano significativi ($ps \leq .004$). Un discorso analogo vale per i concetti illustrati, $F(5, 60) = 432.5, p < .001, R^2 = .97$, dove tutte le interazioni e gli effetti principali sono significativi ($ps \leq .036$) e per ragionamento con matrici, $F(5, 60) = 1109, p < .001, R^2 = .99$ ($ps \leq .001$)

Si noti che, anche se i cambiamenti in termini di R^2 tra i modelli lineari e quadratici sono modesti, le curve sono tuttavia riprodotte con un'accuratezza quasi perfetta nel modello quadratico, con indici che si approssimano a 1 (Figura 1 in basso).

Inserire qui figura 1

Nelle analisi svolte su cifrario e ricerca di simboli, non riportate, le differenze tra i due gruppi, seppur statisticamente significative, sono modeste.

Conclusioni

L'obiettivo del presente studio era quello di confrontare la standardizzazione italiana con quella del Regno Unito. I risultati dimostrano come vi siano punti di contatto e differenze tra le due standardizzazioni: in particolare un individuo inglese con una performance in norma, se confrontato con la standardizzazione del Regno Unito, ottiene risultati spesso diversi se il confronto viene effettuato con norme italiane.

Se ci riferiamo ai criteri di Cohen per la stima della dimensione degli effetti, notiamo delle particolarità (Cohen, 1988): nello specifico, rispetto ai subtest di velocità di elaborazione (cifrario e ricerca di simboli) le differenze sono molto piccole e inferiori ad un punto; ciò significa che se un bambino inglese viene confrontato con uno italiano ha in media una performance sovrapponibile, con scarti minimi e del tutto trascurabili che vanno in entrambe le direzioni (a volte maggiori e a volte minori). La situazione cambia però se consideriamo gli indici di ragionamento visuoperceptivo: in questo caso, le differenze sono piccole (di circa un punto) se consideriamo disegno con cubi e concetti illustrati mentre risultano maggiori nel caso di ragionamento con le matrici (circa due punti). Queste differenze diventano importanti per il calcolo dell'indice di ragionamento visuoperceptivo: con le norme italiane infatti, un bambino che in Inghilterra avrebbe 100 in Italia otterrebbe punteggi decisamente più alti (in media di 8 punti); se consideriamo poi la stima del quoziente intellettivo totale, le differenze tra le due popolazioni sono piccole ma apprezzabili (in media 3 punti).

Non di meno, il dato più importante sembra essere legato alle fasce d'età in cui le differenze sono più marcate: i valori, infatti, tendono ad essere simili tra le due standardizzazioni nei bambini più piccoli e nei ragazzi più grandi, ma a divergere nelle fasce centrali (tra gli 11 e i 14 anni). Se consideriamo tali fasce, i risultati sono dunque più evidenti: un individuo con un punteggio in media in Inghilterra, confrontato con norme italiane avrebbe un indice di ragionamento percettivo compreso tra 111 e 117 (tra $\frac{2}{3}$ e 1 deviazioni standard) e di QI tra 105 e 106 ($\frac{1}{3}$ di deviazione standard).

Nel complesso, se consideriamo solo il QI, i risultati, in termini di dimensione dell'effetto, sono modesti: un bambino inglese con un QI nella media otterrebbe, infatti, un punteggio leggermente più alto (circa 3 punti) se utilizzassimo norme italiane. I clinici dovrebbero sempre fare riferimento agli intervalli di confidenza che al loro interno includono la possibilità di variazioni dovute ad errori di misura. Tuttavia -a titolo di esempio e assumendo per semplicità un intervallo di 5 punti al 95% ed una sovrastima di 3 punti- se un ragazzo ha 81 [76 - 86] alla WISC-IV, ma questo punteggio è sovrastimato di tre punti, allora il suo QI "reale" sarà 78 [75 - 83] con le implicazioni cliniche che questo comporta. Allo stesso modo se un ragazzo ha un punteggio di 67 [62- 72], quindi non è possibile escludere al 95% che il punteggio includa 70, il suo punteggio "reale" sarà di 64 [59 - 69], punteggio che non include la soglia critica di 70. Piccole differenze hanno ovviamente ripercussioni sull'intervallo che in un caso include e nell'altro esclude la soglia critica. Tali differenze, infatti, assumono un valore tutt'altro che trascurabile in termini di "significatività clinica" (Kazdin, 1999). Infatti, 3 punti possono fare rientrare o meno un bambino nella fascia di disabilità intellettiva o possono influire su una diagnosi di disturbo dell'apprendimento (DSA; Istituto superiore di sanità, 2011).

A questo punto occorre domandarsi il perché di queste differenze: nella quarta edizione della WISC sono stati aggiunti dei nuovi subtest, tra i quali ragionamento con le matrici, in cui la differenza tra le norme UK e quelle italiane è marcata. Questo dato evidenzia che una differenza esiste ma non spiega il perché di questo scarto. Una possibilità è che i bambini inglesi abbiano punteggi più alti nei test d'intelligenza rispetto a quelli italiani (ossia siano più intelligenti); questo risultato, se vero, confermerebbe l'ipotesi secondo cui gli individui delle regioni del meridione italiano hanno un QI inferiore rispetto alla media nazionale (si veda per una revisione Cornoldi, Giofrè, & Martini, 2013), essendo il campione di standardizzazione della WISC-IV rappresentativo di tali regioni (principalmente Lazio e Campania).

I risultati delle curve di crescita, però, smentiscono questo dato: i gruppi in realtà risultano piuttosto simili se confrontati in determinate fasce d'età e le differenze si riscontrano soprattutto

nelle curve ipotetiche di crescita; le norme UK e quelle italiane sono state probabilmente ottenute tramite interpolazione lineare dei dati ma se i parametri utilizzati per il calcolo delle norme cambiano, allora le norme saranno necessariamente differenti. Infatti, l'andamento italiano sembra si adatti meglio ad una funzione lineare mentre l'andamento UK sembra più coerente con un andamento quadratico (ossia curvilineo). Questo significa che, in determinate fasce, precisamente quelle centrali, le performance degli italiani e dei britannici tendono ad allontanarsi.

Va notato che le norme italiane sono basate su un ampio campione in cui si è tenuto conto di variabili importanti come età, genere, istruzione e professione dei genitori. Nelle standardizzazioni USA e UK, nonostante gli autori abbiano per esempio controllato per regione geografica e razza, questa attenzione non vi è stata. Da ciò si potrebbe concludere che i tre campioni non siano direttamente confrontabili. Questo potrebbe, inoltre, portare a pensare che i diversi andamenti siano attribuibili a problemi legati alla standardizzazione USA e UK. Tuttavia le norme USA e UK sono molto simili tra di loro. Inoltre, gli andamenti evolutivi sono tipicamente curvilinei in prove di intelligenza fluida come le matrici di Raven (ad es. Fry & Hale, 2000). L'andamento curvilineo è infatti particolarmente evidente nei dati USA e UK, ma poco pronunciato in quelli Italiani e questo è difficilmente spiegabile da fattori culturali. D'altro canto, le differenze potrebbero non essere dovute ad errori di campionamento o ad assunzioni scorrette nel calcolo delle norme ma riflettere reali differenze tra queste popolazioni. Al momento, però, non è di fatto possibile stabilire con esattezza la causa di queste differenze.

Il presente studio non è tuttavia esente da limiti. Il limite principale è il mancato utilizzo dei dati grezzi, poiché non accessibili; di conseguenza, gli autori dell'articolo, sono stati costretti ad utilizzare le tabelle di conversione, derivando in alcuni casi i valori di riferimento. Questa scelta ha potuto in qualche modo influire sulla precisione delle stime ottenute che, seppur ragionevoli, possono presentare qualche imprecisione. Va notato inoltre che misure di intelligenza fluida misurano anche altri costrutti e non sono esenti da errori di misura (Jensen, 1998). Inoltre, possono esserci ampie differenze culturali anche in prove considerate culture-free o d'intelligenza

fluida (e.g., Flynn, 1987; Rosselli & Ardila, 2003; Rushton & Jensen, 2005; Valseschini & Del Ton, 1981). Anche se il Regno Unito e l'Italia fanno parte dei paesi cosiddetti occidentali, non possiamo escludere che le differenze riscontrate siano attribuibili a fattori culturali come l'uso di strategie diverse. Per questa ragione gli adattamenti dei test sono estremamente utili e rimangono fondamentali in questo campo. Infine, nel presente lavoro ci siamo limitati ad analizzare cinque dei dieci subtest principali. Riteniamo che l'analisi degli altri subtest, seppur rischiosa, possa portare a risultati interessanti ed essere oggetto di ricerche future.

Di recente è apparsa nei paesi anglofoni la nuova WISC-V (Wechsler, 2014); l'introduzione di questa nuova batteria sarebbe auspicabile anche in Italia poiché la sua standardizzazione renderebbe possibile il confronto dei dati delle due edizioni della batteria. Se le differenze si dovessero manifestare anche nella WISC-V, dovremmo assumere che queste riflettano discrepanze reali tra le due popolazioni; in caso contrario, il confronto tra le curve di sviluppo di diversi Paesi, permetterebbe di armonizzare i dati italiani con quelli internazionali. Al momento, infatti, le ragioni delle differenze messe in evidenza nel presente articolo non sono chiare e dovranno essere approfondite in ricerche future .

Riferimenti

- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Cornoldi, C., Giofrè, D., & Martini, A. (2013). Problems in deriving Italian regional differences in intelligence from 2009 PISA data. *Intelligence, 41*, 25–33. doi:10.1016/j.intell.2012.10.004
- Evers, A., Muñiz, J., Bartram, D., Boben, D., Egeland, J., Fernández-Hermida, J. R., ... Urbánek, T. (2012). Testing practices in the 21st century. *European Psychologist, 17*, 300–319. doi:10.1027/1016-9040/a000102
- Flynn, J. R. (1987). Massive IQ gains in 14 nations: What IQ tests really measure. *Psychological Bulletin, 101*, 171–191. doi:10.1037/0033-2909.101.2.171
- Fry, A. F., & Hale, S. (2000). Relationships among processing speed, working memory, and fluid intelligence in children. *Biological Psychology, 54*, 1–34. doi:10.1016/S0301-0511(00)00051-X
- Istituto superiore di sanità. (2011). *Consensus conference on Learning Disabilities*.
- Jensen, A. R. (1998). *The g factor: The science of mental ability*. Westport, CT: Praeger.
- Kazdin, A. E. (1999). The meanings and measurement of clinical significance. *Journal of Consulting and Clinical Psychology, 67*, 332–339. doi:10.1037/0022-006X.67.3.332
- Orsini, A., & Pezzuti, L. (2014). L' indice di abilità generale della scala WISC-IV [The WISC-IV General Ability Index]. *Psicologia Clinica Dello Sviluppo, 18*, 301–310. doi:10.1449/77640
- Orsini, A., Pezzuti, L., & Picone, L. (2012). *WISC-IV: Contributo alla taratura Italiana. [WISC-IV Italian Edition]*. Florence, Italy: Giunti O. S.
- R Core Team. (2014). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.
- Rosselli, M., & Ardila, A. (2003). The impact of culture and education on non-verbal neuropsychological measurements: A critical review. *Brain and Cognition, 52*, 326–333. doi:10.1016/S0278-2626(03)00170-2
- Rushton, J. P., & Jensen, A. R. (2005). Thirty years of research on race differences in cognitive ability. *Psychology, Public Policy, and Law, 11*, 235–294. doi:10.1037/1076-8971.11.2.235
- Valseschini, S., & Del Ton, F. (1981). *Le matrici progressive di Raven. Contributo alla taratura su 1123 soggetti e considerazioni sulla validità e attendibilità della prova*. Firenze, Italy: Organizzazioni Speciali.
- Wechsler, D. (2003). *Wechsler intelligence scale for children—Fourth UK Edition (WISC-IV) administration and scoring manual*. San Antonio, TX: Pearson Assessment.
- Wechsler, D. (2014). *Wechsler Intelligence Scale for Children—Fifth Edition*. San Antonio, TX: NCS Pearson.
- Zachary, R. A., & Gorsuch, R. L. (1985). Continuous norming: Implications for the WAIS-R. *Journal of Clinical Psychology, 41*, 86–94. doi:10.1002/1097-4679(198501)41:1<86::AID-JCLP2270410115>3.0.CO;2-W

Tabella 1. Differenze tra la standardizzazione britannica e quella italiana in ciascuna delle fasce d'età: punteggio ponderato e punteggio di QI indicativo che otterrebbe con le norme italiane un bambino che con le norme inglesi ottiene un punteggio di 10 a un subtest e 100 negli indici (si veda il testo per la descrizione della procedura; gli indici sono stati ottenuti sommando i punteggi ponderati e convertendo il risultato con le tabelle di conversione italiane).

Età ⁺	Subtest					Indici				Indici Complessivi	
	DC	CI	RM	CR	RS	IRPpp	IRP	IVEpp	IVE	QIpp	QI
6.1	9.6	9.5	10.7	9.6	9.8	30	100	19	97	99	99
6.2	9.8	10.9	12.2	9.9	10.2	33	106	20	100	103	102
6.3	10.8	10.9	12.5	9.7	10.7	34	108	20	100	105	104
7.1	10.2	10.0	12.5	9.3	10.2	33	106	20	100	102	102
7.2	10.2	9.6	12.4	9.3	10.7	32	104	20	100	102	102
7.3	10.7	10.9	12.4	9.0	9.8	34	108	19	97	103	102
8.1	10.2	10.9	12.4	9.1	8.3	33	106	17	91	101	101
8.2	9.8	10.9	11.6	9.4	9.1	32	104	19	97	101	101
8.3	10.2	10.0	12.3	9.8	9.1	32	104	19	97	101	101
9.1	9.8	10.9	12.0	9.6	9.0	33	106	19	97	101	101
9.2	10.3	10.0	12.8	10.0	9.5	33	106	19	97	103	102
9.3	11.4	10.4	12.6	9.6	8.9	34	108	19	97	103	102
10.1	11.2	10.9	12.1	9.3	9.0	34	108	18	94	102	102
10.2	11.0	11.4	13.4	9.2	8.5	36	113	18	91	103	102
10.3	10.7	11.4	13.4	9.7	9.2	35	111	19	97	104	103
11.1	11.9	11.4	12.5	10.0	9.5	36	113	19	97	105	104
11.2	11.5	11.4	13.1	10.3	10.3	36	113	21	103	107	106
11.3	12.5	11.4	13.2	10.2	10.0	37	115	20	100	107	106
12.1	12.2	11.8	13.5	10.4	9.8	38	117	20	100	108	106
12.2	12.0	11.8	12.9	10.1	9.5	37	115	20	100	106	105
12.3	11.8	11.8	12.9	10.3	9.1	37	115	19	97	106	105
13.1	12.5	11.0	12.9	10.7	10.2	36	113	21	103	107	106
13.2	12.4	11.0	12.9	11.2	9.8	36	113	21	103	107	106
13.3	12.5	11.0	12.2	10.7	10.0	36	113	21	103	106	105
14.1	12.0	11.0	12.6	10.5	10.0	36	113	21	103	106	105
14.2	11.5	12.0	11.8	10.1	10.0	35	111	20	100	105	104
14.3	11.8	12.2	11.8	10.4	9.6	36	113	20	100	106	105
15.1	11.4	11.0	11.1	10.6	9.6	33	106	20	100	104	103
15.2	11.1	11.1	11.1	10.9	9.5	33	106	20	100	104	103
15.3	11.0	11.0	10.7	10.9	9.3	33	106	20	100	103	102
16.1	10.6	11.0	10.7	11.4	10.5	32	104	22	106	104	103
16.2	10.0	11.0	10.3	11.5	10.0	31	102	22	106	103	102
16.3	9.4	11.0	9.6	11.3	9.5	30	100	21	103	101	101
M	11.02	10.99	12.15	10.13	9.64	34.16	108.67	19.77	99.27	103.93	103.12

Note. DC = disegno con cubi, CI = concetti illustrati, RM = ragionamento con matrici, CR = cifrario, RS = ricerca di simboli, IRP = indice di ragionamento visuoperceptivo; IVE = indice di velocità di elaborazione, pp = somma punteggi ponderati.

⁺ .1 = 0 mesi e 0 giorni a 3 mesi e 30 giorni; .2 = 4 mesi e 0 giorni a 7 mesi e 30 giorni; .3 = 8 mesi e 0 giorni a 11 mesi e 30 giorni.

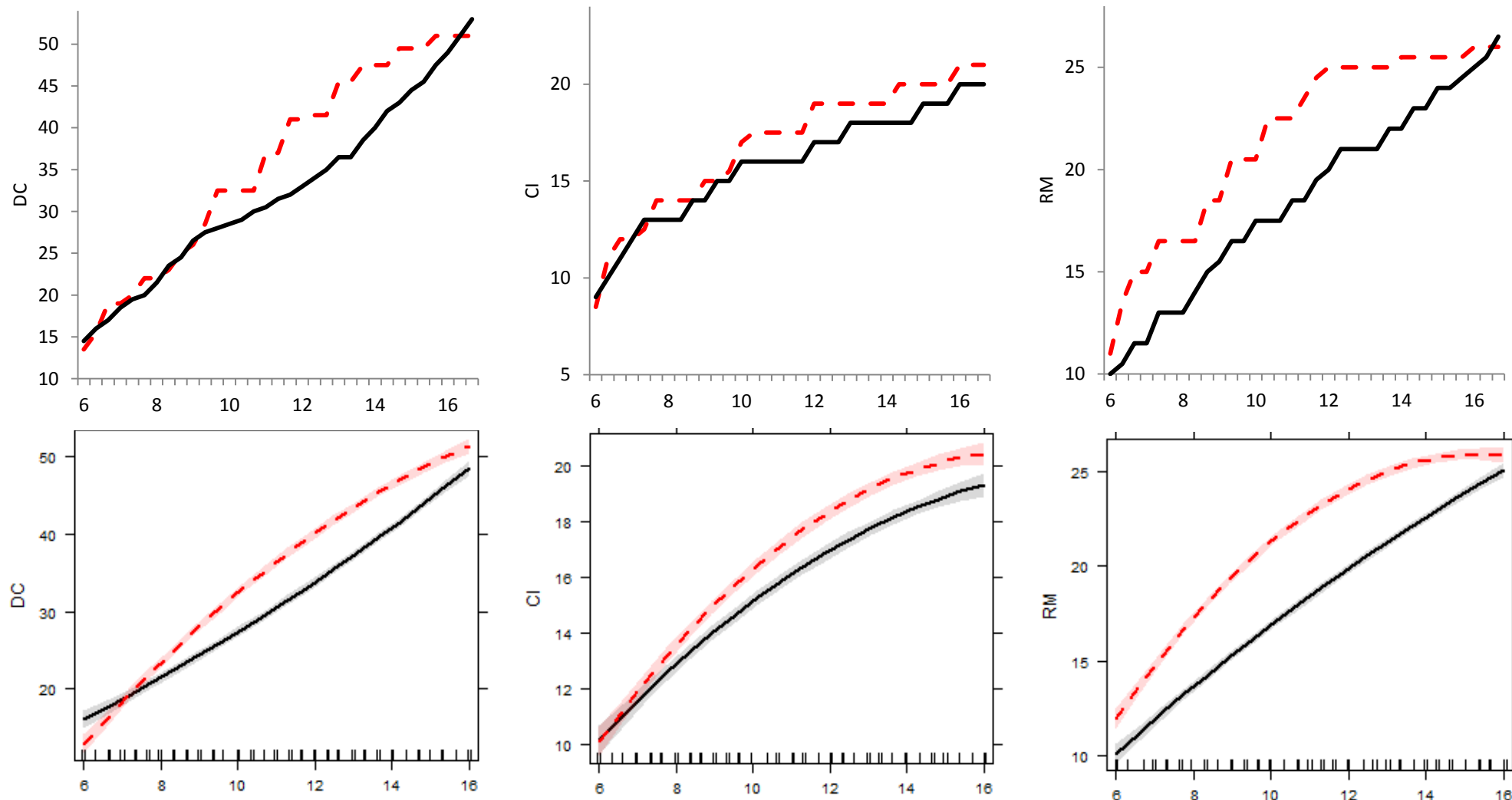


Figura 2. Progressione dei punteggi grezzi (corrispondenti ad un punto ponderato di 10) nei vari subtest in base all'età. Dati reali (in alto) e curve predette dal modello quadratico (in basso). Norme UK, linea tratteggiata, e ITA, linea continua. Le barre d'errore rappresentano intervalli di confidenza al 95%. DC = disegno con cubi; CI = concetti per immagini; RM = ragionamento con matrici.