

Confirmatory Factor Analysis and Reliability of the Mental Health Inventory for Australian Adolescents¹

Bernd G. Heubeck and James T. Neill
The Australian National University

Key words: Mental Health Inventory, MHI, Australian, Adolescents

¹ This project was partially supported by a grant from the Australian Commonwealth Department of Health and Aged Care under the National Mental Health Strategy. Please address correspondence to Bernd Heubeck, Division of Psychology, School of Life Sciences, Faculty of Science, The Australian National University, Canberra 0200, Australia or e-mail (Bernd.Heubeck@anu.edu.au).

Summary

Ostroff, Woolverton, Berry, and Lesko (1996) examined the adolescent subsample of Veit and Ware's 1983 normative data for the Mental Health Inventory and recommended a two-factor rather than the original five-factor model for the assessment of adolescents' mental health. Analysis of a 30 item version in a new, independent sample of 878 adolescents in another English-speaking country supported a two-factor model of psychological well-being and distress for boys and girls. Internal consistency was $>.9$, and scores were stable ($\sim.7$) over a 10-week period. Boys reported slightly better mental health than girls, as in the original American research.

The mental health of adolescents is of major concern to researchers, mental health practitioners, educators, parents and service planners alike. The end of the twentieth century saw a move towards a greater emphasis on early detection and intervention as well as service planning for entire populations (rather than only responding to clinical emergencies) and this trend is likely to continue in the twenty-first century. Such broad population based strategies should be based on quality information gained from large scale epidemiological studies. There is, however, a relative paucity of standardized instruments available for the assessment of mental health in nonclinical adolescent populations (Ostroff, Woolverton, Berry & Lesko, 1996). Current mental health instruments typically reflect a psychopathological orientation and consequently are not able to discriminate among the many adolescents who score at the nonpathological end of these scales.

One instrument that has been specifically designed for assessing the mental health of nonclinical (as well as clinical) samples is the Mental Health Inventory (Veit & Ware, 1983). The MHI is relatively short (38-items) and includes an assessment of psychological well-being as well as signs of psychopathology. Veit and Ware (1983) studied data from over 5000 participants in the Rand Health Insurance Experiment using exploratory and confirmatory factor analytic methods. They concluded that responses to the MHI can be conceptualised at several levels of differentiation. At the most general level they described a global mental health factor. In addition, they presented a two-factor (Psychological Distress and Psychological Well-being) and a five-factor formulation of mental health responses (Anxiety, Depression, Loss of Behavioral/Emotional Control, Emotional Ties, and General Positive Affect). These factors were not only found in the overall sample but replicated in five regional subsamples as well. Extensive reliability and validity information is available for the MHI (e.g. Cassileth *et al.*, 1984; Rosenthal, *et al.*, 1991; Veit & Ware, 1983; Ware, Davies-Avery, & Brook, 1980; Ware, Manning, Duan, Wells, & Newhouse, 1984).

Despite this impressive background, researchers or clinicians interested in adolescent mental health actually find that there is a paucity of information on the MHI which is specific to the adolescent years. Veit and Ware's (1983) analyses collapsed data across a wide age range (adolescence to old age), thus obfuscating the particular structure underlying adolescent responses to the MHI. A study of the factor structure of a 26-item version of the MHI in Taiwan included about 100 adolescents (Liang, Wu, Krause, Chiang, & Wu, 1992), but unfortunately collapsed all age groups (from 14 to 70+), just as Veit and Ware had done.

Ostroff, *et al.* (1996) reexamined the adolescent subsample ($N=953$) of Veit and Ware's (1983) American normative data and reported finding only weak support for the one-factor model, and no support for the five-factor model. However, principal component analyses with varimax rotation appeared to support a two-factor model (Psychological Distress and Psychological Well-being) similar to the original MHI two-factor model. Based on these results Ostroff, *et al.* (1996, p.107) recommended a Psychological Distress and a Psychological Well-being scale for the assessment of adolescent mental health, "particularly in research studies in which comparison to a normative adolescent population is indicated".

The generality of this conclusion, however, may be limited, as no sample outside the Rand Health Insurance Experiment has been studied and the data is now at least 20 years old. In addition, the question of sex differences in the composition of mental health factors during adolescence was not addressed, although mean differences in psychological distress and well-being were found between boys and girls (Ostroff, *et al.*, 1996). In Australia recent discussions about national mental health outcome measures, including measures for

adolescents, have lead to renewed interest in the MHI. However, no data has been published on the psychometric properties of the MHI with Australian adolescents. Researchers in other English-speaking countries will also be interested in a further evaluation of the instrument before they adopt Ostroff *et al.*'s (1996) recommendation.

The current study examined the factor structure underlying adolescent responses in a new sample, drawn from another English-speaking country, with particular attention to the similarity of responses from boys and girls. This study also set out to examine two aspects of reliability, the internal consistency of responses to the Mental Health Inventory as well as their stability over a period of two and a half months.

METHOD

Participants

Data were available for 878 students (51% boys and 49% girls) who attended ten high schools in four Australian capital cities and some country areas. The schools included coeducational government schools and single-sex and coeducational private schools. Homogeneous samples are problematic in factor analytic work because the resulting factors often do not generalize to other populations. The range of schools included in the current study was intended to mitigate against such bias to some extent. However, despite its size and diversity, the sample cannot be claimed to be representative of high school students in Australia. Most students were in Grade 9 (73%), the others were in Grade 10 (12%) and 11 (15%). The average age was 14.7 years ($SD=.9$). All students responded to the Mental Health Inventory during a normal classroom period at their desks under the supervision of a researcher or a teacher. They had been invited to contribute to a study on “adolescent well-being” and there were fewer than five percent refusals or noncompletions.

Measure

The Mental Health Inventory used in the current study was based on the 38-item instrument constructed by Veit and Ware (1983). Each item asked respondents to rate on a six point frequency or intensity scale how they had been feeling during the past month. One item (felt depressed) was rated on a five point scale. An example of a positive affect item is “How much of the time, during the past month, have you felt calm and peaceful?” Possible answers included 1 = all of the time, 2 = most of the time, 3 = a good bit of the time, 4 = some of the time, 5 = a little of the time, 6 = none of the time. Eight items, however, were dropped from the original MHI, creating a 30-item instrument. The three ‘Emotional Ties’ items were dropped because one item (Time felt lonely) had unacceptable cross-loadings (Ostroff *et al.*, 1996), which left an insufficient number of items to form a distinct scale. Five items from the Loss of Behavioral/Emotional Control scale were dropped (Concern about losing control of mind; Control of behavior, thoughts, feelings; Felt emotionally stable; Better off if dead; Think about taking own life) given (a) the severity of their content and problems in gaining ethics committee approval for their use in schools (e.g., ‘Better off if dead’, ‘Think about taking own life’); (b) very skewed distributions in previous studies, and (c) sizable cross-loadings (Ostroff *et al.*, 1996).

RESULTS

Confirmatory Factor Analyses

Floyd and Widaman (1995) suggested that exploratory factor analysis is most appropriate in the initial stages of model development whereas confirmatory factor analysis provides a more powerful tool in the second stage of research when a model has already been established. In the case of the MHI clearly enunciated models were already available in the literature. Consequently we chose the confirmatory over the exploratory method. Four models of mental health were tested. Each of these models was considered for males and females separately and for the total sample. Covariance matrices¹ were prepared with Prelis 2 and their fit to the four models assessed using the maximum likelihood estimation provided by Lisrel 8 (Jöreskog & Sörbom, 1993a, 1993b).

As no single universally accepted criterion exists to judge model fit several widely accepted goodness-of-fit indices were computed (cf. Table 1). While frequently reported, the χ^2 statistic was not used in the evaluation of model fit as it is known to be strongly dependent on sample size (e.g. Marsh, Balla, & McDonald, 1988). The main criteria used to judge model fit included Bentler's (1990) comparative fit index (CFI), a nonnormed index which incorporates a correction for model complexity, the TLI (Tucker & Lewis, 1973), an index that takes degrees of freedom into account, the parsimony normed fit index PNFI (James, Mulaik, & Brett, 1982), an index based on the population discrepancy function, the root mean square error of approximation (RMSEA; Steiger & Lind, 1980) and the standard root mean square residual (Jöreskog & Sörbom, 1993b). A Monte Carlo study of incremental fit indices conducted by Marsh, Balla, and Hau (1996) supported the CFI and the TLI in the assessment of model fit. Many researchers judge fit to be acceptable for models with CFI and TLI greater than .90, while RMSEA is expected to be less than .08 and SRMR .05 or less. While all fit indices were considered in the overall evaluation, the following presentation highlights the TLI results.

The one-factor model exhibited a poor fit overall (TLI=.75) with a better fit for the girls (TLI=.80) than for the boys (TLI=.69). The two-factor model specified a Psychological Distress and a Psychological Well-being factor. The target items for this model were derived from the factor structure described by Ostroff, *et al.* (1996). An orthogonal two-factor model was tested initially because Ostroff, *et al.* reported loadings based on a varimax rotation. This orthogonal model fitted better than the one-factor model (TLI=.85). However, a correlated two-factor model improved the fit still further (TLI=.88). In addition, this model seemed to apply equally well (cf. Tables 1 and 2) to boys (TLI=.88) and to girls (TLI=.87). Overall the correlation between the Psychological Distress and the Psychological Well-Being factor was substantial (.69), and higher for girls (.84) than for boys (.54).

The three-factor model distinguished Psychological Well-being from Anxiety and these two from a factor called Depression/Loss of Behavioral and Emotional Control. A further improvement in model fit was observed (TLI=.90) and this model seemed to apply equally well to boys (TLI=.90) and to girls (TLI=.89). However, a correlation of over .9 between the Anxiety and Depression factor suggested considerable redundancy when treating these as separate factors for adolescents.

TABLE 1
GOODNESS-OF-FIT INDEXES FOR MHI MODELS BY GENDER

Males	<i>df</i>	χ^2	CFI ^d	TLI ^e	PNFI ^f	RMSEA ^g	SRMR ^h
1 factor	405	2587	.71	.69	.63	.170	.120
2 factor orthogonal ^a	405	1354	.87	.86	.77	.076	.180
2 factor correlated ^a	404	1223	.89	.88	.79	.074	.065
3 factor correlated ^b	402	1129	.90	.90	.79	.067	.060
4 factor correlated ^c	399	1117	.90	.90	.79	.067	.060
Females	<i>df</i>	χ^2	CFI	TLI	PNFI	RMSEA	SRMR
1 factor	405	2040	.82	.80	.73	.130	.073
2 factor orthogonal ^a	405	1909	.83	.82	.74	.100	.300
2 factor correlated ^a	404	1480	.88	.87	.78	.097	.061
3 factor correlated ^b	402	1287	.90	.89	.80	.079	.059
4 factor correlated ^c	399	1271	.90	.89	.79	.079	.059

Note. ^a Psychological Well Being (PWB) and Psychological Distress (PD); ^b Depression and Loss of Behavioral/Emotional Control scales were treated as one-factor, Anxiety and PWB were the other factors; ^c Depression, Loss of Behavioral/Emotional Control, Anxiety, and PWB; ^d Comparative Fit Index, ^e Tucker-Lewis Index; ^f Parsimony Normed Fit Index; ^g Root Mean Square Error of Approximation, ^h Standardized Root Mean Square Residual

Finally, a four-factor model was tested which included Depression, Loss of Behavioral and Emotional Control, Anxiety, and Psychological Well-being. This model closely resembled Veit and Ware's (1983) original five-factor model, except that the three-item 'Emotional Ties' scale had been dropped. No overall improvement in fit was observed (TLI=.90 for boys and girls combined). As a result, parsimony-based goodness-of-fit indices showed a plateau or a drop in fit compared to the three-factor models (see Table 1). In addition, correlations in excess of .9 among the three distress factors were found and the Lisrel phi matrix of factor variances and covariances estimated for this model was not positive definite. This indicated that the very high correlations between the factors in this model were problematic and probably represented overextraction.

PD = Personal Distress, PWB = Personal Well-being. While Ostroff *et al.*'s (1996) report mentioned that they performed confirmatory factor analyses, the details of these analyses were not included. Instead they presented a table of factor loadings for the preferred two factor model based on a principal component analysis with varimax rotation. Their first factor accounted for 35% and the second for an additional 9% of variance. This compares to 44% and 11% in the Australian two factor solution. Table 2 shows the corresponding principal component loadings in the Australian adolescent sample for the reader interested in a direct comparison with Ostroff *et al.*'s results. Of course, a certain lack of match between the two analyses remains due to the fact that the Australian data excluded eight of the 38 items. However, as only two factors were extracted from 30 items, a clear correspondence was expected nonetheless. The same Table 2 shows how the Australian sample extends these results to subsamples of boys and girls in a completely confirmatory paradigm. All 30 loadings for boys and 28 loadings for girls were estimated above .5 demonstrating good factor saturation.

TABLE 2
LOADINGS FOR THE TWO FACTOR MODEL AFTER PRINCIPAL COMPONENT ANALYSIS (PCA) AND CONFIRMATORY FACTOR ANALYSIS (CFA)

Item Content	PCA Australia (USA)		CFA Australia only			
	Boys and Girls		Boys		Girls	
	PD	PWB	PD	PWB	PD	PWB
Low spirits	.65 (.70)	-	.80	0	.85	0
Down in dumps	.64 (.58)	-	.78	0	.84	0
Rattled, upset	.72 (.67)	-	.78	0	.82	0
Nervousness	.81 (.69)	-	.76	0	.62	0
Moody, brooded	.73 (.68)	-	.74	0	.78	0
Nothing turned out	.62 (.42)	- (-.32)	.74	0	.73	0
Downhearted, blue	.61 (.69)	-	.73	0	.83	0
Anxious, worried	.73 (.69)	-	.73	0	.77	0
Tense, high strung	.73 (.73)	-	.73	0	.75	0
Difficulty calm down	.73 (.64)	-	.73	0	.70	0
Nothing look forward to	.53 (.43)	.31 (-.42)	.73	0	.68	0
Felt depressed	.57 (.65)	-	.72	0	.74	0
Very nervous person	.82 (.73)	-	.71	0	.59	0
Restless, fidgety	.73 (.59)	-	.67	0	.67	0
Strain, stress, pressure	.59 (.63)	-	.63	0	.72	0
Felt like crying	.60 (.67)	-	.59	0	.76	0
Hands shake	.66 (.56)	-	.53	0	.44	0
Nervous or jumpy	.67 (.64)	-	.51	0	.39	0
Enjoyed things	-	.75 (.74)	0	.77	0	.78
Happy person	-	.70 (.70)	0	.74	0	.82
Calm and peaceful	-	.63 (.60)	0	.72	0	.78
Cheerful, lighthearted	-	.72 (.72)	0	.71	0	.85
Relaxed, free of tension	-	.60 (.70)	0	.69	0	.74
Happy, satisfied	-	.74 (.53)	0	.66	0	.74
Interesting day	-	.77 (.67)	0	.66	0	.59
Relax without difficulty	-	.55 (.61)	0	.64	0	.75
Living wonderful	-	.78 (.67)	0	.63	0	.73
Daily life interesting	-	.71 (.71)	0	.63	0	.67
Future hopeful	-	.71 (.64)	0	.61	0	.59
Wake up fresh, rested	-	.57 (.60)	0	.57	0	.62

Note. For PCA $N = 878$ in Australia and 953 in USA. Loadings $> .3$.

For CFA $N = 449$ boys, 429 girls.

Taking into account the original conceptualization of one, two, and five factors by Veit and Ware (1983), Ostroff *et al.*'s (1996) rejection of the one factor and five factor models, factor loadings, cross-loadings, and fit indices in the current study, the two-factor model emerged as the model of choice. The main strength of this model is that it appears to be replicable across the two nations and across male and female adolescents in Australia.

Descriptive Statistics for the Australian Adolescent Sample

In the next step scale scores were computed for the 18 item Psychological Distress scale, the 12 item Psychological Well-Being scale as well as a 30 item Total Mental Health Index. The well-being items were reversed, so that all measures were scored in the direction of mental health, that is higher scores indicated less distress and greater well-being (cf. Table 3). Adolescent boys reported significantly less Psychological Distress ($t_{876} = 3.63, p < .001$), significantly more Psychological Well-Being ($t_{876} = 2.40, p < .02$) and consequently better overall mental health (Total MHI-30) than adolescent girls ($t_{876} = 3.48, p < .002$). These differences match the American results reported by Ostroff *et al.* (1996).

TABLE 3
MEANS AND STANDARD DEVIATIONS FOR THE 30-ITEM
MENTAL HEALTH INVENTORY

	Psychological Distress		Psychological Well-Being		Mental Health Index		N
	M	SD	M	SD	M	SD	
Boys	80.02	16.64	55.59	10.92	135.61	23.90	449
Girls	76.07	15.55	53.84	10.75	129.90	24.71	429
All	78.09	16.23	54.74	10.87	132.82	24.45	878

Reliability and Internal Consistency

The estimates of internal consistency for the Mental Health scales were high. Cronbach's coefficient alpha for the 18 Personal Distress items was .94 and for the 12 Personal Well-being items .92. One hundred and eighty five adolescents were retested after approximately 10 weeks. Test-retest correlations indicated moderate stability for reports of Psychological Distress ($r_{12} = .71$), Psychological Well-being ($r_{12} = .69$), as well as the 30-item Total MHI score ($r_{12} = .73$).

CONCLUSION

There was reasonable support for the conclusion that a two-factor model of Psychological Distress and Well-being can represent adolescents' responses to the MHI in Australia as well as in the USA. Further, this study found that this model applies similarly well to male and female adolescents in Australia, although the female factors were more correlated than the male ones. Ostroff *et al.* (1996) did not report separate analyses and therefore nothing can be said at this point in time regarding the similarity of the two factors for American boys and girls. A second issue arises in both countries and concerns the wording of the items. All positively worded items form one factor, whereas negatively worded items form the other factor. To avoid the criticism that the two-factor structure is simply a function of item wording further study is needed which demonstrates the distinct validity of both dimensions. A more general validity issue concerns the role of the scales in a broader network of constructs. Given that most of the validity information available for the MHI pertains to adults (e.g. Ware *et al.*, 1984), validity studies are now needed which specifically focus on the adolescent years.

The Emotional Ties scale was not tested in this study as the three relevant items were not included. Studies of social support and mental health should drop these items from the measurement of mental health anyway in order to prevent confounding. This study found little support for the differentiation of the negative factor (Personal Distress) into separate anxiety, depression, and loss of control factors. The absence of five items from the loss of control scale may have weakened this factor too much for it to surface as a distinct entity in the current study. Clearly, further study is needed to settle the question of whether such a factor can be isolated in adolescence. In the meantime Ostroff et al's (1996) finding, which did not support the five-factor solution, should be taken as the best available evidence that the loss of control factor is not distinct from the depression and anxiety factor in adolescence. Finally, both studies provide evidence that the differentiation by self-report of depression and anxiety is difficult during adolescence, a finding which is consistent with other research (e.g. Achenbach, 1991).

In the current study the MHI was well accepted and well understood by male and female adolescents. Given its reliability and known factor structure, the MHI may be useful for research in general adolescent populations and for cross-national projects which focus on the two main factors (Personal Distress and Personal Well-being) which have now been replicated. However, further study with adolescents in non-English speaking countries is still needed to assess the full extent of the cross-cultural generalizability of these two factors.

Footnote 1. Available from the authors on request.

REFERENCES

- Achenbach, T. M. (1991) *Manual for the Youth Self-report and 1991 Profile*. Burlington, VT: University of Vermont, Department of Psychiatry.
- Bentler, P. M. (1990) Comparative fit indices in structural models. *Psychological Bulletin, 107*, 238-246.
- Cassileth, B. R., Lusk, E. J., Strouse, T. B., Miller, D. S., Brown, L. L., Cross, P. A., & Tenaglia, A. N. (1984) Psychosocial status in chronic illness. *The New England Journal of Medicine, 311*, 506-511.
- Floyd, F. J., & Widaman, K. F. (1995) Factor analysis in the development and refinement of clinical assessment instruments. *Psychological Assessment, 7*, 286-299.
- James, L. R., Mulaik, S. A., & Brett, J. M. (1982) *Causal analysis: Assumptions, models and data*. Beverly Hills: Sage.
- Jöreskog, K. G., & Sörbom, D. (1993a) *PRELIS 2 User's Reference Guide*. Chicago, IL: Scientific Software International.
- Jöreskog, K. G., & Sörbom, D. (1993b) *LISREL 8 User's Reference Guide*. Chicago, IL: Scientific Software International.
- Liang, J., Wu, S. C., Krause, N. M., Chiang, T. L., & Wu, H. Y. (1992) The structure of the Mental Health Inventory among Chinese in Taiwan. *Medical Care, 30*, 659-676.
- Marsh, H. W., Balla, J. R., & Hau, K. (1996) An evaluation of incremental fit indices: A clarification of mathematical and empirical properties (pp.315-353). In G. A. Marcoulides and R. E. Schumacker (Eds.), *Advanced structural equation modeling: Issues and techniques*. Mahwah, NJ: Lawrence Erlbaum.
- 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.
- Ostroff, J. S., Woolverton, K. S., Berry, C., & Lesko, L. M. (1996) Use of the Mental Health Inventory with adolescents: a secondary analysis of the Rand Health Insurance Study. *Psychological Assessment, 8*, 105-107.
- Rosenthal, T. L., Downs, J. M., Arheart, K. L., Deal, N., Downs, A. F. D., & Rosenthal, R. H. (1991) Similarities and differences on five inventories among mood and anxiety disorder patients. *Behavior Research and Therapy, 29*, 239-247.
- Steiger, J. H., & Lind, J. C. (1980, May) *Statistically-based tests for the number of common factors*. Paper presented at the annual meeting of the Psychonomic Society, Iowa City, IA.
- Tucker, L. R., & Lewis, C. (1973) A reliability coefficient for maximum likelihood factor analysis. *Psychometrika, 38*, 1-10.
- Veit, C., & Ware, J. (1983) The structure of psychological distress and well-being in general populations. *Journal of Consulting and Clinical Psychology, 51*, 730-742.
- Ware, J. E., Davies-Avery, A., & Brook, R. H. (1980) *Conceptualization and measurement of health for adults in the Health Insurance Study* (Vol. 6, Analysis of relationships among health status measures; R-1987/6-HEW). Washington D.C.: U.S. Government Printing Office.
- Ware Jr., J. E., Manning Jr., W. G., Duan, N., Wells, K. B., & Newhouse, J. P. (1984) Health status and the use of outpatient mental health services. *American Psychologist, 39*, 1090-1100.