Research in Nursing & Health, 1987, 10, 185–195

The Derivation and Validation of Six Multidimensional Health Locus of Control Scale Clusters

Daniel L. Rock, Beth E. Meyerowitz, Stephen A. Maisto, and Kenneth A. Wallston

The purpose of this research was twofold. First, to determine if eight Multidimensional Health Locus of Control types hypothesized by Wallston and Wallston (1982) existed. Second, to assess the reliability, validity, and clinical utility of the control types. Two investigations were conducted based on research procedures designed for the discovery of clusters. The first study involved the derivation and replication of MHLC clusters with a sample of 400 healthy undergraduate men and women. The results of Study 1 suggested the existence of six MHLC clusters: (a) pure internal; (b) double external; (c) pure chance; (d) yea sayer; (e) nay sayer, and (f) believer in control. The sample for Study 2 consisted of 90 female undergraduate and graduate students. The six-cluster solution was replicated in this second cross-validation study. Furthermore, construct validity of the clusters was established through computer simulation. Finally, relationships between clusters and the Krantz Health Opinion Survey suggested the clusters have a theoretical as well as a empirical foundation. The implications of these findings for clinical practice, for the development of the health locus of control construct, and for future research on the MHLC Scale are discussed.

A number of investigators have emphasized the possible importance of Rotter's social learning theory (Rotter, 1954, 1975) in explaining factors that influence health-related behaviors. In particular, the generalized expectancy construct, locus of control (LOC), has been examined for its relevance in understanding fluctuations in health status and health behaviors. Early reviews of applications of the LOC construct to the health domain suggested that such applications might lead to theoretically and empirically meaningful results (see Strickland, 1978), but that specificity of the construct within the health domain might increase the construct's usefulness in predicting health behavior.

To improve the predictive ability of LOC measures, researchers have developed scales to assess specific beliefs regarding control over health, rather than generalized LOC beliefs (Wallston, Wallston, Kaplan, & Maides, 1976; Wallston, Wallston, & DeVellis, 1978). The Multidimensional Health Locus of Control Scale (MHLC) developed by Wallston et al. (1978) is one of the more widely used and psychometrically sound examples of these scales. The MHLC scale consists of three subscales representing internal, chance, and powerful others dimensions of LOC beliefs (Levenson, 1981). The internal scale (IHLC) assesses the degree to which an individual believes that his/her own behavior is responsible for health or illness; the chance scale (CHLC) assesses beliefs that an individual's level of health or illness is a function of luck, chance, fate, or uncontrollable factors; and the powerful others scale (PHLC) assesses an individual's beliefs that the degree of health or illness is determined by important figures such as physi-

Study two was supported in part by a Spencer Foundation grant to Beth E. Meyerowitz.
Requests for reprints should be sent to Daniel L. Rock, Department of Psychiatry, 534 Clinical Drive, Indiana University Medical Center, Indianapolis, Indiana 46223.

cians, other health professionals, parents, or friends.

Over the past several years numerous studies have been performed in which the MHLC scales have been used as independent, dependent, and correlational variables in research on a wide variety of medical disorders (e.g., Burish, Carey, Wallston, Stein, Jamison, & Lyles, 1984; DeVellis, DeVellis, Wallston, & Wallston, 1980; Horlick, Cameron, Firor, Bhalerao, & Baltzan, 1984; Ingle, Burish, & Wallston, 1984; Nagy, & Wolfe, 1983). In fact, health LOC is one of the more frequently measured individual difference variables in health psychology research. Clearly, the health LOC construct has captured the interest of researchers and clinicians in the health psychology area.

Despite the attention that this variable has received, empirical evidence is mixed concerning the importance of the LOC construct for explaining health-related behaviors (see reviews by Wallston, & Wallston, 1981, 1982). Wallston and Wallston (1982) offered a possible explanation for the inconsistent and modest findings relating to health LOC. They suggested that the way in which the construct has been conceptualized empirically may have masked meaningful results. Up to now, most research has focused on determining the relationships between the variables of interest (e.g., adherence to health care recommendations) and scores on each of the LOC scales (i.e., internal, powerful others, chance). This approach overlooks the possibility that an individual's pattern of responses across the three scales may be more predictive than his or her scores on each of the scales separately. For example, a person who is high on both internal and powerful others LOC may engage in different behaviors from a person who scores high on only one of those scales. To allow for such distinctions, Wallston and Wallston (1982) recommended individuals be categorized into one of eight possible types based on whether they score high or low on each of the three MHLC scales. Specifically, the following types have been postulated: (a) pure internal (high internal, low chance and powerful others); (b) pure powerful others external (high powerful others, low internal and chance); (c) pure chance external (high chance, low internal and powerful others); (d) double external (high chance and powerful others, low internal); (e) believer in control (high internal and powerful others, low chance); (f) yea-sayer (high on all scales); (g) nay-sayer (low on all scales); and (h) an unnamed type referred to as Type VI (high internal and chance, low powerful others). As the Wallstons pointed out, the last type, while possible empirically, is not meaningful conceptually and is unlikely to exist. A person with a Type VI configuration would be stating that he or she was internally oriented (i.e., having primary control over his or her health) yet simultaneously stating that *chance* factors are primarily responsible for his or her health status. The other seven types, however, do have intuitive appeal.

The purpose of this research was to determine whether these seven potentially meaningful health LOC configurations exist as types or clusters, and to assess their reliability, validity, and clinical utility. This typology could serve as a new taxonomy for describing more precisely the influence of LOC on health-related behaviors and may also provide an explanation for some of the mixed results of previous research. Moreover, these types may lead to more accurate and meaningful predictions of health-related behaviors.

The first of two studies was designed to discover and replicate the existence of MHLC typologies through the use of cluster analysis. Cluster analysis techniques are empirical procedures designed to create classifications from a data set by grouping together individuals who have similar response patterns. The techniques have been used successfully by a number of investigators to develop useful taxonomies (e.g., for examining the relationship between MMPI scores and reports of pain within health psychology (Bernstein, & Garbin, 1983; Bradley, Prokop, Margolis, & Gentry, 1978; Bradley, & Van der Heide, 1984; McGill, Lawlis, Selby, Mooney, & McCoy, 1983; Prokop, Bradley, Margolis, & Gentry, 1980; Wack, & Turk, 1984). To provide a conservative test of the types that result from Study 1, a second study was designed to cross-validate the cluster solution derived in Study 1 by use of another sample of subjects. Construct validity of the clusters and their external validity on health-related variables also were examined in Study 2.

STUDY 1—DERIVATION AND REPLICATION OF MHLC CLUSTERS

METHOD

Sample

The sample consisted of 400 undergraduate students enrolled in education at a northwestern university. The 190 males and 210 females were preservice teachers, primarily Caucasian, unmarried, middle- to upper-middle-class, and aged 19 to 31 years ($M = 21.5$, $SD = 4.9$). Subjects reported they were free of acute or chronic illnesses or major injury at the time of their participation. All participants were volunteers.

Instruments

Subjects were administered Form B of the Multidimensional Health Locus of Control (MHLC) Scales (Wallston et al., 1978). The MHLC consists of three 6-item scales, rated on six-point Likert scales ranging from strongly disagree to strongly agree. The three scales represent the internal, chance, and powerful others dimensions of LOC beliefs (Levenson, 1981).

Procedure

Cluster analysis is a form of numerical taxonomy. The term cluster analysis does not refer to one standard statistical procedure (e.g., analysis of variance). Instead, there are a variety of clustering strategies, at least 100, which are used to group subjects (e.g., animals, diagnoses, bacteria) across attributes. Cluster analysis is occasionally confused with factor analysis. The confusion arises from the fact that both procedures are used for grouping or classification. The primary conceptual distinction between the two procedures is that factor analysis (except for Q-type) groups attributes across subjects, whereas, cluster analysis groups subjects across attributes. Two of the most important purposes of clustering for health research are: (a) to identify natural clusters within a set of subjects; and (b) to develop useful (i.e., for description and prediction) conceptual models for classifying individuals. The primary disadvantage of cluster analysis for applied researchers is that they are faced with a variety of decisions (e.g., selection of subjects, selection of variables, selection of similarity matrix, selection of clustering procedures, problems of validation, and the process of interpretation) without the benefit of agreement among numerical taxonomy experts regarding how and which decisions should be made. The use of cluster analysis in the present study was in accordance with the strategies outlined by Blashfield (1980), and Morey, Blashfield, and Skinner (1983). Study 1 consisted of two phases: derivation and replication.

The *derivation phase* objectives were: (a) to develop a set of MHLC clusters that are statistically and substantively meaningful, and, (b) to perform the derivation analyses on two independent samples so that the similarity of solutions could be evaluated in the replication phase.

K-Means cluster analysis (Everitt, 1974) was used to identify homogeneous MHLC profile subgroups within the two samples. This clustering technique was selected because it groups together entities, in this case individuals, with similar profiles across attributes, in this case MHLC subscale scores. The starting value for the initial number of clusters to be extracted was set at seven. We excluded the Type VI cluster for reasons described above. By using a conceptually meaningful starting value, the probability of overfitting the data with statistically significant, but unreproducible, clusters is reduced. The euclidean distance measure was used as the indicator of goodness of fit. This metric was selected because it provides readily interpretable information about within and between cluster distances (e.g., average between cluster distance is greater than average within cluster distance).

To reduce the possibility of investigator bias in selecting the number and interpretation of clusters, an a priori set of decision rules was created to aid in the selection and interpretation of clusters. A cluster solution was accepted if each cluster was represented by at least 10 subjects and if the interpretation of MHLC subscale profiles within each cluster was unambiguous. The latter criterion was operationally defined a priori by setting minimum level and configuration criteria for the pattern of MHLC subscale scores. The basic level and configuration rules were: (a) pure internal [internal (I) > mean (M); chance (C), powerful others (P) < M]; (b) double external ($I < M$; $C, P > M$); (c) pure powerful others ($I < M$; $C < M$; $P > M$); (d) pure chance ($I < M$; $C > M$; $P < M$); (e) yea sayer ($I, C, P > M$); (f) nay sayer ($I, C, P < M$); and (g) believer in control ($I > M$; $C < M$; $P > M$). Criteria also were established for the distance between each MHLC subscale and its mean within each cluster. The criteria involved the magnitude and sign of the z score calculated for each MHLC scale-mean comparison (e.g., pure internal = $I > M$; $C, P < M$). A cluster met the criteria for an unambiguous interpretation if the three z scores within each cluster were in the predicted direction, and two of the three z score comparisons reached statistical significance, ($p < .10$, 1-tailed).

Meeting the stability objective involves performing the same set of analyses on two independent sets of data. Specifically, one type of reliability assessment for a cluster analysis involves assessing the degree of replicability of a cluster solution across independent samples. To prepare data for the replication phase, the sample of 400 subjects was randomly divided into two samples of 200 subjects each. All aforementioned cluster derivation, selection, and interpretation criteria were then applied in an identical manner in each sample.

Because reliability is a necessary condition for validity, a reliability or replication assessment of a cluster solution is necessary if subsequent validation studies are planned. The replication procedure used in Study 1 involved three steps. The 200 subjects in the first sample (selected in the derivation phase) were assigned to clusters derived from the 200 subjects in the second sample. Classifications to the second sample were made through the nearest centroid method (Morey et al., 1983). The assignment of subjects in the first sample to clusters in the second was then compared to the original cluster solution derived for the first sample.

The degree of similarity between cluster solutions (i.e., the degree of reliability) derived from the first and second samples was assessed with the modified Rand statistic (Morey & Agresi, 1984). The Rand, modified Rand, and *Kappa* statistics are the three most frequently used measures of the relative number of agreements concerning membership in a cluster between two solutions to the same data. The modified Rand was used in the present study because, of the three options, it provides the most conservative estimate of observed agreement. The modified Rand is relatively more conservative than the Rand and *Kappa* statistics because the modified Rand permits the researcher to compare several different classifications while adjusting for the contribution of chance to the observed agreement.

RESULTS AND DISCUSSION

Preliminary Data Analysis

To facilitate the interpretation of MHLC scale scores within the cluster analyses, all raw total scale scores were converted to T scores ($M = 50$, $SD = 10$). Prior to this conversion, a series of t-tests was performed to assess if males and females differed significantly on any of the three MHLC scales. Results revealed no significant differences between males and females on any of the scales. Consequently, MHLC scale scores from all 200 subjects were used to derive internal, chance, and powerful others T score conversions.

Derivation Phase

A seven-group solution was initially derived for each sample. Each cluster was represented by a unique configuration of MHLC subscale scores. Unfortunately, the seven-group solution failed to meet all of the derivation and interpretation decision rules. One of the seven clusters, pure powerful others, was represented by two subjects only. Interpretation of six of the seven clusters was clear, (i.e., pure internal, pure chance, double external, believer in control, yea-sayer, and nay-sayer), and consistent with six of the eight typologies posited by Wallston & Wallston (1982).

Because the entire set of decision rules was not satisfied, a subsequent hierarchical K-means cluster analysis was performed on both samples where five- and six-cluster solutions were derived instead of seven. The six-group solution in both samples was superior to the five- and seven-group typologies on all cluster derivation and interpretation decision rules. No cluster was represented by fewer than 10 subjects, and each cluster met the aforementioned operational definition for cluster interpretations.

Replication Phase

Based on (first sample) subjects' MHLC scores and the cluster centroids derived in the second sample, subjects in the first sample were assigned to clusters in the second sample. Their classifications in the second sample were compared to their original classifications in the first sample. The modified Rand statistic for this comparison was 0.8541, indicating the difference between the number of observed agreements and chance agreements was 85.41% of the maximum difference. Therefore, a high degree of reliability was achieved.

The replication phase showed a substantial degree of similarity between the cluster solutions in each sample. Consequently, the two samples of 200 subjects were each combined, and a third *k*-means cluster analysis was performed on the aggregated sample of 400 subjects.

The results of this analysis are presented in Table 1, where descriptive statistics are displayed for each MHLC subscale within each of the six clusters, (i.e., pure internal, double external, pure chance, yea sayer, nay sayer, and believer in control). Each of the six clusters was represented by substantially more than 10 subjects. The cluster interpretation criteria were applied to the six-group solution, and each of the six clusters met the criteria for unambiguous interpretation.

The results of Study 1 revealed a number of interesting findings. Application of a conservative cluster derivation and replication scheme developed a priori evidenced six MHLC subgroups that were substantively and statistically meaningful. The existence of relatively homogeneous subgroups suggests the three theoretical dimensions of health beliefs (internal, chance, powerful others) are relatively independent, and, consequently, may vary differently in elevation and pattern (profile). The homogenity of subgroups also indicates the variation between MHLC subscales within an individual is not random, (i.e., individuals with similar profiles can be grouped together in theoretically meaningful clusters that are stable and statistically reliable). In addition, individuals who report the same health status (absence of acute or chronic disease) maintain different patterns and levels of MHLC beliefs. Subjects' differential responses, while not experiencing any illness, suggests the relationship between MHLC health beliefs and health or illness status may not be isomorphic.

Although the subgroup findings of Study 1 have implications for the health LOC construct, the results thus far provide little information concerning the validity of the cluster solution. In keeping with the methodological literature on cluster analysis, a thorough initial cluster validity study would include cross validation as well as construct and external validation of the cluster solution. Cross validation addresses generalizability, in that it attempts to determine the extent to which a cluster solution derived from Sample X generalizes to Sample Y, where Sample Y was drawn from a population unrelated to Sample X. Construct validation in this context refers to information concerning the interpretabil-

Table 1. Means and Standard Deviations for Multidimensional Health Locus of Control (MHLC) Subscales by Cluster—Study 1

MHLC Cluster	MHLC Subscale Scores						Subjects in Cluster	
	Internal		Chance		Powerful Other			
Label	*M*	*SD*	*M*	*SD*	*M*	*SD*	Number	Percent of Sample
Pure Internal	63.90	5.83	40.46	6.59	41.08	7.15	112	28.00
Double External	42.52	4.21	55.58	5.33	61.62	6.84	56	14.00
Pure Chance	37.22	6.09	60.01	6.81	47.98	5.36	64	16.00
Yea Sayer	54.20	3.60	56.46	5.89	58.03	5.38	72	18.00
Nay Sayer	41.75	5.71	40.73	5.53	39.79	6.45	56	14.00
Believer in Control	64.60	5.33	45.64	5.58	60.47	6.53	40	10.00
						Total	400	100.00

Note: MHLC scale values were converted to *T* scores ($M = 50$, $SD = 10$)

ity of the clusters and addresses the question: Are these clusters a result of relationships between MHLC subscales, the data, or a product of the computing algorithm such that the same clusters could be derived from any data set? External validation involves examining the relationship between a cluster solution and additional variables theoretically related to the cluster variables. For example, the health locus of control construct theoretically is connected to a variety of health related attitudes, beliefs, expectancies, values, and behaviors; yet to date, only limited empirical evidence is available to support these propositions (Wallston, Smith, King, Forsberg, Wallston, & Nagy, 1983; Wallston, & Wallston, 1982). A substantial increment in external validity could be documented for the health LOC construct if a MHLC cluster solution could produce predictions of other health beliefs that were more precise than predictions derived through the traditional method of analysis (i.e., zero-order correlations between each MHLC scale and a measure of other health beliefs). Consequently, the purpose of Study 2 was to conduct a cross, construct, and external validation of the MHLC cluster solution derived in Study 1.

STUDY 2—CROSS VALIDATION AND EXTERNAL VALIDATION OF MHLC CLUSTERS

METHOD

Sample

A sample of 90 women was recruited from graduate and undergraduate classes at a southeastern university. All subjects were volunteers participating in a larger study of reactions to cancer in a parent. Half of the participants had a parent who was diagnosed as having cancer and the remaining half had no immediate family members with cancer. Subjects ranged in age from 18–30 years (M = 20.5, SD = 2.8).

Instruments

In addition to the MHLC scale, subjects were administered the Krantz Health Opinion Survey (KHOS) (Krantz, Baum, & Wideman, 1980). In the present study the 16 KHOS items were interspersed with the MHLC items to form a 34-item questionnaire. Subjects responded to each of the 16 KHOS items on a six-point scale ranging from strongly disagree–strongly agree. The measure was designed to assess subjects' attitudes toward different approaches to medical treatment (e.g., self-care, professionally assisted self-care, professional care). The KHOS contains two subscales: a nine-item Behavioral Involvement subscale assessing ". . . attitudes toward self-treatment and active behavioral involvement of patients in medical care," and a seven-item subscale called information which measures ". . . the desire to ask questions and wanting to be informed about medical decisions" (Krantz et al., 1980, pp. 978–979). The KHOS was selected for this study because factor analytic studies (e.g., Wallston et al., 1983) have revealed a moderately high degree of overlap between the MHLC and KHOS. The communality between the measures provides a sensitive test of the external validity of the cluster solution.

Procedure

Assessment of the generalizability of the six types involved three stages (Blashfield, 1980; Morey et al., 1983). First, MHLC scores from subjects in Study 2 were analyzed using k-means cluster analysis to derive a six-group solution. Through an independent set of analyses, subjects in Study 2 were assigned to the clusters obtained in Study 1. Classifications were made using the nearest centroid method. The degree of convergence between the solutions derived from Study 1 and Study 2 was then estimated using the modified Rand statistic. A high degree of convergence between solutions obtained from independent samples would provide evidence in support of the cross-validation of the six groups for these samples.

Construct Validation. The results of the cluster analysis in Study 1 were interpreted as supporting a six-cluster solution. The purpose of examining construct validity was to assess the strength of the evidence in support of the six-cluster solution by comparing the clusters derived from the empirical data to those derived from computer simulation. Use of computer simulation was necessary to determine if the six clusters derived in Study 1 and Study 2 were the result of chance or a characteristic of the cluster algorithm. That is, if the cluster solution derived from empirical data is not significantly different from the cluster solution derived from normally distributed, but random data (i.e., data generated from the computer simulation), then no legitimate claim can be made for the construct validity of the six-cluster solution. Using k-means analysis, a six-group solution derived from computer simulated data was compared to the cluster solution derived from the empirical data presented in Study 2. The simulation study was performed to create a cluster solution that embodied the null hypothesis. That is, in the simulation, data were generated such that only one cluster existed in the data set. Then using k-means analysis, a six-group solution was derived from the simulated data. If the solutions derived from the empirical and simulated data were comparable with respect to interpretability, then the construct validity of the clusters derived from the empirical data would be challenged.

While replicability, cross and construct validity are important psychometric properties of a cluster solution, the relationship of the clusters to a clinically meaningful set of health beliefs is important for demonstrating the clinical utility of the cluster solution. If the clusters are not related to changes in health behaviors, or if the solution adds nothing to the manner in which patients are treated as a function of their MHLC scores, then the utility of the cluster solution is questionable. In this phase of the study, two facets of external validity were investigated. First, we hypothesized that the six-cluster solution would have a significant relationship with health information seeking and behavioral involvement in health care as measured by the KHOS. Second, we hypothesized that as compared to the cluster-analytic solution described in Study 2, the frequently cited method of correlating (i.e., zero order correlations) MHLC subscales with external validity criteria (e.g., KHOS) would lead to more frequent erroneous predictions concerning the health information seeking and health care involvement preferences of subjects.

RESULTS

Cross Validation

The k-means cluster analysis of the MHLC data yielded a six-group cluster solution. Interpretation of the types derived in Study 2 was consistent with the description of the types derived in Study 1. As presented in Table 2, the clusters are stable (i.e., the number of subjects in each cluster and the magnitude of the standard deviations relative to the means met the a priori specified criteria).

Subjects in Study 2 were assigned to one of six MHLC types independent of the aforementioned cluster analysis and by the use of the cluster solution derived in Study 1. The degree of similarity between the results of the initial cluster analysis in Study 2 and the classification of MHLC scores based on the Study 1 solution was subsequently assessed. The results of this analysis demonstrated that the two solutions were highly similar (modified Rand statistic = .893), suggesting that the six-group MHLC cluster solution is equivalent in two independent samples.

Table 2. Means and Standard Deviations for Multidimensional Health Locus of Control (MHLC) Subscales by Cluster—Study 2

MHLC Cluster	MHLC Subscale Scores						Subjects in Cluster	
	Internal		Chance		Powerful Other			
Label	*M*	*SD*	*M*	*SD*	*M*	*SD*	Number	Percent of Sample
Pure Internal	59.19	4.85	42.11	7.05	42.01	5.79	22	22.44
Double External	38.20	6.62	55.12	6.45	58.21	5.02	18	20.00
Pure Chance	47.18	2.09	50.18	7.39	45.07	3.09	11	12.22
Yea Sayer	54.54	5.93	58.96	7.20	58.89	4.17	13	14.44
Nay Sayer	45.52	5.88	45.30	5.44	41.41	6.12	17	18.90
Believer in Control	59.36	5.82	42.54	5.83	62.55	3.81	9	10.00
						Total	90	100.00

Note: MHLC scale values were converted to T scores (M = 50, SD = 10)

Construct Validation

The computer simulation phase involved creating three distributions of scores (one each for: internal, chance, and powerful others MHLC subscales), and the derivation of six clusters from the simulated data using K-means cluster analysis (Dixon, 1981). Three independent, normally distributed, standardized ($M = 0$, $SD = 1$), samples of data ($N = 90$) with known statistical parameters were created. Each distribution represented one of three distributions of MHLC data such that each of the 90 entities had a score on each of the subscales.

To determine if the six clusters derived from the empirical data in Study 2 resulted from chance or some characteristic of the cluster analysis computing algorithm, it was necessary to derive six clusters from the simulated data and apply the aforementioned cluster classification rules [i.e.: (a) the three z scores within each cluster were in the predicted direction; (b) two of the three z scores reached statistical significance ($p < .10$ 1-tailed)] to the clusters derived from that data. Results of the K-means analysis yielded six clusters of which five failed both criteria and one failed the first but passed the second criterion.

In summary, the analysis comparing the cluster solutions from the simulated and empirical data support the conclusion that the clusters derived from Study 2 evidence an adequate level of construct validity, and therefore, are not a result of chance or some peculiarity of the computing algorithm.

External Validation

Using cluster membership as the independent variable, a multivariate analysis of variance was performed on the vector of KHOS means (i.e., information seeking and behavioral involvement). As predicted, differences in health information seeking and behavioral involvement in personal health care varied significantly as a function of cluster membership [F (10, 166.00) = 108.91, $p < .0001$]. Thus, as indicated in Table 3, this analysis[1] indicated that differences in health behaviors, as measured by the KHOS, can be predicted from MHLC cluster membership.

To determine which method of analysis (i.e., cluster or zero-order correlations) yielded the most clinically and theoretically meaningful interpretation of the relationship between MHLC and KHOS, it was necessary to perform the following analyses. First, the MHLC data from Study 2 were aggregated across the six clusters, and the MHLC and KHOS subscales were intercorrelated. The purpose of this analysis was to assess the degree of relationship between the MHLC subscales and health behaviors as measured by the KHOS. No significant correlations ($p < 0.05$) were found between the MHLC subscales and the information-seeking subscale of the KHOS. However, two of the three correlations between the MHLC scale and KHOS behavioral involvement scale were significant (Internal with Behavioral Involvement, $r = 0.31, p < 0.05$; powerful others with behavioral involvement, $r = -0.479$, $p < 0.05$). The correlation between

Table 3. Means and Standard Deviations for Krantz Health Opinion Survey (KHOS) Subscales by Multidimensional Health Locus of Control (MHLC) Clusters Ranked in Decending Order of Magnitude of Mean for Each KHOS Subscale

	KHOS Subscale				
	Information Seeking		Behavioral Involvement		
MHLC Cluster	*M*	*SD*	*M*	*SD*	MHLC Cluster
Pure Chance	32.8	5.1	37.7	5.0	Yea Sayer
Double External	31.8	7.4	33.3	7.8	Pure Chance
Yea Sayer	31.1	6.7	32.8	6.1	Pure Internal
Pure Internal	28.6	3.8	32.6	4.4	Double External
Nea Sayer	28.6	6.7	29.4	5.3	Nea Sayer
Believer in Control	26.7	8.6	29.3	7.4	Believer in Control

internal and powerful others was nonsignificant, indicating that each subscale taps a different dimension of health beliefs, yet both subscales are moderately correlated with behavioral involvement in health as measured by the KHOS.

This finding is difficult to interpret. What could explain the result that internal MHLC is moderately positively correlated with behavioral involvement, and powerful others MHLC is moderately negatively correlated with behavioral involvement, yet internal and powerful others are uncorrelated? One interpretation is that the behavioral involvement subscale is factorially complex, and that internal and powerful others subscales are assessing the divergent dimensions within behavioral involvement subscale. However, if this explanation were true, one would expect a factor analysis of the KHOS to demonstrate that the behavioral involvement subscale was other than unidimensional. The psychometric data presented by Krantz et al., (1980) reveal that the behavioral involvement subscale is not factorially complex, therefore, some other explanation is needed to account for this finding.

Another interpretation is that naturally occuring homogeneous subgroups of subjects exist within the data. If the relationships between MHLC and KHOS variables were different for different subgroups, then one could explain this unlikely finding. That is, while the relationship between the MHLC and KHOS is consistent across MHLC subgroups, the relationship between MHLC subscales within each group is different. Therefore, the existence of a non-normal and multimodal multivariate distribution attenuates the correlation between internal and powerful others MHLC. If this explanation were true, then continued use of zero-order correlations could not be supported for the MHLC and KHOS because more variance could be accounted for by examining the relationship between MHLC subgroups and the KHOS, than by simply intercorrelating individual MHLC and KHOS subscale scores. The purpose of the next set of analyses was to examine the relationship between MHLC subgroups and the behavioral involvement subscale.

If predictions of KHOS behavioral involvement scores based on the cluster analysis are more precise than predictions based on the zero-order correlations of internal and powerful others with behavioral involvement, then significant differences in behavioral involvement scores should be found between clusters that are high on Internal (i.e., pure internal, yea-sayer, believer in control), and between clusters that are low on powerful others (i.e., pure internal, nay-sayer, pure chance). Essentially, these analyses test the hypothesis that considering differences in pattern and level of MHLC subscale scores results in more precise predictions of KHOS health behaviors than predictions based on individual zero-order correlations of each MHLC subscale with behavioral involvement.

As predicted, inclusion of the chance and powerful others subscale scores in clusters with high internal scores yielded significant differences in behavioral involvement between pure internal, yea sayer, and believer in control clusters [F (2,41) = 7.271, $p <$ 0.003]; and, inclusion of internal and chance subscale scores in clusters with low powerful others scores yielded significant differences in behavioral involvement between pure internal, nay sayer, and pure chance clusters [F (2,47) = 3.273, $p < 0.03$]. Thus, differences in behavioral involvement in health care as measured by the KHOS are more accurately predicted from a vector of MHLC subscales in a cluster formation than predictions based on an individual MHLC subscale score (i.e., internal or powerful others).

CONCLUSIONS

Data from these studies support the existence of six of the MHLC subgroups initially hypothesized by Wallston, & Wallston (1982). The cluster solution displayed acceptable levels of reliability, as well as construct and external validity. The high degree of replicability for the six-cluster solution as demonstrated in Study 1, suggests that these six clusters are stable and can be reliably derived from healthy undergraduate student populations in the future. Data from the construct validity phase of Study 2 revealed that the six clusters derived from the empirical data were substantially different from the clusters derived from the simulated data. Thus, the clusters derived from the empirical data are likely to have occurred because of systematic rather than random variation in the pattern and elevation of MHLC subscales scores. Similarly, the results of the external validation phase of Study 2 revealed, as compared to zero-order correlations, the cluster solution can provide greater precision in predictions of behavioral involvement in personal health care. This finding suggests, for example, that a clinician could use knowledge of a patient's cluster membership as a predictor of patient interest and ability to perform self-care behaviors.

The lack of empirical evidence in support of the pure powerful others cluster (i.e., low internal & chance, high powerful other) is not surprising given the "healthy" nature of our subject sample. Unlike individuals diagnosed with a serious chronic disease, these subjects had no need to rely solely on a health care professional for the maintenance of their current level of health. In this regard, Wallston, and Wallston (1981) reported that the mean powerful other MHLC score for a group of cancer patients undergoing treatment was 1.5 standard deviations higher than the mean powerful others MHLC score for a group of healthy individuals. Thus, the lack of evidence for the pure powerful other cluster is more likely a function of the healthy nature of our sample, than ambiguities in the definition of the pure powerful other cluster per se.

These findings can benefit future MHLC research in a number of ways. For example, the procedures used in these studies can serve as a new heuristic for formulating questions concerning the MHLC construct. Since the statistical independence of the MHLC subscales can result in complex patterns of relationships among subscales (cf., MMPI), differences in patterns and elevations of clusters may account for more variance in an outcome measure than simply considering the separate contribution of each subscale as has been done in the past. While this finding is beneficial to the development of a MHLC theory, the presence of homogeneous subgroups within a data set tends to attenuate statistical relationships with other variables (Nunnally, 1976). Before the present studies were completed, MHLC profiles usually were not considered in empirical research using the MHLC scale. Therefore, according to the results of these studies, it is possible that investigators who previously conducted MHLC research may have obtained attenuated statistical relationships due to the existence of MHLC clusters in their data.

The most important conclusion to be drawn from these studies is that previous conceptualizations of the MHLC construct may have oversimplified its complexity, as there seem to be at least six robust clusters that are consistent with more recent MHLC theoretical speculations (Wallston, & Wallston, 1982). Although the clusters in the present studies have acceptable degrees of reliability and validity, there is no evidence to suggest that these same clusters exist in all populations of subjects. In fact, as mentioned previously, it is likely that different clusters may exist in different populations, and that the clusters may vary as a function of the health status (e.g., absence or presence of a disease, type and severity of the disease, factors influencing the delivery of treatment) of that population. An objective of future research should be to identify these clusters and clarify their relationship with an individual's health attitudes, beliefs, and behaviors.

REFERENCES

Bernstein, I. H., & Garbin, C. P. (1983). Hierarchical clustering of pain patients' MMPI profiles: A republican note. *Journal of Personality Assessment, 47,* 171–172.

Blashfield, R. K. (1980). Propositions regarding the use of cluster analysis in clinical research. *Journal of Consulting and Clinical Psychology, 48,* 456–459.

Bradley, L. A., Prokop, C. K., Margolis, R., & Gentry, W. D. (1978). Multivariate analyses of the MMPI profiles of low back pain patients. *Journal of Behavioral Medicine, 1,* 253–272.

Bradley, L. A., & Van der Heide, L. H. (1984). Pain-related correlates of MMPI profile subgroups among back pain patients. *Health Psychology, 3,* 157–174.

Burish, T. G., Carey, M. P., Wallston, K. A., Stein, M. J., Jamison, R. N., & Lyles, J. N. (1984). Health locus of control and chronic disease: An external orientation may be advantageous. *Journal of Social and Clinical Psychology, 2,* 326–332.

DeVellis, R. F., DeVellis, B. M., Wallston, B. S. & Wallston, K. A. (1980). Epilepsy and learned helplessness. *Basic and Applied Social Psychology, 1,* 241–253.

Dixon, W. (Ed.), (1984). *BMDP Biomedical computer programs.* Los Angeles: University of California Press.

Everitt, B. S. (1974). *Cluster analysis.* London: Halstead Press.

Horlick, L., Cameron, R., Firor, W., Bhaleroo, U., & Baltzan, R. (1984). The effects of education and group discussion in the post myocardial infarction patient. *Journal of Psychosomatic Research, 28,* 485–492.

Ingle, R. J., Burish, T. G., & Wallston, K. A. (1984). Conditionability of cancer chemotherapy patients. *Oncology Nursing Forum, 11*(4), 97–102.

Krantz, D. S., Baum, A., & Wideman, M. (1980). Assessment of preferences for self-treatment and information in health care. *Journal of Personality and Social Psychology, 39,* 977–990.

Levenson, H. (1981). Differentiating among internality, powerful others, and chance. In H. Lefcourt (Ed.), *Research with the locus of control construct* (Vol. 1, pp. 15–63). New York: Academic Press.

McGill, J., Lawlis, F., Selby, D., Mooney, V., & McCoy, C. W. (1983). The relationship of Minnesota Multiphasic Personality Inventory (MMPI) profile clusters to pain behaviors. *Journal of Behavioral Medicine, 6,* 77–92.

Morey, L. C., Blashfield, R. K., & Skinner, H. A. (1983). A comparison of cluster analysis techniques within a sequential validation framework. *Multivariate Behavioral Research, 18,* 309–329.

Nagy, V. T., & Wolfe, G. R. (1983). Chronic illness and health locus of control beliefs. *Journal of Social and Clinical Psychology, 1,* 58–65.

Nunnally, J. C. (1978). *Psychometric theory* (2nd. Ed.). New York: McGraw-Hill.

Prokop, C. K., Bradley, L. A., Margolis, R., & Gentry, W. D. (1980). Multivariate analyses of the MMPI profiles of multiple pain patients. *Journal of Personality Assessment, 44,* 246–252.

Rotter, J. B. (1954). *Social learning and clinical psychology.* Englewood Cliffs, NJ: Prentice-Hall.

Rotter, J. E. (1975). Some problems and misconceptions related to the construct of internal vs. external control of reinforcement. *Journal of Consulting and Clinical Psychology, 43,* 56–57.

Strickland, B. R. (1978). Internal-external expectancies and health-related behaviors. *Journal of Consulting and Clinical Psychology, 46,* 1192–1211.

Wack, J. T., & Turk, D. C. (1984). Latent structure of strategies used to cope with nociceptive stimulation. *Health Psychology, 3,* 27–43.

Wallston, B. S., Wallston, K. A., Kaplan, G. D., & Maides, S. A. (1976). Development and validation of the health locus of control (HLC) scale. *Journal of Consulting and Clinical Psychology, 44,* 580–585.

Wallston, K. A., Wallston, B. S., & DeVellis, R. (1978). Development of the Multidimensional Health Locus of Control (MHLC) scales. *Health Education Monographs, 6,* 160–170.

Wallston, K. A., & Wallston, B. S. (1981). Health locus of control scales. In H. Lefcourt (Ed.), *Research with the locus of control construct* (Vol. 1, pp. 189–243). New York: Academic Press.

Wallston, K. A., & Wallston, B. S. (1982). Who is responsible for your health? The construct of health locus of control. In G. Sanders & J. Suls (Eds.), *Social psychology of health and illness.* (pp. 65–95), Hillsdale, NJ: Erlbaum.

Wallston, K. A., Smith, R. A., King, J. E., Forsberg, P. R., Wallston, B. S., & Nagy, V. T. (1983). The relationship between expectancy for control of one's health and desire for control of one's health care. *Personality and Social Psychology Bulletin, 9,* 377–385.

FOOTNOTES

[1] Subsequent univariate analyses of variance of individual cluster means for Information Seeking and Behavioral Involvement were not performed. The primary purpose of this research was to emphasize the psychometric, rather than substantive, nature of the clusters. Presently, no adequate theory or rational basis is available for making specific a priori predictions of univariate analyses of variance for individual cluster means of Information Seeking and Behavioral Involvement subscales. In our opinion, the latter questions are issues for future research after additional data concerning the reliability and validity of MHLC clusters are available.