A STATISTICAL EXAMINATION OF THE HEALTH UTILITY INDEX AS A MEASURE OF HEALTH STATUS FOR USE IN THE NATIONAL POPULATION HEALTH SURVEY

Chris Richardson

B. Sc.(Hon.), University of Guelph, 1993

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE

REQUIREMENTS FOR THE DEGREE OF MASTER OF SCIENCE

in

COMMUNITY HEALTH

THE UNIVERSITY OF NORTHERN BRITISH COLUMBIA

January 1999

Abstract

The purpose of this study was to examine the ability of the Health Utility Index Mark III (HUI-Mark III) to provide a summary measure of health status for the 1994 National Population Health Survey (NPHS). The sample consisted of 838 (402 male, 436 female) randomly selected residents of Prince George, British Columbia. An exploratory factor analysis was conducted using the measures of physical, psychological and social well-being included in the NPHS. The saved factor scores were then regressed on to the HUI-Mark III in an attempt to determine the proportion of variance in the HUI-Mark III accounted for by the factor scores. The results suggest the HUI-Mark III is unable to discriminate between the many different levels of positive health experienced by the vast majority of the general population and that the HUI-Mark III is more or less insensitive to variation in key indicators of mental well-being.

TABLE OF CONTENTS

List of Tables

List of Figures

The realisation that universal, easily accessible medical care does not automatically abolish health inequalities together with the rising costs of today's health care systems have forced health planners and evaluators to critically examine current approaches to achieving a healthy population (Marmor, Barer & Evans, 1994). Traditionally, these evaluations have been based on mortality and morbidity data acquired through vital statistics and hospital databases. However, the increasing realisation that behavioural and social factors represent the keys to preventing or controlling today's leading causes of death and disability has resulted in information needs that cannot be met by existing mortality and morbidity information sources (Green, 1990).

In response to these new information requirements the health measurement field has opened up tremendously. There are now hundreds of measures that can be used to monitor the many different aspects of human health and well-being (Millar & Hull, 1997). While a large variety of specialised measures have been developed, the lack of a generally accepted summary measure of population health has forced researchers to seek comprehensiveness through the use of multiple instruments (Read, Quinn & Hoefer, 1987). Unfortunately, evidence from one measure often appears to conflict with that of other measures. Such situations frequently pose major problems for decision makers, who typically develop solutions based on informal judgements about the relative importance of one measure over another (Read et al., 1987).

Another, more fundamental problem with current sources of health information lies in the substantial imbalance between the high availability and comprehensiveness of data describing the resources used by health care systems and the relatively limited amount of data on the

health and psycho-social status of the populations being served. Such inconsistent approaches to information gathering have contributed to the belief that much of the existing data and information on health and health care is unreliable, fragmentary and incongruent (Wolfson, 1994). As a result, today's health planners require no less than the development of a new integrated system of health statistics that includes reliable measures of population health status (Wolfson, 1994; Health Canada, 1998).

In response to the demand for population level health information, Statistics Canada initiated the National Population Health Survey (NPHS) (Statistics Canada, 1995). In addition to assessing the health status of Canadians, the NPHS was also designed to gather information on many potential determinants of health. To accomplish these goals the survey examined a large number of indicators associated with the physical, psychological and socio-economic well-being of Canadians. The instrument selected to provide an overall measure of individual health status was the Health Utility Index-Mark III (HUI-Mark III) developed by Torrance, Zhang, Feeny, Furlong & Barr (1992). Data collection for the first wave began in June 1994 and finished in June 1995.

The purpose of this thesis was to examine the ability of the HUI-Mark III to provide a summary measure of health status compatible with the information objectives of the NPHS. To accomplish this goal, the investigation identified the dimensions of health measured by the NPHS, and then examined the ability of the HUI-Mark III to represent said dimensions. This information was then used in conjunction with current research on population health

assessment to comment on the appropriateness of the HUI-Mark III as an indicator of health status for the NPHS data.

II: LITERATURE REVIEW

The primary consideration when choosing a measure of health status is ensuring that the information provided by the measure corresponds to the goals of the study (i.e. will the data provided by the measure address the necessary research questions?). The first step in the selection of an appropriate measure is therefore inevitably tied to two inter-linked issues: how 'health' is defined and the purpose for which the measure is required (Noack & Abelin, 1987). Once a set of information requirements has been developed, including a clear definition of health, the next step is to generate a list of compatible measures. The most appropriate instrument may then be chosen by assessing the practicality, reliability and validity of each potential measure.

A Definition of Health

Clearly, attempts to measure health status must be founded in some concept of what constitutes health, or in the conditions, signs and experiences that might indicate ill-health (Hunt et al, 1986). For the past 150 years, rising expectations have changed the definition of health in the developed world from "survival" to "freedom from disease" to "an ability to perform daily activities" to "a sense of happiness and well-being." (Larson, 1991). Health and human wellness are now often used interchangeably and there is an increasing trend towards

the inclusion of more positive indicators of health status, such as self-esteem, in research on health and its determinants (Millar & Hull, 1997).

Fortunately, within the field of community health a degree of consensus exists around the three dimensional conception of health offered by the World Health Organization (WHO) (Bergner & Rothman, 1987; British Columbia. Provincial Health Officer; 1994; Wolinski & Zusman, 1980). In this definition, health is considered a " state of complete physical, mental, and social well-being and not merely the absence of disease or infirmity" (WHO, 1948). In annual reports on the health of British Columbians, the Provincial Health Officer explicitly states that the measures and data used for assessing population health are dependent on the above mentioned WHO definition (British Columbia Provincial Health Officer, 1994). While there is general acceptance of the WHO definition, there remains considerable disagreement over the operationalization of the accompanying dimensions, with the supporters of one dimension often pursuing its measurement at the expense of the remaining dimensions (Wolinski & Zusman, 1980).

Objectives of National Population Health Survey

The NPHS is a longitudinal survey, and will continue to collect data from the same panel of respondents every two years for up to two decades. Data collection for the first wave began in June 1994 and finished in June 1995.

As stated by Catlin and Will (1992, p. 313), the specific objectives of the NPHS are:

a) To aid in the development of public policies designed to improve health, by providing measures of the level, trend and distribution of the health status of the population.
b) To provide data for analytic studies that will assist in understanding the determinants of health.
c) To collect data on the economic, social, demographic, occupational and environmental correlates of health.
d) To increase the understanding of the relationship between health status and the use of health services, not only in the traditional sense, but also in areas such as home care, self-medication and self-care.
e) To provide panel data that will reflect the dynamic process of health and illness and produce periodic cross-sectional estimates.

The Determinants of Health

A determinant of health is not a component of health per se, but rather is a factor known to influence one or more aspects of health. For example, poverty has been linked to many forms of premature mortality and morbidity (Millar, 1994). One of the first widely accepted reports offering a conceptual framework for the determinants of health was the Lalonde Report (Canada, 1974). In this report, the key factors identified as health determinants were lifestyle, environment, human biology and health services. Since then, a growing amount of research

has supported and at the same time revised or expanded upon the health determinants identified in the Lalonde Report (Health Canada, 1994).

More recently, the British Columbia Ministry of Health (1994) identified the following factors as important determinants of health: the physical environment, human biology, individual behaviours, economic factors such as employment, poverty levels, income distribution, and social factors such as the care of infants and children, education, housing, social supports, crime, discrimination, violence, and abuse. Furthermore, the Ministry concluded that the combined influence of the above mentioned determinants has considerably more impact on the health of Canadians than the services provided by the existing medical care industry (British Columbia. Provincial Health Officer, 1994).

The HUI-Mark III as a Measure of Health Status

Central to all of the NPHS objectives is a summary measure of individual health status, both as a descriptor and to facilitate research into health determinants. In light of these objectives, researchers at Statistics Canada selected the HUI-Mark III to provide "a realistic appraisal of individual health status" for use in the NPHS (Statistics Canada, 1995, p. 28).

Over the past two decades, researchers at the Centre for Health Economics and Policy Analysis at McMaster University have put forth a series of three measures of health status that use multi-attribute utility functions to generate a Health Utility Index (HUI) (Torrance et al., 1992). There are two main components to the HUI approach. The first is a health status

classification system that describes functioning on a number of different health attributes (e.g. cognition and eyesight). Each attribute is composed of several different functional levels, so that at any point in time, each individual can be classified into only one level. Each different combination of levels, one from each attribute, thus represents a unique health state (Berthelot, Roberge & Wolfson, 1992).

The second component in the HUI approach involves integrating the health states obtained from the population under study with predetermined societal preferences for said states to produce an overall index score for each respondent. In this way, the index score is said to represent a measure of health status embodying the current views of society (Statistics Canada, 1995).

While the large number of potential health states associated with multi-attribute systems prevents the direct measurement of preferences for all possible outcomes (e.g., the seven attribute HUI-Mark II health status classification system contains a total of 24 000 different states), the application of multi-attribute utility theory enables the prediction of multi-attribute outcome preferences as a function of the preferences of each attribute measured individually (Torrance, 1982). The basic approach is to develop preference functions for each attribute separately and then determine an equation that expresses the overall preference score as a function of the single attribute scores (Torrance et al., 1992). The total number of preferences required to develop the utility function is therefore determined by the sum and not the product of each attribute's number of levels. This substantially reduces the number of preferences required to develop the utility function.

The first index in the series, the HUI-Mark I, was designed to evaluate outcomes associated with neonatal intensive care of very-low-birth-weight infants. Health status was classified using the following four attributes: physical function, role function, socio-emotional function and health problem, each with four to eight levels of functioning (Fenny et al., 1992). The second index in the series, the HUI-Mark II, was developed for use in a cost-utility analysis of childhood cancer treatments (comparing aggressive with intent to cure versus palliative treatments). It described health status using the following set of attributes: sensation, mobility, emotion, cognition, self-care, pain and fertility. (see Appendix A for a complete description of HUI-Mark II attribute levels). The third and most recent index, the HUI-Mark III, has been labelled a measure of functional health status for the general population. The eight attributes selected to describe health status for the HUI-Mark III are vision, hearing, speech, ambulation, dexterity, emotion, cognition and pain (see Appendix A for a description of attribute levels). Over and above its use in the 1994/5 NPHS, the HUI-Mark III has been included in such population level surveys as the 1990 Ontario Health Survey (n = 61 239) and cycle 6 of the 1991 Canada-wide General Social Survey (11 924 households) (Boyle, Torrance, Sinclair & Horwood, 1983).

Despite being extensively used to measure population health status for over five years, a multi-attribute utility function specific to the HUI-Mark III has yet to be developed. Instead, each of the HUI-Mark III attribute scores have been translated into corresponding scores on the HUI-Mark II attribute system and the HUI-Mark II utility function then applied to generate a HUI-Mark III score (Berthelot, Roberge & Wolfson,

1992). While the translation process was based on the best estimate of the McMaster research team, the differences between the two attribute classification systems resulted in a number of compromises. First, the three HUI-Mark III sensation variables (vision, hearing and speech) had to be combined into a single sensation variable. Second, the closest counterpart to the HUI-Mark III dexterity attribute was the self-care attribute. Lastly, the HUI-Mark III does not contain an attribute comparable to fertility (Berthelot, Roberge & Wolfson, 1992).

The preference ratings required for the development of the HUI-Mark II utility function were also developed as part of the aforementioned childhood cancer study (Feeny et al., 1992). In this study, parents were asked to imagine themselves as a child of 10 who would have to live in a variety of health states until death at age 70. Standard gamble and visual analogue scale techniques were used to obtain state preferences or utility weights for each attributes levels and a limited selection of multi-attribute health states (Berthelot et al., 1992).

The visual analogue scale technique utilises a vertical thermometer-shaped scale, 55 cm long, numerically scaled in units from 0-100. Individuals are asked to place printed cards representing attribute levels along the scale. The most preferred attribute level is placed at the top of the scale while the least preferred is placed at the bottom. The spacing between the remaining intermediary attribute levels on the scale is then used to judge the relative utility of one level over another.

In the standard gamble technique, the subject is presented with two alternatives, one consisting of an intermediate health state outcome with certainty (e.g., fully blind), the other a

gamble with specified probabilities of two possible outcome health states. The two outcomes for the gamble represent the best (fully healthy state) and the worst (immediate death) possible health outcomes (see Figure 1). The probabilities of the best and worst states are then changed from decision to decision until the individual is indifferent with respect to his/her preference towards the chronic health state or gamble state (e.g., subject places equal value on being blind or dying immediately). At this point of indifference, the preference of the intermediate state equals the probability of the gamble (i.e., preference rating of being blind is 0) (Berthelot et al., 1992).

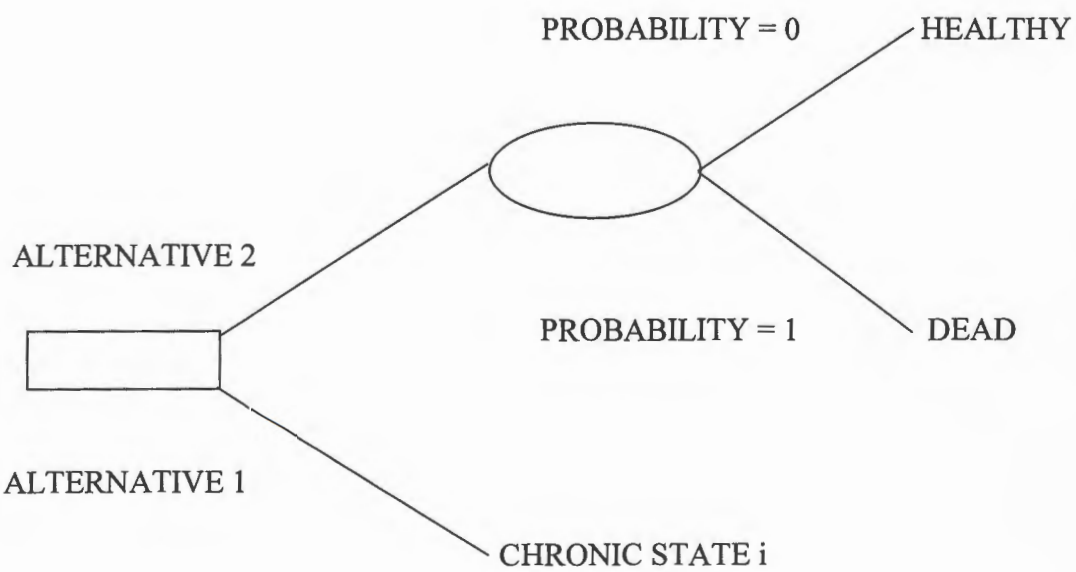

Figure 1. Standard Gamble diagram for the choice between living for sixty years in a chronic health state (blindness) or a 100 percent chance of immediate death, from Torrance (1986), p. 20.

The utility weights (i.e., preferences obtained from visual analogue scale and standard gamble tasks) associated with a selection of single attribute and full health state scores were then used to develop the following multi-attribute utility function:

$$HSI_i = \left\{ \alpha * \prod_{k=1}^{7} \mathrm{W}_k(h_{i,k}) \right\} - \beta,$$

where HSI_i represents the product of the preference weights, $\mathrm{W}_k(h_{i,k})$, associated with the observed health status level $h_{i,k}$ of the i^{th} individual for attribute k. The values α and β are parameters determined by the data and provide a range for the index scores. For the HUI-Mark II, α was 1.06 and β was 0.06, such that the resulting health status index scores ranged from -0.02 (i.e. a state worse than death) to 1.0 (i.e. a fully healthy state). For a more detailed discussion of the development of the multi-attribute utility function see Torrance et al. (1992).

While the NPHS contains a detailed assessment of the psycho-social components of health (e.g. social support, self-esteem, perceived stress), the HUI-Mark III relies on an extremely limited selection of health status indicators. Of the 31 questions used to derive the HUI-Mark III attribute scores, only 3 questions directly assess mental functioning and none appear to examine social functioning (Statistics Canada, 1995). The apparent paucity of items assessing mental and social well-being raises doubts concerning the ability of the HUI-Mark III to provide a valid assessment of health status for use in the NPHS. More specifically, it appears

that the HUI-Mark III is primarily a measure of physical functioning incapable of adequately assessing the mental and social dimensions of health.

Given the dimensionality of the WHO definition of health and recent literature citing stress, self-esteem and social support as the most important factors in explaining today's health gradients (Millar & Hull, 1997), it seems appropriate to incorporate indicators of mental and social well-being into the summary measure of health status used in the NPHS. The failure to include such indicators would not only limit the ability of the summary measure to provide a comprehensive description of health status, but more importantly, would substantially limit its ability to provide information on a large proportion of health determinants thought to achieve their effects through changes in mental and/or social functioning.

In order to comment formally on the use of the HUI-Mark III as a summary measure of health status, it is necessary to gather evidence on what the HUI-Mark III test scores mean or represent (Hubley & Zumbo, 1996). One route to assembling this type of information is through an examination of the reliability and validity associated with the use of the HUI-Mark III in the NPHS.

Reliability

Reliability is an important property of a measure because it is a necessary prerequisite to validity. That is, it is a necessary but not sufficient condition for determining whether an

instrument measures what it intends to measure (Bergner and Rothman, 1987). Boyle, Furlong, Feeny, Torrance and Hatcher (1995) examined the test-retest reliability of the HUI-Mark III on a random sample of 506 individuals interviewed over the telephone during August and September, 1991 and again 1 month later. The study concluded that the individual questions, attributes and provisional index scores provided by the HUI-Mark III generally provided reliable information on health status. The overall test-retest reliability (intra-class correlation coefficient) of the provisional index score was 0.767 (Boyle et al., 1995).

Validity

Of all the concepts in testing and measurement, it has been argued that validity is the most basic and far-reaching; for without validity, a test, measure or observation and any inferences made from it are meaningless (Hubley & Zumbo, 1996). However, despite the widespread use of economic indicators to measure health status (e.g., the HUI-Mark III), very little effort has been devoted to a systematic discussion and evaluation of their validity within the field of Economics (MacPhail, 1998).

In contrast to Economics, many other social sciences have carefully defined and documented the historical evolution of the concept of validity (see, for example, Hubley and Zumbo, 1996). According to the Standards for Educational and Psychological Testing (APA, AERA, & NCME, 1985), validity refers to "the appropriateness, meaningfulness, and usefulness of the specific inferences made from test scores" (p. 9). This approach to validity is that of a

process of disciplined inquiry that tries to address two major threats to the inferences made from our measures or observations. The first is construct underrepresentation, or the idea that the measure fails to include important dimensions or facets of the construct. The second threat to our inferences is construct-irrelevant variance, which has three sources: (a) the measure is too broad and contains excess reliable variance associated with other distinct constructs; (b) reliable variance that is due to the manner in which the measure is obtained (i.e., method variance); and (c) unreliable or error variance that is often quantified by some index (or coefficient) of reliability (Messick, 1990).

While in the past, showing that the test predicted some criterion was often more important than showing what a test actually measured (Angoff, 1988), this new approach to validation takes the form of disciplined inquiry in which plausible, alternative inferences from the test scores are disproved. Given the suspicion that scores on the HUI-Mark III are an underrepresentation of the full range and dimensionality of health measured in the NPHS, a review of content and construct validity appears to be a good starting point in any appraisal of the appropriateness, meaningfulness, and usefulness of inferences associated with the NPHS HUI-Mark III scores.

Content Validity

Content validity addresses the question, "How representative of the construct of health status are the items that comprise the measure?" Two aspects of content validity are sampling validity and face validity. Sampling validity consists of identifying the important dimensions

of health status (e.g. physical, psychological and social well-being) and including items in the instrument that assess each of the relevant categories within those dimensions (Bergner et al, 1987). A measure with content validity will usually exhibit "face validity" which is the simple appearance that the items are related to the construct of interest. However, face validity alone is not sufficient for claiming content validity (Kaplan, Bush, and Berry, 1976).

Therefore, in the context of measuring health status, two features of the WHO definition of health as a complete state of physical, mental and social well-being and not merely the absence of disease or infirmity are crucial: namely, the dimensionality of health and the full spectrum of health states ranging from disease to well-being. Further evidence for the aforementioned dimensions can be found in a recent review of ten widely used general health surveys. In this study, Ware (1995) examined the content of the following instruments: Quality of Well-Being Scale, Sickness Impact Profile, RAND Health Insurance Experiment surveys, Nottingham Health Profile, Quality of Life Index, Dartmouth Function Charts, European Quality of Life Index, Duke Health Profile, Medical Outcomes Study (MOS) Functioning and Well-Being Profile and the MOS 36-Item Short-Form Health Survey. Based on these instruments, Ware (1995) concluded that the minimum set of concepts necessary for comprehensiveness were physical functioning, mental health, limitations in social and role functioning due to health problems and general health perceptions.

Construct Validity

A construct is a theoretically derived notion of the domain we want to measure. An understanding of the construct will lead to expectations about how an instrument should behave if it is valid. Construct validity involves comparisons between measures with the intention of assembling evidence supporting the logical relationships that should exist between a measure and characteristics of the population under study (Guyatt, Feeny & Patrick, 1993). The limitations of conventional indexes have led virtually all health statisticians to advocate the development of measures of health status that are more sensitive to changes in the factors determining health, mainly the economic and social environment (Noack & Abelin, 1987).

Evidence supporting construct validity may take two forms. Convergent evidence attempts to verify the strength and direction of the relationship between a single existing measure and the proposed measure based on existing knowledge of the construct. Discriminant evidence indicates that the proposed measure correlates better with a second measure accepted as more closely related to the construct than it does with a third more distantly related measure (Kaplan et al., 1976).

III: METHODS

The main objectives of the NPHS are to provide a measure of the level, trend and distribution of population health status and to provide information on the effects of many potential health determinants (Catlin et al., 1992). It follows that the instrument selected to provide a summary measure of health status should therefore reflect the level of physical, mental and social well-being in the sampled population.

Research Objectives

The main objective of this thesis was to examine statistically the ability of the HUI-Mark III to provide a measure of health status compatible with the information objectives of the NPHS. The first step in accomplishing this task was to perform an exploratory factor analysis on the indicators of physical, mental and social well-being contained in the survey. Based on the comprehensive nature of the NPHS and similar research by Ware (1987; 1995), it was expected that the factor analysis would identify the following dimensions of health status: physical health, mental health, social and role functioning, and general perceptions of well-being.

The relationship between HUI-Mark III scores and the dimensions uncovered in the factor analysis was then examined using multiple regression. This analysis determined the extent to which scores on the HUI-Mark III were representative of the dimensions of health status

measured in the NPHS (i.e. How much of the variation in the HUI-Mark III was accounted for by scores on the health status dimensions measured in the NPHS?).

Instrument Description

National Population Health Survey

The NPHS was designed to gather information on the health status of Canadians and additionally, to gather information on many suspected determinants of health. To meet these objectives, researchers integrated a variety of indicators assessing health status, health care utilisation, risk factors and socio-economic status into the NPHS (Statistics Canada, 1995).

HUI-Mark III

The questions used to determine the HUI-Mark III can be found in the Health Status section of the NPHS. A complete list of the questions used to determine HUI-Mark III scores can be found in Appendix B.

Indicators Selected for Exploratory Factor Analysis

The following indicators of physical, mental and social well-being were selected for the exploratory factor analysis. The health status vision, hearing, speech, mobility, dexterity,

emotion, cognition and pain and discomfort variables were derived from responses to the HUI-Mark III questionnaire found in the health status section of the NPHS.

Health Status: Vision Attribute: The vision attribute was measured using a 5-point scale ranging from no vision problems to uncorrected problems seeing both near and far. A higher score indicates more severe problems.

Health Status: Hearing Attribute: The hearing attribute was measured using a 3-point scale ranging from no hearing problems to hearing problems that have not been corrected. A higher score indicates more severe problems.

Health Status: Speech Attribute: The speech attribute was measured using a 2-point scale ranging from no speech problems to partially or not understood. A higher score indicates more severe problems.

Health Status: Mobility Attribute: The mobility attribute was measured using a 4-point scale ranging from no mobility problems to the inability to walk. A higher score indicates more severe problems.

Health Status: Dexterity Attribute: The dexterity attribute was measured using a 3-point scale ranging from no limitations associated with the use of hands and fingers to help needed. A higher score indicates more severe problems.

Health Status: Emotion Attribute: The emotion attribute was measured using a 5-point scale ranging from being happy and interested in life to feeling unhappy/life not worthwhile. A higher score indicates less perceived happiness.

Health Status: Cognition Attribute: The cognition attribute was measured using a 5-point scale ranging from no cognitive problems to being very forgetful. A higher score indicates more severe problems.

Health Status: Pain and Discomfort Attribute: The pain and discomfort attribute was measured using a 4-point scale ranging from no pain or discomfort to severe pain or discomfort. A higher score indicates more severe problems.

Adjusted Specific Chronic Stress Index: This index measures the total number of stressors respondents were exposed to from the following areas; personal stress, financial problems, relationship problems, child problems, environmental problems and family health. Scores range from 0 to 14 and are adjusted as if all the items were relevant to each respondent. A higher score indicates a greater number of chronic stressors.

Work Stress Index: This index reflects respondents' perceptions about various dimensions of their work including job security, social support, monotony, physical effort required and extent of participation in decision-making. Higher scores indicate greater work stress.

Self Esteem Index: This index reflects the amount of positive feelings an individual holds about him/herself. Scores on the index are based on a subset of items from the Rosenburg (1969) self esteem scale. The six items factored into one dimension in the factor analysis done by Pearlin and Schooler (1978). Higher scores indicate greater self esteem.

Mastery Index: This index, based on the work of Pearlin and Schooler (1978), measures the extent to which individuals believe that their life-chances are under their control. Higher scores indicate superior mastery.

Sense of Coherence Scale: The 13-item version of the sense of coherence scale developed by Antonovsky was used in the NPHS. It denotes the extent to which individuals perceive events as comprehensible, manageable and meaningful. Higher scores indicate a stronger sense of coherence.

Mental Distress Scale: The index is based on a subset of items from the Composite International Diagnostic Interview designed to produce diagnoses according to the definitions and criteria of both DSM-III-R and the Diagnostic Criteria for Research of the ICD-10. Higher scores indicate more distress.

Social Support Index: The perceived social support index uses a 4-point scale in assessing whether respondents feel that they have someone they can confide in, someone they can count on, someone who can give them advice and someone who makes them feel loved. A higher score indicates greater perceived social support.

Average Frequency of Contact Index: The average frequency of contact index measures the average number of contacts in the past 12 months with family members and friends who are not part of the household and with neighbours. A higher score indicates more contacts.

Derived Health Description Index: This index assesses the respondents' perceived general health. The 5-point scale ranges from poor to excellent general health. A higher score indicates better health.

Sampling

Statistics Canada (1995, pp. 9-11) provided the following description of the sampling methodology. The target population of the NPHS consists of household residents in all provinces and territories, except persons living on Indian reserves, on Canadian Forces bases, or in some remote areas. An institutional component covers long term residents of hospitals and residential care facilities.

The survey collected most of the information from a single household member. Interviewing one respondent simplifies the longitudinal follow-up. Each time the respondent is re-surveyed, the same basic health-related information will be collected from all members of the household in which he or she is then living.

To enhance the representativeness of the panel, a reductive technique was applied. If households had been randomly selected , an individual's chances of being included in the

panel would be inversely related to the number of persons in that household. The panel would thus tend to underrepresent people in large households, typically parents and dependent children, and overrepresent people in small households, who are often single or elderly. The rejective approach was applied by identifying a portion of the sample households for screening, and dropping those that did not have at least one member under age 25.

The sample used in this study consisted of 838 (402 male, 436 female) residents of the Prince George/Northern Interior Health Region of British Columbia. Participants were selected from a random-digit dialling sample of telephone numbers. The final response rate was approximately 88 %.

IV: ANALYSIS

The analysis section of this study was separated into two sections. First an exploratory factor analysis of NPHS health status indicators was conducted to identify the dimensions of health measured in the survey. Then the HUI-Mark III was regressed on to the saved factor scores representing each dimension to determine the relative importance of each factor in predicting HUI-Mark III scores.

Exploratory Factor Analysis

Screening the Data

A screening of the database found no obvious coding errors. A review of variable histograms indicated that several of the health status measures derived from the HUI-Mark III questionnaire showed very little variation in response. This lack of variation was most pronounced in the hearing, speech, mobility and dexterity variables where the proportion of respondents who received the highest score possible was 94, 99, 96 and 99 percent respectively (see Figures 3 - 6 in Appendix C).

Since the data was discrete and exploratory factor analysis allows for the assumptions of normality, linearity and equality of variances to be relaxed, the analysis proceeded without removing outliers or transforming any scores (Tabachnick & Fidell, 1996). It should be noted that the lack of variability in responses reduces the potential for correlation and therefore may suppress the expression of some factors (Tabachnick & Fidell, 1996). The histograms also revealed that the Northern Interior data contained 672 missing values.

While every variable in the analysis contained missing scores, the vast majority of the missing values were in the work stress index (n = 303), chronic stress index (n=96) and the sense of coherence scale (n=101). There were no significant age or sex differences between subjects with and without missing scores on the work stress index ($F(1,837) = .051$, $p = .82$; Chi-square $= .275$, $p = .600$ respectively), the chronic stress index ($F(1,837) = .002$, $p = .965$;

Chi-square = .052, p = .819 respectively) and the sense of coherence scale (F(1,837) = .000, p = .985; Chi-square = .009, p = .924 respectively). While the missing scores in the work stress index are probably related to the fact that employment is a prerequisite for work stress, it seems reasonable to assume that the other missing values are missing at random. It is therefore justifiable to substitute the variable mean for these missing values in order to maintain sample size.

The possible trend underlying missing work stress index scores and lack of employment requires that several additional approaches to dealing with missing data be considered. The first alternate approach is to simply delete all cases with missing work stress index values. Unfortunately, such an approach would have resulted in a substantially reduced sample size that contained only employed persons. A second approach involves removing variables with missing values from the analysis; however, this would seriously jeopardise the ability of the study to comment on the dimensions of health measured in the NPHS.

A more feasible solution to the problem of missing data involves replacing the work stress index missing scores with mean scores. Such an approach allows for the maximum amount of information to be extracted from the data (i.e., all of the participants' responses would be included in the analysis) without artificially changing the mean work stress index score. On the negative side, the variance of the work stress index will be reduced and so will the correlation it has with other variables. This artificially reduced correlation will suppress the contribution of the work stress index variable to the factor analytic solution. Since the

benefits associated with the mean replacement of work stress scores appear to outweigh the costs, missing work stress scores were replaced with mean scores.

Factorability of the Correlation Matrix

The correlation matrix contained numerous correlations in excess of .30 supporting the use of factor analysis. The Kaiser-Meyer Olkin measure of sampling adequacy yielded a value of .80 which is more than adequate (Kaiser, 1970). Bartlett's Test of Sphericity was used to determine whether the correlation matrix differed from the identity matrix. The test supported the use of factor analysis (Chi-Square (136) = 1635.42, p = .00).

Factor Extraction and Rotation

Maximum likelihood factor extraction was selected because it allows for a test of the goodness-of-fit between the data and the factor model. Oblique rotation (direct oblimin with a delta value of zero) was used because it allows for factors to be reasonably correlated.

Determining the Number of Factors

The criteria used to determine the appropriate number of factors consisted of the scree test, the number of eigenvalues greater than one, the Chi-squared goodness-of-fit test, the presence of residuals greater than .05 between observed and reproduced correlation matrices and the overall interpretability of the solution (Tabachnick & Fidell, 1996). Using the scree plot to

determine the approximate number of factors in a solution involves finding the point at which a line drawn through the data points, starting at the bottom right side of the graph, suddenly changes slopes (see Figure 2).

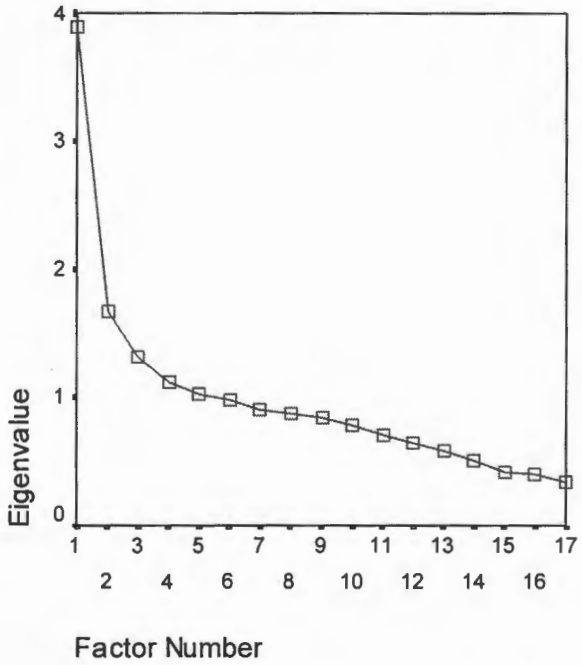

Figure 2. Scree plot with a noticeable change in slope at factor four.

The scree plot in Figure 2 suggests that there are approximately four factors in the solution. However, under the less than optimal conditions associated with this data (e.g. lack of variability in some variables), the scree test is said to be accurate within one or two factors (Tabachnick and Fidell, 1996). The scree test therefore suggests that anywhere from 2 to 6 factors are present in the solution.

The number of eigenvalues greater than one is also used as a rough approximation of the number of factors in a solution. From Table 1, it can be seen that there are five factors with eigenvalues greater than one.

Table 1

Total Variance Explained

Factor	Eigenvalue	Percent of Variance	Cumulative Percent
1	3.896	22.9	22.9
2	1.675	9.8	32.8
3	1.315	7.7	40.5
4	1.125	6.6	47.1
5	1.023	6.0	53.1
6	.974	5.7	58.9
7	.897	5.3	64.1
8	.872	5.1	69.3
9	.837	4.9	74.2
10	.784	4.6	78.8
11	.710	4.2	83.0
12	.639	3.8	86.7
13	.584	3.4	90.2
14	.513	3.0	93.2
15	.420	2.5	95.7
16	.404	2.4	98.0
17	.333	2.0	100.0

A second factor analysis extracting only five factors was performed. This solution accounted for 53.1 % of the variation in original variables, passed the goodness-of-fit test (Chi-square (61) = 75.16, p = .105), and had five residuals with absolute values greater than .05, none of which were greater than .1. In the interest of parsimony, a four factor extraction was attempted. The solution accounted for 47.1 % of the variation in original variables, failed the

goodness-of-fit test (Chi-Square (74) = .002) and had eleven residuals greater than .05, none of which were greater than .1.

The results of the scree test, the number of eigenvalues greater than one and the goodness-of-fit test all support the interpretation of the five factor solution. While interpretation of the four factor solution is supported by the scree test, it failed the goodness-of-fit test (i.e. there was a significant difference between the observed and reproduced correlation matrices) and had more than twice as many residuals greater than .05 as the five factor solution. The five factor solution was therefore selected for interpretation.

Interpretation of the Factor Solution

Intrafactor correlations cause the correlations between variables and factors in the structure matrix to be inflated. However, the pattern matrix represents a measure of the unique relationship between variables and factors. The pattern matrix was therefore selected for interpretation (See Table 2). Since there is no generally accepted cut off level for variable consideration, a cut off level of .200 was selected to simplify interpretation of the pattern matrix while still including all variables in the final solution.

Table 2

Pattern Matrix

	Factor				
Indicator	Physical Impairment	Mental Ill-health	Mental Well-being	General Health Impairment	Social Well-being
Health Status: Dexterity	1.0				
Health Status: Speech	.447				
Sense of Coherence		-.685			
Mental Health: Distress		.665			
Adjusted Chronic Stress		.657			
Self Esteem Index			.838		
Mastery Index		-.240	.565		
Work Stress Index			-.208		
Health Description				-.527	
Pain/Discomfort				.451	
Health Status: Mobility				.437	
Health Status: Hearing				.284	
Health Status: Vision				.257	
Health Status: Cognition				.244	
Frequency of Contacts					.476
Health Status: Emotion		.230	-.246		-.422
Social Support Index					.355

Note: Indicators with loadings less than .200 have not been listed.

Factor 1: Physical Impairment

Indicators of dexterity and speech load primarily and positively on factor 1. This factor seems to represent a measure of impairment associated with physical functioning.

Factor 2: Mental Ill-health

Indicators of distress, chronic stress and emotional functioning load positively on factor 2 while indicators of mastery and sense of coherence load negatively. This factor could be interpreted as an indicator of mental ill-health related to levels of stress/distress.

Factor 3: Mental Well-being

The third factor seems to represent a measure of positive mental health primarily determined by self-esteem and mastery.

Factor 4: General Health Impairment

The fourth factor appears to assess the impact of mobility and pain problems and to a lesser extent hearing, vision and cognition problems on the respondents' description of his/her general health.

Factor 5: Social Well-being

The fifth factor appears to represent a measure of social well-being related to level of emotional functioning (See Table 2).

The factor interpretations are supported by similar findings in the literature indicating that the health is composed of the following five dimensions; physical health; mental health; social functioning; role functioning and general health perceptions (Ware, 1987; 1995).

Additional support for the interpretation of factors lies in the factor correlation matrix (See Table 3). Ratings on physical impairment and general health impairment were positively correlated. Mental ill-health was negatively correlated with mental well-being and social well-being. General health impairment was negatively correlated with social well-being. Lastly, mental well-being was negatively correlated with general health impairment and positively correlated with social well-being. The pattern of these correlations makes

intuitive sense and provides additional support for the interpretation.[1]

Table 3

Factor Correlation Matrix

Factor	1.	2.	3.	4.	5.
1. Physical Impairment	1	.012	-.015	.252	-.129
2. Mental Ill-health		1	-.496	.186	-.306
3. Mental Well-being			1	-.175	.309
4. General Health Impairment				1	-.196
5. Social Well-being					1

Section 2: Multiple Regression

This section of the study regressed the HUI-Mark III on to the five factors identified in section one in order to determine the relative proportion of variation in the HUI-Mark III accounted for by each factor. The resulting regression equation was statistically significant, $F(5,827) = 392.2$, $p < .01$ and had an R-squared of .702.

To determine the relative contribution of each explanatory variable (i.e., each factor) to the regression equation, a relative Pratt index (Thomas, Hughes & Zumbo, in press) was generated (see Table 4). The relative Pratt index quantifies the relative contribution each explanatory variable makes to the overall regression equation by partitioning the model R^2

[1] A second exploratory factor analysis, with cases containing missing scores deleted, was performed to see if the factor solution differed from the analysis using mean substitution. The five factor solution contained a similar pattern of results to the factor solution derived from the data with mean substitution.

into that proportion attributable to each explanatory variable. The scores are additive and will therefore sum to 1.0.

Physical impairment, mental ill-health and mental well-being can be considered as having a negligible or "unimportant" contribution to scores on the HUI-Mark III (Thomas, Hughes & Zumbo, in press). However, general health impairments related to mobility and pain problems accounted for 72.2 % of the HUI-Mark III variance explained by the factors and social well-being accounted for 22.2 % of the HUI-Mark III variance explained by the factors (see Table 4).

Table 4

Relative Pratt Index Scores

Factor	Corr. With HUI-Mark III	Beta-weight	Relative Pratt Score
Physical Impairmnet	-.301	-.083	.0355
Mental Ill-health	-.403	-.008	.0046
Mental Well-being	.379	.031	.0167
General Health Impairment	-.775	-.655	.7221
Social Well-being	.518	.301	.2218

R-squared = .703

Cut off for importance (Relative Pratt Index Score) = .10

V. DISCUSSION

The main objective of the NPHS is to provide a measure of the level, trend and distribution of health status for the general population of Canada. In the context of this objective, health is best defined as a complete state of physical, mental and social well-being and not merely the absence of disease (British Columbia. Provincial Health Officer, 1994; Ware, 1987). Key to this definition is the multidimensional nature of health (i.e. health is composed of physical, mental and social dimensions) and the extension of its range into positive states. The instrument selected to provide a summary measure of health status should therefore accommodate both the dimensionality and full range of health as described in the aforementioned definition.

However, the distributions of the HUI-Mark III and its' derived attributes contain very little variation. The vast majority of respondents scored at or near the highest health level possible. Such distributions suggest that the HUI-Mark III is primarily focused on describing variability associated with negative health states and appears unable to differentiate among the many different levels of positive well-being. This is not surprising given that the health status classification system and the utility function used in the HUI-Mark III were originally developed to assess the impact of paediatric cancer treatments on physical functioning (e.g. comparing aggressive treatment with intent to cure verses palliative care).

While a large proportion of paediatric cancer patients are likely to exhibit signs of functional impairment, only some 15 percent of general population samples will have chronic physical

limitations, and only 10 to 20 percent will have substantial psychiatric impairment. The use of health status indicators focussed on describing the negative aspects of health will therefore generate little or no information on the health of the remaining 70 to 80 percent of the population (Ware et al., 1981). While such a scenario would explain the lack of variation and apparent inability of the HUI-Mark III to differentiate among the many different levels of positive health, it also brings into question the usefulness and appropriateness of the HUI-Mark III as a measure of health status in a survey developed to assess the health status of the general population of Canada.

After examining the distributions of NPHS health status indicators, an exploratory factor analysis was conducted in an attempt to identify the health status dimensions measured in the NPHS. The following five dimensions emerged from the factor analysis; physical impairment; mental ill-health; mental well-being; general health impairment and social well-being. These factors are similar to those identified by Ware (1987; 1995) and suggest that the NPHS does indeed contain indicators of physical, mental and social well-being.

The regression analysis indicated that the five factors identified in the factor analysis account for 70 % of the explained variation found in the HUI-Mark III. However, the vast majority of this variance was accounted for by general health impairment related to pain and mobility problems and, to some extent, social well-being. Mental ill-health and mental well-being together accounted for only 2% of the explained variation in HUI-Mark III scores. This suggests that while the NPHS contains indicators of physical, mental, and social well-being, scores on the HUI-Mark III primarily reflect the impact of physical and to a much lesser

extent social problems on everyday life. Furthermore, the index is insensitive to changes in mental well-being related to stress, self-esteem and sense of mastery.

The importance of considering indicators of stress, self esteem and mastery when measuring population health status can not be understated. A recent publication identified the balance between the stress people experience on the one hand and the sense of self esteem, mastery and social support on the other hand as the most important factor in explaining existing health gradients in British Columbia (Millar & Hull, 1997). The lack of any relationship between scores on the HUI-Mark III and measures of stress, self-esteem and mastery support the suspicion that HUI-Mark III scores underrepresent the concept health status by not adequately sampling mental well-being. This raises serious doubts concerning the meaningfulness of HUI-Mark III scores in the arena of general population health assessment. For example, an individual or demographic group (e.g., single mothers living in poverty) might report experiencing high levels of chronic stress, have low of self-esteem and sense of mastery yet be classified as being very healthy by their scores on the HUI-Mark III. The very real possibility of such results indicate that the 'summary measure of health status' provided by the HUI-Mark III should be interpreted with extreme caution, especially when investigating the influence of health determinants thought to achieve all or part of their effects through changes in levels of mental well-being.

VI. IMPLICATIONS

The results of this study suggest that the problems associated with the HUI-Mark III can be traced to two specific areas. First, the HUI-Mark III does not appear to discriminate between the many different levels of positive health experienced by the vast majority of the general population. Second, the HUI-Mark III is more or less insensitive to variation in key indicators of mental well-being. While the failure of researchers at the Centre for Health Economics and Policy Analysis to develop a utility function for the HUI-Mark III might appear to be part of the problem, the obvious lack of variation in the derived attribute scores lead this author to conclude that nothing short of developing a new classification system is required to adequately sample the concept of health status. This revised classification system would need to include additional indicators of mental well-being and possess an extended range for all attributes that reaches further into positive health states.

From a broader perspective, the results of this study also bring into question the appropriateness of not only the HUI-Mark III, but any instrument relying on a single score to describe health status. Even if a measure tapped the full range and dimensionality of health, the interpretability of scores would remain problematic due to the inherent inability to identify the domains in which improvement or deterioration occurs. This limitation raises serious questions concerning the usefulness of such measures in the investigation of today's health determinants and the subsequent development of appropriate policy interventions. It is therefore highly recommended that the HUI-Mark III not be used as the sole summary measure of health status when attempting to meet the primary objectives of the NPHS; those

being to describe the level, trend and distribution of population health and to investigate the effects of potential health determinants.

VII. FUTURE RESEARCH DIRECTIONS

While it is not possible increase the range and dimensionality of scores on the 1994 NPHS HUI-Mark III, researchers could integrate HUI-Mark III scores and additional indicators of mental well-being into a health profile composed of several dimension specific scores. Based on the results of this study, such a profile would provide a more comprehensive description of the multi-dimensional concept of health status measured by the NPHS and better facilitate research into the identification and understanding of mechanisms associated with the many different determinants of health. The information obtained from the profile could also be compared with that provided by the HUI-Mark III in order to further clarify exactly what types of inferences can be appropriately drawn from scores on the HUI-Mark III.

REFERENCES

American Psychological Association (1974). *Standards for educational and psychological tests*. Washington, DC.

American Psychological Association, American Educational Research Association, & National Council on Measurement in Education (1985). *Standards for educational and psychological testing*. Washington, DC: American Psychological Association.

Angoff, W. H. (1988). Validity: An evolving concept. In H. Wainer & H. Braun (Eds.), *Test validity* (pp. 19-32). Hillsdale, NJ: Erlbaum.

Bergner M. & Rothman M. L. (1987). Health status measures: an overview and guide for selection. *Annual Review of Public Health*, Vol. 8, 191-210.

Berthelot J. M., Roberge R. & Wolfson M. C. (1992). *The calculation of health-adjusted life expectancy for a Canadian province using a multi-attribute utility function: A first attempt*. Statistics Canada, Catalogue 96-01093, Ottawa, Ontario.

Boyle M. H., Furlong W., Feeny D., Torrance G. W. & Hatcher J. (1995). Reliability of the Health Utilities Index-Mark III used in the 1991 cycle 6 Canadian General Social Survey Health Questionnaire. *Quality of Life Research*, Vol. 4, 249-257.

Breslow, L. (1972). A quantitative approach to the World Health Organization definition of health: Physical, mental and social well-being. *International Journal of Epidemiology*, Vol. 1, No. 4, 347-355.

British Columbia. Provincial Health Officer. (1994). *A report on the health of British Columbians: Provincial Health Officer's annual report 1994*. Victoria, BC: Ministry of Health and Ministry Responsible for Seniors.

Canada (1974). *A New Perspective on the health of Canadians* (Lalonde Report). Ottawa: Department of National Health and Welfare.

Catlin, G. & Will, P. (1992). The National Population Health Survey: Highlights of initial developments. *Health Reports*, Vol. 4, No. 3: 313-319.

Feeny, D., Furlong, W., Barr, R. D., Torrance, G. W., Rosenbaum, P. & Weitzman (1992). A comprehensive multiatribute system for classifying the health status of survivors of childhood cancer. *Journal of Clinical Oncology*, Vol 10, No 6, 923-928.

Green, L. W. (1990). *Community health*. Times Mirror/Mosby College Publishing.

Guyatt, G. H., Kirshner, B. & Jaeschke, R. (1992). Measuring health status: What are the necessary measurement properties? *Journal of Clinical Epidemiology*, Vol. 45, No. 12, 1341-1345.

Guyatt, G. H., Feeny, D. H. & Patrick, D. L. (1993). Measuring health-related quality of life. *Annals of Internal Medicine*, Vol 118, 622-629.

Health Canada (1994). *Strategies for population health.* Health Canada, Ottawa.

Health Canada (1998). *Health Canada request for synthesis proposals: measuring the health of populations*. Health Canada, Ottawa.

Hubley, A. M. & Zumbo, B. D. (1996). A dialectic on validity: Where we have been and where we are going. *The Journal of General Psychology*, Vol. 123, No. 3, 207-215.

Hunt, S. J., McEwen, J., McKenna, S. P. (1986). *Measuring health status*. Croom Helm Ltd.

Kaplan, R. M., Bush, J.W. & Berry, C. C. (1976). Health status: Types of validity and the index of well-being. *Health Services Research*, Vol 11, No. 4, 478-507.

Larson, J.S. (1991). *The measurement of health.* New York: Greenwood Press.

Levy, E. (1985). The search for health indicators. *International Social Science Journal*, 15: 11: 433-463.

MacPhail, F. (1998). Moving beyond statistical validity in economics. *Social Indicators Research*, Vol. 45, 119-149.

Marmor, T. R., Barer, M. L. & Evans, R. G. (1994). The determinants of a population's health: What can be done to improve a democratic nation's health status? Pp.217-230 in *Why are some people healthy, and others not? The determinants of health of populations*, edited by Evans R. G., Barer, M. L. & Marmor, T. R.. Aldine de Gruyter. New York.

Messick, S. (1990). *Validity of test interpretation and use*. Research Rep. No. 90-11. Princeton, NJ: Educational Testing Service.

Millar, J. S. (1994). A consensus on the causal role of social factors in population health: Implications for action. Pp. 201-206 in *The determinants of population health: A critical assessment*, edited by Hayes, M. V., Foster, L. T. & Foster, H. D.. Western Geographical Series.

Millar, J. S. & Hull, C. (1997). Measuring human wellness. *Social Indicators Research*, 40, 147-158.

Noack, H. & Abelin, T. (1987). Conceptual and methodological aspects of measurement in health and health promotion. pp. 89-102. In *Measurement in health promotion and protection*, edited by Abelin, T., Brzezinski, Z. J. & Carstairs, V. D. L.. World Health Organization Regional Publications. Copenhagen.

Read, J. L., Quinn, R. J. & Hoefer, M. A. (1987). Measuring overall health: An evaluation of three important approaches. *Journal of Chronic Diseases*, Vol. 40, Suppl. 1, 7S-21S.

Statistics Canada (1995). *National Population Health Survey*, 1994-95. Statistics Canada. Ottawa.

Thomas, D. R., Hughes, E., & Zumbo, B. D. (in press). On variable importance in linear regression. *Social Indicators Research: An International and Interdisciplinary Journal for Quality-of-Life Measurement.*

Torrance, G. W. (1986). Measurement of health state utilities for economic appraisal. *Journal of Health Economics*, Vol. 5, 1-30.

Torrance, G. W., Zhang, Y., Feeny, D., Furlong, W. & Barr, R. (1992). Multi-attribute preference functions for a comprehensive health status classification system. *Centre for Health Economics and Policy Analysis Working Paper Series*, No. 92-18.

Ware, J. E. (1987). Standards for validating health measures: definition and content. *Journal of Chronic Diseases*, Vol. 40, 473-80.

Ware, J. E. (1995). The status of health assessment 1994. *Annual Review of Public Health*, Vol. 16, 327-354.

Ware, J. E., Brook, R. H., Davies, A. R. and Lohr, K. N. (1981). Choosing measures of health status for individuals in general populations. *American Journal of Public Health*, Vol. 71, 6, 620-625.

Ware, J. E. & Sherbourne, C. D. (1992). The MOS 36-item short health survey (SF-36). Conceptual framework and item selection. *Medical Care*, Vol. 30, 473-483.

Wolfson, M. C. (1994). Social proprioception: measurement, data, and information from a population health perspective. Pp.287-316 in *Why are some people healthy, and others not? The determinants of health of populations*, edited by Evans R. G., Barer, M. L. & Marmor, T. R.. Aldine de Gruyter. New York.

Wolinski, F.D. & Zusman, M. E. (1980). Toward comprehensive health status measures. *The Sociological Quarterly*, Vol. 21: 607-621.

World Health Organization (1948). Constitution of the World Health Organization. In *World Health Organization* (1992). Basic documents (39th Ed.). World Health Organization. Geneva.

APPENDIX A

Table A1.
Attribute Levels for HUI-Mark II Health Status Classification System

Attribute	Level	Description
Sensation	1	Able to see, hear and speak normally for age.
	2	Requires equipment to see or hear or speak.
	3	Sees, hears, or speaks with limitations even with equipment.
	4	Blind, deaf or mute
Mobility	1	Able to walk, bend, lift, jump and run normally for age.
	2	Walks, bends, lifts, jumps or runs with some limitations but does not require help.
	3	Requires mechanical equipment (such as canes, crutches, braces or wheelchair) to walk or get around independently.
	4	Requires the help of another person to walk or get around and requires mechanical equipment as well.
	5	Unable to control or use arms and legs.
Emotion	1	Generally happy and free from worry.
	2	Occasionally fretful, angry, irritable, anxious, depressed, or suffering night terrors.
	3	Often fretful, angry, irritable, anxious, depressed, or suffering night terrors.
	4	Almost always fretful, angry, irritable, anxious, depressed.
	5	Extremely fretful, angry, irritable or depressed usually requiring hospitalization or psychiatric institutional care.
Cognition	1	Learns and remembers school work normally for age.
	2	Learns and remembers school work more slowly than classmates as judged by parents and/or teachers.
	3	Learns and remembers very slowly and usually requires special educational assistance.
	4	Unable to learn and remember.
Self-care	1	Eats, bathes, dresses and uses toilet normally for age.
	2	Eats, bathes, dresses and uses toilet independently but with difficulty.
	3	Requires mechanical equipment to eat, bathe, dress and use the toilet independently.
	4	Requires the help of another person to eat, bathe, dress or use the toilet.
Pain	1	Free of pain and discomfort.
	2	occasional pain. Discomfort relieved by non-prescription drugs self-control activity without disruption of normal activities.
	3	Frequent pain. Discomfort relieved by oral medicines with occasional disruption of normal activities.
	4	Frequent pain; frequent disruption of normal activities. Discomfort requires prescription narcotics for relief.
	5	Severe pain. Pain not relieved by drugs and constantly disrupts normal activities.
Fertility	1	Ability to have children with a fertile spouse.
	2	Difficulty in having children with a fertile spouse.
	3	Unable to have children with a fertile spouse.

Table A2.
Attributes for the HUI-Mark III Health Status Classification System

Attribute	Level	Description
Vision	1.	Able to see well enough to read ordinary newsprint and recognize a friend on the other side of the street, without glasses or contact lenses.
	2.	Able to see well to read ordinary newsprint and recognize a friend on the other side of the street, but with glasses.
	3.	Able to read ordinary newsprint with or without glasses but unable to recognize a friend on the other side of the street, even with glasses.
	4.	Able to recognize a friend on the other side of the street with or without glasses but unable to read ordinary newsprint, even with glasses.
	5.	Unable to read ordinary newsprint and unable to recognize a friend on the other side of the street, even with glasses.
Hearing	1.	Able to hear what is said in a group conversation with atleast three other people, without a hearing aid.
	2.	Able to hear what is said in a conversation with one other person in a quiet room without a hearing aid, but requires a hearing aid to hear to hear what is said in a group conversation with atleast three other people.
	3.	Able to hear what is said in a conversation with one other person in a quiet room with a hearing aid, and able to hear what is said in a group conversation with at least three other people with a hearing aid.
	4.	Able to hear what is said in a conversation with one other person in a quiet room without a hearing aid, but unable to hear what is said in a group conversation with at least three other people even with a hearing aid.
	5.	Able to hear what is said in a conversation with one other person in a quiet room with a hearing aid, but unable to hear what is said in a group conversation with at least three other people even with a hearing aid.
	6.	Unable to hear at all.
Speech	1.	Able to be understood completely when speaking with strangers or friends.
	2.	Able to be understood partially when speaking with strangers but able to be understood completely when speaking with people who know me well.
	3.	Able to be understood partially when speaking with strangers or people who know me well.
	4.	Unable to be understood when speaking with strangers but able to be understood partially by people who know me well.
	5.	Unable to be understood when speaking to other people (or unable to speak at all).
Ambulation	1.	Able to walk around the neighbourhood without difficulty, and without walking equipment.
	2.	Able to walk around the neighbourhood with difficulty; but does not require walking equipment or the help of another person.
	3.	Able to walk around the neighbourhood with walking equipment, but without the help of another person.

Table A2. *Continued*

Attribute	Level	Description
Ambulation (con't)	4.	Able to walk only short distances with walking equipment, and requires a wheelchair to get around the neighbourhood.
	5.	Unable to walk alone, even with walking equipment. Able to walk short distances with the help of another person, and requires a wheelchair to get around the neighbourhood.
	6.	Cannot walk at all.
Dexterity	1.	Full use of two hands and ten fingers.
	2.	Limitations in the use of hands or fingures, but does not require special tools or help of another person.
	3.	Limitations in the use of hands or fingers, is independent with use of special tools (does not require the help of another person).
	4.	Limitations in the use of hands or fingers, requires the help of another person for some tasks (not independent even with use of special tools).
	5.	Limitations in the use of hands or fingers, requires the help of another person for most tasks (not independent even with use of special tools).
	6.	Limitations in the use of hands or fingers, requires the help of another person for all tasks (not independent even with use of special tools).
Emotion	1.	Happy and interested in life.
	2.	Somewhat happy.
	3.	Somewhat unhappy.
	4.	Very unhappy.
	5.	So unhappy that life is not worthwhile.
Cognition	1.	Able to remember most things, think clearly and solve day to day problems.
	2.	Able to remember most things, but has a little difficulty when trying to think or solve day to day problems.
	3.	Somewhat forgetful, but able to think clearly and solve day to day problems.
	4.	Somewhat forgetful, and has a little difficulty when trying to think or solve day to day problems.
	5.	Very forgetful, and has great difficulty when trying to think or solve day to day problems.
	6.	Unable to remember anything at all, and unable to think or solve day to day problems.
Pain	1.	Free of pain and discomfort.
	2.	Mild to moderate pain that prevents no activities.
	3.	moderate pain that prevents a few activities.
	4.	Moderate to severe pain that prevents some activities.
	5.	Severe pain that prevents most activities.

APPENDIX B

HUI-Mark III Questionnaire

Health Status

HSTAT-INT The next set of questions ask about ... (r/'s) day to day health. The questions are **not** about illnesses like colds that affect people for short periods of time. They are concerned with a person's usual abilities. You may feel that some of these questions do not apply to you/him/her, but it is important that we ask the same questions of everyone.

Vision

HSTAT-Q1 Are/Is ... *usually* able to see well enough to read ordinary newsprint *without* glasses or contact lenses?

___ Yes (Go to HSTAT-Q4)
___ No
DK, R (Go to HSTAT-Q6)

HSTAT-Q2 Are/Is you/he/she *usually* able to see well enough to read ordinary newsprint *with* glasses or contact lenses?

___ Yes (Go to HSTAT-Q4)
___ No

HSTAT-Q3 Are/Is you/he/she able to see at all?

___ Yes
___ No (Go to HSTAT-Q6)
DK, R (Go to HSTAT-Q6)

HSTAT-Q4 Are/Is you/he/she able to see well enough to recognize a friend on the other side of the street *without* glasses or contact lenses ?

___ Yes (Go to HSTAT-Q6)
___ No
DK, R (Go to HSTAT-Q6)

HSTAT-Q5 Are/Is you/he/she *usually* able to see well enough to recognize a friend on the other side of the street *with* glasses or contact lenses?

___ Yes
___ No

HUI-Mark III questionnaire con't

Hearing

HSTAT-Q6 Are/Is ... *usually* able to hear what is said in a group conversation with at least three other people *without* a hearing aid?

___ Yes (Go to HSTAT-Q10)
___ No
DK, R (Go to HSTAT-Q10)

HSTAT-Q7 Are/Is you/he/she *usually* able to hear what is said in a group conversation with at least three other people *with* a hearing aid?

___ Yes (Go to HSTAT-Q8)
___ No

HSTAT-Q7a Are/Is you/he/she able to hear at all?

___ Yes
___ No (Go to HSTAT-Q10)
DK, R (Go to HSTAT-Q10)

HSTAT-Q8 Are/Is you/he/she *usually* able to hear what is said in a conversation with one other person in a quiet room *without* a hearing aid ?

___ Yes (Go to HSTAT-Q10)
___ No
R (Go to HSTAT-Q10)

HSTAT-Q9 Are/Is you/he/she *usually* able to hear what is said in a conversation with one other person in a quiet room *with* a hearing aid?

___ Yes
___ No

Speech

HSTAT-Q10 Are/Is ... *usually* able to be understood *completely* when speaking with strangers in your own language?

___ Yes (Go to HSTAT-Q14)
___ No
R (Go to HSTAT-Q14)

HUI-Mark III questionnaire con't

HSTAT-Q11 Are/Is you/he/she able to be understood *partially* when speaking with strangers?

___ Yes
___ No

HSTAT-Q12 Are/Is you/he/she able to be understood *completely* when speaking with those who know you/him/her well?

___ Yes (Go to HSTAT-Q14)
___ No
R (Go to HSTAT-Q14)

HSTAT-Q13 Are/Is you/he/she able to be understood *partially* when speaking with those who know you/him/her well?

___ Yes
___ No

Getting Around

HSTAT-Q14 Are/Is ... *usually* able to walk around the neighbourhood *without* difficulty and *without* mechanical support such as braces, a cane or crutches?

___ Yes (Go to HSTAT-Q21)
___ No
DK, R (Go to HSTAT-Q21)

HSTAT-Q15 Are/Is you/he/she able to walk at all?

___ Yes
___ No (Go to HSTAT-Q18
DK, R (Go to HSTAT-Q18)

HSTAT-Q16 Do/Does you/he/she require mechanical support such as braces, a cane or crutches to be able to walk around the neighbourhood?

___ Yes
___ No

HSTAT-Q17 Do/Does you/he/she require the help of another person to be able to walk?

___ Yes
___ No

HUI-Mark III questionnaire con't

HSTAT-Q18 Do/Does you/he/she require a wheelchair to get around?

___ Yes
___ No (Go to HSTAT-Q21)
DK, R (Go to HSTAT-Q21)

HSTAT-Q19 How often do/does you/he/she use a wheelchair?
(Read list. Mark one only.)

___ Always
___ Often
___ Sometimes
___ Never

HSTAT-Q20 Do/Does you/he/she need the help of another person to get around in the wheelchair?

___ Yes
___ No

Hands and Fingers

HSTAT-Q21 Are/Is ... *usually* able to grasp and handle small objects such as a pencil and scissors?

___ Yes (Go to HSTAT-Q25)
___ No
DK, R (Go to HSTAT-Q25)

HSTAT-Q22 Do/Does you/he/she require the help of another person because of limitations in the use of hands or fingers?

___ Yes
___ No (Go to HSTAT-Q24)
DK, R (Go to HSTAT-Q24)

HSTAT-Q23 Do/Does you/he/she require the help of another person with:
(Read list. Mark one only.)

___ Some tasks?
___ Most tasks?
___ Almost all tasks?
___ All tasks?

HUI-Mark III questionnaire con't

HSTAT-Q24 Do/Does you/he/she require special equipment, for example, devices to assist in dressing because of limitations in the use of hands or fingers?

___ Yes
___ No

Feelings

HSTAT-Q25 Would you describe yourself/... as being *usually*:
(Read list. Mark one only.)

___ Happy and interested in life?
___ Somewhat happy?
___ Somewhat unhappy?
___ Unhappy with little interest in life?
___ So unhappy that life is not worthwhile?

Memory

HSTAT-Q26 How would you describe your/his/her *usual* ability to remember things? Are/Is you/he/she:
(Read list. Mark one only.)

___ Able to remember most things?
___ Somewhat forgetful?
___ Very forgetful?
___ Unable to remember anything at all?

Thinking

HSTAT-Q27 How would you describe your/his/her *usual* ability to think and solve day to day problems? Are/Is you/he/she:
(Read list. Mark one only.)

___ Able to think clearly and solve problems?
___ Having a little difficulty?
___ Having some difficulty?
___ Having a great deal of difficulty?
___ Unable to think or solve problems?

HUI-Mark III questionnaire con't

Pain and Discomfort

HSTAT-Q28 Are/Is ... *usually* free of pain or discomfort?

___ Yes (Go to next section)
___ No
DK, R (Go to next section)

HSTAT-Q29 How would you describe the *usual* intensity of your/his/her pain or discomfort? (Read list. Mark one only.)

___ Mild
___ Moderate
___ Severe

HSTAT-Q30 How many activities does your/his/her pain or discomfort prevent? (Read list. Mark one only.)

___ None
___ A few
___ Some
___ Most

APPENDIX C

Histograms of NPHS Health Status Indicators

Index

Figure C1. HUI Mark III
The HUI-Mark III is a generic health status index that provides a description of an individual's overall functional health ranging from -.02 (i.e. a state worse than death) to 1 (i.e. fully healthy).

Figure C2. Health Status: Vision Attribute
The vision attribute was measured using a 5-point scale ranging from no vision problems to uncorrected problems seeing both near and far. A higher score indicates more severe problems.

Figure C3. Health Status: Hearing Attribute
The hearing attribute was measured using a 3-point scale ranging from no hearing problems to hearing problems that have not been corrected. A higher score indicates more severe problems.

Figure C4. Health Status: Speech Attribute
The speech attribute was measured using a 2-point scale ranging from no speech problems to partially or not understood. A higher score indicates more severe problems.

Figure C5. Health Status: Mobility Attribute
The mobility attribute was measured using a 4-point scale ranging from no mobility problems to the inability to walk. A higher score indicates more severe problems.

Figure C6. Health Status: Dexterity Attribute
The dexterity attribute was measured using a 3-point scale ranging from no limitations associated with the use of hands and fingers to help needed. A higher score indicates more severe problems.

Figure C7. Health Status: Emotion Attribute
The emotion attribute was measured using a 5-point scale ranging from being happy and interested in life to feeling unhappy/life not worthwhile. A higher score indicates more severe problems.

Figure C8. Health Status: Cognition Attribute
The cognition attribute was measured using a 5-point scale ranging from no cognitive problems to being very forgetful. A higher score indicates more severe problems.

Figure C9. Health Status: Pain and Discomfort Attribute
The pain and discomfort attribute was measured using a 4-point scale ranging from no pain or discomfort to severe pain or discomfort. A higher score indicates more severe problems.

Figure C10. Adjusted Specific Chronic Stress Index
This index measures the total number of stressors respondents were exposed to from the following areas; personal stress, financial problems, relationship problems, child problems, environmental problems and family health. Scores range from 0 to 14 and are adjusted as if all the items were relevant to each respondent. A higher score indicates a greater number of chronic stressors.

Figure C11. Work Stress Index
This index reflects respondents' perceptions about various dimensions of their work including job security, social support, monotony, physical effort required and extent of participation in decision-making. Higher scores indicate greater work stress.

Figure C12. Self Esteem Index
This index reflects the amount of positive feelings an individual holds about his/herself. Scores on the index are based on a subset of items from the self esteem Rosenburg scale (1969). The six items factored into one dimension in the factor analysis done by Pearlin and Schooler (1978). Higher scores indicate greater self esteem.

Figure C13. Mastery Index
This index, based on the work of Pearlin and Schooler (1978), measures the extent to which individuals believe that their life-chances are under their control. Higher scores indicate superior mastery.

Figure C14. Sense of Coherence Scale
The 13-item version of the sense of coherence scale developed by Antonovsky was used in the NPHS. It denotes the extent to which individuals perceive events as comprehensible, manageable and meaningful. Higher scores indicate a stronger sense of coherence.

Figure C15. Mental Distress Scale
The items and scoring used to derive the distress score are based on the work of Kessler and Mroczek. The index is based on a subset of items from the Composite International Diagnostic Interview designed to produce diagnoses according to the definitions and criteria of both DSM-III-R and the Diagnostic Criteria for Research of the ICD-10. Higher scores indicate more distress.

Figure C16. Social Support Index
The perceived social support index uses a 4-point scale in assessing whether respondents feel that they have someone they can confide in, someone they can count on, someone who can give them advice and someone who makes them feel loved. A higher score indicates greater perceived social support.

Figure C17. Average Frequency of Contact Index
The average frequency of contact index measures the average number of contacts in the past 12 months with family members and friends who are not part of the household and with neighbours. A higher score indicates more contacts.

Figure C18. Derived Health Description Index
This index assesses the respondents' perceived general health. The 5-point scale ranges from poor to excellent general health. A higher score indicates better health.

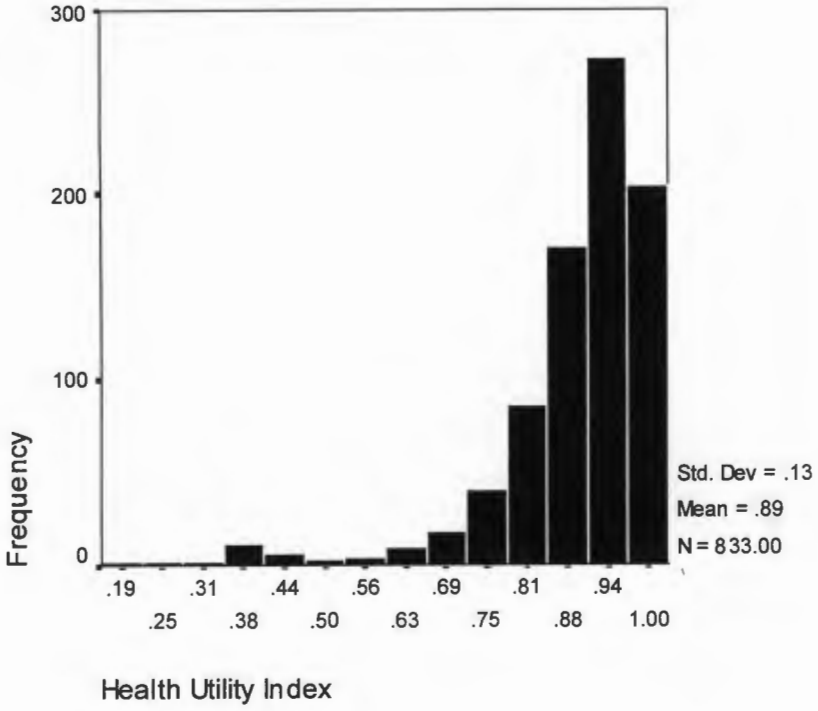

Figure C1. HUI-Mark III Score

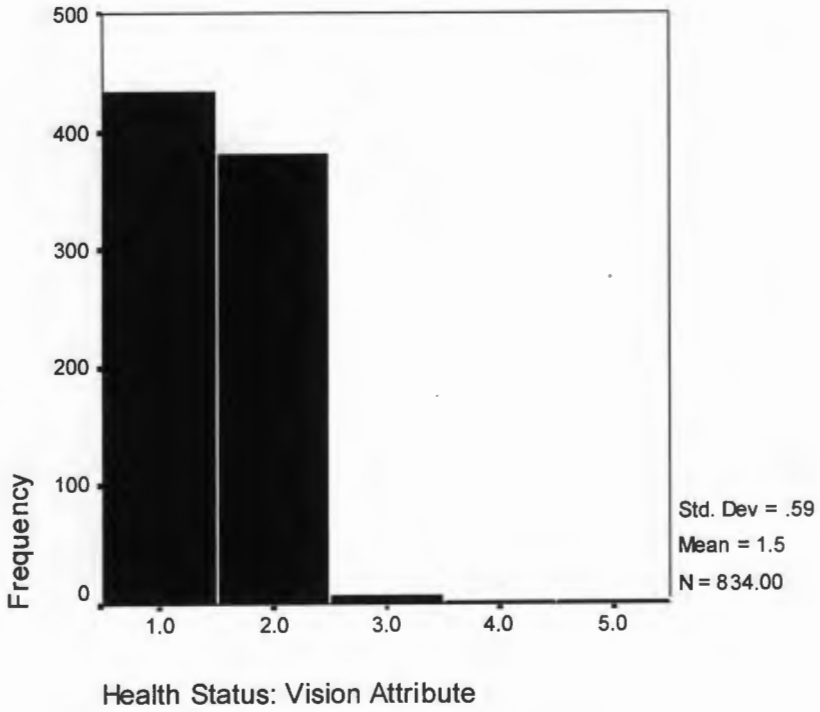

Figure C2. Health Status: Vision Attribute

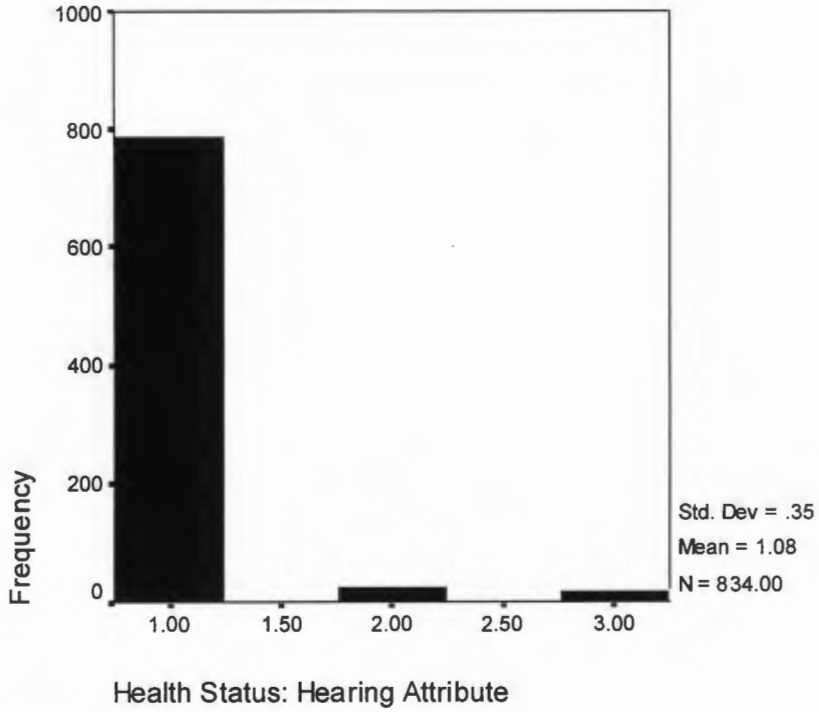

Figure C3. Health Status: Hearing Attribute

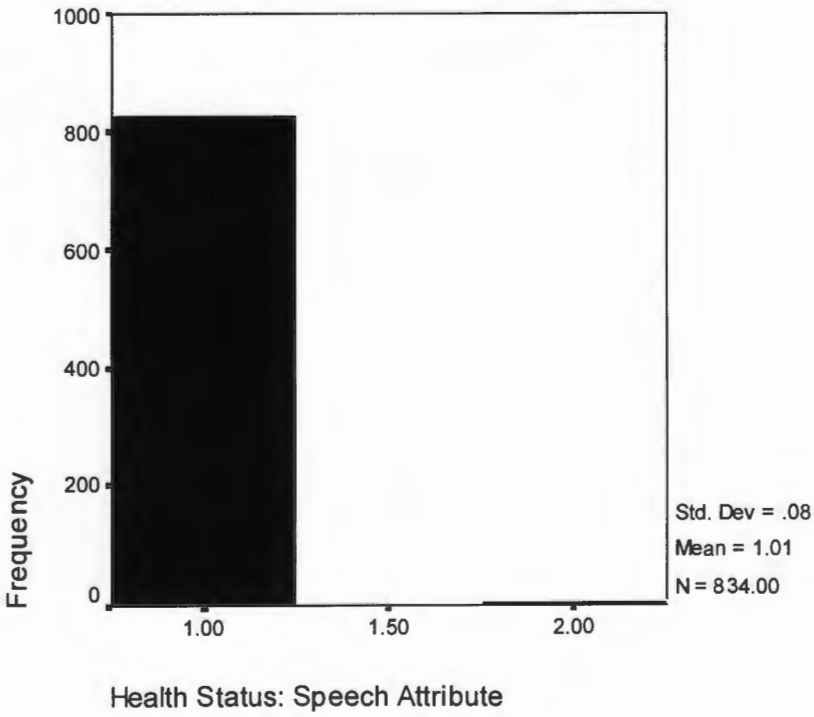

Figure C4. Health Status: Speech Attribute

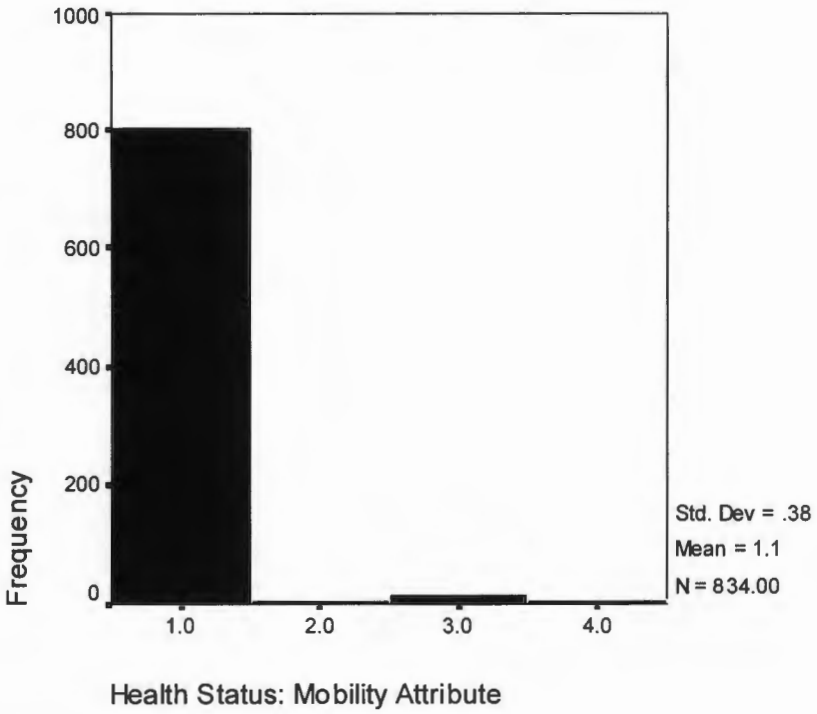

Figure C5. Health Status: Mobility Attribute

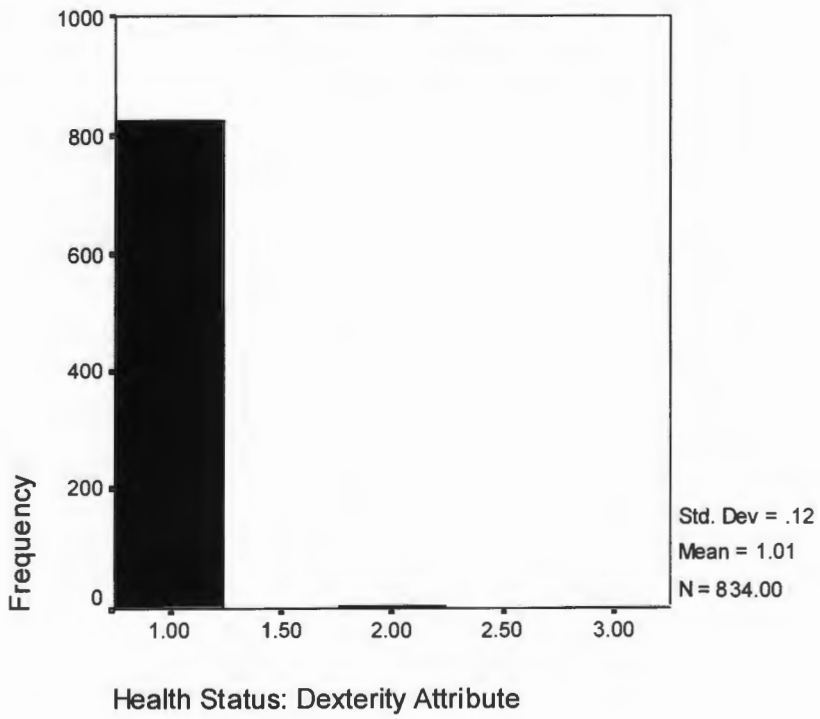

Figure C6. Health Status: Dexterity Attribute

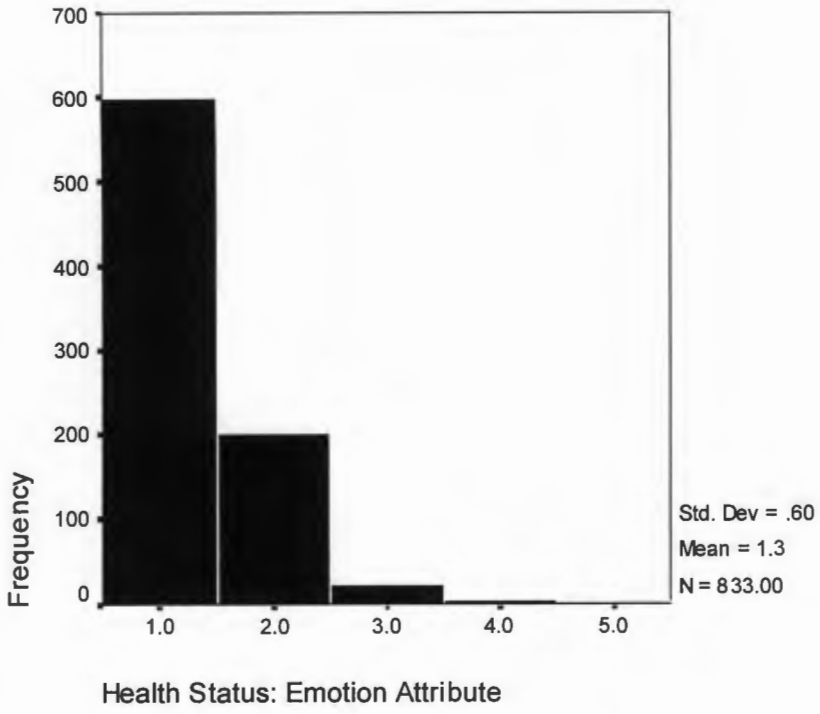

Figure C7. Health Status: Emotion Attribute

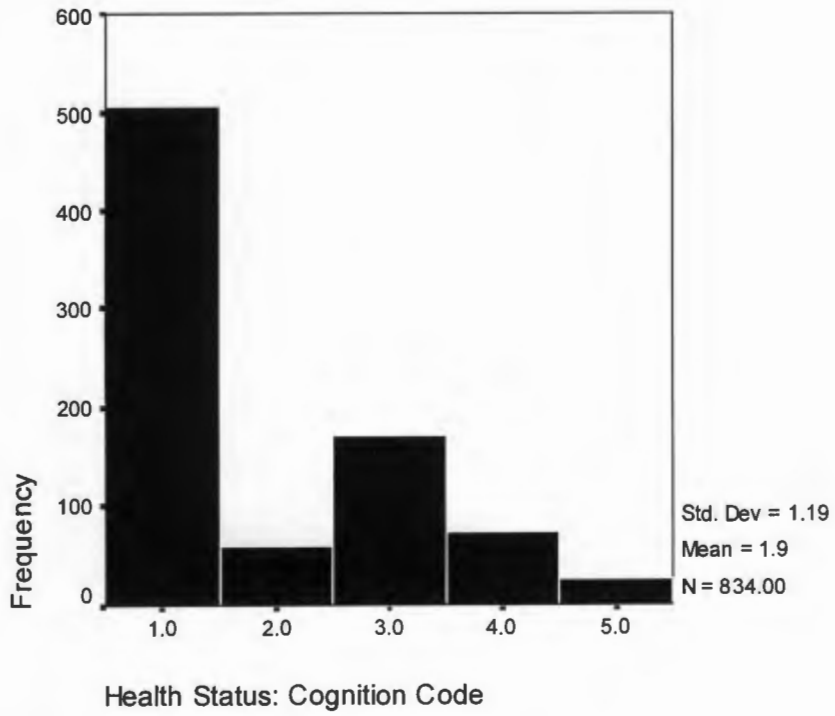

Figure C8. Health Status: Cognition Attribute

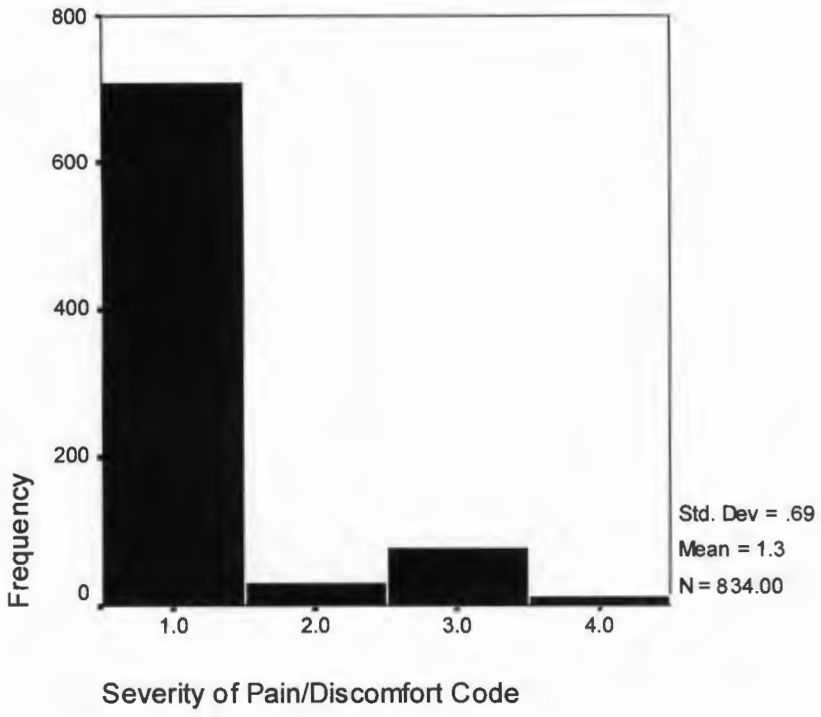

Figure C9. Health Status: Pain and Discomfort Attribute

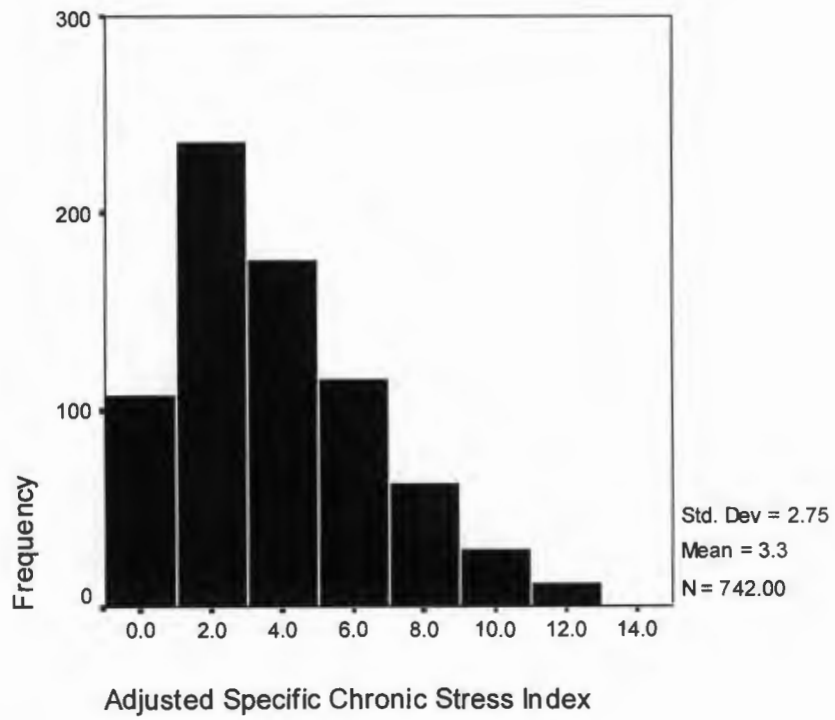

Figure C10. Adjusted Specific Chronic Stress Index

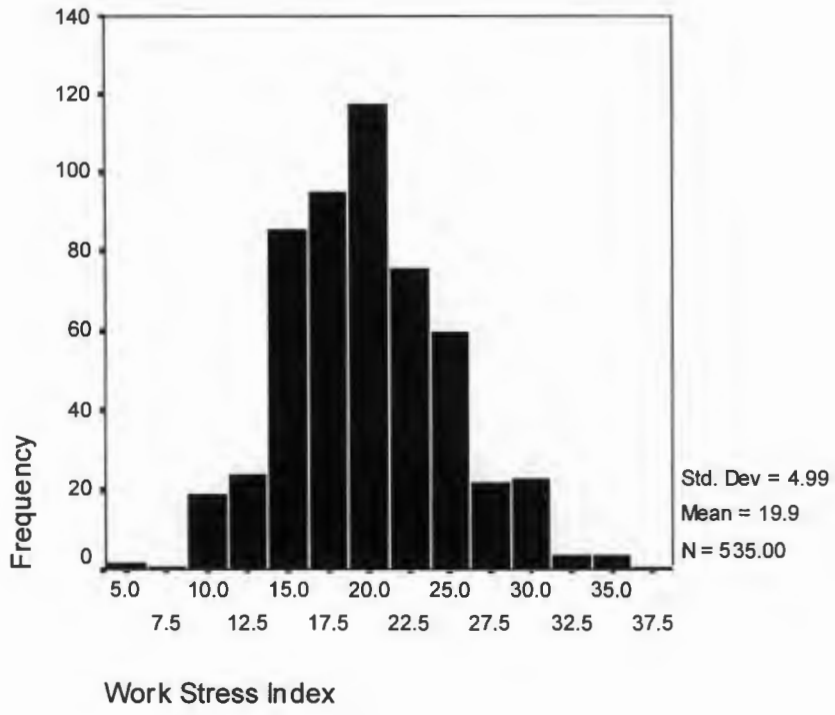

Figure C11. Work Stress Index

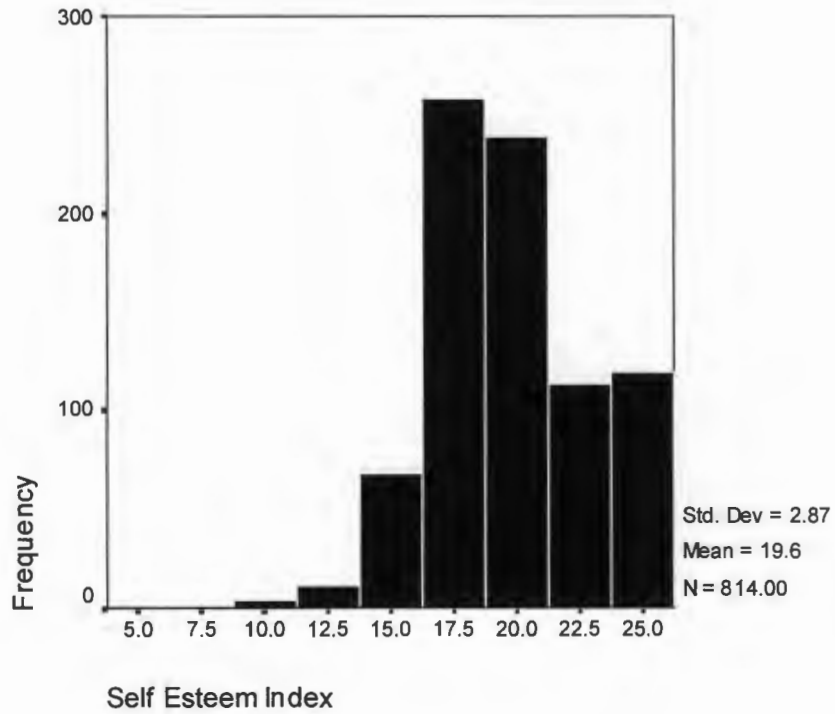

Figure C12. Self Esteem Index

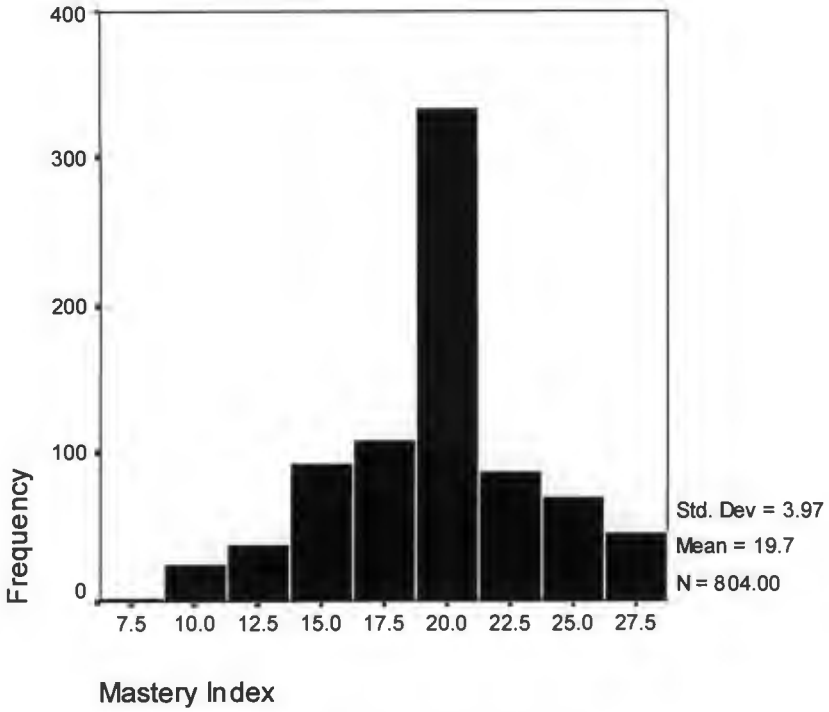

Figure C13. Mastery Index

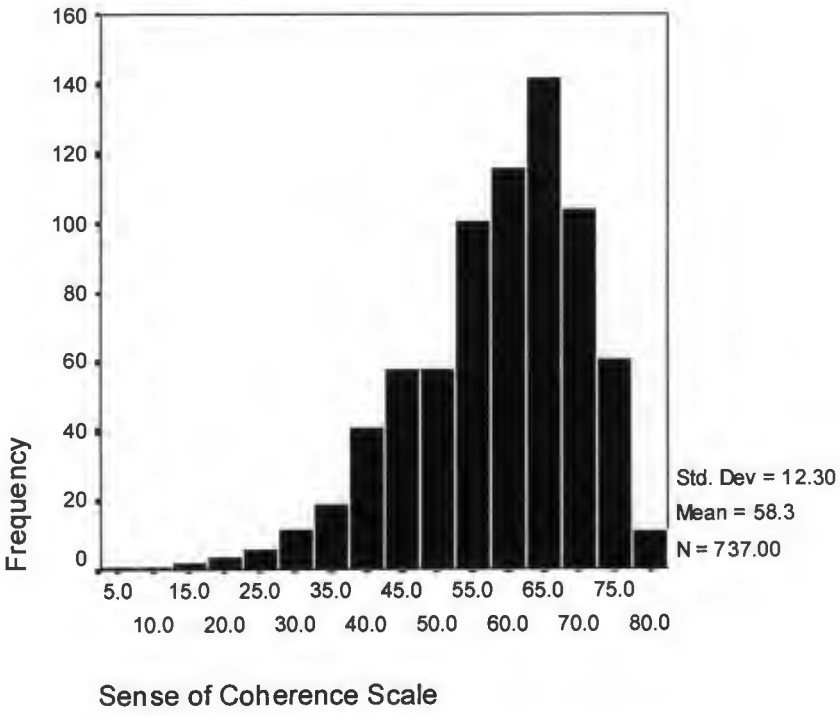

Figure C14. Sense of Coherence Scale

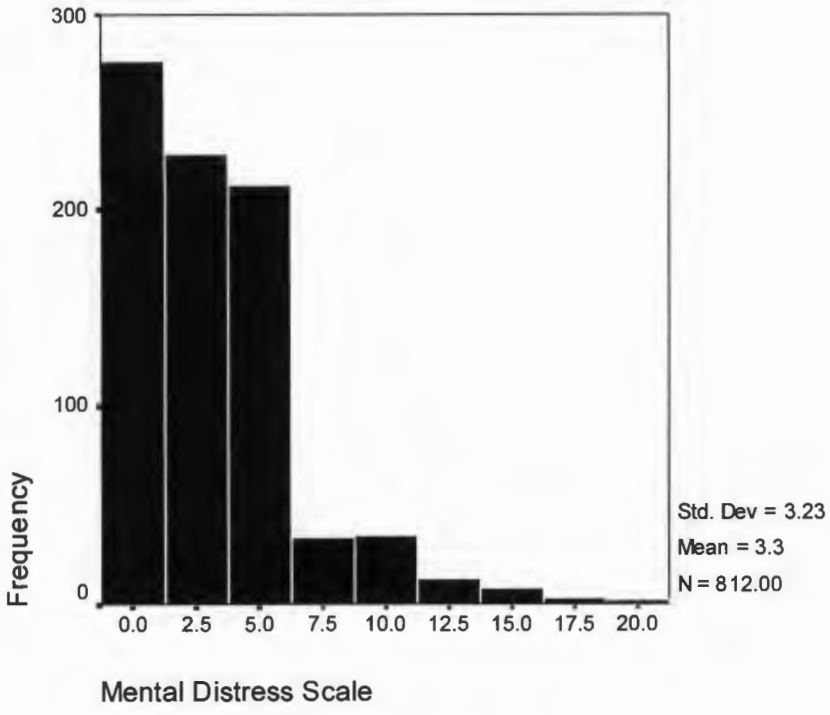

Figure C15. Mental Distress Scale

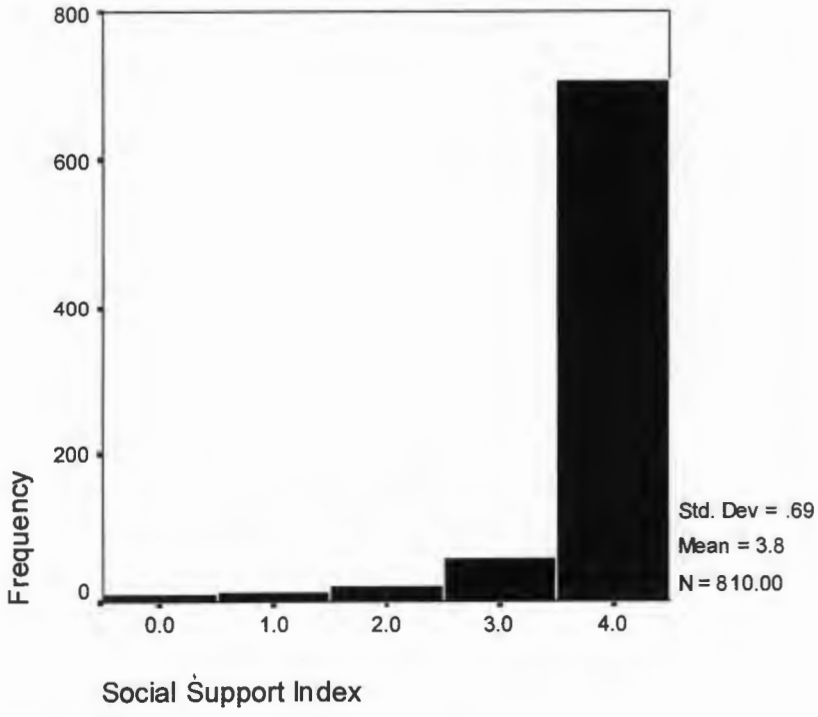

Figure C16. Social Support Index

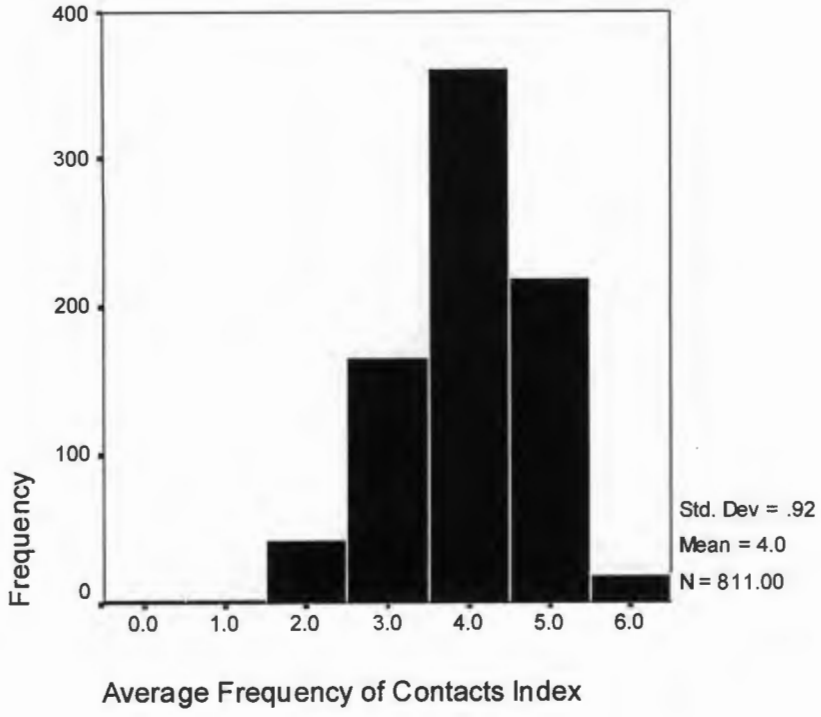

Figure C17. Average Frequency of Contacts Index

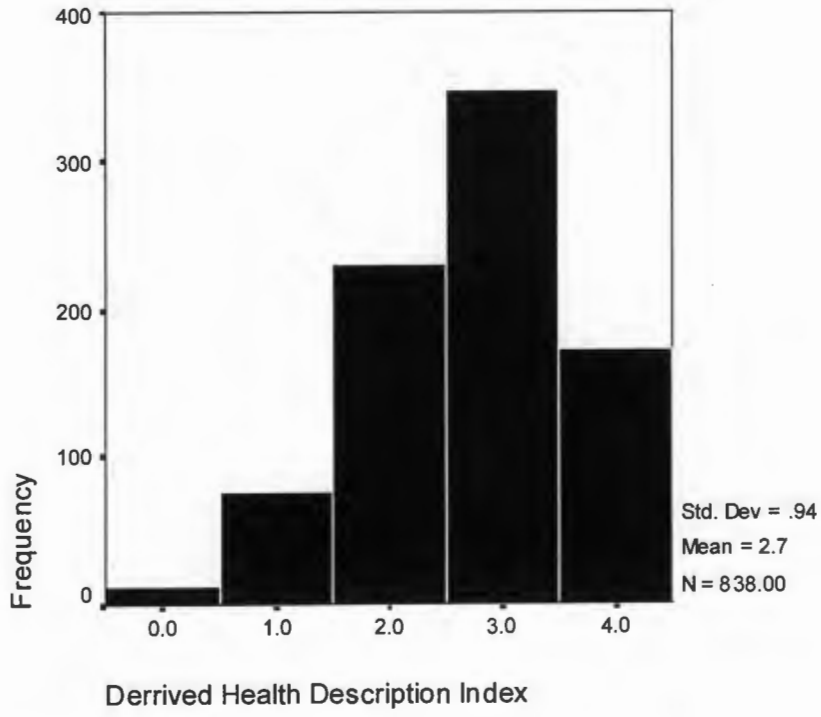

Figure C18. Derived Health Description Index