MULTIVARIATE EXPERIMENTAL CLINICAL RESEARCH

VOLUME 11 NUMBER 3 1997

A Journal for Basic Behavioral Research into Personality Dynamics and Clinical Psychology

Published Three Times a Year by the Psychology Press, from the Department of Psychology, Wichita State University, Wichita, Kansas, U.S.A.

POLICY STATEMENT

The primary aim of *Multivariate Experimental Clinical Research* is to provide a publication outlet for research in the areas covered and indicated currently by the terms personality study, clinical diagnosis and therapy, extending into the learning, social, physiological, applied and developmental aspects of these. Although due representation is given to theoretical articles which may have a methodological basis, the journal is not one of multivariate statistical methods. Although multivariate in outlook, both manipulative and non-manipulative research is accepted. In fact preference is given to dynamic, manipulative and time-sequential studies. Particular encouragement is provided for pioneer experimental attacks on what is designated personality dynamics and motivation, as well as the natural expansion thereof into structured learning theory.

Multivariate Experimental Clinical Research, Volume 11, No. 3, 1997, 109-120

A LONGITUDINAL, STRUCTURAL EQUATION ANALYSIS OF STRESS, HARDINESS, SOCIAL SUPPORT, DEPRESSION, AND ILLNESS

Jeffrey B. Brookings
Department of Psychology
Wittenberg University

Brian Bolton
Arkansas Research and Training Center
in Vocational Rehabilitation
University of Arkansas

ABSTRACT

College undergraduates (45 males, 88 females) completed measures of life stress, depression, and physical illness on two different occasions, along with measures of hardiness at Time 1 and perceived social support at Time 2. Structural equation modeling analyses provided only modest support for the hypothesis that hardiness functions prospectively to buffer the effects of stress on depression, and gave no support for the assumed mitigating effects of hardiness on the stress/physical illness relationship. Time 2 social support had the largest effect on Time 2 depression, while the best predictors of Time 2 physical illness were Time 1 physical illness (i.e., the stability coefficient) and Time 2 stress, respectively.

INTRODUCTION

Over the past 25 years, numerous studies have documented the deleterious effects of life stress on physical health and psychological well-being (for reviews, see Dohrenwend & Dohrenwend, 1978, 1981; Sarason & Sarason, 1984). Equally well-documented, however, are the relatively modest zero-order correlations between measures of stress and indexes of distress (Rabkin & Struening, 1976); that is, stressful life events typically account for no more than 9% of the variance in the dependent measures. More recently, therefore, researchers have turned their attention to a variety of personality and social characteristics — termed "resistance resources" by Antonovsky (1979) — that might buffer or moderate stress effects.

Kobasa (1979) proposed hardiness as a personality characteristic that helps protect or buffer individuals from the effects of life stress. According to Kobasa, hardy persons believe that they control events relevant to their own experience, feel committed to their activities, and view change as a stimulus or challenge to

personal growth. Studies by Kobasa and her colleagues (e.g., Kobasa, Maddi, & Kahn, 1982; Kobasa, Maddi, Puccetti, & Zola, 1985; Kobasa, Maddi, & Zola, 1983) showed that hardiness functions prospectively to buffer the effects of life stress on physical health, and Wiebe (1991) reported that hardiness moderated the effects of a stress-inducing task on affect and physiological measures. However, other reports indicate that hardiness has direct (but not stress-buffering) effects on physical illness (Wiebe & McCallum, 1986), and that hardiness predicts psychological distress, but not physical illness (Nowack, 1989).

Prompted perhaps by the inconsistent findings, several reviews of the hardiness literature have been published recently, and the reviews have been quite critical. Funk and Houston (1987) were unable to confirm the Hardiness Scale's proposed three-subscale structure, and they argued that the statistical procedures used in many hardiness studies (e.g., analysis of variance) are inappropriate because they do not allow the researcher to assess the contribution of hardiness, while controlling for the influence of other predictors with which it is typically correlated (e.g., social support). Hull, Van Treuren, and Virnelli (1987) found that the challenge component was unrelated to health outcomes, while (lack of) commitment and control had direct, rather than stress-buffering effects, and they identified several psychometric problems with the Hardiness Scale. Finally, Cohen and Edwards (1989) concluded that the evidence for the stress-buffering properties of hardiness is mixed, in part because the correlations among the scales used to measure the three hardiness components are not large enough to justify the calculation of a total score.

Typically, hardiness researchers have used prospective designs, which take prior symptom levels into account when predicting symptoms at a later date. Unfortunately, the statistical techniques used in most of these studies (e.g., analysis of variance, multiple regression, path analysis) do not allow determination of the extent to which non-significant findings are attributable to psychometric deficiencies in the Hardiness Scale (or in whatever scale the researcher uses to operationalize the hardiness construct), the absence of relationships between hardiness and well-being, or both.

A more appropriate data-analytic technique for addressing these issues is structural equation modeling (SEM). Briefly, SEM involves the specification and evaluation of: 1) a measurement model, representing relationships between measured variables (e.g., questionnaire items) and latent variables (i.e., factors); and 2) a structural model, which posits causal relations among the latent variables (i.e., constructs). The principal strength of SEM is its incorporation of latent variables represented by multiple indicators, which removes measurement error from the regression coefficients estimated in the structural model.

The purpose of this study was to investigate the effects of life stress, hardiness, and social support on psychological depression and physical illness, using a prospective design. Structural equation modeling was used to examine the relative contributions of the predictor variables.

METHOD

SUBJECTS

The sample consisted of 133 college undergraduates (45 male, 88 female) at a midwestern liberal arts college, who provided complete data on two different

occasions. Subjects participated at Time 1 as part of an introductory psychology course requirement, and participation at Time 2 was voluntary. Subjects ranged in age from 17 to 23 ($\underline{M}$ = 18.1), and 98% (130) were caucasian.

MEASURES

At Time 1, subjects completed the following self-report measures:

1. The revised Hardiness Scale (Kobasa & Maddi, 1982) consists of 36 items in a four-point, agree-disagree format. Item scores are standardized (z-scores) and summed to generate scores for the composite Hardiness Scale and three subscales: Control (16 items), Commitment (12 items), and Challenge (8 items).

2. The student version of the Life Experiences Survey (LES; Sarason, Johnson, & Siegel, 1978) lists 44 stressful life events. Subjects checked the events that they had experienced within the past 12 months and on a seven-point scale estimated the perceived positive or negative impact of each event upon them. Total stress scores were calculated by summing the impact scores of those experienced events perceived as having negative effects.

3. The Center for Epidemiological Studies Depression Scale (CES-D; Radloff, 1977) lists 20 depression-related symptoms. Subjects indicated how frequently they had experienced each symptom during the past week, ranging from "Rarely or none of the time" (0) to "Most or all of the time" (3).

4. The Seriousness of Illness Rating Scale (SIRS; Wyler, Masuda, & Holmes, 1968) is a weighted checklist measure of physical illness. Subjects checked those illnesses (from a list of 121) they had had during the past 12 months. Then, total illness scores were computed by summing the weights associated with the checked illnesses.

At Time 2, subjects again completed the LES, CES-D, and SIRS, along with the college version of the Interpersonal Support Evaluation List (ISEL; Cohen & Hoberman, 1983), which consists of 48 true-false items scored on four subscales (Tangible, Belonging, Appraisal, Self-Esteem) and a total support scale (see Brookings & Bolton, 1988). Descriptive statistics for the scales and correlations among them are shown in Tables 1 and 2, respectively.

PROCEDURE

Subjects completed the Time 1 battery in small-group sessions, and the Time 2 tests individually. Followup intervals ranged from 6 to 12 months. There were no significant differences on any of the Time 1 or Time 2 measures as a function of followup interval, and multiple regression analyses indicated that length of followup interval had no effect on the predictor/criterion relationships. Females scored higher than males on Time 1 depression, but there were no differences on Time 1 stress and hardiness, or on Time 2 stress, social support, or depression. Therefore, the structural equation modeling analyses were performed on the correlation matrix for the combined groups, using LISREL VI (Jöreskog & Sörbom, 1984).

TABLE 1
MEANS, STANDARD DEVIATIONS, AND INTERNAL CONSISTENCY RELIABILITIES FOR THE MEASURED VARIABLES

Variable	No. of items	Time 1 M	Time 1 SD	Time 1 r[a]	Time 2 M	Time 2 SD	Time 2 r[a]
1. Hardiness[b]							
Control	16	.40	6.97	.72	-	-	-
Commitment	12	.28	5.34	.65	-	-	-
Challenge	8	-.11	3.69	.49	-	-	-
Total	36	.63	10.71	.71	-	-	-
2. LES	44	8.88	6.31	-	8.01	8.95	-
3. CES-D	20	16.30	9.39	.89	14.71	9.84	.91
4. ISEL							
Tangible	12	-	-	-	10.52	1.91	.69
Belonging	12	-	-	-	9.20	2.07	.61
Appraisal	12	-	-	-	10.69	2.11	.83
Self-Esteem	12	-	-	-	8.90	1.95	.64
Total	48	-	-	-	39.32	6.10	.86
5. SIRS	121	1332.40	806.42	-	1103.85	813.65	-

Note. N = 133. LES — Life Experience Scale; CES-D — Center for Epidemiological Studies Depression Scale; ISEL — Interpersonal Support Evaluation List; SIRS — Seriousness of Illness Rating Scale.

[a] Reliabilities for the Hardiness and CES-D scales are alpha coefficients; reliabilities for the ISEL subscales and ISEL total are KR20s. Because the LES and SIRS are checklist measures, internal consistency reliabilities are not meaningful.

[b] Means and standard deviations for the Hardiness Scale and subscales are based on sums of standardized item scores.

STRUCTURAL EQUATION MODELING (SEM) ANALYSES

The first analysis consisted of an assessment of the measurement model, which specified relationships between the latent variables and their respective measured indicators. Then, for the structural model, an initial model with all paths fixed at zero was modified by adding causal paths in a pre-determined sequence. First, causal paths connected Time 1 stress and depression factors to the corresponding factors at Time 2, to assess the stability of these constructs. Next, all other Time 1/Time 2 causal paths were assessed. Finally, within-time causal paths were hypothesized, based on the time referents of the measures (see Figure 1).

TABLE 2
CORRELATIONS AMONG THE MEASURED VARIABLES AT TIMES ONE (1) AND TWO (2)

Variable	1	2	3	4	5	6	7	8
1. LES (1)	-							
2. Hardiness (1)[a]	-193	-						
3. CES-D (1)	313	-274	-					
4. SIRS (1)	270	-132	284	-				
5. LES (2)	474	-048	302	239	-			
6. ISEL (2)	-264	140	-503	-172	-332	-		
7. CES-D (2)	406	-182	541	263	564	-554	-	
8. SIRS (2)	270	-087	207	615	511	-244	391	-

Note. Decimals are omitted. Correlations > .170 are significant at .05 (two-tailed); correlations > .230 are significant at .01. N = 133.

[a] The Challenge subscale is not included in this composite.

After each analysis, non-significant paths were fixed at zero in subsequent analyses. (For a detailed description of this model-building strategy, see Aneshensel & Frerichs, 1982).

Correlations between the depression and illness measures (CES-D and SIRS) were statistically significant ($p < .05$) but modest in magnitude. Therefore, the depression and illness data were analyzed separately, and the presentation of findings will focus primarily on depression.

PRELIMINARY ANALYSES

Because the Hardiness Challenge subscale was uncorrelated with the Control and Commitment subscales, and research by Hull, et al. (1987) indicated that Challenge is unrelated to health outcomes, this subscale (8 items) was deleted from the Hardiness composite. (Deleting these items increased the internal consistency reliability of the Hardiness composite from .71 to .79.) Then, to ensure a sufficient number of measured indicators for each latent variable, while maintaining an acceptable ratio of subjects to variables: 1) the 28 Control and Commitment items were allocated to nine within-subscale miniscales or "parcels" for the SEM analyses; 2) the 20 CES-D items were allocated to four subscales (Impaired Motivation, Positive Affect, Negative Affect, Impaired Relations), based on the factor structure reported by Harlow, Newcomb, and Bentler (1986); and 3) the social support construct was represented jointly by the four ISEL subscales (inter-subscale r's ranged from .30 to .85). This resulted in a total of 23 measured varia-

bles (listed in Table 3) for the SEM analyses. Finally, the Time 1 and Time 2 stress and illness constructs were each represented by only a single indicator (LES and SIRS total scores, respectively). Therefore, following the recommendations of Herting (1985), the factor loadings and error/uniquenesses of these indicators were fixed at .90 and .19, respectively.

FIT INDEXES

Because no single index is an infallible indicator of model adequacy, the measurement model was evaluated using the following three indexes: 1) the χ^2/df ratio (values of 2 or less indicate adequate fit); 2) the Root Mean Square Residual (RMSR), reflecting the average residual difference between the model-generated and sample correlation matrices (values less than .10 suggest adequate fit); and 3) the Tucker-Lewis Index (TLI), which compares the fit of a model relative to a null model (one which assumes that the measured variables are uncorrelated). TLI values > .90 are generally regarded as indicating good fit. Of these indexes, the χ^2/df ratio is most influenced by sample size (Marsh, Balla, & McDonald, 1988).

For the structural models, fit was assessed using the χ^2/df ratio, RMSR, and a new fit index derived by Mulaik, et al. (1989), called the relative normed fit index (RNFI). The RNFI provides for the assessment of structural models — independent of the fit (or lack of fit)of the measurement model — by comparing the fit of the model relative to a "latent variable null model" that constrains correlations among the latent variables to zero. This makes it possible to evaluate assumed causal relations between hardiness and other constructs (e.g., depression, illness), while taking account of possible psychometric deficiencies in the Hardiness Scale.

RESULTS AND DISCUSSION

MEASUREMENT MODEL

For the depression data analyses, the fit indexes indicated that the overall fit of the six-factor model to the data was adequate (χ^2/df = 1.49, RMSR = .08, TLI = .88), and the loadings of all measured variables on their respective factors (see Table 3) were statistically significant ($p < .05$). The fit of the six-factor measurement model was adequate for the illness data as well (χ^2/df = 1.17, RMSR = .07, TLI = .95).

STURCTURAL MODELS

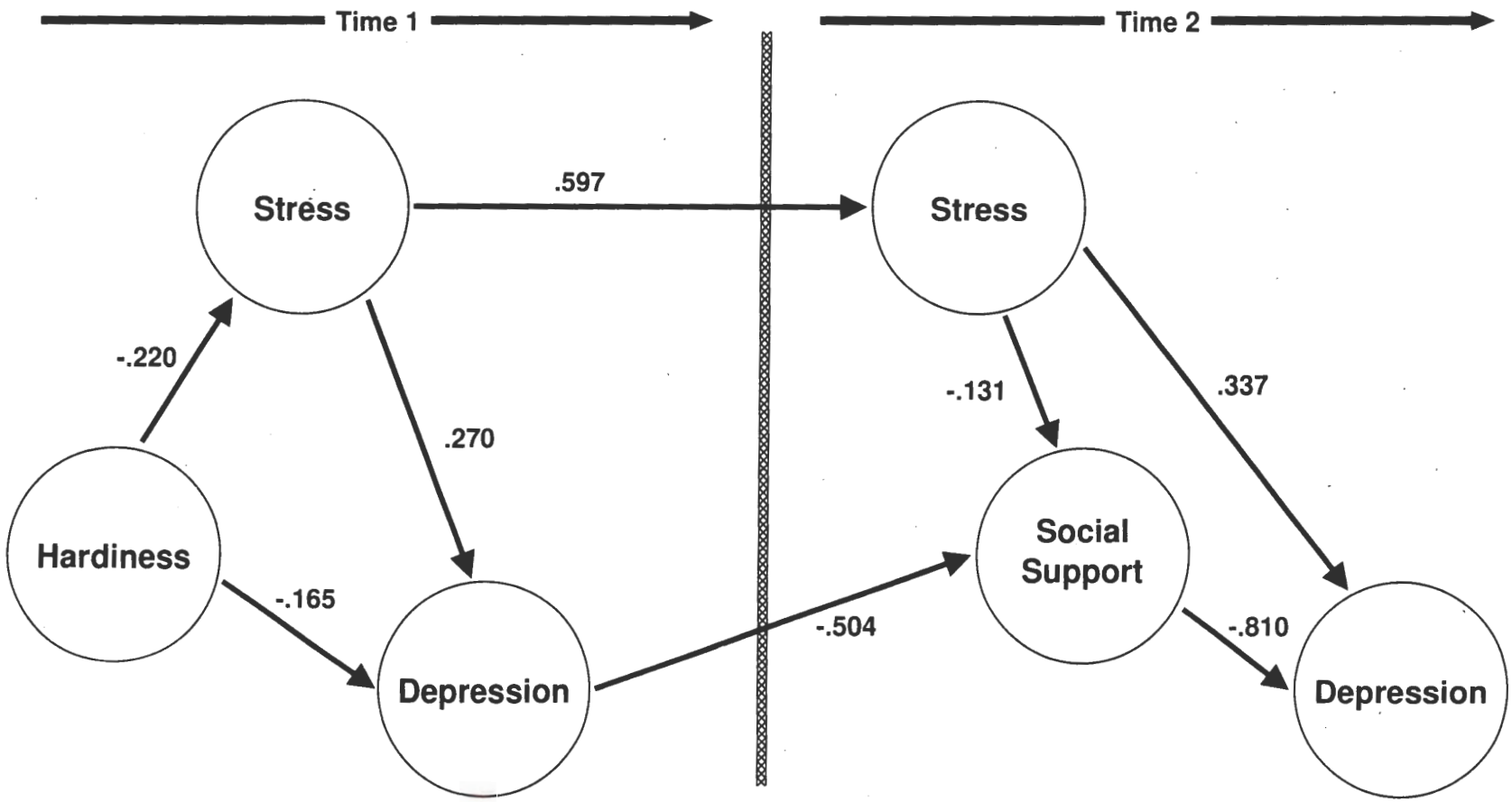
For the depression data, unstandardized regression coefficients for statistically significant ($p < .05$), direct causal effects among the latent variables are shown in Figure 1 for the best-fitting structural model (χ^2/df = 1.50, RMSR = .09, RNFI = .98). Time 1 stress and hardiness had significant effects on Time 1 depression, and the effect of hardiness on Time 1 stress was significant as well. At Time 2, stress and perceived social support had significant effects on depression, and there was a significant effect of stress on social support. The only significant Time 1/Time 2 direct effects were for Stress-1 on Stress-2 and Depression-1 on Support-2. In contrast to previous longitudinal studies, Time 1 depression had no direct effect on Time 2 depression, although it had a large indirect effect via Time 2 social support.

TABLE 3
ESTIMATED FACTOR LOADINGS AND ERROR/UNIQUENESSES FOR THE MEASUREMENT MODEL

	Factor Loadings						
	Time 1			Time 2			
Variable	Hardy	Stress	Depress	Stress	Support	Depress	Error/ Uniquenesses
1. Control_P1	.50						.75
2. Control_P2	.48						.77
3. Control_P3	.57						.68
4. Control_P4	.44						.81
5. Control_P5	.46						.79
6. Commit_P1	.51						.74
7. Commit_P2	.58						.66
8. Commit_P3	.52						.73
9. Commit_P4	.58						.66
10. LES (1)		.90					.19
11. Motivate (1)			.79				.40
12. Pos. Affect (1)			.71				.45
13. Neg. Affect (1)			.83				.42
14. Relations (1)			.76				.56
15. LES (2)				.90			.19
16. Tangible					.58		.67
17. Belonging					.75		.44
18. Appraisal					.57		.68
19. Esteem					.77		.41
20. Motivate (2)						.79	.36
21. Pos. Affect (2)						.71	.56
22. Neg. Affect (2)						.83	.25
23. Relations (2)						.76	.33

Note. Blanks indicate parameters fixed at zero. Variables 1-9 are the Hardiness Scale parcels, variables 11-14 and 20-23 are the CES-D subscales for Times 1 and 2, respectively, and variables 16-19 are the ISEL subscales. Loadings of the stress measures — LES (1) and LES (2) — on their respective factors were fixed at .90, and their error uniquenesses were fixed at .19 (see Herting, 1985). Loadings of corresponding CES-D subscales on the Time 1 and Time 2 Depression factors were constrained to equality; none of these constraints degraded significantly the fit of the model.

Figure 1. Unstandardized regression coefficients for the final structural model.



For the best-fitting structural model of the illness data ($\chi^2/df = 1.21$, RMSR = .09, RNFI = .97), the best predictors of Time 2 illness were Time 1 illness (.678) and Time 2 stress (.596). There were no statistically significant within- or across-time effects for hardiness, and no effect of Time 2 social support on Time 2 illness.

Direct, indirect, and total effects of the latent predictor variables on depression at Times 1 and 2 are summarized for the best-fitting models in Table 4. Hardiness had relatively modest total effects on depression at Times 1 and 2. Time 1 stress and depression had substantial total effects on Time 2 depression, even though neither had significant direct effects, while social support had — by far — the largest direct effect on Time 2 depression. For physical illness, the picture was simpler; the largest total effect on Time 2 illness was the direct effect of Time 1 illness (i.e., the stability coefficient), while lesser contributions were made by Time 2 and Time 1 stress.

TABLE 4
DIRECT, INDIRECT, AND TOTAL EFFECTS OF THE LATENT PREDICTOR VARIABLES ON DEPRESSION AT TIMES 1 AND 2

	Depression-1			Depression-2		
Latent Predictor Variable	Direct	Indirect	Total	Direct	Indirect	Total
Hardiness-1	-.165	-.060	-.225		-.150	-.150
Stress-1	.270		.270		.375	.375
Depression-1					.409	.409
Stress-2				.337	.106	.443
Support-2				-.810		-,810

Note. Effects were calculated from the unstandardized solution. Blanks indicate non-significant paths, which were fixed at zero in the final model.

CONCLUSIONS

The results of this study provided only modest support for the hypothesis that hardiness functions prospectively to mitigate the effects of stress on depression, and gave no support for the assumed mitigating effects of hardiness on the stress/physical illness relationship. The only significant causal paths involving hardiness were within-time effects on stress and depression at Time 1, and the total effects of hardiness on Time 2 depression (all indirect effects) were quite small. The total effects of stress and Time 1 depression on Time 2 depression

were considerably larger, in part because of indirect links via social support, while the strongest predictors of Time 2 physical illness were Time 1 physical illness and — to a slightly lesser extent — Time 2 stress.

These results do not appear to be attributable to reliability problems with the Hardiness Scale; as noted earlier, the SEM procedure estimates relationships among the latent variables, assuming error-free measurement. Rather, the findings indicate that hardiness — as operationalized by the revised Hardiness Scale — explained little of the variance in depression and illness, once the effects of other constructs (prior depression and physical illness, social support) were taken into account.

These results should be interpreted with caution, for at least three reasons. First of all, the stress (LES) and social support (ISEL) measures used in this study assess the perceived impact of stressful life events and the perceived availability of social support, respectively. Consequently, the substantial total effects of stress and social support on depression may reflect — in part — redundancy in the predictor and criterion measures (see Dohrenwend & Shrout, 1985; Monroe & Steiner, 1986). Secondly, statistical confirmation of a model does not imply rejection of all possible competing models (Cliff, 1983), and correlational data — even data obtained from longitudinal designs — are inadequate for drawing unambiguous causal conclusions (Breckler, 1990). Finally, because repeated statistical analyses on a data set raise a variety of logical and statistical issues and may not, in any event, lead to discovery of the "correct" model (see MacCallum, 1986), cross-validation of our best-fitting models (Cudeck & Browne, 1983) would strengthen the findings.

REFERENCES

Aneshensel, C. S., & Frerichs, R. R. (1982). Stress, support, and depression: A longitudinal causal model. *Journal of Community Psychology*, *10*, 363-376.

Antonovsky, A. (1979). *Health, stress, and coping*. San Francisco, CA: Jossey-Bass.

Breckler, S. J. (1990). Applications of covariance structure modeling in psychology: Cause for concern? *Psychological Bulletin*, *107*, 260-273.

Brookings, J. B., & Bolton, B. (1988). Confirmatory factor analysis of the Interpersonal Support Evaluation List. *American Journal of Community Psychology*, *16*, 137-147.

Cliff, N. (1983). Some cautions concerning the application of causal modeling methods. *Multivariate Behavioral Research*, *18*, 115-126.

Cohen, S., & Edwards, J. R. (1989). Personality characteristics as moderators of the relationship between stress and disorder. In R. W. J. Neufeld (Ed.), *Advances in the investigation of psychological stress* (pp. 235-283). New York: Wiley.

Cohen, S., & Hoberman, H. (1983). Positive events and social supports as buffers of life change stress. *Journal of Applied Social Psychology*, *13*, 99-125.

Cudeck, R., & Browne, M. W. (1983). Cross-validation of covariance structures. *Multivariate Behavioral Research*, *18*, 147-167.

Dohrenwend, B. S., & Dohrenwend, B. P. (1978). Some issues in research on stressful life events. *Journal of Nervous and Mental Diseases*, *166*, 7-15.

Dohrenwend, B. S., & Dohrenwend, B. P. (1981). *Stressful life events and their contexts*. New York: Neale Watson.

Dohrenwend, B. P., & Shrout, P. E. (1985). "Hassles" in the conceptualization and measurement of life stress variables. *American Psychologist*, *40*, 780-785.

Funk, S. C., & Houston, B. K. (1987). A critical analysis of the Hardiness Scale's validity and utility. *Journal of Personality and Social Psychology*, *53*, 572-578.

Harlow, L. L., Newcomb, M. D., & Bentler, P. M. (1986). Depression, self-derogation, substance abuse, and suicide ideation: Lack of purpose in life as a mediational factor. *Journal of Clinical Psychology*, *42*, 5-21.

Herting, J. R. (1985). Multiple indicator models using LISREL. In H. M. Blalock, Jr. (Ed.), *Causal models in the social sciences* (2nd ed.) (pp. 263-319). New York: Aldine.

Hull, J. G., Van Treuren, R. R., & Virnelli, S. (1987). Hardiness and health: A critique and alternative approach. *Journal of Personality and Social Psychology*, *53*, 518-530.

Jöreskog, K. G., & Sörbom, D. (1984). LISREL VI: *Analysis of linear structural relationships by the method of maximum likelihood*. Mooresville, IN: Scientific Software, Inc.

Kobasa, S. C. (1979). Stressful life events, personality, and health: An inquiry into hardiness. *Journal of Personality and Social Psychology*, *37*, 1-11.

Kobasa, S. C., & Maddi, S. R. (1982, November). [*The Revised Hardiness Scale*]. (Available from Suzanne Kobasa, Department of Psychology, CUNY Graduate School & University Center, 33 W. 42nd St., New York, NY 10036)

Kobasa, S. C., Maddi, S. R., & Kahn, S. (1982). Hardiness and health: A prospective study. *Journal of Personality and Social Psychology*, *42*, 168-177.

Kobasa, S. C. O., Maddi, S. R., Puccetti, M. C., & Zola, M. A. (1985). Effectiveness of hardiness, exercise and social support as resources against illness. *Journal of Psychosomatic Research*, *29*, 525-533.

Kobasa, S. C., Maddi, S. R., & Zola, M. A. (1983). Type A and hardiness. *Journal of Behavioral Medicine*, *6*, 41-51.

MacCallum, R. (1986). Specification searches in covariance structure modeling. *Psychological Bulletin*, *100*, 107-120.

Marsh, H. W. Balla, J. R., & McDonald, R. P. (1988). Goodness-of-fit indexes in confirmatory factor analysis: The effect of sample size. *Psychological Bulletin*, *103*, 391-410.

Monroe, S. M., & Steiner, S. C. (1986). Social support and psychopathology: Interrelations with preexisting disorder, stress, and personality. *Journal of Abnormal Psychology*, *95*, 29-39.

Mulaik, S. A., James, L. R., Van Alstine, J., Bennett, N., Lind, S., & Stilwell, C. D. (1989). Evaluation of goodness-of-fit indices for structural equation models. *Psychological Bulletin*, *105*, 430-445.

Nowack, K. M. (1989). Coping style, cognitive hardiness, and health status. *Journal of Behavioral Medicine*, *12*, 145-158.

Rabkin, J., & Struening, E. (1976). Life events, stress, and illness. *Science*, *194*, 1013-1020.

Radloff, L. S. (1977). The CES-D scale: A self-report depression scale for research in the general population. *Applied Psychological Measurement*, *1*, 385-401.

Sarason, I. G., Johnson, J. H., & Siegel, J. M. (1978). Assessing the impact of life changes: Development of the Life Experiences Survey. *Journal of Consulting and Clinical Psychology*, *46*, 932-946.

Sarason, I. G., & Sarason, B. R. (1984). Life changes, moderators of stress, and health. In A. Baum, S. E. Taylor, & J. E. Singer (Eds.), *Handbook of psychology and health: Social psychological aspects of health* (Vol. 4) (pp. 279-299). Hillsdale, NJ: Erlbaum.

Wiebe, D. J. (1991). Hardiness and stress moderation: A test of proposed mechanisms. *Journal of Personality and Social Psychology*, *60*, 89-99.

Wiebe, D. J., & McCallum, D. M. (1986). Health practices and hardiness as mediators in the stress-illness relationship. *Health Psychology*, *5*, 425-438.

Wyler, A. R., Masuda, M., & Holmes, T. H. (1968). Seriousness of Illness Rating Scale. *Journal of Psychosomatic Research*, *11*, 363-375.

Author Notes

This research was supported by a Faculty Research Grant from Wittenberg University to the first author.

An earlier version of this article was presented at the American Psychological Association meeting in Boston, MA, on August 13, 1990.

Address requests for reprints and other correspondence to Jeffrey B. Brookings, Psychology Department, Wittenberg University, P.O. Box 720, Springfield, OH 45501. email: jbrook@wittenberg.edu.

Multivariate Experimental Clinical Research, Volume 11, No. 3, 1997, 121-130.

CONFIRMATION OF THE FACTOR PATTERN OF A BATTERY OF INTELLIGENCE AND ACHIEVEMENT INDICATORS FOR A CHILD CLINICAL POPULATION

Ralph Mason Dreger
Louisiana State University

ABSTRACT

An exploratory factor analysis of the data for 50 clinical children, from a pool of 210, on a battery of tests required of clients in a private practice clinic, including four intelligence indicators, three formal indicators of achievement, and Factor II from the Children's Behavioral Classification Project Inventory, revealed a factor pattern loading intelligence indicators on one factor and achievement indicators on another. This result was confirmed in an independent sample of 46 children by confirmatory factor analysis. As assessed by chi-square and specific indices of the EQS program, the empirically-derived covariance matrix and hypothesized covariance matrix were found to be essentially congruent. In addition, traditional matrix comparison coefficients supported the relative equality of the two factor matrices.

INTRODUCTION

A previous article, "A research-oriented private psychological clinic" (Dreger, 1990), reported among other findings one relating Factor II of the Children's Behavioral Classification Project (CBCP) Inventory (described below) to measured intelligence and scholastic achievement. Factor II is labeled among the CBCP factors, "Intellectual and Scholastic Retardation vs. Alert, Socialized Scholastic Achievement." According to current research, the factor is mislabeled.

The *generally* positive relation among intelligence tests is well known (Sattler, 1988).[1] However, that a similar relation holds between parental responses to a *behavioral* Inventory from which Factor II is derived and children's responses to standardized intelligence and achievement measures has yet to be demonstrated. It is true (Dreger, 1990), to be sure, that scores on Factor II derived from the CBCP Inventory predict almost as well to the Reading subtest of the Wide Range Achievement Test as does the Full Scale IQ of the child Wechsler tests.

This latter finding suggested that Factor II is related more closely to the achievement tests than to intelligence measures. In the exploratory factoring of the test battery (also described below), two factorial dimensions were distinctly identified as "Intelligence" and "Achievement." Factor II was much better correlated with the second factor than with the first one. Thus, the factor pattern

which emerged from the analysis consisted of two factors, each primarily identified by four tests in the battery, four intelligence indicators and four achievement indicators.

The purpose of the present investigation was then to attempt to confirm the factor pattern in independent samples of the child clinical population. To ascertain, first, in one sample the factor pattern underlying this battery, i.e., its "causes" in latent variable analysis terms, and second, to confirm its existence in an independent sample are thus important and necessary tasks.

METHOD

PARTICIPANTS

Parents who applied for service in PRS signed informed consent forms for their children and they themselves responded to the research instruments required by the clinic, the author's private practice, including the CBCP Inventory describing their children's behaviors.

An original sample of 210 clinical children in a private practice clinic was reduced to 96 by eliminating all those whose protocols were incomplete, i.e., lacking even one of the tests required for research and service purposes. Even under the strict requirements of the clinic for administration of all tests to all children, for various reasons some children did not have all tests. Use of "missing values" routines was eschewed in favor of completeness, lowering, to be sure, the degrees of freedom available but guaranteeing no artifactual increase in the degrees of freedom. The children ranged in age from 4 to 13 and were almost exclusively Caucasian. The gender ratio was approximately 3 to 1, male to female, a common distribution for psychological clinics.

The 96-participant sample was divided into subsamples of 50 and 46 respectively and analyses carried out on each sample, exploratory factor analysis with the first and confirmatory analysis with the second.[2] These subsamples were deemed comparable inasmuch as the only selective factor was temporal, the first set taken from earlier cases and the second from later cases. There is, however, a significant difference between their mean ages, CA 10.58 and 9.52 respectively, though the ranges are the same for both. What effect the age differential might have on the factor patterns of the two groups is unknown. The proportions of male to female subjects are almost identical in the subsamples: .7 to .3.

INSTRUMENTS

In the author's private practice, purposely named Psychological Research and Services, a minimal test battery was required, in addition to whatever instruments the clinician needed for specific individuals. This is the multimethod approach to psychological evaluation, endorsed appropriately by Sattler (1988) among others, which the author has termed the "principle of redundant validity" (Dreger, 1993). According to this principle, regardless of what incremental validity is contributed by any particular test *for groups*, it is necessary *for individuals* and especially clinical individuals to have more than one indicator of a characteristic to make important decisions about such individuals. Accordingly, several both direct and indirect assessments of intellectual functioning were built into the

battery, as well as, for children, several tests of academic achievement. Included in the battery, therefore, were the following tests:

WISC, *WISC-R*, or *WPPSI* and *Goodenough-Harris Drawing Test (GH)*. These are so well known, no need exists to describe them here.

Early School Personality Questionnaire or *Children's Personality Questionnaire* (Cattell, 1973). Whatever validity may or may not be found in these tests as a whole (e.g., Gough, 1978; Sines, 1978; Thorndike, 1978), *Factor B (ESPQ B)* which purports to measure Spearman's g or general intelligence factor does correlate significantly with the three other intelligence tests in the battery, its highest r being .53 with the Wechsler tests.

Full-Range Picture Vocabulary Test (Ammons & Ammons, 1948). This test has a four-choice picture mode comparable in part to the better-known Peabody. According to Sattler (1988), it has satisfactory reliabilities and correlates well with more extensive intelligence scales.

Wide Range Achievement Test (WRATR, WRATS, WRATA) (1965 revision) (Jastak, Jastak, & Bijou, 1965). This test in the versions used for this study has received quite negative reviews (e.g., Merwin, 1972; Thorndike, 1972), although Thorndike states that the test may have some validity in clinical and research settings with individual administration. This latter consideration, as well as the fact that the WRAT does tap important areas of educational attainment, in subtests of *Reading*, *Spelling*, and *Arithmetic*, and the brevity of the scales prompted the team of researchers/clinicians to choose the instrument for the battery. Although the reported split-half reliabilities verge on the unreal (.98 being the lowest r), more realistic coefficients using the two levels of the test for reliability purposes (.90 for Reading and Spelling and .85 for Arithmetic) give some assurance of internal consistency. Merwin (1972) questions whether the test distinguishes between intelligence and achievement; the factor pattern of the present study tends moderately to support the implied criticism.[3] Despite its shortcomings, the WRAT was included in the required test battery, though often supplemented in clinical practice by full-length achievement tests.

Children's Behavioral Classification Project Inventory (CBCP. Full descriptions of the Children's Behavioral Classification Project and its Inventory, from which *Factor II* was derived, are found in Dreger (1977, 1981, 1996) and Dreger and Dreger (1962). Briefly, the Inventory is a 274-item behavioral questionnaire to which parents or parents-surrogate respond "True" or "False" as to whether the behaviors have been observed in the past six months. The behavioral items were developed by an interdisciplinary team of clinicians and the research design and instrument were thoroughly critiqued and refined by a much broader interdisciplinary Technical Assistance Project sponsored by NIMH. The guiding principle for item creators was "What is it that the child does which prompts someone to say, 'He has a problem'?" Numerous reliability and validity studies have been carried out with the CBCP Inventory with samples from special populations, like "minimal brain dysfunction," blind, psychotic, brain damaged, EMR, sickle cell anemic, diabetes mellitus; in Spanish translation it has been standardized in Venezuela. Factor analyses with samples ranging from 341 to 1379 have been accomplished, with comparable factor composition from study to study.

Factor II (FACTII) of the Inventory is one of 30 derived by both principal factor and principal component analyses. It has 14 scored items and 22 other significantly-weighted identifying items. Scored items are usually those with the highest structure weights in variable rows; no item is scored on more than one factor, in order to avoid index correlation. Items particularly apropos to *Factor II* may be represented by such as "Makes failing grades in arithmetic, makes many mistakes with numbers, or says he does not like arithmetic," "Spells poorly," and "Reads poorly." Although undoubtedly influenced by the child's reports of school achievement or lack thereof, *Factor II*, derived from parental reports, is an assessment of scholastic achievement independent of the test data derived from the child.

PROCEDURE AND ANALYSES

Principal factor analysis using the SAS FACTOR procedure was employed with the 50-participant raw score matrix consisting of scores from the eight tests. Varimax and Promax rotations were applied to the two-factor solution dictated by Scree and eigenvalues greater than one. (A third latent root of slightly over 1.00 was tentatively ignored.) The rotated factor structure was examined first. Weights for the several intelligence tests tended to be higher on the first factor; and for the several achievement tests, including *Factor II*, they were higher on the second factor, but substantial loadings were found on the alternate factor in each case. However, both the reference vector structure and factor pattern clearly distinguished between the two dimensions: significant weights on the first factor for intelligence tests and on the second factor for achievement tests, with insignificant weights on the alternate factor, except for *WRATA* (Arithmetic subtest) which had a statistically significant but practically negligible weight on the first factor. Consequently, the factor pattern was employed for prediction in the confirmatory factor analysis. This choice was dictated, however, not only by the clarity of pattern but also by the fact that structural equation modeling parallels the factor pattern consisting of standardized regression weights.

Based on the two-factor pattern,[4] then, equations were written for carrying out confirmatory analysis on the data of the 46-participant sample's data, by means of LISREL (Joreskog & Sorbom, 1989) and EQS (Bentler, 1989), both of which default to the use of a covariance matrix.

The measurement model was that described in LISREL,

$$x = \Lambda_x \xi + \delta$$

where x is a vector of test scores, usually standardized, Λ_x is a matrix of factor weights, ξ is a vector of latent factors, and δ is a vector of unique elements consisting of specific and error variance. As explained in the Footnotes, LISREL could not complete the analyses.[5] Accordingly, analyses proceeded with EQS. EQS does give appropriate warnings, for example, that a matrix may not be positive definite, but ordinarily continues the analyses after suggesting that the results may not be trustworthy. Since EQS provides start values for parameters to be estimated, the user-supplied estimates for the lambda matrix were not included in the equations. Values for the theta delta diagonal were set to .7 and for the phi

TABLE 1
TESTS OF TWO HYPOTHESIZED LATENT VARIABLES AS ASSESSED BY EIGHT MANIFEST VARIABLES

Statistic or Test	1st Sample (n = 50)	2nd Sample (n = 46)
Determinant of covariance matrix	.116924D+15	.92560D+14
Residuals (S - Sigma)		
Average, absolute standardized	.0357	.0585
Average, off-diagonal absolute standardized	.0459	.0753
Range of distribution	5 - 8	5 - 8
Goodness of fit		
Chi-square (17 df)	18.819	28.772
p	.34	.04
Normal theory reweighted LS chi-square	18.095	23.438
Satorra-Bentley scaled chi-square	16.837	25.256
p	.47	.09
Normed Fit Index	.915	.785
Non-normed Fit Index	.985	.816
Comparative Fit Index	.991	.889

5 = -.2 to -.1; 6 = -.1 to .0; 7 = .0 to .1; 8 = .1 to .2.

matrix at .6. The maximum likelihood method with the ROBUST backup for possible non-normal distributions was ordered.

For making direct comparisons of factor patterns from one set of data to the other, both of the data sets were subjected to EQS, even though the exploratory analysis had originally been done with SAS. It would have been desirable to utilize LISREL, since it has some advantages over EQS, especially in providing squared multiple correlations, a plot of standardized residuals, and modification indices for changing parameter estimates (as well as the lambda matrix in its familiar factor form!). Nevertheless, EQS, too, has its advantages, including especially more easily allowing equality constraints. Thus, for the test of equality of factor patterns in the two data sets, equality constraints were imposed, and tested by the Lagrange multiplier test, corresponding to the hypothesized model of two factors measured by four variables apiece, i.e., WISC, GH, ESPQ B, and FRPVT for the first factor and FACTII, WRATR, WRATS, and WRATA for the second factor.

In addition to the structural equation modeling, more traditional methods of matrix comparison were employed: the Ahmavaara transformation (Ahmavaara, 1954), the congruence coefficient (Burt, 1948), and the Salient Variable Similarity Index (Cattell, 1949).

TABLE 2
FACTOR PATTERNS OF INDEPENDENT DATA SETS FROM EQS ANALYSES

Variable	1st Sample			2nd Sample		
	Fact 1	Fact 2	Uniqueness	Fact 1	Fact 2	Uniqueness
Factor II	.000	-.541	.841	.000	-.481	.876
WISC, WISC-R, WPPSI	.957	.000	.290	.826	.000	.564
GH	.673	.000	.740	.727	.000	.687
ESPQ B	.559	.000	.829	.581	.000	.814
FRPVT	.802	.000	.598	.473	.000	.881
WRATR	.000	.911	.411	.000	.811	.584
WRATS	.000	.915	.402	.000	.781	.624
WRATA	.000	.645	.764	.000	.618	.787

RESULTS

Table 1 reveals the results of testing the hypothesis of a two-factor solution applied to the second data set. (The comparable EQS analysis of the first set is included in Table 1 for a direct comparison.) Both the average of absolute standardized residuals as a whole and of the same for off-diagonal elements are satisfactorily small, indicating a close fit between obtained and model covariance matrices. These figures fairly represent the entire distributions of residuals, in that the largest absolute standardized residual is .180 (compared in the first sample to .164). Likewise, the non-significant chi-squares suggest that there is a fairly close approximation between theory and empirical reality. On the other hand, the three fit indices (above and beyond the chi-squares which in this usage constitute goodness-of-fit indices) give reason to conclude that the model fits less well than it ought to. None of the coefficients, though, indicates a truly "bad" fit. Table 1 does not include figures on the univariate and multivariate skewnesses and kurtoses, none of which, however, obviate the results given in the table.

Table 2 presents the results of the modeling in terms of factor matrices in more familiar form; these represent the standardized measurement equations. As postulated, the zero entries imply negligible elements; in actuality, in exploratory factor analyses these elements may achieve statistical significance and can be used in regression equations to predict scores more accurately on the original variables. But from a structural standpoint (and in Cattell's, 1978, and Gorsuch's, 1983, all-or-nothing weighting), these elements constitute "noise." Even at this point, though, before formally testing comparability of patterns, it is evident that they are reasonably similar. It must be emphasized that these similarities did not arise from inserting initial start values in the EQS equations, for no such values were given; they came entirely from the data themselves.

TABLE 3
EQUALITY OF FIRST AND SECOND FACTOR PATTERNS DETERMINED BY JOINT MODELLING AND COMPARISON COEFFICIENTS

Statistic or Test	1st Sample (n = 50)	2nd Sample (n = 46)
Average, absolute standardized residual	.0553	.0962
Average, off-diagonal absolute standardized residual)	.0596	.1060

Constraints, to test equality hypothesis 1st Sample		2nd Sample		Chi-square	Probability
FACTOR II, F2	-	FACTOR II, F2	= 0	.016	.900
WISC, F1	-	WISC, F1	= 0	.532	.466
GH, F1	-	GH, F1	= 0	1.049	.306
ESPQ B, F1	-	ESPQ B, F1	= 0	1.505	.220
FRPVT, F1	-	FRVPT, F1	= 0	1.226	.268
WRATR, F2	-	WRATR, F2	= 0	.332	.565
WRATS, F2	-	WRATS, F2	= 0	.731	.392
WRATA, F2	-	WRATA, F2	= 0	.242	.623

Comparison Method	Coefficient
Congruence Coefficient	.9885
Ahmavaara Transformation	1.0000
Salient Variable Similarity Index	.6957

Table 3 brings together the tests of equality of factor patterns derived from the separate and wholly independent data sets. Again, it is evident that the model fits the first set better than it does the second, as seen in the average absolute standardized residuals in the upper part of the table. In the second set, although values clustering about 0.1 are not large by the usual criteria, especially in consideration of the fact that the largest absolute standardized residual in either matrix is only .238, a closer fit in the second data would make for a closer parallel between the two sets. In the middle portion of the table, the Lagrange method tests whether the actual values obtained for the relation between the manifest and latent variables are the same for both data sets. From the small chi-squares and associated p values, it can be seen that the differences are statistically non-significant. At

the bottom of Table 3, the more traditional comparison coefficients provide evidence that the two pattern matrices are essentially quite similar.

DISCUSSION

By several different tests the comparability of the two sets of factor patterns, each derived from an independent sample of clinical children, has been demonstrated. Thus, the purpose of this study, to confirm the factor pattern of the two-factor solution by revealing it in independent samples, appears to have been accomplished.

There are several additional comments called for. In Table 2, it might seem as if no factor weights would be found on the alternate factor when the variable is more strongly associated with one factor. No doubt there could be some "significant" weights found on alternate factors if the predictive equations included elements for the alternate. The fact, however, that the exploratory factor analysis so clearly delineated a pattern in which no significant weights were found on the alternate factor virtually required the equations to be set up with only one factor apiece (plus an error term in each instance).

It is both interesting and important to note that both the Lagrange test and the traditional comparison methods (Table 3) yield the same information; examination of Table 2 provides essentially that same information in different form.

Finally, attention should be paid to the Footnote in which the significance of this research is stressed: Revealing the underlying "causes" of the battery of well-standardized tests appears to be an important accomplishment.

REFERENCES

Ahmavaara, Y. (1954). Transformation analysis of factorial data and other new analytical methods of differential psychology with their application to Thurstone's basic studies. *Annales Academiae Scientiarum Fennicae, Sarja-Ser. B nide-tom, 88*, 2.

Ammons, R. B., & Ammons, H. S. (1948). *Full-Range Picture Vocabulary Test.* Missoula, Montana: Psychological Test Specialists.

Bentler, P. M. (1989). *EQS Structural equations program manual.* Los Angeles: BMDP Statistical Software.

Bentler, P. M. (1993). *EQS Structural equations program manual.* Los Angeles: BMDP Statistical Software.

Bollen, K. A. (1989). *Structural equations with latent variables.* New York: Wiley.

Burt, C. (1948). The factorial study of temperamental traits. *British Journal of Psychology, Statistical Section, 1*, 178-203.

Cattell, R. B. (1949). A note on factor invariance and the identification of factors. *British Journal of Psychology, 2*, 134-138.

Cattell, R. B. (1973). *Personality and Mood by Questionnaire*. San Francisco: Jossey-Bass.

Cattell, R. B. (1978). *The scientific use of factor analysis in behavioral and life sciences*. New York: Plenum.

Dreger, R. M. (1953). Different IQ's for the same individual associated with different intelligence tests. *Science, 118*, 594-595.

Dreger, R. M. (1981). The classification of children and their emotional problems. *Clinical Psychology Review, 1*, 415-430.

Dreger, R. M. (1991). The latent structure of the Salvation Opinionnaire, a measure of religious attitudes in the American Christian tradition. *Educational and Psychological Measurement, 51*, 707-719.

Dreger, R. M. (1993). Methodology. In J. L. Matson (Ed.), *Handbook of hyperactivity in children*. Boston: Allyn & Bacon.

Dreger, R. M. (1996). Background for and development of a brief Children's Behavioral Classification Project Inventory. *Psychological Reports, 79*, 963-977.

Dreger, R. M., & Dreger, G. E. (1962). *Proceedings of the Technical Assistance Project*. Behavioral Classification Project Report No. 1. Jacksonville, Florida: Behavioral Classification Project, Jacksonville University.

Gough, H. F. (1978). Review of *Children's Personality Questionnaire, 1975 Edition*. In O. K. Buros (Ed.), *Eighth mental measurements yearbook*. Lincoln, Nebraska: University of Nebraska Press.

Guilford, J. P. (1964). Zero intercorrelations among tests of intellectual abilities. *Psychological Bulletin, 61*, 401-404.

Hayduk, L. A. (1987). *Structural equation modelling with LISREL: Essentials and Advances*. Baltimore and London: Johns Hopkins University Press.

Jastak, J., Jastak, S. R., & Bijou, J. W. (1965). *Wide-Range Achievement Test*. (Rev. ed.) New York: Psychological Corporation.

Joreskog, K. G., & Sorbom, D. (1989). *LISREL 7 user's reference guide*. Mooresville, Indiana: Scientific Software.

Merwin, J. C. (1972). Review of *Wide Range Achievement Test*. In O. K. Buros (Ed.), *Seventh mental measurements yearbook*. Lincoln, Nebraska: University of Nebraska Press.

Messick, S. (1988). Validity. In R. L. Linn (Ed.), *Educational Measurement*. 3rd ed. New York: Macmillan.

Sattler, J. M. (1988). *Assessment of children*. 3rd ed. San Diego: Jerome M. Sattler, Publisher.

Sines, J. O. (1978). Review of *Early School Personality Questionnaire*. In O. K. Buros (Ed.), *Eighth mental measurements yearbook*. Lincoln, Nebraska: University of Nebraska Press.

Thorndike, R. L. (1972). Review of *Wide Range Achievement Test*. In O. K. Buros (Ed.), *Eighth mental measurements yearbook*. Lincoln, Nebraska: University of Nebraska Press.

Thorndike, R. L. (1978). Review of *Early School Personality Questionnaire*. In O. K. Buros (Ed.), *Eighth mental measurements yearbook*. Lincoln, Nebraska: University of Nebraska Press.

Wothke, W. (1992). Nonpositive definite matrices in structural modeling. In K. A. Bollen and S. Long (Eds.), *Testing structural models*. Beverly Hills, California: Sage.

Footnotes

1. Guilford (1964) reported that he had found over 7,000 correlation coefficients in the literature between and among intelligence tests, almost one-fourth of which were not significantly different from zero.
2. Forty-six subjects is lower than Bollen (1989) advocates as a minimum, i.e., 50. He states that he knows no hard and fast rule, but suggests several cases per each free parameter. In the present instance, there are 20 free parameters to be estimated. Consequently, it must be recognized that the chi- squares generated may be too large. However, Hayduk (1987) reports that he has run factor models with as few as 22 cases without discernible problems. Since the solutions in the present instance converged well within the EQS' specified limit of 30 iterations and the goodness-of-fit indexes are fairly satisfactory, the presumption is that the case is similar to Hayduk's.
3. Only for the Arithmetic subtest, however, which distributes its variance between the two factors in the factor structure with a lesser but still significant weight on the intelligence factor.
4. A three-factor model was also tested, utlizing EQS. Although the Bentler fit indices were larger than those for the two factors, the other indices of fit were far inferior to those seen in Table 1. E.g., the average standardized absolute residuals, both including and excluding diagonals, are .3330 and .3033 respectively, as compared to the very small average residuals for two factors. Eigenvalues for the first data set ($\underline{n}$ = 50) are 3.70, 1.37, 1.02, 0.64, 0.56, 0.35, 0.25, and 0.12; for the second data set ($\underline{n}$ = 46), they are 3.44, 1.46, 0.93, 0.79, 0.58, 0.34, 0.26, and 0.21.
5. That LISREL could not produce results stems in part from its sensitivity to non-positive definite matrices and underidentification of parameters. Both Joreskog and Sorbom (1989) and Wothke (1992) suggest ways to overcome lack of positive definiteness in LISREL. Several of these ways were attempted but without success. On the other hand, though EQS does give warnings about lack of identification and non-positiveness, it ordinarily continues the analyses after suggesting that the results may not be trustworthy.

Multivariate Experimental Clinical Research, Volume 11, No. 3, 1997, 131-144.

FACTORS OF THE BECK HOPELESSNESS SCALE: FACT OR ARTIFACT?

Robert A. Steer
Department of Psychiatry
University of Medicine and Dentistry of New Jersey
School of Osteopathic Medicine
Stratford, NJ

Aaron T. Beck
Gregory K. Brown
Department of Psychiatry
University of Pennsylvania Medical School
Philadelphia, PA

ABSTRACT

Weighted least-squares factor analyses of the tetrachoric correlations, principal components analyses of the G-index correlations, and principal components analyses of the phi correlations among the 20 items of the Beck Hopelessness Scale were conducted with 1,126 outpatients who were diagnosed with primary mood disorders and 732 who were diagnosed with primary anxiety disorders. Two dimensions reflecting *pessimism* about the future and *resignation* to the futility of changing the future were found in both diagnostic groups, regardless of factor analytic method or type of data transformation.

INTRODUCTION

The Beck Hopelessness Scale (BHS; Beck, Weissman, Lester, & Trexler, 1974; Beck & Steer, 1993) is a self-report instrument that is frequently employed to measure hopelessness or pessimism about the future and has consistently been reported to be positively related to suicidal behavior and ideation (Beck, 1986). The BHS was specifically developed by Beck et al. (1974) to measure depressed patients' negative views about the future, and pessimism about the future is one of the constructs contained in the negative cognitive triad that Beck (1967, 1976) proposed in his cognitive model of depression. The other two constructs stress negative views of oneself and the world. However, psychiatrists rank hopelessness as the most important risk factor for judging suicidal risk (Truant, O'Reilly, & Donaldson, 1991).

In the original principal-components analysis of the BHS with 294 adult inpatient-suicide attempters, Beck et al. (1974) identified three dimensions which they

called "Feelings about the Future," "Loss of Motivation," and "Future Expectations," but there is controversy about whether the BHS reflects more than one dimension. Nekanda-Trepka, Bishop, and Blackburn (1983) conducted a principal components analysis with 86 depressed outpatients and found five components, whereas Mendonca, Holden, Mazmanian, and Dolan (1983) concluded that a one component solution was preferable to a three component solution with 78 patients presenting to a psychiatric crisis center. More recently, Hill, Gallagher, Thompson, and Ishida (1988) extracted three principal components for 120 older psychiatric outpatients and reported that their components were comparable to those found by Beck et al. (1974). Hill et al. (1988) labeled their components, "Hopelessness about the Future," "Giving Up," and "Future Anticipation."

A principal components analysis of the BHS was also conducted by Steer, Kumar, and Beck (1993) with 108 adolescent inpatients between 12 and 17 years old who were diagnosed with mixed psychiatric disorders. They found three components whose item compositions were comparable to those previously reported by both Beck et al. (1974) and Hill et al. (1988) for adults and called them, "Rejection of the Possibility of a Hopeful Future (Rejection)," "Acceptance of the Inevitability of a Hopeless Future (Acceptance)," and "Resignation to the Futility of Changing the Future (Resignation)."

Finally, Steer, Iguchi, and Platt (1994) administered the BHS to 2,379 intravenous (IV) drug users who were not currently enrolled in a treatment program and seeking HIV testing and counseling, and the Resignation, Rejection, and Acceptance components again emerged. Furthermore, the compositions of these three components were comparable to those originally described by Beck et al. (1974), and the three components displayed different patterns of relationships with the IV drug users' background and clinical characteristics. For example, the Rejection dimension was positively related to testing HIV positive, whereas the other two dimensions were not (Steer et al., 1994).

Although Beck and Steer (1993) suggested that the number of BHS dimensions was dependent upon the diagnostic compositions of samples being studied, we observed in the six principal components studies that were described above that three dimensions have consistently emerged across the samples which contain the same sets of marker items. There was a component that had high loadings for the following three items: item #9 ("I just don't get the breaks, there's no reason to believe I will in the future."), item #16 ("I never get what I want so it is foolish to want anything."), and item #20 ("There's no use in really trying to get something I want because I probably won't get it."). A second component was composed of items #1 ("I look forward to the future with hope and enthusiasm.") and item #15 ("I have great faith in the future."), and a third component contained items #4 ("I can't imagine what my life would be like in 10 years.") and item #18 ("The future seems vague and uncertain to me."). The three marker items representing the first dimension were found in the "Loss of Motivation" component that was identified by Beck et al. (1974), in the "Giving Up" component by Hill et al. (1988), and in the "Resignation to the Futility of Changing the Future" component by Steer et al. (1993, 1994). The items in the two latter sets reflected, respectively, the Rejection and Acceptance components that Steer et al. (1993, 1994) had discussed as being sensitive to the positive and negative wording of the BHS items.

Approaching the question about the dimensionality of the BHS from the perspective of item-response theory (Lord & Novick, 1956), Young, Halper, Clark, Scheftner, and Fawcett (1992) conducted full-information factor analyses (Bock, Gibbons, & Muraki, 1988) with the tetrachoric correlations among the 20 BHS responses of 730 adult outpatients who volunteered for the National Institute of Mental Health (NIMH) Collaborative Study of the Psychobiology of Depression and a mixed sample of 257 patients and normal adults. Although they reported that there was statistical support for the existence of more than one maximum-likelihood factor, Young et al. (1992) also found that none of the additional factors explained more than 6% of the total variance in either sample. The first factors, respectively, explained 53.0% and 55.5% of the NIMH and mixed samples' total variances. Young et al. (1988, p. 585) concluded that the BHS constituted a single dimension of hopelessness and stated, "More complex multidimensional models, while demonstrating some statistical superiority, offered no additional explanatory or conceptual advantage."

The purpose of the present study is to compare the factor structures of the BHS in outpatients diagnosed with either primary mood disorders or primary anxiety disorders to determine whether different methods of factor analysis and data transformations yield comparable dimensions of hopelessness for these two diagnostic groups. We chose these two broad diagnostic groups because Beck and Steer (1993) had found that outpatients diagnosed with primary mood disorders had higher mean BHS scores than those diagnosed with primary anxiety disorders, and we wished to compare the factor structures of the BHS for patients with different mean levels of hopelessness. Young et al. (1992, p. 585) cautioned that the BHS was "relatively insensitive in measuring lower levels of hopelessness", and we wanted to ascertain whether the BHS would yield fewer factors in patients with less severe hopelessness.

METHOD

SAMPLES

For the purposes of the present analyses, we selected 1,126 (52.0%) outpatients diagnosed with primary mood disorders and 732 (33.8%) outpatients diagnosed with primary anxiety disorders from a total sample of 2,165 outpatients who were evaluated at the Philadelphia Center for Cognitive Therapy (CCT) between January, 1986, and October, 1992, for whom complete BHS data were available.

All of the outpatients were diagnosed with the Structured Clinical Interview for DSM-III-R (SCID; Spitzer, Williams, & Gibbon, 1987). Only primary DSM-III-R (American Psychiatric Association, 1987) Axis I diagnoses were focused upon here because prior research comparing patients by Axis I primary, secondary, tertiary, and Axis II (personality) disorders had found that the primary diagnostic group was the most important classification for differentiating among outpatients with respect to severity of hopelessness (Beck & Steer, 1993). A detailed description of the SCID training procedures used at the CCT and of the interjudge reliabilities that have been obtained with these procedures are presented by Riskind, Beck, Berchick, Brown, and Steer (1987). However, no interjudge agreement study was conducted with respect to the present diagnoses.

MOOD DISORDERS

The outpatients with mood disorders were composed of 478 (42.5%) men and 648 (57.5%) women. There were 1,055 (93.7%) Whites, 42 (3.9%) Blacks, and 29 (2.6%) Asians. The mean age was 37.29 (SD = 11.80) years old. There were 254 (22.6%) with single-episode major depression, 529 (47.0%) recurrent-episode major depression, 149 (13.2%) dysthymia, 82 (7.3%) depression not-otherwise-specified (NOS), and 112 (9.9%) bipolar disorders.

ANXIETY DISORDERS

For the outpatients with anxiety disorders, there were 310 (42.3%) men and 422 (57.7%) women. There were 690 (94.3%) Whites, 28 (3.8%) Blacks, and 14 (1.9%) Asians. The mean age was 35.66 (SD = 11.22) years old. There were 184 (25.1%) panic with agoraphobia, 153 (20.9%) panic without agoraphobia, 161 (22.0%) generalized anxiety, 94 (12.8%) social phobia, 51 (7.0%) obsessive compulsive, 52 (7.1%) anxiety NOS, 27 (3.7%) simple anxiety, 7 (1.0%) agoraphobia without a history of panic, and 3 (0.4%) post-traumatic stress disorders.

SAMPLE COMPARISONS

The proportions of men and women in the mood and anxiety samples were comparable, X^2 (1, N = 1,858) = 0.0, ns., and the proportions of Whites, Blacks, and Asians reflected in both samples were also comparable, X^2 (2, N = 1,858) = 0.87, ns. The outpatients diagnosed with mood disorders were older than those diagnosed with anxiety disorders, t (1,856) = 2.96, $p < .01$.

PROCEDURE

After signing voluntary consent forms, the outpatients were routinely administered the BHS as part of the standard intake battery of psychological tests and psychiatric rating scales given to all outpatients being evaluated at the CCT. The SCID-derived diagnoses were also made at this time.

FACTOR ANALYTIC CONSIDERATIONS

A major methodological problem incurred in assessing the factor structure of the BHS is that this scale is composed of dichotomous items. There are 20 true-false items assessing the expectation that one will not be able to overcome an unpleasant life situation or attain things that one values. Nine of the items are keyed false, and 11 are keyed true with 1s being assigned to negative expectations and 0s being assigned to positive expectations. The item responses are summed to yield total scores ranging from 0 to 20. Although the approximate balancing between positively and negative scored items affords some protection against item response biases (Mendonca et al., 1983), there are problems incurred in using dichotomous variables in factor analyses (See Comrey (1988) and Gorsuch (1983) for detailed discussions.)

Briefly, the marginal distributions in the responses of the BHS items would impose limits on the magnitudes of the phi correlations that can be achieved. Spurious difficulty factors can emerge because items with similar marginal distributions will tend to be more highly correlated with one another than those with dissimilar marginal distributions. Gorsuch (1983) suggests that difficulty factors may be identified by factors that are composed of items with similar mean values.

Various data transformations have been proposed to adjust for disparate marginal distributions and to limit the emergence of difficulty factors. For example, Holley and Guilford (1964) proposed the use of the G index, which transforms every item into having a mean of 0.50 and standard deviation of 0.50 by creating an mirror image of the existing data matrix. The item responses for every respondent are reversed as a new set of data, and this reversed set is then appended to the existing set for final analysis. However, Gorsuch (1974) cautions that the extraction of factors based on the G index is oriented toward the combined sample, and the resultant factors are often difficult to interpret.

In discussing the factor analysis of dichotomous data, Joreskog and Sorbom (1988, pp. 202-204) presented an approach using a weighted least-squares factor analysis in which the asymptotic covariance matrix of the tetrachoric correlations among the dichotomous variables is employed. The full-information factor analysis approach used by Young et al. (1992) uses tetrachoric correlations. However, Gorsuch (1983) warns that matrices based upon tetrachoric correlations are often singular and cannot be entered into maximum-likelihood factor-analyses, unless statistical adjustments, such as those incorporated into the full-information factor analysis (Bock et al., 1983), are employed to control for singular correlation matrices (Heywood cases). Gorsuch (1983, pp. 296-297) has suggested, "Phi and point biserial correlation coefficients are better choices for factor analytic work because they are always within normal ranges and give Gramian matrices. The product-moment coefficients also represent relationships in the actual data rather than attempting to generalize to a theoretical variable that has not been observed."

Wilkinson (1985, p. 264) concluded that factor solutions derived from principal components and common factor analyses "rarely differ enough to matter", and Parry and McArdle (1991) found that more sophisticated psychometric models for conducting factor analyses of dichotomous variables were not "markedly" superior to less sophisticated approaches.

PRESENT FACTOR ANALYTIC APPROACHES

Our first approach here involved performing principal components analyses of the phi correlations among the 20 BHS items. A principal components analysis of the G-index correlations among the 20 BHS items represented the second approach, and a weighted least-squares factor analysis of the tetrachoric correlations using the asymptotic covariance matrix reflected the third approach. The latter analysis employed PRELIS (Joreskog & Sorbom, 1986) to calculate the two sets of tetrachoric intercorrelations and asymptotic covariance matrices for the outpatients with either mood or anxiety disorders.

CALIS (SAS Institute, 1990) was used to conduct the exploratory weighted-least-squares factor analyses. A weighted least-squares factor analysis was

employed, instead of a maximum likelihood approach, because (1) the tetrachoric correlation matrix for the outpatients diagnosed with anxiety disorders was singular and (2) Joreskog and Sorbom (1988) indicated that the weighted least-squares approach would yield better estimates of the standard errors for dichotomous data than a maximum-likelihood approach would.

FACTOR (SAS Institute, 1990) was employed to perform the principal components analyses. Varimax rotations were employed for all of the principal components and factor analyses. Kaiser, Hunka, and Bianchini's (1971) factor similarity index was used to compare the structures between the outpatients diagnosed with either mood or anxiety disorders. This index represents the mean cosine between the two sets of components or factors and is interpreted as a correlation coefficient. The correlations among the principal components and factor scores were used to compare factor analytic techniques.

RESULTS

The mean BHS total scores for the outpatients diagnosed with either mood or anxiety disorders were 11.13 (SD = 5.24) and 7.52 (SD = 5.32), respectively. The mean BHS score of the mood disorder sample was higher than that for the anxiety disorder sample, t (1,856) = 14.43, $p < .001$. According to Beck and Steer's (1993) diagnostic ranges, the outpatients who were diagnosed with mood disorders had described moderate hopelessness, whereas the outpatients who were diagnosed with anxiety disorders had reported mild hopelessness. The mean BHS score for the mood disorder group was higher than the mean BHS scores that were reported by Young et al. (1992) for their major affective [N = 730, M = 6.28, SD = 5.52, t (1,854) = 19.07, $p < .001$] and mixed diagnostic [N = 257, M = 7.00, SD = 5.80, Welch's t' (357) = 19.07, $p < .001$] groups. The mean BHS score of the anxiety disorder group was higher than that found in Young et al.'s (1992) major affective sample, t (1,460) = 4.37, $p < .001$, but comparable to that found in their mixed diagnostic sample, Welch's t' (417) = 1.26, ns.

ITEM ANALYSES

The Kuder-Richardson 20 (KR-20) for the BHS was 0.93 in the outpatients diagnosed with mood disorders, and the KR-20 was 0.90 for the outpatients diagnosed with anxiety disorders. All of the corrected item-total correlations of the 20 BHS items shown in Table 1 for both diagnostic samples were significant beyond the .001 level, one-tailed test, even after Bonferroni adjustments (alpha / 20) were used to control for the familywise error rate. Using Fisher Z' transformations, we calculated the mean corrected-item total correlations for the BHS in the diagnostic groups. The mean correlation of 0.51 (SD = 0.16) for the mood disorder sample was comparable to the mean correlation of 0.54 (SD = 0.16) for the anxiety disorder sample, t (38) = 0.76, ns.

The means and standard deviations of the BHS items presented in Table 1 for the outpatients with either mood or anxiety disorders support the concerns about whether the previously reported principal components for the BHS might be artifacts of disproportional splits in the BHS items. For both the mood and anxiety samples, the BHS items with the lowest means within each respective sample were for items #2, #9, #16, and #20 (Table 1). Three of these four items are the

TABLE 1

MEANS, STANDARD DEVIATIONS, AND CORRECTED ITEM-TOTAL CORRELATIONS FOR THE BECK HOPELESSNESS SCALE BY DIAGNOSTIC GROUP

	Mood			Anxiety		
Item	M	SD	r	M	SD	r
1	0.73	0.44	0.55	0.48	0.50	0.59
2	0.21	0.41	0.47	0.10	0.29	0.42
3	0.46	0.50	0.45	0.28	0.45	0.46
4	0.77	0.42	0.31	0.63	0.48	0.45
5	0.59	0.49	0.26	0.50	0.50	0.24
6	0.55	0.50	0.60	0.33	0.47	0.61
7	0.62	0.49	0.63	0.37	0.48	0.66
8	0.74	0.44	0.39	0.62	0.49	0.43
9	0.39	0.49	0.51	0.24	0.43	0.47
10	0.62	0.49	0.35	0.45	0.50	0.45
11	0.49	0.50	0.64	0.30	0.46	0.65
12	0.68	0.47	0.62	0.44	0.50	0.67
13	0.45	0.50	0.57	0.32	0.47	0.51
14	0.69	0.46	0.54	0.45	0.50	0.64
15	0.78	0.41	0.53	0.55	0.50	0.62
16	0.22	0.41	0.47	0.08	0.27	0.42
17	0.39	0.49	0.65	0.17	0.38	0.63
18	0.89	0.31	0.36	0.73	0.45	0.52
19	0.58	0.49	0.66	0.35	0.48	0.68
20	0.29	0.46	0.55	0.15	0.35	0.53

Note — n for mood disorders = 1,126, and n for anxiety disorders = 732. All of the corrected item-total correlations were significant beyond the .001 level, one-tailed test (Bonferroni alpha / 20).

items that were described above as constituting the Resignation dimension that was observed by us to occur in all six previously reported principal-components analyses of the BHS. In contrast, items #1, #4, #15, and #18 displayed some of the highest means in Table 1, and these items represent the other dimensions that were discussed above as reflecting the Rejection and Acceptance dimensions.

STRUCTURES

To determine the number of dimensions underlying the BHS responses in both diagnostic groups, we first examined the goodness of fit indices for the weighted least-squares factor analyses of the tetrachoric correlations shown in Table 2. The

chi-square tests revealed that four factors were statistically sufficient to explain the covariation in the mood disorder sample, whereas three factors were statistically sufficient to describe the covariation in the anxiety disorder group. However, the magnitudes of the root-mean-squares given in Table 2 indicated that only small amounts of the total residual variance in both samples were accounted for beyond that explained by the first factors in both samples. Additional factors afforded < 5% reductions in the root-mean-squares for both samples.

TABLE 2

CHI-SQUARE STATISTICS AND ROOT-MEAN-SQUARES BY NUMBER OF WEIGHTED LEAST-SQUARES FACTORS

Number of factors	Chi-Square	(df)	Root-Mean-Square
Mood			
One	792.24*	170	.08
Two	354.65*	151	.06
Three	209.03*	133	.04
Four	135.74	116	.03
Anxiety			
One	437.13*	170	.08
Two	213.25*	151	.05
Three	125.21	133	.04

Note — n for mood disorders = 1,126, and n for anxiety disorders = 732.
* $p < .001$

To explore whether any of the BHS dimensions were robust with respect to type of factor analysis and data transformation, we first calculated three-factor solutions for the tetrachoric, phi, and G-index correlations in both diagnostic samples. As mentioned previously, the weighted least-squares approach was used with the tetrachoric correlations, and the principal components approach was used with the phi and G-index correlations. Although visual inspection of the magnitudes of the salient (≥ .40) loadings for the resultant three-factor solutions suggested that the factors extracted for the tetrachoric and phi correlations were similar, only one component, identified by items #9, #16, and #20 (Resignation), existed across all three types of analysis. The salient loadings for the other two principal components based on the G-index correlations were dissimilar to those that were based on the tetrachoric and phi correlations. Therefore, we calculated two factor solutions to ascertain whether this number of factors displayed more congruency with respect to similar patterns of salient loadings not only between samples using the same factor analytic techniques, but also across factor analytic techniques.

TABLE 3

VARIMAX-ROTATED FACTOR LOADINGS FOR THE BECK HOPELESSNESS SCALE BY TYPE OF CORRELATION AND DIAGNOSTIC GROUP ALONG WITH CORRELATIONS AMONG THE DIFFERENT TYPES OF FACTOR SCORES

Item	Pessimism			Resignation		
	Tetrachoric	G-Index	Phi	Tetrachoric	G-Index	Phi
Mood						
1	0.78	0.75	0.67	0.27	0.08	0.19
2	0.42	-0.25	0.11	0.63	0.76	0.67
3	0.55	0.25	0.40	0.28	0.47	0.32
4	0.42	0.66	0.46	0.23	-0.18	0.04
5	0.29	0.36	0.26	0.20	0.09	0.16
6	0.74	0.53	0.62	0.31	0.41	0.30
7	0.69	0.64	0.61	0.46	0.35	0.37
8	0.39	0.64	0.41	0.38	-0.05	0.20
9	0.28	0.12	0.20	0.70	0.63	0.62
10	0.43	0.48	0.43	0.20	0.09	0.11
11	0.60	0.42	0.46	0.57	0.59	0.54
12	0.52	0.69	0.52	0.67	0.25	0.44
13	0.72	0.33	0.50	0.35	0.58	0.42
14	0.43	0.65	0.42	0.66	0.20	0.45
15	0.81	0.79	0.69	0.24	-0.04	0.13
16	0.18	-0.25	0.03	0.89	0.77	0.75
17	0.61	0.20	0.36	0.63	0.77	0.67
18	0.64	0.78	0.60	0.19	-0.34	-0.06
19	0.73	0.61	0.61	0.46	0.44	0.41
20	0.29	-0.07	0.12	0.85	0.79	0.77

Type	Correlations among Different Types of Factor Scores					
Tetra	1.00			1.00		
G-Index	0.89	1.00		0.80	1.00	
Phi	0.93	0.95	1.00	0.93	0.92	1.00

Table 3, continued

Item	Pessimism			Resignation		
	Tetrachoric	G-Index	Phi	Tetrachoric	G-Index	Phi
Anxiety						
1	0.81	0.65	0.71	0.23	0.26	0.14
2	0.46	-0.12	0.13	0.60	0.87	0.63
3	0.52	0.20	0.39	0.36	0.63	0.35
4	0.46	0.62	0.52	0.43	-0.11	0.14
5	0.22	0.26	0.22	0.25	0.12	0.18
6	0.73	0.45	0.63	0.36	0.57	0.27
7	0.75	0.56	0.68	0.43	0.51	0.29
8	0.37	0.58	0.48	0.50	-0.07	0.17
9	0.26	0.15	0.26	0.71	0.69	0.53
10	0.45	0.47	0.50	0.38	0.26	0.17
11	0.61	0.36	0.50	0.59	0.69	0.52
12	0.59	0.61	0.61	0.64	0.41	0.40
13	0.63	0.28	0.42	0.34	0.61	0.40
14	0.47	0.58	0.54	0.72	0.38	0.43
15	0.83	0.73	0.74	0.28	0.14	0.14
16	0.21	-0.15	0.03	0.93	0.88	0.76
17	0.64	0.12	0.35	0.64	0.87	0.70
18	0.74	0.74	0.70	0.36	-0.29	0.02
19	0.81	0.52	0.66	0.39	0.58	0.35
20	0.35	0.01	0.17	0.84	0.87	0.77
Type	Correlations among Different Types of Factor Scores					
Tetra	1.00			1.00		
G-Index	0.91	1.00		0.64	1.00	
Phi	0.95	0.97	1.00	0.87	0.90	1.00

Note. n for mood disorders = 1,126, and n for anxiety disorders = 732. The loadings of the marker items used in identifying the factors are underlined.

STRUCTURE COMPARISONS

Table 3 presents the varimax-rotated loadings for the BHS items for the outpatients diagnosed with either mood or anxiety disorders. Kaiser's Measures of Sampling Adequacy (Dziuban & Shirkey, 1974) were 0.93 and 0.94, respectively, for the outpatients diagnosed with either mood or anxiety disorders. Both of these MSA values were considered by Kaiser (1970) to be "marvelous." The first principal components based on the phi correlation matrices of the outpatients

diagnosed with either mood or anxiety disorders explained, respectively, 33.5% and 36.1% of the total variance, whereas the second components explained, respectively, 7.7% and 7.8% of the total variance.

The patterns of salient (≥ .40) loadings for the two factor solutions shown in Table 3 were quite similar. The correlations among the different types of factor and principal components scores reflecting the two dimensions shown in Table 3 were highly positive and significant (ps < .001), and the three mean-factor similarity indices (Kaiser et al., 1971) that were calculated between the factor and principal components structures for the mood and anxiety groups were all equal to 0.99.

We next inspected the magnitudes of the loadings shown in Table 3 and observed that there were two sets of marker items in all six factor structures that displayed salient loadings on the same dimension and had low loadings on the opposite dimension. For operational purposes, we identified these sets of marker items by looking for items that had three loadings ≥ .40 on the same content dimension across all six factor analyses and also had three loadings < .40 on the opposite content dimension. The first principal components or factors shown in Table 3 were thus composed of item #1 ("I look forward to the future with hope and enthusiasm."), item #10 ("My past experiences have prepared me well for the future."), item #15 ("I have great faith in the future."), and item #18 ("The future seems vague and uncertain to me."). These four items reflected general *pessimism* about the future.

A second dimension that occurred across all six factor structures given in Table 3 was composed of item #9 ("I just don't get the breaks, there's no reason to believe I will in the future."), item #16 ("I never get what I want so it is foolish to want anything."), and item #20 ("There's no use in really trying to get something I want because I probably won't get it."). This set of marker items denoted the dimension that Beck et al. (1974) referred to as "Loss of Motivation," Hill et al. (1988) considered to reflect "Giving Up," and Steer et al. (1993, 1994) called "Resignation to the Futility of Changing the Future." We preferred the latter designation here.

DISCUSSION

There are several conclusions that can be drawn from the different types of factor analyses which were performed with the BHS for outpatients diagnosed with either primary mood or primary anxiety disorders. First, regardless of the factor analytic approach or type of data transformation that was employed, the first factor or component that emerged in both diagnostic groups was general *pessimism* about the future and the second factor or component was *resignation* to the futility of changing the future. Young et al.'s (1992) finding that the structure of the BHS was not dependent on the diagnostic composition of the sample was supported. We found that the compositions of our *pessimism* and *resignation* dimensions were comparable for (1) outpatients who were diagnosed with anxiety disorders and described mild hopelessness and (2) outpatients who were diagnosed with mood disorders and reported moderate hopelessness.

Second, there was support for Young et al.'s (1992) finding that the amount of variance contributed to the explanation of the overall set of BHS responses by the

extraction of additional factors was small compared to that afforded by the first factor by itself. The *pessimism* dimension here explained approximately eight times the total amount of variance that was described by the *resignation* dimension in either diagnostic group. However, although the relative importance of the second dimension was small, we disagree with the Young et al.'s (1992, p. 585) conclusion that additional factors of the BHS, such as *resignation*, afford "no additional explanatory or conceptual advantage."

The *resignation* dimension stresses how futile a patient might feel about making any attempt to change his or her future, whereas the *pessimism* dimension emphasizes his or her beliefs about the future. The former dimension addresses motivational issues, and the latter dimension highlights dysfunctional attitudes and cognitions. Clinically, these two aspects of hopelessness may require different types of psychotherapeutic interventions. For example, a behavioral reinforcement approach might be employed to reward an unemployed patient who actually applies for a job or enrolls in a job training program, whereas cognitive therapy might be used to help the same patient replace incorrect beliefs, premises, and expectations about potentially working with more realistic ones.

Third, the present findings have implications for the use of different factor-analytic approaches with dichotomous data, such as those contained in the BHS. The common factor and principal components analyses produced comparable results, regardless of whether a tetrachoric or G-index transformation was employed. Our analyses thus supported Gorsuch's (1983) recommendation about using phi correlations for the analysis of dichotomous data, and the principal components approach yielded results that were comparable to those afforded by the more sophisticated, weighted-least-squares factor analysis. However, the present item analyses for the BHS stress the need for a strong empirical framework, such as that provided by the six previous principal-components analyses that we identified above, against which to judge whether spurious difficulty factors have arisen.

REFERENCES

American Psychiatric Association. (1987). *Diagnostic and statistical manual of mental disorders* (3rd ed., rev.). Washington, DC: Author.

Beck, A. T. (1967). *Depression: Clinical, experimental, and theoretical aspects.* New York: Hoeber.

Beck, A. T. (1976). *Cognitive therapy and the emotional disorders.* New York: International Universities Press.

Beck, A. T. (1986). Hopelessness as a predictor of eventual suicide. In J. J. Mann & M. Stanley (Eds.), *Psychobiology of suicide behavior.* New York: New York Academy of Sciences. Pp. 90-96.

Beck, A. T., & Steer, R. A. (1993). *Manual for the Beck Hopelessness Scale.* San Antonio, TX: Psychological Corporation.

Beck, A. T., Weissman, A., Lester, D., & Trexler, L. (1974) The measurement of pessimism: The Hopelessness Scale. *Journal of Consulting and Clinical Psychology*, *42*, 861-865.

Bock, R. D., Gibbons, R., & Muraki, E. (1988). Full-information item factor analysis. *Applied Psychological Measurement*, *12*, 261-280.

Comrey, A. L. (1988). Factor-analytic methods of scale development in personality and clinical psychology. *Journal of Consulting and Clinical Psychology*, *56*, 754-761.

Gorsuch, R. L. (1974). *Factor analysis*. Philadelphia, PA: W. B. Saunders.

Gorsuch, R. L. (1983). *Factor analysis* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Hill, R. D., Gallagher, D., Thompson, L. W., & Ishida, T. (1988). Hopelessness as a measure of suicidal intent in the depressed elderly. *Psychology and Aging*, *3*, 230-232

Holley, J. W., & Guilford,, J. P. (1964). A note on the G index of agreement. *Educational and Psychological Measurement*, *24*, 749-753.

Joreskog, K. G., & Sorbom, D. (1986). *PRELIS: A preprocessor for LISREL*. Mooresville, IN: Scientific Software.

Joreskog, K. G., & Sorbom, D. (1988). *LISREL 7: A guide to the program and applications*. Chicago, IL: SPSS, Inc.

Kaiser, H. F. (1970) A second generation little jiffy. *Psychometrika*, *35*, 401-415.

Kaiser, H. F., Hunka, S., & Bianchini, J. (1971) Relating factors between studies based upon different individuals. *Multivariate Behavioral Research*, *6*, 409-422.

Lord, F. M., & Novick, M. R. (1956). *Statistical theories of mental test scores*. Reading, MA: Addison-Wesley.

Mendonca, J. D., Holden, R. R., Mazmanian, D., & Dolan, J. (1983). The influence of response style on the Beck Hopelessness Scale. *Canadian Journal of Behavioral Science*, *15*, 237-247.

Nekanda-Trepka, C. J. S., Bishop, S., & Blackburn, I. M. (1983). Hopelessness and depression. *British Journal of Clinical Psychology*, *22*, 49-60.

Parry, C. D., & Mc Ardle, J. J. (1991). An applied comparison of methods of least-squares factor analysis of dichotomous variables. *Applied Psychological Measurement*, *15*, 35-46.

Riskind, J. H., Beck, A. T., Berchick, R. J., Brown, G., & Steer, R. A. (1987). Reliability of DSM-III diagnoses of major depression and generalized anxiety using the Structured Clinical Interview for DSM-III (SCID). *Archives of General Psychiatry*, *44*, 817-820.

SAS Institute, Inc. (1990) *SAS/STAT user's guide* (ver. 6 , 4th ed., vol. 1) Gary, NC: Author.

Spitzer, R. L., Williams, J. B., & Gibbon, M. (1987). *Instruction manual for the Structured Interview for the DSM-III-R (SCID-OP)*. New York, NY: Biometrics Research Department, New York Psychiatric Institute.

Steer, R. A., Iguchi, M. Y., & Platt, J. J. (1994). Hopelessness in IV drug users not in treatment and seeking HIV testing and counseling. *Drug and Alcohol Dependence*, *34*, 99-103.

Steer, R. A., Kumar, G., & Beck, A. T. (1993). Hopelessness in adolescent psychiatric inpatients. *Psychological Reports*, *72*, 559-564.

Truant, G. S., O'Reilly, R., & Donaldson, L. (1991) How psychiatrists weigh risk factors when assessing suicide risk. *Suicide and Life-Threatening Behavior*, *21*, 106-114.

Young, M. A., Halper, I. S., Clark, D. C., Scheftner, W. & Fawcett, J. (1992). An item-response theory evaluation of the Beck Hopelessness Scale. *Cognitive Therapy and Research*, *16*, 579-587.

Wilkinson, L. (1985). *SYSTAT manual*. Evanston, IL: SYSTAT, Inc.

Author Note

1. Address all correspondence to Robert A. Steer, Ed.D., UMDNJ-SOM, Dept. of Psychiatry, 40 East Laurel Road, PCC 218, Stratford, NJ 08084-1350.

Multivariate Experimental Clinical Research, Volume 11, No. 3, 1997, 145-150.

A TEST OF POLYMORPHOUS-PERVERSE INSTINCTUAL DEVELOPMENT

Raymond B. Cattell
University of Hawaii

ABSTRACT

It is first shown from the factor analytic studies of eight investigators that groups of attitudes, objectively measured, mark some six ergs ("instincts") as unitary factors. It is then shown that in addition to the high markers certain other attitudes load these factors in anomalistic low degrees. These indicate connections that may not be clinically easy to perceive or verify. A theory is developed that a more "dense" data base could yield primary factors below the general ergic factor, that represent Freudian and other concepts of polymorphously perverse instinctual fixations. A basis of separation of ordinary low loaded attitudes from such perversion manifestations is thus a suggested possibility.

INTRODUCTION

Since psychoanalysis began to look for experimental support, the results have not been particularly successful (Eysenck, 1992). The most general reason for this is that the experimentalists followed the mainstream Wundtian concept of bivariate experiment, tested by ANOVA. Meanwhile multivariate, factor analytic methods were regarded, with a few exceptions like Kline (1972) and Cattell and Child (1975), as an irrelevant mathematical development. But one cannot experimentally represent a broad concept by the single variable in a bivariate experiment. For example, many pointless experiments have represented anxiety by a measure of skin resistance, whereas, as the OA Battery (Cattell & Schuerger, 1978) shows, at least eight diverse variables are needed to measure the factor analytically defined unitary concept.

A search for the nature and number of primary unitary instincts shows not one or two as Freud suggested, but seven or eight ; sex, fear, narcism, assertiveness, parental protectiveness, gregariousness, curiosity, pugnacity, and possibly rest-seeking and laughter. In scientific psychology these supersede the intuitive lists of Freud, Murray, Drover, McDougall and others.

To avoid getting entangled in concepts of instinct for which we have no experimental basis, we have called those forces ergs, permitting the adjective ergic. Each points to an area of ultimate satisfaction to which our attitudes are subsidiated (to use Murray's term). In the factor analysis of a large number of human at-

titudes in our culture, the nature of these goals is indicated by the highest-loaded attitudes in each. For example, "I want to satisfy my sexual needs" loads mating (sex) .60, and "We must destroy our enemies" loads pugnacity .80.

The new dynamics, incidentally, does not measure attitudes by questionnaire but by objective test devices. The validity of these, such as projection, fantasy, learning rate, perseveration, systolic blood pressure change, decision speed and GSR, can be read in the opening researches (Cattell & Child, 1975). One should also note that the same ergic patterns have been found in the study of attitudes as states, in P-technique (Cattell, 1950; Cattell & Cross, 1952; Cattell & Horn, 1963).

THE THEORY OF POLYMORPHOUS DEVELOPMENT

For reasons not yet understood (or capable of being understood at the advent of measurement) the ultimate goals to which ergs are naturally directed are sometimes subject to change. Sublimation may be considered such a case, though perhaps different in cause and nature from the early life "misdevelopments" which Freud described. They are perhaps due to sudden blockage of the normal outlets or to conflict with other ergs; but this question does not concern us at the present point. Our theory departs from Freud's only in assuming that the changes are not peculiar to sex, but can take place in any erg, and notably in fear, narcism, and parental protectiveness (compassion).

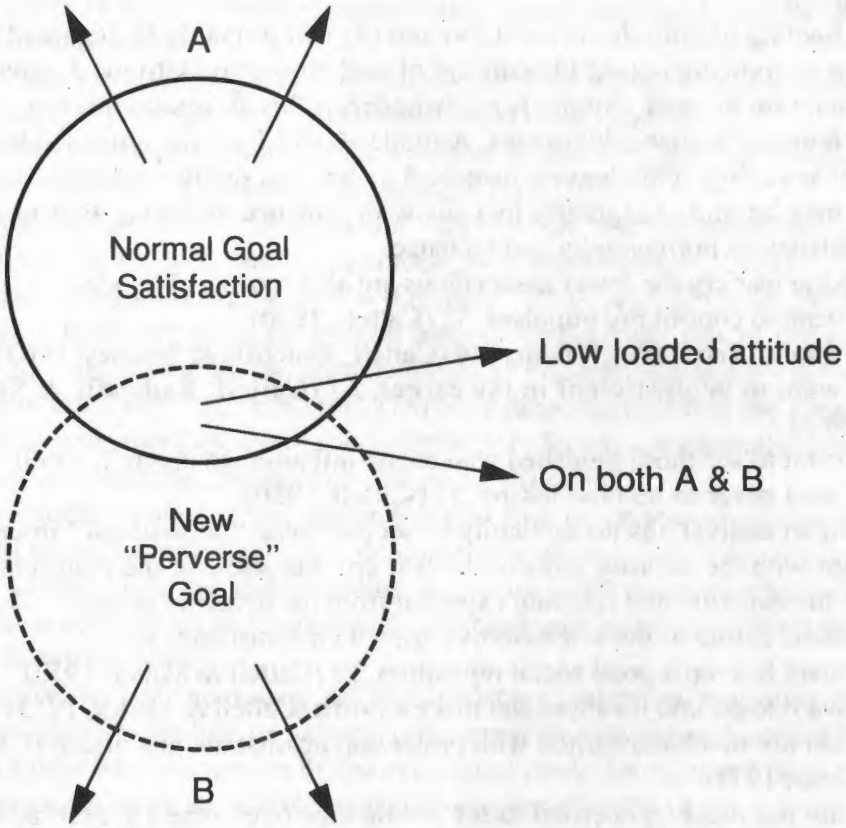
Assuming that the change is gradual, we should soon have, in general, a situation as in Figure 1, in which the departing new center of desires, B, still overlaps with the old A. We should now expect to find among the attitudes loading A some of low loading, because they arise only from the remnant of A still in B.

It is necessary to consider this point also in factor analytic terms. When we find an attitude with a low but significant loading on a factor it may be due (a) to other factors accounting for much of its communality, (b) to an intrinsically weak action of the factor on that attitude, or (c) to the choice of variables being so restricted that the factor which really loads it does not appear on the factor list. The last can probably be ruled out in the experiments to be cited, because the experimenters were deliberately led by the personality-sphere rule. That is to say, they included variables touching every region of the interests of an individual in our culture.

The choice of (b) is to be based partly on a psychological judgment that the nature of the low attitude does not fully express A, and partly on the assumption of small overlap of the factors, which causes only a part of the variance of A to be active.

There is no solution yet to be obtained from knowing what new desires operate in B. In fact it might be argued that if B really exists, we should already have picked it up as an attitude loading pattern. But in our experiments no pathological attitudes were used to pick up B, and it would be likely to appear in only a minority of subjects, thus escaping factor analytic representatives in the general public. We must therefore look for significantly but low loaded attitudes in A which appear to have no very good connection, psychologically, with A.

Figure 1



EXPERIMENTAL RESULTS

We may now examine a variety of experiments in which low loadings have occurred in addition to the high loadings marking the erg. The instrument used in these experiments was the MAT (Cattell & Child, 1975). The samples of normal adults used varied from research study to research study. One study involved 160 married couples, another study used 199 Air Force cadets, and in another study, 277 normals were used. The factor analytic procedures laid particular emphasis on rotation to maximum simple structure. A full summary of studies to 1975 is found in Cattell and Child (1975).

A classical instance is the notoriously polymorphously perverse sex erg, which, in addition to unquestionable sex markers (See Cattell & Child, 1975, p. 38) has shown the following loadings:

1. I want to indulge in smoking .34 (Cattell & Baggaley, 1958)
2. I want to indulge in drinking .45 (Cattell & Baggaley, 1958)
3. I want to travel and see the world .32 (Cattell, 1950)
4. I want to visit more good restaurants .34 (Cattell, 1950)
5. I want to see more violent movies .25 (Cattell & Child, 1975)
6. I do not care about my general reputation .34 (Cattell & Child, 1975)
7. I am not concerned with a sense of duty .29 (Cattell & Child, 1975)

8. I want an alcoholic drink .27 (Birkett & Cattell, 1988)

From now on, for ease of reference, attitudes will be designated by their numbers in this article.

The loading of attitudes (1), (2), (4) and (8) will probably be regarded by any Freudian as confirmation of the concept of oral eroticism. Attitude 3 agrees with the association in most mammals of "wandering" with sexual interest, and the leaving home of human adolescents. Attitudes 6 and 7 are the typical side effects of a free sexuality. This leaves attitude 5 as beyond ready explanation, unless movies may be said to regularly mix sex with violence, but adds evidence to the recent debates on pornography and violence.

With the fear erg the lower associations are also reconcilable, viz:

9. I want to control my impulses .32 (Cattell, 1950)
10. I want to know myself better .44 (Cattell, Radcliffe & Sweney, 1963)
11. I want to be proficient in my career .37 (Cattell, Radcliffe & Sweney, 1963)
12. I want to see those punished who create inflation .44 (Cattell, 1950)
13. I want never to become insane .37 (Cattell, 1950)

Again an analyst has no difficulty in seeing these "anomalous" interests as consistent with the security aims of the fear erg, but some of the connections are scarcely the experimental learning expected from the dynamic lattice.

The same is true of the self-assertive erg which contributes to:

14. I want to keep a good social reputation .28 (Cattell & Miller, 1952)
15. I want to go into business and make a profit (Cattell & Miller, 1952)
16. I am not much concerned with protection against nuclear attack (Cattell & Child, 1975)
17. I am not much concerned to see public vice overcome (Cattell & Baggaley, 1958)
18. I want to have an easy time .47 (Cattell & Baggaley, 1958)

In this case we scent a slightly different quality from that popularly assigned to self-assertion, namely a self-centered dismissal of public responsibilities.

An interesting discovery in the early days of the factorial experiments in dynamic calculus was that Freud was right in distinguishing two forms of the sex drive, namely, the ordinary heterosexual drive and what I have shortened in past writings as narcism. Its revelatory attitudes, beyond those of the primary, self-centered sensuality items are:

19. I want to rest late in bed .86 (Cattell & Baggaley, 1958)
20. I do not particularly want to stick to my job .61 (Cattell & Baggaley, 1958)
21. I want to see divorce made easier (Cattell & Child, 1975)
22. I want to see more violent movies .55 (Cattell & Child, 1975)
23. I want more time to myself .43 (Cattell & Child, 1975)
24. I want to drink more alcohol .58 (Birkett & Cattell, 1978)

In this case, the arguments for subsidiation to a narcistic goal are obscure in the cases of attitudes 21 and 22. But the data shows definitely that a connection must be found, in the average person. We observe, however, that the narcistic erg has given a few anomalous results in associated experimental work. For example, among 250 medical doctors, it was at a significantly higher level in those who held presidential, secretarial, treasurer, etc. positions in local medical societies. It is possibly an erg that has a special facility to sublimate. Besides its correlation

of .58 with the ordinary sex drive, it has a correlation of .42 with autism (M Factor in the 16PF test) and of .35 with the protective pitying erg (Cattell & Horn, 1963). There is room for more intensive study of the Narcism factor, in both its U and its I components

We next approach pugnacity, a component often mixed with self-assertion in "aggression". There are two significantly loaded non-markers:

25. I want more protection against a nuclear bomb .45 (Cattell & Child, 1975)
26. I want to take life easy .31 (Cattell & Child, 1975)

The parental protective "compassion" erg also has two non-markers:

27. I want to see birth control expanded .34 (Cattell & Child, 1975)
28. I do not mind damaging my "self respect" .29 (Cattell & Child, 1975)

This points to the erg being the general basis of altruism and lack of self-concern.

For the busy clinician the time demands of P-technique make it a "deluxe" treatment, but Cattell and Birkett (1978) have demonstrated that the time on a P-technique experiment can safely be powerfully reduced, to a clinically practicable series of occasions, say 50, without loss of the structure.

The general upshot of this research is that firm correlational evidence can be given to trace the particular ergic depth origins of most interests, and with secondary assumptions, for understanding the nature of perversions which clinicians have been compelled to recognize as changed goals on intuitive procedures (Fogel & Myers, 1990).

The problem of separating the low loading of attitudes that arise from the attitudes sharing loading on several other ergic factors from those which arise from a sublimated distortion of the ergic goal itself can perhaps be tackled by other observations than we have primarily suggested in (b) above. Meanwhile we have a method for throwing the attitude (or symptom) into one of a two class category. Possibly more developed factor analysis will present a better way to recognize these ergic goal variants.

The experiment required to check this adequately would be to take a more refined, closely-packed data base of attitudes (much as Goldberg & Norman used in the personality field) obtaining subprimaries which placed the original 16PF primaries as if they secondaries (second order factors). We hypothesize that in sex there would appear individual differences in each of several "perverse" (or undeveloped) satisfactions in such primaries as oral erotic sex, anal erotic sex, homosexuality, and so on.

This experiment is of the highest importance in the developing field of the dynamic calculus (Cattell, 1985). It is true that an early pioneer study of the oral erotic personality by Kline (1972) failed to give a definite, positive result in finding unity factors for Freudian concepts. But this study was undertaken in an early stage of the art of factor analysis and should not discourage opening up the whole question of determining the structure of dynamic factors in terms of factor analytic defined areas of satisfaction.

The above experiments are mainly carried out by R-technique (correlations across people) but the agreement is good with P-technique (Birkett & Cattell, 1978) results treating ergic levels within an individual.

In summary we have shown (1) that factor analytic correlational methods can reveal the ergs and the ergic goals of any specific course of action (and therefore

of symptoms), and (2) that lower loadings in certain conditions, can reveal that an attitude (symptom) derives from a changed form of an ergic satisfaction pattern, and is what Freud described as polymorphously perverse.

REFERENCES

Birkett, H. & Cattell, R. B. Diagnosis of the dynamic roots of a clinical symptom by P-technique: A case of episodic alcoholism. *Multivariate Experimental Clinical Research,* 1978, *3*, 173-194.

Cattell, R. B. The discovery of ergic structure in man in terms of common attitudes. *Journal of Abnormal & Social Psychology*, 1950, *45*, 958-618.

Cattell, R. B. *Human Motivation and the Dynamic Calculus.* New York: Praeger, 1985.

Cattell R. B. & Baggaley, A. R. A confirmation of ergic and engram structure on attitudes measured objectively. *Australian Journal of Psychology.* 1958, *10*, 227-318.

Cattell, R. B. & Birkett, H. Can P-technique be practically shortened? *Multivariate Experimental Clinical Research,* 1980, *5*, 1-16.

Cattell, R. B. & Child, D. *Motivation and Dynamic Structure*, New York: Wiley, 1975.

Cattell, R. B. & Cross, K. Comparison of the ergic and self-sentiment structures found in R- and P-techniques. *Journal of Personality*, 1952, *21*, 250-270.

Cattell, R. B. & Horn, J. L. An integrating study of the factor structure of adult attitudes and interests. *Genetic Psychology Monographs*, 1963, 91-149.

Cattell, R. B. & Miller, A. A confirmation of the ergic and self-sentiment patterns among dynamic traits by R-technique. *British Journal of Psychology*, 1952, *43*, 280-294.

Cattell, R. B., Radcliffe, J. & Sweney, A. B. The nature and measurement of components of motivation. *Genetic Psychology Monographs*, 1963, *68*, 49-211.

Cattell, R. B. & Schuerger, J. M. *Personality Theory in Action: the O-B Battery.* Champaign, IL: IPAT, 1978.

Eysenck, H. J. *The Decline and Fall of the Freudian Empire*. Washington, D.C.: Scott Foresman, 1992.

Fogel, G. T. & Myers, W. A. *Perversions and Near Perversions in Clinical Practice.* New Haven, CT: Yale University Press, 1990.

Hess, E. H. *Science*. 1959, *130*, 133-141.

Kline, P. *Fact and Fantasy in Freudian Theory*, London: Methuen, 1972.

MANUSCRIPT INFORMATION

Submission and Review. Manuscripts should be prepared in triplicate following the instructions contained in the most recent edition of the *Publication Manual of the American Psychological Association.* Manuscripts will be forwarded for reading by two members of the Board of Consulting Editors selected for their expertise in the area covered by the paper. Editorial decisions should be received by the author(s) as soon as possible.

Length. Although papers should be as concise as clarity permits, and tables, figures, and graphs kept as few as possible, there is no standard limit on length. "Research notes" of 500 words or less (including references) will be considered for publication in a special section.

Abstract. All manuscripts should be preceded by an abstract not to exceed 150 words.

Reprints. Fifty reprints of each manuscript accepted for publication are provided free of charge. Additional reprints in quantities of fifty may be ordered on the form sent out with proofs, provided the order is returned within seven days of receipt of proofs.

Tables, Figures, and Graphs. Tables, figures, and graphs should, except in rare cases, be camera ready. The final, revised copy of any manuscript accepted for publication should be submitted in ASCII format on any size diskette.

General. Submission of any manuscript will be taken to imply that it is unpublished and is not being considered for publication elsewhere. Papers published in *Multivariate Experimental Clinical Research* may not be reprinted or published in translation without permission from the Editor. This restriction is waived for authors who wish to reproduce their own articles and when material is reproduced in limited quantity for instructional purposes. Contributions are welcomed from researchers of all nationalities, but must be written in English.

SUBSCRIPTIONS

Subscriptions are available on a per volume basis. The subscription rates are $29 (U.S.) per volume for subscribers in the United States and $35 (U.S.) for subscribers outside of the United States. Single issue price when available is $10.00 (U.S.) for the U.S. and $12.00 (U.S.) for out of U.S. Prices are subject to change without notice.

ADDRESS FOR SUBSCRIPTIONS

Multivariate Experimental Clinical Research, Department of Psychology ¥ Box 34, Wichita State University, Wichita, KS 67260-0034, U.S.A.

ADDRESS FOR MANUSCRIPT SUBMISSION

John E. Cornell, Ph.D.
Review Editor
GRECC/ALMMVH (182)
7400 Merton Minter Blvd
San Antonio, TX 78284, U.S.A.

TABLE OF CONTENTS