

Estensione della Global Scale of English (GSE) alla Global Scale of Languages (GSL)

Parte 1: Allineamento degli obiettivi di
apprendimento dell'italiano alla GSL

Settembre 2023

Ying Zheng, University of Southampton

Catherine Doyle, Pearson

David Booth, Pearson

Mike Mayor, Pearson

Indice

Sintesi	4
1. Introduzione alla GSE e agli obiettivi di apprendimento della GSE	5
2. Obiettivo dello studio	5
3. Metodologia	6
3.1 Giudizio comparativo e sue applicazioni	6
3.2 Progetto di studio	7
3.2.1 Traduzione degli obiettivi di apprendimento: dall'inglese all'italiano	8
3.2.2 Selezione del valutatore	8
3.2.3 Descrizione del set di dati	9
4. Risultati	9
4.1 Statistiche di non adattamento del valutatore.....	9
4.2 Statistiche di non adattamento degli obiettivi di apprendimento	11
4.2.1 Ascolto	11
4.2.2 Lettura.....	13
4.2.3 Conversazione	15
4.2.4 Scrittura.....	17
4.3 Correlazioni tra le stime della GSE e quelle di giudizio comparativo	19
4.4 Equazione di trasformazione per ogni competenza	20
5. Discussione e conclusioni	21
Glossario	22
Riferimenti	22
Appendice: Dati demografici del valutatore	24

Sintesi

La Global Scale of English (GSE) offre uno strumento più dettagliato per descrivere e valutare i progressi e le prestazioni degli studenti di lingua inglese. Pearson ha condotto una ricerca approfondita (**Pearson**) sull'utilizzo degli obiettivi di apprendimento della GSE come scala di riferimento per estendere la serie di indicatori di capacità del Quadro comune europeo di riferimento (QCER) del 2001 al fine di rispondere alle esigenze di un maggior numero di studenti.

Questo studio ha confrontato l'ordine gerarchico degli obiettivi di apprendimento della GSE tradotti in lingua italiana per stabilire se i valori esistenti della GSE sono applicabili agli studenti adulti di italiano come lingua straniera (in inglese IFL). Sono stati tradotti in italiano 320 obiettivi di apprendimento. Un gruppo di 20 valutatori qualificati, scelti da un pool di insegnanti IFL, è stato invitato a condurre 25 confronti di giudizio comparativo per ogni obiettivo di apprendimento, ottenendo così 16.000 valori. È stata condotta una serie di analisi, tra cui le statistiche del valutatore e le statistiche di adattamento degli elementi, per valutare la difficoltà degli obiettivi di apprendimento dell'inglese esistenti nei logit del modello di Rasch e confrontarli con le stime di giudizio comparativo delle versioni inglese e italiana per quattro abilità linguistiche. Da questi confronti sono state ricavate equazioni di trasformazione per allineare i risultati degli obiettivi di apprendimento dell'italiano con l'attuale GSE, portando infine alla creazione di una nuova Global Scale of Italian.

Per ulteriori informazioni sulla Global Scale of Languages, visita [pearson.com/languages](https://www.pearson.com/languages).

1. Introduzione alla GSE e agli obiettivi di apprendimento della GSE

La GSE è una scala di competenza standardizzata da 10 a 90 che misura la conoscenza della lingua inglese ed è psicometricamente allineata al Quadro comune europeo di riferimento per le lingue (QCER, Consiglio d'Europa, 2001). È stata sviluppata una serie di obiettivi di apprendimento della GSE per descrivere le competenze degli studenti in ogni punto della scala, incorporando ed estendendo l'insieme dei descrittori del QCER. Questi obiettivi di apprendimento sono stati valutati dagli insegnanti di inglese come lingua straniera (in inglese EFL) e calibrati rispetto alla Global Scale of English (de Jong, Mayor e Hayes, 2016). A differenza del QCER e di altre scale che descrivono i risultati raggiunti a livelli più generalizzati, la Global Scale of English individua ciò che uno studente è in grado di fare in ogni punto della scala per quanto riguarda le abilità di conversazione, ascolto, lettura e scrittura, per fornire una descrizione più dettagliata della competenza linguistica in crescita. Il lavoro di sviluppo degli obiettivi di apprendimento della GSE si basa sulla ricerca condotta da Brian North e dal Consiglio d'Europa per la creazione del QCER (North, 2000) e la estende. Gli obiettivi di apprendimento della GSE sono stati sviluppati da Pearson English nel corso di diversi anni in collaborazione con oltre 6.000 insegnanti, autori di materiali didattici per la lingua inglese (in inglese ELT) ed esperti linguistici di tutto il mondo.

2. Obiettivo dello studio

L'obiettivo di questo studio è quello di confrontare l'ordine gerarchico degli obiettivi di apprendimento della GSE tradotti in italiano per vedere se i valori esistenti della GSE sono applicabili agli studenti adulti di italiano come lingua straniera, cioè se possono essere messi sulla stessa scala. L'ipotesi alla base del progetto è la seguente: dato che la GSE si basa sul QCER, che è di per sé linguisticamente neutrale, si ritiene che l'ordine complessivo sarà altamente correlato sia alla GSE che al QCER. Questo progetto si propone di verificare tale ipotesi utilizzando l'approccio del giudizio comparativo.

3. Metodologia

3.1 Giudizio comparativo e sue applicazioni

Il giudizio comparativo (CJ) prevede il giudizio olistico di coppie di lavori di studenti da parte di un gruppo di valutatori indipendenti che determinano quale dei due lavori presenta il costrutto globale meglio specificato. Il risultato è una matrice decisionale binaria del “vincitore” e del “perdente” di ogni coppia, che viene poi adattata al modello Bradley-Terry (Bradley e Terry, 1952) per produrre i valori dei parametri (punteggi) e gli errori standard per ogni lavoro degli studenti. Il valore dei parametri consente di costruire un ordine gerarchico dei lavori degli studenti, dal “migliore” al “peggiore”, utilizzabile per scopi di valutazione come la classificazione.

Oltre all'uso nelle commissioni d'esame britanniche per esaminare la comparabilità tra le schede (ad esempio, Fearnley, 2000; Gray, 2000), la comparabilità degli standard nel tempo e il mantenimento degli standard (ad esempio, Chambers e Cunningham, 2022), il giudizio comparativo è stato applicato anche a diversi contesti educativi. Ciò include la valutazione tra pari di relazioni di progetti di Design Thinking universitari (Mentzer et al., 2021), prove scritte sulla comprensione concettuale di un corso di matematica (Jones e Alcock, 2014), valutazioni degli insegnanti su statistica sommativa e prove di inglese (Marshall et al., 2020), saggi (Steedle e Ferrara, 2016) e testi argomentativi (Lesterhuis et al., 2022). Pearson ha usato il giudizio comparativo per allineare gli obiettivi di apprendimento della Global Scale of English (GSE) per i giovani studenti alla Chinese Scale of English proficiency (CSE), confrontando la difficoltà dei descrittori in ogni standard (Pearson, 2020).

La base psicologica del giudizio comparativo è che gli esseri umani sono abili nel confrontare un oggetto con un altro, ma inaffidabili quando valutano gli oggetti singolarmente (Gill e Bramley, 2013; Thurstone, 1927). Gli approcci analitici tradizionali prevedono che gli insegnanti valutino il lavoro degli studenti individualmente in modo assoluto utilizzando delle rubriche, il che può portare a diverse interpretazioni e applicazioni dei descrittori di tali rubriche, nonché alla possibilità di attingere alla loro percezione del lavoro degli altri studenti. Al contrario, il giudizio comparativo minimizza l'influenza comparativa delle rubriche dettagliate e specifiche (Pollitt, 2004) e sfrutta direttamente l'aspetto comparativo della valutazione, rinunciando alle rubriche e alle correzioni. La letteratura precedente ha illustrato come il giudizio comparativo soddisfi elevati standard di validità, affidabilità ed efficienza.

3.2 Progetto di studio

Per condurre questo studio è stato utilizzato *NoMoreMarking* (Wheadon, 2019), uno strumento di giudizio comparativo. Il numero di volte in cui un determinato oggetto viene giudicato rispetto a un altro è un elemento importante in uno studio di giudizio comparativo. Verhavert, et al, (2019) consigliano di avere da 10 a 30 confronti per oggetto per garantire un'affidabilità accettabile. In linea con questa raccomandazione, sono stati raccolti 25 confronti per obiettivo di apprendimento per garantire la validità del progetto.

In questo studio abbiamo selezionato 320 obiettivi di apprendimento della GSE per adulti, che rappresentano il 30% del numero totale disponibile. In termini di dimensioni e selezione del campione, il 20% è generalmente la sovrapposizione minima necessaria per allineare le scale (Kolen e Brennan, 2004). Il campione è strutturato in modo da essere rappresentativo sia del numero di obiettivi di apprendimento in ciascuna delle quattro competenze, sia del numero in ciascun livello del QCER (vedi Tabella 1).

Tabella 1: Distribuzione degli obiettivi di apprendimento

QCER/GSE	Ascolto	Letture	Conversa- zione	Scrittura	TOTALE	% del database
Inferiore a A1 (10-21)	3	3	10	4	20	34%
A1 (22-29)	5	5	14	8	32	27%
A2 (30-35)	6	6	17	10	39	30%
A2+ (36-42)	6	6	16	10	38	27%
B1 (43-50)	7	7	18	11	43	35%
B1+ (51-58)	7	7	18	11	43	33%
B2 (59-66)	7	7	19	11	44	28%
B2+ (67-75)	5	5	14	8	32	27%
C1 (76-84)	3	3	10	4	20	28%
C2 (85-90)	1	1	4	3	9	47%
TOTALE	50	50	140	80	320	30%
% del database	26%	35%	28%	33%	30%	

Si è tenuto conto anche della diversità e dell'ampiezza delle funzioni linguistiche e si è evitato di selezionare obiettivi di apprendimento piuttosto simili. Sono stati selezionati anche alcuni descrittori originali del QCER (che sono anche inclusi nella GSE) per fornire correlazioni statistiche con il modello QCER/North.

3.2.1 Traduzione degli obiettivi di apprendimento: dall'inglese all'italiano

I 320 obiettivi di apprendimento della GSE sono stati tradotti in italiano da un'agenzia di traduzione. All'agenzia sono state fornite come linea guida le traduzioni ufficiali dall'inglese all'italiano del QCER redatte dal Consiglio d'Europa. Dopo le traduzioni iniziali, è stato chiesto a due membri del team editoriale di Pearson in Italia di controllare a campione il lavoro. Sono stati individuati e corretti alcuni problemi minori nella traduzione prima di utilizzare la versione finale degli obiettivi di apprendimento in italiano.

Per garantire che le versioni in italiano e in inglese fossero valutate con lo stesso quadro di riferimento, le versioni in italiano e in inglese dello stesso obiettivo di apprendimento sono state inserite in un unico gruppo di valutazione (diviso per competenze), in modo che i valutatori vedessero due versioni in italiano, due in inglese o una per ciascuna lingua.

3.2.2 Selezione del valutatore

I valutatori sono stati reclutati da un gruppo di insegnanti di lingua italiana come lingua straniera che erano o erano stati impiegati come esaminatori della GCSE e/o di italiano di livello A (qualifiche di scuola secondaria/università nel Regno Unito) dalla commissione d'esame Pearson Edexcel. 138 persone hanno espresso interesse a partecipare alla ricerca e hanno fornito alcune informazioni di base. In base alla loro esperienza nell'insegnamento a studenti adulti e alla loro familiarità con il QCER, sono stati selezionati 20 valutatori per il progetto. Si è tenuto conto anche della necessità di creare un gruppo di valutatori il più eterogeneo possibile in termini di genere, nazionalità ed esperienza. Inoltre, tutti i valutatori avevano esperienza nell'insegnamento di almeno un'altra lingua oltre all'italiano (si veda l'appendice per i dati demografici dei valutatori).

Ai valutatori sono state fornite istruzioni scritte sul compito e sulla piattaforma prima di chiedere loro di effettuare il giudizio comparativo basato su questa domanda: "Quale di questi obiettivi di apprendimento descrive la competenza più difficile per uno studente di lingua?"

3.2.3 Descrizione del set di dati

Tabella 2: Numero di obiettivi di apprendimento e confronti per ciascuna competenza

Competenza	Obiettivi di apprendimento dell'inglese	Obiettivi di apprendimento dell'italiano	Numero totale di giudizi
Ascolto	50	50	2500
Lettura	50	50	2500
Conversazione	140	140	7000
Scrittura	80	80	4000
TOTALE	320	320	16000

4. Risultati

Nel giudizio comparativo, la scala di affidabilità della separazione (in inglese SSR) viene utilizzata come indicatore di affidabilità, in questo caso, l'affidabilità dell'ordine gerarchico degli obiettivi di apprendimento ottenuti con il metodo di giudizio comparativo. La scala di affidabilità della separazione si esprime da 0 a 1, con valori superiori a 0,90 che indicano una scala di giudizio comparativo altamente affidabile. La Tabella 3 mostra la scala di affidabilità della separazione per tutte e quattro le competenze.

Tabella 3: Scala di affidabilità della separazione

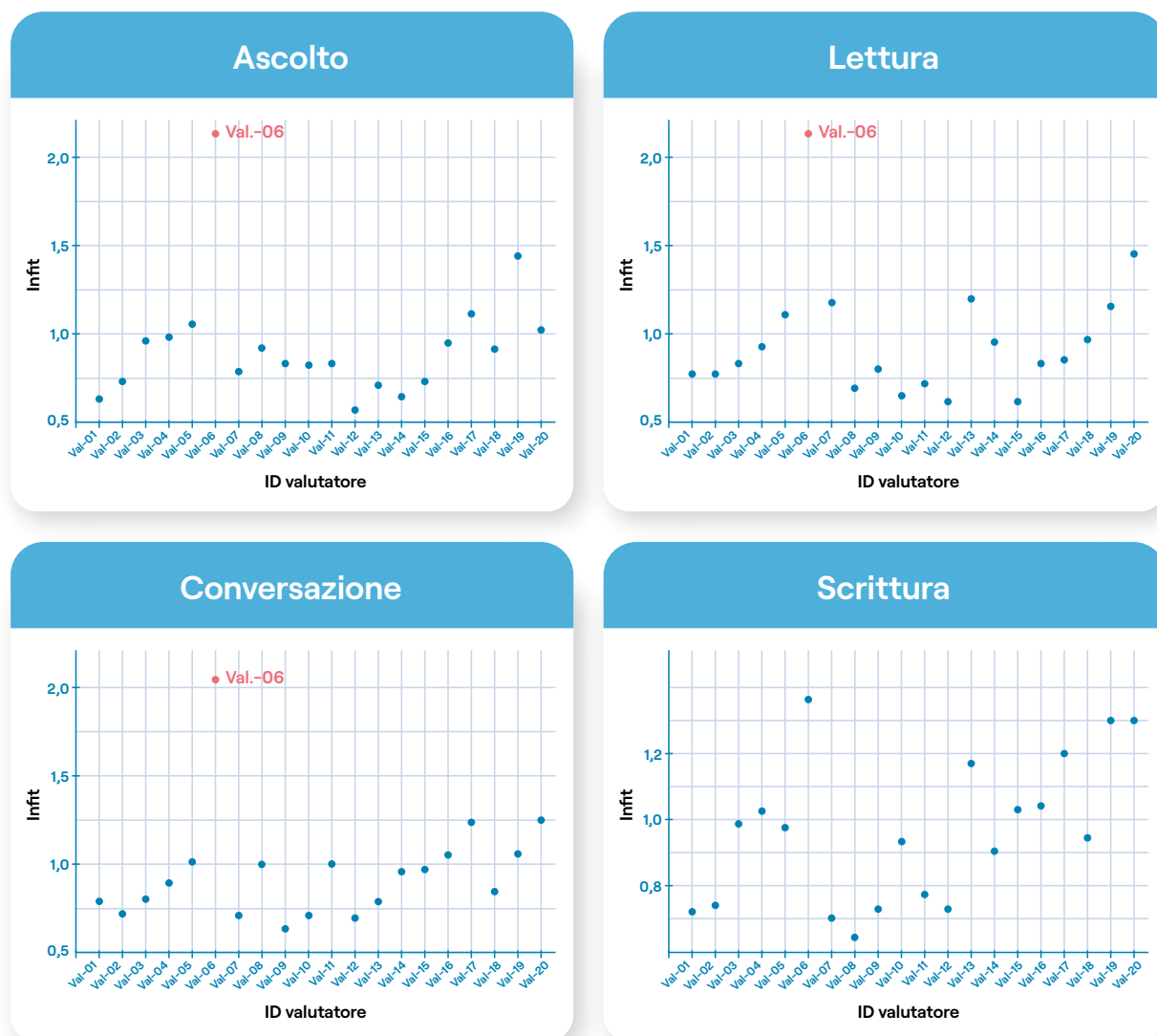
Ascolto	Lettura	Conversazione	Scrittura
0,939	0,938	0,938	0,943

4.1 Statistiche di non adattamento del valutatore

Le statistiche di adattamento sono state calcolate sia per i valutatori che per gli elementi (ad esempio, gli obiettivi di apprendimento) utilizzati in questo esercizio di giudizio comparativo. I valutatori con un infit superiore a due deviazioni standard al di sopra dell'infit medio sono stati esclusi, poiché ciò indica che potrebbero aver giudicato in modo incoerente o non allineato con l'opinione comune degli altri valutatori. 19 su 20 (95%) dei valutatori presentavano statistiche di non adattamento accettabili per tutti e quattro gli esercizi. Un valutatore ha mostrato un disadattamento nell'ascolto, nella lettura e nella conversazione (forse a causa di un'errata interpretazione del compito). Sebbene le statistiche di non adattamento di questo particolare valutatore per la parte di scrittura rientrassero nell'intervallo accettabile, le sue statistiche di disadattamento di valutazione per questo compito erano comunque le più alte tra tutti i valutatori.

Pertanto, i suoi dati di valutazione su tutti e quattro i compiti sono stati esclusi da ulteriori analisi. Rimuovendo i dati di questo valutatore, la correlazione complessiva tra i valori della GSE esistenti e i punteggi di giudizio comparativo è migliorata da 0,79 a 0,93. Vedi la Figura 1 per una rappresentazione visiva delle statistiche di non adattamento del valutatore.

Figura 1: Statistiche di non adattamento del valutatore (quattro competenze)



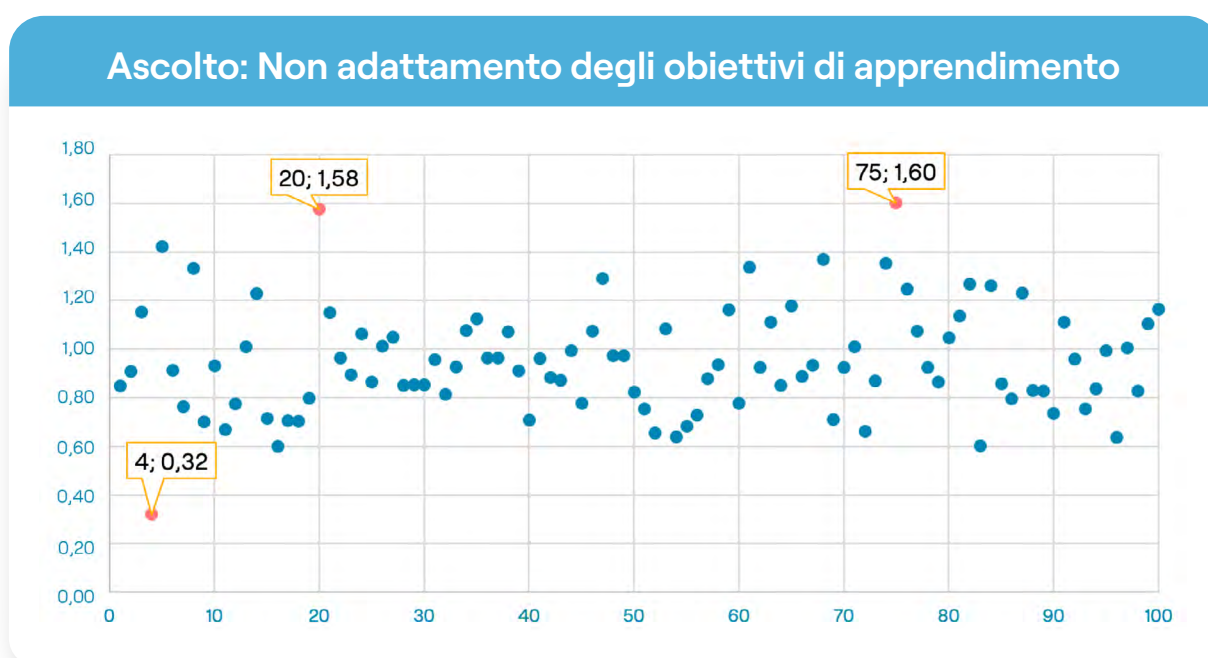
4.2 Statistiche di non adattamento degli obiettivi di apprendimento

Le seguenti sezioni riportano le statistiche di non adattamento degli obiettivi di apprendimento per le quattro competenze. Le Figure 2, 4, 6 e 8 mostrano i grafici a dispersione per ciascuna competenza, con l'asse Y che indica le statistiche di non adattamento degli elementi e l'asse X che indica il numero di elementi. Ci sono 100 elementi per l'ascolto, 100 elementi per la lettura, 160 elementi per la scrittura e 260 elementi per la conversazione.

4.2.1 Ascolto

La correlazione complessiva tra i valori della GSE esistenti degli obiettivi di apprendimento relativi all'ascolto (sia in inglese che in italiano) e i valori di giudizio comparativo è **0,931**.

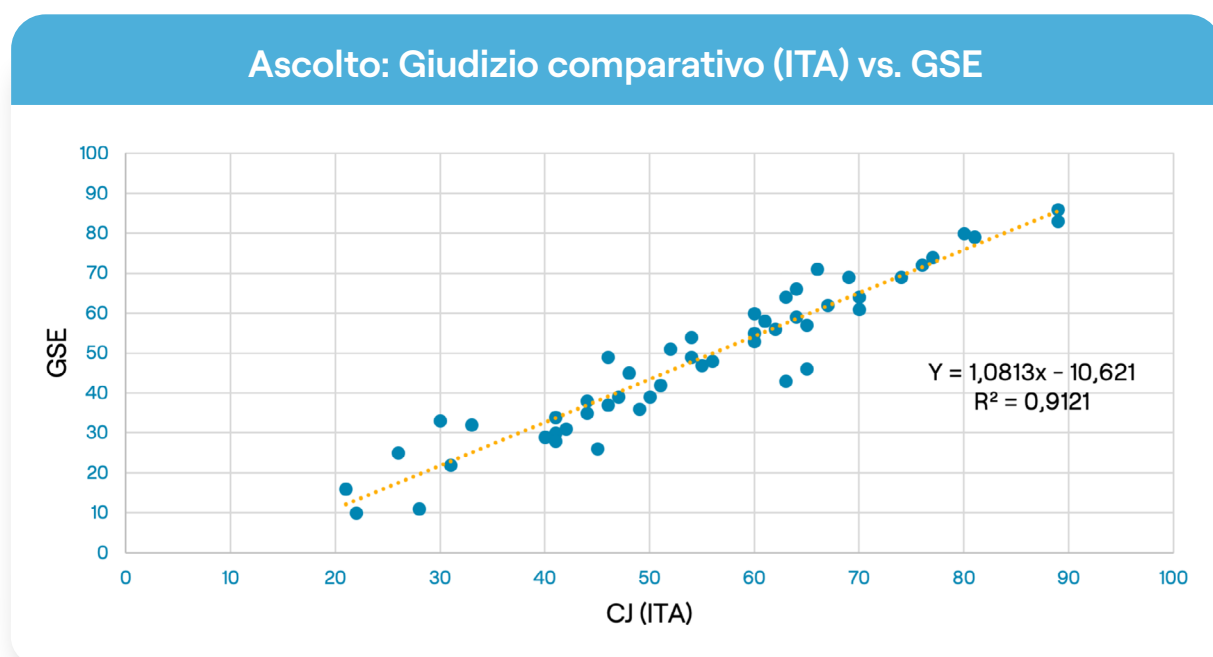
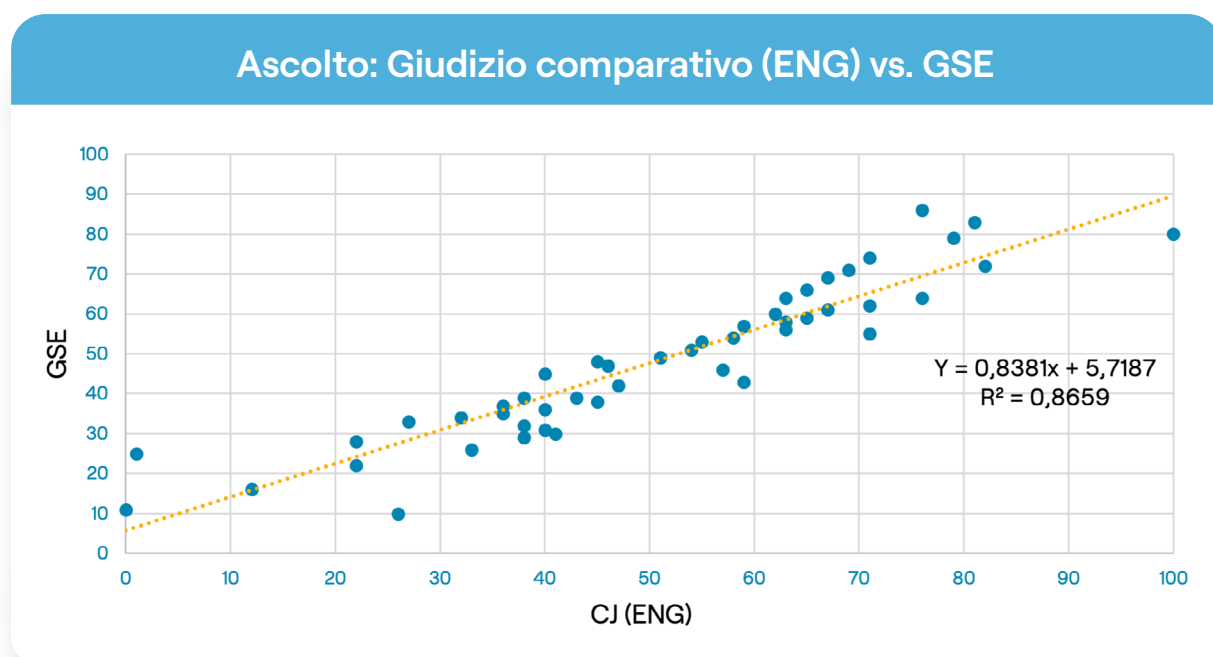
Figura 2: Ascolto – Statistiche di non adattamento degli obiettivi di apprendimento



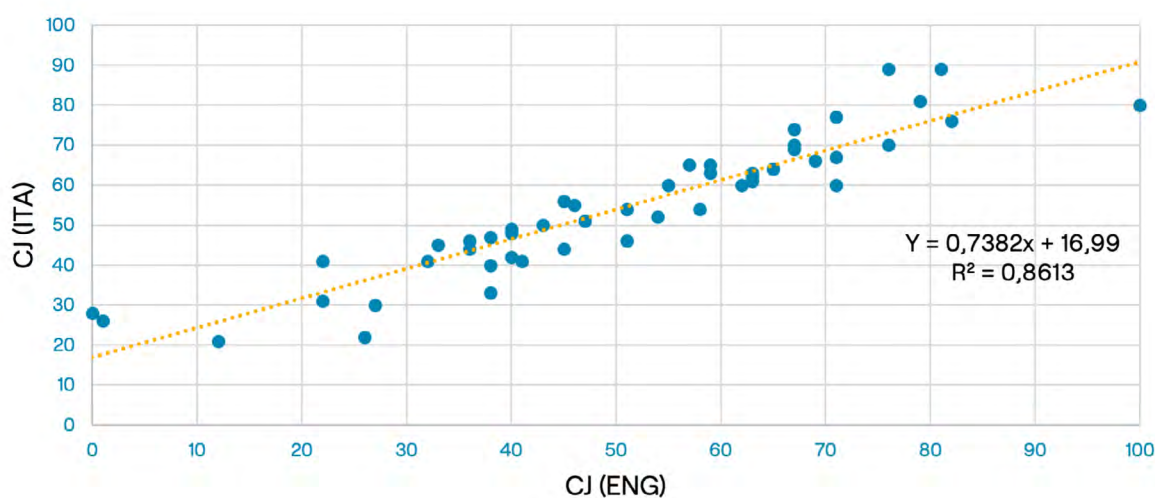
I tre obiettivi di apprendimento evidenziati nella Figura 2 sono stati segnalati per ulteriori controlli qualitativi. Poiché rientravano nell'intervallo di non adattamento accettabile, sono stati inclusi nello studio. [Nota: nella figura precedente, i due numeri sopra ai puntini rossi indicano l'ID dell'elemento seguito dalla sua statistica di non adattamento]

I tre grafici (Figura 3) sottostanti mostrano i confronti dei valori della GSE esistenti per quanto riguarda l'ascolto con le stime di giudizio comparativo degli obiettivi di apprendimento dell'inglese e con le stime di giudizio comparativo degli obiettivi di apprendimento dell'italiano, nonché le due stime di giudizio comparativo degli stessi obiettivi di apprendimento delle due lingue. Per il confronto sono stati utilizzati i punteggi scalari di giudizio comparativo forniti da *NoMoreMarking*. I risultati indicano un elevato livello di concordanza tra questi confronti.

Figura 3: Ascolto – Confronto delle difficoltà degli obiettivi di apprendimento esistenti con le stime di giudizio comparativo



Ascolto: Giudizio comparativo (ENG) vs. giudizio comparativo (ITA)

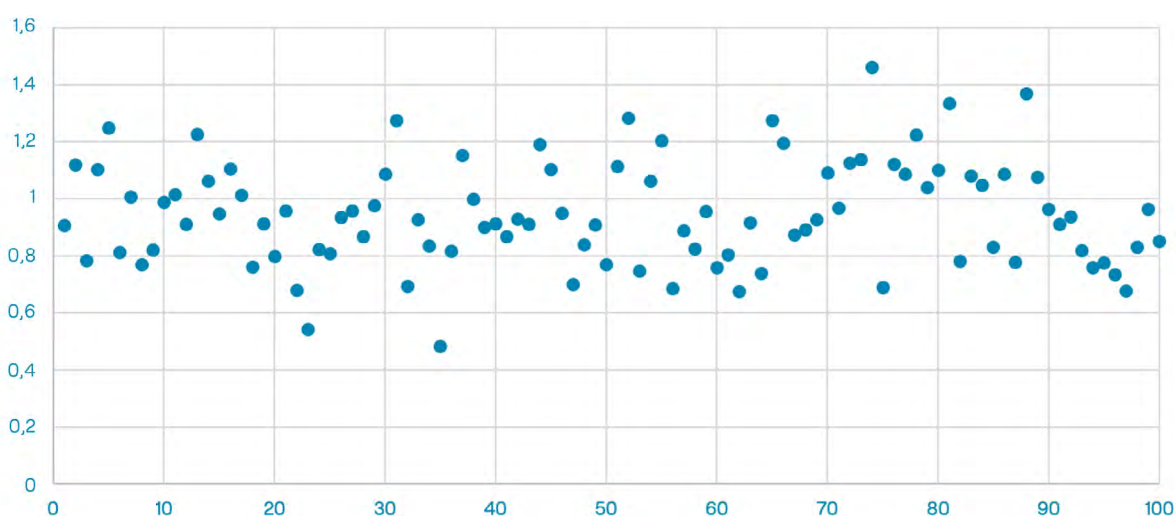


4.2.2 Lettura

La correlazione complessiva tra i valori della GSE esistenti degli obiettivi di apprendimento relativi alla lettura (sia in inglese che in italiano) e i valori di giudizio comparativo è **0,923**. Nessun obiettivo di apprendimento è stato segnalato per ulteriori indagini (Figura 4).

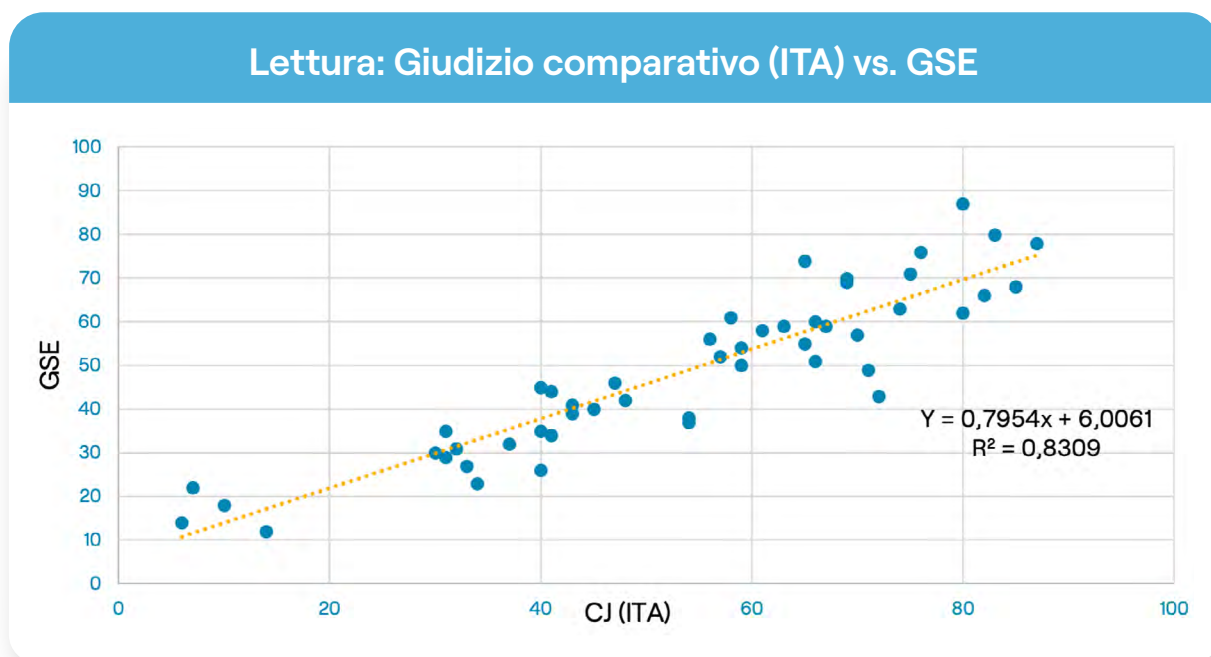
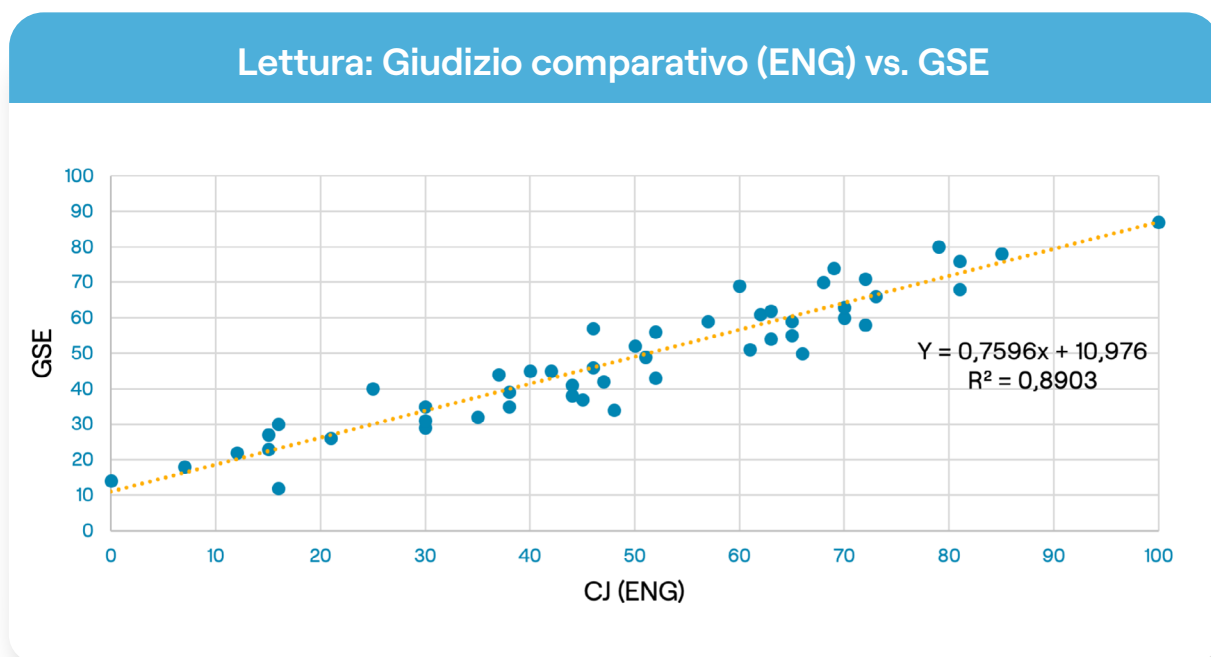
Figura 4: Lettura – Statistiche di non adattamento degli obiettivi di apprendimento

Lettura: Non adattamento degli obiettivi di apprendimento

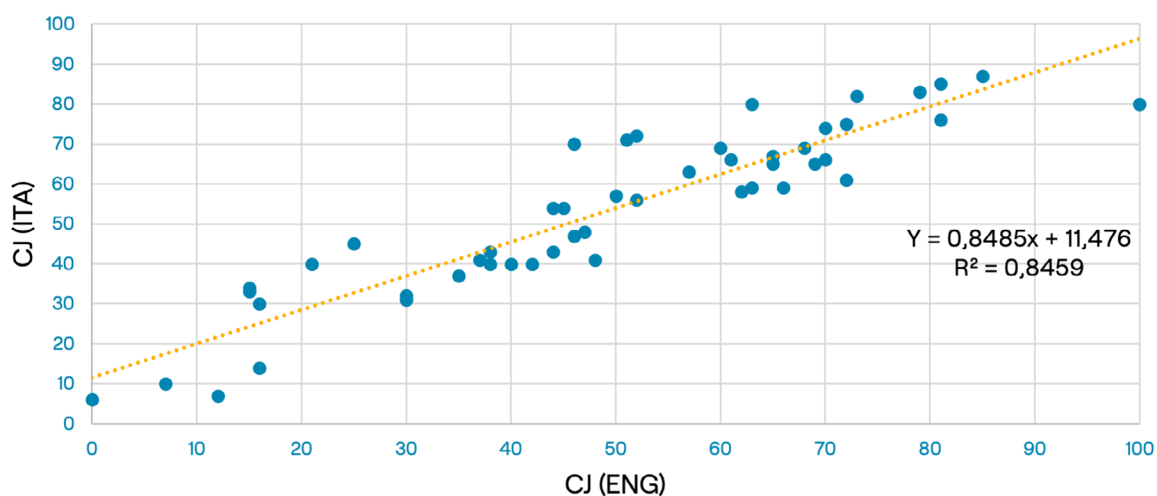


I tre grafici sottostanti (Figura 5) mostrano i confronti dei valori della GSE esistenti per quanto riguarda la lettura con le stime di giudizio comparativo degli obiettivi di apprendimento in inglese e con le stime di giudizio comparativo degli obiettivi di apprendimento in italiano, nonché le due stime di giudizio comparativo degli stessi obiettivi di apprendimento nelle due lingue. I risultati indicano un elevato livello di concordanza tra questi confronti.

Figura 5: Lettura – Confronto delle difficoltà degli obiettivi di apprendimento esistenti con le stime di giudizio comparativo



Lettura: Giudizio comparativo (ENG) vs. giudizio comparativo (ITA)

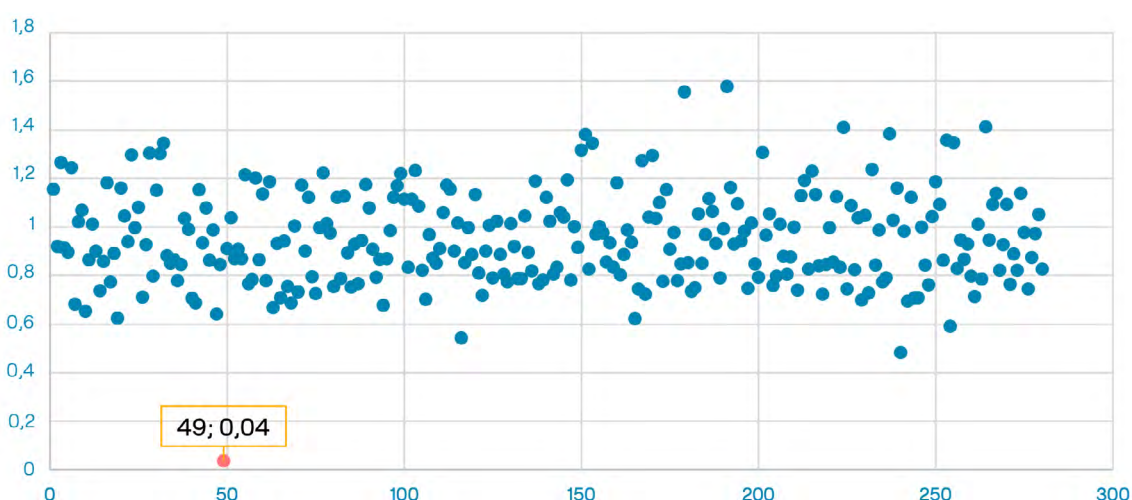


4.2.3 Conversazione

La correlazione complessiva tra i valori della GSE esistenti degli obiettivi di apprendimento relativi alla conversazione (sia in inglese che in italiano) e i valori di giudizio comparativo è **0,917**. Un obiettivo di apprendimento è stato segnalato per ulteriori indagini (Figura 6) [Nota: i due numeri sopra il puntino rosso indicano l'ID dell'elemento seguito dalla sua statistica di non adattamento].

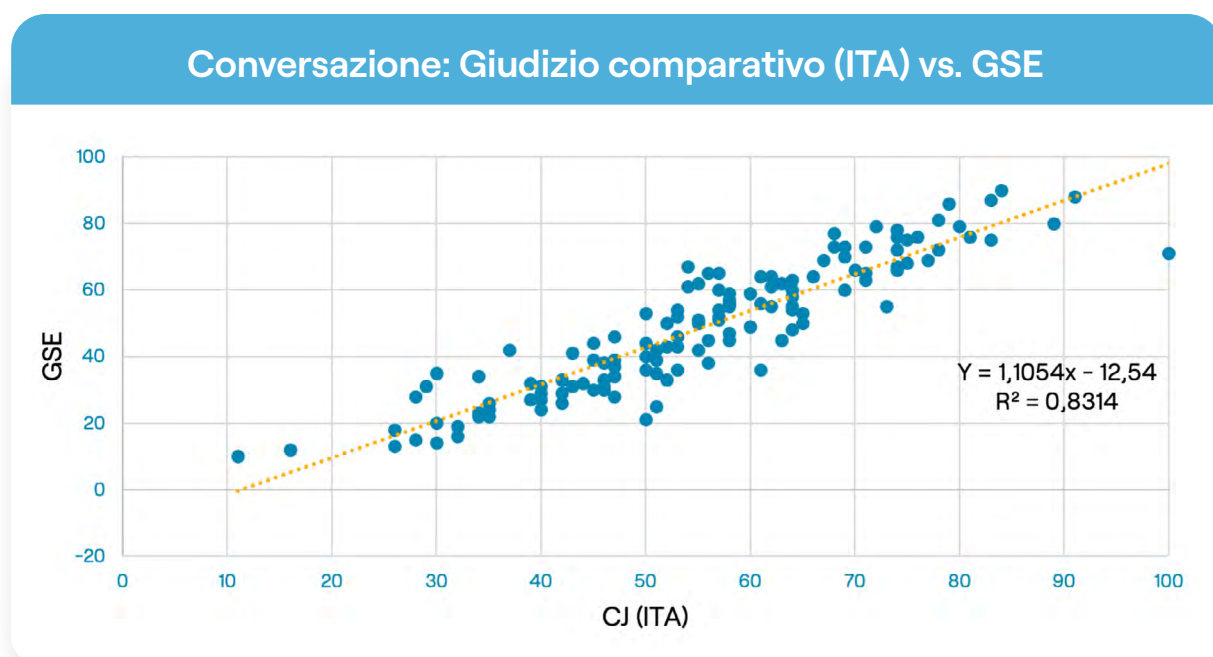
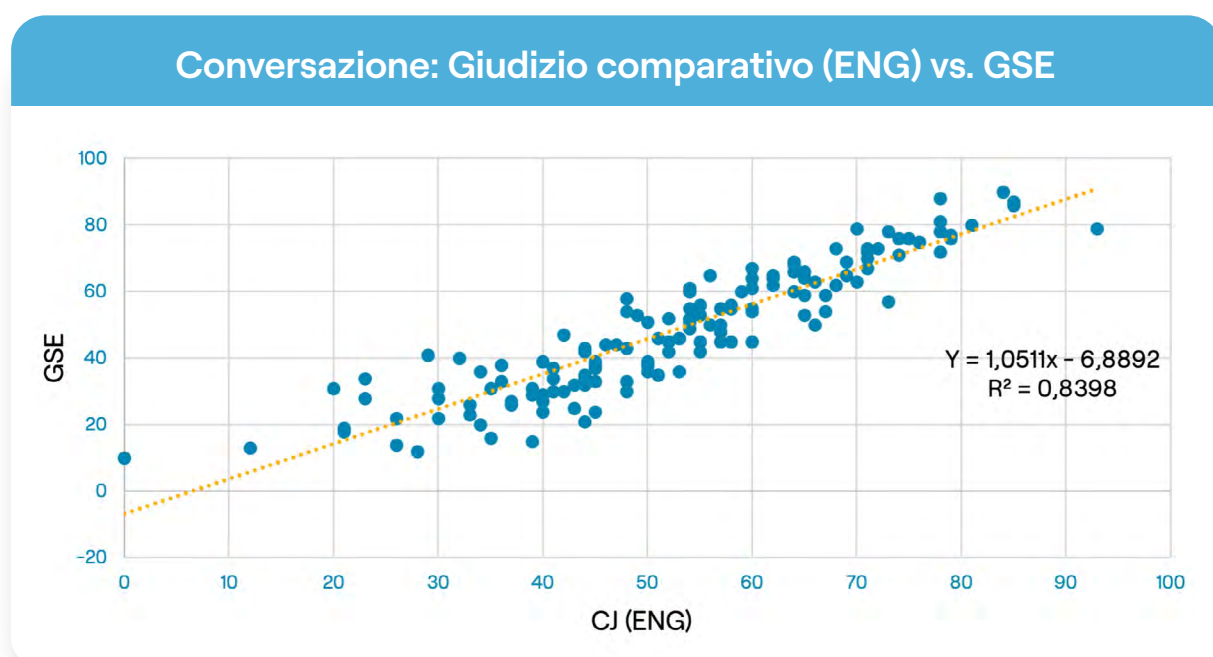
Figura 6: Conversazione – Statistiche di non adattamento degli obiettivi di apprendimento

Conversazione: Non adattamento degli obiettivi di apprendimento

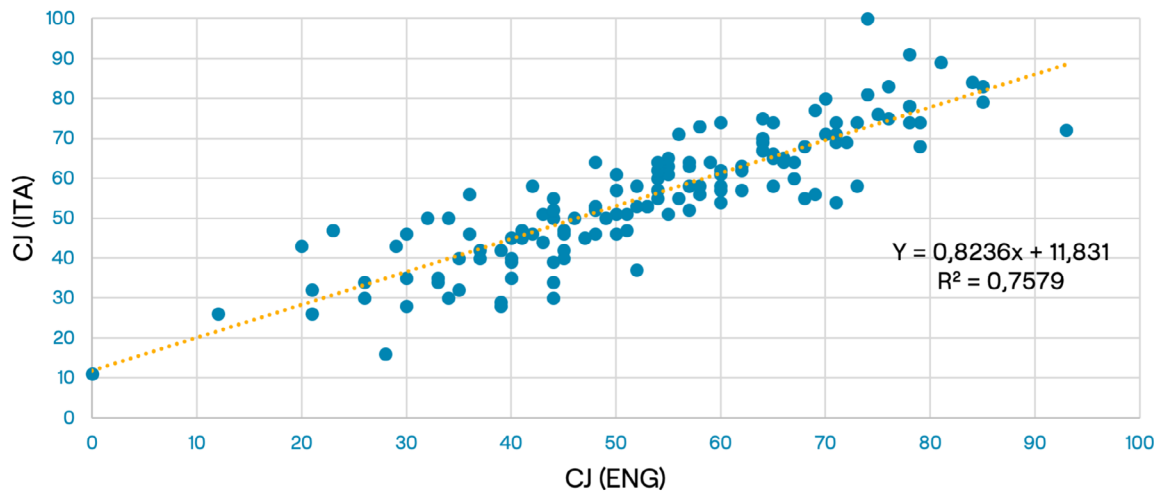


I tre grafici sottostanti (Figura 7) mostrano i confronti dei valori della GSE esistenti per quanto riguarda la conversazione con le stime di giudizio comparativo degli obiettivi di apprendimento in inglese e con le stime di giudizio comparativo degli obiettivi di apprendimento in italiano, nonché le due stime di giudizio comparativo degli stessi obiettivi di apprendimento nelle due lingue. I risultati indicano un elevato livello di concordanza tra questi confronti.

Figura 7: Conversazione – Confronto delle difficoltà degli obiettivi di apprendimento esistenti con le stime di giudizio comparativo



Conversazione: Giudizio comparativo (ENG) vs. giudizio comparativo (ITA)

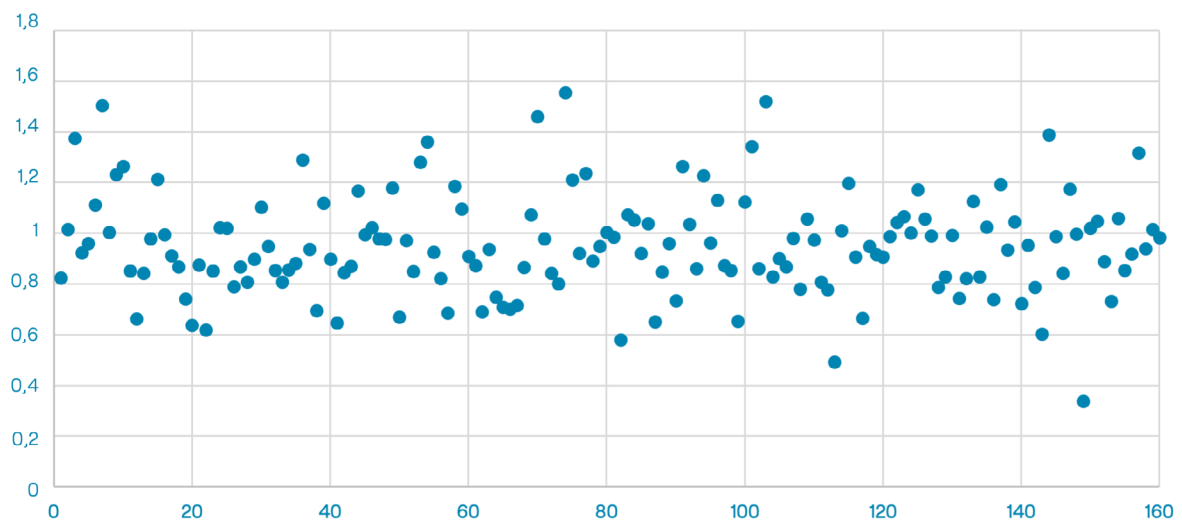


4.2.4 Scrittura

La correlazione complessiva tra i valori della GSE esistenti degli obiettivi di apprendimento relativi alla scrittura (sia in inglese che in italiano) e i valori di giudizio comparativo è **0,925**. Nessun obiettivo di apprendimento è stato segnalato per ulteriori indagini (Figura 8).

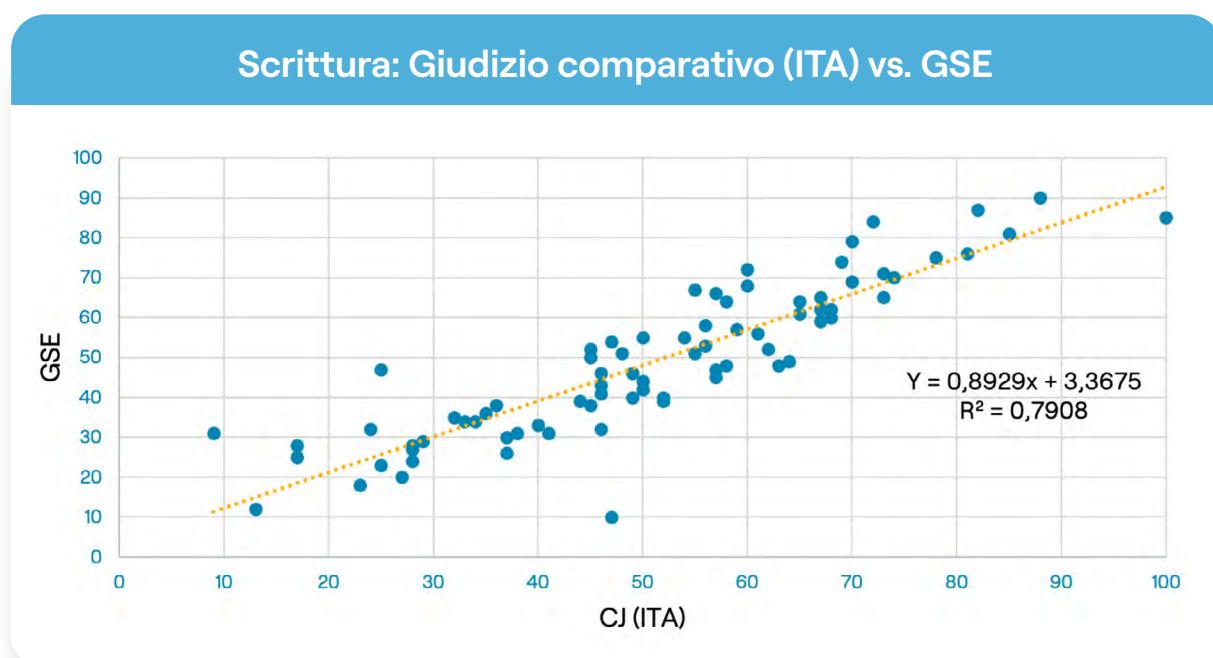
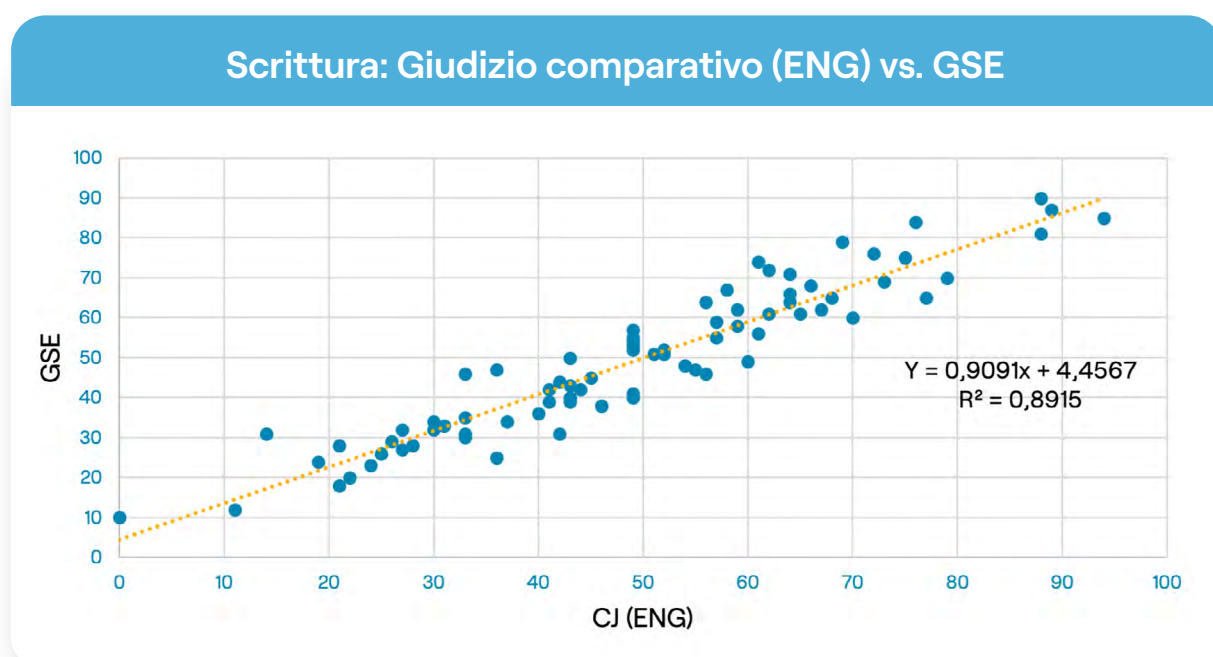
Figura 8: Scrittura – Statistiche di non adattamento degli obiettivi di apprendimento

Scrittura: Non adattamento degli obiettivi di apprendimento

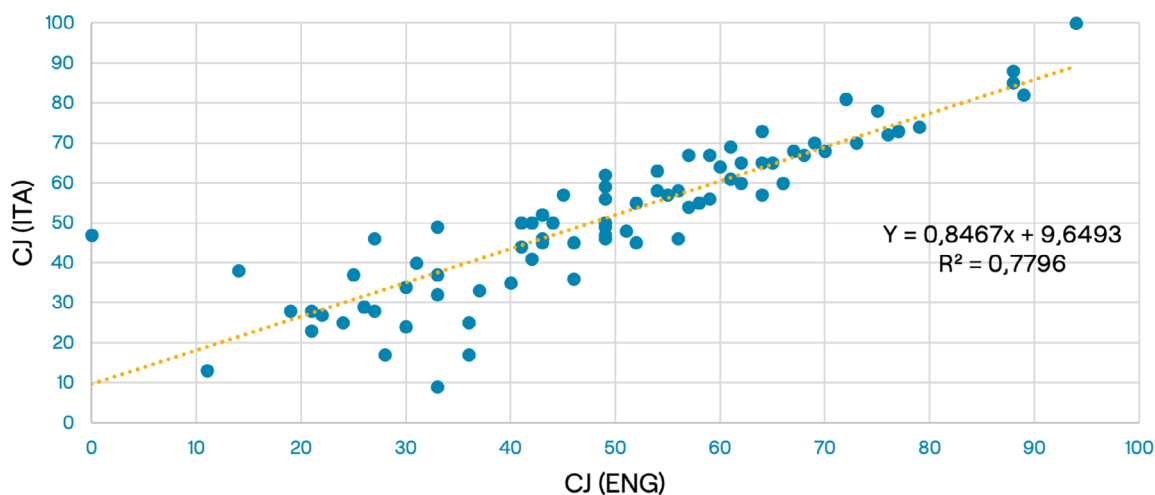


I tre grafici sottostanti (Figura 9) mostrano i confronti dei i valori della GSE esistenti per quanto riguarda la scrittura con le stime di giudizio comparativo degli obiettivi di apprendimento in inglese e con le stime di giudizio comparativo degli obiettivi di apprendimento in italiano, nonché le due stime di giudizio comparativo degli stessi obiettivi di apprendimento nelle due lingue. I risultati indicano un elevato livello di concordanza tra questi confronti.

Figura 9: Scrittura – Confronto delle difficoltà degli obiettivi di apprendimento esistenti con le stime di giudizio comparativo



Scrittura: Giudizio comparativo (ENG) vs. giudizio comparativo (ITA)



4.3 Correlazioni tra le stime della GSE e quelle di giudizio comparativo

Le correlazioni tra i valori esistenti degli obiettivi di apprendimento della GSE e i valori prodotti nello studio di giudizio comparativo sono presentate nella tabella seguente. Sono stati effettuati confronti tra i valori di giudizio comparativo degli obiettivi di apprendimento della GSE esistenti e quelli delle traduzioni in italiano. Con una sola eccezione (per la scrittura), tutte le correlazioni hanno raggiunto valori superiori a 0,9 nei confronti.

Tabella 4: Correlazioni tra i valori della GSE esistenti degli obiettivi di apprendimento e i valori emersi nello studio di giudizio comparativo per lingua e per competenza

	Valori della GSE esistenti rispetto ai valori di giudizio comparativo (inglese)	Valori della GSE esistenti rispetto ai valori di giudizio comparativo (italiano)
Ascolto	0,931	0,955
Lettura	0,944	0,912
Conversazione	0,916	0,912
Scrittura	0,944	0,889

4.4 Equazione di trasformazione per ogni competenza

Sulla base dei risultati soddisfacenti ottenuti finora, sono state generate equazioni di trasformazione per ciascuna competenza linguistica. La Tabella 5 mostra le equazioni di trasformazione dalle stime di giudizio comparativo ai valori della GSE.

A scopo dimostrativo, vengono mostrate due equazioni per ciascuna competenza, la prima per trasformare i valori di giudizio comparativo dell'inglese nella GSE e la seconda per trasformare i valori di giudizio comparativo dell'italiano nella GSL. Per l'uso effettivo, è necessaria solo la seconda equazione per allineare gli obiettivi di apprendimento dell'italiano alla Global Scale 10-90.

Tabella 5: Equazione di trasformazione dalle stime di giudizio comparativo ai valori della GSE

X= punteggio scalare di giudizio comparativo; Y= GSE

Ascolto: Giudizio comparativo (ENG) vs. GSE

$$y = 0,8381x + 5,7187$$

Ascolto: Giudizio comparativo (ITA) vs. GSL

$$y = 1,0813x - 10,621$$

Lettura: Giudizio comparativo (ENG) vs. GSE

$$y = 0,7596x + 10,976$$

Lettura: Giudizio comparativo (ITA) vs. GSL

$$y = 0,7954x + 6,0061$$

Conversazione: Giudizio comparativo (ENG) vs. GSE

$$y = 1,0511x - 6,8892$$

Conversazione: Giudizio comparativo (ITA) vs. GSL

$$y = 1,1054x - 12,54$$

Scrittura: Giudizio comparativo (ENG) vs. GSE

$$y = 0,9091x + 4,4567$$

Scrittura: Giudizio comparativo (ITA) vs. GSL

$$y = 0,8929x + 3,3675$$

5. Discussione e conclusioni

I risultati di questo studio sul giudizio comparativo mostrano un'elevata correlazione tra il livello di competenza degli stessi obiettivi di apprendimento in inglese e italiano. Come ulteriore convalida, vengono fornite specifiche equazioni di trasformazione per ciascuna competenza, che alla fine portano a definire la concordanza tra le stime di difficoltà dell'obiettivo di apprendimento dell'italiano rispettivamente sulla GSE e sul QCER.

Il QCER stesso è un quadro di riferimento linguisticamente neutrale che "può essere adattato e utilizzato per molteplici contesti e applicato a tutte le lingue" (**Consiglio d'Europa**) e, dal suo sviluppo nel 2001, è stato tradotto in 40 lingue (*ibidem*). Il lavoro di Pearson per estendere il QCER e creare la GSE è stato originariamente concepito in un contesto di inglese come lingua straniera (EFL), ma si è ritenuto che questa estensione potesse essere rilevante e utile anche per gli insegnanti e gli studenti di altre lingue. Lo studio di giudizio comparativo descritto in questo articolo fornisce prove a sostegno dell'opinione che gli atti linguistici comunicativi e funzionali espressi nelle affermazioni Can-do ("essere in grado") sia in inglese che in italiano hanno un valore comparabile in termini di competenza, cioè possono essere collocati entrambi sulla stessa scala.

Data la somiglianza tra l'italiano e altre lingue romanze come il francese e lo spagnolo, siamo convinti che questi risultati possano essere estrapolati e applicati anche in questi contesti, allo stesso modo in cui il QCER è applicabile a queste lingue europee. Sono in corso ulteriori studi sulle lingue non romanze, che andranno ad arricchire il corpus di prove di validazione della Global Scale of Languages.

Glossario

QCER: Quadro comune europeo di riferimento per le lingue

CJ: Giudizio comparativo

CSE: China Scale of English

GSE: Global Scale of English

GSL: Global Scale of Languages

Riferimenti

Bradley, R. A. e Terry, M. E. (1952). *Rank analysis of incomplete block designs. I. The method of paired comparisons*. *Biometrika* 39 324–345.

Chambers, L., & Cunningham, E. (2022). *Exploring the Validity of Comparative Judgement: Do Judges Attend to Construct-Irrelevant Features?* *Frontiers in Education* (7).

Council of Europe (2001). *Common European Framework of Reference for Languages: learning, teaching, assessment*. Cambridge: Cambridge University Press.

De Jong, J., Mayor, M., & Hayes, C. (2016). *Developing Global Scale of English Learning Objectives aligned to the Common European Framework*. Disponibile su: <https://www.pearson.com/languages/why-pearson/the-global-scale-of-english/resources.html>

Fearnley, A. (2000). A comparability study in GCSE mathematics. A study based on the summer 1998 examination. In *Assessment and Qualifications Alliance* (Northern Examinations and Assessment Board). Manchester: Joint Forum for the GCSE and GCE.

Gill, T., & Bramley, T. (2013). How accurate are examiners' holistic judgements of script quality?. *Assessment in Education: Principles, Policy & Practice*, 20(3), 308–324.

Gray, E. (2000). *A comparability study in GCSE science 1998. A study based on the 1998 summer examination*. Organizzato da Oxford, Cambridge e RSA Examinations (Midland Examining Group) per conto del comitato congiunto di GCSE e GCE.

-
- Jones, I., & Alcock, L. (2014). Peer assessment without assessment criteria. *Studies in Higher Education*, 39(10), 1774–1787.
- Kolen, M. J., & Brennan R. L. (2004). *Test equating, scaling, and linking: Methods and practices*. 2nd. New York: Springer.
- Lesterhuis, M., Verhavert, S., Coertjens, L., Donche, V., & De Maeyer, S. (2017). Comparative judgement as a promising alternative to score competences. In *Innovative practices for higher education assessment and measurement* (pp. 119–138). IGI Global.
- Marshall, N., Shaw, K., Hunter, J., & Jones, I. (2020). Assessment by comparative judgement: An application to secondary statistics and English in New Zealand. *New Zealand Journal of Educational Studies*, 55, 49–71.
- Mentzer, N., Lee, W., & Bartholomew, S. R. (2021). Examining the Validity of Adaptive Comparative Judgment for Peer Evaluation in a Design Thinking Course. In *Frontiers in Education* (p. 492). Frontiers.
- North, B. (2000). *The development of a common framework scale of language proficiency*. New York: Peter Lang.
- Pollitt, A. (2004). *Let's stop marking exams*, International Association for Educational Assessment Conference. Philadelphia PA.
- Steedle, J. T., & Ferrara, S. (2016). Evaluating comparative judgment as an approach to essay scoring. *Applied Measurement in Education*, 29(3), 211–223.
- Relazione tecnica Pearson (2020): Aligning Global Scale of English-Young Learner to the CSE. Disponibile su <https://m.i21st.cn/elt/15934.html>
- Thurstone, L. L. (1927). A law of comparative judgment. *Psychological review*, 34(4), 273.
- Verhavert, S., Bouwer, R., Donche, V., & Maeyer, S. D. (2019). A meta-analysis on the reliability of comparative judgement. *Assessment in Education: Principles, Policy & Practice*, 26(5), 541–562.
- Wheadon, C. (2019). *No More Marking [Computer Software]*. Estratto da <https://www.nomoremarking.com/>

Appendice: Dati demografici del valutatore

Nazionalità	Conteggio
Britannica	8
Francese	1
Italiana	1
Messicana	2
Spagnola	8
TOTALE	20

Sesso	Conteggio
Uomo	7
Preferisco non dichiarare	1
Donna	12
TOTALE	20

Anni di insegnamento dell'italiano	Conteggio
>10 anni	16
5-10 anni	4
TOTALE	20

Familiarità con il QCER	Conteggio
Conoscenza approfondita	5
Comprensione generica	8
Ne sono a conoscenza	7
TOTALE	20

Altre lingue di insegnamento	Conteggio*
Francese	16
Inglese	6
Tedesco	2
Italiano	2
Latino	2
Greco antico	1
Giapponese	1
Catalano	1

* Tutti avevano insegnato almeno un'altra lingua

Fascia/e di età degli studenti (italiano)	Conteggio
Adulti (18+)	20
Secondaria superiore/college/6° livello (15-19)	18
Secondaria inferiore (12-15);	18
Primaria superiore (9-12);	12
Primaria inferiore (6-9)	3
Scuola dell'infanzia (3-5)	2



Global
Scale of
Languages

Fast-track your progress