

LA VALUTAZIONE CERTIFICATORIA E  
L'IMPORTANZA DELLA RIPETIZIONE:  
QUANTI E QUALI TASK PER GARANTIRE LA  
RAPPRESENTATIVITÀ DEL COSTRUTTO?<sup>1</sup>

INTRODUZIONE

La validità e l'affidabilità sono qualità che contribuiscono all'utilità di un test (Bachman/Palmer 1996). Il fatto che un test sia valido e affidabile ci garantisce che i punteggi siano correttamente interpretati e indicativi dell'abilità o della competenza linguistica dei candidati. Questo, di conseguenza, ci assicura che vengano prese decisioni giuste – ed etiche – sulla base di tali interpretazioni (Shohamy 2001; McNamara *et al.* 2019).

Affinché un processo di valutazione linguistica e i suoi strumenti siano validi ed affidabili, è innanzitutto necessario che il costrutto che ne sta alla base sia chiaramente definito e venga adeguatamente rappresentato dai diversi test di cui si fa uso (Fulcher/Davidson 2007).

Infatti, fra i diversi aspetti che definiscono la validità di un test, la rappresentativi-

---

<sup>1</sup> Il testo è frutto di una riflessione comune e condivisa. Per quanto riguarda la responsabilità dei singoli paragrafi sono di Aisha Nasimi i parr. 3, 4 e *Conclusioni e prospettive future*; di Giulia Peri i parr. *Introduzione*, 1 e 2.

tà del costrutto è tra i più significativi (Messick 1989, 1996; Cumming 1996), poiché garantisce che il test misuri ciò che effettivamente si è proposto di misurare (Hughes 1989; Bachman 2000). A tale scopo, all'interno di un test la costruzione e la ripetizione di più task per la verifica della medesima abilità linguistica costituiscono un'operazione fondamentale, in quanto ogni task favorisce l'elicitazione di singoli tratti che nel loro insieme definiscono una specifica abilità.

In questo contributo, dopo aver illustrato il quadro teorico di riferimento per il nostro lavoro nel par. 1, nel par. 2 prenderemo in esame il caso CILS al fine di discutere nel par. 4 i risultati di uno studio di validazione del costrutto (Messick 1996; McNamara/Roever 2006) condotto su diversi task per la verifica dell'abilità di scrittura nell'esame CILS B1 Adolescenti. L'obiettivo è indagare la misura in cui il costrutto del test viene rappresentato attraverso tali task. La metodologia utilizzata – che prevede la codifica degli input proposti e la loro successiva analisi – è illustrata nel par. 3. Infine, argenteremo le conclusioni e rifletteremo su alcune prospettive future.

## 1. IL QUADRO TEORICO DI RIFERIMENTO: VALIDITÀ E AFFIDABILITÀ DEL TEST

Cronbach e Meehl (1955) individuavano tra i tre principali tipi di validità quella del costrutto. Il costrutto può essere visto come la teoria di riferimento che definisce la conoscenza e le abilità che vogliamo misurare attraverso il test (Bachman 1990; Green 2013). La definizione del costrutto è, infatti, il primo passo che dobbiamo compiere nella costruzione di un test ma, allo stesso tempo, è anche l'operazione più complessa, in quanto ha a che fare con la definizione del modello di competenza che prendiamo come riferimento. Solo quando si è definito il modello di competenza che sta alla base del test, possiamo “operazionalizzarlo” – metterlo in pratica tramite i task del test stesso – e misurarlo. In questo modo il costrutto non è più solo un concetto teorico, ma qualcosa di misurabile.

Nonostante l'approccio di Cronbach e Meehl (1955) sia stato criticato e superato (Fulcher/Davidson 2007), l'importanza del costrutto e della sua validità ha avuto grande risonanza e possiamo dire che non si è mai persa nello sviluppo della letteratura relativa alla validità, tutta focalizzata sull'inglese L2 (cfr. Messick 1989; Bachman/Palmer 1996; Weir 2005).<sup>2</sup>

A tale proposito, Messick (1989: 20) afferma che le evidenze relative al contenuto e al criterio contribuiscono alla significatività del punteggio, e per questo possono essere definite come parte della validità del costrutto. Questa sua affermazione, perciò, porta all'individuazione di un solo tipo di evidenze di validità, ossia quelle relative al costrutto.

---

<sup>2</sup> Per approfondimenti sull'evoluzione negli ultimi decenni delle riflessioni sul concetto di validità e di validazione nel campo della verifica e valutazione delle competenze linguistiche, vedi Chapelle/Voss (2014).

Messick (1989), inoltre, individua due principali minacce alla validità collegate direttamente al costrutto: la *construct-irrelevant variance* e la *construct under-representation*. La prima si riferisce al fatto che nel processo di valutazione vengano coinvolti elementi non previsti dal costrutto, per esempio quando si richiede a un candidato di utilizzare la sua capacità di memorizzazione per portare a termine un compito linguistico. La seconda è legata alla non rappresentatività delle evidenze raccolte in merito al costrutto di riferimento, quindi alla non inclusione di elementi rilevanti per le conoscenze e le abilità che vogliamo misurare. Per esempio, se vogliamo testare l'abilità di scrittura in ambito accademico di un candidato, costruire degli input che invitano a sviluppare una performance su argomenti esclusivamente legati alla sfera personale farebbe venire meno l'appropriatezza dell'argomento dell'input.

A tale scopo e come già si diceva poco sopra, la costruzione e le ripetizioni, all'interno di un test, di più task per la verifica della medesima abilità linguistica costituiscono un'operazione fondamentale, in quanto ogni task favorisce l'elicitazione di singoli tratti che nel loro insieme definiscono una specifica abilità.

Per garantire ai destinatari del test e a tutti i soggetti coinvolti che il costrutto venga adeguatamente rappresentato e “operazionalizzato” nel test, e dunque che la proposta di più task per la verifica di una stessa abilità sia un'operazione che contribuisce ad innalzare l'utilità di un test, si rende necessario ciò che si definisce “processo di validazione”. Tramite questo processo, potenzialmente senza fine (Cronbach/Meehl 1955) o comunque reiterativo, è possibile ricavare delle evidenze che dovrebbero andare a supportare la validità del test.

## 2. IL CASO CILS: L'ESAME B1 ADOLESCENTI

L'esame CILS B1 Adolescenti è costituito da prove che misurano la competenza nelle abilità linguistiche di base (ascolto, lettura, produzione scritta e produzione orale) e si articola in 5 sezioni:

- Ascolto;
- Comprensione della lettura;
- Analisi delle strutture di comunicazione;
- Produzione scritta;
- Produzione orale.

La tab. 1 fornisce la descrizione del formato dell'esame.

Prove	Ascolto	Comprensione della lettura	Analisi delle strutture della comunicazione	Produzione scritta	Produzione orale
I	Dettato- <i>cloze</i>	Testo divulgativo/informativo con test a scelta multipla	<i>Cloze</i> mirato su articoli e preposizioni articolate	Prova a tema: descrizione	Dialogo
II	Dialogo con test a scelta multipla	Testo informativo/regolativo con test a individuazione di informazioni	Test a completamento su forme verbali	Prova strutturata	Monologo
III	Monologhi con test a individuazione di informazioni	Testo narrativo con test a ricostruzione testuale	Test a completamento lessicale con scelta multipla		
IV			Test pragmatico con scelta multipla		

**tab. 1.** *Il formato dell'esame CILS B1 Adolescenti (Centro CILS 2009: 45)*

Come anticipato nell'Introduzione, la nostra ricerca si inserisce nel processo di validazione operato ciclicamente sugli esami CILS ed analizza il test di Produzione scritta dell'esame B1 Adolescenti. Questo si compone di due prove per la verifica della medesima abilità: la prova 1 (cfr. fig. 1) che richiede al candidato la scrittura di un testo descrittivo (es. descrizione di persone o luoghi familiari; cfr. ivi: 31) o narrativo (racconto di una storia con personaggi dati, narrazione diaristica di una gita, di un avvenimento o di un episodio particolare; *ibid.*); la prova 2 (cfr. fig. 2) che richiede al candidato l'esecuzione di una prova semi-strutturata (lettera informale a parenti o amici per ringraziare, chiedere informazioni, raccontare esperienze evidenziando i punti salienti; *ibid.*), quindi la risposta ad un input che funziona da guida più dettagliata per l'esecuzione della prova stessa. Sia per la prova 1 che per la prova 2 vengono fornite indicazioni sul numero di parole richiesto.

### 3. LA METODOLOGIA

La validazione del test di Produzione scritta dell'esame CILS B1 Adolescenti prende le mosse dalle seguenti domande di ricerca:

1. Il numero e la tipologia dei task proposti consente di misurare in maniera adeguata la competenza linguistico-comunicativa prevista per il livello B1?
2. Il numero dei task proposti incide sul grado di validità del test di Produzione

scritta, garantendo cioè una più adeguata rappresentazione del costrutto dello stesso test?

La risposta a tali domande comporta innanzitutto, seguendo il modello proposto da Kunnan (2008: 237) e riportato in tab. 2, la riflessione sulla rappresentatività dei task, sulla rappresentatività degli argomenti e sui tratti del costrutto presi in considerazione.

<b><u>Main quality</u></b>	<b><u>Main focus of analysis</u></b>
I Validity	
Content representativeness and relevance	Representativeness of items, tasks, topics
Construct or theory-based validity	
Criterion-related validity	Construct/underlying trait
Reliability	Score comparison with external criteria
	Internal consistency, inter-rater, and alternate forms

**tab. 2.** La sezione del Fairness Framework di Kunnan (2008: 237) dedicata alla validità (Carr 2011: 155)

Per rilevare i dati necessari a rispondere alle precedenti domande di ricerca, si è tenuto conto, in primo luogo, dei riferimenti alla base del modello di lingua e della sua operazionalizzazione nell'esame CILS B1 Adolescenti relativamente alla sezione dedicata alla Produzione scritta, nello specifico i descrittori relativi al livello B1 del QCER (CoE 2001: 61-62; 83-84) e del Sillabo CILS<sup>3</sup> (Sillabo CILS: 1-3).

Si è ritenuto opportuno prendere in considerazione anche il Companion Volume (CoE 2018a: 75-76; 93-97), il documento che integra ed amplia il QCER (CoE 2001). Per la selezione dei descrittori pertinenti, si sono presi a riferimento i corrispondenti descrittori del QCER, includendo anche i descrittori rilevanti per il pubblico degli *young learners* (CoE 2018b).

In secondo luogo, sono stati selezionati 30 input della Produzione scritta dell'esame e del livello considerato, di cui 15 per la prova 1 e 15 per la prova 2.

L'obiettivo di questa selezione è consistito nello stabilire:

- il grado di corrispondenza tra ognuno dei descrittori appena menzionati e gli input;

<sup>3</sup> «La certificazione CILS è progettata e realizzata tenendo conto dei parametri di certificazione promossi dal Consiglio d'Europa, contenuti nel [...] QCER, e dalle altre istituzioni europee che si occupano della diffusione delle lingue e delle culture» (Centro CILS 2009: 12). Gli esami CILS sono allineati al QCER (Barni *et al.* 2010), ma il Centro CILS opera un processo di validazione costante sugli esami CILS e sul Sillabo CILS, entro cui si colloca anche la presente ricerca.

- il grado di corrispondenza tra le funzioni linguistiche elicitate dai 30 input e ciascun descrittore;
- il numero di occorrenze di ogni singolo descrittore, sia per prova che per input.

La raccolta dei dati e la loro successiva analisi ha richiesto l'attribuzione ad ogni input di un codice. Il codice consiste di lettere e numeri: ad ogni lettera, numero o coppia di numeri corrisponde un'informazione relativa all'input stesso. Per esempio, la stringa P1\_06\_16\_01\_01 equivale a:

P1 = prova 1

06 = mese della sessione d'esame CILS in cui l'input è stato proposto

16 = anno della sessione d'esame in cui l'input è stato proposto

01 = numero progressivo sul totale degli input per prova

01 = numero progressivo sul totale degli input analizzati

La raccolta dei dati è anche consistita nella rilevazione dei domini e dei contesti d'uso elicitati dagli input presi in considerazione.

Analogamente a quanto operato per gli input, ogni descrittore selezionato è stato codificato.<sup>4</sup> Per esempio, prendendo in considerazione un descrittore del Sillabo CILS, la stringa CILS\_P1\_SC\_1 corrisponde a:

CILS = descrittore del Sillabo CILS

P1 = prova 1

SC = scrittura creativa

Tale operazione è stata ripetuta per ognuno dei tre documenti presi a riferimento. Per quanto riguarda la codifica dei descrittori del Companion Volume (CoE 2018a), sono stati individuati i descrittori rilevanti per il pubblico degli *young learners* (CoE 2018b).

## 4. L'ANALISI DEI DATI

### 4.1 I descrittori del Sillabo CILS

Dall'analisi dei dati, raccolti ed organizzati secondo la metodologia sopra illustrata, è emerso che in relazione agli input della prova 1 (cfr. fig. 3), questi sono maggiormente rappresentativi del descrittore CILS\_P1\_SC\_1 (33%) – “è in grado di scrivere descrizioni lineari e precise su una gamma di argomenti familiari che rientrano nel suo campo d'interesse” – seguito dal descrittore CILS\_P1\_SC\_2 (28%) – “è in grado

<sup>4</sup> Per gli scopi della codifica, quando possibile, per ogni abilità sono state isolate le relative sotto-abilità, progressivamente numerate. Per semplicità, utilizzeremo più spesso in questa sede il termine “descrittore”.

di scrivere resoconti di esperienze, descrivendo sentimenti e impressioni in un semplice testo coeso”. I descrittori meno rappresentati dagli input, quindi con il minor numero di occorrenze, sono il CILS\_P1\_SC\_3 – “è in grado di descrivere un avvenimento, un viaggio recente – reale o immaginario” – e il CILS\_P1\_SC\_4 – “raccontare una storia”.

Per ciò che concerne gli input della prova 2 (cfr. fig. 4), il descrittore maggiormente rappresentato è il CILS\_P2\_IS\_2 – “in grado di scrivere lettere e appunti personali per chiedere o dare semplici informazioni di interesse immediato, riuscendo a mettere in evidenza ciò che ritiene importante” – seguito dal CILS\_P2\_AMM\_1 – “è in grado di scrivere messaggi con informazioni di interesse immediato da trasmettere ad amici, persone di servizio, insegnanti e altre persone frequentate nella vita di tutti i giorni, riuscendo a far comprendere i punti che ritiene importanti”, seguito a sua volta dal CILS\_P2\_IS\_1, descrittore ancora una volta relativo all’interazione scritta – “è in grado di trasmettere informazioni e idee su argomenti sia astratti sia concreti, verificare le informazioni ricevute, porre domande su un problema o spiegarlo con ragionevole precisione”.

#### *4.2 I descrittori del QCER*

L’analisi che ha interessato i descrittori del QCER (cfr. fig. 5) ha evidenziato che il descrittore B1\_PSG\_r – “produzione scritta generale” è quello maggiormente rappresentato dagli input presi in esame. Gli altri descrittori contano almeno la metà delle occorrenze rispetto al precedente. Nel dettaglio, si tratta dei descrittori “scrittura creativa” (B1\_SC\_1\_r; B1\_SC\_2\_r; B1\_SC\_3\_r; B1\_SC\_4\_r) e “interazione scritta” (B1\_ISG\_1\_r; B1\_ISG\_2\_r), B1\_AMM\_2\_r – “scrittura di appunti, messaggi e moduli”, seguiti dal descrittore B1\_C\_3\_p – “corrispondenza”. La “produzione di relazioni e saggi” (B1\_SRS\_1\_r; B1\_SRS\_2\_r) è il descrittore meno rappresentato dagli input.

Per ciò che concerne le singole prove, l’analisi evidenzia (cfr. fig. 6) che gli input della prova 1 sono rappresentativi in maniera quasi prevalente dei descrittori relativi alla “scrittura creativa”. Al contrario, quelli della prova 2 sono rappresentativi di un numero maggiore di descrittori quali: “interazione scritta”, “scrittura di appunti, messaggi e moduli” e “produzione scritta generale”.

#### *4.3 I domini e i contesti d’uso della lingua*

I dati relativi ai domini d’uso della lingua (cfr. fig. 7) hanno evidenziato come il dominio personale (PE) e quello educativo (ED) siano quelli maggiormente rappresentati dagli input. Tale dato ci sembra risultare coerente con quanto affermato nella descrizione del profilo di competenza del livello CILS B1 Adolescenti (Centro CILS, 2009: 31; cfr. infra par. 5.1). Nettamente minore è il numero di input riconducibili al dominio pubblico (PU); nessun input è invece, come era facile attendersi, rappresentativo del dominio occupazionale.

Considerando invece le singole prove (cfr. fig. 8), gli input della prova 1 sono rappresentativi per lo più del dominio personale, mentre quelli della prova 2 (cfr. fig. 9) sono maggiormente rappresentativi del dominio educativo.

Per quanto riguarda i contesti d'uso della lingua – che forniscono ulteriori indicazioni circa l'adeguatezza degli input per il pubblico di riferimento e quindi sul loro grado di rappresentatività del costruito – sul totale (36) dei contesti rappresentati dagli input, il contesto educativo (scuola 11%; insegnanti 9%) risulta essere quello maggiormente rappresentato. Inoltre, seppur con percentuali inferiori, gli input risultano essere rappresentativi anche dei contesti relativi alla sfera sociale (amicizie 7%, visite e scambi 6%).

#### *4.4 La rappresentatività degli input: le funzioni linguistiche*

Lo studio qui presentato ha previsto anche una riflessione più dettagliata relativa alle funzioni linguistiche (Mariani 2014) attivate dagli input presi in esame. L'obiettivo è quello di controllare la misura in cui gli input presi in considerazione elicitano adeguatamente la competenza dei candidati in relazione all'abilità di produzione scritta attraverso una sufficiente varietà di funzioni linguistiche a essi ricollegati.

I dati raccolti evidenziano come gli input della prova 1 siano rappresentativi per lo più dei descrittori riconducibili alle funzioni relative alla “scrittura creativa”, che ricorre nella totalità dei casi, pur con sotto-abilità diverse. Diverso è il caso degli input della prova 2, più diversificati, da ricollegare ai descrittori relativi alle funzioni della “corrispondenza” (CILS\_P2\_C\_1; CILS\_P2\_C\_2), alla capacità di “scrivere messaggi e prendere nota” (CILS\_P2\_AMM\_1) e all’“interazione scritta” (CILS\_P2\_IS\_1; CILS\_P2\_IS\_2). Tale diversificazione non è inaspettata ma riflette la tipologia delle prove: la prima è una “prova a tema”, la seconda è riconducibile alla tipologia “lettera” (Sillabo CILS). I descrittori del QCER B1\_SRS\_1\_r – “è in grado di scrivere saggi brevi e semplici su argomenti che lo/le interessano” e B1\_SRS\_3\_p – “è in grado di scrivere relazioni molto brevi su un modello standard per trasmettere informazioni fattuali sulle routine abituali e spiegare le ragioni delle azioni”, invece, non sono stati mai presi in considerazione per la costruzione delle prove. Tuttavia, questi non fanno parte dei descrittori del Sillabo CILS, nel quale rientra invece, per la categoria “Scrivere relazioni e saggi”, soltanto il descrittore B1\_SRS\_2\_r – “è in grado di scrivere con discreta sicurezza per riassumere e riferire informazioni fattuali sulle routine abituali e su altri aspetti del vivere quotidiano che lo/la riguardano esprimendo la propria opinione.” In conclusione, seppur gli input di entrambe le prove lavorino in maniera pressoché costante sulle funzioni linguistiche dei descrittori più rappresentati, tralasciandone altri, tali funzioni risultano essere adeguatamente coerenti con il profilo del gruppo target (Sillabo CILS; cfr. infra par. 5.1).

Per ciò che concerne le funzioni linguistiche in base ai domini maggiormente rappresentati dagli input delle prove prese in esame (dominio personale maggiormente rappresentato nella prova 1; dominio educativo nella prova 2), i dati raccolti (cfr. in-



fra par. 4.3) sono a loro volta rappresentativi del descrittore del livello B1 del QCER che evidenzia i compiti comunicativi legati per lo più alla sfera personale: “sa scrivere semplici testi coerenti su argomenti noti e di suo interesse. Sa scrivere lettere personali esponendo esperienze e impressioni” (CoE 2001). I domini personale ed educativo sono perciò adeguati rispetto a quanto dichiarato nel Sillabo CILS relativamente al profilo del livello B1 Adolescenti per l’abilità della scrittura: “Il candidato è in grado di produrre testi scritti con strutture semplici, ma che trasmettano le informazioni in modo chiaro e efficace dal punto di vista comunicativo su argomenti noti o di suo interesse” (Centro CILS 2009: 34).

#### *4.5 La rappresentatività del costrutto: il profilo di competenza CILS B1 Adolescenti*

In estrema sintesi, gli input analizzati elicitano le conoscenze e competenze previste per il livello B1. Il costrutto dell’esame CILS B1 Adolescenti viene dunque adeguatamente rappresentato dagli input delle prove di Produzione scritta, costruite con l’obiettivo di elicitarne la competenza dell’«apprendente autonomo del QCER, e di verificare inoltre le capacità comunicative necessarie per usare la lingua italiana in maniera autonoma e in modo adeguato nei contesti e nelle situazioni comunicative di un adolescente in Italia» (Centro CILS 2009: 45). Inoltre, i dati raccolti ci permettono di supportare la dichiarazione proposta per il profilo CILS B1 Adolescenti che fa riferimento ad un pubblico compreso nella fascia d’età 14-18 anni, inserito nei contesti formativi scolastici, in Italia e all’estero.

### 5. CONCLUSIONI E PROSPETTIVE FUTURE

Il numero e la tipologia degli input analizzati fornisce un interessante riscontro sulla capacità di questi di attivare ed elicitarne le competenze linguistico-comunicative del gruppo target. Come discusso poco sopra, le domande di ricerca (cfr. infra par. 3) hanno dato un esito generalmente positivo: il test di Produzione scritta dell’esame CILS B1 Adolescenti consente di misurare la competenza linguistico-comunicativa prevista per il livello B1 nel rispetto dei requisiti di validità. La ripetizione degli input per la Produzione scritta, che si attua nella proposta di due task, permette di rappresentare in maniera adeguata le diverse funzioni linguistiche, senza però replicarle (cfr. infra par. 4.1; 4.2), fornendo quindi evidenze diversificate (e più rappresentative) per una medesima abilità. Inoltre, il dichiarato allineamento tra i descrittori del Sillabo CILS e del QCER (Centro CILS 2009: 12) e i domini e contesti maggiormente rappresentati negli input di una o dell’altra prova, denota la rilevanza degli stessi per il gruppo target. In ultimo, se annoveriamo tra i contesti le situazioni comunicative di un adolescente in Italia e all’estero quelli relativi alla sfera educativa e alla vita quotidiana, possiamo di nuovo affermare come gli input analizzati siano in linea con i compiti che il candidato, come attore sociale, dovrà essere in grado di portare a termine in tali contesti.

I dati raccolti evidenziano infine la necessità di un adeguamento del Sillabo a quanto proposto dal Companion Volume, ad esempio in relazione ad alcuni descrittori considerati rilevanti per il gruppo degli *young learners* (CoE 2018b), quali i descrittori della produzione scritta e interazione online che ben riflettono i bisogni linguistico-comunicativi degli apprendenti adolescenti di oggi, che agiscono in contesti sempre più pervasi da modalità comunicative differenti e simultanee (Palermo 2017).

## BIBLIOGRAFIA

- Bachman 1990 = Lyle F. Bachman, *Fundamental Considerations in Language Testing*, Oxford, Oxford University Press.
- Bachman 2000 = Lyle F. Bachman, *Modern language testing at the turn of the century: assuring that what we count counts*, in «Language Testing», 17 (1), pp. 1-42.
- Bachman/Palmer 1996 = Lyle F. Bachman / Adrian S. Palmer, *Language Testing in Practice Designing and Developing Useful Language Tests*, Oxford, Oxford University Press.
- Barni *et al.* 2010 = Monica Barni / Anna Maria Scaglioso / Sabrina Machetti, *Linking the CILS examinations to the CEFR: the A1 speaking test*, in Waldemar Martyniuk (a cura di), *Aligning Tests with the CEFR. Reflections on using the Council of Europe's draft manual*, Cambridge, Cambridge University Press.
- Carr 2011 = Nathan Carr, *Designing and Analyzing Language Tests*, Oxford, Oxford University Press.
- Centro CILS 2009 = Monica Barni *et al.*, *Linee guida CILS, Certificazione di Italiano come Lingua Straniera*, con la presentazione di Massimo Vedovelli, Perugia, Guerra Edizioni.
- CoE 2001 = Council of Europe, *Common European Framework of Reference for Languages: Learning, Teaching, Assessment*, Cambridge, Cambridge University Press.
- CoE 2018a = Council of Europe, *Common European Framework of Reference for Languages: Learning, Teaching, Assessment. Companion Volume with new descriptors*, Strasbourg, Council of Europe.
- CoE 2018b = Council of Europe, *Collated representative samples of descriptors of language competences developed for young learners*, Strasbourg, Council of Europe.
- Cronbach/Meehl 1955 = Lee J. Cronbach / Paul E. Meehl, *Construct validity in psychological tests*, in «Psychological Bulletin», 52, pp. 281-302.
- Cumming 1996 = Alister H. Cumming, *Introduction: the concept of validation in language testing*, in Alister H. Cumming / Richard Berwick (ed), *Validation in language testing*, Clevedon, Multilingual Matters, pp. 1-14.
- Fulcher/Davidson 2007 = Glenn Fulcher / Fred Davidson, *Language Testing and Assessment. An advanced resource book*, Routledge Applied Linguistics.
- Green 2013 = Anthony Green, *Exploring Language Assessment and Testing: Language in Action*, London, Routledge.
- Hughes 1989 = Arthur Hughes, *Testing for Language Teachers*, Cambridge, Cambridge University Press.
- Kunnan 2008 = Antony J. Kunnan, *Towards a model of test evaluation: Using the Test Fairness and Test Context Frameworks*, in Taylor Lynda / Cyril J. Weir (eds.), *Multilingualism and Assessment: Achieving Transparency, Assuring Quality, Sustaining Diversity: Proceedings of*

- the ALTE Berlin Conference*, May 2005, Cambridge, UCLES, Cambridge University Press, pp. 229-251.
- Mariani 2014 = Luciano Mariani, *Il Quadro Comune Europeo di Riferimento e la sua valenza formativa*, in Adriana Arcuri / Egle Mocciano (a cura di), *Verso una didattica linguistica riflessiva. Percorsi di formazione iniziale per insegnanti di italiano lingua non materna*, Palermo, Università degli Studi ([https://www.researchgate.net/publication/268981394\\_Il\\_Quadro\\_Comune\\_Europeo\\_di\\_Riferimento\\_e\\_la\\_sua\\_valenza\\_formativa](https://www.researchgate.net/publication/268981394_Il_Quadro_Comune_Europeo_di_Riferimento_e_la_sua_valenza_formativa)).
- McNamara/Roever 2006 = Tim McNamara / Carsten Roever, *Language testing: the social dimension*, London, Wiley-Blackwell.
- Messick 1989 = Samuel Messick, *Validity*, in Robert L. Linn (ed.), *Educational measurement* (3<sup>rd</sup> ed.), New York, NY: American Council on education and Macmillan, pp. 13-104.
- Messick 1996 = Samuel Messick, *Validity and Washback in Language Testing*, in «Language Testing», 13 (3), pp. 241-256.
- McNamara *et al.* 2019 = Tim McNamara / Ute Knoch / Jason Fan, *Fairness, Justice and Language Assessment*, Oxford, Oxford University Press.
- Palermo 2017 = Massimo Palermo, *Italiano scritto 2.0. Testi e ipertesti*, Roma, Carocci.
- Shohamy 2001 = Elana Shohamy, *The power of tests: A critical perspective on the uses of language tests*, Harlow, Longman.
- Sillabo CILS = Documento ad uso interno del Centro CILS, *Produzione Scritta B1 Adolescenti*, pp. 1-3.
- Weir 2005 = Cyril J. Weir, *Language testing and validation: an evidence based approach*, Palgrave Macmillan.

UNIVERSITÀ PER STRANIERI DI SIENA  
ESAME CILS LIVELLO B1 - modulo adolescenti

**PRODUZIONE SCRITTA - prova N.1**

Parla di un libro che hai letto e che hai trovato interessante. Lo consiglieresti a un amico? Spiega perché. Devi scrivere da 100 a 120 parole.

fig. 1. Esempio prova 1.

UNIVERSITÀ PER STRANIERI DI SIENA  
ESAME CILS LIVELLO B1 - modulo adolescenti

**PRODUZIONE SCRITTA - prova N.2**

Un tuo amico ti ha scritto una mail per informarti che vuole organizzare una festa prima delle vacanze. Rispondi al tuo amico con una mail dove:

- accetti di aiutarlo
- dai alcuni consigli sull'organizzazione della festa (luogo, cibo, musica...).

Devi scrivere da 50 a 80 parole.

fig. 2. Esempio prova 2.

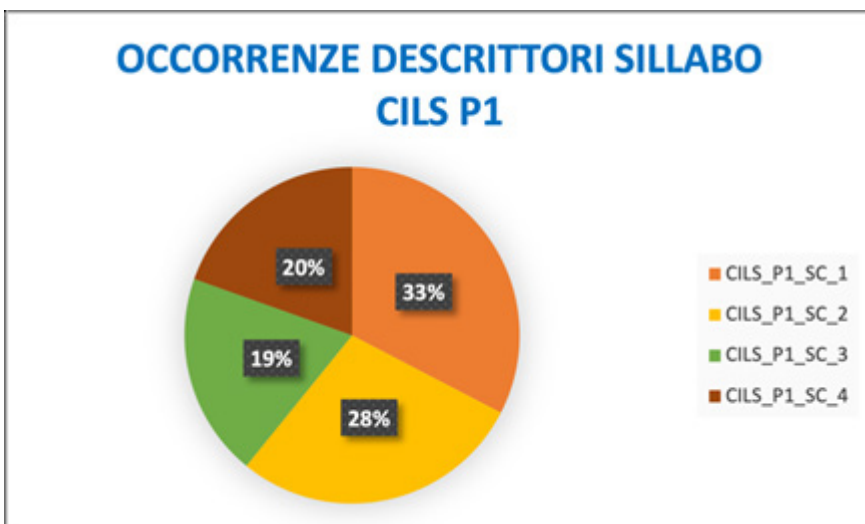


fig. 3. Rappresentatività descrittori Sillabo CILS per la prova 1 (% occorrenze).

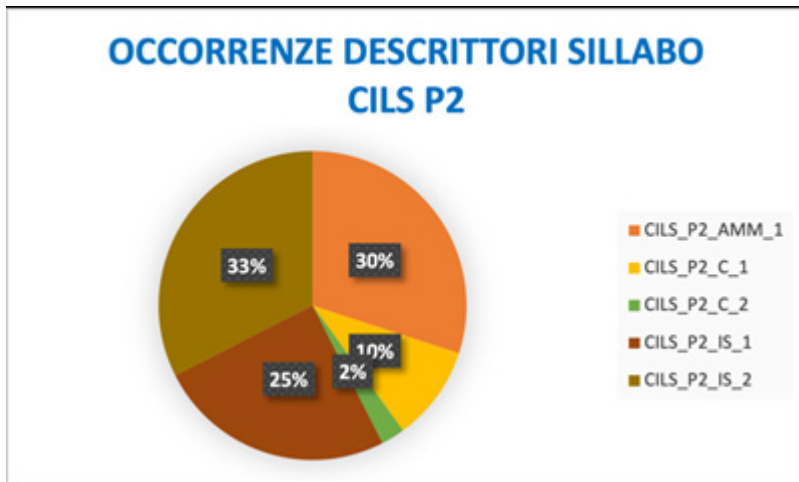


fig. 4. Rappresentatività descrittori Sillabo CILS per la prova 2 (% occorrenze).

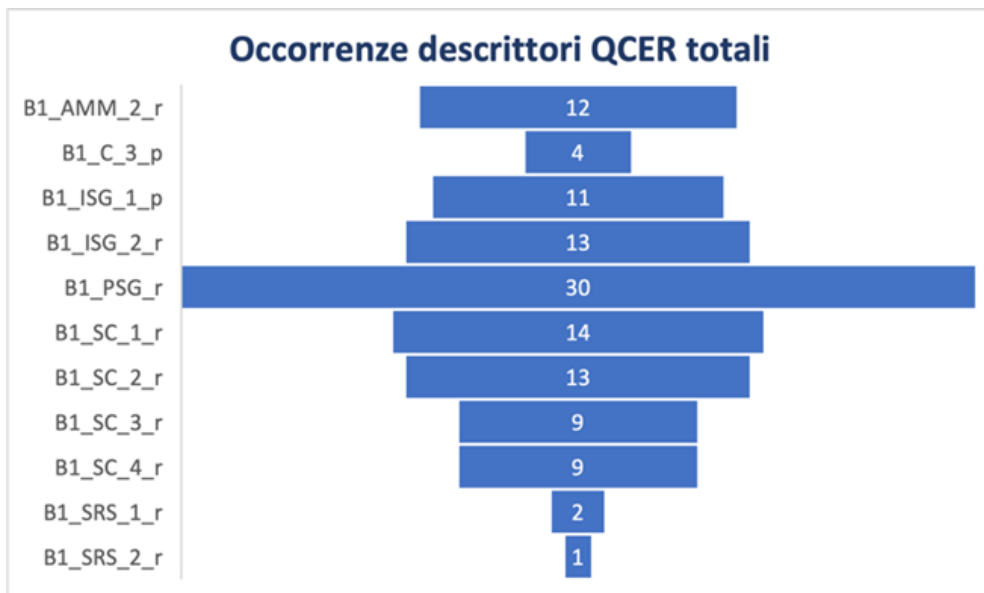


Fig. 5. Rappresentatività descrittori QCER per gli input in totale (nr. occorrenze).

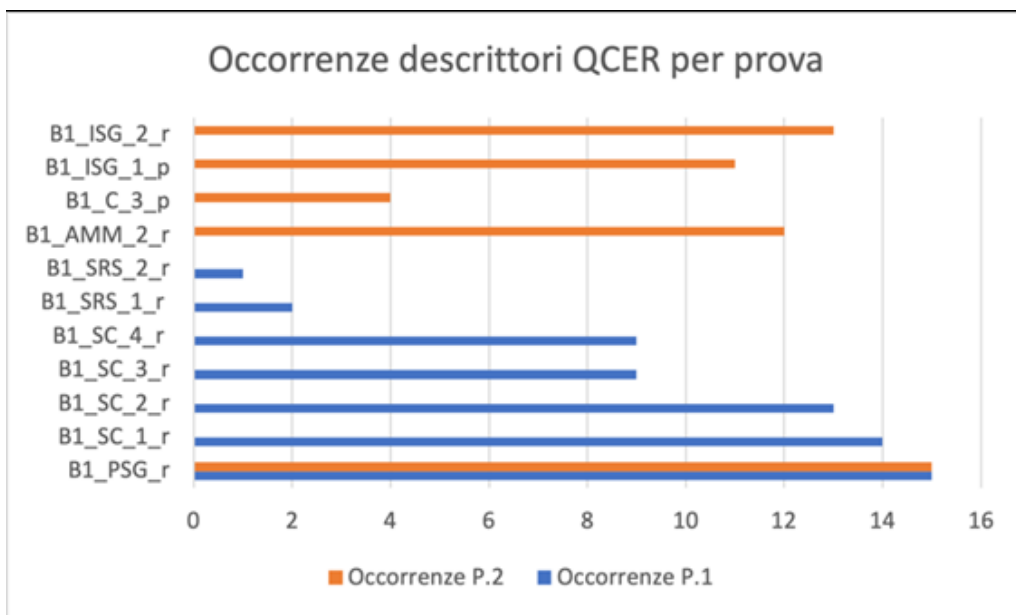


fig. 6. Rappresentatività descrittori QCER per prova (nr. occorrenze).



fig. 7. Occorrenze domini totali.

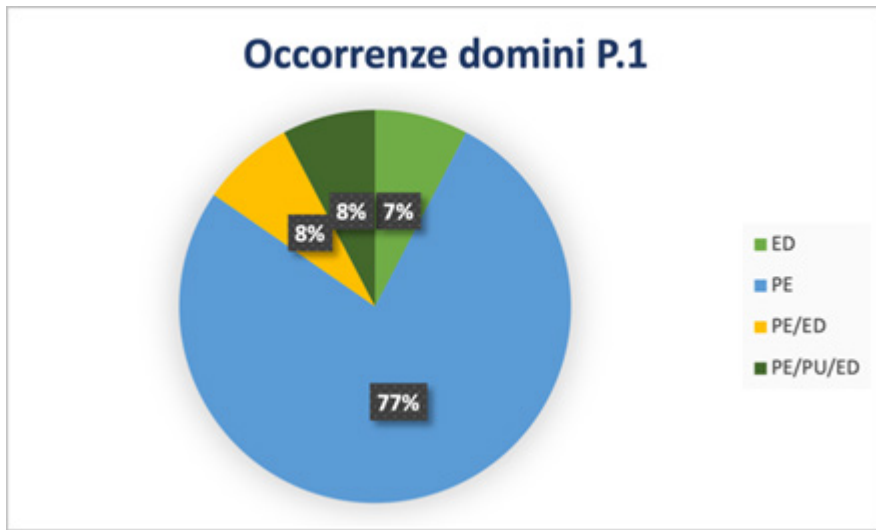


fig. 8 Rappresentatività domini per prova 1 (% occorrenze).

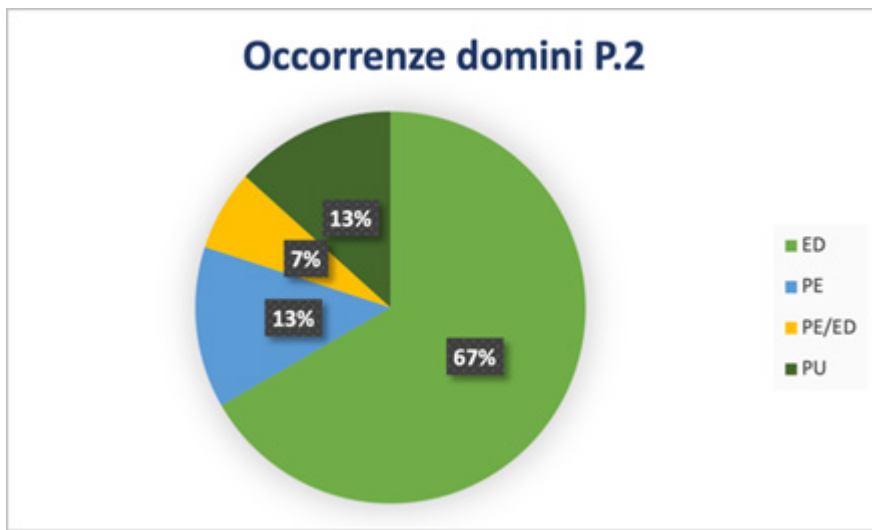


fig. 9 Rappresentatività domini per prova 2 (% occorrenze).