

ANALISI FATTORIALE

In questo documento presentiamo alcune opzioni analitiche della procedura di analisi fattoriale di SPSS che non sono state incluse nel testo pubblicato. Si tratta di opzioni che, pur non essendo utilizzate correntemente per effettuare l'analisi fattoriale, consentono di completare lo spettro delle possibilità di analisi offerte da SPSS. In particolare, il documento completa la descrizione della procedura di analisi fattoriale con in seguenti paragrafi:

1. *Scale validity coefficient (coefficiente di validità fattoriale)*
2. Il metodo di Everett
3. I coefficienti di congruenza fattoriale di Tucker
4. Calcolo degli indici alternativi di bontà dell'adattamento
5. L'Analisi Fattoriale con il linguaggio SINTASSI
6. Opzioni di analisi ottenibili solo attraverso SINTASSI
 - Stime diverse per le comunalità
 - Rappresentazione grafica delle variabili nello spazio fattoriale
 - Salvataggio di punteggi fattoriali con nomi specificati dall'utente
 - Analisi diretta di matrici

1. Scale validity

I punteggi fattoriali stimati dalla soluzione fattoriale attraverso un metodo come la regressione (vedi pag. 77 del testo) possono essere correlati con le scale definite teoricamente prima di effettuare l'analisi fattoriale stessa. Il coefficiente di correlazione così ricavabile viene definito coefficiente di validità fattoriale (*scale validity coefficient*, Cattell & Tsujoka, 1964) e può essere utilizzato come indice di validità della scala teorica, ovvero come indice che informa il ricercatore se la scala definita teoricamente a priori trova conforto nei risultati dell'analisi fattoriale. Nell'esempio

discusso nel testo abbiamo calcolato i punteggi nella Coscienziosità e nell'Energia sommando i 3

item relativi a ciascuna delle due scale con la seguente sintassi:

***** Calcolo delle scale "teoriche" C ed E.**

```
compute C= agg20 +agg21 +agg2  
compute E= agg9 +agg11 +agg8
```

La tabella 1 presenta i risultati della correlazioni tra le due serie di punteggi, quelli teorici e quelli empirici. Abbiamo preferito considerare gli item originali per mantenere la stessa unità di misura degli altri item non ricodificati: tuttavia, correlazioni sovrapponibili a queste si ottengono considerando gli item 2 e 8 ricodificati come indicato nel testo. Un'altra alternativa possibile per effettuare la somma degli aggettivi mantenendo la stessa unità di misura era quella di standardizzare prima gli aggettivi e poi sommarli). Le correlazioni tra punteggi omologhi (riportate in grassetto) sono molto elevate, e questo conferma la bontà dell'interpretazione dei fattori data sopra: il primo fattore correla fortemente con la scala teorica di Coscienziosità, il secondo con la scala teorica di Energia. L'interpretazione di queste correlazioni risulta ancora più forte se invece di considerare le correlazioni di ordine zero si considerano le correlazioni parziali (infatti, sia i punteggi fattoriali, sia le scale teoriche sono fortemente correlate, come è evidente nella tabella 1). La correlazione parziale tra il primo fattore e la Coscienziosità, controllando il secondo fattore e l'Energia, è pari a .96. Quella tra il primo fattore e l'Energia, controllando il secondo fattore e la Coscienziosità, è invece pari a .20. La correlazione parziale tra il secondo fattore e l'Energia, controllando il primo fattore e la Coscienziosità, è pari a .99. Quella tra il secondo fattore e la Coscienziosità, controllando il primo fattore e l'Energia, è invece pari a .31.

Di seguito presentiamo le linee di SINTASSI utilizzate per il calcolo delle correlazioni parziali:

```

**** CORREL. PARZIALE TRA F1 C CONTROLLANDO F2 E ***.
PARTIAL CORR
/VARIABLES= fac1_1 C BY E fac2_1.

**** CORREL. PARZIALE TRA F1 E CONTROLLANDO F2 C ***.
PARTIAL CORR
/VARIABLES= fac1_1 E BY fac2_1 C.

**** CORREL. PARZIALE TRA F2 E CONTROLLANDO F1 C ***.
PARTIAL CORR
/VARIABLES= E fac2_1 BY C fac1_1.

**** CORREL. PARZIALE TRA F2 C CONTROLLANDO F1 E ***.
PARTIAL CORR
/VARIABLES= fac2_1 C BY fac1_1 E.

```

Tabella 1. Correlazioni tra punteggi fattoriali e scale “teoriche”

Correlazioni

		FAC1_1 REGR factor score 1 for analysis 1	FAC2_1 REGR factor score 2 for analysis 1	C	E
FAC1_1 REGR factor score 1 for analysis 1	Correlazione di Pearson Sig. (2-code) N	1 .3075	,657** ,000 3075	,981** ,000 3075	,610** ,000 3075
FAC2_1 REGR factor score 2 for analysis 1	Correlazione di Pearson Sig. (2-code) N	,657** ,000 3075	1 .3075	,619** ,000 3075	,993** ,000 3075
C	Correlazione di Pearson Sig. (2-code) N	,981** ,000 3075	,619** ,000 3075	1 .3138	,565** ,000 3137
E	Correlazione di Pearson Sig. (2-code) N	,610** ,000 3075	,993** ,000 3075	,565** ,000 3137	1 .3140

** . La correlazione è significativa al livello 0,01 (2-code).

2. Il metodo di Everett

I fattori “validi” di una soluzione fattoriale sono quelli che risultano più facilmente replicabili su campioni diversi da quelli nei quali sono stati individuati. I fattori “spuri”, invece, risultano poco generalizzabili e sono determinati sostanzialmente dall’errore campionario (a questo riguardo si veda Cattell, 1978). Sulla base di questo ragionamento Everett (1983) ha sviluppato una procedura che serve per validare le soluzioni fattoriali effettuate su campioni diversi e che utilizza i punteggi fattoriali. Tale procedura si basa sulla selezione dei fattori che vengono replicati attraverso gruppi diversi di soggetti. La procedura da seguire è schematicamente la seguente:

- a) bisogna dividere il campione in due sottocampioni A e B ricavando così due matrici di dati standardizzati, Z_A e Z_B ;
- b) quindi si deve effettuare l’analisi fattoriale separatamente su Z_A e su Z_B estraendo da m a k fattori (di solito $m=1$), e derivare i coefficienti dei punteggi fattoriali (*factor scores coefficients*) specifici per i due campioni A e B (questi coefficienti sono contenuti nelle matrici B_A e B_B , che hanno tante righe quante sono le variabili e tante colonne quanti sono i fattori estratti);
- c) si devono utilizzare questi due insiemi di coefficienti per calcolare su tutti i soggetti i punteggi nei fattori generati da A e da B (ovvero, $F_A=ZB_A$, e $F_B=ZB_B$);
- d) infine, si devono correlare i fattori corrispondenti nelle matrici F_A e F_B e tenere solo quelli che presentano una correlazione pari almeno a .9.

Il metodo di Everett, pur se introdotto originariamente per esaminare la stabilità della soluzione fattoriale, ha trovato valide applicazioni come metodo per stabilire il numero di fattori nella soluzione (ad esempio, Borkenau e Ostendorf, 1990, e McCrae et al., 1996). Anche se non è molto diffuso, la sua implementazione tramite SPSS è piuttosto semplice. Dopo aver selezionato le variabili da fattorializzare, definito le opzioni opportune nei metodi di estrazione e rotazione, e selezionato un metodo per salvare i punteggi fattoriali, nella finestra di dialogo principale per l’analisi fattoriale si inserisce nel box “*Variabile di selezione*” la variabile che serve a stratificare il campione nei diversi gruppi. Con un clic del mouse sul pulsante “Valore” si accede alla finestra “*Imposta valore*” che serve per specificare il valore che deve essere soddisfatto affinché il soggetto

sia filtrato per l'analisi. Quindi si procede con l'analisi che verrà effettuata solo sui casi che nella variabile di selezione soddisfano il valore specificato in "Imposta valore": se le saturazioni fattoriali e i coefficienti dei punteggi fattoriali sono calcolati considerando solo questi soggetti, i punteggi fattoriali sono però calcolati *anche* per i soggetti non filtrati. Si ripete l'analisi specificando questa volta un altro valore in "Imposta valore" e si avrà così un secondo set di punteggi fattoriali calcolati considerando i coefficienti solo dei soggetti che soddisfano la nuova condizione. Si procede finché tutti i valori della variabile di selezione non sono esauriti. A questo punto si correlano i punteggi fattoriali ottenuti nelle diverse analisi: questi saranno i coefficienti di invarianza di Everett.

Consideriamo l'esempio di analisi fattoriale che abbiamo discusso nel testo. Esaminiamo se la stessa struttura fattoriale emerge considerando il gruppo degli elettori di centro-sinistra (SCH = 1) e il gruppo degli elettori di centro-destra (SCH = 2). Dopo aver specificato le opzioni necessarie per ottenere la soluzione fattoriale e i punteggi fattoriali, specifichiamo la variabile di selezione (figura 1) inserendo la prima volta il valore 1 e mandando in esecuzione la procedura, e poi il valore 2 e ripetendo l'analisi con la nuova selezione. Al termine di questa doppia analisi avremo due insiemi di punteggi fattoriali, come specificato nella figura 2: i punteggi fac1_2 e fac2_2 sono stati ricavati per tutti i soggetti considerando i coefficienti dei punteggi fattoriali ricavati dall'analisi degli elettori di centro sinistra (tabella 2, sezione di sinistra), i punteggi fac1_3 e fac2_3 sono stati ricavati per tutti i soggetti considerando i coefficienti dei punteggi fattoriali ricavati dall'analisi degli elettori di centro destra (tabella 2, sezione di destra). Correlando i due insiemi di punteggi fattoriali si ottengono i valori riportati nella tabella 3 (vedi i valori in grassetto). Emerge una forte corrispondenza tra i punteggi in questi due insiemi: il primo fattore per gli elettori del centro-sinistra corrisponde al secondo fattore per gli elettori del centro-destra, e viceversa. Le matrici dei modelli (tabella 4) evidenziano questa inversione: mentre negli elettori di centro-sinistra il primo fattore corrisponde alla energia e il secondo alla coscienziosità, per gli elettori di centro destra avviene esattamente il contrario.

Figura 1. Finestre di dialogo “imposta valore”

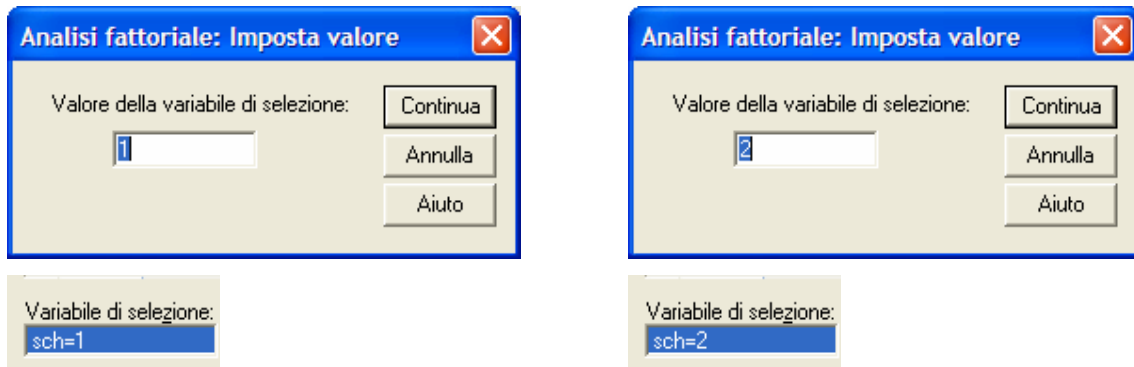


Figura 2. Estratto del data-file con le due serie di punteggi fattoriali

ana_fatt.sav - Editor dei dati SPSS

File Modifica Visualizza Dati Trasforma Analizza Grafici Strumenti Finestra ?

1 :

	fac1_2	fac2_2	fac1_3	fac2_3	var	var
91	,09552	1,94481	-,72270	,98705		
92	,25043	2,09249	-,64331	1,19280		
93	,76409	-1,07574	1,14253	,17768		
94	-,87789	2,06965	-1,65973	,18271		
95	-,90462	2,09405	-1,69419	,17035		
96	-,39503	-,02444	-,34900	-,36126		
97	,63642	-,09056	,61636	,52174		
98	-,42052	-1,97644	,44035	-1,28962		

Tabella 2. Matrici dei coefficienti dei punteggi fattoriali nei due sottocampioni

Matrice dei coefficienti di punteggio fattoriale

	Fattore	
	1	2
AGG9 Dinamico	,416	,007
AGG11 Energico	,304	,062
AGG8_B Determinato	,229	,046
AGG20 Responsabile	,027	,571
AGG21 Scrupoloso	,035	,258
AGG2_B Affidabile	,063	,118

Metodo estrazione: fattorizzazione dell'asse principale.
 Metodo rotazione: Promax con normalizzazione di Kaiser.
 Metodo punteggi fattoriali: regressione.

a. Nella fase di analisi sono utilizzati solo i casi per cui SCH = 1.

Matrice dei coefficienti di punteggio fattoriale

	Fattore	
	1	2
AGG9 Dinamico	-,010	,428
AGG11 Energico	,053	,305
AGG8_B Determinato	,033	,277
AGG20 Responsabile	,706	-,019
AGG21 Scrupoloso	,162	,082
AGG2_B Affidabile	,122	,029

Metodo estrazione: fattorizzazione dell'asse principale.
 Metodo rotazione: Promax con normalizzazione di Kaiser.
 Metodo punteggi fattoriali: regressione.

a. Nella fase di analisi sono utilizzati solo i casi per cui SCH = 2.

Tabella 3. Correlazioni tra i punteggi fattoriali

Correlazioni

		FAC1_2 REGR factor score 1 for analysis 2	FAC2_2 REGR factor score 2 for analysis 2	FAC1_3 REGR factor score 1 for analysis 3	FAC2_3 REGR factor score 2 for analysis 3
FAC1_2 REGR factor score 1 for analysis 2	Correlazione di Pearson Sig. (2-code) N	1 . 3075	,693** ,000 3075	,675** ,000 3075	,998** ,000 3075
FAC2_2 REGR factor score 2 for analysis 2	Correlazione di Pearson Sig. (2-code) N	,693** ,000 3075	1 . 3075	,998** ,000 3075	,655** ,000 3075
FAC1_3 REGR factor score 1 for analysis 3	Correlazione di Pearson Sig. (2-code) N	,675** ,000 3075	,998** ,000 3075	1 . 3075	,635** ,000 3075
FAC2_3 REGR factor score 2 for analysis 3	Correlazione di Pearson Sig. (2-code) N	,998** ,000 3075	,655** ,000 3075	,635** ,000 3075	1 . 3075

** . La correlazione è significativa al livello 0,01 (2-code).

Tabella 4. Soluzioni fattoriali (matrici pattern) nei due sottocampioni

Matrice dei modelli^{a,b}

	Fattore	
	1	2
AGG9 Dinamico	,896	,000
AGG11 Energico	,805	,086
AGG8_B Determinato	,771	,074
AGG20 Responsabile	-,017	,943
AGG21 Scrupoloso	,046	,826
AGG2_B Affidabile	,209	,595

Metodo estrazione: fattorizzazione dell'asse principale.
Metodo rotazione: Promax con normalizzazione di Kaiser.

- a. La rotazione ha raggiunto i criteri di convergenza in 3 iterazioni.
- b. Nella fase di analisi sono utilizzati solo i casi per cui SCH = 1. CENTRO SINISTRA**

Matrice dei modelli^{a,b}

	Fattore	
	1	2
AGG9 Dinamico	-,096	,873
AGG11 Energico	,130	,730
AGG8_B Determinato	,012	,761
AGG20 Responsabile	,991	-,073
AGG21 Scrupoloso	,801	,094
AGG2_B Affidabile	,799	,001

Metodo estrazione: fattorizzazione dell'asse principale.
Metodo rotazione: Promax con normalizzazione di Kaiser.

- a. La rotazione ha raggiunto i criteri di convergenza in 3 iterazioni.
- b. Nella fase di analisi sono utilizzati solo i casi per cui SCH = 2. CENTRO DESTRA**

Di seguito presentiamo le linee di SINTASSI utilizzate per il metodo di Everett.

*** analisi sugli elettori di centro-sinistra.**

```
FACTOR
/VARIABLES agg9 agg11 agg8_b agg20 agg21 agg2_b
/MISSING LISTWISE
/ANALYSIS agg9 agg11 agg8_b agg20 agg21 agg2_b
/SELECT=sch(1)
/PRINT INITIAL EXTRACTION ROTATION FSCORE
/CRITERIA FACTORS(2) ITERATE(25)
/EXTRACTION PAF
/CRITERIA ITERATE(25)
/ROTATION PROMAX(4)
/SAVE REG(ALL)
/METHOD=CORRELATION .
```

*** analisi sugli elettori di centro-destra.**

```
FACTOR
/VARIABLES agg9 agg11 agg8_b agg20 agg21 agg2_b /MISSING
LISTWISE
/ANALYSIS agg9 agg11 agg8_b agg20 agg21 agg2_b
/SELECT=sch(2)
/PRINT INITIAL EXTRACTION ROTATION FSCORE
/CRITERIA FACTORS(2) ITERATE(25)
/EXTRACTION PAF
/CRITERIA ITERATE(25)
/ROTATION PROMAX(4)
/SAVE REG(ALL)
/METHOD=CORRELATION .
```

*** correlazione tra i punteggi fattoriali.**

```
CORRELATIONS
/VARIABLES=fac1_2 fac2_2 fac1_3 fac2_3
/PRINT=TWOTAIL NOSIG
/MISSING=PAIRWISE .
```


3. I coefficienti di congruenza fattoriale di Tucker

Le soluzioni ottenute su gruppi diversi di soggetti con le stesse variabili possono essere confrontate utilizzando appositi coefficienti che vengono definiti di solito “coefficienti di congruenza”. Uno degli indici di congruenza più utilizzati è il coefficiente ϕ di Tucker (1951) che è calcolabile nel modo seguente:

$$\phi_{pq} = (\sum_j a_{1jp} a_{2jq}) / [(\sum_j a_{1jp}^2)(\sum_j a_{2jq}^2)]^{.5}$$

dove a_{1jp} a_{2jq} sono le saturazioni della variabile j relative al fattore p del campione 1, e al fattore q del campione 2. Solitamente coefficienti superiori a .90 indicano una raggiunta convergenza (ovvero i fattori nei 2 campioni possono essere considerati identici), mentre coefficienti inferiori a .90 indicano una corrispondenza problematica tra gli stessi fattori nei 2 gruppi diversi (vedi Van de Vijver e Leung, 1997).

Non è possibile calcolare direttamente i coefficienti di Tucker tramite i menù di SPSS, ma soltanto tramite il linguaggio SINTASSI, utilizzando MATRIX, una procedura sviluppata per analizzare dati in formato matriciale. Supponiamo ad esempio di voler confrontare la similarità tra le matrici pattern che emergono dalle analisi fattoriale effettuate separatamente sugli elettori di centro sinistra e su quelli di centro destra (tabella 4). La prima cosa da fare è creare un file in formato testo (**.TXT**) che contiene le due diverse matrici: in questi file però l'ordine dei fattori deve essere specificato in modo tale che il primo fattore del primo file corrisponda al primo del secondo file, ecc. Nella tabella 4 è evidente che i fattori sono emersi in ordine diverso nei due campioni: il primo fattore negli elettori di CS corrisponde al secondo per gli elettori di CD, e viceversa. La tabella 5 presenta i due file ordinati nel modo corretto. Di seguito invece presentiamo le linee di SINTASSI per calcolare i coefficienti di congruenza:

****** Coefficiente di congruenza fattoriale di Tucker.**

```
MATRIX.  
READ A/FILE='C:\CS.TXT'/SIZE={6;2}/FIELD = 1 TO 12 BY 6.  
PRINT A.  
READ B/FILE='C:\CD.TXT'/SIZE={6;2}/FIELD = 1 TO 12 BY 6.  
PRINT B.  
COMPUTE C=A*B.  
COMPUTE D=A*A.  
COMPUTE E=B*B.  
COMPUTE C1=CSUM(C).  
COMPUTE D1=CSUM(D).  
COMPUTE E1=CSUM(E).  
COMPUTE DEN=SQRT(D1*E1).  
COMPUTE R=C1/DEN.  
PRINT R.  
END MATRIX.
```

Il comando `MATRIX.` segnala l'inizio di una serie di linee di comando in cui si utilizzerà il linguaggio di programmazione `MATRIX.`

Il comando `READ A/FILE='C:\CS.TXT'` indica il percorso dove SPSS trova la prima matrice di saturazione alla quale verrà data l'etichetta `A`.

Il comando `/SIZE={6;2}/` indica che nel file abbiamo una matrice con 6 righe e 2 colonne.

Il comando `FIELD = 1 TO 12 BY 6` indica che i dati sono nelle colonne che vanno da 1 a 12, e che ogni valore (cioè ogni saturazione della pattern matrix) occupa 6 colonne. Viene chiesto di stampare tale matrice tramite il comando `PRINT A.`

I comandi `COMPUTE C=A*B.` `COMPUTE D=A*A.` `COMPUTE E=B*B.` calcolano rispettivamente le matrici **C**, **D** ed **E**, tramite la moltiplicazione matriciale elemento per elemento, ovvero, vengono moltiplicati tra loro gli elementi corrispondenti delle due matrici termini.

I comandi `COMPUTE C1=CSUM(C).` `COMPUTE D1=CSUM(D).` `COMPUTE E1=CSUM(E).` calcolano rispettivamente i vettori riga **C1**, **D1** ed **E1**, che ha lo stesso numero di colonne della matrice argomento: ogni colonna del vettore riga contiene la somma degli elementi nella corrispondente colonna della matrice argomento.

Il comando `COMPUTE DEN=SQRT (D1&*E1)` . calcola il vettore riga **DEN** che è uguale alla radice quadrata del prodotto elemento per elemento dei due vettori diga **D1** ed **E1**.

Il comando `COMPUTE R=C1&/DEN` . calcola il vettore riga **R** che è uguale al rapporto elemento per elemento dei due vettori diga **C1** ed **DEN**. Questo vettore conterrà i valori dei coefficienti di Tucker, che verranno “stampati” nell’output tramite il comando `PRINT R` .

Il comando `COMPUTE DEN=SQRT (D1&*E1)` . calcola il vettore riga **DEN** che è uguale alla

Il comando `END MATRIX` . segnala la fine della serie di linee di comando nel linguaggio di programmazione `MATRIX`.

La tabella 6 riporta, nella sezione di sinistra, il risultato della procedura per il calcolo dei coefficienti di Tucker. I valori sotto le etichette A e B sono semplicemente le due matrici pattern che sono state lette da SPSS ed elaborate per calcolare i coefficienti di congruenza. Questi sono i valori in grassetto sotto l’etichetta R: tali coefficienti risultano molto elevati anche considerando che il numero di variabili esaminate è molto basso. La sezione di destra della tabella 6 riporta alcuni valori di probabilità che possono essere utilizzati per interpretare il coefficiente di Tucker (Cattell, 1978). Si tratta di valori generati da Schneewind e Cattell (e riportati in Cattell, 1978) simulando con dati artificiali il comportamento di pattern fattoriali tipici. A seconda del numero di variabili considerate è possibile definire un intervallo di confidenza (per i seguenti diversi livelli di probabilità: .001, .01, .025, .05) al di fuori del quale vengono a trovarsi i valori significativi di ϕ . Questi valori vanno comunque interpretati con cautela, mentre riteniamo più prudente attenersi conservativamente al valore-limite indicativo di .90.

I coefficienti di congruenza ottenuti risultano assai simili a quelli ottenuti con il metodo di Everett. A questo riguardo, i coefficienti ricavabili con il metodo di Everett , o “coefficienti di *invarianza*” (vedi anche Pinneau e Newhouse, 1964), vengono solitamente interpretati facendo riferimento agli stessi valori-limite utilizzati per i coefficienti di congruenza. In particolare Ten Berge (1986a) ha dimostrato che un’elevata congruenza coincide con un’elevata invarianza; egli inoltre ha sottolineato che le procedure basate sul calcolo dei coefficienti di congruenza/invarianza “sono totalmente protette dalla capitalizzazione sul caso, poiché le analisi sono effettuate

indipendentemente. Qualsiasi evidenza venga ottenuta da questa procedura essa è confermativa”

(Ten Berge, 1986b, p. 44).

I coefficienti di Tucker e di Everett, così come altri coefficienti di congruenza fattoriale che non abbiamo riportato (e per i quali si rimanda a Van de Vijver e Leung, 1997) sono utili per avere un’idea a livello globale della similarità delle strutture fattoriali ottenute dalle stesse variabili su campioni diversi. Van de Vijver e Leung (1997) sottolineano comunque che questi coefficienti “non sono sufficientemente accurati per identificare item anomali e differenze sottili nella composizione e nel significato dei fattori attraverso i gruppi” (p. 93). A questo riguardo, metodi di analisi più adeguati che possono indagare la congruenza a livello di singoli item sono quelli basati sulle tecniche di analisi fattoriale confermativa multigruppo.

Tabella 5. Matrici pattern nel formato utilizzabile per il calcolo dei coefficienti di Tucker

*** MATRICE CENTRO SINISTRA (file "CS.TXT")		*** MATRICE CENTRO DESTRA (file "CD.TXT")	
,896	,000	,873	-,096
,805	,086	,730	,130
,771	,074	,761	,012
-,017	,943	-,073	,991
,046	,826	,094	,801
,209	,595	,001	,799

Tabella 6. Risultato dell'applicazione del programma per il calcolo dei coefficienti di Tucker

		LIVELLI DI PROBABILITA' PER INTERPRETARE I COEFFICIENTI DI TUCKER (CATTELL, 1978).		
Run MATRIX procedure:		NVAR= NUMERO DI VARIABILI		
		NVAR	p	PHI
A				
	,8960000000 ,0000000000			
	,8050000000 ,0860000000			
	,7710000000 ,0740000000	10	.001	-.87, .91
	-,0170000000 ,9430000000		.01	-.75, .78
	,0460000000 ,8260000000		.025	-.67, .70
	,2090000000 ,5950000000		.05	-.59, .63
B		20	.001	-.76, .82
	,8730000000 -,0960000000		.01	-.63, .68
	,7300000000 ,1300000000		.025	-.53, .57
	,7610000000 ,0120000000		.05	-.47, .50
	-,0730000000 ,9910000000			
	,0940000000 ,8010000000	30	.001	-.68, .73
	,0010000000 ,7990000000		.01	-.55, .58
			.025	-.48, .51
			.05	-.38, .43
R				
	,9874765270 ,9893093740			
-		40	.001	-.51, .65
			.01	-.38, .46
			.025	-.33, .39
			.05	-.27, .32
	----- END MATRIX -----			
		50	.001	-.42, .54
			.01	-.34, .39
			.025	-.28, .32
			.05	-.23, .27

4. Calcolo degli indici alternativi di bontà dell'adattamento

Nel caso della soluzione di massima verosimiglianza il test di bontà dell'adattamento può risultare significativo anche in presenza di una matrice riprodotta sostanzialmente identica alla matrice osservata (si veda il paragrafo 2.3.2 del testo). Questo dipende dalla grandezza del campione che rende troppo potente il test statistico. Sono stati sviluppati allora degli indici alternativi di bontà dell'adattamento che risultano meno influenzati dall'ampiezza campionaria. Nelle pagine seguenti presentiamo alcuni di questi indici discutendo il modo in cui essi possono essere calcolati con il foglio elettronico Excel utilizzando l'output di SPSS

4.1. Indici incrementali o comparativi

Gli indici incrementali o comparativi (come il *Nonnormed Fit Index* - NNFI, e il *Comparative Fit Index* - CFI) valutano l'adeguatezza del modello confrontando quale è il miglioramento che esso fornisce nello spiegare i dati osservati, rispetto ad un modello nullo in cui si ipotizza che non ci sia alcuna relazione tra le variabili.

Indice di Tucker e Lewis (Tucker e Lewis, 1973)

$$TLI = [(\chi^2_{\text{nullo}} / df_{\text{nullo}}) - (\chi^2_{\text{target}} / df_{\text{target}})] / [(\chi^2_{\text{nullo}} / df_{\text{nullo}}) - 1]$$

dove χ^2_{nullo} è il limite inferiore dell'adattamento realizzato dal modello in cui non viene ipotizzata alcuna struttura sui dati: in pratica, l'indice χ^2_{nullo} viene derivato dalla funzione di adattamento relativa alle stime dei parametri dove si ipotizza che non ci sia relazione tra le variabili. L'indice χ^2_{target} invece rappresenta il valore del test di bontà dell'adattamento relativo al modello ipotizzato dal ricercatore (ovvero, il modello in cui si ipotizza la presenza di k fattori). I valori df_{target} e df_{nullo} , infine, sono rispettivamente i gradi di libertà del modello target e del modello nullo. Un coefficiente vicino a 1 (e comunque superiore a .9) indica un buon adattamento del modello ai dati empirici.

Poiché la sua grandezza può risultare anche superiore a 1, l'indice TLI viene chiamato anche “indice *non* normalizzato di adattamento” (*Non-normed Fit Index*, NNFI). Per stimare il chi quadrato del modello nullo si può utilizzare il chi quadrato ottenuto dal test di sfericità di Bartlett con gradi di libertà pari a $n(n-1)/2$, dove n = numero delle variabili osservate.

La figura 3 presenta le impostazioni del foglio elettronico Excel per calcolare l'indice TLI: la formula corrispondente alla cella J4 (evidenziata nella barra della formula sotto le barre degli strumenti) consente di calcolare l'indice sulla base dei valori del chi-quadrato e dei gradi di libertà desunti dalle soluzioni ML a 1 e a 2 fattori (ovviamente la scelta del numero di fattori dipende dal data set particolare che viene analizzato) e dalla soluzione *nulla* (in questo caso chi-quadrato e gradi di libertà sono quelli ricavabili dal test di sfericità di Bartlett). Nel nostro esempio la soluzione a 1 fattore risulta fortemente inadeguata, mentre quella a 2 fattori presenta un TLI assolutamente adeguato (.99).

Figura 3. Foglio Excel per il calcolo dell'indice di Tucker e Lewis

	A	B	C	D	E	F	G	H	I	J	K
1	TLI = $[(\chi^2_{\text{nullo}} / df_{\text{nullo}}) - (\chi^2_{\text{target}} / df_{\text{target}})] / [(\chi^2_{\text{nullo}} / df_{\text{nullo}}) - 1]$										
2											
3		χ^2_{nullo}	df_{nullo}		χ^2_{target}	df_{target}		Soluzione ML		TLI	
4		11769,4296	15,0000		2883,993	9		1 fattore		0,592353	
5					35,5108	4,0000		2 fattori		0,989947	
6											

Comparative Fit Index (CFI, Bentler, 1990)

$$CFI = 1 - [\max(\chi^2_{\text{target}} - df_{\text{target}}, 0) / \max(\chi^2_{\text{nullo}} - df_{\text{nullo}}, \chi^2_{\text{target}} - df_{\text{target}}, 0)]$$

dove i termini sono uguali a quelli definiti nel caso del TLI. Rispetto all'indice TLI, il CFI stima la bontà dell'adattamento *nella popolazione* e non su un campione specifico di essa. Anche in questo caso valori vicini a 1 (e comunque superiori a .9) indicano un buon adattamento ai dati. Il CFI presenta il vantaggio, condiviso anche dal TLI, di riflettere il grado di adeguatezza relativamente bene per campioni di qualsiasi grandezza.

La figura 4 presenta le impostazioni del foglio elettronico Excel per calcolare l'indice CFI: la formula corrispondente alla cella J4 consente di calcolare l'indice sulla base dei valori del chi-quadrato e dei gradi di libertà desunti dalle soluzioni ML a 1 e a 2 fattori e dalla soluzione *nulla* (vedi sopra). Nel nostro esempio la soluzione a 1 fattore risulta fortemente inadeguata anche per il CFI, mentre quella a 2 fattori presenta un indice comparativo assolutamente adeguato (.99).

Figura 4. Foglio Excel per il calcolo dell'indice CFI

	A	B	C	D	E	F	G	H	I	J	K
1	CFI = 1 - [max(χ^2_{target} -df $_{target}$, 0)/max(χ^2_{nullo} -df $_{nullo}$, χ^2_{target} -df $_{target}$, 0)]										
2											
3		χ^2_{nullo}	df $_{nullo}$		χ^2_{target}	df $_{target}$		Soluzione ML		CFI	
4		11769,4296	15,0000		2883,993	9		1 fattore		0,755412	
5					35,5108	4,0000		2 fattori		0,997319	
6											

4.2. Il Root Means Square Error of Approximation (RMSEA, Steiger, 1990; Steiger e Lind, 1980)

Il chi quadrato valuta l'ipotesi nulla che $\mathbf{R} = \hat{\mathbf{R}}$ e presuppone che esista un modello vero nella popolazione, ovvero un modello che è perfettamente consistente con i dati empirici, ovvero che rappresenta una "fotografia" perfetta della realtà. Spesso questo è irrealistico. I modelli al più possono fornire un'immagine approssimativa della realtà. Allora, l'RMSEA valuta quanto errore commettiamo nell'approssimare la realtà con il nostro modello. Se l'errore è contenuto, il nostro modello non sarà una fotografia perfetta della realtà ma almeno la approssima sufficientemente bene. Se l'errore è grande, il nostro modello non può considerarsi nemmeno un'approssimazione della realtà. Tra le diverse formule proposte per il calcolo dell'RMSEA presentiamo la seguente che è facilmente implementabile considerando il valore del chi-quadrato ricavabile dalle soluzioni ML e GLS ottenute con SPSS:

$$RMSEA = [(\chi^2 - df)/N/df]^{1/2},$$

dove df= gradi di libertà, N= numero dei soggetti.

L'indice RMSEA (*Root Means Square Error of Approximation*, Steiger, 1990; Steiger e Lind, 1980) è un indice molto importante che fornisce una stima dell'errore di approssimazione che si commette quando la matrice delle correlazioni osservate \mathbf{R} viene riprodotta tramite la matrice $\hat{\mathbf{R}}$

ricavata dalle saturazioni di un determinato modello fattoriale. Valori inferiori a .05 suggeriscono che l'errore di approssimazione è minimo, valori compresi tra .05 e .08 indicano un errore di approssimazione accettabile, mentre valori superiori a .08 indicano che il modello fattoriale non tiene (ovvero $\hat{\mathbf{R}}$ non può nemmeno essere considerata come una approssimazione di \mathbf{R}). Questo indice rappresenta una stima della bontà dell'adattamento del modello nella popolazione, ponderata per i gradi di libertà del modello, e quindi è una misura che tiene in considerazione la parsimonia del modello.

La figura 5 presenta le impostazioni del foglio elettronico Excel per calcolare l'indice RMSEA: la formula corrispondente alla cella H4 consente di calcolare l'indice sulla base dei valori del chi-quadrato e dei gradi di libertà desunti dalle soluzioni ML a 1 e a 2 fattori (vedi sopra). Nel nostro esempio anche l'RMSEA evidenzia una soluzione a 1 fattore fortemente inadeguata, mentre quella a 2 fattori presenta un indice più basso, comunque accettabile (il valore è infatti appena al di sopra della prima soglia di .05). Rispetto ai 2 indici comparati visti sopra questo indice penalizza maggiormente il modello perché si tratta di un modello con pochi gradi di libertà, quindi poco parsimonioso.

Figura 5. Foglio Excel per il calcolo dell'indice RMSEA

	A	B	C	D	E	F	G	H	I	J	K
1	RMSEA = $[(\chi^2 - df)/N/df]^{1/2}$,										
2											
3		χ^2	df	N		Soluzione ML		RMSEA			
4		2883,993	9	3075		1 fattore		0,322311			
5		35,5108	4			2 fattori		0,050615			
6											

4.3. Lo SRMSR (*Standardized Root Mean Square Residual*, Jöreskog e Sörbom, 1984)

L'indice SRMSR (*Standardized Root Mean Square Residual*) fornisce una stima della media della correlazione residua, cioè non spiegata dal modello. Esso indica un buon adattamento se il suo valore è piccolo (vengono raccomandati valori inferiori a .08 e preferibilmente inferiori a .05). Può essere calcolato con la seguente formula:

$$\text{SRMSR} = [2\sum_i\sum_j(r_{ij} - \hat{r}_{ij})^2/(q(q+1))]^{1/2}$$

dove q è il numero totale delle variabili, r_{ij} e \hat{r}_{ij} sono rispettivamente la correlazione osservata e la correlazione riprodotta tra le variabili i e j . La figura 6a presenta le impostazioni del foglio elettronico Excel per calcolare l'indice SRMSR: la formula corrispondente alla cella C17 consente di calcolare l'indice sulla base dei valori delle correlazioni residue desunte dalla soluzione ML a 1 fattore. In particolare le celle nel quadrato C3-H8 presentano i valori delle correlazioni residue, mentre le celle nel quadrato C10-H15 presentano i valori delle correlazioni residue elevate al quadrato (questa operazione può essere effettuata all'interno di Excel tramite le funzioni aritmetiche). Nel nostro esempio anche l'SRMSR evidenzia una soluzione a 1 fattore fortemente inadeguata (l'indice è uguale a .12), mentre quella a 2 fattori presenta un indice molto più basso (0.007, vedi figura 6b) inferiore alla soglia di .08 raccomandata in letteratura. L'indice per la soluzione a 2 fattori può essere ricavato direttamente dal foglio elettronico impostato come in figura 6 semplicemente sostituendo nel quadrato C3-H8 i valori delle correlazioni residue della soluzione a 1 fattore con quelli della soluzione a 2 fattori.

Figura 6a. Foglio Excel per il calcolo dell'indice SRMR (soluzione a 1 fattore)

	A	B	C	D	E	F	G	H	I	J
1										
2										
3		Residui		0,3170	-0,0964	-0,0707	-0,0714	0,3196		
4			0,3170		-0,0881	-0,0487	-0,0481	0,2552		
5			-0,0964	-0,0881		0,0599	0,0543	-0,0966		
6			-0,0707	-0,0487	0,0599		-0,0078	-0,0677		
7			-0,0714	-0,0481	0,0543	-0,0078		-0,0369		
8			0,3196	0,2552	-0,0966	-0,0677	-0,0369			
9										
10				0,1005	0,0093	0,0050	0,0051	0,1022		
11		Residui	0,1005		0,0078	0,0024	0,0023	0,0651		
12		al quadrato	0,0093	0,0078		0,0036	0,0029	0,0093		
13			0,0050	0,0024	0,0036		0,0001	0,0046		
14			0,0051	0,0023	0,0029	0,0001		0,0014		
15			0,1022	0,0651	0,0093	0,0046	0,0014			
16										
17			0,1237							
18										

Figura 6b. Foglio Excel per il calcolo dell'indice SRMR (soluzione a 2 fattori)

	A	B	C	D	E	F	G	H	I	J
18										
19										
20										
21	Residu	AGG9 Dinamico		0,0009	0,0052	-0,0073	-0,0156	0,0024		
22		AGG11 Ener	0,0009		-0,0047	0,0118	0,0053	-0,0049		
23		AGG20 Resp	0,0052	-0,0047		0,0004	0,0009	-0,0031		
24		AGG21 Scru	-0,0073	0,0118	0,0004		-0,0048	-0,0022		
25		AGG2_B Affi	-0,0156	0,0053	0,0009	-0,0048		0,0210		
26		AGG8_B De	0,0024	-0,0049	-0,0031	-0,0022	0,0210			
27										
28				0,0000	0,0000	0,0001	0,0002	0,0000		
29			0,0000		0,0000	0,0001	0,0000	0,0000		
30			0,0000	0,0000		0,0000	0,0000	0,0000		
31			0,0001	0,0001	0,0000		0,0000	0,0000		
32			0,0002	0,0000	0,0000	0,0000		0,0004		
33			0,0000	0,0000	0,0000	0,0000	0,0004			
34										
35			0,0070							
36										
37										

4.4. Gli indici di Cliff

Cliff (1987) ha proposto alcuni indici, che definisce *indici di accuratezza*, i quali possono fornire informazioni rilevanti sul numero di fattori presenti nella soluzione, e quindi sulla bontà dell'adattamento del modello di analisi fattoriale. Questi indici si basano sulla corrispondenza tra autovalori e varianza spiegata; si tratta dunque di indici applicabili eminentemente ai risultati dell'analisi delle componenti principali e dell'analisi in fattori principali (vedi Barbaranelli, 2003). In particolare, è dimostrabile che in una soluzione di ACP o di AFP che prevede r componenti/fattori e q variabili, la somma degli ultimi $q-r$ autovalori al quadrato di una matrice di correlazione \mathbf{R} completa o di una matrice \mathbf{R}_h (con stime delle communalità nella diagonale principale)

equivale alla somma dei residui al quadrato (cioè la somma dei valori al quadrato nella matrice dei residui), ovvero:

$$\sum_i \sum_j (r_{ij} - \hat{f}_{ij})^2 = \sum_{m=r+1}^q \lambda_m^2.$$

Questa equazione fornisce un indice generale della bontà dell'adattamento intesa come la capacità della soluzione fattoriale di spiegare (ovvero di riprodurre) la matrice delle correlazioni, tramite il rapporto della somma dei primi r autovalori al quadrato, rispetto alla somma totale degli autovalori al quadrato, ovvero:

$$g_r = \frac{\sum_{m=1}^r \lambda_m^2}{\sum_{m=1}^q \lambda_m^2}, \text{ dove } r = \text{numero di fattori estratti, } q = \text{numero di variabili.}$$

Dalla equazione precedente può essere ottenuto facilmente un altro indice che rappresenta la media dei valori nella matrice di correlazione residua:

$$d_r = \left(\sum_{m=r+1}^q \lambda_m^2 / q^2 \right)^{1/2} = \left[\sum_i \sum_j (r_{ij} - \hat{f}_{ij})^2 / q^2 \right]^{1/2}$$

I due indici g_r e d_r sono entrambi indici di accuratezza della soluzione fattoriale e risultano inversamente proporzionali. In un buon modello fattoriale è augurabile che l'indice g_r sia maggiore di .95, e che l'indice d_r sia minore di .05. Questi indici sono sicuramente appropriati nel caso in cui si effettua un'analisi delle componenti principali o dei fattori principali, ovvero procedimenti di fattorializzazione basati direttamente sul calcolo degli autovalori. L'applicazione a tecniche come quelle basate sui minimi quadrati e sulla massima verosimiglianza potrebbe non essere completamente appropriata, poiché queste tecniche non si basano direttamente sulla scomposizione di \mathbf{R} tramite i suoi autovalori e autovettori.

Le figure 7 e 8 presentano le impostazioni del foglio elettronico Excel per calcolare gli indici di Cliff: le formule corrispondenti alle celle E8 ed E9 consentono di calcolare gli indici sulla base dei valori degli autovalori delle matrici completa e ridotta per le soluzioni da 1 a 6 componenti (nel caso in cui si utilizza la matrice completa) e da 1 a 5 fattori (nel caso in cui si utilizza la matrice ridotta: occorre ricordare qui quanto detto in merito agli autovalori della matrice ridotta nel paragrafo 5.2). Nel nostro esempio l'indice g_r è adeguato nella soluzione a 2 fattori sia nel caso

dell'ACP sia nel caso dell'AFP. L'indice d_r invece è adeguato a partire dalla soluzione a 2 fattori nel caso della AFP, mentre per la ACP solo con 5 componenti estratte l'indice diventa inferiore a .05, quindi anche se si estraggono 3 o 4 componenti le correlazioni residue continuano ad essere elevate. Si tratta chiaramente di un ulteriore esempio di come l'analisi delle componenti principali risulti inadeguata per ricostruire le correlazioni tra le variabili, ovvero ciò che le variabili condividono (si veda al riguardo quanto detto nel testo alle pagine 103 e 104).

Figura 7. Foglio Excel per il calcolo dell'indice g_r di Cliff

	A	B	C	D	E	F	G	H	I	J	K	L
1												
2												
3	$g_r = \frac{\sum_{m=1}^r \lambda_m^2}{\sum_{m=1}^q \lambda_m^2}$											
4												
5		Matrice R completa					Matrice R ridotta					
6												
7	Autovalori	Autov. Quadr	Soluzione	gr	Autovalor	Autov. Quadr	Soluzione	gr				
8	3,8243	14,6252	1	0,9100	3,5876	12,8705	1	0,9495				
9	1,0545	1,1119	2	0,9792	0,8224	0,6764	2	0,9994				
10	0,3655	0,1336	3	0,9875	0,0730	0,0053	3	0,9998				
11	0,3205	0,1027	4	0,9939	0,0411	0,0017	4	0,9999				
12	0,2564	0,0657	5	0,9980	0,0298	0,0009	5	1,0000				
13	0,1789	0,0320	6	1,0000								
14												
15												

Figura 8. Foglio Excel per il calcolo dell'indice dr di Cliff

Microsoft Excel - indici.fit.xls

File Modifica Visualizza Inserisci Formato Strumenti Dati Finestra ? Adobe PDF

Rispondi con modifiche... Termina revisione...

E8 $\text{=RADQ(SOMMA(B9:B13)/(6^2))}$

$$d_r = \left(\sum_{m=r+1}^q \lambda_m^2 / q^2 \right)^{1/2}$$

Matrice R completa				Matrice R ridotta			
Autovalori	Autov. Quadr	Soluzione	dr	Autovalor	Autov. Quadr	Soluzione	
3,8243	14,6252	1	0,2004	3,5876	12,8705	1	
1,0545	1,1119	2	0,0963	0,8224	0,6764	2	
0,3655	0,1336	3	0,0746	0,0730	0,0053	3	
0,3205	0,1027	4	0,0521	0,0411	0,0017	4	
0,2564	0,0657	5	0,0298	0,0298	0,0009	5	
0,1789	0,0320	6	0,0000				

4.5. Considerazioni conclusive sugli indici alternativi di bontà dell'adattamento

In conclusione, gli indici alternativi di bontà dell'adattamento possono dare informazioni supplementari molto importanti per valutare la bontà di un modello fattoriale. Mentre gli indici comparativi e l'RMSEA sono ricavabili solo per le soluzioni ML e GLS (infatti si basano sull'indice chi-quadrato), e gli indici di Cliff sono ricavabili solo per le soluzioni che si basano sul calcolo di autovalori e autovettori (quindi componenti e fattori principali), l'indice SRMSR invece è ricavabile per *qualsiasi* soluzione poiché si basa sulle correlazioni residue che sono sempre ottenibili da ogni soluzione fattoriale.

5. L'analisi fattoriale con il linguaggio SINTASSI

Di seguito presentiamo i programmi SINTASSI per le analisi descritte nel testo.

**** Analisi dei fattori principali

```
FACTOR
/VARIABLES agg9 agg11 agg8_b agg20 agg21 agg2_b
/MISSING LISTWISE
/ANALYSIS agg9 agg11 agg8_b agg20 agg21 agg2_b
/PRINT UNIVARIATE INITIAL CORRELATION
SIG DET KMO REPR AIC EXTRACTION
/PLOT EIGEN
/CRITERIA FACTORS(2) ITERATE(1)
/EXTRACTION PAF
/ROTATION PROMAX(4)
/METHOD=CORRELATION .
```

Illustriamo brevemente i comandi specificati nella procedura FACTOR.

Il comando VARIABLES definisce le variabili che possono essere sottoposte ad analisi.

Il comando /MISSING defisse il metodo utilizzato per il trattamento dei valori mancanti (in questo caso è lasciata l'opzione di default LISTWISE).

Il comando /ANALYSIS specifica quale saranno le variabili che effettivamente verranno sottoposte ad analisi fattoriale (se è assente verranno utilizzate le variabili indicate in /VARIABLES).

Il comando /PRINT serve a specificare le opzioni relative agli elementi che si troveranno nell'output.

Il comando /PLOT serve a specificare le opzioni relative alle rappresentazioni grafiche (in questo caso si richiede lo scree plot degli autovalori).

Il comando /CRITERIA serve a specificare alcune opzioni aggiuntive che in questo caso riguardano il numero di fattori che verranno estratti e poi ruotati, FACTORS(2), e il numero di iterazioni massime per ottenere la soluzione iniziale, ITERATE(25).

Il comando /EXTRACTION specifica il metodo utilizzato per estrarre i fattori, che in questo caso è l'analisi in fattori principali (PAF, Principal Axis Factoring).

Il comando /ROTATION invece specifica il metodo di rotazione utilizzato che in questo caso è la rotazione obliqua PROMAX: il valore tra parentesi indica il grado di obliquità che verrà tenuto nella soluzione, e il valore 4 rappresenta una soluzione di “media” obliquità (per una descrizione del comando si veda il testo alla pagina 76).

Il comando /METHOD specifica il tipo di matrice che viene analizzata, in questo caso si tratta dell’opzione di default ovvero una matrice di correlazioni.

****** Analisi dei fattori principali - Calcolo dei punteggi fattoriali.**

```
FACTOR
/VARIABLES agg9 agg11 agg8_b agg20 agg21 agg2_b /MISSING LISTWISE
/ANALYSIS agg9 agg11 agg8_b agg20 agg21 agg2_b
/PRINT INITIAL EXTRACTION ROTATION FSCORE
/PLOT ROTATION
/CRITERIA FACTORS(2) ITERATE(25)
/EXTRACTION PAF
/CRITERIA ITERATE(25)
/ROTATION PROMAX(4)
/SAVE REG(ALL)
/METHOD=CORRELATION .
```

Il comando “/SAVE REG(ALL)” consente di calcolare e salvare i punteggi fattoriali dei 2 fattori estratti e ruotati nell’analisi dei fattori principali.

****** Analisi delle componenti principali.**

```
FACTOR
/VARIABLES agg9 agg11 agg8_b agg20 agg21 agg2_b /MISSING LISTWISE
/ANALYSIS agg9 agg11 agg8_b agg20 agg21 agg2_b
/PRINT INITIAL EXTRACTION ROTATION
/CRITERIA FACTORS(2) ITERATE(25)
/EXTRACTION PC
/CRITERIA ITERATE(25)
/ROTATION PROMAX(4)
/METHOD=CORRELATION .
```

Specificando il comando “/EXTRACTION PC” viene effettuata l’analisi delle componenti principali. Vale la pena ricordare che si tratta del metodo di default di SPSS.

****** Analisi di massima verosimiglianza con rotazione
Varimax.**

```
FACTOR  
  /VARIABLES agg9 agg11 agg8_b agg20 agg21 agg2_b  /MISSING LISTWISE  
  /ANALYSIS agg9 agg11 agg8_b agg20 agg21 agg2_b  
  /PRINT INITIAL REPR EXTRACTION ROTATION  
  /CRITERIA FACTORS(2) ITERATE(25)  
  /EXTRACTION ML  
  /CRITERIA ITERATE(25)  
  /ROTATION VARIMAX .
```

Specificando il comando “/EXTRACTION ML” viene utilizzato il metodo della massima verosimiglianza per effettuata l’analisi fattoriale.

6. Opzioni di analisi ottenibili solo attraverso SINTASSI

Alcune funzionalità e opzioni di analisi della procedura di analisi fattoriale di SPSS non sono accessibili tramite il menù, ma richiedono una programmazione nel linguaggio SINTASSI. Illustreremo alcune applicazioni di questo genere di programmazione che riteniamo particolarmente interessanti. Esse rappresentano un’integrazione di quanto ottenibile utilizzando i menù e le finestre di dialogo fornite dal programma.

6.1. Stime diverse per le comunalità

Come abbiamo visto, nei metodi che richiedono stime iniziali delle comunalità i valori di default sono rappresentati dal coefficiente di correlazione multipla al quadrato tra una variabile osservata e le altre variabili osservate in analisi. Limitatamente all’analisi dei fattori principali (*Principal Axis Factoring*, PAF) è possibile specificare valori diversi per le stime iniziali delle comunalità tramite il comando **DIAGONAL**. Questo comando viene invece ignorato se il metodo di estrazione specificato è diverso da PAF. Nell’esempio descritto nel testo potremmo utilizzare, come stime iniziali delle comunalità, il valore della correlazione più elevata per ogni variabile. Dalla matrice

presentata nella tabella 2.5 a pagina 86 è facile verificare che questi valori sono i seguenti: .736, .736, .699, .808, .808, .729. Le linee di programma SINTASSI per utilizzare questi valori sono le seguenti:

```
**** Stime diverse per le comunalità.

FACTOR
/VARIABLES agg9 agg11 agg8_b agg20 agg21 agg2_b
/MISSING LISTWISE
/ANALYSIS agg9 agg11 agg8_b agg20 agg21 agg2_b
/DIAGONAL=.736 .736 .699 .808 .808 .729
/PRINT INITIAL EXTRACTION ROTATION
/CRITERIA FACTORS(2) ITERATE(25)
/EXTRACTION PAF
/CRITERIA ITERATE(25)
/ROTATION PROMAX(4)
/METHOD=CORRELATION .
```

/DIAGONAL consente di specificare i valori per le stime iniziali delle comunalità: questi vanno messi rispettando l'ordine delle variabili specificato sopra, quindi .736 fa riferimento alla variabile agg9, .736 alla variabile agg11 e così via. Poiché in /DIAGONAL abbiamo valori contigui che vengono ripetuti più volte, è possibile segnalare tali ripetizioni utilizzando il carattere *. Allora il comando diventa: **/DIAGONAL= 2*.736 .699 2*.808 .729** Questo sta a significare che il valore .736 e il valore .808 sono ripetuti due volte.

6.2. Rappresentazione grafica delle variabili nello spazio fattoriale

Abbiamo visto che è possibile richiedere la rappresentazione grafica della posizione delle variabili nello spazio fattoriale utilizzando come coordinate le saturazioni nella matrice dei modelli se la soluzione è obliqua, ovvero nell'unica matrice ruotata presente in output se la soluzione è ortogonale. Se vengono estratti più di due fattori il programma automaticamente costruisce dei plot tridimensionali sulla base delle saturazioni nei primi tre fattori. Spesso questo non è il modo migliore per rappresentare graficamente le soluzioni fattoriali, invece un metodo più efficace è rappresentato dai grafici bidimensionali ottenibili incrociando due fattori alla volta. Supponiamo ad

esempio di aver estratto 3 fattori nel nostro esempio: i comandi SINTASSI da utilizzare per ottenere i plot bidimensionali sono illustrati di seguito.

****** Grafici bidimensionali dello spazio fattoriale.**

```
FACTOR
/VARIABLES agg9 agg11 agg8_b agg20 agg21 agg2_b
/MISSING LISTWISE
/ANALYSIS agg9 agg11 agg8_b agg20 agg21 agg2_b
/PRINT INITIAL EXTRACTION ROTATION
/PLOT ROTATION (1,2) (1,3) (2,3)
/CRITERIA FACTORS(3) ITERATE(25)
/EXTRACTION PAF
/CRITERIA ITERATE(25)
/ROTATION PROMAX(4)
/METHOD=CORRELATION .
```

Tramite il comando /PLOT vengono effettuate le rappresentazioni grafiche bidimensionali relative agli spazi individuati dalle seguenti coppie di fattori: fattore 1 e fattore 2, fattore 1 e fattore 3, fattore 2 e fattore 3.

6.3. Salvataggio di punteggi fattoriali con nomi specificati dall'utente

I punteggi fattoriali nei fattori estratti possono essere salvati tramite la finestra di dialogo presentata in figura 2.7 pagina 78. In questo modo, tuttavia, il programma fornisce nomi di default per i punteggi fattoriali che vengono creati. L'utente può utilizzare nomi ad hoc utilizzando linee di programma SINTASSI come le seguenti:

****** Nomi diversi per i punteggi fattoriali.**

```
FACTOR
/VARIABLES agg9 agg11 agg8_b agg20 agg21 agg2_b
/MISSING LISTWISE
/ANALYSIS agg9 agg11 agg8_b agg20 agg21 agg2_b
/PRINT INITIAL EXTRACTION ROTATION
/PLOT ROTATION
/CRITERIA FACTORS(2) ITERATE(25)
/EXTRACTION PFA
/CRITERIA ITERATE(25)
/ROTATION PROMAX(4)
/SAVE REG (2 F_).
```

Il programma salverà due punteggi fattoriali con i seguenti nomi: F_1 e F_2.

6.4. Analisi diretta di matrici

E' possibile leggere e scrivere file in formato SPSS che contengono una matrice tramite il comando **MATRIX**. Questo tipo di file vengono chiamati "*matrix data files*". Le linee di comandi SINTASSI che seguono consentono di effettuare l'analisi fattoriale direttamente su una matrice di correlazioni opportunamente salvata come matrix data file SPSS.

****** Analisi diretta di una matrice di correlazioni.**

```
FACTOR
MATRIX IN(COR=*)
/CRITERIA=FACTORS(2)
/EXTRACTION=PAF
/ROTATION=PROMAX.
```

Il comando **MATRIX** deve essere sempre specificato per primo senza bisogno di specificare il nome delle variabili. E' possibile specificare solo una opzione **IN** e una opzione **OUT** nell'ambito di un unico comando **MATRIX**. Mentre l'opzione **IN** consente di leggere matrici, l'opzione **OUT** consente di scriverle su un file esterno. Le opzioni **IN** e **OUT** nel comando **MATRIX** possono essere specificate in qualsiasi ordine. Se queste opzioni vengono specificate più di una volta, la

procedure **FACTOR** non viene eseguita. L'utente deve indicare il tipo di matrice che viene letta o che viene scritta nelle opzioni **IN** o **OUT**. Nella procedura **FACTOR** i tipi di matrice ammissibili sono: **COR** per indicare una matrice di correlazione, **COV** per indicare una matrice di covarianza, e **FAC** per indicare una matrice di saturazioni fattoriali.

MATRIX=IN non può essere utilizzato se non è stato già caricato un file SPSS. Per leggere una matrice in un file esterno all'inizio di una sessione SPSS, bisogna prima caricare il file, e poi specificare **IN(COR=*)**, **IN(COV=*)** o **IN(FAC=*)** nel comando **MATRIX**. L'utente deve indicare il tipo di matrice tra parentesi immediatamente prima di fornire nome e percorso per il file che contiene la matrice [ad esempio, **IN(COR=NOMEFILE)**]. Nel nostro esempio abbiamo chiesto al programma di analizzare una matrice di correlazioni (**COR**) e poiché la matrice è stata già caricata nel programma (quindi rappresenta il nostro file attivo) abbiamo semplicemente scritto un asterisco nello spazio riservato per indicare il nome e il percorso del file. In questo caso ovviamente il file attivo è costituito da una matrice e non da dati nel solito formato Casi X Variabili. Se si legge una matrice da un file esterno questa non rimpiazza il file attivo in quel momento.

L'opzione **OUT (nome del file)** consente di salvare un matrix data file. Anche in questo caso è necessario specificare il tipo di matrice (**COR**, **COV**, or **FAC**) e il nome del file tra parentesi [es., **OUT(COR=NOMEFILE)**]. Se si specifica un asterisco invece del nome del file [es., **OUT(COR=*)**] la matrice che viene creata sostituirà il file attivo in quel momento e non verrà salvata finché l'utente non lo farà utilizzando il menù file o i comandi **SAVE** o **XSAVE**. Ovviamente, non è possibile salvare in **OUT** una matrice di varianze/covarianze se la matrice di input è una matrice di saturazioni o una matrice di correlazioni senza deviazioni standard.

Mentre il file che contiene una matrice di correlazioni e di covarianze può includere anche il numero di soggetti, le medie e le deviazioni oltre alle correlazioni e alle covarianze, il file che contiene la matrice di saturazioni fattoriali può includere solo i valori delle saturazioni.

Poiché l'opzione **VARIABLES** non può essere utilizzata con **MATRIX IN**, l'utente tramite il comando **ANALYSIS** può specificare su quali variabili effettuare l'analisi. Ad esempio, per

restringere l'analisi alle sole variabili agg9 agg11 agg20 agg21 è possibile specificare la seguente linea di comando: /ANALYSIS agg9 agg11 agg20 agg21. Quando viene analizzata una matrice di saturazioni, inoltre, neppure questa opzione è possibile, quindi l'analisi va effettuata su tutte le variabili presenti nel file. L'opzione di default per il comando ANALYSIS prevede l'utilizzo di tutte le variabili nella matrice.

Caratteristiche di un matrix data file

La figura 9 presenta la matrice di correlazione analizzata tramite i comandi specificati nel quadro precedente. Una matrice di correlazioni o di covarianze del tipo leggibile da SPSS deve contenere due variabili "speciali" che sono create da SPSS nel momento in cui il file viene salvato in formato matrix. **ROWTYPE_** è una variabile di tipo stringa che serve per codificare il tipo di parametri presenti nelle righe e nelle colonne della matrice: MEAN indica che la riga corrispondente conterrà le medie delle variabili, STDDEV indica che la riga corrispondente conterrà le deviazioni standard delle variabili, N indica che la riga corrispondente conterrà il numero di soggetti sui quali sono stati calcolati i coefficienti di correlazione, CORR indica che la riga corrispondente conterrà dei coefficienti di correlazione di Pearson. In una matrice di varianze/covarianze verrà utilizzato il valore COV per indicare che indica che la riga corrispondente conterrà delle covarianze. **VARNAME_** è una variabile di tipo stringa che viene utilizzata per indicare il nome delle variabili incluse nella matrice. Per poter effettuare correttamente un'analisi basta che nella matrice ci siano i coefficienti di correlazione: l'analisi infatti viene effettuata anche se nella matrice *non* sono presenti le medie, le deviazioni standard e il numero di casi. Quando vengono utilizzati i metodi di estrazione ML e GLS l'informazione relativa al numero dei soggetti è necessaria per poter effettuare il test del chi-quadrato: l'analisi comunque viene condotta anche se tale informazione non è presente, ma non verrà effettuato il test di bontà dell'adattamento.

Figura 9. Matrix data file: matrice delle correlazioni con medie, deviazioni standard e numerosità dei soggetti

	rowtype_	varname_	agg9	agg11	agg8_b	agg20	agg21	agg2_b	var
1	MEAN		5,102764	4,992846	1,465017	4,686504	4,693984	,4810946	
2	STDDEV		1,602447	1,607141	,4771177	1,835969	1,763418	,2643651	
3	N		3075,000	3075,000	3075,000	3075,000	3075,000	3075,000	
4	CORR	AGG9	1,000000	,7358239	,6992088	,4400856	,4595316	,4066734	
5	CORR	AGG11	,7358239	1,000000	,6729541	,5024184	,5349310	,4781489	
6	CORR	AGG8_B	,6992088	,6729541	1,000000	,4385452	,4612381	,4399815	
7	CORR	AGG20	,4400856	,5024184	,4385452	1,000000	,8075501	,7283793	
8	CORR	AGG21	,4595316	,5349310	,4612381	,8075501	1,000000	,6585110	
9	CORR	AGG2_B	,4066734	,4781489	,4399815	,7283793	,6585110	1,000000	
10									

La figura 10 presenta una matrice di saturazioni che può essere analizzata tramite i comandi specificati nel quadro seguente.

****** Analisi diretta di una matrice di saturazioni.**

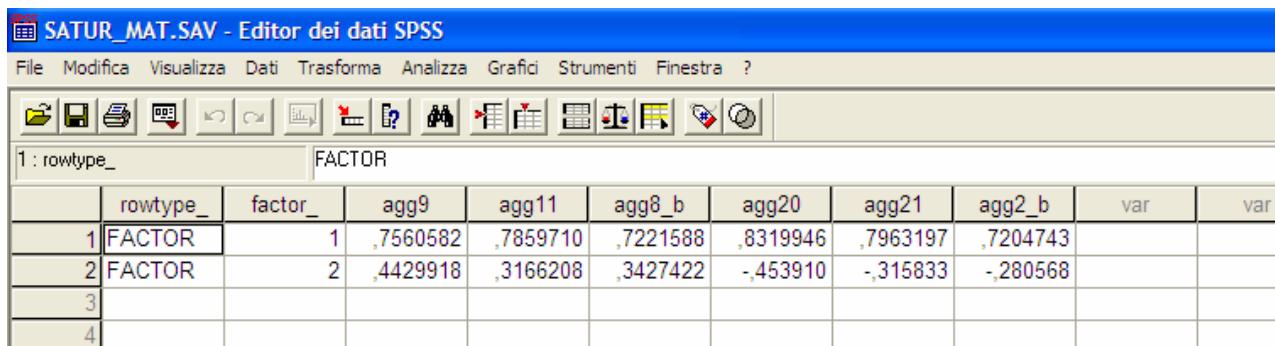
```

FACTOR
  MATRIX IN(FAC=*)
  /ROTATION=PROMAX.
    
```

La variabile di sistema **ROWTYPE_** contiene il valore **FACTOR** ad indicare che la riga corrispondente conterrà le saturazioni delle variabili nel fattore specificato nella variabile di sistema **VARNAME_**. Non è possibile specificare un comando **EXTRACTION** quando si utilizza come input una matrice di saturazioni. Ovviamente non è possibile modificare il metodo di estrazione dei fattori con questo tipo di input: invece si può cambiare il metodo di rotazione dei fattori, ma comunque verranno ruotati tutti i fattori presenti nella matrice. Non bisogna quindi specificare il

numero di fattori perché è quello previsto dal numero di righe nella matrice (nell'esempio della figura 10 i fattori sono due).

Figura 10. Matrix data file: matrice delle saturazioni fattoriali



The screenshot shows the SPSS data editor window titled 'SATUR_MAT.SAV - Editor dei dati SPSS'. The menu bar includes 'File', 'Modifica', 'Visualizza', 'Dati', 'Trasforma', 'Analizza', 'Grafici', 'Strumenti', and 'Finestra ?'. The toolbar contains various icons for file operations and data manipulation. The data grid shows a matrix with columns for 'rowtype_', 'factor_', 'agg9', 'agg11', 'agg8_b', 'agg20', 'agg21', 'agg2_b', 'var', and 'var'. The first two rows represent factors 1 and 2, with their respective loadings on the variables.

	rowtype_	factor_	agg9	agg11	agg8_b	agg20	agg21	agg2_b	var	var
1	FACTOR	1	,7560582	,7859710	,7221588	,8319946	,7963197	,7204743		
2	FACTOR	2	,4429918	,3166208	,3427422	-,453910	-,315833	-,280568		
3										
4										

Anche quando si analizza una matrice di correlazione o di covarianze è possibile richiedere gli indici che consentono di valutare la fattorializzabilità della matrice (KMO, test di Bartlett). Inoltre, anche in questo caso è possibile calcolare i coefficienti dei punteggi fattoriali e i factor scores determinancy (vedi sopra). I comandi specificati nel quadro seguente possono essere utilizzati per richiedere questo genere di informazioni.

****** Analisi diretta di una matrice di correlazioni
Calcolo degli indici di fattorializzabilità e
Dei coefficienti dei punteggi fattoriali.**

```
FACTOR  
MATRIX IN(COR=*)  
/PRINT DEFAULT FSCORE KMO  
/CRITERIA=FACTORS(2)  
/EXTRACTION=PAF  
/ROTATION=VARIMAX.
```

6.4.1. Come creare un input file SPSS in formato matriciale

Se il ricercatore ha a disposizione i dati grezzi non ha bisogno di analizzare una matrice ma utilizzerà direttamente il file di dati nel solito formato Casi X Variabili come input per l'analisi. A volte può capitare che il ricercatore debba analizzare dati che non dispone nel formato originale, ma per i quali ha soltanto informazioni "sintetiche" come ad esempio una matrice di correlazioni o una matrice di covarianze. A volte inoltre può essere necessario o interessante rianalizzare dati pubblicati (nelle riviste scientifiche, oppure resi disponibili su siti internet). Se le matrici che devono essere sottoposte ad analisi sono già nel formato SPSS allora il ricercatore può utilizzare i comandi definiti sopra per effettuare le analisi di suo interesse. Se invece le matrici non sono nel formato matrix data file di SPSS si pone il problema di costruire, a partire dai dati a disposizione, una matrice adeguata per SPSS. La cosa più semplice è aprire un file dati SPSS vuoto (FILE⇒NUOVO⇒DATI), e inserire i dati specificando le variabili di sistema necessarie. Ad esempio per creare una matrice di correlazione, dopo aver aperto un file dati nuovi, inseriamo in Visualizzazione Variabili nella colonna Nome i nomi delle due variabili di sistema **rowtype_** e **varname_** e definiamole come variabile stringa (figura 11). Inseriamo sempre nella colonna Nome i nomi delle sei variabili che sono contenute nella matrice delle correlazioni vera e propria e settiamo a 3 il numero di decimali (figura 12). Possiamo salvare il file con il nome **MATRICE_DI_CORRELAZIONE.SAV** e passare a "Visualizzazione dati" dove troveremo un file di dati completamente vuoto (figura 13). Inseriamo nella prima riga il carattere N nella prima colonna e nelle colonne 3-8 il numero 3075, che rappresenta il numero di soggetti sui quali sono state calcolate le correlazioni (figura 14). Ora inseriamo nelle righe 2-7 della prima colonna il carattere CORR, inseriamo nelle righe 2-7 della seconda colonna le etichette delle 6 variabili, e copiamo i valori dei coefficienti di correlazione presenti nella tabella 7 nelle celle che incrociano le righe 2-7 alle colonne 3-8: otterremo alla fine un file dati come quello della figura 15. Una volta salvato, questo file può essere sottoposto ad analisi fattoriale tramite le linee di programma definite sopra.

Tabella 7. Matrice delle correlazioni

1,000	,736	,699	,440	,460	,407
,736	1,000	,673	,502	,535	,478
,699	,673	1,000	,439	,461	,440
,440	,502	,439	1,000	,808	,728
,460	,535	,461	,808	1,000	,659
,407	,478	,440	,728	,659	1,000

Figura 11. Editor dei dati in “Visualizzazione variabili” parte 1

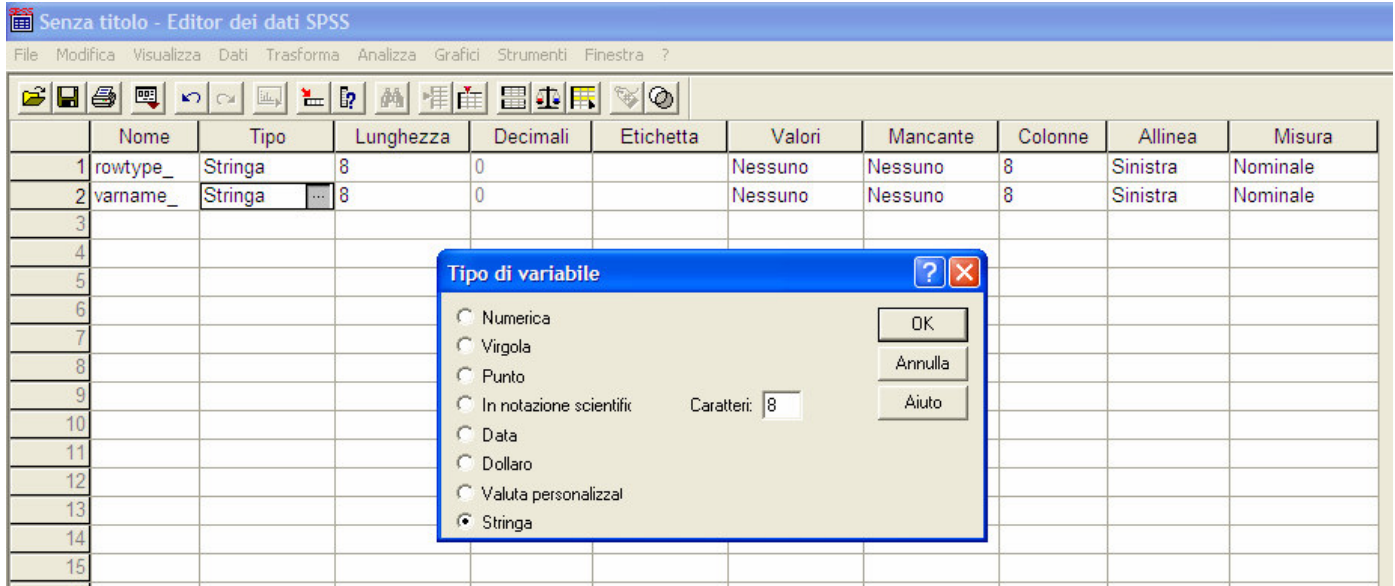


Figura 12. Editor dei dati in “Visualizzazione variabili” parte 2

The screenshot shows the SPSS 'Editor dei dati' window for a file named 'Senza titolo'. The window title bar includes the file name and the application name. Below the title bar is a menu bar with options: File, Modifica, Visualizza, Dati, Trasforma, Analizza, Grafici, Strumenti, Finestra, and ? (Help). A toolbar with various icons is located below the menu bar. The main area contains a table with columns for variable properties:

	Nome	Tipo	Lunghezza	Decimali	Etichetta	Valori	Mancante	Colonne	Allinea	Misura
1	rowtype_	Stringa	8	0		Nessuno	Nessuno	8	Sinistra	Nominale
2	varname_	Stringa	8	0		Nessuno	Nessuno	8	Sinistra	Nominale
3	agg9	Numerica	8	3		Nessuno	Nessuno	8	Destra	Scala
4	agg11	Numerica	8	3		Nessuno	Nessuno	8	Destra	Scala
5	agg8_b	Numerica	8	3		Nessuno	Nessuno	8	Destra	Scala
6	agg20	Numerica	8	3		Nessuno	Nessuno	8	Destra	Scala
7	agg21	Numerica	8	3		Nessuno	Nessuno	8	Destra	Scala
8	agg2_b	Numerica	8	3		Nessuno	Nessuno	8	Destra	Scala
9										
10										

Figura 13. Editor dei dati in “Visualizzazione dati” parte 1

The screenshot shows the SPSS 'Editor dei dati' window for a file named 'MATRICE_DI_CORRELAZIONE.sav'. The window title bar includes the file name and the application name. Below the title bar is a menu bar with options: File, Modifica, Visualizza, Dati, Trasforma, Analizza, Grafici, Strumenti, Finestra, and ? (Help). A toolbar with various icons is located below the menu bar. The main area shows a data grid with the following columns:

	rowtype_	varname_	agg9	agg11	agg8_b	agg20	agg21	agg2_b	var
1									
2									
3									
4									
5									
6									
7									
8									

Figura 14. Editor dei dati in “Visualizzazione dati” parte 2

	rowtype_	varname_	agg9	agg11	agg8_b	agg20	agg21	agg2_b	var
1	N		3075,000	3075,000	3075,000	3075,000	3075,000	3075,000	
2									
3									
4									
5									
6									
7									
8									

Figura 15. Editor dei dati in “Visualizzazione dati” parte 3

	rowtype_	varname_	agg9	agg11	agg8_b	agg20	agg21	agg2_b	var
1	N		3075,000	3075,000	3075,000	3075,000	3075,000	3075,000	
2	CORR	agg9	1,000	,736	,699	,440	,460	,407	
3	CORR	agg11	,736	1,000	,673	,502	,535	,478	
4	CORR	agg8_b	,699	,673	1,000	,439	,461	,440	
5	CORR	agg20	,440	,502	,439	1,000	,808	,728	
6	CORR	agg21	,460	,535	,461	,808	1,000	,659	
7	CORR	agg2_b	,407	,478	,440	,728	,659	1,000	
8									

E' possibile anche leggere una matrice salvata in formato testo in modo da convertirla in un matrix data file di SPSS. La tabella 8 presenta la matrice di correlazione che è contenuta nel file “MAT_COR.TXT”. Nel file i valori delle correlazioni devono essere separati con uno o più spazi: non si possono utilizzare tab o altri segni per separare le correlazioni. I comandi specificati nel quadro seguente possono essere utilizzati per importare la matrice della tabella 7 in SPSS. Il risultato sarà quello di ottenere un file dati esattamente identico a quello nella figura 15. Il file può

essere salvato in formato .SAV in modo da poter essere aperto in SPSS senza bisogno di lanciare la sintassi fornita nel quadro successivo.

**** Comandi per leggere una matrice di correlazioni scritta in formato TXT.**

```
MATRIX DATA VARIABLES=  
  AGG9 AGG11 AGG8_B AGG20 AGG21 AGG2_B  
  /FILE=  
    'C:\MAT_COR.TXT'  
  /FORMAT=FULL  
  /N=3075  
  /CONTENTS=CORR.
```

Nelle linee di sintassi definite sopra i comandi VARIABLES, FILE e N si interpretano nel modo solito. Il comando FORMAT fa riferimento al formato della matrice che viene letta. FULL sta a significare che viene letta una matrice completa. Altre opzioni possono essere LOWER (viene letta solo la parte sotto la diagonale principale) e UPPER (viene letta solo la parte sopra la diagonale principale). Di default il programma si aspetta che ci siano anche gli elementi sulla diagonale principale (nel nostro caso sono tutti uguali a 1): per segnalare l'assenza di questi elementi occorre specificare in FORMAT l'opzione NODIAG. Il comando CONTENTS si riferisce al contenuto nelle celle della matrice: l'opzione di default è CORR (coefficienti di correlazione di Pearson). Se si analizza una matrice di covarianze allora bisogna specificare COV. Se oltre alle correlazioni si inseriscono anche le medie e le deviazioni standard bisognerà aggiungere anche MEAN e STDDEV. Per le altre opzioni il lettore può fare riferimento ai manuali del programma. Ad esempio per importare una matrice come quella nella tabella 8 è necessario utilizzare i comandi SPSS specificati nel quadro seguente. Il risultato sarà quello di ottenere un file dati identico a quello nella figura 9 (a meno di errori di arrotondamento).

**** Comandi per leggere una matrice di correlazioni scritta in formato TXT che contiene anche medie e dev standard.**

```
matrix data variables=
agg9 agg11 agg8_b agg20 agg21 agg2_b
/file=
'C:\mat_M_DS_COR.txt'
/format=full
/n=3075
/contents=mean stddev corr.
```

Tabella 8. Matrice delle correlazioni con medie e deviazioni standard

5,102764	4,992846	1,465017	4,686504	4,693984	,4810946
1,602447	1,607141	,477117	1,835969	1,763418	,2643651
1,000	,736	,699	,440	,460	,407
,736	1,000	,673	,502	,535	,478
,699	,673	1,000	,439	,461	,440
,440	,502	,439	1,000	,808	,728
,460	,535	,461	,808	1,000	,659
,407	,478	,440	,728	,659	1,000

Riferimenti bibliografici

Bentler, P.M. (1990). Comparative fit indexes in structural equations models. *Psychological Bulletin*, 107, 238-246.

Borkenau, P., e Ostendorf, F. (1990). Comparing exploratory and confirmatory factor analysis: A study on the 5-factor model of personality. *Personality and Individual Differences*, 11, 515-524.

Cattell, R.B. (1978). *The scientific use of factor analysis*. New York: Plenum Press.

Cattell, R. B., & Tsujioka, B. (1964). The importance of factor-trueness and validity, versus homogeneity and orthogonality, in test scales. *Educational and Psychological Measurement*, 24, 3-30.

Cliff, N. (1987). *Analyzing multivariate data*. San Diego: Arcourt Brace Jovanovich.

Everett, J.E. (1983). Factor comparability as a means of determining the number of factors and their rotation. *Multivariate Behavioral Research*, 18, 197-218.

Jöreskog, K.G., e Sörbom, D. (1984). *LISREL VI: Analysis of Linear Structural Relationships by Maximum Likelihood, Instrumental Variables, and Least Squares Methods*. Mooresville, IN: Scientific Software Inc.

McCrae, R.R., Zonderman, A.B., Costa, P.T., Bond, M.H., & Paunonen, S.V. (1996). Evaluating replicability of factors in the revised NEO Personality Inventory: Confirmatory Factor Analysis and Procrustes Rotation. *Journal of Personality and Social Psychology*, 70, 552-565.

Pinneau, S.R., e Newhouse, A. (1964). Measures of invariance and comparability in factor analysis for fixed variables. *Psychometrika*, 29, 271-281.

Steiger, J.H. (1990). Structural model evaluation and modification: An interval estimation approach. *Multivariate Behavioral Research*, 25, 173-180.

Steiger, J.H., e Lind, J.M. (1980). *Statistically based tests for the number of common factors*. Paper presented at the annual meeting of the Psychometric Society, Iowa City, IA.

Ten Berge, J.M.F. (1986a). Some relationships between descriptive comparisons of components from different studies. *Multivariate Behavioral Research*, 21, 29-40.

Ten Berge, J.M.F. (1986b). Rotation to perfect congruence and the cross-validation of component weights across populations. *Multivariate Behavioral Research*, 21, 41-64.

Tucker, L.R. (1951). *A method for synthesis of factor analysis studies*. (Personnel Research Section Report No. 984). Washington, D.C.: Dept. of the Army.

Tucker, L.R., & Lewis, C. (1973). The reliability coefficient for maximum likelihood factor analysis. *Psychometrika*, 38, 1-10.

Van de Vijver, F.J.R., & Leung, K. (1997). *Methods of data analysis for cross-cultural research*. Thousand Oaks, CA: Sage.