

# ANALISI DELLA PREDITTIVITÀ DI ITEM TRATTI DA PROVE DI COMPrensIONE DELLA CERTIFICAZIONE. IT PER LA FASCIA DI COMPETENZA DELL'AUTONOMIA

## ANÁLISE DE PREVISIBILIDADE DE ITENS EM PROVAS DE COMPREENSÃO DE CERTIFICAÇÃO LINGUÍSTICA. IT PARA A FAIXA DE COMPETÊNCIA DE AUTONOMIA

**Paolo Torresan**

---

### **RIASSUNTO**

Nell'articolo introduciamo un valore percentuale (*indice di predittività*) adottabile nell'ambito dei processi di validazione di un test di competenza (*proficiency test*). L'indice ci informa sul grado di predittività di *item* di prove di comprensione esterni al testo oggetto di comprensione (i test di riordino sono esclusi). L'analisi si conduce somministrando gli *item* a un campione rappresentativo, senza fornire però i testi di ascolto/lettura corrispondenti. Al campione vien detto di provare a rispondere affidandosi alla logica e all'intuito. Orientativamente, abbiamo stabilito che un valore superiore a 50 indichi margini di predittività significativi e costringa, di conseguenza, l'*item writer* a ragionare sulla/e possibile/i causa/e che facilitano la risposta immediata. Nel saggio, il calcolo dell'*indice di predittività* viene applicato a 8 prove di comprensione, quattro di livello B1 e quattro di livello B2, tratte dalla certificazione di italiano per stranieri .IT, dell'Università di Studi Roma Tre (Roma), per un totale di 49 *item*. Le conclusioni cui giungiamo è che l'indice, secondo i parametri espressi, spieghi fenomeni di forte predittività, ascrivibili a diversi fattori (ovvietà delle risposte, facilità a rispondere sulla base di preconcoscenze specifiche, correlazioni tra *item*, implausibilità dei distrattori). In due casi si sono evidenziati *item* con predittività significativa pur se con un indice appena al di sotto della soglia convenzionalmente definita; in un'altra circostanza, invece, un *item* presenta un'innegabile predittività pur con un indice piuttosto basso (è un caso specifico di *predittività regressiva*).

**Parole-chiave:** *testing* linguistico, validazione degli *item*, prove di comprensione.

### **RESUMO**

Introduzimos o artigo com o valor percentual (índice de previsibilidade) adotado no âmbito dos processos de validação de um teste de competência linguística (*proficiency test*). O índice, em especial, nos informa o grau de previsibilidade das respostas a quesitos de provas de

---

\* Centro di ricerca sulla didattica delle lingue, Università Ca' Foscari, Venezia, Italia. email: esh@unive.it

compreensão com *itens* externos ao texto de leitura (estão excluídos os testes de reorganização textual). A análise é conduzida aplicando os itens em um grupo representativo, sem fornecer, entretanto, os testes auditivo e de leitura correspondentes. Pede-se ao grupo para tentar responder confiando na lógica e na intuição. Estabelecemos, como orientação, que um valor superior a 50 indique margens de previsibilidade significativas e obrigue, conseqüentemente, que o *item writer* raciocine sobre a/s possível/eis causa/s que facilitam a resposta imediata. No artigo, o cálculo do *índice de previsibilidade* é aplicado a 8 provas de compreensão, quatro de nível B1 e quatro de nível B2, extraídas de uma certificação linguística de italiano para estrangeiros, .IT, da Universidade Roma Tre (Roma), com um total de 49 *itens*. Concluímos que o índice, segundo os parâmetros expressos, explica fenômenos de forte previsibilidade, atribuídos a diversos fatores (obviedade das respostas, facilidade para responder a partir de base de pré-conhecimento específico, correlações entre *itens*, implausibilidade dos distratores). Em dois casos evidenciaram-se *itens* com previsibilidade significativa, embora com um índice um pouco abaixo do limite convencionalmente definido; em outra circunstância, ao contrário, um *item* apresenta uma inegável previsibilidade, embora com um índice mais baixo (é um caso específico de *previsibilidade regressiva*).

**Palavras-chave:** *testing* linguístico; validação de *itens*; provas de compreensão.

## 1. L'INDICE DI PREDITTIVITÀ

Per *predittività delle risposte* ci riferiamo al fatto che a *item* di prove di lettura/ascolto si possa rispondere correttamente pur senza aver letto/ascoltato il brano oggetto di comprensione (cfr. POWERS, WILSON 1993). Si tratta di un fattore che agisce a detrimento della validità del test, dal momento che la risposta corretta, in sede di esame, può non dipendere dall'abilità oggetto di valutazione (lettura/ascolto), quanto, invece, da *fattori* esterni. Tra essi citiamo:

- il *senso comune*, nel caso di *ovvietà* (OV), del tipo : “Il sole è: a) una stella, b) un pianeta, c) un satellite”; per rispondere al quesito, non occorre essere esperti di astrofisica
- *preconoscenze del soggetto* (PC); si tratta di risposte meno scontate, cui si arriva per via di conoscenze condivise tra quanti coltivano un certo interesse o più in generale sono informati su un determinato argomento, del tipo “Dove nasce il melodramma?”: chi è appassionato di storia e cultura italiana o di canto lirico vi può rispondere. Altro esempio “In che squadra gioca Messi?”; in generale, si suppone che le conoscenze calcistiche in possesso favoriscano i maschi rispetto alle femmine
- cattivo *design* della prova, e in particolare:
  - o risposta inferibile sulla base di *item* precedenti/successivi, ovvero *correlazione* (CR, o *interdipendenza*) che può essere

- *forte*, nel caso la chiave si presenti automatica e ogni dubbio sia fugato
- *debole*, nel caso in cui i rimandi di altri *item* rendono altamente probabile, ma non assolutamente certa, la risposta
- risposta inferibile sulla base del *layout* (*risalto*: RS); in un quesito a scelta multipla, gli *item writer* con meno esperienza spesso curano con maggiore dettaglio la chiave, rispetto ai descrittori, con il risultato che la prima *risalta* sui secondi per lunghezza e/o per ricercatezza del lessico o complessità sintattica; lo studente strategico, pur non sapendo rispondere, può scegliere, a ragione, quest'opzione rispetto alle altre
- *implausibilità dei distrattori* (IM); i distrattori si autoconfutano tra loro o vengono negati per via dei rimandi interni tra gli *item*; in entrambi i casi, l'opzione rimanente emerge come alternativa possibile. Anche in questa circostanza, possiamo avere un'implausibilità
  - *forte*, con distrattori assurdi e la chiave che si impone come la sola possibile
  - *debole*, quando i distrattori presentano una bassissima probabilità, mentre è notevole la plausibilità della chiave

L'analisi della predittività può essere condotta in due modi:

- mediante il giudizio critico di una persona esterna alla prova. Questi ragiona in termini di logica e avvisa l'*item writer* dei quesiti cui è possibile rispondere intuitivamente
- su base matematica. Si tratta di adottare un indice percentuale, definibile come "indice di predittività" [IP], riferibile alla percentuale di un campione rappresentativo ( $n \geq 100$ , e di competenza linguistica pari alla competenza *target*) che risponde correttamente all'*item*, pur se sprovvisto dei testi di riferimento (audio/scritti)

L'oggetto di questo saggio è la presentazione di questo valore, applicato all'analisi della predittività di prove di comprensione tratte da una certificazione di italiano come LS, .IT, elaborata dall'Università degli Studi Roma Tre.

Convenzionalmente, riteniamo che se tale indice supera il valore di 50, il certificatore è messo in allarme circa possibili *bias*, difetti cioè dell'*item* riconducibili ai casi citati sopra: *ovvietà, preconoscenze, correlazione, risalto, implausibilità dei distrattori*.

L'analisi statistica può integrarsi con quella affidata al puro ragionamento: può mettere in luce aspetti che la seconda non rileva (anche se è vero il contrario,

come avremo modo di vedere); inoltre può restituirci *quanto* sia facile rispondere all'*item* a prescindere del testo di lettura/ascolto.

L'*analisi della predittività* contribuisce, in generale, alla validità di una prova: è un ulteriore elemento a disposizione dell'*item writer* per accertare che la prova presenti la minor approssimazione possibile. In altre parole, in ambito certificatorio, nel ciclo di un test di comprensione, e più in particolare nel processo di validazione, l'*analisi della predittività* costituisce una tappa significativa (SHOHAMY 1985; WEIR 1990; 1993, LYNCH, DAVIDSON 1994; ALDERSON, CLAPHAM, WALL, 1995; McNAMARA 2000; CARR, 2011), come evidenziato nella fig. 1.

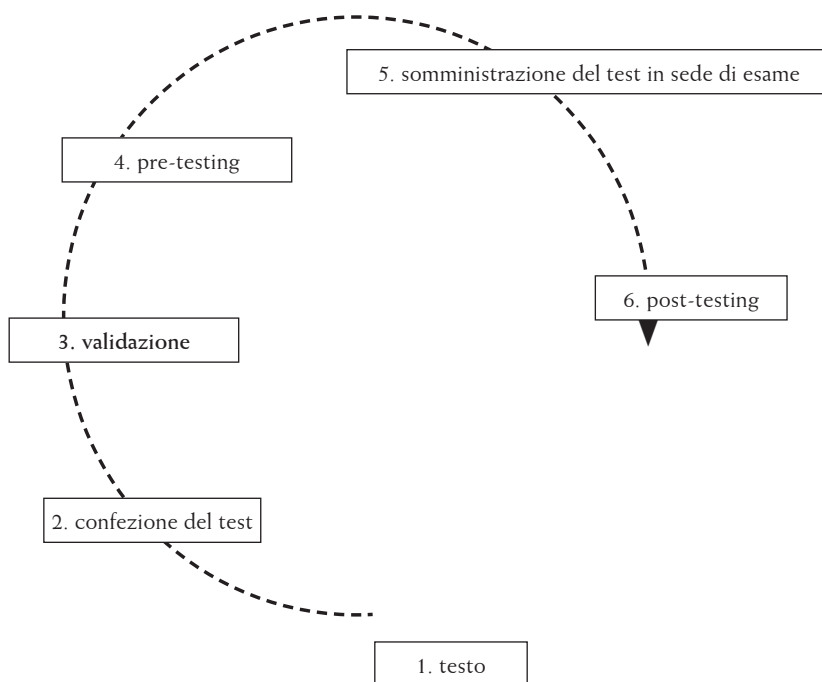


Fig. 1. Il ciclo di un test di comprensione

Nella tab. 1 rappresentiamo in dettaglio fasi e sottofasi del ciclo di un test di comprensione, in sede certificatoria; l'elenco è corredato delle domande a cui un *item writer* deve dar risposta, fase per fase. In neretto illustriamo dove, all'interno della fase di validazione, si colloca l'*analisi della predittività*.

FASE	SOTTOFASI	DOMANDE
1. Testo	Scelta, adattamento o confezione del testo	<ul style="list-style-type: none"> <li>a. Il testo è motivante?</li> <li>b. Il testo è libero da diritti d'autore?</li> <li>c. Il testo è rapportato al livello, sia in termini di dominio che in termini di complessità linguistica? Ho consultato le specifiche circa le caratteristiche previste?</li> <li>d. Il testo rischia di favorire alcuni studenti rispetto ad altri, in termini di pre-conoscenze?</li> <li>e. La lingua è sorvegliata?</li> <li>f. Ci sono dei <i>bias</i> culturali?</li> <li>g. Ci sono frasi o passaggi che fanno riferimento ad altri testi?</li> <li>h. Il testo presenta riferimenti culturali che rischiano di risultare opachi a uno straniero?</li> <li>i. Vengono affrontati argomenti o è impiegato un linguaggio che può suscitare imbarazzo o tensione?</li> <li>j. Ci sono passaggi concettualmente oscuri, imprecisi, contraddittori?</li> <li>k. Se il testo è audio, la registrazione è buona? Ci sono effetti di disturbo che possono compromettere la validità della prova?</li> </ul>
	Controllo	<ul style="list-style-type: none"> <li>l. Eventuali modifiche apportate a una fonte rischiano di compromettere coesione e coerenza?</li> </ul>
2. Confezione del test	<p>Scelta del <i>format</i></p> <p>Individuazione dei nuclei informativi e decisione in merito al tipo di strategie da sollecitare</p>	<ul style="list-style-type: none"> <li>m. Il <i>format</i> è familiare al candidato?</li> <li>n. Il <i>format</i> è valido?</li> <li>o. Il <i>format</i> è affidabile?</li> <li>p. Il <i>format</i> è somministrabile in diversi contesti?</li> <li>q. Ho esperienza di redazione di <i>test</i> simili? Mi sono documentato sulle possibili insidie?</li> <li>r. Considerata la struttura del testo e il tipo di strategie che intendo attivare (comprensione di dettagli, di nuclei concettuali salienti, di inferenze), quali sono le informazioni <i>target</i> verso le quali intendo dirigere l'attenzione del candidato?</li> <li>s. Ci sono immagini che corredano il testo? Hanno una funzione decorativa?</li> <li>t. Le istruzioni sono chiare, sintetiche, efficaci?</li> </ul>

2. Confezione del test	Redazione di istruzioni, scelta del punteggio, definizione di un item esemplificativo, scelta del layout	<ul style="list-style-type: none"> <li>u. Il punteggio è adeguato? Evito punteggi negativi (pedagogicamente e docimologicamente discutibili)?</li> <li>v. Presento un <i>item</i>-esempio?</li> <li>w. Prevedo una rosa il più ampia possibile di chiavi?</li> <li>x. Nel caso di test di ascolto, concedo un tempo sufficiente affinché i candidati leggano gli <i>item</i> prima del primo ascolto, tra il primo e un eventuale secondo ascolto, dopo l'eventuale secondo ascolto?</li> <li>y. In un testo scritto, se il testo è autentico, rispetto il <i>layout</i> della fonte?</li> <li>z. Se il testo è autentico, cito la fonte?</li> </ul>
3. Validazione		<ul style="list-style-type: none"> <li>aa. Mi sono documentato sulle buone pratiche di <i>item writing</i> nello specifico della prova che adottato?</li> <li>bb. Ho evitato doppie chiavi o distrattori parzialmente veri?</li> <li>cc. Ho avviato un'analisi circa la predittività degli <i>item</i>?</li> <li>dd. Ho somministrato la prova a nativi esperti, per raccogliere il loro <i>feedback</i>?</li> <li>ee. Ho accertato che gli <i>item</i> siano calibrati?</li> <li>ff. Ho somministrato la prova a un campione-pilota?</li> <li>gg. Ho evitato una sovrapposizione tra gli <i>item</i> (il fatto cioè che <i>item</i> differenti si riferiscano a una stessa informazione)?</li> <li>hh. Ho evitato <i>item</i> ambigui?</li> </ul>
4. Pre-testing	Analisi statistiche	<ul style="list-style-type: none"> <li>ii. Ho somministrato la prova a un campione rappresentativo, per indagare la difficoltà degli <i>item</i> (p-value) e se e come gli <i>item</i> discriminano (Alpha di Cronbach)?</li> <li>jj. Ho analizzato altri valori statistici, quali: mediana, moda, media, deviazione standard, correlazione punto biseriale?</li> <li>kk. Nel caso di scelta multipla, ho eseguito un'analisi dei distrattori?</li> </ul>

4. Pre-testing	Eventuale riscrittura delle prova	ll. Dall'analisi precedente risultano <i>item</i> da scartare, da riscrivere, da modificare? mm. Se ho agito come descritto sopra, ho ripetuto delle misurazioni per accertare la qualità della prova (un secondo <i>pre-testing</i> ?)
6. Somministrazione delle prove		nn. Mi sono assicurato che la prova sia somministrata in condizioni di sicurezza e di garanzia? oo. Mi sono assicurato che le condizioni logistiche dell'istituzione di riferimento siano tali da non inficiare l'affidabilità dei risultati (es. area urbana rumorosa, rimbombo nelle aule, ecc.)?
7. Post-testing		pp. Una volta gestiti statisticamente i dati relativi alla somministrazione, quali conseguenze posso trarre?

Tab. 1. Fasi e sottofasi relative al ciclo di un test di comprensione

## 2. L'OGGETTO DI STUDIO

Il nostro studio ha per oggetto prove di comprensione della fascia dell'autonomia (livello B, secondo le indicazioni del *Quadro Comune Europeo di Riferimento delle Lingue*; COUNCIL OF EUROPE 2001), tratte dalla certificazione di italiano per stranieri .IT, dell'Università degli Studi Roma Tre.

Si tratta di *quesiti a scelta multipla* a tre opzioni [QSM] e di esercizi di *vero/falso* [VF], mediante i quali il certificatore intende misurare l'abilità di lettura e di ascolto in italiano LS da parte di studenti di livello B1 e di livello B2. Le prove sono attinte dagli esempi caricati sul sito dell'Università (cfr. sitografia).

Il campione di riferimento è costituito da studenti di livello rispettivamente B1 per le prove di livello B1 (n= 120) e di livello B2 per le prove di livello B2 (n=134). Si tratta di studenti di italiano facenti capo a istituzioni scolastiche (in prevalenza università), presenti in diversi contesti linguistici e culturali (Vietnam, Inghilterra, Stati Uniti, Messico, Serbia). Ai colleghi che ci hanno trasmesso i dati va il nostro ringraziamento.

Nella tab. 2 rappresentiamo la struttura delle prove ed evidenziamo, prova per prova, il numero di *item* con IP > 50 (20).

Livello	Format	Abilità	Rimando nell'articolo	Numero item		numero item con IP > 50
				Singole prove	Tot	
B1	QSM	ascolto	§ 3.1.1	5	19	2
		lettura	§ 3.1.2	4		2
	VF	ascolto	§ 3.1.3	5		4
		lettura	§ 3.1.4	5		4
B2	QSM	ascolto	§ 3.1.1	5	20	1
		lettura	§ 3.2.2	5		3
	VF	ascolto	§ 3.2.3	5		1
		lettura	§ 3.2.4	5		3
				49	20	

Tab. 2. Le prove oggetto di studio e gli *item* con IP > 50

Nei paragrafi a seguire realizziamo un'analisi dettagliata di ciascun *item* che presenta un IP > 50.

### 3. ANALISI DELLA PREDITTIVITÀ RIFERITA AGLI ITEM DI PROVE DI COMPrensIONE DELLA CERTIFICAZIONE .IT (LIVELLI DELL'AUTONOMIA)

Consideriamo il comportamento del campione in riferimento alla *predittività* delle risposte per ciascun *item* delle otto prove. Presentiamo una prova alla volta, cominciando da quelle relative al livello B1.

#### 3.1. Livello B1

Analizziamo, *in primis*, le prove relative al livello B1.

##### 3.1.1. QSM ascolto

La prova è costituita da cinque quesiti a tre entrate. Viene presentato un *item*-esempio, che indica il genere di testo audio somministrato.



Esempio:

- 0 ~~Hai appena ascoltato...~~  
~~a. ... un servizio radiofonico.~~  
~~b. ... una pubblicità.~~  
~~c. ... un discorso importante.~~

Il tema dell'ascolto, facile da evincere attraverso la lettura degli *item*, è il *turismo naturalistico*. Qui sotto, il lettore ha modo di accedere all'intera prova con le percentuali relative alle opzioni scelte *item* per *item*. Le chiavi sono evidenziate in grassetto.

- 1 Nel testo ascoltato si parla...  
a. ... dei vantaggi e degli svantaggi delle vacanze in natura. (37,3)  
b. ... delle iniziative del WWF Italia per le prossime vacanze. (9,00)  
c. ... dei motivi per scegliere di trascorrere delle **vacanze in natura**. (53,7)
- 2 Secondo Roberto Furlani...  
a. ... l'attività fisica fa bene al nostro cervello. (38,2)  
b. ... la cromoterapia consiglia le vacanze in natura. (30,9)  
c. ... il verde e l'azzurro aiutano a rilassarsi. (30,9)
- 3 Secondo Roberto Furlani, **le vacanze in natura...**  
a. ... sono perfette per chi ha spirito di avventura. (64,0)  
b. ... sono ideali per chi usa la tecnologia. (14,7)  
c. ... non sono adatte ai ragazzi di oggi. (21,3)
- 4 Secondo Roberto Furlani, i giovani d'oggi...  
a. ... sono appassionati di avventura e natura (17,7)  
b. ... conoscono l'avventura solo attraverso i videogiochi. (41,9)  
c. ... prenotano le vacanze con computer, *tablet* o *smartphone*. (40,4)
- 5 Secondo Roberto Furlani, **le vacanze in natura...**  
a. ... ci permettono di visitare ambienti nuovi e inesplorati. (38,9)  
b. ... aumentano la nostra conoscenza su specie animali e ambienti naturali. (41,0)  
c. ... ci danno la possibilità di visitare ambienti che stanno scomparendo. (20,1)

Di seguito analizziamo i due *item* con IP > 50:

- # 1 (IP=53,7)
- # 3 (IP=64,0)

*Item # 1*

Problema: *correlazione*

L'alto IP (53,7) si può spiegare in virtù del rimando interno tra gli *item*. Come si vede attraverso le frecce nella prova riportata sopra, la chiave [c] è facilmente intuibile, dal momento che lo stesso contenuto è dato per scontato (e quindi confermato) negli *stem* degli *item* ## 3 e 5.

Si potrebbe dire, che gli stessi rimandi potrebbero giustificare, in verità, anche il distrattore [a] (in effetti la % di consensi non è esigua: 37,3%); tuttavia, a una lettura attenta dell'intera prova non risulta alcun accenno al tema degli "svantaggi", di cui si dà invece menzione nel distrattore [a], appunto.

*Item # 3*

Problema: *implausibilità dei distrattori; ovvietà*

Il motivo dell'alto IP (64) pare legato al concorso di due fattori: l'implausibilità dei distrattori (e quindi la loro inefficacia) e l'ovvietà della chiave ("le vacanze in natura sono perfette per chi ha spirito di avventura").

### 3.1.2. QSM lettura

La prova è costituita da quattro quesiti a tre entrate. Viene presentato un *item*-esempio, che indica il genere di testo somministrato.

Esempio:

- 0  Hai appena letto...
- ... un articolo.
- ... un'intervista.
- ... una pubblicità.

Il tema del brano riguarda la *Pet Therapy*, come si evince dalla lettura degli *item*. Qui sotto, il lettore ha modo di accedere all'intera prova con le percentuali relative alle opzioni scelte *item* per *item*. Le chiavi sono evidenziate in grassetto.

21 L'articolo parla principalmente...	
a. ... della storia di Luca e Jack.	(28,9)
b. ... della storia della <i>Pet Teraphy</i> .	(20,7)
c. ... della <i>Pet Teraphy</i> in Italia.	(50,4)
22 Secondo l'articolo...	
a. ... Luca è guarito dalla sua malattia.	(13,4)
b. ... Luca si avvicina più spesso a Jack.	(6,00)
c. ... Luca ha fatto progressi grazie al cane Jack.	(80,6)
23 Secondo l'articolo...	
a. ... ci sono progetti di <i>terapia assistita con animale</i> in tutta Italia.	(34,6)
b. ... stanno nascendo centri di <i>Pet Teraphy</i> su tutto il territorio nazionale.	(36,7)
c. ... si comincia a lavorare con <i>terapia assistita con animale</i> .	(28,7)
24 Secondo l'articolo...	
a. ... tutti gli animali sono di aiuto nella <i>Pet Teraphy</i> .	(53,4)
b. ... gli animali maturi sono i migliori per la <i>Pet Teraphy</i> .	(33,8)
c. ... il gatto è l'animale più indicato per questo tipo di terapia.	(12,8)

Di seguito analizziamo i due *item* con un IP > 50:

- # 21 (IP=50,4)
- # 22 (IP=79,4)

#### Item # 21

Problema: *correlazione; prenoscenze*

L'attrattività della chiave (IP= 50,4) può essere spiegata per via dei rimandi che provengono dagli altri *item*. In questo caso, non sono gli *stem* a pilotare le ipotesi, come nell'*item* # 1 della prova precedente, quanto invece la frequenza del concetto di "Pet Therapy" tra le opzioni degli *item* ## 23 e 24, in opposizione alla vicenda di Luca e Jack, attorno a cui ruota solamente l'*item* # 22 (argomento che invece il distrattore [a] descrive come tema dell'articolo). Stando così le cose, vien da intuire che l'argomento dell'articolo è quello a cui più *item* danno rilievo; così come, le proprie conoscenze in materia di articoli a carattere divulgativo, e in particolare del loro stile espositivo, ci fanno supporre che, in questo caso, come in molti altri, si parta da un esempio (la vicenda di Luca e Jack) per introdurre il tema generale, trattato poi in un secondo momento (nel nostro caso, la terapia attraverso gli animali). Gli stessi rimandi interni fanno escludere che l'opzione [b] (secondo la quale nell'articolo si parla della "storia della Pet Therapy") sia la chiave (abbiamo

solo un 20,7 % di consensi): si constata, di fatto, una totale assenza di date né si menziona altro luogo che non sia l'Italia.

### Item # 22

Problema: *implausibilità dei distrattori*

L'altissimo IP (80,6) si lega all'implausibilità dei distrattori. Secondo quanto detto sopra, l'argomento risulta essere chiaro: si parla di animali e di terapia (per molti, abbiamo visto, è proprio l'argomento "principale"; in ogni caso per l'intero campione è evidente che, se anche non lo fosse, è comunque un tema di cui si tratta).

È logico supporre che Jack sia effettivamente un cane (tra l'altro il nome proprio, non italiano, suffraga l'ipotesi); il fatto che compaia tra i distrattori (in questo *item*, come nel precedente) pare suggerire che la menzione di questo animale sia un dato significativo (e quindi esclude il distrattore [a], che fa riferimento ad una guarigione senza specificare – dato invece importante – in virtù della compagnia di quale *mascotte*); inoltre, proprio perché l'animale ha una funzione terapeutica, la relazione tra Luca e Jack non è fine a se stessa (come vuole il distrattore [b]) ma è lo strumento mediante il quale le abilità sociali di Luca vengono potenziate.

### 3.1.3. VF ascolto

La prova è costituita da cinque *item*. Viene presentato un *item*-esempio, che indica il genere di ascolto e il tema: "una squadra di calcio".

Esempio:

0 Il servizio radiofonico parla di una squadra di calcio  F

Qui sotto, il lettore ha modo di accedere all'intera prova con le percentuali relative alle opzioni scelte *item* per *item*. Le chiavi sono evidenziate in grassetto.

11 I giocatori del KOA Bosco sono tutti immigrati	V	F (32,4)
12 I rapporti tra i giocatori e la popolazione di Rosarno sono migliorati	V	F (73,5)
13 Mettere su la squadra è stato semplice e veloce	V	F (64,7)
14 Uno dei giocatori vive ancora in una tendopoli	V	F (55,1)
15 Due giocatori sono stati assunti da aziende locali e hanno lasciato la squadra	V	F (76,5)

Di seguito analizziamo i quattro *item* con un IP > 50:

- # 12 (IP=73,5)
- # 13 (IP=64,7)
- # 14 (IP=55,1)
- # 15 (IP=76,5)

Consideriamo, peraltro, anche l'\*item # 11, il quale, benché presenti un basso IP, consente – a una lettura attenta – di predire la chiave.

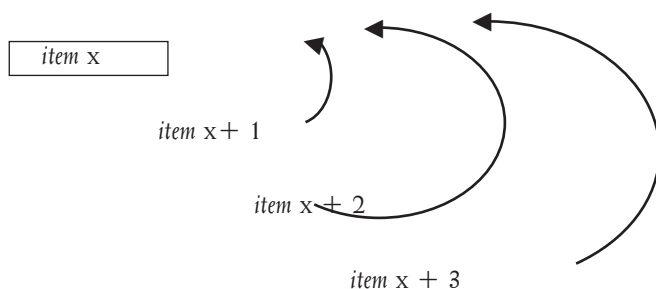
#### \*Item # 11

Problema: *correlazione*

L'item # 11, nonostante il basso IP (32,2), ha potenzialmente un buon margine di predittività. Nell'item # 12 infatti, "i giocatori della squadra" vengono presentati in opposizione alla "popolazione locale"; quindi, dovrebbe essere immediato dedurre che i "giocatori" sono tutti allotocni. Tale deduzione non è stata operata però in maniera consistente dal campione. C'è da supporre, in tal senso, che la predittività, nell'ambito della correlazione tra gli item, possieda maggiore forza se ha un carattere *progressivo* e minore forza se ha un carattere *regressivo*. Spieghiamo meglio. La correlazione, sia essa forte o debole, pare avere un maggiore impatto in termini di predittività quando si sta per rispondere ad un item (x) e si procede a una lettura cursoria degli item successivi (x+1, x+2, x+3, ecc.) o eventualmente precedenti (x-1, ecc.). Viceversa, se all'item "x" abbiamo già dato risposta, eventuali dati in entrata, che una lettura attenta degli item successivi può provvedere, hanno uno scarso impatto nel costringere a riformulare e ristrutturare le ipotesi che ci hanno fatto considerare quell'item (x) in un certo modo. Ciò, secondo noi, è dovuto al particolare tipo di testo che una batteria di item rappresenta: la lista (di frasi, in genere). In una lista, la lettura è continuamente 'spostata in avanti' e si concentra sul nuovo elemento verso cui, di volta in volta, si posa l'attenzione del lettore. Il procedere 'avanti e indietro' (o 'su e giù'), tipico del lettore esperto alle prese con testi scritti a struttura sintattico-testuale complessa, viene meno, una volta che siamo esposti a liste. Illustriamo il concetto alla fig. 2:

- il caso a) rappresenta il contributo che item successivi hanno nello stabilire la risposta corretta, in termini di predittività, dell'item iniziale;
- il caso b) invece rappresenta lo scarso impatto (*predittività regressiva*) che i dati che provengono da altri item esercitano su un item a cui è già stata data risposta, e che quindi non rientra più nel focus dell'attenzione del lettore (si veda la curva tratteggiata).

Caso a)



Caso b)

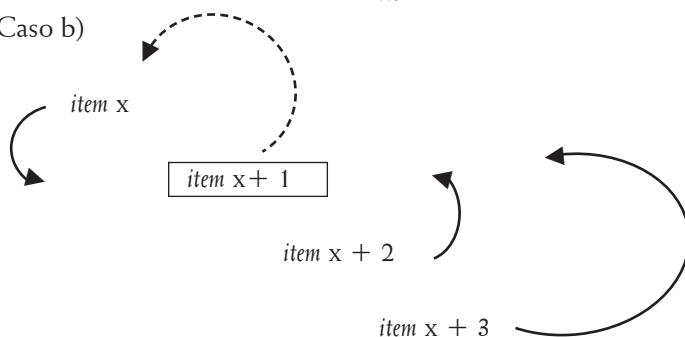


Fig. 2. La predittività regressiva

I casi di *predittività regressiva* possono *non* essere rilevati dall'analisi statistica della predittività; sfuggono cioè al calcolo dell'IP, costituendo una predittività 'nascosta'. Tuttavia, vanno tenuti in seria considerazione, dal momento che non possiamo escludere che un candidato possa rispondere a questo genere di *item* servendosi della sola logica, senza quindi far riferimento al testo oggetto di comprensione.

*Item # 12*

Problema: *correlazione tra gli item; prenoscenze*

Dall'*item* precedente (# 11) risulta evidente che la "squadra di calcio" che l'*item*-esempio presenta come tema, sia costituita da un certo numero di giocatori immigrati. Rientra nel senso comune, poi, l'idea secondo la quale, se si danno occasioni di incontro (e quindi di reciproca conoscenza) tra persone di culture diverse, è più facile l'integrazione, specie da parte degli allotocni rispetto agli autotocni, e quindi si ha un miglioramento delle relazioni tra le comunità. Questi due fattori, quindi, la correlazione tra gli *item* e le prenoscenze, spiegano l'IP elevato dell'*item # 12* (=73,5).

*Item # 13*

Problema: *preconoscenze, implausibilità dei distrattori*

L'item recita: "Mettere su la squadra è stato semplice e veloce". Al lettore risulta chiaro che la squadra ("KOA Bosco") è composta di "immigrati" (non sa in che percentuale, ma il dato è certo); la sua esperienza lo fa supporre che contribuire alla formazione di un gruppo di persone provenienti da paesi e contesti diversi possa non essere un'impresa immediata. In più, a falsificare l'affermazione rappresentata dall'item, contribuisce l'esagerazione implicita nell'endiadi "semplice e veloce": la "semplicità" nel "mettere su la squadra" equivale per forza di cose alla "rapidità"; dispositivi linguistici, come in questo caso, che prevedono elementi pleonastici (cioè ridondanti, e quindi inutili) mettono in guardia dal ritenere attendibile il contenuto veicolato dall'item. In sostanza, in un VF, la proposizione falsificabile (come l'item in sé e di per sé, in questo caso) vale da distrattore, perché è come se avessimo una scelta multipla a due entrate:

Mettere su la squadra

- a) è stato semplice e veloce
- b) non è stato semplice e veloce

La falsificazione della prima opzione (e quindi l'implausibilità dell'item stesso) viene inferita dalla costruzione lessicale ingiustificata.

*Item # 14*

Problema: *preconoscenze, correlazione*

Le preconoscenze giocano a favore della predittività della chiave. La cronaca ci informa che un immigrato, in un paese europeo, si trova spesso in condizioni precarie, tanto di lavoro (come del resto si intuisce nell'item # 15) quanto di sistemazione. È quindi altamente probabile che un giocatore immigrato viva "ancora in una tendopoli". D'altronde, se la frase fosse falsa, bisognerebbe chiedersi perché vi è l'accento alla "tendopoli"; forse – si potrebbe congetturare – nel brano si dice che il giocatore (o più di un giocatore) viveva ma non vive più in una "tendopoli", ma ciò supporrebbe che il KOA Bosco sia una squadra professionale, che dà di che vivere ai suoi giocatori, il che però viene smentito da altri rimandi interni agli item (primo tra tutti l'item # 15, che il campione ritiene come altamente plausibile, secondo il quale, con l'occasione di un impiego, due giocatori hanno abbandonato la squadra).

## Item # 15

Problema: *preconoscenze, correlazione*

I rimandi interni tra gli *item*, uniti all'esperienza del lettore, spingono a ritenere plausibile l'*item* in questione. Sappiamo che alla "squadra" appartengono immigrati (# 11): non sono giocatori professionisti; è plausibile che i rapporti, prima della formazione della squadra, con la gente del posto non fossero tra i più rosei (# 12; e ciò conferma la situazione di difficoltà nella quale versano i giocatori), può essere che addirittura uno di loro viva ancora in un campo allestito (# 14); di conseguenza è probabile che, una volta assunti da aziende locali, due giocatori abbiano lasciato la squadra, vista la difficoltà a conciliare sport e lavoro.

### 3.1.4. VF lettura

La prova è costituita da cinque *item*. Viene presentato un *item*-esempio, che indica il genere di testo esibito al lettore.

Esempio:

0 Hai appena letto un articolo di giornale ~~V~~ F

Scorrendo gli *item*, si intuisce l'argomento: *una mostra dedicata alla scienza*. Qui sotto, il lettore ha modo di accedere all'intera prova con le percentuali relative alle opzioni scelte *item* per *item*. Le chiavi sono evidenziate in grassetto.

- 31 La mostra apre per celebrare la morte di Galileo Galilei. V F (41,9)
- 32 La mostra racconta gli **errori dei grandi scienziati del passato** V F (61,8)
- 33 Secondo il Dott. Masiero **gli errori nella scienza** sono fondamentali V F (68,4)
- 34 Einstein **teorizzò la costante cosmologica e capì fin da subito la sua importanza** V F (52,2)
- 35 Ci sono video e installazioni che aiutano a capire alcune scoperte scientifiche V F (80,1)

Di seguito analizziamo i quattro *item* che presentano un IP > 50:

- # 32 (IP=73,5)
- # 33 (IP=64,7)
- # 34 (IP=55,1)
- # 35 (IP=76,5)



Item## 32, 33, 34

Problema: *correlazione, implausibilità dei distrattori* (per il # 33)

Gli item ## 32, 33, 34 sono interdipendenti e si rinforzano a vicenda. Nel # 32 si parla di "errori", come pure nel # 33. È nel # 33 che l'ipotesi adottata nel precedente (cioè che la mostra racconti "gli errori" compiuti nel corso di sperimentazioni scientifiche) viene validata: se l'intervistato, il dott. Masiero (probabile curatore della mostra?), avesse detto che gli errori nella scienza non sono fondamentali, avrebbe affermato una banalità; quindi è probabile che abbia detto il contrario, e di conseguenza prende forma l'ipotesi che la mostra tratti proprio di questo genere di errori (# 32). Segue, secondo logica, la risposta all'item # 34: proprio in accordo con il tema dell'"errare" (nel senso di sperimentare, senza sapere esattamente qual è il destino della propria ricerca), vien da sé che è plausibile che, nell'atto di teorizzare la "costante cosmologica", Einstein non fosse in grado di prevedere la portata della propria scoperta.

Item # 35

Problema: *preconoscenze*

Considerato che la mostra tratta (anche) di Galileo (non sappiamo se è una commemorazione della nascita, della morte o di un evento significativo della sua carriera di scienziato), e che comunque tratta di scienza, l'esperienza ci dice che è assai probabile vi siano "video" e "installazioni" che consentano di apprezzare meglio "alcune scoperte scientifiche".

## 3.2. Livello B2

Consideriamo le prove relative al livello B2.

### 3.2.1. QSM ascolto

La prova è costituita da cinque quesiti a tre entrate. Viene corrisposto un *item-esempio*, che illustra l'argomento di cui parla il brano: "l'ospedale delle bambole", vale a dire "una bottega dove si aggiustano le bambole".

Esempio:

0 L'ospedale delle bambole è...

- a. ...un centro di divertimento per bambini.
- b. ...un negozio dove si vendono giocattoli.
- ~~c. ...una bottega dove si aggiustano bambole.~~

Qui sotto, il lettore ha modo di accedere all'intera prova con le percentuali relative alle opzioni scelte *item per item*. Le chiavi sono evidenziate in grassetto.

- 1 L'ospedale delle bambole si trova...
  - a. ... nella piazza principale di Napoli. (18,2)
  - b. ... nel centro storico di Napoli.** (41,3)
  - c. ... nella periferia di Napoli. (40,5)
  
- 2 La famiglia di Tiziana si occupa...
  - a. ... della riparazione di bambole.** (57,9)
  - b. ... della produzione artigianale di bambole. (33,9)
  - c. ... della vendita *online* di bambole. (8,20)
  
- 3 Spesso, gli abiti delle bambole provengono da...
  - a. ... negozianti di giocattoli. (31,4)
  - b. ... rivenditori di tessuti. (33,9)
  - c. ... clienti della bottega.** (34,7)
  
- 4 Il nome "ospedale delle bambole" è stato suggerito da...
  - a. ... il bisnonno di Tiziana. (50,4)
  - b. ... il padre di Tiziana. (26,5)
  - c. ... un abitante del quartiere.** (23,1)
  
- 5 Tiziana spera che in futuro...
  - a. ... i bambini visiteranno la bottega.** (41,3)
  - b. ... i figli lavoreranno in bottega. (49,8)
  - c. ... i figli venderanno la bottega. (8,90)

Solo l'*item # 2*, che analizziamo qui sotto, presenta un IP > 50.

#### *Item # 2*

Problema: *correlazione*

L'alta predittività della chiave (IP=57,9) è dovuta al fatto che l'informazione veicolata è già contenuta nell'*item*-esempio. Lo studente, quindi, può rispondere senza ascoltare il brano corrispondente. Ci si potrebbe aspettare un IP ancora più alto, in verità, il fatto però che l'incrocio di informazioni avvenga con un *item* esterno alla prova, qual è l'*item*-esempio, può giustificare la svista da parte di coloro che non hanno riconosciuto la sovrapposizione.

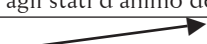
### 3.2.2. QSM lettura

La prova è costituita da quattro quesiti a tre entrate. Viene presentato un *item*-esempio, che illustra l'argomento: "una ricerca su cosa tiene unite le coppie".

Esempio:

- 0 Hai appena letto...
- a. ...un'inchiesta sui divorzi in Italia.
  - ~~b. ...una ricerca su cosa tiene unite le coppie.~~
  - c. ...alcune tecniche per vivere insieme.

Qui sotto, il lettore ha modo di accedere all'intera prova con le percentuali relative alle opzioni scelte *item* per *item*. Le chiavi sono evidenziate in grassetto.

- 21 Gli studi di John Gottman...
- a. ...**sono continuati per molti anni.** (31,7)
  - b. ...hanno permesso il ricongiungimento di molte coppie. (44,2)
  - c. ...non sono applicabili in tutti i paesi. (24,1)
- 22 Le categorie *Master* e *Disaster*...
- a. ...hanno caratteristiche simili al loro interno. (18,5)
  - b. ...**distinguono le coppie ancora unite da quelle separate.** (68,9)
  - c. ...raccontano per lo più fatti di vita piacevoli. (12,6)
- 23 Gli studi del 1990...
- a. ...hanno riguardato coppie sposate da molti anni. (35,5)
  - b. ...**hanno osservato momenti di vita quotidiana di neosposi.** (33,1)
  - c. ...hanno considerato le reazioni delle mogli di fronte a passioni dei mariti. (31,4)
- 24 Per un amore durevole...
- a. ...è necessario condividere momenti brutti insieme. (37,2)
  - b. ...è indispensabile uscire più spesso con il partner. (9,90)
  - c. ...è importante partecipare agli stati d'animo del compagno. (52,9)
- 25 Generosità ed empatia sono...
- ...**requisiti fondamentali per stare insieme.** (60,3)
  - ...doti sicuramente innate nell'individuo. (12,6)
  - ...caratteristiche presenti anche nella categoria *Disaster*. (27,1)
- 

A seguire analizziamo i tre *item* con un IP > 50:

- # 22 (IP=68,9)
- # 24 (IP=52,9)
- # 25 (IP=60,3)

#### *Item # 22*

Problema: *ovvietà; correlazione; implausibilità di un distrattore*

L'alto IP (68,9) si deve a una deduzione logica, possibile grazie alla conoscenza, pur se elementare, dell'inglese. "Master" e "Disaster" appaiono forme in opposizione, da cui la confutazione dell'assimilazione, implicitamente veicolata dai distrattori [a] e [c]. Per di più, considerato che il tema, come trasmesso dall'*item*-esempio, riguarda il cosa "tiene unite le coppie", viene immediato associare un comportamento virtuoso alla categoria "Master", e uno, al contrario, negativo alla categoria "Disaster". In questo senso, a favore della predittività, agisce anche una correlazione con l'*item*-esempio. Il distrattore [c], oltretutto, si presenta assurdo.

#### *Item # 24*

Problema: *correlazione*

Le risposta corretta [c] può essere derivata dallo *stem* dell'*item* successivo, come evidenziato nella prova riportata sopra. La chiave di # 24 è, infatti, la perifrasi di uno degli attributi ("empaticità" – termine che lascia adito, del resto, a dubbi in quanto a correttezza) riportati nello *stem* di # 25 (e quindi dato come presente nel corpo del testo).

#### *Item # 25*

Problema: *ovvietà; correlazione; implausibilità di un distrattore*

Mentre il distrattore [c] è **improbabile** (è un dato noto che l'assenza di empatia, e quindi di sana comunicazione, è causa di problemi di relazione), il distrattore [b] vanta una certa plausibilità, a seconda della visione 'filosofica' che uno ha della vita (una visione ottimistica lascia supporre che ci sia un'empatia innata, rappresentabile neurologicamente, peraltro, attraverso i circuiti cerebrali che fanno capo al funzionamento dei neuroni-specchio). A tradire il certificatore, è in ogni caso l'aggiunta, nel corpo del distrattore stesso, di un avverbio pleonastico "sicuramente"; inoltre, a rendere implausibile il distrattore, concorre il fatto che esso sia fuori-tema: che l'empatia sia una dote innata, è una tesi che ci si potrebbe aspettare all'interno di un testo filosofico, ma non in uno psicologico, dedicato all'unità della coppia. Ed è, infatti, la coerenza che l'informazione veicolata dalla chiave [a] dell'*item* # 25 ha

con il tema dell'argomento (la vita di coppia), unita alla sovrapposizione semantica con l'item precedente (## 24 e 25 comunicano, infatti, uno stesso messaggio) e all'ovvietà del dato espresso (gentilezza e disponibilità generano benessere), che giustifica l'alto IP (60,3).

### 3.2.3. VF ascolto

La prova è costituita da cinque *item*. Viene presentato un *item*-esempio, che indica l'argomento del brano: "Il nuovo disco di Fiorella Mannoia", intitolato "Sud".

Esempio:

0 Il nuovo disco di Fiorella Mannoia s'intitola Sud ~~V~~ F

Qui sotto, il lettore ha modo di accedere all'intera prova con le percentuali relative alle opzioni scelte *item* per *item*. Le chiavi sono evidenziate in grassetto.

11 Il nuovo disco "Sud" è una raccolta di canzoni sull'Italia del sud	<b>V</b> F (35,8)
12 La fonte d'ispirazione del disco è un libro di Pino Aprile	V <b>F</b> (41,7)
13 Nel disco, la cantante ha collaborato con musicisti internazionali	V F (62,5)
14 Secondo la cantante, si è sempre valorizzato il sud del mondo	V F (49,2)
15 <i>Vuelvo al Sur</i> è l'ultima canzone del disco cantata in spagnolo da Fiorella Mannoia	V <b>F</b> (36,7)

Un solo *item* presenta un IP > 50, # 13 (62,5). Oltre ad esso, a seguire rivolgiamo una particolare menzione all'\**item* 14, il quale, a nostro parere, presenta margini di predittività significativi, nonostante l'IP sia appena al di sotto della soglia da noi stabilita.

#### Item # 13

Problema: *implausibilità del distrattore*

L'item è così formulato: "Nel disco, la cantante ha collaborato con musicisti internazionali". Si potrebbe presumere che sia stato il titolo della canzone in spagnolo, nell'item finale # 15, a far decidere il campione verso il consesso internazionale di musicisti (correlazione). Eppure c'è da riconoscere che l'IP relativo a quest'ultimo è basso (36,7) e che, d'altronde, se anche effettivamente avesse avuto il potere di orientare il gruppo a rispondere correttamente all'item in questione, avrebbe dovuto dimostrare pure la 'falsità' dell'item # 11 (secondo il quale la cantante si

sarebbe limitata a un repertorio nazionale, e non internazionale), ma così non è stato.

Di conseguenza, riteniamo non sia stata tanto una correlazione tra gli *item* quanto l'implausibilità dell'ipotesi contraria a quanto affermato a spingere il campione a giudicare il contenuto della frase come attendibile. Vogliamo dire, il fatto che l'artista si sia misurata con un gruppo di musicisti internazionali è degno di nota nel testo di una recensione; il contrario, e cioè che un'artista italiana abbia collaborato con musicisti italiani, non è una notizia significativa, perché descrive ciò che è normale e scontato; la sua presenza in un testo scritto da nativi per nativi sarebbe quanto meno poco realistica.

Anche in questo caso, intendiamo per distrattore, all'interno di una prova VF, tutte le possibili affermazioni falsificabili, costituite dal distrattore e dalle varianti sull'asse paradigmatico. Nel caso specifico, sull'asse paradigmatico abbiamo una sola alternativa (internazionali/nazionali), facile da escludere, come abbiamo detto, non già perché assurda, ma perché altamente improbabile.

#### \*Item 14

Problema: *ovvietà*

Benché all'*item* # 14 corrisponda un IP inferiore alla soglia stabilita (49,2), tale *item* presenta dei margini di predittività legati all'ovvietà dell'informazione veicolata. L'*item* recita "Secondo la cantante si è sempre valorizzato il sud del mondo". Ora, noi non sappiamo cosa pensi Fiorella Mannoia della questione, ma il fatto contrario a quanto dichiarato, e cioè che il sud del mondo non sia mai stato valorizzato (da tutti i punti di vista: economico, culturale, politico, ecc.), è un'evidenza. Tra gli altri esperti di *testing*, Cangelosi insiste sulla cautela che l'*item writer* deve avere nella scrittura di distrattori che sono falsi, stando al testo, ma veri nella realtà delle cose (CANGELOSI 1990<sup>3</sup>); la stessa attenzione dev'essere prestata, aggiungiamo noi, nel caso contrario (*item* veri secondo il testo ma falsi nella realtà delle cose) e pure quando gli *item* veicolano informazioni false su entrambi i piani.

#### 3.1.4. VF lettura

La prova è costituita da cinque *item*. A differenza delle precedenti, non viene fornito alcun *item*-esempio. Scorrendo gli *item*, si intuisce che il testo ruota attorno al "Calintri Sponz Festival".

Qui sotto, il lettore ha modo di accedere all'intera prova con le percentuali relative alle opzioni scelte *item* per *item*. Le chiavi sono evidenziate in grassetto.

31 Il festival "Sponz Fest" si svolge in Irpinia	V F (56,7)
32 Il Calitri Sponz Festival si occupa esclusivamente di musica e fotografia	V F (49,2)
33 Secondo Capossela, il festival è un'occasione per rilassarsi e divertirsi	V F (78,3)
34 Il festival è organizzato grazie al sostegno economico dei finanziatori	V F (40,8)
35 Il festival terminerà a mezzanotte	V F (50,8)

Analizziamo di seguito i tre *item* con IP > 50:

- # 31 (IP=56,7)
- # 33 (IP=78,3)
- # 35 (IP=50,8)

Diamo pure menzione all'\**item* # 32; benché l'IP sia di poco al di sotto della soglia definita, lascia presagire margini di predittibilità significativi.

#### *Item* # 31

Problema: *implausibilità del distrattore*

L'apprendente viene informato che "Il festival «Sponz Fest» si svolge in Irpinia".

Se l'*item* fosse falso, occorrerebbe che nel testo corrispondente (il brano di lettura) venisse indicata un'altra regione, per via diretta (es. Sicilia) o per via indiretta, attraverso perifrasi (es. l'isola più a sud d'Italia) o una città o un paese che non facciano parte dell'Irpinia (es. Milano). In realtà, proprio perché l'Irpinia è un'area poco nota, la falsificazione con un'eventuale informazione divergente presente nel testo sarebbe troppo difficile, proprio perché il candidato non ha modo di prefigurarsi dove sia la zona in questione (è una microarea all'interno di una regione, come lo è la Marca Trevigiana all'interno del Veneto? È una macroarea tra regioni, come lo è la Pianura Padana? È il sinonimo di un'altra regione, come lo è la Lucania rispetto alla Basilicata?). Ogni falsificazione comporterebbe, insomma, un compito improbo da parte del candidato.

#### \**Item* # 32

Problema: *correlazione, implausibilità del distrattore*

Benché l'IP sia al di sotto della soglia da noi ammessa (49,2), l'*item* # 32 presenta dei margini di predittività a nostro giudizio significativi, per via delle inferenze che si possono trarre da informazioni veicolate da altri *item*. L'*item* recita "Il Calitri Sponz Festival si occupa esclusivamente di musica e fotografia"; le possibilità alternative, che falsificherebbero l'affermazione, sono date dalle seguenti varianti:

- i. "Il Calitri Sponz Festival si occupa di x, oltre che di musica e di fotografia"
- ii. "Il Calitri Sponz Festival si occupa di musica"
- iii. "Il Calitri Sponz Festival si occupa di fotografia"

Che si occupi di musica, è confermato da informazioni che provengono da *altri* item (il # 33, che prefigura la possibilità di "rilassarsi e divertirsi" e il # 35, che afferma la possibilità di una chiusura notturna, alquanto improbabile per un festival di sola fotografia). L'alternativa iii) è dunque falsificata. Rimangono le altre due opzioni. A decidere a favore della prima, vi è l'uso di un avverbio dal valore pleonastico: "esclusivamente". Se, in effetti, nel testo si dicesse che queste sono le arti attorno a cui ruota il festival, sarebbe improbabile che l'autore si fosse servito di un avverbio che indica limitazione "esclusivamente musica e fotografia", che lascia intendere che vi è un'opposizione ad altro; piuttosto avrebbe enunciato "si serve di musica e fotografia" e non avrebbe aggiunto altro. L'uso sospetto di "esclusivamente" fa rilevare, per i lettori più accorti e intuitivi, una forzatura linguistica che manifesta l'*item* come un distrattore, vale a dire un'informazione che non rispecchia l'informazione contenuta nel testo.

#### Item # 33

Problema: *ovvietà*

Benché il lettore possa non esser in grado di decifrare esattamente quale sia l'oggetto del "Calitri Sponz Festival", risulta evidente che un festival (termine inglese, di origine latina, che ha stesso etimo di "festa") è un'occasione per "rilassarsi e divertirsi". Improbabile, il contrario: che una manifestazione sia pensata per innervosire o annoiare i partecipanti.

#### Item # 35

Problema: *preconoscenze, implausibilità del distrattore*

La maggioranza del campione ha ritenuto falsa l'informazione veicolata da quest'*item*, secondo il quale il festival termina a mezzanotte. Probabile molti abbiano ragionato in termini di preconoscenze: i festival letterari, filosofici, ecc. terminano prima; i festival musicali terminano dopo. Considerata, oltretutto, l'alta probabilità che si tratti di un festival (in parte) musicale<sup>1</sup>, si può pensare che l'ora di chiusura vada ben al di là di quanto espresso nell'*item*.

---

<sup>1</sup> Ad ogni modo, benché non sia un nome molto noto all'estero, il riferimento al cantante Capossela, nell'*item* # 33, eliminerebbe ogni dubbio.



#### 4. DISCUSSIONE

Nell'analisi condotta, risulta che le prove di comprensione della certificazione .IT, in riferimento al livello oggetto di studio (B), presentano un numero consistente di *item* che consentono di individuare la risposta corretta senza comprendere il testo di riferimento. Gli *item* con  $IP > 50$  sono 20, quindi il 40% del totale. Se aggiungiamo pure gli *\*item* con IP inferiore ma che comunque costituiscono delle insidie alla validità delle prove per i margini di predittività ammessi, raggiungiamo la percentuale ragguardevole del 47% (Tab. 3). Insomma, quasi la metà degli *item* delle prove prese in considerazione dovrebbe/avrebbe dovuto essere rivista, al fine di presentare margini di predittività più contenuti. In particolar modo, è necessario /sarebbe stato necessario che l'*item writer* presti /avesse prestato una particolare attenzione a *bias* dovuti alla correlazione tra gli *item* e all'implausibilità dei distrattori – fattori di predittività che si sono dimostrati tra i più ricorrenti.

LI-VELLO	FOR-MAT	ABILITÀ	ITEM	IP								PROBLEMI									
				20	30	40	50	60	70	80	OV	PC	CR	IM							
B1	QSM	ASCOLTO (\$ 3.1.1)	1				53,7												•		
			2		30,9																
			3					64								•					•
			4			41,9															
			5	20,1																	
		LETTURA (\$ 3.1.2)	21				50,4													•	
			22									80,6									
			23		34,6																
			24		33,8																
			11		*32,2																
		ASCOLTO (\$ 3.1.3)	12								73,5								•		
			13								64,7									•	
			14				55,1												•		
			15								76,5									•	
	VF		31					41,9													
		LETTURA (\$ 3.1.4)	32								61,8										•
			33								68,4										•
			34						52,2												•
			35												80,1				•		

Legenda

OV: ovvietà

PC: preconoscenze

CR: correlazione

IM: implausibilità dei distrattori

B2	QSM	ASCOL- TO	1			41,3																				
			2											57,9												
		(\$ 3.2.1)	3				34,7																			
			4			23,1																				
			5					41,3																		
	LETTURA	21							31,7																	
		22																			68,9					
		23																								
		(\$ 3.2.2)	24																			52,9				
			25																				60,3			
	VF	ASCOL- TO	11																							
			12											41,7												
		(\$ 3.2.3)	13																							
			14																							
			15																							
	LETTURA	31																								
		32																								
		(\$ 3.2.4)	33																							
			34																							
		35																								

Tab. 3. Sintesi degli IP

## 5. CONCLUSIONI

Attraverso questo saggio abbiamo introdotto un indice, l'*indice di predittività*, che torna utile nel processo di validazione di una prova di comprensione con *item* separati dal testo. L'indice permette all'*item writer* di accertarsi che gli *item* non consentano ai candidati di scoprire quale sia la chiave, per via di processi inferenziali, indipendenti dalla lettura o dall'ascolto del brano di riferimento.

La soglia da noi definita corrisponde al valore di 50; l'analisi condotta ci dimostra l'opportunità di considerare anche *item* con IP di poco inferiori, oltre ad altri che presentano problemi di *predittività regressiva*, non rilevabili dall'indice.

Considerata retroattivamente e con occhio critico, l'analisi effettuata avrebbe potuto avere un maggior rigore se avessimo:

- a. informato previamente il campione sulle possibili strategie che si possono implementare durante la lettura degli *item* senza testo; ciò ci avrebbe assicurato di disporre del maggior numero possibile di comportamenti strategici
- b. fornito uno spazio per eventuali commenti metacognitivi; alcune inferenze da noi addotte avrebbero, plausibilmente, trovato maggiore riscontro.

Ulteriori indagini che si possono svolgere in tal senso potranno tener conto di queste indicazioni.

---

## REFERÊNCIAS BIBLIOGRÁFICAS

- ALDERSON, C.; CLAPHAM, C.; WALL, D. (1995). *Language test construction and evaluation*. Cambridge: CUP.
- CANGELOSI, J. S. (1990<sup>3</sup>). *Designing tests for evaluating student achievement*. White Plains, NY: Longman.
- CARR, N. T. (2011). *Designing and analyzing language tests*. Oxford: OUP.
- COUNCIL OF EUROPE (2001). *The common European framework of reference for languages: learning, teaching, assessment*. Cambridge: CUP.
- LYNCH, B. K.; DAVIDSON, F. G. (1994). Criterion-referenced language tests development: linking curricula, teachers and tests. *TESOL Quarterly*, 28, 4, pp. 727-743.
- McNAMARA, T. (2000). *Language testing*. Oxford: OUP.
- POWERS, D. E.; WILSON, S. T. (1993). Passage dependence on the new SAT reading comprehension questions. *College Board Report*, 93, 3, pp. 1-18. Disponibile online: <<http://research.collegeboard.org/>>.

Analisi della predittività di item tratti da prove di comprensione della certificazione...

SHOHAMY, E. (1985). *A practical handbook in language testing for the second language teacher*. Tel Aviv: Tel Aviv University.

WEIR, C. (1990). *Communicative language teaching*. Prentice Hall: London.

WEIR, C. (1993). *Understanding and developing language tests*. Prentice Hall: London.

Sitografia

Prova B2 .IT

<<http://www.certificazioneitaliano.uniroma3.it/B2-intlIT.aspx>>  
18.01.2016)

(consultazione:

Recebido: 27/01/2016

Aceito: 23/03/2016