

Il Mini Questionario di Qualità Percepita (QQP): uno studio sulle differenze di genere

ANNA COLUCCIA, FABIO FERRETTI, FRANCESCA LORINI ¹, RAFFAELE CIOFFI²

Introduzione

Il *Mini Questionario di Qualità Percepita (MQQP)* (Coluccia A., Ferretti F., Cioffi R., Lorini F, in corso di stampa 2), ricavato dal *Questionario di Qualità Percepita (MQQP)* (Coluccia, Ferretti, Lorini, Calamai, 2002; Coluccia A, Ferretti F, Cioffi R., Lorini F, Vidotto G, in corso di stampa 1; Coluccia A, Cioffi R, Ferretti F, Lorini F, Vidotto G, in corso di giudizio) si caratterizza come uno strumento realizzato per il monitoraggio della qualità percepita del servizio sanitario ricevuto durante la degenza attraverso il giudizio dei degenti, il cui utilizzo trova forti collegamenti con gli aspetti gestionali delle politiche di qualità

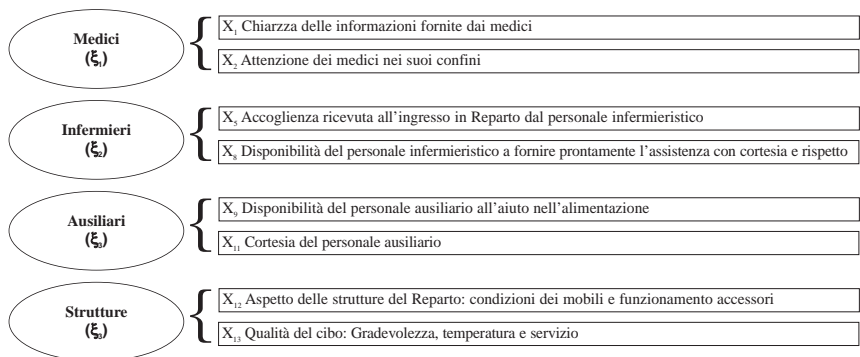


Figura 1
classificazione degli item del MQQP in relazione al costrutto teorico di appartenenza

¹ Centro di Soddifazione dell'Utenza Policlinico Universitario "Le Scotte" Via Le Scotte - 53100 Siena (ITALY) 0039 577 586439

² Università degli Studi di Chieti

interne all'azienda. Il test misura la qualità percepita del servizio sanitario su un gruppo di costrutti teorici riferiti non tanto al sistema di erogazione del servizio (accoglienza, degenza, dimissioni), ma piuttosto sui supporti organizzativi che lo producono all'interno del reparto (medici, infermieri, ausiliari, strutture), proprio a causa delle esigenze gestionali sopra descritte. Lo strumento è strutturato su 8 item (ai quali si risponde su scala likert), rappresentanti i quattro costrutti teorici sopra indicati; la figura 1 illustra la classificazione degli item rispetto alle aree d'osservazione.

Per la costruzione del test, gli Autori hanno fatto ricorso alla tecnica delle equazioni strutturali (Jöreskog e Sörbom, 1979) che permettono una migliore gestione della complessità dei fenomeni in osservazione. I modelli di equazioni strutturali, infatti, consentono di testare un'ipotesi fattoriale precisa imponendo la presenza/assenza di un item in un raggruppamento e, conseguentemente, di verificare la bontà del modello stesso ricorrendo all'analisi degli indici di adattamento.

Per la valutazione dell'adattamento complessivo hanno considerato quindi i seguenti indici di fit: c^2 (e rispettivi gradi di libertà, df), *Goodness of Fit Index (GFI)* (Jöreskog & Sörbom, 1984), *Adjusted Goodness of Fit Index (AGFI)* (Jöreskog & Sörbom, 1984; Tanaka & Huba, 1985), *Comparative Fit Index (CFI)* (Bentler, 1989, 1990), *Nonnormed Fit Index (NNFI)* (Tucker & Lewis, 1973; Bentler & Bonnett, 1980), *Root-Mean-Square Error of Approximation (RMSEA)* (Steiger & Lind, 1980, Steiger, 1989). *GFI*, *AGFI*, *CFI* e *NNFI* sono indici che variano tra 0 e 1, ed esprimono un adattamento tanto migliore quanto più si avvicinano ad 1. *RMSEA* invece indica un buon adattamento quando è inferiore a 0,06 (Hu & Bentler, 1999).

Inoltre, per dimostrare il reale miglioramento del passaggio da fattori ortogonali a fattori obliqui, hanno stimato la significatività della differenza dei χ^2 (Satorra & Bentler, 1999) e dei rispettivi gradi di libertà dei modelli. Abbiamo posto le seguenti ipotesi:

H_0 = la differenza fra i modelli è casuale ($\chi^2_a - \chi^2_b = \emptyset$)

H_1 = la differenza fra i modelli è di natura non casuale ($\chi^2_a \geq \chi^2_b$).

Per la riduzione degli item, in prima battuta, le quattro variabili latenti esogene – fattori (Medici (x_1), Infermieri (x_2), Ausiliari (x_3), Strutture

(x_4) sono state analizzate separatamente, in quattro differenti modelli. Sono stati calcolati, per ciascun item, i coefficienti strutturali fra le X e le x (?) e degli errori stocastici delle X (?). Per l'eliminazione degli item sono stati utilizzati due criteri: un criterio statistico ed uno più strettamente clinico. Seguendo il primo criterio sono stati eliminati gli item con valori di saturazione più bassi, seguendo il secondo sono stati preservati gli item che avevano una semantica più coerente con la teoria del fattore considerato.

Per la scelta della metodologia di risposta, gli Autori del test di sono rifatti alla seguente letteratura.

In accordo con Bond & Fox (2001), una considerevole quantità di ricerche nella letteratura psicometrica esamina la questione su come determinare il numero di categorie nelle scale di risposte agli item. Generalmente, il criterio per stabilire un ottimale numero di risposte incide sull'affidabilità della tecnica delle risposte stesse. I risultati mostrano plurime conclusioni, con diverse attribuzioni di attendibilità alle scale: ciò è indipendente dal numero delle categorie di risposta utilizzate (Bendig, 1953; Brown, Widing, & Coulter, 1991; Komorita, 1963; Cicchetti, Showalter & Tyrer, 1985), oppure che è massimizzato con una scala a 7 punti di valutazione (Finn, 1972; Nunnally, 1967; Ramsey, 1973; Symonds, 1924), una a 7 +/- 2 (Miller, 1956), una scala a 5 scelte (Lisitz & Green, 1975), una a 4 alternative oppure, infine una a 3 (Bendig, 1954). Linacre (1995), Wright & Linacre (1992a, 1992b), Lopez (1996) sostengono che un preciso aumento delle alternative di risposta incrementa la fedeltà del test. Tale aumento potrebbe introdurre una sorta di errore dovuto alle differenze linguistiche (Chang, 1997). Certo è che non esiste un numero ottimale di categorie di risposta che possa essere applicato univocamente alle scale di un test (Lopez, 1996).

Obiettivi del lavoro

Nel presente lavoro abbiamo ripreso quindi il modello così come proposto dagli Autori (Coluccia A., Ferretti F., Cioffi R., Lorini F, in corso di stampa 2), ricorrendo poi all'analisi multi-gruppo, verificando un'eventuale differenza dei parametri strutturali rispetto alla variabile genere.

La tecnica di “*Multi-sample analysis*” si propone di confrontare tra loro dei modelli di equazioni strutturali applicati a diversi campioni che possono differenziarsi per determinate caratteristiche (ad esempio sesso, età, differente trattamento sperimentale ecc.). L’idea di base è che la stima simultanea di una serie di parametri consenta di testare una serie di ipotesi sull’omogeneità dei gruppi. Considerato un insieme di G gruppi, si assume che il modello LISREL, espresso dalle tre equazioni

$$\eta = B\eta + \Gamma\xi + \zeta \quad (1)$$

$$x = \Lambda_x\xi + \delta \quad (2)$$

$$y = \Lambda_y\eta + \delta \quad (3)$$

sia lo stesso per ciascun gruppo considerato. Per ciascuno dei G gruppi saranno definite le otto matrici di parametri: $L_x^{(g)}$, $L_y^{(g)}$, $B^{(g)}$, $\Gamma^{(g)}$, $\Phi^{(g)}$, $\Psi^{(g)}$, $\Theta_\varepsilon^{(g)}$, $\Theta_\delta^{(g)}$ in cui g indica il generico gruppo, con $g = 1, 2 \dots, G$. Le ipotesi imposte al modello riguarderanno i vari parametri delle matrici sottoposti a determinati vincoli. A partire dall’ipotesi più generale, l’invarianza delle strutture di covarianza tra le variabili osservate, si passa in successione alla definizione di vincoli sempre più restrittivi circa l’uguaglianza delle matrici di parametri. Le ipotesi dipendono dagli obiettivi della ricerca e dal tipo di modello utilizzato ma, in generale seguono un percorso definito che può riassumersi in tre punti: 1) confronto tra le strutture fattoriali, in particolare tra i valori delle matrici Λ ; 2) confronto tra le varianze e covarianze dei fattori, cioè tra le matrici Φ ; 3) confronto tra i coefficienti di regressione (matrici B e Γ). Altri confronti, forse meno interessanti, possono comprendere le varianze d’errore (Θ_ε , Θ_δ , Ψ).

Campione

Il campione utilizzato è composto da 1600 Pazienti dell’Ospedale Le Scotte di Siena afferenti da molteplici reparti:

- 800 Maschi (M): età ? = 51,1 con ? = 21,3;
- 800 Femmine (M) età: ? = 48,1 con ? = 20,9.

Tu. Per somministrare il test sono stati contattati i Dirigenti di Reparto facendo esplicita richiesta di somministrazione del test ai loro pazienti. La sua compilazione è stata in forma anonima. Al fine di condurre le analisi dei dati sono stati ricavati due sotto-raggruppamenti (estrazione casuale con pareggiamento per sesso ed età) di uguale numerosità:

- C_1 = gruppo sperimentale (800 soggetti: 400 donne, 400 uomini),
- C_2 = gruppo di controllo (800 soggetti: 400 donne, 400 uomini).

Risultati

Sono state condotte diverse analisi fattoriali, come proposto nel seguente schema:

Gruppo Sperimentale (C_1): Test con Fattori Obliqui (Maschi + Femmine),
 Gruppo Sperimentale (C_1): Test con Fattori Ortogonali (Maschi + Femmine),
 Gruppo di Controllo (C_2): Test con Fattori Obliqui (Maschi + Femmine),
 Gruppo di Controllo (C_2): Test con Fattori Ortogonali (Maschi + Femmine),
 Gruppo Sperimentale (C_1): Test con Fattori Obliqui (Maschi),
 Gruppo Sperimentale (C_1): Test con Fattori Ortogonali (Maschi),
 Gruppo di Controllo (C_2): Test con Fattori Obliqui (Maschi),
 Gruppo di Controllo (C_2): Test con Fattori Ortogonali (Maschi),
 Gruppo Sperimentale (C_1): Test con Fattori Obliqui (Femmine),
 Gruppo Sperimentale (C_1): Test con Fattori Ortogonali (Femmine),
 Gruppo di Controllo (C_2): Test con Fattori Obliqui (Femmine),
 Gruppo di Controllo (C_2): Test con Fattori Ortogonali (Femmine),

In quest'articolo, al fine di confrontare i modelli ottenuti, riporteremo unicamente gli indici di fit delle soluzioni 3, 4, 7, 8, 11, 12.

Tabella 1

Struttura Testata	Chi²	gl	p	GFI	AGFI	RMSEA	RMSEA 90%	CFI	NNFI	
Modello 3	37,62	14	<0,001	0,99	0,98	0,032	0,020	0,065	0,99	0,99
Modello 4	2704,86	20	<0,001	0,67	40	0,31	0,300	0,320	0,58	0,42
Modello 7	27,15	14	<0,001	0,99	0,98	0,032	0,013	0,052	0,99	0,99
Modello 8	1341,77	20	<0,001	0,67	0,40	0,31	0,300	0,330	0,58	0,42
Modello 11	27,70	14	<0,001	0,99	0,98	0,035	0,015	0,054	0,99	0,99
Modello 12	1386,59	20	<0,001	0,67	0,40	0,31	0,300	0,330	0,58	0,42

Sull'intero campione, per quanto concerne la stima degli indici di fit, (tabella 1), si sono ottenuti i seguenti risultati:

modello 3: $\chi^2 = 37,62$; $gl = 14$; $p < 0,001$; $GFI = 0,990$; $AGFI = 0,980$; $RMSEA = 0,032$ (*int.conf.* 90% = 0,020 – 0,065); $CFI = 0,990$; $NNFI = 0,990$;

modello 4: $\chi^2 = 2704,86$; $gl = 20$; $p < 0,001$; $GFI = 0,670$; $AGFI = 0,400$; $RMSEA = 0,310$ (*int.conf.* 90% = 0,300 – 0,320); $CFI = 0,580$; $NNFI = 0,420$.

Dopo aver confrontato i risultati, nella tabella 2 sono stati riportati i valori delle saturazioni (λ^x) fra le variabili latenti esogene ξ_i ($\xi_1 =$ Medici, $\xi_2 =$ Infermieri, $\xi_3 =$ Ausiliari, $\xi_4 =$ Struttura) e le variabili osservate esogene X (gli item) per il modello 3. I λ^x sono superiori al 0,35 (Jöreskog & Sörbom, 1984) ed oscillano da un minimo di 0.68 ($\lambda^x_{13,4}$) ad un massimo di 0.86 ($\lambda^x_{2,1}$).

Tabella 2

Item	ξ_1	ξ_2	ξ_3	ξ_4
X_1	0.84 ($\lambda^x_{1,1}$)			
X_2	0.86 ($\lambda^x_{2,1}$)			
X_5		0.81 ($\lambda^x_{5,2}$)		
X_8		0.85 ($\lambda^x_{8,2}$)		
X_9			0.82 ($\lambda^x_{9,3}$)	
X_{11}			0.85 ($\lambda^x_{11,3}$)	
X_{12}				0.70 ($\lambda^x_{12,4}$)
X_{13}				0.68 ($\lambda^x_{13,4}$)

Per quanto concerne il test della differenza di χ^2 , in tabella 3, inoltre, si nota che: $\chi^2_{diff} = 2667,24$ ($df = 6$, $p < 0,01$). Questi dati propendono per accettare l'ipotesi alternativa (H_1) a scapito dell'ipotesi nulla (H_0).

Tabella 3

χ^2	g.d.l.	p
2704,86 - 37,62 = 2667,24	20 - 14 = 6	0,001

Facendo riferimento a Bentler & Bonett (1980) e Jöreskog & Sörbom (1979, 1993, 1996), ed osservando i dati riportati nelle tabelle 1, 2 e 3, riteniamo il modello a 4 fattori correlati il migliore. In tabella 4, infine, sono riportate le correlazioni fra i fattori. Si notano le seguenti correlazioni:

Medici (ξ_1) – Infermieri (ξ_2) = 0,70 (ϕ_{12}),
 Medici (ξ_1) – Ausiliari (ξ_3) = 0,66 (ϕ_{13});
 Medici (ξ_1) – Struttura (ξ_4) = 0,72 (ϕ_{14});
 Infermieri (ξ_2) – Ausiliari (ξ_3) = 0,87 (ϕ_{23});
 Infermieri (ξ_2) – Struttura (ξ_4) = 0,69 (ϕ_{24});
 Ausiliari (ξ_3) – Struttura (ξ_4) = 0,72 (ϕ_{34}).

Tabella 4

	ξ_1	ξ_2	ξ_3	ξ_4
ξ_1	1.00 (ϕ_{11})			
ξ_2	0.70 (ϕ_{12})	1.00 (ϕ_{22})		
ξ_3	0.66 (ϕ_{13})	0.87 (ϕ_{23})	1.00 (ϕ_{33})	
ξ_4	0.72 (ϕ_{14})	0.69 (ϕ_{24})	0.72 (ϕ_{34})	1.00 (ϕ_{44})

Al fine di paragonare i modelli, in funzione della variabile sesso siamo ricorsi ad un'analisi multi – gruppo (*Multi-sample analysis*). Il campione totale, quindi, è stato scisso in due sotto raggruppamenti (in base alla variabile sesso) composto da 800 soggetti ciascuno. Di seguito riportiamo le analisi per ogni campione.

Sul campione dei Soggetti Maschi, per quanto concerne la stima degli indici di fit, quindi (si faccia ancora riferimento alla tabella 1), si sono ottenuti i seguenti risultati:

modello 3: $\chi^2 = 27,15$; $gl = 14$; $p < 0,001$; $GFI = 0,990$; $AGFI = 0,980$; $RMSEA = 0,032$ (*int.conf. 90%* = 0,013 – 0,052); $CFI = 0,990$; $NNFI = 0,990$;

modello 4: $\chi^2 = 1341,77$; $gl = 20$; $p < 0,001$; $GFI = 0,670$; $AGFI = 0,400$; $RMSEA = 0,310$ (*int.conf. 90%* = 0,300 – 0,330); $CFI = 0,580$; $NNFI = 0,420$.

Dopo aver confrontato gli indici di fit, nella tabella 5 sono stati riportati i valori delle saturazioni (I^x) fra le variabili latenti esogene x e le variabili osservate esogene X per il modello 7. Ancora una volta I^x sono superiori al 0,35 (Jöreskog & Sörbom, 1984) ed oscillano da un minimo di 0.62 ($I^x_{13,4}$) ad un massimo di 0.87 ($I^x_{2,1}$).

Tabella 5

Item	ξ_1	ξ_2	ξ_3	ξ_4
X_1	0.85 ($\lambda^x_{1,1}$)			
X_2	0.87 ($\lambda^x_{2,1}$)			
X_5		0.82 ($\lambda^x_{5,2}$)		
X_8		0.85 ($\lambda^x_{8,2}$)		
X_9			0.86 ($\lambda^x_{9,3}$)	
X_{11}			0.86 ($\lambda^x_{11,3}$)	
X_{12}				0.64 ($\lambda^x_{12,4}$)
X_{13}				0.62 ($\lambda^x_{13,4}$)

Per quanto concerne il test della differenza di χ^2 , in tabella 6 si nota che: $\chi^2_{diff} = 1314,62$ ($df = 6, p < 0,01$). Questi dati, ancora una volta, propendono per accettare l'ipotesi alternativa (H_1) a scapito dell'ipotesi nulla (H_0).

Tabella 6

χ^2	g.d.l.	p
$1341,77 - 27,15 =$ 1314,62	$20 - 14 =$ 6	0,001

Ancora una volta, Facendo riferimento a Bentler & Bonett (1980) e Jöreskog & Sörbom (1979, 1993, 1996), ed osservati dati riportati nelle tabelle 1, 5 e 6, riteniamo il modello a 4 fattori obliqui il migliore. In tabella 7, infine, sono state proposte le correlazioni fra i fattori. Si notano le seguenti correlazioni:

$$\begin{aligned} \text{Medici } (\xi_1) - \text{Infermieri } (\xi_2) &= 0,69 (\phi_{12}), \\ \text{Medici } (\xi_1) - \text{Ausiliari } (\xi_3) &= 0,64 (\phi_{13}); \end{aligned}$$

Medici (ξ_1) – Struttura (ξ_4) = 0,60 (ϕ_{14});
 Infermieri (ξ_2) – Ausiliari (ξ_3) = 0,89 (ϕ_{23});
 Infermieri (ξ_2) – Struttura (ξ_4) = 0,77 (ϕ_{24});
 Ausiliari (ξ_3) – Struttura (ξ_4) = 0,82 (ϕ_{34}).

Tabella 7

	ξ_1	ξ_2	ξ_3	ξ_4
ξ_1	1.00 (ϕ_{11})			
ξ_2	0.69 (ϕ_{12})	1.00 (ϕ_{22})		
ξ_3	0.64 (ϕ_{13})	0.89 (ϕ_{23})	1.00 (ϕ_{33})	
ξ_4	0.60 (ϕ_{14})	0.77 (ϕ_{24})	0.82 (ϕ_{34})	1.00 (ϕ_{44})

Anche sul campione di Soggetti Femmine, sono stati calcolati gli indici di fit (si faccia ancora riferimento alla tabella 1), ottenendo i seguenti risultati:

modello 11: $\chi^2 = 27,70$; $gl = 14$; $p < 0,001$; $GFI = 0,990$; $AGFI = 0,980$; $RMSEA = 0,035$ (*int.conf.* 90% = 0,015 – 0,054); $CFI = 0,990$; $NNFI = 0,990$;

modello 12: $\chi^2 = 1386,59$; $gl = 20$; $p < 0,001$; $GFI = 0,670$; $AGFI = 0,400$; $RMSEA = 0,310$ (*int.conf.* 90% = 0,300 – 0,330); $CFI = 0,580$; $NNFI = 0,420$.

Dopo aver confrontato gli indici di fit, nella tabella 8 sono stati riportati i valori delle saturazioni (λ^x) fra le variabili latenti esogene ξ e le variabili osservate esogene X per il modello 11. I λ^x sono superiori al 0,35 (Jöreskog & Sörbom, 1984) ed oscillano da un minimo di 0.65 ($\lambda^x_{13,4}$) ad un massimo di 0.86 ($\lambda^x_{2,1}$, $\lambda^x_{11,3}$).

Tabella 8

Item	ξ_1	ξ_2	ξ_3	ξ_4
X_1	0.84 ($\lambda^x_{1,1}$)			
X_2	0.86 ($\lambda^x_{2,1}$)			
X_5		0.82 ($\lambda^x_{5,2}$)		
X_8		0.85 ($\lambda^x_{8,2}$)		
X_9			0.84 ($\lambda^x_{9,3}$)	
X_{11}			0.86 ($\lambda^x_{11,3}$)	
X_{12}				0.67 ($\lambda^x_{12,4}$)
X_{13}				0.65 ($\lambda^x_{13,4}$)

Per quanto concerne il test della differenza di χ^2 , in tabella 6 si nota che: $\chi^2_{\text{diff}} = 1358,89$ ($df = 6, p < 0,01$). Questi dati propendono per accettare l'ipotesi alternativa (H_1) a scapito dell'ipotesi nulla (H_0).

Tabella 9

χ^2	g.d.l.	p
$1386,59 - 27,70 =$ 1358,89	$20 - 14 =$ 6	0,001

Ancora una volta, facendo riferimento a Bentler & Bonett (1980) e Jöreskog & Sörbom (1979, 1993, 1996), ed osservando dati riportati nelle tabelle 1, 8 e 9, riteniamo il modello a 4 fattori obliqui il migliore. In tabella 10, infine, sono riportate le correlazioni fra i fattori. Si notano le seguenti correlazioni:

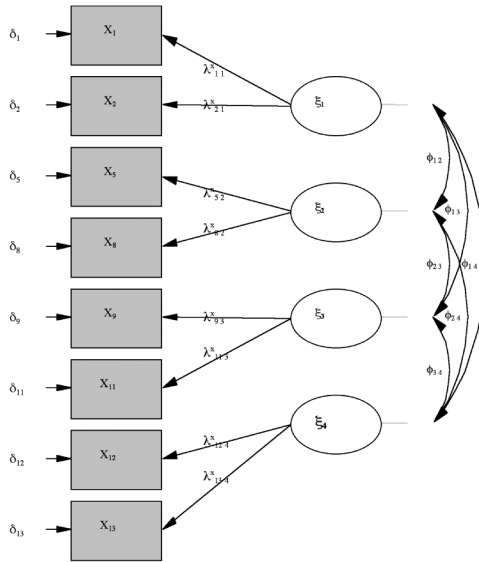
- Medici (ξ_1) – Infermieri (ξ_2) = 0,69 (ϕ_{12});
- Medici (ξ_1) – Ausiliari (ξ_3) = 0,65 (ϕ_{13});
- Medici (ξ_1) – Struttura (ξ_4) = 0,66 (ϕ_{14});
- Infermieri (ξ_2) – Ausiliari (ξ_3) = 0,88 (ϕ_{23});
- Infermieri (ξ_2) – Struttura (ξ_4) = 0,73 (ϕ_{24});
- Ausiliari (ξ_3) – Struttura (ξ_4) = 0,77 (ϕ_{34}).

Tabella 10

	ξ_1	ξ_2	ξ_3	ξ_4
ξ_1	1.00 (ϕ_{11})			
ξ_2	0.69 (ϕ_{12})	1.00 (ϕ_{22})		
ξ_3	0.65 (ϕ_{13})	0.88 (ϕ_{23})	1.00 (ϕ_{33})	
ξ_4	0.66 (ϕ_{14})	0.73 (ϕ_{24})	0.77 (ϕ_{34})	1.00 (ϕ_{44})

Il modello di riferimento è riportato in figura 2. Si tratta, come si può vedere di un modello a quattro fattori obliqui.

Figura 2



Dopo aver riportato le analisi per singolo raggruppamento (campioni maschile e campione femminile), passiamo al confronto degli indici. Il confronto tra gruppi deve seguire una serie di ipotesi sempre più restrittive (Jöreskog, 1971) a partire dal confronto tra le matrici di covarianze osservate. Ciascun'ipotesi più restrittiva ha senso solo se la precedente risulta non significativa. Unica eccezione il confronto tra le matrici di covarianze, che dovrebbe risultare significativo, altrimenti vuol dire che i gruppi non si differenziano tra loro e pertanto il confronto fra gruppi non ha senso (Jöreskog, 1971; Byrne, 1998). In tabella 11 sono riportati gli esiti delle due ipotesi testate.

Tabella 11

			Chi^2	df	p	CFI	NNFI	RMSE A	Confronto Fra Matrici
$H_0 : ?_1 ? ?_2 ? ?_3 ? ?_4$			44,599	36	0,029	0,999	0,998	0,017	
			Chi^2	GFI	SRM R				
maschi	400		21,442	0,994	0,041				
femmine	400		23,156	0,993	0,041				
			Chi^2	df	p	CFI	NNFI	RMSE A	
$H_0 : k_1 ? k_2 ? k_3 ? k_4$			43,85	28	0,029	0,998	0,995	0,026	Confronto Fra Strutture
			Chi^2	GFI	SRM R				
maschi	400		16,149	0,995	0,010				
femmine	400		27,703	0,991	0,011				

L'ipotesi di invarianza tra le matrici di covarianza (I) ha dato come risultato una differenza non significativa tra i due gruppi, e questo ci fa propendere per l'esistenza di due strutture differenti del modello base nei due gruppi. Tale ipotesi viene in qualche modo confermata nel secondo test che abbiamo condotto (ipotesi II), quello che mette a confronto le strutture del modello nei due gruppi. Dalle analisi ora condotte, si evince quindi che il modello non subisce cambiamenti secondo il sesso del soggetto. Tale risultato è stato approfondito con un'analisi della varianza (ANOVA, tabella 12) indicando come livello di significatività $p < 0,05$.

Tabella 12

ξ	Media	Dev.St.	Lim. Inf.	Lim. Sup.	p
ξ_1 (maschi)	4,2119	,75200	4,1597	4,2641	0,293
ξ_1 (femmine)	4,1719	,76828	4,1186	4,2252	
ξ_1 (totale)	4,1919	,76021	4,1546	4,2292	
ξ_2 (maschi)	4,1694	,73258	4,1185	4,2202	0,096
ξ_2 (femmine)	4,1056	,79723	4,0503	4,1610	
ξ_2 (totale)	4,1375	,76601	4,0999	4,1751	
ξ_3 (maschi)	4,0144	,73665	3,9633	4,0655	0,768
ξ_3 (femmine)	4,0031	,78450	3,9487	4,0576	
ξ_3 (totale)	4,0088	,76074	3,9714	4,0461	
ξ_4 (maschi)	3,8288	,78114	3,7745	3,8830	0,529
ξ_4 (femmine)	3,8044	,76748	3,7511	3,8576	
ξ_4 (totale)	3,8166	,77420	3,7786	3,8545	

Conclusioni

Le proprietà psicometriche di uno strumento studiato per misurare la qualità percepita dei servizi sanitari da parte degli utenti non forniscono soltanto la garanzia di ottenere misure attendibili di ciò che viene osservato, ma rappresentano la base per un utilizzo generalizzato e standardizzato anche all'interno di strutture sanitarie che offrono servizi clinici anche molto eterogenei fra di loro.

Un questionario come il MiniQQP si colloca idealmente in quella categoria di strumenti particolarmente idonei ad essere integrati nel sistema di gestione e controllo, assumendo il ruolo di strumento di monitoraggio a supporto delle decisioni strategiche delle aziende in termini di politiche sulla qualità.

Le analisi presentate mostrano come le buone proprietà del MiniQQP, fatte intuire dalle analisi confermate, siano state confermate dal presente lavoro. La costruzione teorica della scala basata su quattro fattori, non sembra differire se analizzata considerando la dimensione sesso.

Queste conclusioni lasciano ipotizzare che il concetto di soddisfazione

ne del paziente genericamente misurato dallo strumento, sia scomponibile in quattro aree concettuali riferite ad ambiti organizzativi e gestionali del servizio ben identificate. Tale risultato è in linea con le ipotesi teoriche del QQP (Coluccia, Ferretti, Lorini, Calamai, 2002).

Quest'ulteriore caratteristica porterebbe a considerare che: ad un'indispensabile informazione sulla qualità percepita da parte dei degenti in relazione al servizio ricevuto, si è in grado di associare una serie di indicatori gestionali specifici (medici, infermieri, ausiliari e strutture) ai fini del miglioramento della qualità.

Riassunto

Nel presente lavoro abbiamo analizzato le proprietà psicometriche di un questionario italiano sulla misura della qualità percepita in ambito sanitario (Mini Questionnaire of Perceived Quality; Coluccia, Ferretti, Cioffi, Lorini, *in itinere*). I soggetti devono rispondere a 14 domande suddivise in 4 fattori (Soddisfazione riguardo ai Medici, Infermieri, Ausiliari, Strutture dell'Ospedale). Abbiamo somministrato il test a 1600 pazienti dell'Ospedale "Le Scotte" di Siena, Facendo riferimento metodo delle equazioni strutturali, abbiamo studiato la dimensionalità del questionario usando un'analisi fattoriale confermativa e, successivamente, abbiamo studiato le dimensioni nella variabile genere (maschi / femmine) usando un'Analisi Multi-Gruppo. I risultati non mostrano alcuna significativa differenza del genere rispetto alle dimensioni del test.

Summary

Mini Questionnaire of Perceived Quality (QQP): a study on the differences of gender.

In the present paper we analyze the psychometric properties of an Italian questionnaire measuring perceived quality of sanitary service (Mini Questionnaire of Perceived Quality; Coluccia, Ferretti, Cioffi, Lorini, *in itinere*). Subjects answer 14 questions subdivided in four dimensions (i.e. Satisfaction towards Medical Doctors, Nurses,

Auxiliary Staff, and Hospital Structure). We administered the questionnaire to 1600 patients in the Hospital of Siena “Le Scotte”. According to structural equations modeling, we studied the dimensionality of the questionnaire using confirmatory factor analysis and, later on, we studied differences in gender using Multi-sample analysis. Results does not show significant differences in gender for test dimensions.

Bibliografia

- Bendig A.W. (1954). Reliability and the Number of Rating Categories, *Journal of Applied Psychology*, 38, 38-40.
- Bendig, A. W. (1953). The Reliability of Self-Ratings as a Function of the Amount of Verbal Anchoring and the Number of Categories on the Scale. *The Journal of Applied Psychology*, 1, 37.
- Bentler P M & Bonett Douglas G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin*, 88, 588-606.
- Bentler, P.M. (1989). *EQS structural equations program manual*. Los Angeles: BMDP Statistical Software.
- Bentler, P.M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, 107, 238-246.
- Bond T. & Fox C. (2001). *Applying the Rasch Model: Fundamental Measurement in the Human Sciences*. Lawrence Erlbaum Associates, Mahwah NJ.
- Byrne, B.M. (1998). *Structural Equation Modeling with LISREL, PRELIS and SIMPLIS: Basic Concepts, Applications, and Programming*. Lawrence Erlbaum Associates, Inc., Mahwah, N.J.
- Chang, L. (1997). Dependability of anchoring labels of Likert-type scales. *Educational and Psychological Measurement*, 57, 800-7.
- Cicchetti, Domenic V., Donald Showalter & Peter J. Tyrer (1985). The Effect of Number of Rating Scale Categories on Levels of Interrater Reliability: A Monte Carlo Investigation. *Applied Psychological Measurement*, 9, 31-36.
- Coluccia A., Ferretti F. (2001). *Sull'importanza delle indagini di customer satisfaction anche come risorsa economica*. in N. Di Paolo, U. Buoncristiani (a cura di), *Tecniche Nefrologiche e dialitiche 2001*, Editoriale Bios.
- Coluccia A., Ferretti F., Cioffi R., Lorini F. (in corso di stampa 2). Ottimizzazione e Riduzione del test Questionario di Qualità Percepita (QQP); *Minerva Medica*.
- Coluccia A., Ferretti F., Loré C. (2000). Qualità dei servizi sanitari e Customer Satisfaction: primi risultati di uno studio pilota. *Difesa Sociale*, 5/6, 7-16.
- Coluccia A., Ferretti F., Lorini F., Calamai M. (2002). Il questionario di qualità percepita (QQP). *Mondo Sanitario*, 10, 14-19.
- Coluccia A., Cioffi R., Ferretti F., Lorini F., Vidotto G., (in corso di giudizio).

- Questionario di Qualità Percepita (QQP): uno studio sulle differenze di genere. *Giornale Italiano di Medicina del Lavoro ed Ergonomia*.
- Coluccia A., Ferretti F., Cioffi R., Lorenzi L., Lorini F., Vidotto G. (in corso di stampa 1). Qualità Percepita e Soddisfazione dell'Utenza: uno studio sul Q.Q.P. *Rivista Italiana di Medicina Legale*.
- Finn, R. H. (1972). Effects of Some Variations in Rating Scale Characteristics on the Means and Reliabilities of Ratings. *Educational and Psychological Measurement*, 32, 255-265.
- Hu, L. & Bentler, P.M. (1999). Cutoff Criteria for Fit Indexes in Covariance Structure Analysis: Conventional Criteria Versus New Alternatives. *Structural Equation Modeling*, 6, 1-55.
- Jöreskog K. G., Sörbom D. (1979). *Advances in Factor Analysis and Structural Equation Models*. Cambridge, Massachusetts, Abstract Books.
- Jöreskog, K. G & Sörbom, D (1993). *Structural Equation Modeling with the SIMPLIS Command Language*. Chicago: Scientific Software International.
- Jöreskog, K. G. & Sörbom, D. (1996). *LISREL 8 User's Reference Guide*. Chicago: Scientific Software International.
- Jöreskog, K.G. (1971). Simultaneous factor analysis in several populations. *Psychometrika*, 36, 409-426.
- Komorita, S.S. (1963). Attitude content, intensity and the neutral point on a Likert scale. *Journal of Social Psychology*, 61, 327-334.
- Linacre J. M. (1995). Categorical misfit statistics. *Rasch Measurement Transactions*, 9, 450-451.
- Lissitz, Robert W. & Samuel B. Green (1975). Effect of the Number of Scale Points on Reliability: A Monte Carlo Approach. *Journal of Applied Psychology*, 60, 10-13.
- Lopez, W. (1996). Communication validity and rating scales. *Rasch Measurement Transactions*, 10, 482-483.
- Miller, George A. (1956). The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information. *The Psychological Review*, 63, 81-97.
- Nunnally J.C. (1978). *Psychometric Theory 2° ed.*. McGraw Hill, New York.
- Ramsey, J. O. (1973). The Effect of Number of Categories in Rating Scales on Precision of Estimation of Scale Values *Psychometrika*, 38, 4.
- Satorra A. & Bentler P.M. (1999). A Scaled Difference Chi-square Test Statistic for Moment Structure Analysis. *Economics Working Papers*, 412, 1-14.
- Schmitt N. (1996). Uses and Abuses of Coefficient Alpha. *Psychological Assessment*, 8, 350-353.
- Steiger, J.H. & Lind, J.C. (1980). *Statistically based tests for the number of*

- common factors*. Paper presented at the annual meeting of the Psychometric Society, Iowa City, IA.
- Steiger, J.H. (1989). *EzPATH: A supplementary module for SYSTAT and SYGRAPH*. Evanston, IL: SYSTAT.
- Symonds, P. M. (1924). On the Loss of Reliability in Ratings Due to Coarseness of the Scale. *Journal of Experimental Psychology*, 7, 456-461.
- Tanaka, J.S. & Huba, G.J. (1985). A fit index for covariance structural models under arbitrary GLS estimation. *British Journal of Mathematical and Statistical Psychology*, 42, 233-239.
- Tucker L.R. & Lewis C. (1973). A Reliability Coefficient for Maximum Likelihood Factor Analysis. *Psychometrika*, 38, 1 -10.
- Wright, B.D., & Linacre, J.M. (1992a). Disordered steps? *Rasch Measurement Transactions*, 6, 22-233.
- Wright, B.D., & Linacre, J.M. (1992b). Combining and splitting categories. *Rasch Measurement*, 6, 233-235.