Paper 383-2008

Factor Analysis of Caregiver Strain Questionnaire in Clinical Research

Hui Liu and David Shen, ClinForce, LLC, King of Prussia, PA

ABSTRACT
Caregiver Strain Questionnaire (CGSQ) is expected to have multiple dimensions. It was the aim of this study to investigate and confirm the dimensional structure of CGSQ measured from subjects. We used rotated principal-components analysis as EPA to explore the dimensional structure of CGSQ. The first principal component explained 86% of the variance in the data. Cumulative variance for 3 factors is 99.3%. Three dimensions (objective caregiver strain, internalized subjective caregiver strain and externalized subjective caregiver strain) were found. Reliability of the latent constructs was measured with Cronbach's alpha (0.943, 0.876, 0.877). Finally, CPA demonstrated the relationship between the observed CGSQ and their underlying latent dimensions. SAS® procedures such as PROC CORR, PROC CALIS, PROC FACTOR were used for this study.

KEY WORDS Caregiver Strain Questionnaire, Factor Analysis, EPA, CPA

INTRODUCTION
Schizophrenia and bipolar disorder are chronic and debilitating psychiatric illnesses that affect the life quality of the patients' and their families. Onset of schizophrenia before the age of 13 years is rare, but the incidence of schizophrenia increases steadily during the adolescent years, and estimates of the lifetime prevalence range from 0.5% to 1.5%. While onset of bipolar disorder most commonly occurs in adolescence or early adulthood, 20% to 40% of adults with bipolar disorder report onset during childhood. The estimated prevalence of bipolar disorder among children and adolescents aged 9 to 17 years is 1.2%. Because schizophrenia and bipolar disorder often manifest early in life and have been associated with poor outcomes, early recognition of these disorders and effective treatments for young patients are needed. The atypical antipsychotics offer important advantages over the typical antipsychotics. This study included male and female patients who had a DSM-IV diagnosis of either bipolar I disorder or schizophrenia.

Unlike many other branches of medicine, the field of psychiatry in general lacks laboratory or other biologic measures that can be used to asses the presence or severity of illness. To more objectively define various types of mental illness and to standardize assessment of these disorders, mental health practitioners have developed a wide array of rating scales for psychiatric illness that can utilize in a variety of settings. A group of rating scales may be selected that identifying how change in a novel drug therapy affects an individual's level of functioning and satisfaction with treatment. Careful selection and appropriate use of rating scales in a given situation will assist in obtaining information of patient outcomes that is accurate and useful for future care planning. In this study, the assessments of Caregiver Strain Questionnaire (CGSQ) were performed at each visit during the treatment. GSQ is a caregiver report to assess the extent to which caregivers are affected by the caring for a child with emotional and behavioral challenges. Caregiver strain is closely related to child functioning. CGSQ the caregiver responded to would be a direct result of the child's emotional/behavioral problems. The intensity of strain from CGSQ should be a consistent predictor for the treatment outcome when the child is under medical therapy.

CGSQ, defined as a set of attributes and criteria for the assessment of caregiver status and quality of life, contains 21 items that assess strain experienced by caregivers.

Items on the CGSQ are rated on a 5-point scale with the following response options:
0=Not at all; 1=A little; 2=Somewhat; 3=Quite a bit; and 4=Very much.
CGSQ14 should be reverse coded before analysis.

The theoretical model characterizes CGSQ as multidimensional. However, different study population may group items into different topic categories as characteristics of population vary. This may serve to assign CGSQ items to factors, so statistical methods such as exploratory factor analysis are helpful to explore multivariable relationships. The aims of this study were: (i) to determine the dimensional structure of CGSQ (ii) to test the reliability of latent dimensions (iii) to confirm the derived dimensions of CGSQ.

Table 1. Caregiver Strain Questionnaire (CGSQ) Profiles

CGSQ01	Personal Time Interrupted
CGSQ02	Missed Work, Neglect Duties
CGSQ03	Family Routines Disrupted
CGSQ04	Family Having to Do Without Things
CGSQ05	Family Suffer Negative Health Effects
CGSQ06	Child Getting Into Trouble
CGSQ07	Financial Strain
CGSQ08	Less Attention to Family Members
CGSQ09	Disrupt Family Relationships
CGSQ10	Disrupt Family Social Activities
CGSQ11	Social Isolation
CGSQ12	Felt Sad or Unhappy
CGSQ13	Felt Embarrassed
CGSQ14	Relate to Child
CGSQ15	Angry Toward Child
CGSQ16	Worried About Child Future
CGSQ17	Worried About Family Future
CGSQ18	Felt Guilty
CGSQ19	Felt Resentful
CGSQ20	Felt Tired or Strained
CGSQ21	Toll on Family

MATERIALS AND METHODS

Subjects

Subjects were 489 participants at different locations in the United States. The sample was designed to be representative of the aged 9–17.

Table 2. Characteristics of Subjects in the Study

	N (%)	Mean (SD)
Age		13.2 (2.15)
Gender		
Female	230 (47.0)	
Male	259 (53.0)	
Race		
Caucasian	380 (77.7)	
Black	67 (13.7)	
Other	42 (8.6)	

All study subjects gave the signed informed consent before the treatment.

Data Collection

Data for the CGSQ were collected in a computer-assisted personal interview. Each of the 21 items describes a specific impact. Responses were made on a Likert-type scale (0-not at all, 1-a little, 2-somewhat, 3-quite a bit, 4-very much). The CGSQ score were recorded and used to characterize the dimensions as a whole.

Data Analysis

To derive the latent dimensions, we subjected the baseline data (before study medication treatment) to a principal-components analysis. CGSQ dimensional structure was explored. First, the structure of the instrument (CGSQ) was developed by principal-components analysis. Retained principal components were varimax-rotated. Items (see table) were assigned to retained rotated principal components when they had a loading on these components (i.e., dimensions) of 0.5 or greater in absolute value. The number of principal components was then determined for all CGSQ items. Cronbach's correlation coefficients was calculated to check the reliability of dimensions and interpret dimensions retained rotated principal components were found to be stable. Confirmatory factor analysis (CFA was conducted to test the hypothesized 3-factor model.

RESULTS AND DISCUSSION

1. The EFA Model is Y = Xβ+ E

where Y is a matrix of measured variables
X is a matrix of common factors
β is a matrix of weights (factor loadings)
E is a matrix of unique factors, error variation

EFA is used to determines the factor structure (model) and explain a maximum amount of variance. Factor analysis seeks to discover common factors. The technique for extracting factors attempts to take out as much common variance as possible in the first factor. Subsequent factors are, in turn, intended to account for the maximum amount of the remaining common variance until, hopefully, no common variance remains.

```
proc factor data=cgsq  method= principal reorder rotate=v scree;
run;
```

```
The FACTOR Procedure
Initial Factor Method: Principal Factors
Eigenvalues of the Reduced Correlation Matrix: Total = 12.1718005  Average = 0.57960955
```

	Eigenvalue	Difference	Proportion	Cumulative
1	10.4781418	9.4049482	0.8609	0.8609
2	1.0731936	0.5342801	0.0882	0.9490
3	0.5389135	0.2184999	0.0443	0.9933
4	0.3204136	0.0795411	0.0263	1.0196
5	0.2408725	0.0255280	0.0198	1.0394
6	0.2153445	0.0371057	0.0177	1.0571
7	0.1782388	0.0757355	0.0146	1.0717
8	0.1025034	0.0367824	0.0084	1.0802
9	0.0657210	0.0165326	0.0054	1.0856
10	0.0491884	0.0534646	0.0040	1.0896
11	-0.0042761	0.0363118	-0.0004	1.0893
12	-0.0405879	0.0087493	-0.0033	1.0859

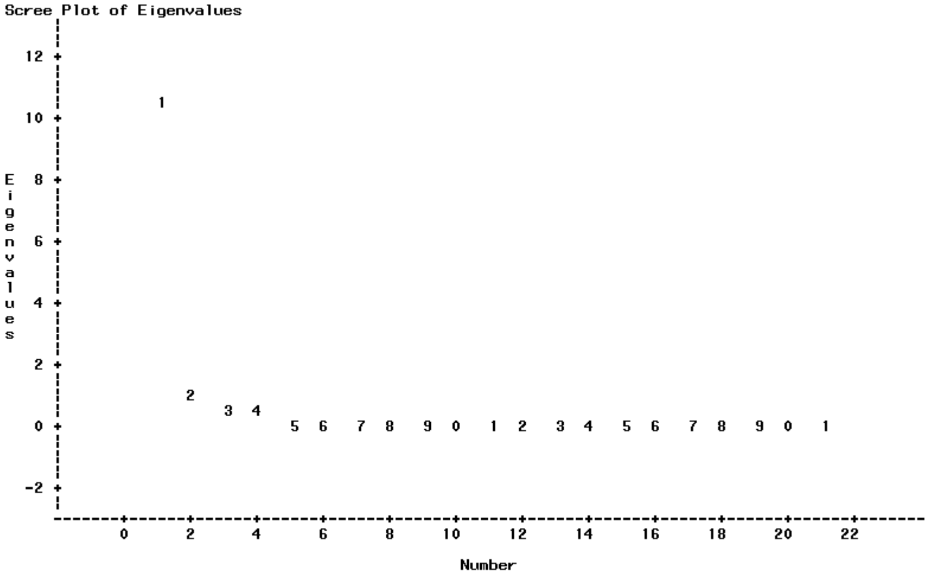
. . . . . .

```
3 factors will be retained by the NFACTOR criterion.


Scree Plot of Eigenvalues
```

The FACTOR Procedure
Initial Factor Method: Principal Factors

Scree Plot of Eigenvalues



```
     Variance Explained by Each Factor

   Factor1         Factor2         Factor3
 10.478142        1.073194        0.538914

Rotation Method: Varimax
```

Rotated Factor Pattern

		Factor1	Factor2	Factor3
CGSQ03	03. Family Routines Disrupted	0.80436	0.29319	0.19237
CGSQ02	02. Missed Work, Neglect Duties	0.78123	0.29154	0.08309
CGSQ08	08. Less Attention to Family Members	0.76933	0.23790	0.28034
CGSQ01	01. Personal Time Interrupted	0.75480	0.31418	0.21904
CGSQ04	04. Family Having to Do Without Things	0.73021	0.18035	0.29268
CGSQ10	10. Disrupt Family Social Activities	0.70554	0.25963	0.34584
CGSQ09	09. Disrupt Family Relationships	0.68151	0.29825	0.34253
CGSQ05	05. Family Suffer Neg. Health Effects	0.65475	0.21994	0.27052
CGSQ07	07. Financial Strain	0.60860	0.26144	0.13833
CGSQ06	06. Child Getting Into Trouble	0.53566	0.31186	0.16596
CGSQ11	11. Felt Isolated	0.53279	0.51800	0.22454
CGSQ16	16. Worried About Child Future	0.29224	0.68451	0.17365
CGSQ12	12. Felt Sad or Unhappy	0.39965	0.64980	0.27253
CGSQ21	21. Toll on Family	0.35509	0.60311	0.26550
CGSQ20	20. Felt Tired or Strained	0.43736	0.59665	0.26086
CGSQ17	17. Worred About Family Future	0.28076	0.58603	0.30012
CGSQ18	18. Felt Guilty	0.11726	0.54676	0.33033
CGSQ14	14. Relate to Child	0.20679	0.31260	0.69055
CGSQ19	19. Felt Resentful	0.22225	0.21912	0.64078
CGSQ15	15. Angry Toward Child	0.26456	0.32232	0.62454
CGSQ13	13. Felt Embarrassed	0.28073	0.39075	0.50549

1). In initial extraction Each factor accounts for a maximum amount of variance that has not previously been accounted for by the other factors. The factors are not correlated. Eigenvalues represent amount of variance accounted for by each factor. In common practice, factor scores are calculated with a mean or sum of measured variables that load on a factor. The exploratory factor analysis resulted in 3 principal components (PC), which together explained 99.33% of the variance. The first component explained 86.3% of the variance, and the second PC added 8.8% to the variance explained. The third explained variances were 4.4%. The preliminary eigenvalues are 10.4781418, 1.0731936 and 0.5389135, which have a significant loading (>0.50). Proportion of variance accounted for keeps a factor if it accounts for a predetermined amount of the variance of about 5%.

2). Scree test Look for an elbow in the scree plot to explore the number of factors. The scree plot supports to take 3 factors.

3). Rotation – a transformation Factors may not provide direct interpretation. Adjustment to the frames of reference by rotation methods improves the interpretation of factor loading by reducing some of the ambiguities which accompany the preliminary analysis. The process of manipulating the reference axes is known as rotation. The results of rotation methods are sometimes referred to as derived solution because they are obtained as a second stage from the results of direct solutions. Rotation applied to the reference axes means the axes are turned about the origin until some alternative position has been reached. The simplest case is when the axes are held at 90o to each other, orthogonal rotation. So the factor loading will illustrate the correlations between items and factors. The REORDER option arranges factors loading by factor from largest to smallest value

Factor1 – CGSQ01 -- CGSQ11
Factor2 – CGSQ12 CGSQ16 CHSQ17 CGSQ18 CGSQ20 CGSQ21
Factor3 – CGSQ13 CGSQ14 CGSQ15 CGSQ19

2. Reliability - Cronbach's Coefficient Alpha
Interrelated items may be summed to obtain an overall score for each participant. Cronbach's coefficient alpha estimates the reliability of this type of scale by determining the internal consistency of the test or the average correlation of items within the test.

With the ALPHA option, the CORR procedure computes Cronbach's coefficient alpha,

```
proc corr data =cgsq2 alpha  nomiss nosimple  nocorr ;
  var cgsq13 cgsq14 cgsq15 cgsq19 ;
run;
```

which is a lower bound for the reliability coefficient for the raw variables and the standardized variables. If the variances of the items vary widely, you can standardize the items to a standard deviation of 1 before computing the coefficient alpha.

Table 3. Cronbach Coefficient Alpha

	CGSQ	Raw Alpha	Standardized alpha
Factor1	1- 11	0.943264	0.944063
Factor2	12 16 17 18 20 21	0.875707	0.879068
Factor3	13 14 15 19	0.876738	0.888917

The larger the overall alpha coefficient, the more likely that items contribute to a reliable scale. Usually 0.70 is suggested as an acceptable reliability coefficient; smaller reliability coefficients are seen as inadequate. The Cronbach's coefficient alpha of 0.943, 0876, 0.877 provide the acceptable lower bound for the reliability coefficient for each dimension.

To determine how each item reflects the reliability of the scale, you calculate a coefficient alpha after deleting each variable independently from the scale. The alpha coefficient provides information on how each variable reflects the reliability of the scale with variables. If the alpha decreases after removing a variable from the construct, then this variable is strongly correlated with other variables in the scale. On the other hand, if the alpha increases after removing a variable from the construct, then removing this variable from the scale makes the construct more reliable.

The "Cronbach Coefficient Alpha with Deleted Variables" table in does not show significant increase or decrease for the alpha coefficients.

Table 4. Cronbach Coefficient Alpha with Deleted Variable

	Raw Variables		Standardized Variables	
Deleted Variable	Correlation with Total	Alpha	Correlation with Total	Alpha
CGSQ13	0.639710	0.887100	0.650382	0.895997
CGSQ14	1.000000	0.753475	0.999186	0.759283
CGSQ15	0.689582	0.862171	0.703723	0.876660
CGSQ19	0.682260	0.862065	0.694895	0.879898

PROC CORR does not automatically use listwise deletion if you specify the ALPHA option. Therefore, the NOMISS should be an option if the data set contains missing values. Otherwise, PROC CORR prints a warning message indicating the need to use the NOMISS option with the ALPHA option.

A simple structure was achieved, because each of the 21 items loaded highly on only one particular dimension and not substantially on the other dimensions. When the questionnaire was scored according to the CGSQ dimensions, pair-wise correlations between dimensions ranged between 0.61 and 0.86 (Table 5). According to guidelines, these are considered good correlations.

Table 5. Matrix of Correlations among the Three Dimensions of CGSQ

	T1	T2	T3
T1	1.00000		
T2	0.86057	1.00000	
T3	0.61333	0.63725	1.00000

3. Confirmatory Factor Analysis

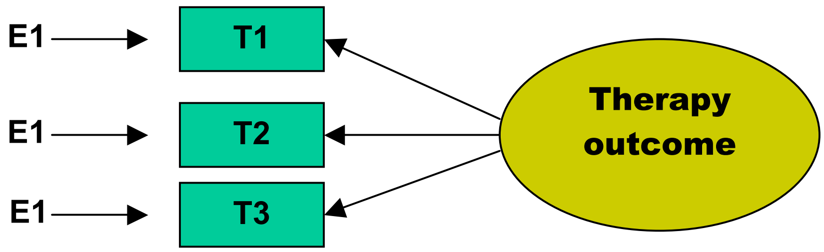


Figure 1. Three-Factor Confirmatory Model

Figure 1 illustrates the hypothesized model based on the exploratory findings. The PROC CALIS procedure (Covariance Analysis of Linear Structural Equations) estimates parameters and tests, the appropriateness of structural equation models using covariance structural analysis. CFA for the 3-factor structure was run with PROC CALIS using the following SAS code.

```
 proc calis data=tcorr corr;
    lineqs
      T1 = p1 F1 + e1,
      T2 = p2 F1 + e2,
      T3 = p3 F1 + e3;
    std
      e1-e3 = vare1-vare3,
      F1 = 1;
    var t1 t2 t3;
run;
```

Fit Statistics
CFA for Structural Equation Modeling (SEM) specifically relies on several statistical tests to determine the adequacy of model fit to the data. The chi-square test indicates the amount of difference between expected and observed covariance matrices. A chi-square value close to zero indicates little difference between the expected and observed covariance matrices. In addition, the probability level must be greater than 0.05 when chi-square is close to zero.

The Comparative Fit Index (CFI) is equal to the discrepancy function adjusted for sample size. CFI ranges from 0 to 1 with a larger value indicating better model fit. Acceptable model fit is indicated by a CFI value of 0.90 or greater (Hu & Bentler, 1999).

Root Mean Square Error of Approximation (RMSEA) is related to residual in the model. RMSEA values range from 0 to 1 with a smaller RMSEA value indicating better model fit. Acceptable model fit is indicated by an RMSEA value of 0.06 or less (Hu & Bentler, 1999).

```
The CALIS Procedure
Covariance Structure Analysis: Pattern and Initial Values

ABSGCONV convergence criterion satisfied.

         Predicted Model Matrix

            T1           T2           T3
T1       1.0000       0.7606       0.5733
T2       0.7606       1.0000       0.6373
T3       0.5733       0.6373       1.0000

Determinant      0.242493    Ln      -1.416782

The CALIS Procedure
Covariance Structure Analysis: Maximum Likelihood Estimation

Fit Function                                          0.0000
Goodness of Fit Index (GFI)                           1.0000
GFI Adjusted for Degrees of Freedom (AGFI)                 .
```

```
Root Mean Square Residual (RMR)                          0.0000
Parsimonious GFI (Mulaik, 1989)                          0.0000
Chi-Square                                               0.0000
Chi-Square DF                                                 0
Pr > Chi-Square                                          <.0001
Independence Model Chi-Square                            599.30
Independence Model Chi-Square DF                              3
RMSEA Estimate                                           0.0000
RMSEA 90% Lower Confidence Limit                              .
RMSEA 90% Upper Confidence Limit                              .
ECVI Estimate                                            0.0286
ECVI 90% Lower Confidence Limit                               .
ECVI 90% Upper Confidence Limit                               .
Probability of Close Fit                 .
Bentler's Comparative Fit Index                          1.0000
Normal Theory Reweighted LS Chi-Square                   0.0000
Akaike's Information Criterion                           0.0000
Bozdogan's (1987) CAIC                                   0.0000
Schwarz's Bayesian Criterion                             0.0000
McDonald's (1989) Centrality                             1.0000
Bentler & Bonett's (1980) Non-normed Index                    .
Bentler & Bonett's (1980) NFI                            1.0000
James, Mulaik, & Brett (1982) Parsimonious NFI           0.0000
Z-Test of Wilson & Hilferty (1931)                            .
Bollen (1986) Normed Index Rho1                               .
Bollen (1988) Non-normed Index Delta2                    1.0000
Hoelter's (1983) Critical N                                   .
```

Results indicated good model fit (chi-square = 0.000, df = 0, p < 0.0001, RMSEA = 0.000, CFI = 1.000, NNFI = 1.000). The CFA analysis has confirmed the factor structure.

SUMMARY

Dimensional Structure of the CGSQ

Based on our results, CGSQ has 3 dimensions. It contains the dimensions of objective caregiver strain, internalized subjective caregiver strain, and externalized subjective caregiver strain. There is a high degree of equivalence between the item assignments in the CGSQ and our findings for these 3 dimensions.
The objective caregiver strain dimension consists of the following 11 items: interruption of personal time; missing work or neglecting duties; child getting into trouble with the community; family member(s) having to do without things; suffering mental or physical health effects; receiving less attention; financial strain; social isolation; disruption of family routines; disruption of relationships; and disruption of social activities. The internalized subjective caregiver strain dimension consists of the following six items: feeling sad or unhappy; worrying about he family's future; worrying about the child's future; feeling guilty; feeling tired and strained; and sensing that a toll had been taken on the family. The externalized subjective caregiver strain dimension consists of four items: resentment, anger, embarrassment, and relating poorly with the child. The reliability test and confirmatory analysis supported this.

Derivation of the Informative CGSQ Summary Score

Summary scores, preferably a simple sum or mean of item responses, are likely to be an informative and efficient way to characterize the construct CGSQ. To calculate a dimension scores, the mean of all the items in a dimension is produced. It does not seem surprising that the global measure of CGSQ as a single construct is the mean of all the items in the CGSQ taken.

REFERENCES

McGrath C, Bedi R (2001). An evaluation of a new measure of oral health related quality of life—OHQoL-UK(W). Community Dent Health 18:138–143.

Norman GF, Streiner DL (2000). Biostatistics—the bare essentials. 2nd ed. Hamilton, ON: B.C. Decker.

Dawis RV (1998). Scale construction. In: Methodological issues & strategies in clinical research. Kazdin AE, editor. Washington, DC: American Psychological Association, pp. 193–213.

Gilbert GH, Duncan RP, Heft MW, Dolan TA, Vogel WB (1998). Multidimensionality of oral health in dentate adults. Med Care 36:988–1001.

Cohen J (1977). Statistical power analyses for the behavioral sciences. New York: Academic Press.

Patrick DL, Erickson P (1993). Health status and health policy—quality of life in health care evaluation and resource allocation. New York: Oxford University Press.

Scientific Advisory Committee of the Medical Outcomes Trust (2002). Assessing health status and quality-of-life instruments: attributes and review criteria. Qual Life Res 11:193–205.

Tapsoba H, Deschamps JP, Leclercq MH (2000). Factor analytic study of two questionnaires measuring oral health-related quality of life among children and adults in New Zealand, Germany and Poland. Qual Life Res 9:559–569.

WHOQOL group (1993). Study protocol for the World Health Organization project to develop a Quality of Life assessment instrument (WHOQOL). Qual Life Res 2:153–159.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the authors at:

Hui Liu, Team Leader of Statistics & Programming, Phase II – IV
David Shen, Statistician
ClinForce, LLC
2200 Renaissance Blvd.
King of Prussia, PA 19406
Work Phone: 484-322-0604
Fax: 484-322-0628