October 3, 2006

Factor Analysis Examples:

Example 1:

Factor Analysis of Gulf War Illness: What Does It Add to Our Understanding of Possible Health Effects of Deployment?

Susan E. Shapiro[1], Michael R. Lasarev[2] and Linda McCauley[2]

American Journal of Epidemiology 2002; 156:578-585.

ABSTRACT:

The authors conducted factor analysis on survey data from 1,779 Persian Gulf War veterans.

Their purposes were to: 1) determine whether factor analysis identified a unique "Gulf War syndrome" among veterans potentially exposed to chemical warfare agents; 2) compare the findings of factor analysis with those from an epidemiologic analysis of symptom prevalence; and 3) observe the behavior of factor analysis when performed on dichotomous data.

The factor analysis identified three factors, but they were not unique to any particular deployment group. A unique pattern of illness was not found for the larger group of veterans potentially exposed to chemical warfare agents; however, veterans who had witnessed the demolition of chemical warfare agents at the Khamisiyah site in Iraq had a greater prevalence of dysesthesia. An analysis of the performance of dichotomous variables in factor analysis showed that the standard criteria used to determine the number of relevant factors and the dominant variables within them may be inappropriate. While Gulf War veterans appear to suffer an increased burden of illness, there is insufficient evidence to identify a unique syndrome in this population of deployed servicemen and women. Furthermore, the results provide evidence that factor analysis may make a limited contribution in this area of research.

- For more than 10 years, scientists have debated the nature and etiology of illness among veterans of the Persian Gulf War.
- Investigators have repeatedly observed an increased prevalence of many health symptoms among troops deployed in the conflict; however, a central question has been whether these symptoms represent a distinct medical entity that can be labeled "Gulf War syndrome.
- In a recent study, McCauley et al. (16) reported on the health status of veterans who had been deployed within a 50-km radius of Khamisiyah.
    - They noted no increased risk of current self-reported symptoms among veterans deployed in the Khamisiyah area compared with those who had been deployed to the Gulf region but not to Khamisiyah.
    - Within the Khamisiyah group, however, veterans close enough to witness the demolition reported significantly more of 16 different symptoms within 2 weeks of the demolition than nonwitnesses, and all but three of these symptoms were consistent with exposure to organophosphate agents.
    - Eight years after the demolition, these same witnesses reported a significant excess of eight health-related symptoms, some of which could plausibly be related to long-term effects of low-dose exposure to chemical warfare agents.

- This paper presents the results of further analysis of the McCauley et al. (16) data

- Our first purpose was to subject the self-reported health data of Khamisiyah troops to a factor analysis to determine whether a unique pattern of symptoms or factors was present and, if so, whether the pattern differed from that of non-Khamisiyah troops and troops not deployed to the Gulf region.
- Our second purpose was to compare the results of the factor analysis with those reported by McCauley et al., especially as they related to the apparent increase in the presence of certain current symptoms among veterans who witnessed the detonation.
- Our last purpose was to examine how factor analysis behaves when the data are dichotomous.

- Study Population
  - veterans on active or reserve duty in US Army or Nat'l Guard during Gulf War (jan 1, 1991 – march 31, 1991).
  - Random sampling of three groups:
    - Those not deployed to gulf region (non-deployed)
    - Those deployed to southwestern Asia, but not within 50km of Khamisiyah (deployed non-Khamisiyah)
    - Thos deployed with 50km of Khamisiyah (deployed Khamisiyah)
  - Telephone interviews
  - 91% of 3200 contacted
  - 500 ineligible
  - 1833 interviews
  - 1779 useable interviews for analysis
    - 516 non-deployed
    - 610 deployed non-Khamisiyah
    - 653 deployed Khamisiyah
      - 162 witnessed munitions detonations (witness subgroup)
      - 405 did not (non-witness subgroup)

- Study Instrument
  - Adapted existing survey used in study of Gulf War veterans
  - Includes more information on
    - troop movements
    - neurologic symptoms
  - Reliability: reported elsewhere (????)
  - Two checklists
    - Health symptoms within 2 weeks of Khamisiyah detonations (not used here)
    - Current health symptoms within past month
  - Unclear the format of data???

- Statistical Analysis
  - Factor analysis:
    - Exploratory factor analysis per group.
    - Principal components extraction
    - Varimax rotation
    - Retained factors with eigenvalues greater than 1
    - Dominant items: loadings greater than 0.60.
  - Comparison of factor analysis results with prevalence of self-reported symptoms
    - Not able to do subgroup analysis of witnesses
    - Used hypothesis test to see if different
    - Number of factors per veteran determined (???)
    - Compared distribution of “factors” with subgroup
  - Dichotomous Variables
    - Simulations to see how factor analysis performs with binary items
    - Factor analysis performed on artificial data
    - Eigenvalues and loadings examined

RESULTS

TABLE 2. Percentages of Gulf War veterans in three different study populations reporting health symptoms used in a factor analysis

Symptom	Khamisiyah (*n* = 653)	Deployed non-Khamisiyah (*n* = 610)	Nondeployed (*n* = 516)
Tingling, burning sensation of pins and needles	16	15	8
Numbness or lack of feeling	17	17	9
Loss of muscle strength in arms/legs	21	21	8
Loss of balance/coordination	12	13	5
Dizzy spells	17	20	9
Changes in memory	51	43	14
Difficulty sleeping	38	38	13
Persistent fatigue, tiredness, or weakness	45	42	11
Depression	26	33	10
Unusual irritability/anger	35	34	11
Mood swings	36	35	10
Problems following directions/instructions	11	14	3
Difficulty concentrating	27	27	8
Cramping, aches, pains, or muscle stiffness	33	30	15
Problems breathing	32	27	11
Increased sensitivity to everyday chemicals	23	20	9
Effects from confined spaces	11	18	7
Continuous eye irritation/sensitivity	22	18	7
Unexplained weight gain of more than 10 pounds (>4.5 kg)	28	24	12

- Twenty-five symptoms, created correlation matrix of 1779 veteran data
- Removed 6 symptoms with correlation <0.3 with all others
- Factor analysis results:
  - 3 factors identified in Khamisiyah group: explained 47% of variance
  - 3 factors identified in non-deployed group: explained 52% of variance
  - 3 factors identified for deployed non-Khamisiyah group: explained 50% of variance

TABLE 3. Factors extracted from three different populations of Gulf War veterans

Symptom	Khamisiyah			Deployed non-Khamisiyah			Nondeployed		
	Factor 1 (cognitive–psychological)	Factor 2 (dysesthesia)	Factor 3 (vestibular dysfunction)	Factor 1 (cognitive–psychological)	Factor 2 (dysesthesia)	Factor 3 (vestibular and other)	Factor 1 (cognitive–psychological)	Factor 2 (neuromuscular)	Factor 3 (mixed)
Unusual irritability/ anger	X			X			X		
Depression	X			X			X		
Difficulty concentrating	X			X			X		
Difficulty following directions							X		
Difficulty sleeping				X					
Fatigue	X						X		
Memory changes	X			X					
Mood swings	X			X			X		
Loss of strength								X	
Numbness		X			X			X	
Tingling		X			X			X	
Chemical sensitivity						X			X
Dizziness			X			X			X
Loss of balance			X			X			
Shortness of breath						X			

- Witness versus non-witness
  - No significant differences in factor scores between witness and non-witness groups by non-parametric testing
  - Factor 2 differed by witness status via log-linear testing

- Performance of factor analysis with dichotomous variables.
  - Our work has shown that application of standard rules to 19 randomly generated and independently created dichotomous variables could result in models containing five factors, which explained approximately 30 percent of the total variance.
  - Even more troubling is the realization that rotated loadings in excess of 0.40, the traditional cutoff used by investigators, occurred more than 95 percent of the time in our randomly generated data set.
  - If, as our simulation demonstrated, similar results can be obtained using randomly generated data, we are forced to reconsider the existence of syndromes found in earlier studies, especially those discovered through factor analysis of dichotomous variables.

- Discussion Points
  - Although our findings are largely consistent with those of most other investigators, direct comparisons are not possible, for several reasons. First, different investigators have used different lists of symptoms in their studies. Second, the sizes, sources, and compositions of the samples used in other reports have varied considerably. Third, investigators have used different factor extraction and rotation techniques and different thresholds for factor loadings to identify their syndromes. While these analytical variations are entirely acceptable (31), this has resulted in disparate findings. Table 4 summarizes these differences.
  - The one consistent theme among these studies is that no investigators other than Haley et al. (5) have identified a unique Gulf War syndrome based on the results of a factor analysis. Our findings concur with those of most other investigators. While we were able to identify clusters of symptoms that appeared to form plausible syndromes, they were not unique to any deployment group, even among veterans who had the greatest acknowledged likelihood of having been exposed to chemical warfare agents.
  - Although the epidemiologic evidence in McCauley et al.'s earlier paper clearly demonstrated an increase in bloody diarrhea in the witness group compared with the nonwitness group (odds ratio = 3.1, 95 percent confidence interval: 1.6, 6.0), we did not include this symptom in our factor analysis, because it failed to correlate with any other symptoms. This was the only

significant symptom in the witness group that was not included in our factor analysis. This illustrates another limitation of using factor analysis for epidemiologic purposes: Factor analysis only identifies *joint associations* among variables and the latent structures that may describe them; it provides no information about individual variables, thereby making it possible to miss an isolated symptom that significantly adds to the burden of illness in a population.

Example 2:

Factor analysis and validity of the Transplant Evaluation Rating Scale in a large bone marrow transplant sample

Flora Hoodin and Karen R. Kalbfleisch

Journal of Psychosomatic Research Volume 54, Issue 5 , May 2003, Pages 465-473

ABSTRACT

Objective: Clinical and methodological challenges are involved in screening bone marrow transplant (BMT) recipients for pretransplant psychosocial adjustment in an attempt to anticipate and prevent behavioral difficulties. Validity of the Transplant Evaluation Rating Scale (TERS), which quantifies the disparate salient elements in a structured clinical assessment, has not been adequately established.

This study comprehensively investigated three questions about convergent, internal–structural, and predictive validity of the TERS:

1) how indicative the TERS is of psychosocial difficulties;
2) whether the TERS is uni- or multidimensional;
3) to what degree the TERS predicts long-range adjustment during recovery posttransplant.

Methods: Pre-BMT, 345 consecutive patients were prospectively assessed and completed the MMPI. TERS ratings were assigned retrospectively by two raters (interrater reliability r=.89).

Results: The TERS showed good convergent validity relative to MMPI subscales, and a clear, simple, two-factor structure accounting for 47% of the variance. On a subset of our sample (n=29), the factor subscales, "Defiance" and "Emotional Sensitivity," exhibited differential predictive validity to functional status at 1 year posttransplant.

Conclusions: This study, the first large-scale statistical investigation of TERS validity, provided evidence for the validity of the TERS on all three questions. The TERS is indeed indicative of psychosocial risk indexed by MMPI behavioral pathology. It has an understandable, clinically

useful factor structure. Its subordinate constructs, Defiance and Emotional Sensitivity, can and should be distinguished conceptually and measured separately. The TERS has clinical utility for specifying behavioral concerns before and guiding proactive intervention after BMT.

- Many bone marrow transplant (BMT) centers assess pretransplant psychosocial adjustment in an attempt to anticipate behavioral difficulties and prevent emotional crises.
- However, the research literature is sparse with respect to which elements of psychosocial adjustment are predictive of behavioral difficulties.
- The Transplant Evaluation Rating Scale (TERS) [1] was devised to rate and identify patients at risk for behavioral complications during and after organ transplantation.
- The TERS is comprised of 10 psychosocial domains, rated on a three-point scale indicating minimal/mild, moderate, and severe concerns, based on clinical interview.
- Preliminary psychometrics of the TERS, assessed on a sample of 35 liver transplant candidates, revealed satisfactory interrater reliability and some evidence for predictive validity in the form of statistically significant correlations between TERS scores and retrospectively assigned visual analogue scale ratings of five outcome variables 1–3 years after transplant for 28 survivors [1].

- The objective of the current study, part of an ongoing project on pre- and post-BMT psychosocial adjustment in a large BMT sample, was to comprehensively examine the validity of the TERS: first, convergent validity; second, internal–structural validity; and third, predictive validity.

- The aims of the present prospective, longitudinal study were to examine the following three questions using psychometric and factor analytic techniques.
  1) How indicative is the TERS of pre-BMT psychosocial difficulties identified by pre-BMT-assessed psychometrically sound measures, as opposed to more limited retrospectively assessed subjective ratings of prior preliminary validity studies?
  2) How internally consistent is the TERS: do TERS items load on separate factors in a clinically meaningful manner, and is the separation of TERS items supported by discriminant validity of the factors that emerge?
  3) How well does the TERS predict quality of life and functional status measured by an empirically validated instrument 1 year after transplant, as opposed to the more limited subjective ratings and lack of uniform time frame of prior investigations into predictive validity?

Study Population

- 345 BMT patients between 1989 and 1994 at a midwestern regional cancer center
- nonrandom sample of 29 used to assess predictive validity

Procedure

- psychological assessment, completed MMPI pre-BMT
- random assignment to one of two raters who retrospectively completed TERS
- high interrater reliabikity (0.89)
- convergent and internal structural validity assessed on all 345 patients.
- Predictive validity on 29 survivors at 1 years with sickness impact profile (SIP)

Measures

- TERS:
  - 10 items, indexes aspects of psychosocial functioning thought to be important in adjusting to transplant
  - items rated by clinician on 3 point scale
  - items are assigned a priori weight ranging from 1-4
  - Range of scores 26.5 – 79.5
  - High scores: more problems
  - TERS severity index (SI) computed by summing items for which subject was rated in most problematic range
- MMPI
  - Test of personality functioning
  - 566 true/false questions
  - Test-retest reliability 0.74
  - Internal consistency 0.87
- SIP
  - Behaviorally based measure of perceived health status and QOL
  - Designed for use in program planning, policy formation, monitoring
  - Test-retest reliability 0.92
  - Internal consistency 0.94
  - Clinical validity (convergent construct) 0.41 – 0.81

Analyses

- Q1: convergent validity of TERS with MMPI
  - Univariate and multivariate ANOVA
- Q2: internal consistency of TERS
  - Principal components
  - Oblique rotation
  - Corrected item total correlations
  - Cronbach's alpha
  - Multivariate ANOVA, multiple regression
- Q3: predictive validity
  - Hierarchical multiple regression

Q1 RESULTS

We performed a tripartite split on the TERS, dividing the frequency table of TERS scores at the 33rd and 66th percentiles. This produced three groups of TERS scores: low; moderate; high Thereafter, only the two extreme groups were retained for analyses since prior research indicates that people scoring in the moderate range on dimensional scales often demonstrate inconsistent responses on other measures.

All of these differences are in the predicted direction and indicate greater psychopathology for the high-scoring TERS group. Thus, those who are identified as having psychosocial risk factors for transplant recovery based on the TERS also exhibit a variety of psychologically pathological tendencies as indexed by the MMPI.

MMPI scale	Mean square	$F(1,234)$	Significance	η^2
L	9.813	0.195	.659	.001
F	4999.298	58.958	.000	.202
K	3123.109	39.909	.000	.146
Hs	3217.162	22.218	.000	.087
D	3436.301	25.876	.000	.100
Hy	526.417	5.691	.018	.024
Pd	4164.617	44.069	.000	.159
Mf	359.793	3.649	.057	.015
Pa	1566.139	20.268	.000	.080
Pt	2700.584	24.424	.000	.095
Sc	6055.526	43.782	.000	.158
Ma	1836.695	14.495	.000	.059
Si	2380.194	28.791	.000	.110

Table 1. Univariate ANOVAs for the individual MMPI scales, comparing the high versus low scorers in the tripartite split of the TERS total score

Q2 RESULTS:

- The 10 items of the TERS scale were initially subjected to a principal components analysis.
- Three eigenvalues were greater than unity (the eigenvalues for the first three factors were 3.4, 1.3, 1.0) and an inspection of a screen plot suggested the presence of two or three factors.
- After two- and three-factor solutions were examined with oblique and orthogonal rotations, the two-factor oblique solution was retained based on interpretability.
- The oblique rotation provided a clear simple structure, which accounted for 47% of the total variance.
- Six items loaded on the first factor, which accounted for 34% of the variance (eigenvalue=3.4). This factor was labeled "Defiance," as the items loading on this factor involve lack of cooperation with medical treatment (e.g., history of substance abuse, poor health behaviors, noncompliance).
- Four items loaded on the second factor, which accounted for 13% of the total variance (eigenvalue=1.3). This factor was labeled "Emotional Sensitivity," as the items loading on this factor involve cognitive and affective dysregulation in coping with the disease. The interfactor correlation was r=.35 (see Table 3).

Table 3. Factor structure of the TERS scale with Oblique (Oblimin) Rotation (n=345)

TERS items	h^2	Factor loadings	
		Defiance	Emotional sensitivity
Substance abuse	.61	.81	−.12
Health behavior	.56	.79	−.17
Compliance	.41	.64	.00
Coping history	.63	.63	.31
DSM-IV Axis II	.35	.49	.20
Family support	.30	.44	.21
DSM-IV Axis I	.61	.11	.74
Affect quality	.50	−.06	.72
Mental status	.38	−.08	.64
Disease coping	.63	.13	.56
Eigenvalue		3.4	1.3
Percent total variance		34.1	13.2

Factor correlation, r=.35.

- TERS subscale scores were formed by summing the item scores that loaded on the same factor
  - Defiance mean (±S.D.)=23.16±6.5, Cronbach's α=.75;
  - Emotional Sensitivity mean (±S.D.)=12.40±3.98, Cronbach's α=.52

Q3 RESULTS

- We investigated whether the TERS would predict quality of life and functional impairment (indicated by SIP Total score) 12 months posttransplant for a subset of 29 patients in our original sample of 345, who had also enrolled in a concurrent study of regimen adherence.
- Recognizing that medical and demographic variables may play major roles in this prediction, first we examined the bivariate correlations of SIP with medical, demographic, and psychological variables.
- The medical variables (type of transplant and diagnosis) showed no relationship, nor did three of the demographic variables (gender, race, marital status). The remaining two demographic variables showed marginal relationships (age r=.25, P=.18; years of education r=.27, P=.16).
- The psychological variables were the best predictors (Defiance r=.49, $P \leq .01$; Emotional Sensitivity r=.34, P=.06).

Table 6. Hierarchical multiple regression analysis with stepwise selection predicting SIP total at 12 months posttransplant (n=29): predictors included, in order of entry

Dependent variable	Predictor	β	t
SIP total 12 months post-BMT [$F(1,28)=9.001$*, $R^2=.243$]	First block: demographic variables		
	Age	.24	1.49
	Years of education	.53	3.58 * **
	Second block: psychological variables		
	Defiance	.49	3.00 * *
	Emotional sensitivity	.09	0.46

DISCUSSION POINTS

- Firstly, with respect to convergent validity, the significant differences in MMPI scale scores between the highest and lowest TERS tertile scores corroborate the TERS scoring system as indicative of behavioral pathology.
  - Patients with higher TERS scores have more psychopathology on the MMPI.
  - More specifically, the MMPI subscales, which account for relatively more of the variance in the TERS, are high Pd, low K, high L, and high Sc scale scores.
- Secondly, the results of this study indicate that the TERS has an understandable and clinically useful factor structure.
  - Factor I, Defiance, accounting for 34% of the variance, is a clearly demarcated behavioral factor comprised of a history of difficulties with substance abuse/use, health self-care, noncompliance, family support, Axis II interpersonal functioning, and general coping.
  - Factor II, Emotional Sensitivity, accounting for 13% of the variance, is composed of items tapping manifest depressed mood, adjustment to illness, mental status, and Axis I diagnoses.
  - The factor structure suggests that as an indicator of pretransplant psychosocial risk, the TERS total score, is actually a multifaceted construct

composed of two subordinate constructs. While related to each other empirically and logically, they can and should be distinguished conceptually and measured separately.

- Thirdly, regarding predictive validity, the TERS factor subscale scores appear to differentially flag survivors at-risk for poor adjustment over time.
  - After controlling for pre-BMT disease, treatment, and demographic variables, a significant amount of variance in functional impairment at 12 months was accounted for by pre-BMT TERS Defiance subscale scores, not by pre-BMT Emotional Sensitivity subscale scores.
  - That much of the variance remains unexplained does not reflect negatively on the validity of the TERS subscale, since a variety of variables, such as quality of aftercare, are also likely to contribute to the variance.