

UN APPROCCIO INNOVATIVO AL TESTING PSICOPATOLOGICO: TALEIA PARTE II: ATTENDIBILITÀ E VALIDITÀ DEL TEST

LUCIA BONCORI¹, ALESSANDRA DE CORO¹,
GIOVANNI CUOMO² E PAOLO RENZI¹

¹ *Sapienza Università di Roma*, ² *Sanità della Polizia di Stato*

Riassunto. Viene verificata attendibilità e validità di TALEIA-400A (*Test for Axis I Evaluation and Interview for Clinical, Personnel, and Guidance Applications*) e delle sue forme parallele (fra cui il Difesa Test). Gli studi sull'attendibilità evidenziano stabilità al retest (in media $r = 0.83$) e parallelismo delle forme (in media $r = 0.74$). ACP insieme a scale che valutano tratti normali evidenziano quattro fattori, il primo dei quali psicopatologico e altri in cui appare continuità fra normalità e patologia. MANOVA fra campioni normali e patologici ($N = 4838$) evidenziano capacità diagnostica in tre ambiti: clinica, selezione del personale, orientamento. Altre analisi evidenziano l'influsso significativo dei fattori Sesso e Condizione d'uso e suggeriscono l'opportunità di elaborare parametri specifici per la condizione concorsuale, utilizzando dati raccolti in situazioni reali.

1. INTRODUZIONE

TALEIA è uno strumento costruito come sussidio per la psicodiagnosi, destinato all'utilizzazione nelle situazioni professionali più comuni: diagnosi clinica, selezione del personale, orientamento e *counselling*. Include un questionario che valuta otto disturbi clinici e dieci di personalità, assumendo come indicatori i criteri diagnostici definiti da DSM-IV e ICD-10 per i disturbi stessi. Le scale riferite all'Asse I sono: S – Schizofrenia, D – Depressione, M – Ipomania e mania, AA – Ansia acuta e attacchi di panico, FO – Sindromi fobiche, SOC – Sindrome ossessivo-compulsiva, AG – Ansia generalizzata, AL – Disturbi da alterato comportamento alimentare. Le scale riferite all'Asse II sono: PP – Disturbo di personalità (d'ora in poi: DP) Paranoide, PSK – DP Schizoide, PSKT – DP Schizotipico, PAS – DP Antisociale, PB – DP Borderline, PI – DP Istrionico, PN – DP Narcisista, PEV – DP Evitante, PD – DP Dipendente, POC – DP Ossessivo-Compulsivo o Anancastico. Le scale di controllo sono: L – Tendenza a presentare un'immagine favorevole di sé (Desiderabilità sociale), F – Tendenza a presentare un'immagine sfavorevole di sé (autoattribuzione di disturbi poco frequenti), INC – Incongruità variabile fra le risposte. I criteri di costruzione applicati mirano a costruire quesiti le cui risposte siano interpretabili non tanto come au-

todescrizioni quanto come comportamenti verbali, il cui rapporto con le variabili misurate viene verificato empiricamente. Il questionario è integrato da un colloquio standardizzato, le cui domande sono selezionate da un software dedicato che automaticamente stampa i quesiti adatti ai singoli profili basati sulle risposte al questionario. In questo articolo diamo un breve e selettivo rendiconto delle ricerche compiute nel corso di circa dieci anni di studio per verificare la rispondenza di TALEIA agli scopi prefissi.

2. ATTENDIBILITÀ

Come si è scritto nella Parte I, TALEIA (Boncori, 2007a) nasce come un insieme di test paralleli fra loro. Ogni singola forma di TALEIA è scarsamente omogenea al suo interno, in quanto articolata in scale differenziate per contenuto, ognuna delle quali è definita contenutisticamente dal riferimento a criteri diagnostici che, se non fossero molteplici e differenziati, non avrebbero motivo di coesistere e non consentirebbero di diagnosticare disturbi diversi tra loro. L'attendibilità va quindi studiata come concordanza tra forme parallele e come stabilità delle misure prese sugli stessi soggetti in tempi diversi.

Per quel che riguarda la stabilità al retest, che riteniamo fondamentale sia per le applicazioni cliniche, sia per quelle all'orientamento e al lavoro, sono stati compiuti vari studi (Boncori, 2007b). Quello, inedito, di cui qui riportiamo i risultati, è stato compiuto somministrando la forma TALEIA-400A a un campione di studenti del primo anno di un Corso di Laurea in Scienze e tecniche psicologiche¹, che non avevano mai frequentato corsi di psicopatologia. La situazione era di tipo didattico: il test era presentato come un esempio, da commentare successivamente, ed era somministrato sotto pseudonimo. Gli studenti non vennero avvertiti che ci sarebbe stato un retest, per evitare di suscitare ansia e dinamiche conscie o inconscie di memorizzazione degli item per dimostrare stabilità. Quando il test venne somministrato di nuovo otto giorni dopo, ovviamente, mancavano alcuni degli studenti presenti alla prima somministrazione, ma il numero degli studenti presenti in ambedue le occasioni ($N = 123$, mentre alla prima somministrazione avevamo $N = 139$) era sufficiente per il calcolo di coefficienti di correlazione stabili (tab. 1).

¹ Il campione della prima somministrazione era costituito da 139 soggetti, di cui 18 maschi e 121 femmine; alla seconda somministrazione erano presenti 18 maschi e 105 femmine.

TAB. 1. *Correlazioni test-retest (una settimana di distanza)*

Scale	S	D	M	AA	FO	SOC	AG	AL	PP	PSK	PSKT
N. di item	53	86	73	29	42	29	51	45	32	32	53
r	0.81	0.86	0.87	0.82	0.88	0.77	0.84	0.83	0.79	0.83	0.86
p	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
Scale	PAS	PB	PI	PN	PEV	PD	POC	L	F	INC	
N. di item	53	61	38	41	33	21	34	81	65	20	
r	0.84	0.80	0.87	0.88	0.87	0.77	0.80	0.79	0.85	0.66	
p	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	

I coefficienti di correlazione fra scale sono risultati alti (in media², $r = 0.83$), soprattutto per un test di personalità. In campioni analoghi sono stati ottenuti risultati simili (Boncori, 2007b).

Quanto alla correlazione tra forme parallele, riportiamo qui, per brevità, i risultati di un solo studio, inedito, che mette a confronto le due forme finora pubblicate: Difesa Test (300 item) e TALEIA-400A (400 item). Anche in questo caso, il campione era costituito da studenti del primo anno di un Corso di Laurea in Scienze e tecniche psicologiche³, che non avevano mai frequentato corsi di psicopatologia, in situazione didattica analoga alla precedente. In questa occasione i due questionari vennero somministrati nella stessa giornata, separando le due somministrazioni con un intervallo di 20 minuti. Dopo la prima somministrazione (Difesa Test) alcuni dei 150 studenti che componevano il campione iniziale non tornarono in aula e il campione si ridusse a 139 studenti (22 M e 117 F) (tab. 2).

Il valore medio dei coefficienti di correlazione fra le due forme parallele, di diversa lunghezza (300 item il DT e 400 il TALEIA-400A), è $r = 0.74$ e tutti i coefficienti sono statisticamente significativi al livello $p < 0.01$.

L'attendibilità di TALEIA, a parità di metodi di misura, appare non inferiore a quella di test di personalità largamente diffusi e pubblicati recentemente. Per quel che riguarda la stabilità nel retest a distanza di tempo, a fronte del valore $r = 0.83$ riguardante TALEIA-400A, il 16-PF-5 (forma statunitense: Russell e Karol, 2001) presenta un valore medio a distanza di due settimane $r = 0.80$ e il MMPI-2, a distanza di una settimana di tempo e sempre nella forma statunitense (Graham, 1990, p. 175), $r = 0.82$ su un piccolo campione maschile ($N = 82$) e $r = 0.79$ su un piccolo campione femminile ($N = 111$). Per quel che

² Le medie tra coefficienti di correlazione r di Bravais-Pearson sono sempre state calcolate dopo la trasformazione dei valori r in valori z di Fisher, ritrasformando poi in r i risultati ottenuti.

³ Il campione della prima somministrazione era costituito da 150 soggetti, di cui 24 maschi e 126 femmine.

TAB. 2. Correlazioni fra il Difesa Test (DT) e TALEIA-400A (N = 139)

Scale	S	D	M	AA	FO	SOC	AG	AL	PP	PAS	PB	PI	PN	PEV	PD	POC	L	F	INC
N. di Difesa Test	33	59	36	25	33	19	27	14	23	34	53	17	27	22	16	26	18	147	20
TALEIA-400A	53	86	73	29	42	29	51	45	32	53	61	38	41	33	21	34	81	65	20
r	0.85	0.91	0.64	0.83	0.83	0.78	0.84	0.54	0.74	0.79	0.78	0.38	0.49	0.82	0.56	0.50	0.83	0.87	0.36
p	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	.01	0.01	0.01	0.01	0.01	0.01



riguarda i coefficienti di correlazione tra forme parallele, mentre per DT/TALEIA-400A il valore medio è $r = 0.74$, la media dei coefficienti di correlazione fra 16-PF-5 e 16 PF-A (forma statunitense: Russell e Karol, 2001; mancano dati su popolazione italiana) è $r = 0.52$ e, ovviamente, non ci sono dati per il MMPI-2 che non ha forme parallele.

3. VALIDITÀ DIAGNOSTICA

La verifica della validità del costrutto di TALEIA, anche se è cominciata da circa dieci anni, può dirsi ancora all'inizio. Si è cercato però di stabilire delle priorità che garantissero le applicazioni pratiche degli strumenti nel momento della loro pubblicazione. Qui viene data una breve sintesi di due ricerche inedite (ciascuna delle quali è stata oggetto di repliche, pubblicate nel manuale del test: Boncori, 2007b) che miravano a verificare:

1. l'ipotesi che in un confronto con altri test di personalità la struttura fattoriale di TALEIA evidenziasse una componente importante in comune con le scale di altri test psicopatologici e una componente di minor rilievo che evidenziasse la continuità fra alcune scale di TALEIA e tratti normali di personalità; non sono stati compiuti studi con tecniche confirmatorie, non essendo disponibili in letteratura dati adeguati a cui riferirsi;

2. l'ipotesi che le scale di TALEIA consentissero una discriminazione significativa tra i punteggi ricevuti da soggetti normali («certificati» come tali), soggetti in situazioni di vita comune, soggetti diagnosticati come patologici, soggetti in situazione di selezione del personale.

3.1. *Analisi delle componenti principali di TALEIA con altri test*

I risultati di TALEIA sono stati sottoposti ad ACP insieme ai risultati ottenuti dagli stessi soggetti in altri test che misuravano tratti di personalità normale o disturbi, per evidenziare quali fattori latenti fossero in comune tra TALEIA e gli altri due tipi di test. Fra i vari studi di questo tipo compiuti (vedi Boncori, 2007b), qui ne presentiamo uno basato su dati raccolti in una situazione reale di orientamento al termine della scuola secondaria superiore. I soggetti erano 299, di cui 156 maschi e 143 femmine, che frequentavano l'ultimo anno di scuola (76%) o il penultimo nella città di Roma o in cittadine della provincia di Roma. Tutti avevano chiesto volontariamente di partecipare a un intervento di orientamento universitario. Insieme a TALEIA (nella forma D3) furono somministrati il 16 PF Forma A (Cattell, 1949-

1980), che dichiaratamente misura tutti e soltanto i tratti della «sfera della personalità normale» e il PNP (Pichot, 1963), che è un test di *screening* psicopatologico generale. Alle scale tradizionali del 16 PF venne aggiunta una scala di Ansia elaborata in seguito ad uno studio su Ss. italiani (Boncori, 1979), mentre venne esclusa dalle analisi la scala B (Intelligenza) del 16 PF. I soggetti che risposero in modo completo a tutti e tre i test erano in tutto 280, con le proporzioni di M e F, ultimo e penultimo anno invariate rispetto al campione originario. I 39 punteggi dei tre test vennero inseriti in un'unica matrice, sottoposta ad ACP. Questa evidenza di dieci fattori con autovalore $\Rightarrow 1$, che in base allo *scree test* vennero ridotti a quattro fattori che nell'insieme spiegavano il 53% della varianza e che furono sottoposti a rotazione Varimax normalizzata (tab. 3).

Le saturazioni nei quattro fattori, riportate nella tabella 3, evidenziano bene la distinzione fra scale riferibili a fattori latenti connessi con disturbi, scale riferite a tratti normali di personalità, e scale che sembrano valutare un continuum fra normalità e patologia.

Limitando l'interpretazione alle saturazioni $> |0,25|$ possiamo descrivere i quattro fattori nel modo seguente.

– *Il Fattore 1°*, che spiega l'11% della varianza, al polo positivo si caratterizza per saturazioni elevate di tutte le scale «cliniche» di TALEIA, ad eccezione della scala POC, della scala di controllo F (che misura la presenza di un elevato numero di sintomi psicopatologici), di tutte e tre le scale «cliniche» del PNP, della scala ANS (Ansietà) del 16 PF; al polo negativo si caratterizza per la presenza di saturazioni elevate nelle due scale L (del PNP e di TALEIA) che valutano la tendenza a mettere in evidenza caratteristiche «positive» di personalità; possiamo quindi definirlo un fattore di «Disagio psicopatologico generale», articolato in un continuum che va dall'Ansietà come tratto «normale» (ANS) ai disturbi di personalità e ai disturbi del I Asse.

Più difficili da interpretare appaiono i rapporti fra Disturbi di personalità e disturbi classificati nell'Asse I da un lato, e dall'altro tratti della personalità normale.

– *Il Fattore 2°*, che spiega il 4% della varianza, al polo positivo è caratterizzato dall'elevata saturazione di scale che prevalentemente descrivono disturbi dello «spettro ansioso»: FO (Fobie), PD (DP Dipendente), PEV (DP Evitante), POC (DP Ossessivo-compulsivo), mentre al polo negativo è caratterizzato prevalentemente da scale che nella teoria di Cattell si riferiscono al «fattore Q_1 di II strato Invia», comunemente noto come Introversione (F, umore cupo; H, mancanza di disinvoltura; E, sottomissione) e che nella nosografia del DSM-IV descrivono valori bassi nella scala M (Mania e Ipomania) e nelle scale PI e PN, il che nell'insieme equivale a dire mancanza di disinvoltura

TAB. 3. *Analisi delle componenti principali di TALEIA e di altri due test di personalità*

Scale	TALEIA A / D3			
	Fattore 1	Fattore 2	Fattore 3	Fattore 4
A	0.03	0.13	0.03	-0.39
C	-0.14	-0.09	-0.59	0.28
E	0.03	-0.33	-0.26	0.49
F	0	-0.62	-0.01	0.19
G	-0.02	0.31	-0.38	-0.09
H	-0.05	-0.6	-0.21	0.12
I	0.02	0.19	0.46	0.05
L	0.16	0.16	0.41	0.36
M	0.11	-0.12	0.57	0.37
N	-0.05	0.02	-0.02	0.29
O	0.19	0.19	0.81	-0.09
Q ₁	0.03	-0.06	0.13	0.5
Q ₂	-0.04	0.06	0.18	0.64
Q ₃	-0.22	0.18	-0.65	-0.08
Q ₄	0.23	-0.01	0.75	0.02
ANS	0.26	0	0.81	0.02
PNP	-	-	-	-
N ₁	0.68	0.26	0.16	-0.08
N ₂	0.5	0.12	0.12	-0.04
P	0.72	0.09	0.1	0.01
L	-0.33	0.24	0.04	-0.24
	Fattore 1	Fattore 2	Fattore 3	Fattore 4
	0.81	-0.04	0.14	0.14
	0.87	0.27	0.16	0.16
	0.54	-0.56	0.11	-0.07
	0.8	0.09	0.22	-0.17
	0.45	0.68	0.06	0.04
	0.7	0.22	0.09	-0.14
	0.83	0.17	0.19	-0.12
	0.58	-0.01	0.14	-0.2
	0.79	0.29	0.07	0.23
	0.82	-0.19	0.08	0.19
	0.93	0.04	0.13	0.12
	0.48	-0.58	0.02	0.04
	0.66	-0.33	0.02	0.07
	0.53	0.63	0.05	0.2
	0.46	0.60	0.12	-0.15
	0.21	0.48	-0.11	0.02
	-0.78	0.08	-0.13	-0.25
	0.94	0.11	0.16	-0.01
	0.52	-0.01	0.19	-0.16
	1.8	3.79	4.02	2.03
	0.28	0.1	0.1	0.05
	Scale	S	D	M
		AA	FO	SOC
		AG	AL	PP
		PAS	PB	PI
		PN	PEV	PD
		POC	L	F
		INC	Var. Sp	Prp.Tot.

sociale. Troviamo quindi continuità fra tratti «normali» e «patologici», con prevalenza dei disturbi del II Asse, in un quadro generale che potremmo definire «Introversione sociale ansiosa contrapposta a Tendenza a mettersi in evidenza».

– Il *Fattore 3°*, che spiega anch'esso il 4% della varianza, si caratterizza per la saturazione elevata soltanto di scale che descrivono tratti della personalità normale, in parte attinenti al fattore Q_{II} di II s. denominato da Cattell «Adattamento-Ansietà» (C, D, Q3, Q4) in parte al fattore Q_{VIII} «Forza del Super-Io» (G, Q3, F) e in parte al fattore Q_{VI} , Realismo (I, M, L). Potremmo denominarlo «Capacità di regolazione degli affetti e degli impulsi».

– Anche il *Fattore 4°*, che spiega il 2% della varianza, si caratterizza esclusivamente per elevate saturazioni in scale che descrivono tratti normali (fra questi possiamo includere anche la tendenza a fare buona impressione misurata dalla scala L di TALEIA). Tutte le scale del 16 PF con saturazioni maggiori di |.25| riguardano in qualche modo la gestione dei rapporti con gli altri: Q_2 (Autosufficienza sociale), Q_1 (Innovazione nelle strutture sociali), E (Dominanza), A (Contatto affettivo con gli altri), L (Sospettosità), M (Mentalità pratica), N (Scaltrezza), C (Forza dell'Io). Potremmo quindi denominarlo «Efficienza sociale».

A margine dell'interpretazione dei singoli fattori osserviamo alcuni dati interessanti, pro e/o contro la continuità fra tratti normali e disturbi:

– A (Affettotimia), che al polo negativo è denominato da Cattell «Sizotimia», in quanto estensione in un tratto normale dei «sintomi negativi» della Schizofrenia, ha la sua saturazione più alta (-0.39) nel 4° fattore, in cui la scala S, Schizofrenia, ha una saturazione bassa e di segno opposto (0,10);

– nella dimensione «Umore» è evidente, all'interno del fattore 2°, la continuità fra dimensione «normale» (16 PF: F e H) e «patologica» (TALEIA: D e M);

– la scala ANS del 16 PF, che dovrebbe valutare il tratto di Ansietà sotto l'aspetto normale, ha una duplice collocazione (com'era prevedibile): la sua saturazione più elevata (0.81) è nel 3° fattore, insieme a vari tratti normali che hanno una componente di ansia, ma anche una componente non trascurabile (0.26) nel 1° fattore, dove si collocano quasi tutte le scale psicopatologiche: c'è quindi una forma di continuità fra ansia patologica e ansietà normale, ma all'interno di una struttura bifattoriale della scala di Ansietà elaborata all'interno del 16 PF;

– E (Dominanza) ha la sua saturazione più alta (0.49) nel 4° fattore, in cui PD (DP Dipendente) ha una saturazione di segno opposto (-0.15), ma di valore modesto; le due scale si contrappongono anche all'interno del 2° fattore, dove E ha una saturazione più mode-

sta (-0.33) e PD più elevata (0.60): questo fa pensare a una duplicità fattoriale della dimensione sottostante, in cui a seconda delle componenti considerate può essere dominante una componente patologica (che abbiamo chiamato «Introversione sociale») o una componente sana, di «Efficienza sociale».

Per quel che riguarda la struttura interna di TALEIA, essa appare sostanzialmente monofattoriale, con l'unica eccezione della scala POC (DP Ossessivo-compulsivo) che ha saturazioni ugualmente importanti in due diversi fattori. Non c'è però una sostanziale distinzione fattoriale fra scale attinenti all'Asse I e all'Asse II. Così pure ci sono vari segni di continuità fra dimensioni normali e patologiche di uno stesso costrutto.

Questa struttura fattoriale è stata sostanzialmente confermata da studi analoghi, talvolta compiuti con test di personalità diversi da quelli a cui si riferiscono i dati nella tabella 3 e in situazioni diverse da quella orientativa (Boncori, 2007b). Verifiche sulla struttura fattoriale del test sono attualmente in corso su versioni in lingua polacca⁴, in lingua spagnola⁵ e ispano-americana⁶ e hanno dato sostanziali conferme, per ora inedite.

3.2. *Validità diagnostica delle scale cliniche e delle scale di controllo «fake bad»*

Il secondo obiettivo delle ricerche sulla validità del costrutto è stato verificare se TALEIA consentisse di distinguere con buon margine di attendibilità tra soggetti patologici e non patologici, includendo nello studio il problema della simulazione della patologia e, implicitamente, il problema della validità delle scale F e INC che dovrebbero contribuire a segnalare. Uno studio che riteniamo insostituibile per illustrare la validità diagnostica del test ha coinvolto 486 persone (382 uomini e 91 donne) con una diagnosi, formulata da medici specializzati operanti in strutture prevalentemente pubbliche, di patologia clinica o di disturbo di personalità. In alcuni casi (49 M) la diagnosi era stata emessa in forma dubitativa in un primo momento (tipicamente: in sede di visita di leva) e successivamente era stata sottoposta a revisione, spesso dopo un soggiorno ospedaliero. Per questo

⁴ Ringraziamo il prof. B. Zawadzky, Università di Varsavia, e la psicologa A. Puchtinska.

⁵ Ringraziamo la prof. C. Jenaro Río, Università di Salamanca, e la psicologa F. Frau.

⁶ Ringraziamo la prof. M. Fernández Liporace, Università di Buenos Aires, e la psicologa C. Tardugno.

motivo il campione «patologico» di 486 soggetti si è ridotto, di fatto, a 436 soggetti, che nell'insieme risultavano affetti da disturbi riferiti a più di 50 diversi codici del DSM-IV⁷. Nel campione, ovviamente, sono meglio rappresentati i disturbi frequenti (forme di ansia, disturbi dell'umore...) e sono assenti o quasi i disturbi più rari, come peraltro avviene anche in campioni più numerosi (es. Samuels e Eaton, 2002). Sono stati volontariamente esclusi Ss. con diagnosi principale di abuso di sostanze, tipicamente oggetto di diagnosi multiple (cfr. Nathan e Langenbucher, 1999), in quanto ci era noto che l'accertamento di questa condizione con analisi biologiche induceva di per sé all'esclusione dal servizio militare e quindi automaticamente rendeva gli operatori meno motivati a formulare una diagnosi completa.

La frequenza delle comorbidità nel nostro campione è più bassa rispetto a quella segnalata nelle ricerche internazionali. Gli incontri con gli operatori ci hanno documentato peraltro una ridotta diffusione dei manuali di riferimento delle nosografie DSM-IV e ICD-10 nelle strutture e l'ampia diffusione della tendenza a ritenere errata l'attribuzione di più di una patologia allo stesso soggetto. Da quel che risulta dagli incontri con gli operatori, questa convinzione di fondo ha indotto molte volte a esprimere una sola diagnosi riferita a una categoria ritenuta più ampia (e quindi implicitamente più valida) dove invece il test evidenzia problematiche più articolate. Ad esempio, spesso a fronte di punteggi alti in più scale dello «spettro ansioso» del I Asse (Ansia acuta, Fobie, Sindrome ossessivo-compulsiva...) e/o del II Asse (disturbo di personalità evitante, dipendente, ossessivo-compulsivo o anancastico) troviamo una diagnosi di «Ansia generalizzata», che parlando con l'operatore risulta formulata così perché ritenuta una sintesi dei vari disturbi evidenziati dalle scale di TALEIA. Questa tendenza «semplificatoria» può essere in rapporto anche con la scarsa presenza di diagnosi riferite all'Asse II: 26% in tutto, quasi sempre come diagnosi aggiuntiva rispetto ad una di I Asse. Peraltro, il fatto che le comorbidità vengano sottostimate dagli operatori è segnalato anche nella ricerca internazionale (cfr. Weaver, Madden, Charles., Stimson e Renton, 2003).

I campioni da mettere a contrasto con il campione clinico appena descritto sono stati: 1) un campione di Ss. che potevano rigorosamente essere definiti «normali certificati» in quanto selezionati come idonei alla visita di Leva in base ai risultati di MMPI, colloquio, visita medica: 773 soggetti, tutti maschi (all'epoca le donne non venivano

⁷ Le diagnosi sono state riscontrate una per una sui dati delle cartelle cliniche da S. Nicole (picologo e psichiatra), che ha ricodificato tutti i disturbi appropriatamente, eliminando gli errori che sarebbero stati indotti dalla presenza di diagnosi «compassionevoli», presenti soprattutto nel campione dei giovani esaminati in sede di visita di Leva.

ТАВ. 4. *Differenze in TALEIA (forma D3) fra campioni di uomini selezionati («Arruolati»), non selezionati («Visita leva» e «Stud.Sup.») e con diagnosi psicopatologica*

Scale	M.A.											
	Arr (1)	Leva RM (2)	StudScSu (3)	Pato (4)	Diff. (1) (2) Arr/Leva	Diff. (1) (2) Arr/Stud	Diff. (1) (3) Arr/Stud	Diff. (1) (4) Arr/Pat	Leva/Stud	Diff. (2) (3) Leva/Stud	Diff. (2) (4) Leva/Pat	Diff. (3) (4) Stud/Pat
AA	34.78	56.41	36.25	54.66	0**	0.59 ^{ns}	0**	0**	0**	0.08 ^{ns}	0**	0**
AG	52.26	77.84	54.99	72.86	0**	0.22 ^{ns}	0**	0**	0**	0**	0**	0**
AL	34.59	45.35	35.4	39.09	0**	0.66 ^{ns}	0**	0**	0**	0**	0**	0**
B	119.38	156.4	131.01	156.28	0**	0**	0**	0**	0**	1 ^{ns}	0**	0**
D	126.07	157.53	135.18	171.53	0**	0**	0**	0**	0**	0**	0**	0**
EV	60.36	74.35	62.15	79.24	0**	0.33 ^{ns}	0**	0**	0**	0**	0**	0**
FO	91.1	117.12	89.72	111.37	0**	0.82 ^{ns}	0**	0**	0**	0**	0**	0**
M	121.64	134.6	129.78	128.66	0**	0**	0**	0**	0.01*	0**	0**	0.89 ^{ns}
OC	65.42	72.54	67.18	73.01	0**	0.38 ^{ns}	0**	0**	0**	0.9 ^{ns}	0**	0**
PAS	90.58	113.81	97.37	112.58	0**	0**	0**	0**	0**	0.63 ^{ns}	0**	0**
PD	51.37	57.59	53.38	63.09	0**	0.1 ^{ns}	0**	0**	0**	0**	0**	0**
PI	39.41	46.34	41.87	41.19	0**	0**	0**	0**	0**	0**	0**	0.77 ^{ns}
PN	70.46	75.93	75.41	71.98	0**	0**	0**	0.26 ^{ns}	0.98 ^{ns}	0**	0**	0.05 ^{ns}
POC	93.55	91.20	90.83	91.36	0**	0.05 ^{ns}	0**	0**	0.98 ^{ns}	0.99 ^{ns}	0**	0.96 ^{ns}
PP	54.05	70.55	58.85	69.49	0**	0**	0**	0**	0**	0.47 ^{ns}	0**	0**
S	65.96	89.04	75.84	89.43	0**	0**	0**	0**	0**	0.97 ^{ns}	0**	0**
F	227.07	339.77	248.04	324.08	0**	0**	0**	0**	0**	0**	0**	0**
L	93.36	84.79	87.30	83.37	0**	0**	0**	0**	0.08 ^{ns}	0.15 ^{ns}	0**	0**
INC	5.26	18.27	5.88	16.44	0**	0.5 ^{ns}	0**	0**	0**	0**	0**	0**

* Statisticamente significativo oltre p = 0.05

** Statisticamente significativo oltre p = 0.01

arruolate nell'Esercito Italiano) che svolgevano il servizio di Leva in quattro diverse città italiane; 2) un campione (N = 386) di ragazzi esaminati in sede di visita di leva, residenti a Roma e provincia, «normali non certificati», in una situazione che tipicamente induceva a simulare patologie per evitare il servizio militare; 3) un altro campione (N = 156 maschi, tratti da un campione complessivo in cui erano presenti anche N = 143 ragazze, escluse per evitare il sesso come variabile di disturbo) di «normali non certificati» che frequentavano l'ultimo o il penultimo della scuola secondaria superiore a Roma e provincia e avevano chiesto di essere sottoposti a test nell'ambito di un servizio di orientamento, e cioè in una situazione che può indurre a cercar di fare «buona figura».

Il confronto tra campioni è stato compiuto mediante una MANOVA (condotta con il software «Statistica»), il cui risultato ha evidenziato un'elevata significatività ($p < 0.001$) del fattore principale. Per $df_1 = 57$ e $df_2 = 4837$, il valore Lambda di Wilks era $\text{Lambda} = 0.09$ e il valore della R di Rao⁸ $R = 104.02$. I campioni hanno varianze disomogenee in tutte le scale, come è prevedibile quando si confrontano tra loro campioni esaminati in situazioni diverse, dato che notoriamente la diversità fra situazioni è in rapporto con la componente della desiderabilità sociale specificamente connessa alle situazioni o *response bias* (Nunnally e Bernstein, 1994 e discussione a questo riguardo nella prima parte del presente lavoro). È stato peraltro dimostrato che la statistica F risulta estremamente robusta rispetto alla violazione di questo assunto (Edwards, 1960; Lindman, 1974, Lindquist, 1953)⁹. Le analisi *post-hoc*, compiute applicando il test HSD (*Honest Significant Difference*) di Tukey corretto per la diversa numerosità dei campioni (test di Spjotvoll-Stoline) ha dato i risultati riportati nella tabella 4.

I risultati nella tabella 4 sono meglio interpretabili se si parte dalle scale di controllo.

La scala F, che assume valori tanto più alti quanto più i Ss. manifestano caratteristiche psicopatologiche, ha i suoi valori medi più alti nel campione esaminato alla visita di leva, congruentemente con la funzione per cui la scala è stata costruita. Più bassi di una decina di punti sono i valori medi del «campione psicopatologico» e a maggior

⁸ La statistica R di Rao equivale a F (con gli appropriati gradi di libertà) quando è calcolata su un modello ANOVA univariato, quando l'analisi si basa un qualunque numero di gruppi con due variabili indipendenti, su due gruppi con un qualunque numero di variabili indipendenti, su tre gruppi con un qualunque numero di variabili. In tutti gli altri casi, si considera un'approssimazione di F per i gdl considerati (Tatsuoka e Lohnes, 1988).

⁹ Per ulteriore garanzia, sono stati anche calcolati tutti i possibili t di Student per campioni con varianze significativamente diverse, ottenendo sempre differenze tra medie statisticamente significative oltre il livello $p = 0.001$.

distanza si collocano i punteggi medi degli studenti e degli arruolati. Tutte le differenze fra subcampioni sono statisticamente significative oltre il livello $p = 0.01$ e in direzione congruente con le ipotesi.

La scala L, che al contrario assume valori tanto più alti quanto più i Ss. evidenziano caratteristiche positive, differenzia significativamente il campione degli arruolati da tutti gli altri e il campione degli studenti dal campione psicopatologico, ma non evidenzia differenze significative tra i giovani alla visita di leva e il campione psicopatologico. Questo dato giustifica la prassi di riferirsi a norme separate per valutare le risposte dei giovani alla visita di leva e in genere in situazione di selezione lavorativa.

La scala INC, che valuta l'attenzione e l'accuratezza nel rispondere (evidenziata dalla proporzione più o meno alta di contraddizioni interne) ha valori medi non significativamente diversi e bassi per arruolati e studenti, significativamente più elevati per giovani alla visita di leva e campione psicopatologico. Anche questo quadro sembra congruente con le caratteristiche dei campioni e delle situazioni considerate.

Con queste premesse, esaminiamo le differenze tra i punteggi medi delle scale cliniche, ponendo come problema centrale le differenze fra campione psicopatologico e subcampioni normali.

I valori medi del campione psicopatologico:

- rispetto al subcampione degli arruolati (selezionati e con elevata tendenza a ben figurare: v. scala L) differiscono significativamente nella direzione attesa in tutte le scale tranne PN (Disturbo di Personalità Narcisistico), con cui la differenza è nella direzione attesa ma di modesta entità e POC (Disturbo di Personalità ossessivo-compulsivo o Anancastico), con cui la differenza è statisticamente significativa, ma in direzione opposta al previsto;

- rispetto al subcampione degli studenti (non selezionati e con più accentuata tendenza a evidenziare i propri mali: v. scala F) in 12 scale le differenze sono statisticamente significative oltre il livello $p = 0.01$ e nella direzione attesa, mentre in quattro scale le differenze sono al di sotto della soglia di significatività $p = 0.05$: M (Mania e ipomania), PI (D.P. Istrionico), PN (D.P. Narcisistico) e POC (Disturbo di Personalità ossessivo-compulsivo o Anancastico);

- rispetto al subcampione dei giovani di leva (non selezionati e con spiccata tendenza a evidenziare caratteristiche psicopatologiche) i valori medi delle scale sono significativamente diversi e nella direzione attesa in quattro scale (D, Depressione; M, Mania e ipomania; PEV, Disturbo di Personalità Evitante), significativamente diversi ma più elevati di quelli del campione patologico in sei scale (AG, Ansia Generalizzata; AL, Disturbi del comportamento alimentare; FO, Disturbi

fobici; PD, Disturbo di Personalità Dipendente; PI, DP Istrionico; PN, DP Narcisistico) e non significativamente diversi in sette scale.

Le poche scale in cui campioni «normali» e «patologici» non si differenziano significativamente sono quelle riferite a diagnosi scarsamente o per nulla rappresentate nel campione dei «patologici». Questo è il caso di PEV (nessuna diagnosi), M (5 diagnosi), PI (10 diagnosi), PN (10 diagnosi), POC (3 diagnosi di questo tipo su 486 Soggetti clinici.). L'assenza di diagnosi esplicita non esclude la presenza di tratti del disturbo, ma comunque consente di inquadrare l'assenza di differenze significative come un elemento non molto rilevante.

Nell'insieme, appare corroborata l'ipotesi della validità diagnostica generale del test, anche se è evidente la sensibilità del test alle situazioni e quindi si conferma la necessità sia di adottare norme separate per situazioni, sia di interpretare per prima cosa i valori delle scale «di controllo».

4. VERIFICA DELLA VALIDITÀ IN AMBITO SELETTIVO E DELLA VALIDITÀ DELLA SCALA DI CONTROLLO *FAKE GOOD*

Il problema della validità di un test per l'uso selettivo è strettamente dipendente dalla capacità del test di differenziare chi risponde riferendosi a se stesso rispetto a chi risponde riferendosi a stereotipi che descrivono il profilo ideale delle persone da selezionare. Dato che TALEIA era costruito anche per questo tipo di applicazione, sono stati progettati più studi per verificarne la validità in situazione selettiva. In questa parte del lavoro è stata fondamentale la collaborazione con la Direzione Generale della Sanità di Pubblica Sicurezza¹⁰. Fra i molti studi compiuti e le successive verifiche (Boncori, 2007b), pubblichiamo qui i risultati di uno studio inedito, articolato in due momenti successivi, in cui ci siamo proposti di verificare i limiti del metodo basato sulla simulazione di situazioni ai fini di valutare l'affidabilità delle risposte ottenute in situazione concorsuale. Il metodo scelto è stato la MANOVA, calcolata con il software Statistica, che in un primo momento è stata applicata alla verifica della rilevanza di due fattori (sexo dei soggetti e istruzioni date nella somministrazione del test) e in un secondo momento è stata applicata al confronto fra le risposte date dal campione di studenti che avevano risposto sotto l'istruzione di simulare la condizione concorsuale e le risposte date da un campione di concorrenti reali.

¹⁰ Oltre al dott. G. Cuomo, coautore del presente lavoro, va un sentito ringraziamento al dirigente superiore dott. L. Lucchetti, neuropsichiatra, e al dott. U. Ungaro, psicologo.

TAB. 5. *Differenze per sesso e istruzioni di somministrazione*



Effetto	Lambda di Wilks	R di Rao	gdl 1	gdl 2	p
{1}	0.758*	7.547*	21*	497*	< 0.0001**
{2}	0.785*	6.478*	21*	497*	< 0.0001**
{12}	0.948	1.308	21	497	0.163

* ~~??~~

** Statisticamente significativo oltre $p = 0.01$.

Nella prima fase dello studio, TALEIA è stato somministrato collettivamente a un campione di studenti che frequentavano l'ultimo o il penultimo anno di scuola secondaria superiore in scuole pubbliche di Viterbo¹¹ (N = 339 F e N = 181 M). È stato chiesto di apporre sul foglio di risposta uno pseudonimo, ritenendo questa istruzione una sufficiente garanzia per la protezione dell'anonimato da un lato e per la sicurezza nell'abbinamento delle risposte date alle due somministrazioni dall'altro. In una prima somministrazione sono state usate le istruzioni di somministrazione standard e in una seconda somministrazione istruzioni con cui si chiedeva di simulare una situazione concorsuale, sintetizzabili in «rispondi pensando che stai partecipando ad un concorso per un posto di lavoro, per esempio per entrare in un grande Ente pubblico, come la Polizia di Stato». Le due istruzioni sono state date in ordine diverso a «blocchi» diversi del campione, per controllare l'effetto dell'ordine in cui le due situazioni si alternavano.

I punteggi ottenuti dagli studenti con le due diverse istruzioni sono stati sottoposti ad una MANOVA per prove ripetute, inserendo come variabile indipendente il sesso e come fattori ripetuti i punteggi nelle 21 scale di TALEIA.

I risultati dell'analisi (tab. 5) evidenziano un effetto significativo ($p < 0.0001$) sia del fattore Sesso sia del fattore Condizioni di somministrazione, e un'interazione fra i due inferiore al livello di significatività $p = 5\%$. In base a questi risultati, nella seconda fase dello studio il confronto tra studenti simulatori e concorrenti reali è stato condotto a parità di sesso, limitatamente al campione femminile. Per quel che riguarda l'effetto delle Condizioni di somministrazione, la significatività delle differenze ($p < 0.0001$) nell'insieme delle 21 scale sembra corroborare la prassi di validare i questionari con il metodo della simulazione.

¹¹ Il test è stato somministrato dal dott. Paolo Salotti, che lo ha utilizzato per uno studio sulle problematiche delle scale di controllo, in un contesto molto più ampio e nel quadro di una discussione critica della letteratura, presentandolo come tesi di Diploma *post lauream* in Valutazione Psicologica: P. Salotti, *Verifiche empiriche di validità delle scale di controllo in questionari psicopatologici*. Roma, Università «La Sapienza», a.a. 2001-02.

TAB. 6. *Differenze fra punteggi di studentesse dopo istruzione di simulare la partecipazione a un concorso e di vere concorrenti a un concorso pubblico (agente PS)*

Scale	Punteggi medi		Significatività differenze	
	Studentesse simulatrici {1}	Concorrenti per Agenti PS {2}	Differenze tra {1} e {2}	p
S	115.14	81.39	33.75	0
D	208.22	144.82	63.4	0
M	204.58	189.75	14.83	0
AA	60.37	35.84	24.53	0
FO	98.03	68.53	29.5	0
SOC	75.82	56.59	19.23	0
AG	120.17	74.37	45.8	0
AL	117.32	97.05	20.27	0
PP	79.68	57.5	22.19	0
PSK	81.86	78.23	3.64	0
PSKT	119.12	81.41	37.71	0
PAS	118.92	85.94	32.98	0
PB	131.03	93.48	37.54	0
PI	99.14	88.19	10.95	0
PN	109.68	92.04	17.64	0
PEV	83.58	63.86	19.72	0
PD	58.45	53.74	4.71	0
POC	104.65	102.23	2.42	0.026
L	297.83	358.01	-60.18	0
F	145.85	92.69	53.16	0
INC	14.72	3.51	11.22	0

Le risposte date dalle studentesse (N = 339) sotto l'istruzione di simulare la condizione concorsuale sono state confrontate con le risposte date da un campione di giovani donne che partecipavano a un concorso reale per posti di agente nella Polizia di Stato (N = 207). I risultati della MANOVA per l'effetto principale sono significativi al livello $p < 0.001$: Lambda di Wilks = 0.31, R di Rao = 55.45. I successivi test post hoc sono stati calcolati applicando il test HSD (*Honest Significant Different*) di Tukey, corretto per la diversa numerosità dei gruppi (test di Spjotvoll-Stoline). I risultati dei test post hoc sono stati riportati nella tabella 6.

Se si confrontano i punteggi delle studentesse con quelli delle concorrenti, si osserva che le differenze riguardanti tutte le scale sono statisticamente significative ($p < 0.001$). Sotto l'aspetto pratico, per le applicazioni alla psicologia del lavoro, la presenza di differenze statisticamente significative fra condizioni di simulazione e di partecipazione reale a una selezione induce a considerare che la simulazione non fornisca informazioni sufficienti, e che sia necessario compiere verifiche per l'affidabilità del test su concorrenti in situazioni reali. Nella tabella 6 si noti anche la diversa grandezza numerica delle differenze tra medie a seconda delle scale, che induce a ritenere l'effetto

distorsivo controllabile solo con l'applicazione di parametri specifici. Ne consegue la necessità di stabilire per la taratura parametri specifici calcolati tenendo conto della situazione concorsuale.

Studi analoghi per impostazione a quello appena descritto e a quello i cui risultati sono riportati nella tabella 6 sono stati compiuti su campioni di persone di ambo i sessi e di diverse fasce d'età, esaminate in situazioni diverse (Boncori, 2007b; ~~Boncori De Coro, Laganà, Nicole e Renzi, 2009~~). In base alle differenze osservate si è deciso di pubblicare il test con norme differenziate per sesso e per situazione (clinica, orientamento, selezione di 1° livello, selezione di 2° livello) e di pianificare l'aggiornamento continuo dei parametri via software (Boncori e Di Marco, 2007).

5. CONCLUSIONI

Le verifiche sull'attendibilità e la validità dello strumento TALEIA sono state compiute sempre in connessione con le sue caratteristiche di contenuto, avendo in mente la sua finalità di sussidio per diagnosi formulate in riferimento alle nosografie internazionali correnti (DSM-IV e ICD-10). Sono stati anche costruiti specifici sussidi per condurre i colloqui integrandoli con i risultati del questionario, in vista di una diagnosi articolata, che risponda ai requisiti della categorizzazione nosografica, ma anche di una valutazione dimensionale della personalità. I sussidi, integrati nel software, rispondono anche all'esigenza di avere colloqui in qualche misura standardizzati, accorgimento particolarmente utile quando le situazioni hanno implicazioni medico-legali. L'articolo presenta solo una parte dell'ampia serie di ricerche svolte per verificare attendibilità e validità del test, focalizzando aspetti ritenuti fondamentali: attendibilità, validità diagnostica e validità nell'ambito della selezione del personale. In questi primi contributi si è cercato soprattutto di dare l'immagine globale dello strumento e delle sue caratteristiche innovative, che lo rendono adatto ad essere inserito in una batteria per la valutazione psicopatologica insieme a test «classici», come l'MMPI-2 e il MCMI, a tecniche proiettive tradizionali nell'uso clinico, come il Rorschach e il TAT, o a strumenti di nuova concezione come la SWAP, l'OPD e il PDM.

BIBLIOGRAFIA

- AMERICAN PSYCHIATRIC ASSOCIATION (1994). *Diagnostic and statistical manual of mental disorders, fourth edition (DSM-IV)*. Washington, DC: APA (trad. it. *DSM-IV. Manuale diagnostico e statistico dei disturbi mentali*. Milano: Masson, 1995).
- BONCORI L. (1979). Indicatori dell'ansietà nel 16 PF di R.B. Cattell. *Orientamenti Pedagogici*, 26, 651-678.
- BONCORI L. (a cura di) (2007a). *TALEIA-400A: Test for Axial Evaluation and Interview for Clinical, Personnel, and Guidance Applications*. Trento: Erickson.
- BONCORI L. (2007b). *TALEIA-400A: Test for Axial Evaluation and Interview for Clinical, Personnel, and Guidance Applications – Manuale*. Trento: Erickson.
- ~~BONCORI L., DE CORO A., LAGANÀ L., NICOLE S., RENZI P. (2009). *Psychopathological assessment in clinical and workplace applications: An Italian measure. International Journal of Testing*.~~
- BONCORI L., DI MARCO M. (2007). *TALEIA-400A Software*. Trento: Erickson.
- CATTELL R.B., SAUNDERS D.R., STICE G. (1949-1980). *Sixteen Personality Factors Questionnaire*. Champaign, IL: IPAT (trad. it. *IPAT, 16 PF, Forma A e D*. Firenze: Organizzazioni Speciali).
- EDWARDS A.D. (1960) *Experimental design in psychological research*. New York: Holt Rinehart & Winston.
- GRAHAM J.R. (1990). *MMPI-2: Assessing Personality and Psychopathology*. Oxford: Oxford University Press.
- LINDMAN H.R. (1974). *Analysis of variance in complex experimental designs*. San Francisco, CA: W.H. Freeman & Co.
- LINDQUIST E.F. (1953). *Design and analysis of experiments in psychology and education*. Boston, MA: Houghton Mifflin.
- NATHAN P.E., LANGENBUCHER J.W. (1999). Psychopathology: Description and Classification. *Annual Review of Psychology*, 50, 79-107.
- ~~NOVAGA M., PEDON A. (1977). *Contributo allo studio della personalità. il 16 P.F. test di Cattell*. Firenze: Organizzazioni Speciali.~~
- ~~PAULHUS D.L., REID D.B. (1991). *Enhancement and denial in socially desirable responding. Journal of personality and Social Psychology*, 60, 307-317.~~
- PICHOT P. (1963). *Il questionario P.N.P.*. Firenze: Organizzazioni Speciali.
- RUSSELL M.T., KAROL D.L. (2001). *16 PF-5. Manuale*. Firenze: Organizzazioni Speciali.
- SAMUELS J., EATON W.W. (2002). Prevalence and correlates of personality disorders in a community sample. *British Journal of Psychiatry*, 180, 536-542.
- TATSUOKA M.M., LOHNES P.R. (1988). *Multivariate analysis* (2nd ed.). New York: Macmillan.
- WEAVER T., MADDEN P., CHARLES V., STIMSON G., RENTON A. (2003). Comorbidity of substance misuse and mental illness in community mental health and substance misuse services. *British Journal of Psychiatry*, 183, 304-313.
- WORLD HEALTH ORGANIZATION (1992). *The ICD-10 Classification of mental disorders and behavioral disorders: Clinical descriptions and diagnostic guidelines*. Geneva: WHO (trad. it. *Decima revisione della classificazione internazionale delle sindromi e dei disturbi psichici e comportamentali: descrizioni cliniche e direttive diagnostiche*. Milano: Masson, 1992).
- WORLD HEALTH ORGANIZATION (1993). *The ICD-10 classification of mental and behavioral disorders: Diagnostic criteria for research*. Geneva: WHO (trad. it. *Decima revisione della classificazione internazionale delle*

sindromi e dei disturbi psichici e comportamentali: criteri diagnostici per la ricerca. Milano: Masson, 1994).

[Ricevuto il 6 febbraio 2009]

[Accettato il 2 agosto 2009]

Innovation in psychopathological testing: TALEIA. Part I: Reliability and structural validity

Summary. Studies reliability and validity of TALEIA (Test for Axial Evaluation and Interview for Clinical, Personnel, and Guidance Applications). Retest (N = 123, one-week interval) shows average $r = 0.83$, and correlations between parallel forms (N = 139) average to $r = 0.74$. An ACP (N = 280) on TALEIA, 16 PF, and PNP shows a four factors structure explaining 53% of the total variance, the first factor having psychopathological content and the others showing continuity between normal personality traits and clinical or personality disorders. A sample of «certified pathological» (N = 436), «certified normals» (N = 773 males), and two samples of «non-certified normals» (N = 386 draftees and N = 156 male students asking for school guidance) were compared. The MANOVA results were significant ($p < 0.001$), and post-hoc analyses (HSD test corrected for different size) showed significant differences related both to clinical status (pathologicals vs. normals) and to different situations (guidance and military draft). These results seem to allow a professional use of TALEIA, if different parameters for different applications are provided.

Keywords:???



La corrispondenza va inviata a Lucia Boncori, Dipartimento di Psicologia, Sapienza Università di Roma, Via dei Marsi 78, 00185 Roma, e-mail: lucia.boncori@uniroma1.it

