

III. Test linguistici per la coesione sociale e la cittadinanza. Linee guida per politici e istituzioni.

(Versione in italiano a cura di Giuliana Grego Bolli e Lorenzo Rocca)

Il documento è stato redatto con il contributo di:

Andrew Balch

Shalini Roppe

Michael Corrigan

Heinrich Rübeling

Sara Gysen

Steffi Steiner

Henk Kuijper

Piet Van Avermaet

Michaela Perlmann-Balme

Beate Zeidler

e con il supporto del gruppo di studio del LAMI (Language Assessment for Migration and Integration)

1. Introduzione

2. Decidere cosa testare

2.1 Panoramica

2.2 Determinare l'obiettivo delle prove e ciò che il 'mondo reale' richiede ai candidati in termini di 'saper fare'

2.3 Determinare i bisogni linguistici

2.4 Determinare il livello di difficoltà

2.5 Produrre le specificazioni delle prove

3. Garantire che le specificazioni del test vengano applicate

3.1 Panoramica

3.2 Criteri di valutazione e formato delle prove

3.3 Produzione degli item

3.4 Pretesting

3.5 Somministrazione delle prove

3.6 Trattamento dei candidati con particolari esigenze

3.7 Attribuzione del punteggio e classificazione

3.8 Monitoraggio

3.8.1 Monitoraggio delle risposte

3.8.2 Monitoraggio dei bias

4. Conclusioni

1. Introduzione

Molti paesi europei stanno introducendo o formalizzando requisiti linguistici ai fini dell'immigrazione, della residenza e della cittadinanza. I governi nazionali sempre più frequentemente richiedono di poter utilizzare test linguistici o altre procedure formali di valutazione.

Lo scopo di questo lavoro non è promuovere l'uso di test linguistici, quanto piuttosto offrire un supporto professionale, basato su principi di buona prassi, laddove tali test siano obbligatori o in procinto di diventarlo, al fine di garantire che le esigenze degli utenti siano soddisfatte e le prove¹ siano eque per i candidati. L'equità delle prove è una caratteristica particolarmente importante proprio nel caso in cui esse siano legate a contesti di immigrazione, al conferimento della residenza o della cittadinanza: prove non eque possono essere infatti considerate lesive dei diritti umani e civili degli immigrati stessi.

Standard facilmente accessibili (v. *ulteriori riferimenti*) forniscono linee guida per lo sviluppo e la gestione dell'equità delle prove. Unitamente al contenuto di questo capitolo, possono risultare particolarmente utili per meglio comprendere come diversi aspetti del processo di produzione di prove linguistiche risultino necessariamente legati ad un quadro etico.

Laddove si richieda una valutazione delle competenze linguistiche, chi è chiamato a prendere decisioni politiche è invitato a considerare con grande attenzione i punti qui di seguito elencati in forma di domanda:

- potrebbe essere più opportuno utilizzare una forma di verifica diversa da un test?
- potrebbe essere più opportuno utilizzare più metodi di verifica?
- quale uso verrà fatto dei risultati?
- quali potrebbero essere le ricadute sociali?
- quale sarà l'impatto sull'immigrato?
- quale sarà l'impatto sulla società di origine dell'immigrato?

Modalità di valutazione

Relativamente alla prima e alla seconda domanda, le autorità responsabili e i politici devono essere consapevoli dell'esistenza di metodi di verifica diversi (dai test) che potrebbero risultare ugualmente appropriati. Sia i test, sia altri metodi di verifica, hanno dei vantaggi, relativamente a caratteristiche specifiche quali l'incidenza sul candidato, l'interpretabilità, la standardizzazione e affidabilità dei risultati, il costo e la praticabilità. È di conseguenza importante che ogni possibile variabile venga considerata con attenzione, al fine di individuare il tipo di valutazione più opportuna. Va inoltre considerato che una combinazione di più metodi di verifica è comunque sempre possibile.

¹ Nel capitolo i termini: test e prova vengono utilizzati alternativamente come sinonimi.

Test correttamente progettati, costruiti e somministrati presentano i seguenti vantaggi:

- i risultati sono standardizzati ed affidabili. Ciò significa che è facile mettere a confronto le performance dei candidati attraverso somministrazioni diverse;
- i candidati vengono valutati, per quanto umanamente possibile, in modo indipendente e oggettivo;
- può essere verificata la competenza di un numero elevato di persone in breve tempo.

In alternativa ai test, si potrebbe pensare ad una forma di valutazione continua durante tutto un corso, o alla valutazione di forme diverse di evidenza della competenza di un candidato durante il corso. Qualora la valutazione sia destinata ad avere una forte valenza formativa, può diventare parte integrante di un corso, aiutando gli studenti ad orientare il proprio apprendimento.

La valutazione può comportare anche l'auto-valutazione, l'osservazione e la valutazione di gruppo.

Oltre al potenziale formativo di tale tipo di valutazione, i vantaggi di un simile approccio possono essere molteplici:

- la valutazione può avvenire in condizioni di minor stress (ad esempio, in classe); ciò può migliorarne la validità, fornendo evidenza delle vere capacità di un candidato;
- tale evidenza può, a sua volta, derivare dall'esecuzione di compiti che riflettano situazioni di vita reale, identificati come particolarmente importanti per immigranti e cittadini;
- la valutazione della performance di un candidato può essere globale, basata sull'abilità sottesa di portare a termine con successo determinati compiti, ponendo meno l'accento sul controllo di elementi linguistici isolati.

Impatto

Nell'affrontare tematiche legate all'impatto di qualsiasi test, va chiarito in via preliminare che nei risultati di un test ci può essere comunque un margine di errore. L'obiettivo di questo capitolo è quello di aiutare a *minimizzare* la possibilità di conseguenze negative derivanti dalle prove, riducendo al massimo tale margine di errore. Va aggiunto inoltre che test su larga scala, non possono facilmente prendere in considerazione caratteristiche personali, percorsi individuali di apprendimento e la storia personale di ogni singolo candidato al momento della verifica e valutazione delle sue capacità.

Una volta selezionata la modalità di verifica, è molto importante considerarne l'uso e le possibili conseguenze. Queste ultime infatti possono risultare profonde e

imprevedibili e, proprio per questo, devono essere attentamente valutate nella pianificazione di un processo di valutazione.

Le conseguenze di cui sopra, possono comportare anche possibili modifiche all'insegnamento e alle pratiche di apprendimento o, ancora, cambiamenti nel sistema di istruzione dei Paesi degli immigrati.

Dopo aver esaminato tutti gli aspetti precedentemente accennati, se si ritiene di dover utilizzare dei test linguistici, è necessario che tutti i soggetti coinvolti, comprese le autorità responsabili e i politici, si assicurino che il test sia stato elaborato in modo tale da soddisfare i bisogni individuati e che funzioni nel modo previsto. Ciò al fine di consentire che le decisioni politiche che ne conseguono possano essere applicate in modo appropriato ed equo.

Il concetto di equità è di fondamentale importanza per tutti i tipi di test linguistici e per tutti i candidati ma, come sopra indicato, lo diventa ancor più in contesto di immigrazione, residenza e cittadinanza, a causa delle profonde implicazioni in termini di diritti civili ed umani. Il lavoro che viene fatto per garantire che un test sia equo avrà inizio durante la fase di pianificazione e dovrà necessariamente continuare per tutto il processo ciclico di elaborazione e somministrazione del test. Tale lavoro consentirà agli utenti di interpretare e utilizzare correttamente i risultati del test stesso. Per aiutare in tal senso autorità e politici, nell'ultima parte di questo capitolo si forniranno indicazioni in merito a tutte le fasi dello sviluppo e della somministrazione del test, alla selezione degli item e all'interpretazione dei risultati, al fine di fornire una guida alla produzione delle prove a livello nazionale/locale .

Come avviene con tutte le prove, il risultato ottenuto grazie all'applicazione di una corretta prassi, descritta nel presente capitolo, contribuirà a garantire che vengano testate le necessarie competenze e le conoscenze, in modo costante per tutti i candidati e in tutte le versioni del test, per garantirne l'affidabilità.

Ulteriori riferimenti:

[ALTE Code of Practice](http://www.alte.org/quality_assurance/index.php) – http://www.alte.org/quality_assurance/index.php

Multilingual Glossary of Language Testing Terms, Studies in Language Testing volume 6, Cambridge University Press (ISBN: 0-521-65877-2)

[ILTA Code of Ethics](http://www.iltaonline.com/code.pdf) – <http://www.iltaonline.com/code.pdf>

[JCTP Code of Fair Testing Practice in Education](http://www.apa.org/science/FinalCode.pdf) – <http://www.apa.org/science/FinalCode.pdf>

2. Decidere cosa testare

Panoramica

In questa sezione, vengono prese in esame le fasi del processo di produzione di test destinate a garantire che il test risponda in modo adeguato all'obiettivo per cui è stato costruito. La prima fase di tale processo è l'identificazione dell'obiettivo del test; segue la determinazione del contenuto e del livello di difficoltà; si passa infine alla stesura delle specificazioni del test, essenziali per le successive fasi di costruzione e revisione del test stesso.

Determinare l'obiettivo del test e ciò che il mondo reale richiede ai candidati

Prima di elaborare qualsiasi test di lingua, è necessario stabilirne l'obiettivo. Non è sufficiente affermare che il test sia funzionale all'immigrazione o alla concessione della cittadinanza in quanto, anche in tale ambito, sono diversi i motivi per cui gli immigranti possono essere sottoposti ad un test: dalla semplice motivazione (per aiutarli a usare e a migliorare la loro attuale competenza nella lingua target), all'accertare se la competenza di cui dispongono sia sufficiente per partecipare a situazioni sociali ben definite (ad esempio, di studio o di lavoro), al gestire con maggior consapevolezza decisioni relative ai propri diritti, al progetto migratorio legato al Paese che li ospita o all'acquisizione della cittadinanza.

Un obiettivo chiaro ed esplicito non solo aiuterà a chiarire le aspettative dei partecipanti, contribuendo all'equità complessiva della prova, ma consentirà anche agli altri attori sociali di interpretare e di utilizzare i risultati dei test in modo appropriato. Questo processo di definizione delle esigenze viene chiamato *analisi dei bisogni*.

Nell'analizzare i bisogni, è necessario tener conto del fatto che ci sono vari sottogruppi di immigrati con necessità specifiche. Sarebbe bene che chi elabora prove linguistiche definisca i contesti rilevanti e le diverse condizioni del gruppo target. Nel pianificare l'analisi dei bisogni i politici dovrebbero, a loro volta, essere sicuri di aver raccolto risorse e dati sufficienti coinvolgendo rappresentanti di diversi settori della società.

Determinare le competenze linguistiche richieste

Una volta identificati, tali bisogni devono essere tradotti in requisiti linguistici, specificando non solo le conoscenze e le competenze, ma anche il livello che possibilmente ogni candidato alle prove dovrebbe avere. Se, per esempio, le prove sono state progettate per valutare se il candidato abbia le competenze linguistiche

necessarie per seguire un corso di formazione professionale, ci si può ragionevolmente aspettare che verifichino la capacità di seguire lezioni e seminari, di comunicare con insegnanti e compagni di studio, di leggere la letteratura pertinente, ecc. Tale tipo di analisi contribuisce a determinare il livello di competenza necessario(o i livelli richiesti per ciascuna delle singole abilità testate). Se, per fare un altro esempio, i test di lingua sono stati progettati per certificare le competenze di traduttori o interpreti, l'analisi dei bisogni relativa all'esercizio di tali professioni, dovrebbe aver dimostrato che il livello di competenza richiesto deve essere elevato; dovrebbe essere inoltre emerso che, nel caso dei traduttori, il livello di competenza in lettura e scrittura dovrà essere maggiore, laddove nel caso degli interpreti dovrebbe risultare più importante la competenza orale. Nel caso degli immigranti e dei potenziali candidati di test per la cittadinanza, il rapporto tra le competenze linguistiche richieste nella lingua ufficiale e la capacità di integrarsi nella società e / o esercitare i diritti connessi alla cittadinanza è molto più difficile da definire. Dopo tutto, se le competenze linguistiche fossero l'unico fattore in gioco, tutti gli abitanti nativi di un Paese dovrebbero essere dei cittadini perfettamente integrati. Dato che questo non avviene, si può dedurre che anche altri fattori svolgono un ruolo importante.

Esami di lingua per motivi di studio e di lavoro, nella maggior parte dei casi, sono somministrati a gruppi di candidati omogenei per quanto riguarda la formazione e le capacità cognitive, mentre i test per l'integrazione e la cittadinanza (test cioè per l'acquisizione di diritti civili) devono soddisfare una vasta tipologia di possibili candidati, e devono perciò essere accessibili sia alle persone con scarsa o del tutto assente alfabetizzazione, sia a quelli con un elevato livello di formazione, anche accademica.

Determinare il livello appropriato di difficoltà

Una volta individuate le competenze linguistiche richieste, chi elabora test linguistici dovrebbe cercare di 'mapparle' sui saper fare del Quadro Comune Europeo di Riferimento per le lingue (QCER) che *fornisce una base comune per l'elaborazione di programmi di lingua(sillabi), linee guida per i curricula, esami, libri di testo, ecc in tutta Europa. Esso descrive che cosa gli studenti di lingua devono imparare per essere in grado di utilizzare la lingua per la comunicazione e quali conoscenze e competenze devono sviluppare in modo da essere in grado di agire efficacemente.* QCER, p1

Il Quadro contiene una serie di scale illustrative (per il parlato, la scrittura, la lettura, l'ascolto e l'interazione), che definiscono livelli di abilità, mostrano ciò che un discente *sa fare* ad ogni livello e fanno sì che il progresso possa essere misurato lungo una scala di sei livelli, dal livello A1 al livello C2.

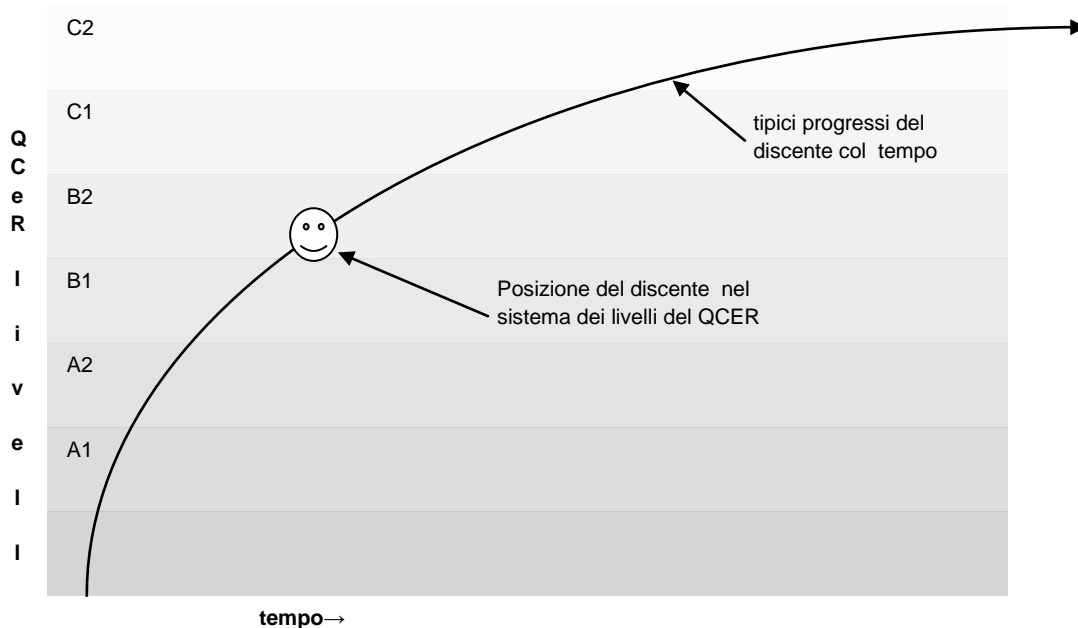


Figura 1. Lo sviluppo nell'apprendimento e i livelli del QCER

La Figura 1 mostra lo sviluppo tipico di uno studente nel corso del tempo in relazione ai livelli del QCER nelle fasce orizzontali. Va osservato che la distanza fra un livello e l'altro nel diagramma, non implica che un equivalente lasso di tempo sia necessario per passare da un livello al successivo. Per questo motivo, la curva che illustra il tipico progresso di un discente è ripida all'inizio e si appiattisce verso la fine. Ciò è dovuto al fatto che la gamma di competenze aumenta ad ogni livello e, di conseguenza, il tempo necessario per passare da un livello all'altro aumenta progressivamente.

L'analisi dei bisogni, non dovrebbe pertanto essere basata sul numero di ore di studio, ma piuttosto su un'attenta mappatura delle esigenze degli utenti relativamente ai saper fare.

Chi elabora test dovrebbe essere inoltre consapevole del fatto che le competenze della maggior parte dei candidati non sono uniformemente distribuite sulle quattro abilità:

leggere, ascoltare, scrivere e parlare, come suggerisce la Figura 1. E' invece più comune che le competenze nel parlare e nell'ascoltare siano più elevate rispetto a quelle nello scrivere e nel leggere. Il discente rappresentato nel diagramma potrebbe avere, quindi, un profilo *frastagliato* nelle quattro abilità linguistiche, come illustrato nella Figura 2.

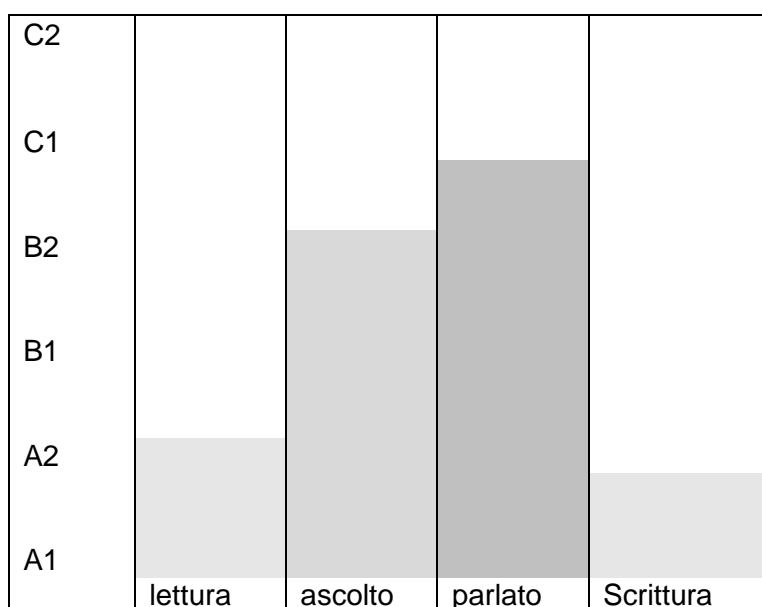


Figura 2 *Profilo frastagliato delle abilità linguistiche di un discente*

Oltre a dare un quadro più preciso delle capacità di un candidato, un approccio modulare alla verifica presenta diversi vantaggi. Se fosse possibile verificare separatamente ogni abilità, i candidati potrebbero sostenere una prova su quelle in cui si sentono più preparati (il che potrebbe essere anche più motivante) piuttosto che su tutte e quattro le abilità in una sola volta. Questa possibilità potrebbe risultare particolarmente adatta per gruppi di immigranti analfabeti o per candidati con poca esperienza di scrittura. Conseguenza di un simile approccio è che i corsi di lingua potrebbero, a loro volta, focalizzarsi sulle abilità da verificare. Nel comunicare i risultati di prove di verifica di competenze parziali, le istituzioni che le organizzano devono chiarire le competenze oggetto di verifica, in modo che non vi sia alcuna confusione fra tali prove e quelle relative a tutte e quattro le abilità. Parallelamente risultati di test o certificati, che riguardano profili di abilità, non dovrebbero riferirsi ad un livello globale, quanto piuttosto al livello raggiunto in ogni singola abilità.

Ulteriori riferimenti:

The [Common European Framework of Reference for Languages: Learning, Teaching, Assessment](http://www.coe.int/lang) – [www.coe.int/lang] http://www.coe.int/t/dg4/linguistic/CADRE_EN.asp

Produrre le specifiche delle prove

La definizione sia della popolazione che dell'obiettivo della prova deve essere documentata nelle specifiche che descrivono, in maniera dettagliata, anche quali tipi di compiti vengono utilizzati, il formato delle prove e altri aspetti pratici. Il tutto deve risultare collegato alla definizione dell'utenza. Le specifiche delle prove si propongono come documento di riferimento per ogni decisione relativa all'uso delle prove stesse.

3. Garantire che le specifiche siano messe in pratica

Panoramica

Una volta definite le specifiche, ulteriori passaggi si rendono necessari per far sì che la prova funzioni come previsto. Devono essere elaborati i criteri di valutazione e sviluppato il formato delle prove; in particolare, gli item devono essere prodotti ed assemblati in conformità alle specifiche.

Una volta prodotta, la prova deve essere somministrata in modo sistematico ed equo. Infine, i dati risultanti dalla somministrazione dovrebbero essere analizzati per confermare che l'esecuzione della prova abbia risposto alle attese.

In tutto questo processo, sono necessari controlli a garanzia della qualità, come descritti in questo documento. Diversi e ulteriori documenti possono risultare utili nell'assistere chi elabora e produce le prove a mettere in pratica quanto indicato dai Codici etici di comportamento elencati qui di seguito.

Ulteriori riferimenti:

[AERA/APA/NCME Standards for Educational and Psychological Testing](http://www.apa.org/science/standards.html) –
<http://www.apa.org/science/standards.html>

[ALTE COP QMS Checklists](http://www.alte.org/quality_assurance/code/checklist.php) – http://www.alte.org/quality_assurance/code/checklist.php

[ALTE Minimum standards for establishing quality profiles in ALTE examinations](http://www.alte.org/quality_assurance/index.php) –
http://www.alte.org/quality_assurance/index.php

[ALTE Principles of Good Practice](http://www.alte.org/quality_assurance/code/good_practice.pdf) – http://www.alte.org/quality_assurance/code/good_practice.pdf

[EALTA Guidelines for Good Practice](http://www.ealta.eu.org/guidelines.htm) – <http://www.ealta.eu.org/guidelines.htm>

[ILTA Draft Code of Practice](http://www.iltaonline.com/ILTA-COP-ver3-21Jun2006.pdf) – <http://www.iltaonline.com/ILTA-COP-ver3-21Jun2006.pdf>

Criteri di valutazione e formato delle prove

L'oggetto della verifica, così come descritto nelle specifiche, dovrà essere suddiviso in singoli, specifici punti in modo da poter essere meglio utilizzato da chi produce gli item. Solo dopo avere chiaramente definito come la performance di un candidato verrà verificata e valutata, potrà iniziare il lavoro effettivo di sviluppo dei compiti da inserire nelle prove, tenendo presente che i candidati devono poter dimostrare di avere la capacità per soddisfare i criteri di valutazioni stabiliti.

Scrivere gli item

Gli *item-writers* devono avere a disposizione linee guida molto chiare che normalmente dovrebbero fornire un quadro generale dell'utenza e chiarire gli obiettivi della prova, dando unitamente consigli di carattere generale, come l'adeguatezza o l'inadeguatezza di alcuni argomenti, la lunghezza degli input (ad esempio il numero di parole in un testo di lettura) e degli output (vale a dire il numero di parole che i candidati devono ad esempio scrivere). Le specifiche devono inoltre chiarire se ed in quale misura i testi sono 'autentici', ecc. Una volta che gli aspetti principali del test sono stati stabiliti, esperti 'esterni' dovrebbero giudicare se le linee guida siano state effettivamente rispettate ed applicate.

Ulteriori riferimenti:

[Item Writer Guidelines](http://www.alte.org/projects/item_writer.php) - http://www.alte.org/projects/item_writer.php

Pretesting

Il giudizio di esperti 'esterni' (di cui al precedente punto) potrebbe di per sé garantire che gli item siano adatti e funzionali agli scopi della verifica, tuttavia per essere certi che gli item funzionino come previsto, (verificando la lingua target, distinguendo i candidati più bravi dai meno bravi e non contenendo *bias* verso nessun tipo di candidati) è necessario pretestare i materiali, nelle stesse condizioni di somministrazione del test finale, su un campione di candidati con un profilo demografico il più possibile analogo a quello della popolazione finale del test.

Il pretest dei test oggettivi (ad esempio, la scelta multipla) viene seguito da una dettagliata analisi statistica, mentre nel caso dei test soggettivi (ad esempio, prove di produzione orale), saranno piuttosto analisi qualitative della produzione dei candidati a determinare se e in quale misura siano state rispettate le aspettative di chi ha

prodotto il test.

Sulla base di tali analisi, item e compiti specifici possono essere accettati per il test dal vivo, corretti e pretestati di nuovo, oppure definitivamente respinti.

Inoltre, nel caso in cui venga creata una banca di item contenente item con caratteristiche anche se di poco diverse, sarà possibile costruire test con caratteristiche ben determinate, in modo tale da coprire una ben determinata serie di livelli.

Nella varie fasi che caratterizzano la produzione degli item e il pretest, è importante che gli esperti rivedano compiti e item. Ciò dovrebbe accadere probabilmente più di una volta, poiché gli item possono essere modificati e rivisti, oppure perché è possibile ottenere nuove informazioni, ad esempio dai risultati del pretest. Alcuni strumenti di supporto in tale opera di revisione sono elencati nella sezione *ulteriori riferimenti* di cui sotto.

Item e compiti, e a volte le stesse risposte dei candidati, possono essere analizzati attraverso tali strumenti in modo da poter stabilire come differiscano l'uno dall'altro e in quale misura i compiti risultino in linea con le specifiche del test.

Le griglie indicate qui di seguito sono state originariamente progettate per aiutare chi produce le prove ad allineare i propri esami al QCER, ma possono essere anche utilizzate con obiettivi leggermente diversi.

Ulteriori riferimenti:

[Content Analysis Checklists](http://www.alte.org/projects/content.php) – <http://www.alte.org/projects/content.php>

Council of Europe: www.coe.int/portfolio

[CEFR Grids for the analysis of test tasks \(listening, reading, speaking and writing\)](http://www.coe.int/T/DG4/Portfolio/?L=E&M=/documents_intro/Manual.html)

[http://www.coe.int/T/DG4/Portfolio/?L=E&M=/documents_intro/Manual.html]

See also: [Illustrations of the European levels of language proficiency](http://www.coe.int/T/DG4/Portfolio/?L=E&M=/main_pages/illustrationse.html)

[http://www.coe.int/T/DG4/Portfolio/?L=E&M=/main_pages/illustrationse.html]

Gestione delle prove

Chi produce le prove deve garantire che vengano somministrate in condizioni eque per tutti i candidati. A tal fine, si raccomanda che le procedure siano progettate in modo da poter ridurre al minimo le differenze nella somministrazione e nella gestione.

Tali procedure devono garantire:

- la disponibilità di centri d'esame accreditati per lo svolgimento delle prove;
- la presenza al loro interno di personale professionale competente, adeguatamente preparato, supportato e monitorato laddove necessario;
- un alto livello di sicurezza e di rispetto della normativa per il trattamento dei dati personali, dall'iscrizione al rilascio dei risultati e dei certificati;
- adeguate condizioni nelle aule degli esami (rumore, temperatura, distanza tra i candidati, ecc.);
- adeguate condizioni per candidati con handicap o con particolari esigenze.

Inserimento di candidati con particolari esigenze

Il sistema di verifica non deve essere discriminatorio nei confronti di candidati con particolari esigenze. Queste possono essere temporanee o permanenti, possono riguardare disturbi mentali, problemi emotivi, o disabilità, disturbi di apprendimento, malattie, analfabetismo in L1 o nella lingua target, regole legate alla religione, situazioni di detenzione o qualsiasi altra circostanza.

Dovrebbero essere approntate misure per:

- decidere se eventuali candidati con particolari esigenze possano essere esentati dal sottoporsi a tutta la prova o a parti di essa;
- prendere misure adeguate al fine di garantire che i candidati con particolari esigenze vengano giudicati in modo equo;
- definire chi debba decidere se un determinato candidato possa svolgere la prova;
- definire quali condizioni applicare caso per caso (ad esempio fascicoli in Braille, stampa a caratteri più grandi, computer con caratteristiche speciali, la presenza di un assistente durante alcune parti della prova, pause supplementari, un interprete del linguaggio dei segni, date o sedi d'esame speciali);
- garantire che sia possibile appellarsi contro decisioni prese in tal senso, informare i candidati come appellarsi ed indicare il modo in cui verrà presa la decisione finale.

Le informazioni relative a tali regole e alle possibili esenzioni devono essere rese pubbliche e accessibili ai candidati.

Ulteriori riferimenti:

"Special Educational Needs in Europe. The Teaching & Learning of Languages. Teaching Languages to Learners with Special Needs", European Commission, DG EAC 23 03 LOT 3, January 2005.

Attribuzione dei punteggi

Item oggettivi possono essere corretti e valutati da macchine (lettori ottici, computer) o da personale appositamente preparato. Item soggettivi richiedono abitualmente personale appositamente preparato per l'attribuzione del punteggio. In tal caso è possibile che si determinino errori nell'espressione del giudizio e nella conseguente attribuzione del punteggio. Tali errori possono riguardare una variabilità nel modo in cui l'esaminatore valuta la risposta del candidato o nel modo in cui l'esaminatore stesso applica i criteri di valutazione. Ciascun esaminatore può divergere nel proprio giudizio rispetto agli altri esaminatori o variarlo, di volta in volta in base all'applicazione dei criteri. Tali incongruenze possono essere notevolmente ridotte sia con una formazione rigorosa e l'applicazione di adeguate procedure di controllo, sia facendo valutare alcune o tutte le produzioni dei candidati da più esaminatori. E' possibile tenere conto dell'eccessiva severità o 'bontà' degli esaminatori, certo è che quando le differenze, in termini di comportamento, diventano eccessive, è necessario prevedere una formazione ulteriore.

Monitoraggio

Così come è importante monitorare il comportamento degli esaminatori, è altrettanto importante che chi elabora test linguistici raccolga e analizzi informazioni sia sulle risposte dei candidati, sia sui candidati stessi: età, sesso, nazionalità, ecc. Tutto ciò è necessario per garantire che:

- ogni prova misuri le abilità che si prefigge di misurare;
- le abilità siano misurate in modo stabile attraverso somministrazioni diverse della stessa tipologia di prova;
- ogni prova funzioni in un modo equo per tutti i partecipanti target al test, indipendentemente dal loro background.

I risultati di un tale monitoraggio possono essere usati per garantire che i risultati del test definiscano l'abilità del candidato. Inoltre, conclusioni emerse dall'analisi possono essere utilizzate nella costruzione della prova, nella sua amministrazione e nel processo di attribuzione dei punteggi, in modo tale che il processo complessivo possa essere costantemente migliorato.

Monitoraggio delle risposte dei candidati

Durante le prove dal vivo, le risposte dei candidati agli item e ai compiti vengono utilizzate per fornire sia il punteggio grezzo, sia le informazioni a chi elabora il test su come compiti e singoli item abbiano funzionato nel misurare la abilità dei candidati.

Gli item dovrebbero essere giudicati non solo per il loro livello di difficoltà, ma anche per come discriminano i candidati ‘forti’ dai candidati ‘deboli’ (dato che l'obiettivo fondamentale della prova è quello di distinguere questi due gruppi).

I dati di queste e di altre analisi statistiche dovrebbero essere conservati e dovrebbe essere effettuato un confronto con quelli emersi da versioni passate e future dello stesso tipo di prova. Ciò consentirà a chi produce le prove di garantire che i risultati di una versione o sessione siano comparabili con quelli di altre.

Laddove la costruzione della prova è basata sul pretest, è probabile che compiti e item funzioneranno come previsto nelle prove dal vivo. Tuttavia, la conferma di ciò attraverso un’analisi diretta è importante e potrebbero essere necessarie indagini ulteriori, nonché una correzione del punteggio, in particolare nel caso in cui le performance non si dimostrino in linea con le aspettative di partenza. E’anche importante che una particolare versione di test dia una rappresentazione dell’abilità dei candidati che risulti coerente con versioni passate e future del test. Per questo motivo, i dati delle risposte dal vivo sono spesso utilizzati per stabilire punteggi di taglio e fasce di punteggi. Se un candidato si sottoponesse a due diverse versioni di una stessa prova, è molto improbabile che riuscirebbe ad ottenere lo stesso punteggio grezzo. Tuttavia è possibile, per chi elabora il test, garantire con ragionevole certezza che il candidato potrebbe ottenere lo stesso grado (inteso come fascia di punteggi), o avere un punteggio di taglio che cada nello stesso ‘range’ in entrambe le prove. Questo può essere fatto attraverso un processo di ‘*equating*’ fra un test e l’altro, stabilendo quale punteggio nella versione B equivalga al punteggio necessario per passare la versione A. Laddove sia stato effettuato il pretest, tale valutazione potrà risultare più facile e probabilmente più accurata.

Monitoraggio dei bias

Tecniche simili a quelle delineate al punto precedente dovrebbero essere applicate anche alle performance di gruppi di candidati con stessi livelli di abilità. Se, per esempio, candidati di una data nazionalità, messi a confronto con candidati con la

stessa abilità, ma di nazionalità diversa sono risultati significativamente migliori o peggiori in un compito o in una serie di compiti, ciò può essere dovuto a bias operati nei confronti di candidati appartenenti a quella particolare nazionalità, determinando un'ingiustizia nei loro confronti. Tuttavia, la causa può essere attribuita anche a differenze / analogie linguistiche tra lingua madre e lingua target, il che non può essere considerato ingiusto.

In entrambi i casi, dopo che un'analisi quantitativa ha rivelato possibili bias, è necessaria un'analisi qualitativa. Se un item dovesse risultare affetto da bias, questo richiederebbe la sua rimozione e / o modifiche alle procedure in base alle quali è stato prodotto.

Conclusioni

L'utilizzo dei test per l'immigrazione e la cittadinanza è notevolmente più complesso di quanto possa sembrare a prima vista. Questo documento ha cercato di delineare gli aspetti problematici da prendere in considerazione, quelli sui quali i politici sono chiamati ad assumere le proprie responsabilità.

In primo luogo vanno affrontati problemi relativi a quale tipo di valutazione sia necessaria sulla base degli obiettivi prefissati e a cosa debba essere verificato. Laddove si sia deciso di utilizzare un test, è di vitale importanza che questo soddisfi i requisiti delineati in questo documento. La prova deve essere continuamente monitorata, al fine di garantirne funzionamento e qualità. Non va dimenticato, inoltre, che i risultati di un test possono avere importanti conseguenze sia per i candidati, sia per la società nel suo complesso. Tra queste conseguenze ci sono anche quelle relative ai diritti civili e umani dei candidati ai test. Per un uso corretto di test di lingua in contesto di immigrazione e di concessione della cittadinanza, coloro che prendono decisioni politiche devono lavorare insieme a chi produce test su vari aspetti, quali la definizione dello scopo della prova e dell'uso che ne verrà fatto, nonché la ripartizione delle risorse necessarie all'espletamento di tutte le fasi del processo di sviluppo di un test. In ogni momento, l'equità della prova deve essere considerata di primaria importanza.