1-1-1999

A revision of the World Health Organisation psychiatric disability assessment schedule

Daniel Rock
Edith Cowan University

Edith Cowan University

Copyright Warning

A Revision of the World Health Organisation - Psychiatric Disability Assessment Schedule

by

Daniel Rock, RN, RMHN, BN 0914166

A Thesis Submitted in Partial Fulfillment of the Requirements for the Award of Master of Nursing.

At the Faculty of Communications, Health and Science, Edith Cowan University, Churchlands.

Date of submission: 21st July 1999

USE OF THESIS

The Use of Thesis statement is not included in this version of the thesis.

Abstract

The burden schizophrenia places on the individual, on the family and on society in general, is determined more by the associated disability rather than the manifest symptoms. Therefore, measurement of disability provides a baseline for evaluating change and directing treatment. This research has two main aims. The primary purpose of this study was to develop and validate a contemporary revision of the World Health Organisation - Psychiatric Disability Assessment Schedule (1988). Data used was based on data from the WHO Collaborative Study on Determinants of Outcome of Severe Mental Disorders (Sartorius et al., 1986). The second aim of this study was to develop a methodology for the construction of a psychometrically rigorous instrument. This involved the use of standard, modified and original data checking and analysis techniques and input from content-experts. Results from this process indicated that the criteria and selection of experts is based upon the considered judgement of the researcher that the sum of the advantages of the expert validation process outweigh the disadvantages. This study has two discrete outcomes: The revised DAS can be used to assess disability in schizophrenia, and a "gold standard" methodology which can be applied during instrument development.

Declaration

I certify that this thesis does not incorporate without acknowledgment any material previously submitted for a degree or diploma in any institution of higher education; and that to the best of my knowledge and belief it does not contain any material previously written by another person except where due reference is made in the text.

Signature__ _____

Date 21. July. 1999

Acknowledgments

Many individuals have been kind enough to help me with this research. To all my colleagues and friends who offered encouragement and were simply pleased for me as I progressed with the work; you know who you are and I am most grateful. Five persons require special mention.

Professor Assen Jablensky for providing the seed of the idea that this research became.

Professor John Cooper who gave me so much of his valuable time and invaluable advice.

Dr Kathy Ahern, my tireless supervisor, for all the help, encouragement, enthusiasm and critical appraisal, but mostly for the timely threats. If I was as good a student as you are a supervisor, I would have done twice as well in half the time. Thank you so very much.

Finally, Susan and Brona. I could not have done it without your support, sacrifice and enduring patience. I am blessed with you both.

Table of Contents

Abstract	2
Declaration	3
Acknowledgments	4
List of Tables	8
List of Figures	9
CHAPTER ONE: Introduction	10
Summary of the problem	13
Significance of the study	13
Purpose	14
Definition of terms	16
Limitations	19
CHAPTER TWO: Review of the literature	20
Conceptual frameworks of disability	27
Classification of disability and social roles	34
Measurement of disability	37
The role of the nurse in assessing disability	41
CHAPTER THREE: Overview of the datasets	44
Data management	50
CHAPTER FOUR: Measurement theory applied to items selection and instrument development	52
Validity	54
Measures of validity	56
Panel of experts	58

Methodology of item selection 60

Dummy data 62

Principal components analysis 63

Data cleaning and reduction 67

First draft 73

Instrument layout and scoring design 74

Coefficient of validity index 79

Second draft 80

CHAPTER FIVE: Reliability issues 85

Probabilistic error checking 86

Procedure for dataset comparison 90

Congruence of data and WHO-DAS scoring scheme 92

Coversheet 94

Memory and recall in schizophrenia 96

Rating period 101

CHAPTER SIX: Discussion, conclusion and comments 108

Datasets 108

Models of disability 111

Panel of experts 112

Lynn's model 113

Dummy data 114

Principal components analysis 116

Recommendations arising and future research 119

LIST OF REFERENCES 122

APPENDIX A: Converting the SAS dataset into an SPSS dataset 156

APPENDIX B: Expert panel 157

APPENDIX C: Dummy data included in the first draft 158

APPENDIX D: Blinded performance table of experts 160

APPENDIX E: Explained variance of each factor for each dataset 161

APPENDIX F: Eigenvalues per factor for each dataset 170

APPENDIX G: Scree plots of the three datasets 173

APPENDIX H: WHO-DAS variables and descriptors 176

APPENDIX I: First draft 178

APPENDIX J: Second draft 183

APPENDIX K: CVI ratings by expert panel members 193

APPENDIX L: Rangeplot of CVI per item 195

APPENDIX M: Draft 2 with CVI scales included 196

APPENDIX N: Final draft 209

APPENDIX O: Instrument development template 219

List of Tables

TABLE 2.1: Main characteristics of the ICIDH 33

TABLE 2.2: Main characteristics of the Cooper classification 34

TABLE 3.1: Number of patients across all collection centres that had disability data collected at each rating period 49

TABLE 4.1: PCA derived factors with a loading ≥0.8 for each dataset 71

TABLE 5.1: Binomial test of the probability of finding and error between datasets if 1850 random fields are sampled and the error rate is fixed at 0.2% 90

TABLE 5.2: Scoring scheme errors and error rate per dataset 93

List of Figures

FIGURE 2.1: Conceptual framework guiding instrument development 42

Chapter 1

Introduction

Background

The aim of this study is to develop a valid instrument to measure the disability associated with schizophrenia. Whilst there are a number of instruments that purport to measure disability in persons with schizophrenia, none of these have been developed using a data analysis perspective which aims to maximise the predictive power of the final instrument based on the analysis of an existing body of disability data. Such an analysis will provide factors that are highly predictive of the variance within the datasets and hence of disability. The incorporation of these factors into the format of an instrument will create an instrument that has quantifiable content validity.

It is important that attempts to quantify and measure disability are made, because patients have a right to receive treatment that is directed towards the best possible outcome (Human Rights & Equal Opportunity Commission, 1993).

Disability per se is such an abstraction. Any measure of disability needs to target observable entities that can be ascertained and considered to be realistic representations of components of the disability. These measures

can be considered "surrogates". The closer these "surrogates" represent the concept of disability, the more valid the measure. From this, it is self evident that if the observable indicant/abstract entity relationship is strong, analysis of the empirical indicants can lead to useful inferences about the underlying concepts; and also about the relationship *between* groups of concepts (Ducharme, Stevens & Rowat, 1994; Starker, 1986). In other words it is the relationship between the scale items and the conceptual model of disability which determines the strength of the instrument. In order to measure role disability, for example, which is abstract and unobservable (Wiersma, 1996), it is necessary to generate valid observable indicants that serve as measures of the underlying concept. The central body of this thesis is the generation of valid indicants of the more abstract and unoberservable entity of psychiatric disability.

Schizophrenia, which is defined in terms of its signs and symptoms, can be both a chronic and disabling disorder. However, recent research indicates it is the disability that seems to determine both the human and financial burden of schizophrenia (Andrews, 1991; Andrews et al., 1985; Biehl, Maurer, Krumm & Jung, 1986; Goldney, Fischer & Walmsley, 1998; McGuire, 1991; Veltro, Magliano, Lobrice, Morosini & Maj, 1994). The increasing focus on disability in schizophrenia may be due, in part, to the increasing effectiveness of the newer novel and atypical antipsychotics in suppressing the psychotic symptomatology, such that the underlying disability is revealed. It may also be due to the increasing recognition that

in order to manage patients with schizophrenia effectively, the disability of the disorder has to be targeted as a legitimate focus for intervention.

Interventions that target disability have been shown to be at least moderately effective (Knapp, 1997; Rice, 1999; Rupp & Kieth, 1993; Waslenyki, 1994; Williams & Dickson, 1995). However, with the increasing emphasis on evidence-based approaches to patient care, clinicians and researchers seek more and more refined measures to determine the effect of any therapeutic approach. In order to produce evidence about the effectiveness of interventions aimed at reducing disability, it is necessary to describe disability in schizophrenia and measure changes.

The problem of classification is further compounded by the number and orientation of the competing conceptual frameworks. There are various disability classifications. The most well known is the International Classification of Impairments Disabilities and Handicaps (ICIDH) (WHO, 1980). This has been the subject of increasing criticism, particularly directed at the lack of specificity of some of the concepts. In addition, this framework is almost 20 years old and has not been revised. One of those involved in developing the original ICIDH (1980) classification, Professor John Cooper, has sought to remedy a number of the faults within the ICIDH (1980), and has produced his own model and classification of disability based on a larger classification of behaviour in general (Cooper, 1994). It is Cooper's model that was used in this study.

Earlier attempts to classify disability have led to a number of instruments being developed that attempt to measure the concept of disability. Among the better known, and certainly the most universally applied, is the WHO-Psychiatric Disability Assessment Schedule (WHO-DAS) (WHO, 1988).The WHO-DAS (1988) although only published in 1989 was developed in the late 1970's to be used in some of the first large scale studies that attempted to quantify the degree of disability associated with schizophrenia (Jablensky, Schwarz & Tomov, 1980).

Summary of the Problem

Disability is an important consideration in the study of schizophrenia and in the planning of clinical care. Both these functions can be predicated on the availability of a valid instrument. The WHO-DAS (WHO, 1988), in its original and so far unrevised formulation, is nearly 20 years old and is in need of revision. A revision of this instrument is the purpose of this study.

Significance of the Study

This study will revise the WHO-DAS (WHO, 1988), based on an analysis of the data collected by the instrument in the WHO Collaborative Study on Determinants of Outcome of Severe Mental Disorders (Sartorius et al., 1986). In addition, this study will be used to develop and test a methodology for the validation of instruments of this type in general. The new methodology can then be used to inform the revision of existing

instruments and guide the creation of new measures. In effect, it is the purpose of this study to develop a "gold standard" validation methodology.

Purpose

The purpose of this study is to develop a single valid instrument that forms an updated revision of the WHO-DAS (1988) which measures the disability associated with schizophrenia. It is based on an existing body of research data, and developed to be congruent with Cooper's (1994) classification. It is expected the revised instrument will facilitate nursing assessment, planning, evaluation and the management of patients with schizophrenia, both in hospital and the community.

In this study, the empirical component of the proposed revision methodology is addressed in the method chosen for the primary data analysis, whilst the theoretical component is drawn from both the theory of instrument development (Coates, 1995; Fullerton, 1993; Lynn, 1986), and the model of disability that provides the instrument framework (Cooper, 1994).

Structure of the Thesis

This thesis is organised in a different manner from the "traditional" thesis because of the complex nature of the instrument development. Therefore, there is no "methodology" chapter per se. Rather, the researcher has

structured the thesis around major issues involved in instrument development, such as validity and reliability. Review of the literature relevant to this broad issue, as well as methodological issues regarding this research, will be included in the pertinent chapters. Chapter 2 will cover the review of the literature pertaining to the classification of disability, including the conceptual frameworks of disability and instrument development. Chapter 3 discusses the procedures involved in origins and cleaning up the data set. Chapter 4 will cover literature and methods used to test the validity of the instrument and the results of various analyses pertaining to instrument validity. Chapter 5 will deal with the literature, procedures and results of reliability studies. Discussion of the issues surrounding the process of development and the application of the instrument will also be discussed in relation to the literature review and the theoretical model in chapter 5. Chapter 6 will also include a summary of the thesis and discussion of the limitations of, and the implications for, this instrument as a research and clinical instrument.

Research Questions

- What are the main concepts that define disability emergent from the original data?

- What is the most efficient way of assessing these concepts within an instrument?

- What processes are required to optimally meet strict validity criteria?

Definition of Terms

Impairment: refers to the functional impairment of organs and limbs only. It does not include abnormalities of anatomical structure (except in so far as these act on function, and where it is the functional deficit that is being rated).

Personal disability: refers to the activity of the person, independent of social role such as activities of daily living.

Role handicap: refers to those behaviours that are role dependent. The definition of "role handicap" depends on a definition of social role performance. Social role performance refers to those interpersonal behaviours *that require a reciprocal social role to be performed by another individual.*

Anchor descriptor: refers to the description that forms the header of each item and is a broad description of the item content.

Descriptive stem: refers to the written description that accompanies each scaling point for each item. It contains both a description of the severity required and examples of cases that would meet the severity requirement.

Scaling frame: refers to the uniform system of classifying and rating severity across all items in the scale.

Item Content: refers to the adequacy of the sum of stem to exemplify all possible stem descriptors that may included under the anchor descriptor.

Item Scaling: refers to whether the stem descriptors have been adequately assigned on the scaling frame.

Present State: in relation to the capture period of the instrument, this refers to the 4 weeks prior to the rating date.

Lifetime Before: in relation to the capture period of the instrument, this refers to the entire period that can be recalled prior to the Present State.

Representative episode: in relation to the capture period of the instrument, this refers to a defined period in the past as determined by the rater.

Lifetime Ever: in relation to the capture period of the instrument, this refers to the entire lifetime rating and can be determined as the aggregated maximum of the Present State plus the Lifetime Before ratings.

Assumptions

It has been assumed that the variance within the original dataset is substantially related to the variance in disability within the original study cohort across time. Any series of factors that predict this data variance should also predict the disability of any individual rated against these extracted factors.

The issue of the equivalence of social role performance across cultures has been contentious (Dickson, 1996). Whilst it is often true that cultural variations in social role performance are more obvious than they are for functions and activities (the ability to climb stairs is functionally similar whether you are an Greek widow in Athens or a Native American building worker in Boston), the difference is still open to objective measurement. On the other hand, the recognition of cultural social role performance differences, and their accommodation in an assessment requires the same type of knowledge bases and judgement from the assessors. For example, a child and adult are physically different, but the assessment of range-of-motion in both requires similar knowledge and judgement of the assessor. Therefore, although different standards and contexts of judgement might be applied to individuals, the ability to make such judgements is assumed to be well within the capability of an appropriately trained health professional. This is the premise on which this revision of the WHO-DAS has been developed as an *international* instrument for assessment of schizophrenia-related disability.

Limitations

The primary limitation of this research was that the data is nearly twenty years old and reflects the content and historical context of the original WHO-DAS (1988). This meant that some of the factors extracted could not be used as items in the revised instrument. In total, five factors were discarded because they were determinations of patient management that are no longer practiced or relate to uncommon situations and therefore would not be rateable for the majority of patients with schizophrenia. Fortunately, these discarded factors were not highly powerful and in fact only "explain" a little over eight percent of the data variance. In addition, the analysis relies on data from only the first two years of the study as this was the WHO's period of involvement. Since that time each of the original participating centres has continued to follow up its original cohorts. These data have not been centrally aggregated. If these data had been available for this analysis it may have provided more robust evidence of enduring disability variance for the factor analysis, as well as providing data on the both those patients whose illness follows a more refractory course to compare with those with good outcome. This would allow the identification of those items that have high positive predictive value, as well as those with high negative predictive value.

Chapter 2

Review of the Literature Pertaining to the Classification of Disability

Introduction

Despite disagreements about the concepts and models of psychiatric disorder, there is a general consensus within the scientific community on the value of classification (Bertelsen, 1999; Jablensky, 1988). However, this consensus is set against a background of psychiatric disorders that are in many cases merely refined conjectures (Ryan, 1991). Without a total knowledge about aetiology, most psychiatric classifications rely on the grouping of disorders according to co-occurring signs and symptoms. Signs and symptoms, because they are perhaps more obvious and cross-sectionally apparent, are often used as convenient measures for the presence of illness. However, the pragmatic decision to use signs and symptoms as grouping variables may not reflect the importance of other, perhaps less easily determined, illness variables. An illness is not simply the sum of the signs and symptoms.. Illness itself, and chronic illness in particular, is a much more complex and difficult to measure phenomenon.

Schizophrenia, for example, is a condition defined only in terms of certain abnormalities of experience and behaviour since there is, as yet, no test or biological marker that can screen for the presence of the disorder (Carpenter & Buchanan, 1994; Coleman & Gillberg, 1997; Provencher, Fournier & Dupuis, 1997). Lacking knowledge of the supposed biological

cause, there is room for a wide range of opinion about what defines the condition (Wing, 1995).

It is useful, however, to classify schizophrenia both in terms of its presentation and its consequences (Johnstone & Lee, 1994). The presentation allows us to recognise broadly the condition. Knowledge of consequences allows us to predict the pattern of the course. However, the problem with classifying the presentation has been likened to attempting to classify clouds; it is pragmatic rather than explanatory (Wing, Sartorius & Wing, 1998). Whilst it is accepted that there is no "conclusively defined disease known as schizophrenia" (Janzarik, 1987), it is important to create the best possible classification, both in terms of usefulness and completeness. Such labelling provides an indispensable basis for communication and further investigation (Frances & Egger, 1999; Jablensky, 1999; Robins & Helzer, 1986; Roman, 1971; Sartorius, 1993; Sartorius et al., 1993).

Classification Systems

The two most common systems used to classify psychiatric disorder are the World Health Organisation's International Statistical Classification of Diseases and Health Related Problems, tenth revision (ICD-10) (1992) and the fourth edition of the American Psychiatric Association's Diagnostic and Statistical Manual (1994). Both systems use a multiaxial diagnostic approach. The principal or Axis I diagnosis, in both systems, relies on

applying the appropriate diagnostic algorithm to the described signs and symptoms. In essence, the patient describes signs and symptoms and the clinician then applies the diagnostic rules, as set out in the particular manual, in order to reach the most likely diagnosis.

In addition to the Axis I diagnosis there is also the facility in both systems to classify illness consequences on other axes. In this manner, disability statements enter the diagnostic formulation as important independent illness dimensions to be considered when making a diagnosis. The need for disability statements comes from the increased recognition of disability as the most prominent clinical feature of many chronic psychiatric disorders such as schizophrenia (Davidson & McGlashen, 1997; Fenton & McGlashen, 1991; Harding, Zubin & Strauss, 1987; Mason et al., 1995).

Importance of Measuring Disability rather than Signs and Symptoms in Mental Health

Numerous attempts have been made to classify disabilities. The most prominent and widely used is the International Classification of Impairments, Disabilities and Handicaps (ICIDH) (World Health Organisation (WHO), 1980).

According to the ICIDH (WHO, 1980, pp. 143) "In the context of health experience, a disability is any restriction or lack (*resulting* from an impairment) of ability to perform an activity in the manner or within the range considered normal for a human".

The importance of considering disability when examining illness can be highlighted using schizophrenia as an example. The morbid risk for schizophrenia is around one percent, and this is essentially constant throughout the world (Sartorius et al., 1972). Despite this low morbidity, patients diagnosed with schizophrenia account for the second largest proportion of hospital admission days and the majority of the medium to high intensity community management resources (Biehl, Maurer, Schubart, Krumm & Jung, 1986; McGuire, 1991). This is mainly due to the patient's inability to function at a normal level, even when the acute episode has passed (Davies & Drummond, 1994; Kirkby, Dadiels, Jones & McInnes, 1995).

Furthermore, the consequences of prolonged or repeated institutionalisation may result in an increased disability burden for the patient even when the illness is in remission (Johnstone, Owens, Crow & Gold, 1981). Therefore, the measurement of disability provides a more global and applicable realistic standard of the *consequences* of having been ill (Janca et al., 1996). This in turn has implications for the wider social community, since being able to measure disability is becoming increasingly important as an indicator of health service funding (Everst & Ament, 1995).

Finally, the development of a measure of disability, rather than signs and symptoms of an illness, is important because disability includes many of the influences that may act on the individual during the course of his/her

illness, rather than a particular pathological process characterised by the medical model of disease (Mechanic, 1995). This provides a more holistic approach to the care of the mentally ill.

Disability as a Measure of Illness Outcome

Disability is the optimal measure for serious mental illness for several reasons. In a landmark study examining schizophrenia, Strauss and Carpenter (1981) found that commonly used outcome measures, such as social function, symptomatology and need for resources, were only moderately correlated. Parker and Hadzi-Pavlovic (1990) found that, in contrast to symptom severity, which was very poorly predictive, disability was a reasonable predictor of relapse within the first four years post index admission. Similarly, Shepherd, Watt, Faloon and Smeeton (1989) found that clinical outcome at year five post index was "highly correlated" with disability at year one. Interestingly, they also found that comorbid depressive symptoms were reduced by 50% over the same period. This supports the view that disability is a predictor of clinical outcome, rather than the view of Glazer, Prusoff, John and Williams (1981), that disability measures are simply correlations of affective state. Furthermore, the level of the post-morbid rather than pre-morbid disability offers the greatest predictive power (Hubschmid et al., 1991).

The conclusion can be drawn that the greatest burden of psychotic illnesses does not seem to be the acute or residual symptoms but the on-

going disability (Crow, McMillan, Johnson & Johnstone, 1986; Glick & Zigler, 1980; May et al., 1981; Veltro, Magliano, Lobrice, Morosini & Maj, 1994). It is this separateness of disability from positive symptoms that is such a prominent feature of schizophrenia (Strauss & Carpenter, 1981).

It is tempting to assume that the signs and symptoms of a psychotic disorder can be taken as convenient surrogates for disability. Whilst signs and symptoms may seem to be highly correlated with disability, in reality they are frequently "desynchronous" (Sheehan, Harnett-Sheehan & Raj, 1989). This separateness is clearly highlighted by Conway, Malzer and Hale (1994) who studied a cohort of 51 schizophrenics who had been out of hospital at least one year but remained psychotic and dysfunctional. Over a three year period a high intensity programme targeted the "high levels" of psychotic symptomatology within the group. The strategy was remarkably successful. However, the researchers had to conclude that despite significantly reducing the proportion of patients with psychotic mental states, improvement in social functioning ("global disability") did not follow. This supports Wiersma's (1996) suggestion that disability is a semi-separate illness outcome measure. For example, Marneros (1989) found that while 10% of persons diagnosed with schizophrenia had no residual disability at three year follow-up, 14% were found to have "extremely severe maladjustment". Such a view is also taken by the WHO who separate signs and symptoms from disability both in conception and aetiology (WHO, 1980).

Patients with chronic mental illness come into hospital and stay, not primarily because of symptoms, but because of symptoms coupled with disability. It is the disability, not the symptoms, that best predicts the length of hospital stay and, as a consequence, cost (Beihl, Maurer, Schubart, Krumm & Jung, 1986; Goldney, Fischer & Walmsley, 1998; Pary, Turns, Stephenson, Tobias & Lippmann, 1992; Rossi et al., 1989). In New South Wales for example, the cost of the in-patient treatment of schizophrenia is half that of myocardial infarction, yet schizophrenia is twelve times less common (Andrews, 1991; Andrews at al., 1985; Hall et al., 1985)! With an estimated twenty million persons worldwide suffering from schizophrenia (Sartorius & de Girolama, 1991) and upwards of 50% of these having a chronic course, the human and financial cost is considerable (Rupp & Keith, 1993).

Disability, Hospitalisation and Treatment Outcomes

Diagnostic Related Groups (DRGs) have been put forward as the best measure of clinical and management performance (Dada, White, Stokes & Kurzeja, 1992). DRGs are heavily weighted on the ability of diagnoses to predict length of stay (LOS) (Schumacher, Namerow, Parker, Fox & Kofie, 1986; Creed, Tomenson, Anthony & Tramner, 1997). However, McCrone and Phelan (1994) found, in a study involving 5482 psychiatric in-patients, that DRGs only predict around three percent of the LOS variance. Ashcraft et al., (1989), reported that combining DRGs with dual diagnoses data could only account for around 11% of the LOS variance. A similar outcome

was found by Davis, Lowell and Davis (1993) using a computerised Artificial Neural Network (ANN) which performed at least as well as a team of independent clinicians.

However, when Lowell and Davis (1994) refined the input data and "retrained" the ANN on a further 1064 cases that included an extensive array of disability items, the ANN was now able to predict between 35-70% of the LOS variance. The improvement was attributed to the addition of disability data. Such a result should not have been unexpected for as early as 1977 Munley et al. (1977) found that by using stepwise multiple regression of disability criteria, both LOS and three month relapse rate could be predicted (R~0.45).

In addition, services that target disability associated with mental illness have been shown to be the most efficacious both in terms of outcome and cost (Knapp, 1997; Rice, 1999; Rupp & Keith, 1993; Wasylenki, 1994; Williams & Dickson, 1995). Even if a health service does not identify or focus resources towards disability, disability will remain a significant factor in service utilisation (Kent, Fogarty & Yellowlees, 1995).

In summary, disability seems to provide a useful measure of patient service utilisation (and cost), as well as offering some reasonable degree of clinical prediction about course of illness. It is important that attempts to quantify and then measure disability are made because patients have a right to receive treatment that is directed towards the best possible outcome

(Human Rights & Equal Opportunity Commission, 1993). Furthermore, measurement provides a baseline for evaluating change and directing treatment (DeJong, Giel, Slooff & Wiersma, 1985; Phelan, Wykes & Goldman, 1994; Smith, Waxman, Snyder & Raphael, 1996; Wiersma, Giel, DeJong & Slooff, 1996).

Conceptual Frameworks of Disability

Before going on to examine specific disability measures in more detail, it is useful to review conceptual frameworks of disability. These can provide a reference against which the disability measures can be judged for completeness of coverage. It is the conceptual framework that sets the scope of the concept which instruments and measures attempt to capture. Two conceptual frameworks will be discussed because the framework used in this research (Cooper's framework) was developed from an original framework devised by the WHO (The International Classification of Impairment, Disability and Handicaps). Both models define disability in terms of illness outcome.

International Classification of Impairment, Disability and Handicap (ICIDH)

The ICIDH (WHO, 1980) has become the standard system for the classification of the consequences of illness. The ICIDH (WHO, 1980) is a tri-partite classification that consists of separate classifications for

impairment, disability and handicap. In the original formulation these are seen as causal, in that illness may cause impairment, which may cause disability, which in turn may cause handicap. In this way, if the ICIDH classification is followed, it is impossible for an individual to have a disability without being impaired.

Limitations of the ICIDH.

Problems have been reported that centre mainly on the overlap between the three concepts of impairment, disability and handicap. There is an apparent discrepancy between concept definitions and some of the specific items listed as being within a particular category. For example, whilst impairment is centered around the structure or function of an *organ*, psychological impairment is manifest and defined in the ICIDH (WHO, 1980) at the level of the *person* (a disability). Dickson (1996) suggests this leads to a circular argument: the presence of a behavioural disability is a measure of the psychological impairment manifest in the behavioural disability.

Much of criticism of the ICIDH (WHO, 1980) relates to the delineation of "disability" from "handicap" at the level of classification by the clinician. Many of the problems arising from the delineation of "disability" and "handicap" can be traced back to the process under which the ICIDH (WHO, 1980) was developed.

Whilst the final draft of the ICIDH (WHO, 1980) was in preparation, the WHO Division of Mental Health was preparing the first version of what was to become the WHO-Psychiatric Disability Assessment Schedule (WHO-DAS) (WHO, 1988). The WHO-DAS (WHO, 1988) was developed for use in the WHO Collaborative Study on Impairments and Disabilities associated with Schizophrenic Disorders (Jablensky, Schwarz & Tomov, 1980).

There was some interchange between the developers of the ICIDH (WHO, 1980) and the WHO-DAS (WHO, 1988), which resulted in several sections of the latter instrument being included in the former classification. For example, the WHO-DAS (WHO, 1988) includes social role performance as a domain of "disability". This, according to the conceptual definition ("fulfilment of social role") meets the criterion for a "handicap" in the ICIDH (WHO, 1980). However the ICIDH (WHO, 1980) developers placed the WHO-DAS (WHO, 1988) items which met the criterion for "handicap" in the "disability" domain (Badley, 1987). These include section 17 "Family role handicap" and section 18 "Occupational role handicap").

Cooper's Classification

In order to overcome this problem of overlap, Cooper (1984) suggested that whilst the concepts "impairment", "disability" and "handicap" are useful, the practical application of a classification of illness *consequences*

should enable consistent agreement between the operational definitions (impairment, disability and handicap) and the categories and content of the framework. Cooper (1984) proposed a similar framework to the ICIDH (WHO, 1980) but used more specific operational definitions for the concepts equivalent to impairment, disability and handicap. He used the term "equivalent concepts" in the sense that the concepts, of illness consequences map the same domain as the ICIDH (WHO, 1980). Copper's classification also broadly encompasses the theoretical underpinnings of the WHO-DAS (WHO, 1988) (at least in regard to "disability" and "social role performance"). Cooper's classification also draws from his 1983 paper argued that disability be placed in the context of a suggested classification of normal behaviour (Cooper, 1983). In other words, disability, by definition is a deficit from the norm.

Cooper's classification is based on the concepts of "impairment", "personal disability" and "role handicap. These three concepts are defined only in terms of how they impact on a person's performance of a task. "Impairment" refers to the functional impairment of organs and limbs only. It does not include abnormalities of anatomical structure (except in so far as these act on function, and where it is the functional deficit that is being rated). "Personal disability" is centered around the activity of the person independent of social role. "Role handicap" is reserved for those behaviours that are role dependent.

The definition of "role handicap" depends on a definition of social role performance. Social role performance refers to those interpersonal behaviours *that require a reciprocal social role to be performed by another individual.* As in the ICIDH (WHO, 1980), "role handicap" is simply a deficiency in expected social role performance. It is the necessity of requiring another person performing a reciprocal social role that separates "role handicap" from "personal disability", in Cooper's classification.

The underlying principle of Cooper's classification is that there should be no overlap between concepts, and that the system it describes is a linked rather than causal. In other words, each of the concepts (impairment, personal disability and role handicap) can be present independently of the others. The boundary between impairment (function based) and personal disability (activity based) can be clarified by specifying that a function does not involve interaction with the environment while a disability does. Cooper's classification (1994)is summarised (Table 2.2) in order to provide a contrast to the summary of the ICIDH (1980) (Tables 2.1).

Table 2.1

Main Characteristics of the ICIDH

	ICIDH		
	Impairment	Disability	Handicap
Definition	Centered around structure or function (organ)	Centered around activity (person)	Centered around social role
In spite of the ICIDH definition the content refers to:	Many structures + Many functions + Some activities + A few social role performances	Some function + Many activities + Some social role performances	Some functions + Some activities + Some social role performances

Table 2.2

Main Characteristics of the Cooper Classification

	Cooper's classification		
	Impairment	Disability	
	Impairment	Personal disability	Role Handicap
Definition	Centered around function	Centered around activity	Centered around social role
According to Cooper's definitions content of the classification refers to:	Functions only	Personal activities only	Social roles only

Classification of Disability and Social Roles.

Although Cooper (1994) places "personal disability" and "role handicap" under the general heading of disability, these are clearly separate concepts, as discussed above. Cooper's (1994) bi-partite concept of disability has an advantage over the ICIDH's separate "disability" and "handicap" classifications because it has a stronger theoretical base and better coverage of role performance than the ICIDH (WHO, 1980). An example may help to illustrate the advantage of Cooper's "personal disability" and "role handicap" classifications over the ICIDH "disability" and "handicap" classifications.

Consider a patient with schizophrenia who is unable to differentiate the voices of those around him from auditory hallucinations (ICIDH "disability in understanding speech"), is neglectful of his self-care (ICIDH "personal hygiene disability") and harasses his neighbours whenever they try to work in their garden (ICIDH "other social role disability"). These are all included in the disability section of the ICIDH (WHO, 1980) although this is inconsistent with the ICIDH definition of disability ("... any restriction or lack (resulting from an impairment) of ability to perform an activity in a manner or within the range considered normal for a human being" (WHO, 1980, pp. 143). Confusion arises because whilst all of the above represent disabilities in the ICIDH, if the person were thought disordered rather than hallucinated he has an "impairment". If the person is disorientated due to sensory hallucinations this would be classified as an "orientation

handicap". If his lack of care suggested that assistance from others could remediate the problem then, this is categorised in ICIDH (WHO, 1980) as a "physical independence handicap". This inconsistency between definitions, categories and content is one of the major criticisms of the ICIDH.

The application of the Cooper (1994) classification to the above example demonstrates how Cooper's classification is more consistent and less confusing: the inability to understand speech, whether or not this is associated with auditory hallucinations or thought disorder, is an "impairment". The poor self-care is a "personal disability" whether or not it can be remediated with external help. The poor relationship with the neighbours represents a "social role handicap" and disorientation would be a "personal disability".

In the above example the ICIDH identifies two social role handicaps. The first is in regard to the relationship with the neighbours. This would also be classified as "role handicap" in the Cooper classification (1994). In addition, however, if the patient's poor self care could be remediated with assistance, and it is difficult to think of an example when this would not be the case, then this is also an example of social role handicap given in the ICIDH (1980). This second example is clearly at variance with the ICIDH definition of handicap as "a disadvantage ... that limits or prevents the fulfilment of a social role that is normal ... for the individual" characterised "by a discordance between the individual's performance or status and the

expectations of the individual ...". In addition, disorientation, which is independent of social role, is also classified as a "handicap" in the ICIDH, presumably because of its possible effect on social roles rather than because it is a social role handicap in itself.

This lack of consistent delineation within the body of the ICIDH makes it very difficult to be used as a conceptual framework around which to develop a measure of disability. However, it is evident from this example that Cooper's classification, unlike the ICIDH (WHO, 1980), allows the performance of an individual to be consistently assessed at all three levels without overlap, thus enabling an examination of how these mutually exclusive domains might be related.

Use of Cooper's system would be of potential interest for clinicians who carry out the detailed clinical assessment of individuals, in which the number and patterns of impairments, personal disabilities and role handicaps for groups of individuals with different diagnoses are examined. For example, schizophrenia will usually produce a number of performance interferences at all three levels suggesting that a pervasive disturbance is present which encompasses the most basic to the most complex levels of task performance. By contrast, phobic reactions may have most entries at the role performance end of the classification.

Conclusion

Copper's classification is clearly an improvement on the ICIDH (WHO, 1980). The definitions of "impairment", "personal disability" and "role handicap" are straightforward, unambiguous and mutually exclusive, and overcome the definitional boundary problems of the ICIDH (WHO, 1980). However, the vast majority of the ICIDH (WHO, 1980) variables, modified where necessary to meet the new conceptual definitions, have been retained by Cooper. For these reasons, Cooper's classification has been used as the conceptual framework of disability used in this revision of the WHO-DAS (WHO, 1988).

Measurement of Disability

Comparing disability measures against Cooper's classification (1994) allowed the researcher to firstly select a model that best suited the research purpose, and then select an instrument that best captured the data which fit the model. Whilst there are many instruments that claim to measure disability, only three will be discussed here as exemplars of type. These are quality of life measures, social disability schedules and psychiatric disability schedules.

Quality of Life Measures

Quality of life (QoL) is an umbrella concept and involves a multiplicity of dimensions, often including components of disability. QoL can include all, or a combination of: personal sense of well-being, distress, satisfaction with housing/income, one's perceived health, feeling safe, social networks, satisfaction with social support and morale. The usefulness of the general concept has gained increasing support (Bullinger, Anderson, Cella & Aaronson, 1993; Greer, 1987; Holmes, 1989; Pocock, 1991; Veldhuyzen, 1991; Zhan; 1992). McDowell and Newell (1987) describe QoL as a "convenient" but "nebulous" term that is "intuitively familiar" and, because of this familiarity, has suffered from a lack of examination. QoL scales as a group have been criticised for being psychometrically weak (Barnett, 1991; Bergner, 1989; Deyo & Patrick, 1989; Jenkins, 1992), with the major criticism being the lack of an operational definition of the concept. Much of the above discussion can be highlighted with an example.

The Lehman Quality of Life Interview (Lehman, 1988) is well known and widely used. It covers current living situation, daily activities and functioning, finances, work/school, legal and safety issues. However, only the "daily activities" section would conform to Cooper's (1994) classification of "personal disability". Whilst the scales have excellent internal reliability and scale independence, the validity of the concepts included is at best moderate (Lehman, Postrado, & Rachuba, 1993) . This

type of scale, whilst containing components of disability, is not suitable as a derived measure of disability.

Social Disability

There are a number of instruments that were specifically designed to measure social function in the psychiatric population. The Groningen Social Disabilities Schedule (SDS) (Wiersma, DeJong & Ormel, 1988) would fall into this category. This is a measure that defines disability in terms of social role deficits analogous to Cooper's concept of role handicap. The various items allow the user to rate performance across a defined range of expected social roles.

However, such scales have been shown to have a clear ceiling effect when used to examine persons with schizophrenia. Many schizophrenics have global deficits in terms of social role, or have such a reduced range of social roles they would be rated as having "maximum disability" for much of the time. Given this, such scales are not the best instruments to assess within group differences in illness outcome, particularly at the more chronic end of the spectrum. For this reason the SDS has found limited uptake in schizophrenia research.

Psychiatric Disability Assessment Schedules

The most widely tested psychiatric disability measure is the WHO – Psychiatric Disability Assessment Schedule (WHO-DAS) (WHO, 1980) and is the instrument upon which this study is based. The WHO-DAS was developed for use in two WHO studies[1] (Jablensky, Schwarz & Tomov, 1980; Sartorius et al., 1986). The WHO-DAS was developed in tandem with, and mirrors the theoretical underpinnings of the ICIDH (WHO, 1980) (Ustun, Cooper, Van Duuren-Kristen, Kennedy, Hendershot & Sartorius, 1995) so that the Cooper classification is also congruent with the items of the WHO-DAS (1988). Furthermore, the WHO-DAS (WHO, 1988) is the instrument of choice for WHO sponsored or collaborative studies, and has added a degree of standardisation to the measurement of disability (Satrorius & Janca, 1996). However, the WHO-DAS (WHO, 1988) requires revision. Like all the other disability measures, it reflects the social circumstances in which it was constructed. Furthermore any revisions to this instrument need to be based on scientific, rather than intuitive, principles (Leckman et al., 1997).

[1]The WHO Collaborative Study on Impairments and Disabilities associated with Schizophrenic Disorders and the WHO Collaborative Study on Determinants of Outcome of Severe Mental Disorders)

The role of the Nurse in Assessing Disability

Barker (1991) puts forward the view that a primary function of mental health nursing lies in the capacity of the nurse to effect change in the patient that is directed towards minimising the social dysfunction associated with mental illness. It is in this area of social management that the nurse, as part of the multidisciplinary team, can create the most sustained change (Bebbington & Kuipers, 1982; Rosen, Slade & Shankar, 1995). There is a mounting body of evidence that such interventions, focusing as they do on disability, produce "favourable" outcomes in patients with serious mental illness (Brooker, Turner, Barraclough, Butterworth & Goldberg, 1992; King & Nazareth, 1996; Prendergast, 1995; Remington, 1995; Starkey & Flannery, 1997). Despite this, the ability of mental health nurses to accurately assess changes in functional status over time has been questioned (Boyd & Luetje, 1992). This in turn might affect nursing interventions in the treatment of schizophrenia, and Gournay (1996a, b) suggests that because of this, the use of psychosocial interventions are more talked about than practised.

With the increasing importance of evidence based practice, it is important that nurses have the opportunity to use measures which examine areas of their practice that are most likely to generate the biggest effect size (Sullivan, 1998). Many studies have shown that the area where mental health nursing produces the greatest effect is in the remediation of illness consequences (Armytage, 1995; Gournay, 1995; Lancashire, 1997). In

other words, nurses should not be judged on their ability to alter symptoms but rather on how they manage their patients' disabilities.

Conceptual Framework

The conceptual framework for this study was developed from the concepts of instrument development reported in the literature and depicted in Figure 2.1 below.

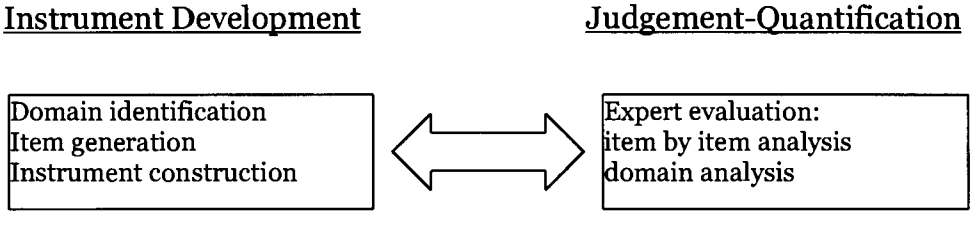


Figure 2.1. Conceptual framework guiding instrument development

Davis (1996) and others (De Vellis, 1991; Gable & Wolf, 1993; Lynn, 1986; Nunnally & Bertstein, 1994) separate instrument construction into two parallel processes. There is the process of instrument development consisting of domain identification, item generation and instrument construction. Concurrent with this runs the separate process of judgement-quantification. This is a process of expert evaluation of the content, both item-by-item and as item groups (De Vellis, 1991; Gable & Wolf, 1993). In reality, there is much communication between these two processes, since the expert evaluation informs and guides instrument construction. The revised instrument is then re-evaluated by the experts.

This process continues until there is consensus by the experts that a valid instrument has been developed. This is the conceptual framework which guided the development of the proposed instrument.

Chapter 3

Overview of the Datasets

Introduction

This chapter describes the origins of the data used in the development of this revision of the WHO-DAS (WHO, 1988). It includes a discussion of the growing awareness, resulting from analyses of the studies from which this data has been drawn, that the course of disability related to schizophrenia is separate from the course of signs and symptoms of schizophrenia.

Original Datasets

The three datasets (identified as dasiu, dasfu1, dasfu2) used in this study were collected, using the pre-publication version of the World Health Organisation - Psychiatric Disability Assessment Schedule (WHO-DAS) (WHO, 1988), as part of the WHO Collaborative Study on Determinants of Outcome of Severe Mental Disorders (Sartorius et al., 1986).). For clarity the three datasets will be referred to according to their original filenames as dasiu, dasfu1 and dasfu2. These names refer to the baseline dataset and the year one and year two follow-up datasets respectively. Whilst the different centres that participated in the original study are still collecting data, the data used for this study were collected during the first two years

of WHO Collaborative Study on Determinants of Outcome of Severe Mental Disorders (Sartorius et al., 1986).

The WHO Collaborative Study on Determinants of Outcome of Severe Mental Disorders (Sartorius et al., 1986) was one of three[1] large scale international multi-centre studies which were developed as a consequence of the WHO International Pilot Study of Schizophrenia (IPSS) (Sartorius, Shapiro, Kimura & Barrett, 1972; World Health Organisation (WHO), 1973). The IPSS (Sartorius et al., 1972; WHO, 1973) was the first large scale international multi-centre study of the epidemiology of schizophrenia and was coordinated by the WHO. It has come to be considered as one of the pre-eminent studies in this area (Crow, 1995). The IPSS (Sartorius et al., 1972; WHO 1973) involved 811 schizophrenic patients at 11 centres within nine countries across the globe. All participants were recruited using the same sampling frame and assessed using instruments from a predefined battery (albeit in different languages). The battery comprised a core of instruments used at all sites and a small set of additional instruments used only at some selected sites.

[1]WHO Collaborative Study on Impairments and Disabilities associated with Schizophrenic Disorders (Jablensky, Schwarz & Tomov, 1980; Schizophrenia: an International follow-up study (WHO, 1979); WHO Collaborative Study on Determinants of outcome of severe mental disorders (Sartorius et al., 1986)

The IPPS (Sartorius et al., 1972; WHO 1973) demonstrated three clear effects. Firstly, schizophrenia is a universal disorder that has the same symptomatic profile across the world. Secondly, the prevalence for the disorder was constant across all centres. Thirdly, there was a clear difference between the centres in terms of outcome, with patients in developing countries having significantly better outcomes in terms of remission and relapse rates (WHO, 1979).

From this third effect it was clear that in order to develop the epidemiological perspective of schizophrenia, indices other than signs and symptoms would need to be explored. Up to this time (1978-9) little work had been done that examined schizophrenia in terms of its consequences such as impairments, disabilities and handicaps. Following the IPSS (Sartorius et al., 1972; WHO 1973), the WHO Division of Mental Health embarked on two further prospective studies that focused on the course and outcome of schizophrenia (Jablensky, 1988). The first study was the WHO Collaborative Study on Impairments and Disabilities associated with Schizophrenic Disorders (Jablensky, Schwarz & Tomov, 1980).

The data used in this thesis comes from the second of these studies, the WHO Collaborative Study on Determinants of Outcome of Severe Mental Disorders (Sartorius et al., 1986). For this large study, cohorts of consecutive new patients (treated-incidence cohorts) from predetermined geographical catchments were created at a number of centres. The prima facie basis for inclusion in the original study was that the individual was

suffering from a first life-time episode of a non-affective functional psychosis.

Patients in the original WHO Collaborative Study on Determinants of Outcome of Severe Mental Disorders (Sartorius et al., 1986) have now been followed up for nearly 20 years. Data from these patients have led to many publications (see for example Cooper et al., 1987; Hambrecht, Maurer, Hafner, Sartorius, 1992; Volavka et al., 1997).

The result of these studies has been that an important body of evidence has been established which has changed perceptions about the course and outcome in schizophrenia. A number of findings clearly demonstrated the importance of outcome as part of the overall picture of schizophrenia. For example, the idea that schizophrenia is a disorder characterised by a chronic course and poor outcome is no longer the accepted view. The manifest differences in outcome that had been suggested in the IPPS (Sartorius et al., 1972; WHO, 1973) was confirmed, with around 30% of all first episode patients having a good symptomatic outcome:- that is, no diagnostically schizophrenic symptomatology at two year follow-up . Also, despite the relapse rate increasing over time, the cumulative psychosis free period was either stable or increased - that is, those patients who remained symptomatic and had relapses, had on average, progressively shorter periods of psychosis. However, patients were as disabled at five years post index as at two (Jablensky et al., 1992). This would suggest that

symptomatic indices and disability do not co-vary; and that whilst symptoms (or schizophrenia onset) may precede disability and could be necessary to the production of disability, they are not sufficient to cause it.

This is clearly illustrated in one of the study centres, Nottingham[1]. At the two year follow-up, where nearly 50% of the original patients no longer met the diagnostic criteria for schizophrenia, the majority of these individuals were still moderately to severely disabled (Cooper & Bostock). There were also a smaller proportion (~7%) of the cohort who at the two year follow up were neither diagnosabley schizophrenic or measurably disabled. Analyses of these and other data reinforce the central role disability, rather than the simple presence of signs and symptoms, has in any perspective of schizophrenia (see for example Bellack, Morrison, Wixted & Mueser, 1990; de Jong, Giel, Slooff & Wiersma, 1985; De Jong, van der Lubbe & Wiersma, 1996; Hafner, Nowotny, Loffler, der Heiden & Maurer, 1995; Maneros, Deister, Rohde, Steinmeyer & Junemann, 1989; Moller & von Zerssen, 1995; Wiersma, Nienhuis, Giel & Slooff, 1998; Wiersma, Nienhuis, Slooff & Giel, 1998).

[1]For reasons external to the study design (principally s lack of staff resources) not all of the centres collected disability data via the WHO-DAS (WHO, 1988). Despite this shortcoming, disability data are available from a sufficient number of diverse centres such that the findings could be generalised.

In all there were 12 catchment areas in 10 countries (Colombia, Czechoslovakia, Denmark, India (2), Ireland, Japan, Nigeria, UK (2), USA and USSR.)

The number of patients rated for disability using the WHO-DAS (WHO, 1988) is given in Table 3.1 below;

Table 3.1

Number of Patients Across All Collection Centres that had Disability Data Collected at Each Rating Period

Dataset	dasiu	dasfu1	dasfu2
(year)	(1980)	(1981)	(1982)
N	447	426	665

Note that there is a variation in numbers across the two years of the study. This relates principally to the availability of suitably trained raters at each of the sites, and does not imply that new subjects were recruited post index. This variation in numbers does not have any effect on the validity of the data collected.

Data management

Procedure

The original data was collected in the form of the original WHO-DAS (WHO, 1988) coding sheets. These were checked for completion at each centre before being passed on to the study coordinator at the Mental Health Division of the WHO in Geneva. The data were then manually copied into machine readable format, which at that time created computer punch-cards which were then read onto DAT tapes. The data was delivered to this study in the form of SAS files copied onto 10 floppy disks. As the researcher only had limited access to a copy of SAS (1991) but routine access to SPSS (1996) it was decided to convert the data into SPSS (1996) files. As SPSS (1996) has no facility to read or convert SAS (1990) files automatically, a computer code was developed by the researcher to perform this conversion. See Appendix A for the executable code.

Summary

The three datasets used in this study were collected using the WHO-DAS (WHO, 1988) over the first two years of the WHO Collaborative Study on Determinants of Outcome of Severe Mental Disorders (Sartorius et al., 1986). These three datasets represent data on around 450 individuals with a diagnosis of schizophrenia. These data are an important record of the variance in disability both across the cohort and over time. Analyses of these and other related data have not only changed perceptions about the

disability associated with schizophrenia, but have also been used to develop better models that explain the relationship between disability and other illness variables.

Chapter 4

Measurement Theory applied to Item Selection and Instrument Development

Introduction

This study is largely based on psychometric data analysis. As such, a discussion of basic measurement theory is required to place the analysis in context. This chapter discusses item selection for this revision of the WHO-DAS (WHO, 1988) in the context of measures of validity.

What is Measurement?

Accurate measurement is at the core of quantitative research (Barrett & Caroselli, 1998; McDowell & Newell, 1987) and provides precision and objectivity (Tiernany et al., 1986) so that information can be shared and replicated (Giuffre, 1995). Stevens (1952) describes measurement as the "assignment of numbers to objects or events according to rules". Zeller and Carmines (1979) describe measurement in the behavioural sciences, as the process of "linking abstract concepts to empirical indicants". This definition, which is in keeping with classical testing theory, suggests that in order to investigate measurement, both the theoretical and empirical components need to be considered (Keats, 1967). In this study, the empirical component is addressed in the method chosen for the primary

data analysis, whilst the theoretical component is drawn from both the theory of instrument development (Coates, 1995; Fullerton, 1993; Lynn, 1986) that has been used as a framework for the process and the model of disability that provides the framework for the product, namely the instrument (Cooper, 1994).

The empirical component of measurement, in the behavioural sciences, focuses on some observable indicant such as a specific behaviour. The response to a question or a mark on a self-rating scale are also examples of observable indicants. The theoretical component of measurement focuses on the underlying unobservable entity that is represented by the indicant, and sets out the relationship between the indicant and the hidden entity (Coates, 1995). The indicant can be considered to be the surrogate of the unobserved entity.

An example of the relationship between an indicant and its unobservable entity frequently used in psychiatry, is the concept of thought disorder. The rating of the presence of formal thought disorder in an individual has important clinical consequences, although formal thought disorder is not directly observable. A characteristic pattern of disordered discourse (the observable indicant) is taken as a surrogate for the assumed underlying pattern of disordered thinking (the unobservable entity); and the disordered speech is considered a mirror of the underlying thought disorder (Andreasen, 1979). This means that measurement and, by default, indicators of measurement performance, must focus on the

relationship between empirically grounded indicants and the underlying unobservable concepts.

Validity

As discussed in Chapter 1, the strength of the relationship between the observed indicant and the inferred entity is accepted as the validity of the measure (Lemke & Wiersma, 1976). Nolan and Behi (1995a) place validity at the core of the research process and validity is the concept at the core of this study.

Validity has been defined by Zeller and Carmines (1979) as the extent to which any measuring instrument measures what it is intended to measure. In other words, validity is a function of the fit between the indicant and the entity. The more abstract the concept, the more difficult it is to establish validity of the instrument measuring it (Morley & Pallin, 1995; Polit & Hungler, 1991; Tilden, 1985). Cronbach (1971) clarifies that it is not the measurement per se that is validated, but rather the interpretation of the data in relation to a specific situation or procedure. In other words, it is not the measuring instrument in isolation that is validated, but the measuring instrument in relation to the *purpose* in which it is being used (Perlow, 1992). In this study, it is not the measurement of disability that is being developed, but the *measurement of disability in persons who have met the predetermined operational criteria for a diagnosis of schizophrenia.*

The role of the expert panel is to provide a contextual link between the factor analysis and the subject experience (Alemi, Stephens, Llorens & Orris, 1995; West & Isenberg, 1997). The panel chosen for this study accomplished this by validating the suggested item headings and the stem anchors. In other words, they provided a mechanism through which the results of the statistical analysis at the centre of this study could be verified against a defined body of human expertise.

Unlike reliability, which is purely empirically grounded and directly calculable, validity concerns the relationship between the indicant and the concept (Duli, 1989; Gaberson, 1997; Heermann & Megel, 1994). For example, the reliability of the National Adult Reading Test (NART) (Nelson, 1982; Nelson & Willison, 1991) can be calculated independently of the use to which it is being applied. To determine the validity of the NART it needs to be fixed in a particular conceptual context.

The performance of some measures are particularly easy to validate. For example, the standard 30cm ruler does not require much use to find that the measurements produced fit with axiomatic concepts of the nature of length. Furthermore the measurements also relate, through known laws, with many other variables. The determination of validity in the case of the ruler would be fairly straightforward. However, the quantification of behavioural and other psychosocial processes is less straightforward. The relationship between the indicant and the entity is usually not well developed and is often not amenable to direct examination (Munro, 1991).

Thus, in relation to psychometric instruments, validity is an evidence based determination which is related directly to the function of a particular measurement (Fullerton, 1993; Gaberson, 1997; Heerman & Megel, 1994; Nolan & Behi, 1995a; Nolan & Behi, 1995b). For example, the validity of the NART (Nelson, 1982; Nelson & Willison, 1991), mentioned previously, would be different, depending whether it is used as a measure of current adult literacy or as a measure of premorbid IQ in persons with schizophrenia. It has been used and shown to be valid for both of these purposes (Beardsall & Brayne, 1990; Jones et al., 1993; O'Carroll et al, 1992; Tracy, McGrory, Josiassen & Monaco, 1996). Although the content of the NART (Nelson, 1982; Nelson & Willison, 1991) is the same in both situations, its validity is due to the content embedded in the context. In being used for two different purposes it is as if two completely separate instruments have been used. The validity in the first situation is independent of the validity in the second.

Measures of Validity

Behavioural and psychosocial measures can be grouped according to three main functions, and the validity of any behavioural or psychosocial instrument can be assessed in relation to these three major functions (Deane, 1991). These functions are:

- The establishment of a statistical relationship with some other variable
- The representation of a specific domain

- The measurement of specific behavioural or psychosocial states or traits

The emphasis given to any of these separate but related validities is dependent upon the use to which the instrument is to be put at any particular time. The three functions listed above correspond to the three standard measures of validity; namely criterion-related, content and construct validity respectively (Nunnally, 1978).

Content Validity.

For the purpose of this thesis, only the concept of content validity will be discussed in depth because it is the core of instrument development (Goodwin, 1997; Lynn, 1986; Nolan & Behi, 1995a). Construct and criterion-related validity are determined from the data the instrument collects in use (Deane, 1991), and as such are beyond the scope of this study.

According to Polit and Hungler (1990), content validity addresses the sampling adequacy of a measure and is concerned with the extent to which it covers the dimensions of the concept under consideration. Lynn (1986) cautions that content validity should not be confused with the less vigorously determined face validity which is simply an opinion that an instrument promises to be relevant.

Gable and Wolf (1993) suggest both conceptual and operational definitions need to be considered when assessing content validity. The conceptual definition(s) define the domain of content that is to be measured. They set the boundaries of the concept under consideration and, as unambiguously as possible, set out what falls within the domain as well as that which falls outside. The operational definitions sub-divide this domain space in as specific a manner as possible (forming the instrument items), as well as set the dimensional depth of the space through the use of appropriate scaling for these items. The primary goal of content validation is to have the expert panel (content judges) offer opinions that support or reject the operational definitions, and the chosen scaling system that define a conceptual domain.

In this case, the predictive ability of the factors is not the major focus of the content judgment. The focus is whether the manner of operationalisation of these factors is satisfactory, comprehensive and congruent with the theoretical perspective proposed.

Panel of Experts.

As previously stated, the use of a panel of experts as a reference group can provide the developer of a new or modified instrument with a powerful mechanism for establishing the necessary content validity during the pre-trial developmental phase of an instrument. Selby-Harrington et al. (1994) suggest that expert panel review is the only adequate strategy for developing content validity. This is proposition is supported by a number

of other authors (Berk, 1990; Fullerton, 1993; Haller, 1990; Lynn, 1986; Rew, Stuppy & Becker, 1988; Slocumb & Cole, 1991).

It is not sufficient simply to confer selected individuals with the title of an expert: expertise should be an evidence based determination. Careful consideration must be given to the mechanism of panel selection. Berk (1990) states that whilst the use of expert panels may be quoted in the literature concerning a particular instrument, the characteristics, qualifications and the process of expert selection is rarely quoted. It has been suggested (Grant & Davis, 1997) that such basic information is necessary if the instrument is to be taken seriously. The selection criteria should include significant published research on the construct under consideration (Grant & Kinney, 1992), and additional clinical expertise when an instrument has clinical application (Grant & Davis, 1997). At least one panel member should have expertise in the theoretical or conceptual framework on which the instrument is based (Davis, 1996). Such suggestions fit in well with Lynn's (1986) view that any expert panel should have at least three members. In addition Lynn (1986) calculated that a minimum number of five experts are needed to control for agreement simply by chance alone. If standard error of proportions are developed, then using seven panelists allows one expert to disagree completely with the other six without the content validity (calculated using Lynn's (1986) methodology), falling below the 0.05 significance level.

Inter-rater Agreement.

Walz, Strickland and Lenz (1991) recommend that before estimating content validity, the inter-rater agreement is calculated at each iteration. Put simply, the panel counts the number of items they agree should be included versus those they feel should be discarded, on each occasion there has been a change made to the number of items. Once the exert panel agrees about which items to include, this inter-rater agreement process can stop. For instruments with a large number of items and a relatively small adjudicating panel such as the one in the present study, it is reasonable to calculate item-group agreement (Bartko, 1991). As a general rule, inter-rater agreement should be no lower than 70% (Davis, 1992; Selby-Harrington, et al., 1994).

Methodology of Item Selection

Panel Selection

The objective criteria for expert panel selection discussed in the previous section are clear. However personal information about all possible candidates was not easily accessible. The short list of possible panel members was based, to a large degree, on a review of the related literature and the knowledge of the researcher in consultation with supervisors and other professional colleagues, who suggested a number of potential candidates. The recruitment of mental health nursing clinical experts was problematic because the mental health nursing infrastructure in Western

Australia is not sufficiently developed outside of universities such that any clinical nurse would have met this study's standard of an expert using the criteria specified previously.

A shortlist of all possible candidates was then drawn up (N=10) and these were approached. Seven of them agreed to being involved in the study. The panel members are listed below.

- Professor Assen Jablensky. Head, University Department of Psychiatry and Behavioural Science, University of Western Australia
- Professor Eamon Shanley. Foundation Chair of Mental Health Nursing, School of Nursing, Edith Cowan University, Western Australia
- Dr David Castle. Senior Lecturer, Department of Psychiatry and Behavioural Science, University of Western Australia. Director of Mental Health Services, Fremantle Hospital, Western Australia
- Emeritus Professor John Cooper. University of Nottingham, UK., Foundation Chair in Psychiatry, Nottingham University Medical School. Member: Expert Panel on Mental Health, WHO. Adviser: ICD-8, 9, 10, ICIDH and ICIDH-2
- Dr Farooq Ahmad. Consultant Psychiatrist Graylands Hospital Western Australia
- Dr Aaron Groves. Deputy Chief Psychiatrist for Western Australia, Health Department, Perth, Western Australia
- Associate Professor Aleksander Janca. University Department of Psychiatry and Behavioural Science, University of Western Australia.

Formerly Division of Mental Health, World Health Organisation, Geneva

A more specific summary of their relevant expertise is presented in Appendix B.

Dummy Data

To provide an extra level of external control over the expert mediated content justification process, the use of "catch trials", using dummy data has been suggested as a method to rate the performance of individuals carrying out a task that requires vigilance and attention (Spiker, 1989). A dummy datum is one item that is deliberately included as a "target" for exclusion as part of the catch trial. It is, in psychometric terms, a deliberate target that the alert rater should notice (Olsson, Bengtsson, Heijl & Rootzen, 1997). In practice, this means a dummy datum, for example a stem descriptor, is deliberately included because it is at variance or inconsistent with the anchor descriptor or vice versa. An example used in this study was the item anchor "Underactivity" which in draft one did not match the content, based on the stem descriptors, which was related to *purposeful* activity (as derived from the factor analysis). In this example the anchor heading was the dummy datum. Any of the selected experts, provided they paid sufficient attention to their review, would be expected to spot such an obvious mistake. Four dummy items were included in draft one. A list of the dummy items is given in Appendix C. None of the

experts were aware that dummy items were included. The dummy items were only included for the review of the first draft.

Whilst such testing may seem to be attacking the integrity and expertise of the panel, in reality and in this situation, such a procedure only reinforced the opinion that the experts had applied their full attention to the review. It would be inappropriate to give the results of named panel members, but it is sufficient to state that the panel as a whole performed very well with most members commenting on the suitability/appropriateness of most dummy variables. A blinded performance table is given in Appendix D.

Principal Components Analysis

Description

A principal components analysis (PCA) was carried out on the three datasets (dasiu, dasfu1, dasfu2). The main applications of factor analytic techniques are (a) to reduce the number of variables and (b) to detect structure in the relationship between variables in an instrument (Harman, 1976; Lindeman, Merenda & Gold, 1980; Muliak, 1972). In PCA, variables are reduced through the combination of two or more variables into one factor. If the correlation between two variables (the correlation matrix) are summarised in a scatter-plot, a regression line can then be drawn that represents a best summary of the linear relationship between the variables. This created third variable (factor) captures most of the meaning of the

two separate variables (Stevens, 1986). In this way two variables are reduced to one factor. The combination of two or more correlated variables into one factor is the basic method of principal components analysis.

Sample Size and Reliable Factors

Rather than become involved in the complex modeling required to calculate the exact minimum sample size necessary to produce reliable factors, Gorsuch (1983, p. 332) suggests that "(an) absolute minimum ratio is five individuals per variable, but not less than 100 individuals for any analysis". For this study each dataset contained 61 variables, so the minimum dataset size needed to contain 305 cases. The smallest dataset, dasfu1, has 426 cases, which is approximately 30% greater than this minimum. This would suggest that the datasets have adequate size to support reliable factor extraction.

Factor Extraction

When more than two variables are involved in the PCA they define a multidimensional variable space. After the first factor has been extracted, the original variable space is rotated so that the original x-axis (if thinking in two-dimensions) approximates the first regression line. This type of rotation is called variance maximising (varimax) because the purpose is to maximise the remaining variance available to the next factor whilst minimising the variance associated with the first extracted factor (Stevens,

1986). In this manner consecutive factors are extracted. Furthermore, because each factor explains variance not captured by the preceding factors, the extracted factors are uncorrelated (orthogonal). In instrument construction terms this means that there are no redundant items.

Variance Reduction

As more factors are extracted, progressively less unexplained variance remains. The point at which to stop the factor extraction may not be clear. However, examination of the properties of the original correlation matrix can allow an objective decision to be made (Stevens, 1986). The variance of all variables within the correlation matrix is equal to one (1.0), and therefore the total variance in the matrix is equal to the number of variables. The variances extracted (explained) by each factor are called eigenvalues. These values are expressed as a percent of the total variance such that the sum of the eigenvalues is equal to the number of variables. The widely cited Kaiser (1960) criterion state that only factors with eigenvalues greater than one (unity) should be retained. That is, unless a factor extracts as least as much variance as the equivalent of one original variable, it is not retained. In this way, the first stage of the questionnaire development used PCA to identify significant factors relating to the prediction of disability from the three datasets (dasiu, dasfu1, dasfu2). If the extracted factors (from the datasets dasiu, dasfu1, dasfu2) explain a large proportion of the variance within each dataset and the degree of explained variance is attributable to the factors in a summative manner,

then any measure that uses these factors as measures of disability in a similar population is inherently valid (at least in terms of content). The advantage of the PCA is also that no dependent variable needs to be identified, although in one sense the dependent variable can be thought of the data variance within the dataset (Stevens, 1986). Also, because the factors explaining most of the variance are orthogonal, they will approach Hempel's (1959) ideal classification described as "mutually exclusive and jointly exhaustive".

Summary

Content validity is the most important determination in instrument development prior to use (Deane, 1991; Lynn, 1986). The process of expert review is enhanced if it is transparent in terms of panel selection, and rigorous in terms of the application of predetermined item inclusion coefficients.

The use of well established factor analytic techniques to extract "disability factors" from well developed datasets provides a sound basis from which to start the necessary process of expert review. PCA was used because it allowed the data to be reduced to a series of orthogonal factors without loosing a large amount of the predictive quality of the original data. That is, the PCA reduced the original items to a lesser number of factors whilst preserving the variance inherent in the original dataset.

Data Cleaning and Reduction

Data Cleaning

Prior to any data analysis, the original datasets were put through a stringent series of checking, cleaning and quality assurance procedures.

The two issues that needed to be satisfied were that the data conformed to the scoring regimen of the WHO-DAS (WHO, 1988); i.e. there were no out-of-scale scores, and secondly, the data translation process (from the supplied SAS format into the used SPSS format) did not corrupt data integrity. In other words, procedures were undertaken to ascertain that there was no difference other than format between the original dataset and the converted dataset used in this study.

Following the conversion from SAS to SPSS the datasets were compared and found to be equivalent. Once this was done the datasets were then processed to ensure the scores represented the range of scores possible for each field according to WHO-DAS (WHO, 1988) scaling. A more detailed explanation of these processes is given in Chapter 5 in the discussion of data cleaning and quality control.

Data Reduction

There are two standard methods used to compile factors once the PCA is complete. Whilst each method has both advantages and disadvantages,

the choice of method is largely dictated by the shape of the scree plot of Eigenvalues versus factors (Cattell, 1966).

Kaiser criterion (1960).

The Kaiser criterion (1960) requires all factors with an eigenvalue greater than unity be retained. The advantage of this method is that the maximum explanatory power from the data is extracted. The disadvantage of this method is that when a series of factors comprising only two or three items can explain even a little more than any single item, these low power factors are retained.

Scree method (Cattell, 1966).

The second method, the scree method (Cattell, 1966), includes all those factors with maximum explanatory power but excludes those factors, that whilst having more explanatory power than any single variable, do not contribute significantly to the explained variance. Essentially factors with eigenvalues approaching unity are excluded. This is done through "eye-balling" the data and making a parsimonious decision about where the rate of change of eigenvalues versus factors is adding factors (in this case to the instrument) without substantially increasing the explanatory power of the compiled factors. As consecutive factors account for less and less of the variance, the decision to stop including them is based on the point where little random variability remains.

Choice of Factor Selection Method

The choice of which method of factor selection to apply is not based any standard criteria as none are readily calculable. Despite both methods being studied in detail (Brown, 1968; Cattell & Jaspers, 1967; Hakstian, Roger & Cattell, 1982; Linn, 1968; Tucker, Koopman & Linn, 1969) the choice of method comes down to which makes sense in the circumstances.

However, it is theoretically possible to generate sets of random data based around a particular number of factors. From such datasets it is possible to estimate which of the criteria (Kaiser (1960) versus Cattell (1966)) have the most accurate factor detection rate under particular conditions. This was not undertaken for two reasons. Firstly, such complex mathematical modeling is beyond the scope of this study and secondly, and perhaps more importantly, the extent to which the results from such a process is interpretable has been questioned (Hakstian, Roger & Cattell, 1982).

In this study, in order to keep the instrument succinct, the decision was made not to include an extra eight factors that increased the explained variance by only 4-5%; and so the stricter scree criteria was applied to the data. Appendix E provides details of the explained variances for each of the three datasets, Appendix F details the eigenvalues per factor; and the scree plots from the analysis of all three datasets (dasiu, dasfu1, dasfu2) are given in Appendix G.

Factor loadings.

In addition, the factor loadings for the factors included in the original (WHO-DAS) instrument were also considered in the process of deciding which factors to retain in the revised instrument. Factor loadings are essentially correlations between the respective factors and the variables. Only factors with loadings greater than or equal to 0.8 were included. This ensures that the factors are interpretable in terms of the original factor groupings. Those factors that meet the scree criterion (Cattell, 1966) or have a factor loading ≥ 0.8 are shown in Table 4.1 below.

Table 4.1

PCA Derived Factors with a Loading ≥ 0.8 for Each Dataset.

Dataset			Factor descriptor
dasiu	dasfu1	dasfu2	
X	**X**		Patient's self care last year
X	**X**	**X**	Purposeful activity last year
X	**X**		Slowness
	X	**X**	Social withdrawal
	X	**X**	Participation in household activities
X	X	**X**	Affective relationship with spouse
X	X	**X**	Sexual relation with spouse
X	X	X	Friction in interpersonal relationships outside household
	X	X	Slowness of movement (ward behaviour)
X	**X**		Underactivity (ward behaviour)
	X	X	Conversation (restricted)
X	**X**	X	Social withdrawal (ward behaviour)
X	**X**	X	Leisure interest (ward behaviour)
	X	X	Hospital work (nurses' opinion)
		X	Possess matches (nurses' opinion)
X	***X***	*X*	*Free to make out of hospital visits (nurses' opinion)*
X	***X***		*Out unsupervised with other patients (nurses' opinion)*
	X	*X*	*Work outside hospital (nurses' opinion)*
X	*X*	*X*	*Work outside ward with varying levels of supervision*
X	***X***		*Hours attending OT*
	X		Need for supervision for security reasons

Note. Bold type indicates factor meets scree test (Cattell, 1966). Italics indicate factors subsequently removed from first draft.

In applying the above inclusion criteria, it needs to be considered that the data represents a cohort measured on three occasions over two years. To take this cross sectionality into account it was decided to include only those factors that met the scree criteria at least once or had a factor loading

≥0.8 *across* the three measurement points. This decision was made on the basis of discussion with a biostatistician (Pranom, personal communication, Feb 1998) as there was no literature available regarding time-series PCA. A full table of all the WHO-DAS (WHO, 1988) variables and their explanatory descriptors is given in Appendix H. For copyright reasons it is not possible to reproduce the entire WHO-DAS (WHO, 1988).

Further Data Reduction

From Table 4.1 it can be seen that some of the factors cover the same area of activity but are either separated into hospital versus community or rater versus nurses' report. Where appropriate these were incorporated into one item.

The final data reduction method does not have a statistical basis, but reflects the changing patterns of patient management in the twenty years since the original WHO Collaborative Study on Determinants of Outcome of Severe Mental Disorders (Sartorius et al., 1986) commenced. Five factors were discarded because they are determinations of patient management that are no longer practised, and relate to the management of long-stay patients in a total institution (Weinstein, 1994). Few current patients would be ratable on such items (as originally defined in the WHO-DAS (WHO, 1988)). These factors are in italics in Table 4.1 above.

First Draft

The factors that met the aforementioned criteria were included in the first draft. The development from a factor list to a draft instrument can be considered as the development of an anchor descriptor and then a series of within item descriptive stems that correspond to a scaling regimen.

Scaling System

By examination of the distribution of scores per item in the original datasets it was ascertained there was no apparent conspicuous ceiling or bottoming effect. This conclusion was achieved through producing a frequency polygon of scores for every WHO-DAS (WHO, 1988) item across each dataset (dasiu, dasfu1, dasfu2) and then "eye-balling" the plots. This determination would suggest that the original scaling was dimensionally adequate, although scale selection is rarely an entirely mathematical consideration and is open to review in the expert panel process (Summers, 1993).

It was therefore decided to follow the original five point scaling system from "no problem" to "maximum problem". This scaling is ordinal and stem descriptors were written such that they reflected both the content of the factors and their severity on the scaling framework.

Stem Content

Given the nature of what are essentially qualitative judgments it was decided to enlist the expertise of one of the panel who has a background in the area qualitative research methods (A S). Following this review, some syntactical changes were incorporated into the stem text.

Instrument Layout and Scoring Design

Scoring System

The original intention had been to use a visual analogue scale (VAS) for scoring the items. This type of scale is familiar to most clinicians (McCormack, Horne & Sheather, 1988) and also provides a reference of all possible scores and had been used in the original WHO-DAS (WHO, 1988) in a modified form. However it was decided not to use a VAS because it was felt that a VAS would erroneously imply, by its linear layout, that the data scored were interval. This is not the case, so rating boxes were used instead. Two boxes were included for each item so that "present state" disability as well as "lifetime before" disability could be rated at the same time and on the same schedule. This multiple period rating is a feature of a number of WHO instruments such as the Schedules for Clinical Assessment in Neuropsychiatry (SCAN) (WHO, 1994).

Other

In addition, no anchor descriptors other than the item heading (factor description) were included in the first draft. The sum of the stem descriptors for any item was considered sufficient information for the expert panel. Further, it was felt that this "looser" structure would reduce "guiding" responses/feedback. It was planned to write more extensive anchor descriptors once the revisions had been returned from the panel. A copy of the first draft is in Appendix I.

Summary

The first draft was constructed as a series of items headed by an anchor descriptor (the item heading), followed by a series of five stem descriptors written to reflect both the scaling regime and the overall anchor description. Scoring was accomplished using scoring boxes adjacent to each item heading (anchor).

Results of the First Draft after Review

In general the first draft was favourably received, with most comments concerning only minor changes. None of the reviewers suggested the exclusion or inclusion of any items, so the aggregated percentage item agreement in this study was 100 %. The suggested improvements could be grouped under three main headings. The first was there was a general

consensus that the items required, in addition to the stem descriptors, an anchor description. This was felt necessary for a number of reasons including ease of use in the interview situation as well as allowing the rater to score responses that did not fall within the limited number of stem descriptors or examples.

Secondly, it was suggested that consideration should be given to changing the overall scaling from the "no problem" to "maximum problem" format to a system that allowed items to rated as "present but not problematic". Those experts that had considerable clinical experience felt that the first draft scaling implied that the presence of a specified behaviour was axiomatically problematic, which they felt was unjustified. The scaling was modified to reflect this for the second draft. The third series of suggested changes concerned the stem descriptors. All of the experts offered clinical examples of disabled patients who would not have been ratable, at least not at a level reflecting their estimation of the patient's disability, on the instrument. This situation was corrected by the addition of one or two extra clauses to the stem descriptors to cover the particular examples. A great deal of consideration was given to this in the discussions with the experts in the development of the second draft. Finally, five of the experts commented that the proposal to rate multiple periods on the same scale was unnecessarily confusing, and that if a rater wished to rate two or more separate periods he/she could simply use additional scoring sheets.

Modifications and Further Developments

All of the proposed changes and additions were discussed with both supervisors and condensed into a lesser number of alterations. Where any of the suggested changes were contrary to the advice of one of the experts, or where a proposed change was a compromise between two or more suggestions, those experts were contacted on an individual basis for further input.

The series of revisions to the first draft were listed and checked for compatibility with the table of suggested changes. Again this was done in consultation with both supervisors. Through this process of alteration, revision and rigorous discussion, the first draft was transformed by stages into the second draft. A copy of the second draft is in Appendix J. The entire basis of this process was to increase the validity of the scale through a process of comprehensive review. Whilst the PCA was used to generate factors from the original data, it was the process of translating these into specific examples that relied on the expertise of the reviewers. It is the quality of this translation that preserved the validity of the original data.

Development of Anchor Descriptors

From the first review the expert panel was unanimous that, in addition to the stem descriptors, an overall item anchor descriptor would have been a useful addition. It had been decided, a priori, that anchor descriptors

would be developed for the second draft. The rationale behind the decision to present the first draft without anchor descriptors is covered previously (p. 75).

The process of developing anchor descriptors was very much guided and informed by one of the panel (JC) who has particular expertise in this area (advising for example on the development and revision of various WHO classifications including the ICD-8 to 10 and the ICIDH).

JC's suggested process was to develop a number of clinical scenarios based on real experiences, and relate them to a particular anchor title or group of stem descriptors within an item. From the specific description that incorporated, for example, a statement that described poor self care, a more general anchor description was developed. This syntax was then "tested" against a second example, adjusted if required and then "re-tested" against the first example; again to ensure agreement was maintained. This process of revision and revaluation was continued until the examination of new scenario examples did not lead to any further changes to the developed anchor descriptor. This method was applied for each item and specific anchor descriptors were developed. In addition, colleagues were used informally as a resource to provide further clinical examples that could be tested against the derived anchor descriptor. In this way anchor descriptors were developed that matched the agreed item content as specified by the stem descriptors.

With the incorporation of this modified scaling frame, a number of the original stem descriptors needed to be reclassified and some new descriptors (principally present but not problematic) had to be generated.

The instrument was recompiled with the retained/modified item headings, newly developed anchor descriptors, modified scaling frame and retained/revised stem descriptors.

Coefficient of Validity Index

Lynn (1986) has suggested that where there is general agreement about content, the expert panel can be used to produce data so that a "coefficient of validity index" (CVI) can be calculated. In order to gather these data, two counterbalanced VAS were added to each item as well as the instrument as a whole. Counter-balancing was used to reduce "rating contamination" were the rating on one scale influences the direction of rating on the second (Pollatsek & Well, 1995). Rather than have a separate CVI sheet, the VAS were incorporated into the second draft. An example is given in Appendix N.

The first VAS was for the experts to record their scores of the item content. The second VAS was to score the item scaling. "Content" as described in the covering letter, referred to the adequacy of the sum of stem descriptors for any item to exemplify all possible stem descriptors that may included

under the anchor descriptor. "Scaling", in contrast, refers to whether the stem descriptors have been adequately assigned on the scaling frame.

The overall instrument CVI allows the researcher to examine whether each expert has been consistent in his/her ratings. It would be expected that the mean of the total item scores would approach the total instrument score. The revised draft with covering letter and instructions was redistributed to the expert panel.

Unfortunately one panel member (DC) had at this stage left his university position and was unavailable for this second review. A substitute was approached and agreed to join the panel. The specific details are given below and full details are given in Appendix B.

- Dr. Peter Wynn Owen, Consultant Psychiatrist, Director of Postgraduate Training, Royal Australian and New Zealand College of Psychiatrists, Perth, Western Australia.

All of the changes discussed above were collated and incorporated into the second draft.

Second Draft

A copy of the second draft is given in Appendix J. The draft submitted for the second expert review was, on the basis of the first review and its

consequent processes, at a much higher level of development than the first draft. Items were fully operationalised, the scaling system standardised and the formatting refined.

In addition to similar instructions to the first review, each panel member was asked to "score" the content validity in a manner that would allow a measure of agreement or consensus (the Coefficient of Validity Index (CVI) (Lynn, 1986)) to be determined.

Because of the consultative process that was undertaken between the return of the first draft and the re-submission of the second draft, it was envisaged that only minor changes would have been suggested and this turned out to have been the case. Of course, for changes that were made on the basis of a compromise between differing judgments, it was expected that no expert would be one hundred per cent satisfied with the balance. What was required was a reasonable level of agreement between all experts and across all items. The predetermined mean coefficient of agreement (mean CVI) was 0.83, both for items and for the instrument as a whole. This figure was determined through the application of Lynn's (1986) methodology to determine content validity.

The item by item mean CVI values are given in Appendix K along with the individual de-identified expert scores. A graphical representation (range-plot), described below, of these data are given in Appendix L. Although individual ratings for some items fell below the predetermined 0.83 level,

it is the mean values that determine the CVI. None of the mean CVI totals fell below the 0.83 threshold. Given this level of content validity for the second draft, there was a question about what to do with the experts' suggested changes. It was decided that where the suggestions were minor, or involved the resolution of some small inconsistency which would not adversely affect the CVI, these would be incorporated into the final draft. This contingency covered all of the additional suggestions.

Development of the CVI Procedure

The CVI procedure was modelled on Lynn's (1986) suggested methodology. A strategy that scored the content and scaling adequacy of each item independently was developed by the researcher. In addition, the expert panel members were asked to give an overall score for both the content and the scaling of the total instrument. In practical terms this meant that panel members were supplied with a copy of the second draft that had two counterbalanced visual-analogue scales (VAS) on which to score each of the items, as well as give a total score. Scoring using VAS is common because VAS's are of particular value in scoring opinions (Lee & Kieckhefer, 1989). The content and scaling scales were counterbalanced, which is a strategy designed to reduce any directional bias of scales that are in close proximity (Pollatsek & Well, 1995). A copy of the second draft with CVI scales can be seen in Appendix M.

Item Scores and Group Procedure

In addition to Lynn's (1986) criterion that the instrument as a whole should achieve a mean CVI of 0.83, it was decided, given that this was a revision based on well developed original data, that this 0.83 coefficient should be applied to each individual item. Only items that attained a CVI ≥ 0.83 would be considered sufficiently well developed and that further modification would be unnecessary. Only items that failed the 0.83 criterion were to be developed further before resubmission to the expert panel. All the items from the second draft attained the 0.83 criterion and so there was no requirement that the instrument be developed further in terms of content validity. A copy of the final instrument can be seen in Appendix N.

Final Paper Scale

Following the satisfactory outcome of the second review in terms of content validity, only a few minor changes were required. As these were only grammatical or layout in nature it was decided that a further submission was unnecessary, and the scale was complete and ready for pilot testing. A copy of this final version is given in Appendix N.

The work involved in the pilot testing phase of the instrument development is beyond the terms of reference of this study and will be the subject of further research by this researcher.

Summary

On the basis of the first expert review no items were suggested for inclusion or deletion. However, a number of changes were made to the instrument based on the feedback from the panel. These included changing the capture period to allow a rater defined interval; changing the scaling frame to include a "present but not problematic" rating, as well as simplifying the problematic rating to a mild-moderate-severe format. In addition, changes were made to simplify the accompanying instructions and provide a more aesthetically pleasing layout.

Finally the instrument was resubmitted to the expert panel for a second review with the aim of quantifying the content validity. The second draft met the predetermined standard for acceptable content validity.

Chapter 5

Reliability Issues

Introduction

This chapter describes the process by which the integrity of the original datasets were maintained during the statistical analyses described in Chapter 4.

Prior to analysing any large dataset containing raw data, it is good practice to institute a series of quality control measures to check the integrity of the data (Benneyan, 1998; Knatterud et al., 1998). Principally the data should conform to the scoring scheme of any instrument that has been used to collect the original data. In addition, as the datasets have been converted from one file format (SAS (1991)) to another (SPSS (1996)), it made sense to ensure no transcription errors had been introduced into the new dataset as a result of the conversion process alone. With the three supplied datasets (dasiu, dasfu1, dasfu2) this quality control procedure was accomplished in two stages.

Stage 1: Determining the Accuracy of the Data Translation

The data received from the WHO had already been converted from the IBM mainframe format in which it is stored, to the SAS (1991) format in which it was delivered. It is not possible to check the accuracy of this

translation, but given that the WHO has developed enormous expertise in handling such datasets and presumably has quality control procedures in place, the standard of this original translation has to be assumed to be high (see for example Pinol, Bergel, Chaisiri, Diaz & Gandeh, 1998). In the first stage, the SPSS (1996) and SAS (1991) datasets were compared to ensure equivalence. There are many data checking and integrity approaches described in the literature (Benneyan, 1998; Issel, Eggers, Hofman, Kruger & Scheinpflug, 1981; Petersen et al., 1996; Schwartz, Weiss & Buchanan, 1985). Most involve either systems to monitor and check for errors as data are entered, or provision of internal logic checks on evolving datasets (Alsop & Langley, 1998; Benneyan, 1998; Chernoff, Wang, Andersen & Felson, 1995; Petersen et al., 1996).

Probabilistic Error Checking

There are also probabilistic methods of checking for errors where the function is not primarily to check for errors, but to link common data groupings across different databases based on an estimation that the records are homogenous (Wagner & Newcombe, 1970). An example would be a probability estimation that an individual on one data base is the same individual on a second database based on matching a number of fields (such as age, sex, and postcode) that are common to both databases. The more fields that match, the greater the probability that the two records are for the same individual. Where it is not possible to create exact concordance, an agreement threshold has to be set above which the match

is accepted and below which it is not. The use of such probabilistic data linkage techniques is commonly used to link large case registers (Jaro, 1995; Roos & Wajda, 1991; Waien, 1997).

Whilst none of the techniques described above on its own, provides a model that can be applied to the WHO datasets, together they can be modified to create a workable data checking system. Each of the WHO datasets are relatively large, each comprising between approximately 27,000 and 40,000 fields. To check for complete data duplication would have required the fields to be manually scanned and compared. This would have been very time consuming and require a level of concentration that would be difficult to sustain for the extended periods necessary. In addition, if the researcher could have had access to SAS (1991) for such an extended period, there would have been no necessity to convert the datasets to a SPSS (1996) format in the first place! Another alternative would have been to break the datasets (SPSS (1996) and SAS (1991)) down, and use a simple comparative technique to check for errors. Whilst a comparative technique was developed for this task, it was not applied since it was considered that the process of database deconstruction and reconstruction was unnecessarily complex and as likely to create errors as detect them. In addition, any database reassembly errors would be undetectable, coming as they would after the data had been checked.

Error detection strategy.

An error detection strategy was developed by the researcher that when used was fast, did not interfere with the data structure of the datasets and did not require further data processing after the technique had been applied. The method was a modification of the probabilistic techniques used to link different datasets (such as those described above) and based around the binomial distribution.

The binomial distribution is useful for describing binomial events, such as the number of male and females in a specific dataset. In this study the binomial event is agreement or error per compared field.

This method is probabilistic rather than absolute. It requires the researcher to set an acceptable maximum error rate and then calculate the number of fields that need to be checked for agreement between the two datasets being compared to test that this maximum error rate is not exceeded. In addition, by basing the analysis on a binomial distribution as well as defining the acceptable error rate (the test proportion), the binomial allows a probability that an error has been missed to be estimated. In other words, once the researcher fixes the maximum acceptable error rate, it is possible to calculate the minimum number of fields that need to be checked. If no errors are detected, it is possible to state that the likelihood of there being undetected errors greater than the fixed maximum error rate is below a certain level of probability. Once the

technique was developed and modelled by the researcher, it was reviewed by a researcher experienced in working with large (genomic) dataset who found it to be methodologically acceptable (Hallmayer, personal communication, May 1998).

The maximum transcription error rate was set at 0.2 percent. On average this would mean that 64 fields could be mis-transcribed and would go undetected. The technique relied on randomly sampling a predetermined number of fields in each SAS (1991) dataset and comparing the data in each identified field with the comparable field in the equivalent SPSS (1996) dataset. The number of fields that needed to be sampled to check for a maximum 0.2 percent error rate is given by the standard formula for the binomial distribution:

$$f(x) = [(n!/x!\ (n-x)!)].\ p^x.q^{n-x}$$

where

p is the probability that the event will occur

q is equal to 1-p

n is the maximum number of trials

This was calculated using the SPSS Binomial Test (1996). Put simply, an iterative series of binomials were calculated with different maximums (number of trials) until the 0.05 significance level was exceeded. Only the results for the computation exceeding the 0.05 significance level is given in Table 5.1 below:

Table 5.1

Binomial test of the Probability of Finding an Error between Datasets if 1850 Random Fields are Sampled and the Maximum Error Rate is Fixed at 0.2%

	Category	N	Observed proportion	Test proportion	Significance
Group 1	1.00	1850	1	0.998	0.048
Total		1850	1		

This means that if the error rate is greater than 0.2 %, which is calculated to be more than 64 errors on average per dataset, the binomial sampling strategy will detect at least one of these errors (p=0.05). In other words, sampling 17 % of each dataset will detect on average at least one error if the error rate exceeds 0.2 %.

Procedure for dataset comparison.

A series of 1850 pseudo-random numbers were generated using the Stastica (1994) random number generator. The upper limit for the random number generation was taken as the product of the number of subjects and the number of variables.

These 1850 numbers identified individual locations within the data matrix based simply on a right to left row-wise count. The data in each of the

1850 identified fields were extracted from the three SPSS (1996) datasets (dasiu, dasfu1, dasfu2) as a string in a text delimited file. In addition, the same random number sequences were used in the same manner to extract the equivalent number sequences from the three supplied SAS (1991) datasets. For ease, both the SPSS (1996) and SAS (1991) random number generated files were imported into SPSS (1996) and checked for equivalence. The separate SAS (1991) and SPSS(1996) were read into separate columns. All that was necessary to check for equivalence between common fields across the datasets was the subtraction of the SAS (1991) field score from the counterpart SPSS (1996) field score. Any transcription errors would be revealed by any non-zero results.

This procedure was carried out on all three datasets (dasiu, dasfu1, dasfu2). Had any errors been detected, then it was planned that the entire translation process and SPSS (1996) import code (Appendix A) be reviewed prior to repeating the translation process and retesting. No errors between the two datasets (SAS (1991) and SPSS (1996) were found, suggesting that the data translation had not introduced any errors (at least above the 0.2 % threshold). It is worth reporting that post analysis this checking procedure was repeated on these datasets (for reasons external to this study) with the maximum error rate set to 0.01 %. This required the checking of 3700 fields for equivalence. Again no errors were found.

Stage 2: Congruence of Data and WHO-DAS (1988) Scoring Scheme

As the data comprised the raw scores copied directly from the original data collection sheets from the various field centres, a degree of error could be expected. The second stage of the data quality control was to ensure that the data were congruent with the scoring regimen of the WHO-DAS. This meant that the scores in any particular field were within the range of scores possible for the item. This second stage of data quality control was accomplished by creating a SPSS query that specified the scores possible per field as set out in the instructions for the WHO-DAS (1988). This query was then applied to each dataset (dasiu, dasfu1, dasfu2) so that out-of-range data were identified and copied to a separate error table that also identified their location within the original dataset from which they were derived. The per dataset error rates were calculated by comparing the errors detected to the size of the original data matrices (rows x columns). These out-of-range data were deleted from the original SPSS (1996) datasets (dasiu, dasfu1, dasfu2) because deletion is considered the best method of dealing with such errors (Offord & Boyle, 1986).

There are a number of statistical techniques available for estimating missing values (Chernoff, Wang, Anderson & Felson, 1995). The specific method applied to these data was included in the description of the data analysis in Chapter 4. The only other method would have been to substitute a zero score. However, this was not the best option because it would have had the effect of attributing a positive rating of absence to a

particular item and could have been a confound in the later data analysis (in terms of decreasing the sensitivity and specificity of the extracted data). The number of individual scores found to be outside the scaling range and the error rate per dataset is given in Table 5.2 below.

Table 5.2.

Scoring Scheme Errors and Error Rate per Dataset

	Dataset		
	dasiu	dasfu1	dasfu2
Errors	126	47	138
Error rate (%)	0.004	0.0015	0.003

Other Data Management Issues

The scale used to rate severity in the WHO-DAS (1988) is ordinal, but there are nominal ratings available to code an item as "not applicable" ("9") and for some items as "impossible to rate" ("8"). Because a PCA gives weight to numbers these ratings had to be removed. Not applicable ratings ("9") were deleted without substituting any value.

However, those items coded as impossible to rate ("8") were recoded as zero, meaning absence. This was done for three reasons. Firstly if a null value was entered it would have implied the item was not applicable or an errorful score had been made. Neither of these situations reflect this rating. Secondly, a conservative approach would suggest a rating of absence in order to minimise false positive scores. Finally, with the very

low number of "8" and "9" ratings it appeared to be unlikely that the differing approaches for each rating had any effect on the analysis. This was in fact the case, as three further exploratory PCA's were performed on all three datasets (dasiu, dasfu1, dasfu2) (where eights and nines were coded zero; eights were coded null and nines coded zero; and where the eights and nines were set as null) without any difference in the factors extracted. Such multiple analyses producing comparable results under different data substitution rules suggests that the original PCA provided reliable (and stable) factors.

Coversheet

For the first draft, the content, rather than the layout, of the coversheet was considered important. The coversheet was designed to carry the minimum amount of information necessary to make the rating meaningful in the clinical or research setting. This content can be broken down into two types of information. The first can be described as "convenience" information whilst the second can be called "validity bound" information.

"Convenience" Information

"Convenience" information is essentially information that any user may wish to record in order to attach the content of the rating to a particular individual in a particular study or clinical environment. This may include subject name, study ID number, date of birth, sex, date of rating etc. This

information can be customised according to the needs of the user. Whether or not this "convenience" information is recorded has no bearing on the psychometric properties of the instrument per se.

"Validity Bound" Information.

In contrast, the "validity bound" information is directly connected to the validity of the instrument. "Validity bound" information includes the theoretical overview of the instrument. Thus a brief description of the model of disability must be included so that the user understands what is being measured. A review of the scaling framework that has been adopted is also required. The other "validity bound" information includes the instructions for use. Whilst such instructions may be seen as a reliability issue, it needs to be born in mind that validity and reliability are not separate entities. It is the reliability that determines the upper limit of validity.

The most important issue to address is the context of the instrument's application. The primary context-based consideration is the data acquisition strategy; with the time period covered by the instrument as a secondary issue. The ordinate source of data when this instrument is used is the patient with schizophrenia, and these data are acquired through retrospective recall. Given this, consideration needs to be given to issues related to recall in schizophrenia as they pertain to the development of this revision of the WHO-DAS (1988).

Memory and Recall in Schizophrenia

Schizophrenics as a group generally perform poorly on most memory related tasks compared to healthy controls (Gold et al., 1994; Saykin, Gur & Gur, 1991). These deficits however are not uniform. Procedural memory, for example, can remain relatively intact (Clare, McKenna, Mortimer & Baddeley, 1993). The memory deficits in chronic schizophrenia have been likened by McKenna (1990) to the loss suffered by persons with frank brain injury. Such a position may be overstating the case, as others (Schroder, Tittel, Stocket & Karr, 1996) have found the memory deficits found in schizophrenia to be relatively uniform across both course or symptom severity. However, it is generally accepted that schizophrenics have poorer recall than controls (Gold, Randolph, Carpenter, Goldberg & Weinberger, 1992; Stip, 1996).

Furthermore, the memory deficit is differentially increased as a function of the recall period (Calev, Berlin & Lerer, 1987; Rushe, Woodruff, Murray & Morris, 1999; Sengel & Lovallo, 1983). Therefore, the longer the time between the event and its recall, the greater the memory deficit between schizophrenics and healthy controls. This is most probably related to neo-cortical organicity (Feinsten, Goldberg, Nowlin & Weinberger, 1998; Mitrushina, Abara & Blumenfeld, 1996; Paulsen et al., 1995) and generally localised to the prefrontal regions (Goldberg, Weinberger, Pliskin, Berman & Podd, 1989; Harvey, Powchik, Mohs & Davidson, 1995).

This situation is complicated further as it seems that whilst the decline in long term memory (greater than two or three weeks) is fairly linear in controls, the decline in schizophrenics follows a negatively skewed u-curve (Feinsten, Goldberg, Nowlin & Weinberger, 1998). This means that the difference in recall between controls and schizophrenics does not increase linearly, but rather in an exaggerated curvi-linear manner. In other words, not only does the recall deficit increase as a function of the recall period, the percentage increase in this deficit also increases according to the same function. Asking schizophrenics to recall events that have occurred in the more remote past is highly problematic, with the magnitude of the problem increasing the further back the recall event.

In addition, schizophrenics have been shown to have a pervasive temporal order memory deficit that is characteristic of a prefrontal brain disorder (Schwartz, Deutsch, Cohen, Warden & Deutsch, 1991). This means that in addition to poorer recall than controls, the temporal sequencing of the recalled information is also impaired. Schizophrenics, in effect, can remember events but not when they occurred (Rizzo, Danion, van der Linden & Grange, 1996). Control subjects show reasonable sequencing for those events they can recall; i.e. that A occurred before B but after C, for example. Such ability to sequence memories can be used to fix recall around specific events in the past (Fuhrman & Wyler, 1988).

Enhancing Recall

Cues have been shown to enhance information retrieval for both schizophrenics and normal controls (Sengel & Lavallo, 1983). In fact, the proportional increase in recall is the same across the two groups (Feinsten, Goldberg, Nowlin & Weinberger, 1998). The performance of schizophrenics remains below that of the controls, but this seems to be a baseline effect, indicating that the schizophrenics are simply starting from a poorer position (Feinstein, Goldberg, Nowlin, Weinberger, 1998). The source of the cue also seems to be important. There is evidence that schizophrenics as a group under-report illness variables when completing self-report measures (Gorman, 1993), and over-report the same illness variables when highly structured interview formats developed for administration by trained lay interviewers are used (Booth, Kirchner, Hamilton, Harrell & Smith, 1998; Sandanger et al., 1999).

In order to develop further strategies to minimise the memory problems outlined above, the literature on life-event research was explored. The life-event literature was chosen because it is almost entirely based on retrospective self report. Despite the size of the literature and the many papers that address recall issue (see for example Betz & Skowronski, 1997; Chaikelson, Arbuckle, Lapidus & Gold, 1994; Goddard, Dritschel & Burton, 1996; Harvey, Bryant & Dang, 1998; Hyman & Loftus, 1998; Lteif & Mavissakalian, 1996; Monroe, 1982; Parks & Balon, 1995; Wilhelm, McNally, Baer & Florin, 1997) relatively little has been written concerning

the issue of recall in patients with schizophrenia. What has been written only confirms the problems addressed in the previous paragraphs. Monroe (1982) found that schizophrenics under-reported life-events, and that this effect was particularly significant for recall periods greater than four months prior to interview. Neugebauer (1983) has questioned the validity of retrospective studies involving schizophrenics that rely entirely on recall, where the recall period is greater than the previous six months. The use of relatives and significant others as supplementary sources of information about events in the patients' past, only increases the "information yield" by around 20%. This lead Dixon and King (1995) to recommend that whilst there are problems with recall in patients with schizophrenia, the patient, rather than significant others, remains the best source of information about their illness, at least in regard to retrospective recall.

The clinical interview based around a rating schedule has been shown to be the most reliable strategy for gathering retrospective data from patients with schizophrenia (Cooper, Peters & Andrews, 1998). In addition, requiring the patient to give plausible concrete examples when he/she describes possible ratings of presence, significantly reduces the false positive rate (Spitzer, 1983).

For the reasons outlined above, it was decided to recommend that the instrument only be administered by trained clinicians, with the schedule forming the basis for a semi structured interview (using the stem examples

as cues) and that specific exemplars should be sought from the patient before making a rating of presence for any item. Furthermore, it was decided that the above direction should be included in the instrument instructions.

Triangulation of sources has been shown to increase the data yield particularly where the point of reference is in the more remote past (Cowman, 1993). The use of multiple sources of information to describe or quantify the same object or event is referred to as triangulation (Fahie, 1994; Nolan & Behi, 1995c). The goal of triangulation is to provide confirmatory or missing information (Begeley, 1996). It is also a common strategy employed in nursing and psychiatric research when the object under investigation is complex and dynamic (Dootson, 1995; Mitchell, 1986; Shih, 1998; Sohier, 1988).

Whilst the "memory" literature suggests the use of the patient as the primary information source and significant others as supplementary sources of information, adopting the broader triangulation methodology suggests accessing all possible sources of information.

For these reasons, whilst the patient is defined as the primary source of information, the instruction to use (unless otherwise specified) all possible sources of information was included in the instrument instructions.

Rating Period

Related to the above discussion on rating sources is the issue of rating period or periods. The characteristics of the schizophrenic memory deficit must be taken into account when establishing the rating period. The ability to rate different time periods is a function of a number of psychiatric research instruments. This may be related to the importance of course and severity in many psychiatric disorders. The four common rating periods used in a number of different psychiatric clinical/research instruments are Present State (PS) which refers to the previous four weeks; Representative Episode (RE) which refers to some fixed index period in the past; Lifetime Ever (LE) which is self-explanatory; and Lifetime Before (LB) which refers to the entire period prior to the PS.

There are several disadvantages in including representative episode, LB and LE in this revision of the WHO-DAS (1988).

Whilst a number of other WHO psychiatric instruments utilise the RE rating period, this has been rejected as an option for this revision of the WHO-DAS. The problems schizophrenics have with temporal sequencing, coupled with recalling the remote past, would suggest that centering recall around some index period in the past would be highly unreliable.

However, LB has been incorporated into this revision. Course and severity are important issues in schizophrenia research and so a lifetime rating has

been incorporated. Whilst it is recognised that recall diminishes across time, sequencing problems are not encountered when making a lifetime rating. In addition, because there is a definite tendency to forget events rather than confabulate them, the size of the Type I error is enhanced at the expense of the Type II. As LB ratings are predominantly used for research purposes, such a situation is tolerable. In addition, whilst recall problems are encountered in making LB ratings, the quality (if not the quantity) of the secondary sources (such as casenotes etc.) should remain relatively intact.

Lifetime ever ratings are not specifically recorded but can be calculated by aggregating the maximum scores from the PS and LB ratings.

Present state ratings.

The fixed four week rating period has been chosen for the rating of PS disability in this instrument. This has a number of advantages. The four week capture period, as mentioned previously, is congruent with many WHO and other rating scales that attempt to rate present state. It is generally accepted that "present state" means at most the last four weeks. Secondly, by using the same rating period as other instruments, when a researcher is using multiple rating instruments with the same subject, confusion related to changing reference periods can be avoided. In addition, a standard rating period allows cross sectional comparisons to be made across instruments. In other words, if the researcher describes a

subject as being both impaired and disabled based on present state, it is more useful if both ratings share a common time frame. Finally, it has been shown that whilst memory declines as a function of time, the period of maximum recall is around the previous two weeks (Jenkins, Hurst & Rose, 1979). The fact that the rating period is four weeks does not in itself diminish the two week recall rate, and also means that the recall rate will be relatively similar across all other instruments.

Changes to Capture Period

The Present State (PS) rating period was felt to a useful time-frame by all of the experts. Whilst the fixed PS rating period was retained, four of the experts suggested that the Lifetime Before (LB) rating was too inflexible. One of the raters suggested that there is a particular problem that could be encountered if using LB ratings: It is possible to accumulate a high total score simply by rating highly on the different items at different times in the past. This in effect can load the scale in the direction of those respondents who experience a range of illness consequences, but not necessarily concurrently. For example, an individual who had repeated or even continuous episodes of behaviours that only score on one item would have an LB disability significantly lower than an individual who has had a few single episodes of behaviours that each score on different items. However, despite the higher LB score, it would be difficult to argue that the second person is significantly more disabled than the first.

Despite this limitation, the evidence presented in Chapter 5 in favour of the LB rating remains convincing. In order to accommodate both conclusions derived from the literature and the expert advice, it was decided to include in the second draft the additional option of allowing the rater to define a discrete rating period, with the instruction that this should only be used as a special case where PS or LB would be insufficient or inappropriate.

Changes to Scaling

A major change incorporated into the second draft was a revision of the general scaling frame. This was based on two concerns that the experts expressed about the original 5-point "problem absent" to "severe problem" scale.

The first concerned the "mild-moderate-marked-severe" breakdown. All of the experts with considerable clinical experience (N=5) expressed doubt about whether they could make such a differentiation in practice. Four of the panel noted a particular difficulty differentiating markedly disabled from severely disabled and felt that the differences in the stem examples were minor. In addition, three of the panel suggested rating items as present along a three point (mild-moderate-severe) presence scale would provide more useful information because most clinical staff would find a three-point rating more easily interpreted. Secondly, two of the panel suggested that including four present categories suggests a degree of

discrimination that they doubted could be achieved even with the provision of stem descriptors. It was also noted by three of the experts that presence alone does not necessarily indicate a problem. Such a view is in keeping with the literature (as reviewed and presented in Chapter 2). It was decided to alter the scaling frame to incorporate and reflect this advice.

Not wishing to make such a major change without collateral psychometric evidence, this researcher reviewed the relevant literature with the aim of finding an existing and suitably tested scale that would accommodate the concerns of the expert panel. The scaling frame for a number of functional instruments was examined to see if any incorporated a "present but not problematic" rating. Such a scale would also have the advantage of being able to give a positive rating of absence.

Whilst there are many examples of different scaling frames (see for example: Derogatis, 1994; Honigfeld, Gillis & Klett, 1965; Lehman, 1988; Rhoades & Overall, 1988; Rosen, Hadzi-Pavlovic & Parker, 1989; Smith, Burnam, Burns, Cleary & Rost, 1995; Tessler & Gamache, 1995; Ware & Sherbourne, 1992) the system developed for the Health of the Nation Scales (HoNOS) (Wing, Curtis & Beevor, 1996) was taken as the exemplar. The HoNOS (1996) was initially developed for the UK Health of the Nation survey (1994) as a broad measure of psychiatric morbidity. It has subsequently found wide use as a standard measure of outcome particularly in the UK and Australia.

A modified version of the HoNOS scaling frame has been adopted for the revised draft. This modification retains the 5-point scale but allows both "absence" and "present but not problematic" to be rated separately. The rating of present and problematic has been simplified to the suggested "mild", "moderate" or "severe" ratings.

Revision of Coversheet

Information on the coversheet about the model of disability on which the instrument was developed was considered unnecessary by the majority (N=6) of experts, and was deleted.

In addition, those experts that commented on the coversheet (N=6), commented the assignment of rating period was considered too inflexible and the accompanying instructions (about how to assign the rating period(s)) were difficult to follow. These comments were acted on.

Rather than use the complex system which allowed the rating of two separate periods on the same scoring sheet, it was decided to adopt one scoring sheet and one rating period. This also meant that the two rating boxes per item were replaced with a single rating box. In addition, the fixed rating periods (PS, LB) incorporated into the first draft were supplemented, as discussed previously, by a period of time set by the rater, and recorded on the score sheet under "from" and "to". The instructions about use were also simplified and analysed with the Microsoft Word for

Windows 95 Flesch-Kincaid Grade level score (Microsoft Corporation, 1996). The current instructions are easily understandable by a person of average education (Year 12). As this level of education is the minimum standard for university entrance, this is congruent with the instrument's expected clinician use.

Finally, as a result of general comments about the layout and general aesthetics of the coversheet, a number of changes were made. Whilst these final changes were made entirely based on personal preference, the comments and suggestions of colleagues who use various instruments as part of their daily practice were sought (See Appendix N).

Chapter 6

Discussion, Conclusion and Recommendations

The two main outcomes arising from this study are firstly a methodologically rigorous revision of the WHO-DAS (1988) that is ready to be field tested. Secondly, and perhaps more importantly, has been the development of a methodology that can be used as a "template", allowing instrument development to proceed from the analysis of appropriate data through to validation against an expert panel. The methodology developed for this study relies on the incorporation of techniques from three sources. Firstly, existing but separate instrument development strategies were amalgamated. Secondly, methods used in other areas of behavioural research, such as catch trials, were adapted to improve the methodological rigour. Finally, some statistical techniques not previously associated with instrument design have been used to provide measures of data integrity. A flow chart of the instrument development process created during this study is given in Appendix O.

Datasets

Original datasets sets.

The single most important advantage in using the WHO datasets was their size. As discussed in Chapter 4, the individual datasets were all of

sufficient size that lack of statistical power was not an issue when interpreting the reliability of the factors extracted. Small sample size leading to lack of statistical power has been cited as a major problem with much of the psychiatric literature (Bartko, 1988).

The major disadvantage in obtaining data from a large organisation such as the WHO is that the size of the organisation necessitates the use of standardised data storage formats. The data used in this study could only be supplied in an IBM mainframe or SAS format. This would not have been a problem where the recipient had access to these same programs. However, in this study this was not the case, and a procedure needed to be developed to transfer the data into SPSS (which was available to the researcher). Such extra processing always introduces the possibility of manufacturing data errors. In this study the authenticity of the data translation was determined using a modification of standard probabilistic data linkage techniques.

Another characteristic of the datasets was that the number of returns varied across the two years of the study. This was reported to be due to the original study design that relied on local clinicians who collected the data in addition to their normal clinical work. This variation in numbers over the three data capture periods did not effect either the choice or outcome of the statistical analyses. This was because each dataset was analysed independently and only the *results* of these separate analyses pooled.

WHO studies.

In addition to creating large datasets, the WHO Collaborative Study on the Determinants of Outcome of Severe Mental Disorders (Sartorius et al., 1988) study had a number of other advantages that enhanced the interpretability of the data collected.

The study was prospective. Prospective studies have many advantages. A particular advantage in psychiatric research is that such studies do not rely on the rating of notes or other "second-hand" evidence, but can obtain data from primary sources. This is particularly important when the issues associated with recall covered in Chapter 5 are considered.

The study was multicentre. This has the advantage of reducing researcher bias, as there are a large number of researchers involved at a range of geographically separate centres. It also has the advantage of reducing selection bias. It is generally acknowledged that relying on patients from a single service will introduce service specific selection effects (Robinson, Woerner, Pollack & Lerner, 1996; Strauss, 1973; Thomas et al., 1997; Thornley & Adams, 1998; Treno, Gruenewald, & Johnson, 1998). This is because few single services cater for the full range of referrals and presentations. Services, either through planning or operation, tend to be orientated towards a particular patient population, such as dual diagnosis patients or patients with specific behavioural problems. Developing a sampling frame that includes a range of centres can reduce such a selection

bias. In addition to being geographically diverse, the original sample is culturally diverse as well. Again, one criticism levelled at many studies particularly in the area of mental health is lack of cultural diversity in the study population which limits the generalisability of the results (Lewis-Fernandez & Kleinman, 1995). Cultural heterogeneity was built into the original sampling frame. This means that the results of this study can be applied to a broader patient population.

The WHO Determinants of Severe Outcome study was a two year multicentre prospective study. For this period the WHO coordinated the study, aggregated the data and produced the final report. Since that time each of the original participating centres has continued to follow up its original cohorts. Unfortunately these data have not been centrally aggregated. If these data had been available for this analysis it may have provided more robust evidence of enduring disability variance for the factor analysis.

Models of Disability

The ICIDH has been criticised for being taxonomically inconsistent (Brandsma, Lakerveld-Heyl, van Ravensberg & Heerkens, 1995; Dickso, 1994; Colvez & Robine, 1985; Schuntermann, 1996; Ustun et al, 1995; Wiersma, D). In defence of the ICIDH, the original authors never intended it to be the definitive conceptual framework for illness consequences (Badley, 1993; Wood, 1989). It was the intention that the ICIDH, as the

first framework of illness consequences, should provoke debate that would lead to more refined models. In this regard the ICIDH has been very successful. It was through examining the ICIDH, and to some extent the Groningen model of Social Disability, that Cooper (1994) formulated his model that has been used in this study.

Panel of Experts

The judgement of what constitutes an "expert" is a complex task made more difficult by the lack of any established operational definition. It is ironic that it is often experts whom are needed to judge who is an expert, and applying the external criteria as suggested by numerous authors and used in this study still requires the arbiter to make a judgement call. This judgement call purports to be objective but in reality is not measurable. The surrogates used to judge expertise, such as longevity of practice, number of publications etc. have the appearance of being reasonable, but these concepts have never been tested.

There is the further irony that a study of the content validity of a new instrument is unable to make any statement about the content validity of the expert selection. All that can be said from the literature in support of the selection of experts is that these criteria have face validity. It would be a very interesting extension of this study to develop an instrument for identifying content experts.

Finally, the expert panel used in this study was based on a convenience sample. This has both advantages and disadvantages.

It may be trite to suggest that the advantage of a convenience sample is that it is convenient, but this is a very important consideration. Convenience, in this study, means the experts were local, accessible and the indicants of their expertise were more readily available. Also, their being known to the researcher could have had a positive influence on their motivation to participate in the study, as well as on their efforts to produce a well considered contribution. This is an important consideration.

In addition, it is difficult to imagine how one could design a method to randomly select a panel of experts from a potential world-wide pool and ensure those selected both agree to participate and are willing to devote their time completing what is in general a rather mundane task.

In summary, the criteria and selection of experts is based upon the considered judgement of the researcher that that the sum of the advantages of the process outweigh the disadvantages, particularly in terms of the pragmatic aspects of required methodology.

Lynn's Model

This research has provided further evidence of the value of Lynn's model for the quantitative determination of content validity. Lynn's model

provides clearly defined statistical criteria for deciding when an instrument has achieved acceptable content validity. It allows the researcher to pre-determine the number of experts needed in the panel and the level of agreement necessary for an item to be judged acceptable or otherwise. It also provides a methodology for calculating the level of agreement required between the panel members for an item to be accepted based on the size of the panel. However, Lynn's methodology is limited in that if there are less than seven experts it only takes one member to disagree about an item for an item to fall below the threshold for inclusion. In this case, an item would have been excluded if it fell below the minimum inclusion index. Therefore, the minimum of seven panel members is required so that one panel member can completely disagree with the others and an item still achieve an acceptable CVI. This will allow for some disagreement to occur among the panel without leading to automatic exclusion of an item. It is therefore recommended that the minimum number of expert judges forming the judgement panel should be set at seven.

Dummy Data

This study is unique from the perspective of an instrument design incorporating the use of dummy data to provide an external measure of the experts' performance during the first draft review. There is good evidence from the psychological testing literature that attests to the value of using such "catch trails" and this has been presented in Chapter 4. The dummy

variables were not intended as a measure of expertise. The judgement of expertise was predetermined against the criteria set out in the literature. The catch trial was a measure of the experts' task performance, in this case identifying inconstancies, errors and omissions. All of the introduced errors were considered by the researcher to be moderately obvious. Since the purpose was testing vigilance rather than expertise, and because such trials form a normal part of the psychological testing armamartum, there was no ethical reason not to include such measures.

It is worth noting that all of the panel were informed at the end of the study about the use of catch trials and none expressed any consternation. Most in fact expressed mild amusement!

There are two issues that must be commented in regard to the catch trials. They were only used on a single occasion because it was assumed that the score on this occasion would be indicative of the individuals' task performance across both reviews. This may not have been the case, but the deleterious effect on the CVI prevented the inclusion of further catch trials during the second review. Furthermore, as one of the experts had left the study and a substitute was found there is no data on the replacement's performance.

The second and more important consideration for this study was what would have happened if one or more of the experts failed to identify the dummy variables. To exclude these individuals would make attaining the

predetermined CVI, discussed above, all the more difficult. Therefore, for future works of this kind, it is recommended that, if possible, the researcher should over-sample and create an expert panel in excess of the minimum necessary. These potential content experts could then be tested using dummy variables embedded in a previously validated instrument and those that "fail" this catch trial could be excluded.

Principal Components Analysis

Principal components analysis has a number of advantages that make it suitable for this study. It allows the researcher to determine factors without the need to specify a dependent variable. This allows the variance of the dataset to be "factored" without the need to anchor each regression line to a single variable. The factors are both orthogonal and hierarchal. The factors explain progressively less of the remaining variance. In addition, as the factors are orthogonal, there is no redundancy when these factors are converted into the instrument items. This means that the minimum number of items needed to encompass the concept need to be included in the instrument.

PCA is a well established statistical technique which provides a transparent account of the analysis in terms of factor and variable loadings. Also, as each factor is extracted, there is little "variance leakage". This means that the factor extracts a discrete amount of variance from the variance pool,

and most of this extracted variance is included in the factor (Stevens, 1986).

The major disadvantage of PCA, like most factor analytic techniques, is the assumption that the relationship between the variables can be described by a series of linear equations. It has been demonstrated that the relationship between such complex variables such as those used for this study are rarely, if ever, linear. Despite this, the use of nests of linear equations to approximate the complex relationship between such variables is a well established technique. For the purposes of this study, where the extracted variables were further validated against a panel of experts, this technique was appropriate. In addition, when the results of this analysis was presented to the panel, no items were added or removed. This is further evidence of the soundness of the PCA as a useful technique for the revision of existing instruments.

The determination of the appropriate cut-off point for factors to be included in the instrument relies on choosing, in the first instance, between two established criteria. There is no truly objective procedure that can be used to make this choice. When this issue is addressed in the literature, the advice is in the form of choosing what makes sense in the circumstances. Again, this relies on the researcher making a qualitative judgement about which criteria to apply in a particular situation. In this study and for the reasons set out in Chapter 4, Catell's criterion (1966) has been used in preference to the Kaiser criterion (1960). Ironically, the

application of Catell's criterion (1966) to the data requires the researcher to make further essentially subjective judgements.

It is a constant theme of this study that despite using a sophisticated and rigorous methodology, the final decision about many of the analyses was based on the judgement of the researcher. However, it is considered that by embedding the analyses in the framework that has been developed for this study, these decision are educated, informed and conservative.

Validity Threats

There were a number of validity threats that needed to be overcome when completing the multistage data processing to ensure that the processing itself did not alter the data content.

The probabilistic data sampling strategy developed to ensure the fidelity of the data translations from SAS to SPSS was appropriate. It gave the researcher confidence when making statements about the analyses of the transformed datasets that would otherwise be impossible to validate. It is recommended that probabilistic data checking be used for any study where large datasets are transformed across formats.

Secondly, using a scoring frame filter to verify that the data conformed to the original scoring frame is also highly recommended, again because of the confidence it built into the datasets in this research.

For administrative purposes, the WHO datasets included a number of nominal ratings. With the low number of these ratings that were included in the original datasets, the various substitution strategies used had no effect on the analysis. However, in cases were the results are different and dependent on the strategies employed expert statistical advice should be sought before proceeding.

Summary

The two main outcomes f this study are the development of a a timely revision of the WHO-DAS and a rigorous methodology to guide development and the process for instrument design.

Recommendations Arising from this Study and Areas for Future Research

Recommendations

- The minimum number of expert judges forming the content-judgement panel should be seven.

- If possible, the researcher should over-sample and create an expert panel pool in excess of the minimum seven necessary. These potential content experts should be tested using "dummy variables" embedded in a previously validated instrument and those that "fail" this catch trial should be excluded from participation in the final panel.

- The probabilistic data checking methodology developed for this study should be used for any situation where large datasets are transformed between formats.

- If data have been collected using a quantitative instrument, a scoring frame filter (query) should be used to verify that data conforms to the original scoring frame.

- Where a number of nominal ratings are present in higher order (ie above nominal) datasets the choice of nominal value substitution strategy (recoded as either null or zero) may have an effect on the analysis. In such situations, all possible permutations of substitute values should be used and the data analysed. Where there is little variance of outcome then it is reasonable to assume that the data substitution has had little effect on the final analysis. However, in cases were the results are clearly different expert statistical advice should be sought before proceeding.

- Choice of expert panel needs to be a transparent process, where a documented audit trail provides a rationale for criteria applied in the selection of content "experts"

Future Research

The next major step in the development of this revised instrument is field testing. This should be designed so that the remaining psychometric properties, criterion-related and construct validity and reliability, can be established.

There is no standard instrument that can be used to identify whether "potential candidates" meet the criteria of "expert". The development of such an instrument would be an interesting extension of this study.

References

Alemi, F., Stephens, R. C., Llorens, S. & Orris, B. (1995). A review of factors affecting treatment outcomes: Expected Treatment Outcome Scale. American Journal of Drug and Alcohol Abuse, 21(4), 483-509.

Alsop, J. C. & Langley, J. D. (1998). Determining first admissions in a hospital discharge file via record linkage. Methods of Information in Medicine, 37(1), 32-7.

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

American Psychiatric Association (1994). Diagnostic and statistical manual of mental disorders (4th ed.). Washington, DC: Author.

Andreasen, N. C. (1979). Though, language, and communication disorders. I. Clinical assessment, definition of terms, and evaluation of their reliability. Archives of General Psychiatry, 36(12), 1315-21.

Andrews, G. (1991). The cost of schizophrenia revisited. Schizophrenia Bulletin, 17(3), 389-394.

Andrews, G., Hall, W., Goldstein, G., Lapsley, H., Bartels, R. & Silove, D. (1985). The economic costs of schizophrenia. Implications for public policy. Archives of General Psychiatry, 42(6), 537-543.

Armitage, P. (1995). Changes in long-term mental health nursing. Nursing Standard, 9(38), 28-30.

Ashcraft, M. L., Fries, B. E., Nerez, D. R., Falcon, S. P., Srivastava, S. V., Lee, C. Z., Berki, S. E. & Errera, P. (1989). A psychiatric patient classification system: An alternative to diagnostic-related groups. Medical Care, 27(5), 543-557.

Badley, E. M. (1987). The ICIDH: format, application in different settings, the distinction between disability and handicap. International Disability Studies, 9(3), 122-125.

Badley, E. M. (1993). An introduction to the concepts and classifications of the international classification of impairments, disabilities, and handicaps. Disability and Rehabilitation, 15(4), 161-178.

Barker, P. (1991). Mental health. Finding common ground. Nursing Times, 87(2), 37-8.

Barker, S., Barron, N., McFarland, B. H., Bigelow, D. A. & Carnahan, T. (1994). A community ability scale for chronically mentally ill consumers: Part II. Applications. Community Mental Health Journal, 30(5), 459-72.

Barnett, D. B. (1991). Assessment of quality of life. American Journal of Cardiology, 67(12), 41-44.

Barrett, E. A. & Caroselli, C. (1988). Methodological ponderings related to the power as knowing in change tool. Nursing Science Quarterly 11(1), 17-22.

Bartko, J. J. (1991). Measurement and reliability: statistical thinking considerations. Schizophrenia Bulletin, 17(3), 483-9.

Betz, A. L. & Skowronski, J. J. (1997). Self-events and other-events: temporal dating and event memory. Memory and Cognition, 25(5), 701-714.

Biehl, H., Maurer, K., Schubart, C., Krumm, B. & Jung, E. (1986). Prediction of outcome and utilization of medical services in a prospective study of first onset schizophrenics. Results of a prospective 5-year follow-up study. European Archives of Psychiatry and Neurological Sciences, 236(3), 139-47.

Booth, B. M., Kirchner, J. E., Hamilton, G., Harrell, R. & Smith, G. R. (1998). Diagnosing depression in the medically ill: validity of a lay-administered structured diagnostic interview. Journal of Psychiatric Research, 32(6), 353-60.

Boyd, M. A. & Luetje, V. (1992). The individual who is severely and persistently mentally ill: directions for research and practice. Issues in Mental Health Nursing, 13(3), 207-218.

Brandsma, J. W., Lakerveld-Heyl, K., van Ravensberg, C. D. & Heerkens, Y. F. (1995). Reflection on the definition of impairment and disability as defined by the World Health Organisation. Disability and Rehabilitation, 17(3/4), 119-127.

Brekke, J. S. (1992). An examination of the relationships among three outcome scales in schizophrenia. Journal of Nervous and Mental Disease, 180(3), 162-7.

Brooker, C., Tarrier, N., Barrowclough, C., Butterworth, A. & Goldberg, D. (1992). Training community psychiatric nurses for psychosocial intervention. Report of a pilot study. British Journal of Psychiatry, 160, 836-44.

Brooker, C., Falloon, I., Butterworth, A., Goldberg, D., Graham-Hole, V. & Hillier, V. (1994). The outcome of training community psychiatric nurses to deliver psychosocial intervention. British Journal of Psychiatry, 165(2), 222-30.

Browne, M. W. (1968). A comparison of factor analytic techniques. Psychometrika, 33, 267-334.

Bullinger, M., Anderson, R., Cella, D. & Aaronson, N. (1993). Developing and evaluating cross-cultural instruments from minimum requirements to optimal models. Quality of Life Research, 2(6), 451-459.

Calev, A., Berlin, H. & Lerer, B. (1987). Remote and recent memory in long-hospitalized chronic schizophrenics. Biological Psychiatry, 22(1), 79-85.

Carpenter, W. T. & Buchanan, R. W. (1994). Schizophrenia. New England Journal of Medicine, 330(10), 681-690.

Cattell, R. B. (1966). The scree test for the number of factors. Multivariate Behavioural Research, 1, 245-276.

Cattell, R. B. & Jaspers, J. A. (1967). A general plasmode for factor analytic exercises and research. Multivariate Behavioural Research Monographs.

Chaikelson, J. S., Arbuckle, T. Y., Lapidues, S. & Gold, D. P. (1994). Measurement of lifetime alcohol consumption. Journal of Studies of Alcohol, 55(2), 133-140.

Chernoff, M. C., Wang, M., Anderson, J. J. & Felson, D. T. (1995). Problems and suggested solutions in creating an archive of clinical trials data to permit later meta-analysis: an example of methotrexate trials in rheumatoid arthritis. Controlled Clinical Trials, 16(5), 342-55.

Clare, L., McKenna, P. J., Mortimer, A. M. & Baddeley, A. D. (1993). Memory in schizophrenia: what is impaired and what is preserved? Neuropsychologia, 31(11), 1225-41.

Coates, V. E. (1995). Measuring constructs accurately: a prerequisite to theory testing. Journal of Psychiatric and Mental Health Nursing, 2(5), 287-93.

Coleman M. & Gillberg, C. (1997). A biological approach to the schizophrenia spectrum disorders. Journal of Neuropsychiatry and Clinical Neuroscience, 9(4), 601-605.

Colvez, A. & Robine, J. M. (1986). Problems encountered in using the concets of impairments, disability, and handicap in a health assessment survey of the elderly in Upper Normandy. International Rehabilitation Medicine, 8(1), 18-22.

Conway, A. S., Melzer, D. & Hale, A. S. (1994). The outcome of targeting community mental health services: evidence from the West Lambeth schizophrenia cohort. British Medical Journal, 308(6929), 627-30.

Cooper, J. (1994). Measurement of disability. Proceedings of the Australian Institute of Health and Welfare. Canberra, ACT, Australia. NH&MRC Social Psychiatry Unit.

Cooper, J. E. & Bostock, J. (1988). Relationship between schizophrenia, social disability, symptoms and diagnosis. In A. S. Henderson & G. D. Burrows (Eds.). Handbook of social psychiatry (pp. 317-330). Amsterdam, The Netherlands: Elsevier.

Cooper, J. E., Goodhead, D., Craig, T., Harris, M., Howat, J. & Korer, J. (1987). The incidence of schizophrenia in Nottingham. British Journal of Psychiatry, 151, 619-26.

Cooper, L., Peters, L. & Andrews, G. (1998). Validity of the Composite International Diagnostic Interview (CIDI) psychosis module in a psychiatric setting. Journal of Psychiatric Research, 32(6), 361-8.

Cowman, S. (1993). Triangulation: a means of reconciliation in nursing research. Journal of Advanced Nursing, 18(5), 788-92.

Creed, F., Tomenseon, B, Anthony, P. & Trammer, M. (1997). Predicting length of stay in psychiatry. Psychological Medicine, 27(4), 961-966.

Cronbach, L. J;, (1971). Test validation. In R. L. Thorndike (Ed.), Educational measurement (2nd ed.). Washington, DC: American Council on Education.

Crow, T. J., MacMillan, J. F., Johnson, A. L. & Johnstone, E. C. (1986). A randomised controlled trial of prophylactic neuroleptic treatment. British Journal of Psychiatry, 148, 120-127.

Dada, M., White, W. D., Stokes, H. H. & Kurzeja, P. (1992). Prospective payment for psychiatric services. Journal of Health Politics, Policy and Law, 17(3), 483-508.

Davidson, L. & McGlashen, T. H. (1997). The varied outcomes of schizophrenia. Canadian Journal of Psychiatry, 42(1), 34-43.

Davies, L. M. & Drummond, M. F. (1994). Economics and schizophrenia: the real cost. British Journal of Psychiatry, 25, 18-21.

Davis, A. E. (1996). Instrument development: getting started. Journal of Neuroscience Nursing, 28(3), 204-7.

Davis, G. E., Lowell, W. E. & Davis, G. L. (1993). A neural network that predicts psychiatric length of stay. MD Computing, 10(2), 87-92.

Deane, D. M. (1991). Content, construct, and criterion-related validity. Diabetes Education, 17(5), 361-2.

de Jong, A., Giel, R., Slooff, C. J. & Wiersma, D. (1985). Social disability and outcome in schizophrenic patients. British Journal of Psychiatry, 147, 631-6.

deJong, A., Giel, R., Slooff, C. J. & Wiersma, D. (1986). Relationship between symptomatology and social disability. Empirical evidence from a follow-up study of schizophrenic patients. Journal of Social Psychiatry and Psychiatric Epidemiology, 21(4), 200-5.

Derogatis, L. R. (1994). Symptom Checklist-90-R (SCL-90-R) administration, scoring and procedures manual (3rd ed.). Minneapolis: National Computer Systems.

de Vellis, R. F. (1991). Scale development: Theory and applications. Newbury Park, CA: Sage.

Deyo, R. A. & Patrick, D. L. (1989). Barriers to the use of health status measures in clinical investigation, patient care, and policy research. Medical Care, 27(3), 254-268.

Dickson, H. G. (1996). Problems with the ICIDH definition of impairment. Disability and Rehabilitation, 18(1), 52-54.

Dixon, M. J. & King, S. (1995). The concordance between symptom information gathered from remitted schizophrenic patients and their relatives. Journal of Psychiatric Research, 29(6), 447-56.

Dootson, S. (1995). An in-depth study of triangulation. Journal of Advanced Nursing, 22(1), 183-7.

Ducharme, F., Stevens, B. & Rowat, K. (1994). Social support: conceptual and methodological issues for research in mental health nursing. Issues in Mental Health Nursing, 15(4), 373-92.

Duli, L. S. (1989). Measurement issues. Orthopaedic Nursing, 8(4), 56-7, 68.

Evers, S. M. & Ament, A. J. (1995). Costs of schizophrenia in The Netherlands. Schizophrenia Bulletin, 21(1), 141-153.

Fahie, V. P. (1994). Triangulation: measurement issues and solution exploration. ABNF Journal, 5(5), 138-41.

Feinstein, A., Goldberg, T. E., Nowlin, B. & Weinberger, D. R. (1998). Types and characteristics of remote memory impairment in schizophrenia. Schizophrenia Research, 30(2), 155-63.

Fenton, W. S. & McGlashen, T. H. (1991). Natural history of schizophrenia subtypes. II. Positive and negative symptoms and long term course. Archives of General Psychiatry 48(11), 978-986.

Frances, A. J. & Egger, H. L. (1999). Wither psychiatric diagnosis. Australian and New Zealand Journal of Psychiatry, 33(2), 161-165.

Fuhrman, R. W. & Wyler, R. S. (1988). Event memory: temporal-order judgements of personal life experiences. Journal of Personal and Social Psychology, 54(3), 365-384.

Fullerton, J. T. (1993). Evaluation of research studies. Part IV: Validity and reliability- concepts and application. Journal of Nurse Midwifery, 38(2), 121-5.

Fyer, A. J., Katon, W., Hollifield, M., Rassnick, H., Mannuzza, S., Chapman, T. & Ballenger, J. C. (1996). The DSM-IV panic disorder field trial: panic attack frequency and functional disability. Anxiety, 2(4), 157-166.

Gaberson, K. B. (1997). Measurement reliability and validity. AORN Journal, 66(6), 1092-4.

Gable, R. K. & Wolf, J. W. (1993). Instrument development in the affective domain: Measuring attitudes and values in corporate and school settings. Boston: Kluwer Academic.

Giuffre, M. (1995). Reading research critically: assessing the validity and reliability of research instrumentation-Part 1. Journal of Post Anesthetic Nursing, 10(1), 33-7.

Glazer, W., Prusoff, B., John, K. & Williams, D. (1981). Depression and social adjustment among chronic schizophrenic outpatients. Journal of Nervous and Mental Disease, 169(11), 712-7.

Glick, M., Marsh, A. & Zigler, E. (1980). Inter-relationship among hospitalization measures of psychiatric outcome. Journal of Mental and Nervous Disease, 168(12), 741-744.

Goddard, L., Dritchel, B. & Burton, A. Role of autobiographical memory in social problem solving and depression. Journal of Abnormal Psychology, 105(4), 609-616.

Goldney, R. D., Fisher, L. J. & Walmsley, S. H. (1998). The Health of the Nation Outcome Scales in psychiatric hospitalisation: a multicentre study examining outcome and prediction of length of stay. Australian and New Zealand Journal of Psychiatry, 32(2), 199-205.

Gold, J. M., Randolph, C., Carpenter, C. J., Goldberg, T. E. & Weinberger, D. R. (1992). Forms of memory failure in schizophrenia. Journal of Abnormal Psychology, 101(3), 487-94.

Gold, J. M., Hermann, B. P., Randolph, C., Wyler, A. R., Goldberg, T. E. & Weinberger, D. R. (1994). Schizophrenia and temporal lobe epilepsy. A neuropsychological analysis. Archives of General Psychiatry, 51(4), 265-72.

Goldberg, T. E., Weinberger, D. R., Pliskin, N. H., Berman, K. F. & Podd, M. H. (1989). Recall memory deficit in schizophrenia. A possible manifestation of prefrontal dysfunction. Schizophrenia Research, 2(3), 251-7.

Goodwin, L. D. (1997). Changing conceptions of measurement. Journal of Nursing Education, 36(3), 102-7.

Gorsuch, R. L. (1983). Factor analysis. Hillsdale, New Jersey: Lawrence Erlbaum.

Gorman, D. M. (1993). A review of studies comparing checklist and interview methods of data collection in life event research. Behavioural Medicine, 19(2), 66-73.

Gournay, K. (1995). Mental health nurses working purposefully with people with serious and enduring mental illness - an international perspective. International Journal of Nursing Studies, 32(4), 341-352.

Gournay, K. (1996a). Schizophrenia: a review of the contemporary literature and implications for mental health nursing theory, practice and education. Journal of Psychiatric and Mental Health Nursing, 3(1), 7-12.

Gournay, K. (1996b). Setting clinical standards for care in schizophrenia. Nursing Times, 92(7), 36-7.

Grant, J. S. & Kinney, M. R. (1992). Using the Delphi technique to examine the content validity of nursing diagnoses. Nursing Diagnostics, 3(1), 12-22.

Grant, J. S., Kinney, M. R. & Davis, L. L. (1993). Using conceptual frameworks or models to guide nursing research. Journal of Neuroscience Nursing, 25(1), 52-6.

Grant, J. S. & Davis, L. L. (1997). Selection and use of content experts for instrument development. Research in Nursing and Health, 20(3), 269-74.

Greer, D. S. (1987). Quality of life measurement in the clinical realm. Journal of Chronic Disease, 40(6), 629-630.

Hafner, H., Nowotny, B., Loffler, W., an der Heiden, W. & Maurer, K. (1995). When and how does schizophrenia produce social deficits? European Archives of Psychiatry and Clinical Neuroscience, 246(1), 17-28.

Hakinstan, A. R., Rogers, W. D. & Cattell, R. B. (1982). The behaviour of numbers of factors rules with simulated data. Multivariate Behavioural Research, 17, 193-219.

Hall, W., Andrews, G. & Goldstein G. (1985). The costs of schizophrenia. Australian and New Zealand Journal of Psychiatry, 19(1), 3-5.

Haller, K. B. (1990). Research instruments: assessing validity. MCN American Journal of Maternal and Child Nursing, 15(3), 214.

Hambrecht, M., Maurer, K. & Hafner, H. (1992). Gender differences in schizophrenia in three cultures. Results of the WHO collaborative study on psychiatric disability. Social Psychiatry and Psychiatric Epidemiology , 27(3), 117-21.

Harding, C. M., Zubin, J. & Strauss, J. S. (1987). Chronicity in schizophrenia: fact, partial fact, or artifact? Hospital and Community Psychiatry, 38(5), 477-486.

Harrison, G., Croudace, T., Mason, P., Glazebrook, C. & Medley, I. (1996). Predicting the long-term outcome of schizophrenia. Psychological Medicine, 26(4), 697-705.

Harman, H., (1976). Modern factor analysis (3rd ed.) (pp. 133-173). University of Chicago Press: Chicago.

Harvey, A. G., Bryant, R. A. & Dang, S. T. (1998). Autobiographical memory in acute stress disorder. Journal of Consulting and Clinical Psychology, 66(3), 500-506.

Harvey, P. D., Powchik, P., Mohs, R. C. & Davidson, M. (1995). Memory functions in geriatric chronic schizophrenic patients: a neuropsychological study. Journal of Neuropsychiatry and Clinical Neuroscience, 7(2), 207-12.

Hempel, C. G. (1959). Introduction to problems in taxonomy. In J. Zubin (Ed.), Field studies in mental disorders. New York: Grune and Stratton.

Henderson, S., Jorm, A. F., MacKinnon, A., Christensen, H., Scott, L. R., Korten, A. E. & Doyle, C. (1993). The prevalence of depressive disorders and the distribution of depressive symptoms in later life: a survey using draft ICD-10 and DSM-III-R. Psychological Medicine, 23(3), 719-729.

Heermann, J. A. & Megel, M. E. (1994). Reliability and validity. Plastic Surgery Nursing, 14(3), 183-4, 189.

Holmes, C. A. Health care and the quality of life: a review. Journal of Advanced Nursing, 14(10), 833-839.

Honigfeld, G., Gillis, R. D. & Klett, C. J. (1966). NOSIE-30: a treatment-sensitive ward behaviour scale. Psychological Reports, 19(1), 180-182.

Hubschmid, T., Perrin, N. & Ciompi, L. (1991). The dynamics of outcome predictors in the early phase of schizophrenic psychoses-an empirical study. Psychiatrische Praxis, 18(6), 196-201.

Hyman, I. E. & Loftus, E. F. (1998). Errors in autobiographical memory. Clinical Psychology Review, 18(8), 933-947.

Jenkins, C. D., Hurst, M. W. & Rose, R. M. (1979). Life changes. Do people really remember? Archives of General Psychiatry, 36(4), 379-84.

Johnstone, E. C., Owens, D. G., Gold, A., Crow, T. J. & MacMillan, J. F. (1981). Institutionalization and the defects of schizophrenia. British Journal of Psychiatry, 139, 195-203.

Johnstone, E. C., Owens, D. G., Gold, A., Crow, T. J. & Macmillan, J. F. (1984). Schizophrenic patients discharged from hospital-a follow-up study. British Journal of Psychiatry, 145, 586-90.

Jones, P. B., Bebbington, P., Foerster, A., Lewis, S. W., Murray, R. M., Russell, A., Sham, P. C., Toone, B. K. & Wilkins, S. (1993). Premorbid social underachievement in schizophrenia. Results from the Camberwell Collaborative Psychosis Study. British Journal of Psychiatry, 162, 65-71.

Kaiser, H. F. (1960). The application of electronic computers to factor analysis. Educational and Psychological Measurement, 20, 141-151

Keats, J. A. (1967). Test theory. Annual Review of Psychology, 18, 217-38.

Kent, S., Fogarty, M. & Yellowlees, P. (1995). A review of studies of heavy users of psychiatric services. Psychiatric Services, 46(12), 1247-1253.

King, M. & Nazareth, I. (1996). Community care of patients with schizophrenia: the role of the primary health care team. British Journal of General Practice, 46(405), 231-237.

Kirkby, K., Daniels, B., Jones, I. & McInnes, M. A survey of the social outcome of schizophrenia in Tasmania. Australian and New Zealand Journal of Psychiatry, 31(3), 405-410.

Knapp, M. (1997). Economic evaluations and interventions for children and adolescents with mental health problems. Journal of Child Psychology and Psychiatry, 38(1), 3-25.

Knapp, M. & Kavanagh, S. (1997). Economic outcomes and costs in the treatment of schizophrenia. Clinical Therapeutics, 19(1), 128-138.

Knatterud, G. L., Rockhold, F. W., George, S. L., Barton, F. B., Davis, C. E., Fairweather, W. R., Honohan, T., Mowery, R. & O'Neill, R. (1998). Guidelines for quality assurance in multicenter trials: a position paper. Controlled Clinical Trials, 19(5), 477-93.

Lancashire, S., Haddock, G., Tarrier, N., Baguley, I., Butterworth, C. A. & Brooker, C. (1997). Effects of training on psychosocial interventions for community psychiatric nurses in England. Psychiatric Services, 48(1), 39-41.

Leckman, J. F., Grice, D. E., Boardman, J., Zhang, H., Vitale, A., Bondi, C., Alsobrook, J., Peterson, B. S., Rasmussen, S. A., Goodman, W. K., McDougle, C. J. & Pauls, D. L. (1997). Symptoms of obsessive-compulsive disorder. American Journal of Psychiatry, 154(7), 911-917.

Lee, K. A. & Kieckhefer, G. M. (1989). Measuring human responses using visual analogue scales. Western Journal of Nursing Research, 11(1), 128-32.

Lehman, A. F. (1988). A quality of life interview for the chronically mentally ill. Evaluation and Program Planning, 11, 51-62.

Lehman, A. F., Postrado, L. T. & Rachuba, L. T. (1993). Convergent validity of quality of life assessments for persons with severe mental illness. Quality of Life Research, 2(5), 327-333.

Lemke, W. & Wiersma, W. (1976). Principles of psychological measurement.(pp. 133-149). Rand McNally: Chicago.

Lewis-Fernandez, R. & Kleinman, A. (1995). Cultural psychiatry. Theoretical, clinical, and research issues. Psychiatric Clinics of North America, 18(3), 433-488.

Lindeman, R. H., Merenda, P. F. & Gold, R. Z. (1980). Introduction to bivariate and multivariate analysis (pp. 245-285). Scott, Foresman and Company: Glenview, Illinois.

Linn, R. L. (1968). The Monte Carlo approach to the number of factors problem. Psychometrika, 33, 37-71.

Lowell, W., E. & Davis, G., E. (1994). Predicting length of stay for psychiatric diagnosis-related groups using neural networks. Journal of the American Medical Informatics Association, 1(6), 459-466.

Lteif, G. N & Mavissakalian, M. R. (1996). Life events and panic disorder/agoraphobia; a comparison at two time periods. Comprehensive Psychiatry, 37(4), 241-244.

Lynn, M. R. (1986). Determination and quantification of content validity. Nursing Research, 35(6), 382-5.

Marneros, A., Steinmeyer, E. M., Deister, A., Rohde, A. & Junemann, H. (1989). Long-term outcome of schizoaffective and schizophrenic disorders: a comparative study. III. Social consequences. Europen Archives of Psychiatry and Neurological Science, 238(3), 135-9.

Mason, P., Harrison, G., Glazebrook, C., Medley, I., Dalkin, T. & Croudace, T. (1995). Characteristics of outcome in schizophrenia at 13 years. British Journal of Psychiatry, 167(5), 596-603.

May, P. R., Tuma, A. H., Dixon, W. J., Yale, C., Thiele, D. A. & Kraude, W. H. (1981). Schizophrenia: A follow up study of the results of five forms of treatment. Archives of General Psychiatry, 38(7), 776-784.

McCormack, H. M., Horne, D. J. & Sheather, S. (1988). Clinical applications of visual analogue scales: a critical review. Psychological Medicine, 18(4), 1007-19.

McCormack, H. M., Horne, D. J. & Sheather, S. (1988). Clinical applications of visual analogue scales: a critical review. Psychological Medicine, 18(4), 1007-19.

McCrone, P. & Phelan, M. (1994). Diagnosis and length of psychiatric in-patient stay. Psychological Medicine, 24(4), 1025-1030.

McDowell, I. & Newell, C. (1987). Measuring health: A guide to rating scales and questionnaires (pp. 12-35). New York: Oxford University Press.

McGuire, T. G. (1991). Measuring the economic costs of schizophrenia. Schizophrenia Bulletin, 17(3), 375-88.

McKenna, P. J., Tamlyn, D., Lund, C. E., Mortimer, A. M., Hammond, S. & Baddeley, A. D. (1990). Amnesic syndrome in schizophrenia. Psychological Medicine, 20(4), 967-72.

Mechanic,D. (1995). Sociological dimensions of illness behaviour. Social Science and Medicine, 41(9), 1207-1216.

Microsoft Corporation. (1996). Microsoft word for windows 95 [Computer software]. Seatle: Author.

Mitchell, E. S. (1986). Multiple triangulation: a methodology for nursing science. Advances in Nursing Science, 8(3), 18-26.

Mitrushina, M., Abara, J. & Blumenfeld, A. (1996). A comparison of cognitive profiles in schizophrenia and other psychiatric disorders. Journal of Clinical Psychology, 52(2), 177-90.

Monroe, S. M. (1982). Assessment of life events: retrospective vs concurrent strategies. Archives of General Psychiatry, 39(5), 606-10.

Morley, S. & Pallin, V. (1995). Scaling the affective domain of pain: a study of the dimensionality of verbal descriptors. Pain, 62(1), 39-49.

Munro, B. H. (1991). Measuring relationships. Clinical Nurse Specialist, 5(3), 149-50.

Muliak, S. A. (1972). The foundations of factor analysis (pp. 173-218). McGraw-Hill: New York.

Munley, P. H., Devone, N., Einhorn, C. M., Gash, I. A., Hyer, L. & Kuhn, K. C. (1977). Demographic and clinical characteristics as predictors of length of hospitalisation and readmission. Journal of Clinical Psychology, 33(4), 1093-1099.

National Health and Medical Research Council. (1992). NHMRC statement on human experimentation and supplementary notes. Canberra, Australian Capital Territory: Author.

Nelson, H. E. (1982). National Adult Reading Test, part I. Nfer-Nelson Publishing Company: Windsor, Berkshire.

Nelson, H. E. & Willison, J. R. (1991). National Adult Reading Test, part II. Nfer-Nelson Publishing Company: Windsor, Berkshire.

Neugebauer, R. (1983). Reliability of life-event interviews with outpatient schizophrenics. Archives of General Psychiatry, 40(4), 378-83.

Nolan, M. & Behi, R. (1995a). Validity: a concept at the heart of research. British Journal of Nursing, 4(9), 530-3.

Nolan, M. & Behi, R. (1995b). Measurement in research. British Journal of Nursing, 4(7), 402-5.

Nolan, M. & Behi, R. (1995c). Triangulation: the best of all worlds? British Journal of Nursing, 4(14), 829-32.

Nunally, J. C. & Bernstein, I. H. (1978). Psychometric theory (2nd. ed.) (pp. 86-116). New York: McGraw-Hill.

O'Carroll, R., Walker, M., Dunan, J., Murray, C., Blackwood, D., Ebmeier, K. P. & Goodwin, G. M. (1992). Selecting controls for schizophrenia research studies: the use of the National Adult Reading Test (NART) is a measure of premorbid ability. Schizophrenia Research, 8(2), 137-41.

Offord, D. R. & Boyle, M. H. (1986). Problems in setting up an executing large-scale psychiatric epidemiological studies. Psychiatric Developments, 4(3), 257-72.

Olsson, J., Bengtsson, B., Heijl, A. & Rootzen, H. (1997). An improved method to estimate frequency of false positive answers in computerized perimetry. Acta Ophthalmologia Scandinavica, 75(2), 181-3.

Parker, G. & Hadzi-Pavlovic, D. (1990). Expressed emotion as a predictor of schizophrenic relapse: an analysis of aggregated data. Psychological Medicine, 20(4), 961-5.

Parks, E. D. & Balon, R. (1995). Autobiographical memory for childhood events: patterns of recall in psychiatric patients with a history of alleged trauma. Psychiatry, 58(3), 199-208.

Parry, R., Turns, D. M., Stephenson, J. J., Tobias, C. & Lippmann, S. (1992). Disability status and length of stay at a VA medical center. Hospital and Community Psychiatry, 43(8), 844-845.

Paulsen, J. S., Heaton, R. K., Sadek, J. R., Perry, W., Delis, D. C., Braff, D., Kuck, J., Zisook, S. & Jeste, D. V. (1995). The nature of learning and memory impairments in schizophrenia. Journal of the International Neuropsychological Society, 1(1), 88-99.

Perlow, M. (1992). Validity and reliability of the PSES. Western Journal of Nursing Research, 14(2), 201-10.

Phelan, M., Wykes, T. & Goldman, H. (1994). Global function scales. Journal of Social Psychiatry and Psychiatric Epidemiology, 29(5), 205-11.

Pinol, A., Bergel, E., Chaisiri, K., Diaz, E. & Gandeh, M. (1998). Managing data for a randomised controlled clinical trial: experience from the WHO Antenatal Care Trial. WHO Antenatal Care Trial Research Group. Paediatric and Perinatal Epidemiology, 12(2), 142-55.

Pocock, S. J. (1991). A perspective on the role of quality-of-life assessment in clinical trials. Controlled Clinical Trials, 12(4), 257-265.

Polit, D. F. & Hungler, B. P. (1989). Essentials of nursing research; Methods, appraisal and utilization (2nd ed.) (pp. 117-148). New York: Lippincott.

Pollatsek, A. & Well, A. D. (1995). On the use of counterbalanced designs in cognitive research: a suggestion for a better and more powerful analysis. Journal of Experimental Psychology, Learning, Memory and Cognition, 21(3), 785-94.

Prendergast, P. J. (1995). Integration of psychiatric rehabilitation in the long-term management of schizophrenia. Canadian Journal of Psychiatry, 40(3), 18-21.

Provencher, H. L., Fournier, J. P. & Dupuis, N. (1997). Schizophrenia: revisited. Journal of Psychiatric and Mental Health Nursing, 4(4), 275-285.

Remmington, G. J. (1995). The long-term treatment of schizophrenia. Canadian Journal of Psychiatry, 40(3), 3-4.

Rew, L., Stuppy, D. & Becker, H. (1988). Construct validity in instrument development: a vital link between nursing practice, research, and theory. ANS Advances in Nursing Science, 10(4), 10-22.

Rhoades, H. M. & Overall, J. E. (1988). The semi-structured BPRS interview and rating guide. Psychopharmacology Bulletin, 24, 101-104.

Rice, D. P. (1999). The economic impact of schizophrenia. Journal of Clinical Psychiatry, 60(1), 4-6.

Rizzo, L., Danion, J. M., van der Linden, M. & Grange, D. (1996). Patients with schizophrenia remember that an event has occurred, but not when. British Journal of Psychiatry, 168(4), 427-31.

Robbins, L. N. & Helzer, J. E. (1986). Diagnosis and clinical assessment: The current state of psychiatric diagnosis. Annual Review of Psychology, 37, 409-432.

Robinson, D., Woerner, M. G., Pollack, S. & Lerner, G. (1996). Subject selection biases in clinical trials: data from a multicenter schizophrenia treatment study. Journal of Clinical Psychopharmacology, 16(2), 170-176.

Roman, P. M. (1971). Labelling theory and community psychiatry: The impact of psychiatric sociology on ideology and practice in American psychiatry. Psychiatry, 34(4), 378-390.

Roos, L. L. & Wajda, A. (1991). Record linkage strategies. Part I: Estimating information and evaluating approaches. Methods of Information in Medicine, 30(2), 117-23.

Rosen, A., Hadzi-Pavlovic, D, & Parker, G. (1989). The life skills profile: a measure assessing function and disability in schizophrenia. Schizophrenia Bulletin, 15(2), 325-337.

Rossi, A., Stratta, P., Bolino, F., Giordano, L. de Leonardis, R., Schiazza, G. & Casacchia, M. (1989). Social disability and the clinical symptomatology in schizophrenia: their impact on compulsory admission. Psychopathology, 22(4), 177-181.

Rupp, A. & Keith, S. J. (1993). The costs of schizophrenia. Assessing the burden. Psychiatric Clinics of North America, 16(2), 413-423.

Rushe, T. M., Woodruff, P. W., Murray, R. M. & Morris, R. G. (1999). Episodic memory and learning in patients with chronic schizophrenia. Schizophrenia Research, 35(1), 85-96.

Ryan, P. M. (1991). Epidemiology, etiology, diagnosis and treatment of schizophrenia. American Journal of Hospital Pharmacotherapy, 48(6), 1271-1280.

Sandanger, I., Moum, T., Ingebrigtsen, G., Sorensen, T., Dalgard, O. S. & Bruusgaard, D. (1999). The meaning and significance of caseness: the Hopkins Symptom Checklist- 25 and the Composite International Diagnostic Interview. II. Social Psychiatry and Psychiatric Epidemiology , 34(1), 53-9.

Sartorius, N. (1993). WHO's work on the epidemiology of mental disorders. Social Psychiatry and Psychiatric Epidemiology, 28(4), 147-155.

Sartorius, N., Jablensky, A., Korten, A., Ernberg, G., Anker, M., Cooper, J. E. & Day, R. (1986). Early manifestations and first-contact incidence of schizophrenia in different cultures. A preliminary report on the initial evaluation phase of the WHO Collaborative Study on determinants of outcome of severe mental disorders. Psychological Medicine, 16(4), 909-28.

Sartorius, N., Kaelber, C. T., Cooper, J. E., Roper, M. T., Rae, D. S., Gulbinat, W., Ustun, T. B. & Regier, D. A. (1993). Progress towards achieving a common language in psychiatry. Results from the field trial of clinical guidelines accompanying the WHO classification of mental and behavioural disorders in ICD-10. Archives of General Psychiatry, 50(2), 115-124.

Sartorius, N., Shapiro, R., Kimura, M. & Barrett, K. (1972). WHO international pilot study of schizophrenia. Psychological Medicine, 2(4), 422-5.

SAS [Computer software]. (1991). Cary, North Carolina: SAS Institute Incorporated.

Saykin, A. J., Gur, R. C., Gur, R. E., Mozley, P. D., Mozley, L. H., Resnick, S. M., Kester, D. B. & Stafiniak, P. (1991). Neuropsychological function in schizophrenia. Selective impairment in memory and learning. Archives of General Psychiatry, 48(7), 618-24.

Schroder, J., Tittel, A., Stockert, A. & Karr, M. (1996). Memory deficits in subsyndromes of chronic schizophrenia. Schizophrenia Research, 21(1), 19-26.

Schumacher, D. N., Namerow, M. J., Parker, B., Fox, P. & Kofie, V. (1986). Prospective payment for psychiatry-feasability and impact. New England Journal of Medicine, 315(21), 1331-1336.

Schuntermann, M. F. (1996). The international classification of impairments, disabilities and handicaps (ICIDH) - results and problems. International Journal of Rehabilitation Research, 19, 1-11.

Schwartz, R. J., Weiss, K. M. & Buchanan, A. V. (1985). Error control in medical data. MD Computing, 2(2), 19-25.

Schwartz, B. L., Deutsch, L. H., Cohen, C., Warden, D. & Deutsch, S. I. (1991). Memory for temporal order in schizophrenia. Biological Psychiatry, 29(4), 329-39.

Selby-Harrington, M. L., Mehta, S. M., Jutsum, V., Riportella-Muller, R. & Quade, D. (1994). Reporting of instrument validity and reliability in selected clinical nursing journals, 1989. Journal of Professional Nursing, 10(1), 47-56.

Sengel, R. A. & Lovallo, W. R. (1983). Effects of cueing on immediate and recent memory in schizophrenics. Journal of Nervous and Mental Disease, 171(7), 426-30.

Sheehan, D. V., Harnett-Sheehan, K. & Raj, B. A. (1996). The measurement of disability. International Clinical Psychopharmacology, 3, 89-95.

Shepherd, M., Watt, D., Falloon, I. & Smeeton, N. (1989). The natural history of schizophrenia: a five-year follow-up study of outcome and prediction in a representative sample of schizophrenics. Psychological Medicine Monograph 15, 1-46.

Shih, F. J. (1998). Triangulation in nursing research: issues of conceptual clarity and purpose. Journal of Advanced Nursing, 28(3), 631-41.

Sibeko, C. R. & Greeff, M. (1995). Psychiatric nurses' communication with psychiatric patients. Curationis, 18(4), 15-19.

Slade, M., Rosen, A. & Shankar, R. Multidisciplinary mental health teams. International Journal of Social Psychiatry, 41(3), 180-189.

Slocumb, E. M. & Cole, F. L. (1991). A practical approach to content validation. Applied Nursing Research, 4(4), 192-5.

Smith, G. R., Babor, T., Burnam, A., Rost, K. M., Drake, R. & Heithoff, K. The Substance Abuse Outcomes Module. (1995). Little Rock: University of Arkansas Press.

Smith, G. R., Waxman, H., Snyder, D. & Raphael, M. (1996). Assessment tools in psychiatry. American Journal of Medical Quality, 11(1), S46-9.

Sohier, R. (1988). Multiple triangulation and contemporary nursing research. Western Journal of Nursing Research, 10(6), 732-42.

Spiker, C. C. (1989). Cognitive psychology: mentalistic or behavioristic? Advances in Child Development and Behavior, 21, 73-90.

Spitzer, R. L. (1986). Psychiatric diagnosis; a clinicians still necessary? Comprehensive Psychiatry, 24(5), 399-411.

SPSS [Computer software]. (1996). Chicago: SPSS Incorporated.

Starker, J. (1986). Methodological and conceptual issues in research on social support. Hospital and Community Psychiatry, 37(5), 485-90.

Starkey, D. & Flannery, R. B. (1997). Schizophrenia, psychiatric rehabilitation, and healthy development: a theoretical framework. Psychiatric Quarterly, 68(2), 155-166.

Statistica [Computer software]. (1994). Tulsa: Statsoft Incorporated.

Stevens, J. (1986). Applied multivarite statistics for the social sciences (pp.337-372). Lawrence Erlbaum: Hillsdale, New Jersey.

Stip, E. (1996). Memory impairment in schizophrenia: perspective's from psychopathology and pharmacotherapy. Canadian Journal of Psychiatry, 41(8, Suppl. 2), S27-34.

Strauss, M. E. (1973). Behavioural differences between acute and chronic schizophrenics: course of psychosis, effects of institutionalization, or sampling biases? Psychological Bulletin, 79(5), 271-279.

Strauss, J. & Carpenter, W. (1981). Schizophrenia. New York: Plenum.

Sullivan, P. (1998). Developing evidence-based care in mental health nursing. Nursing Standard, 12(31), 35-38.

Summers, S. (1993). Establishing the reliability and validity of a new instrument: pilot testing. Journal of Post Anesthetic Nursing, 8(2), 124-7.

Tessler, R.C. & Gamache, G. M. (1995). Toolkit for evaluating family experiences with severe mental illness. Cambridge, MA: Human Research Institute.

Thomas, M. R., Stoyva, J., Rosenberg, S. A., Kassner, C., Fryer, G. E., Giese, A. A. & Dubovsky, S. L. (1997). Selection bias in an inpatient outcomes monitoring project. General Hospital Psychiatry, 19(1), 56-61.

Thornley, B. & Adams, C. (1998). Content and quality of 2000 controlled trials in schizophrenia over 50 years. British Medical Journal, 317(7167), 1181-1184.

Tiernany, A., Closs, J., Atkin, I., Anderson, J., Murphy-Black, T. & Macmillan, M. (1988). On measurement and nursing research. Nursing Times, 84(12), 55-58.

Tilden, V. P. (1985). Issues of conceptualization and measurement of social support in the construction of nursing theory. Research in Nursing and Health, 8(2), 199-206.

Tracy, J. I., McGrory, A. C., Josiassen, R. C. & Monaco, C. A. (1996). A comparison of reading and demographic-based estimates of premorbid intelligence in schizophrenia. Schizophrenia Research, 22(2), 103-9.

Treno, A. J., Gruenewald, P. J. & Johnson, F. W. (1998). Sample selection bias in the emergency room: an examination of the role of alcohol in injury. Addiction, 93(1), 113-129.

Tucker, L. R., Koopman, R. F. & Linn, R. L. (1969). Evaluation of factor analytic research procedures by means of simulated correlation matrices. Psychometrika, 34, 421-459.

Ustun, T. B., Cooper, J. E., van Duuren-Kristen, C., Kennedy, G., Hendershot, G. & Sartorius, N. (1995). Revision of the ICIDH: mental health aspects. Disability and Rehabilitation, 17(3/4), 202-209

Veltro, F., Magliano, L., Lobrace, S., Morosini., P. L. & Maj, M. (1994). Burden on key relatives of patients with schizophrenia vs neurotic disorders: a pilot study. Social Psychiatry and Psychiatric Epidemiology, 29(2), 66-70.

Volavka, J., Laska, E., Baker, S., Meisner, M., Czobor, P. & Krivelevich, I. (1997). History of violent behaviour and schizophrenia in different cultures. Analyses based on the WHO study on Determinants of Outcome of Severe Mental Disorders. British Journal of Psychiatry, 171, 9-14.

Volavka, J., Laska, E., Baker, S., Meisner, M., Czobor, P. & Krivelevich, I. (1997). History of violent behaviour and schizophrenia in different cultures. Analyses based on the WHO study on Determinants of Outcome of Severe Mental Disorders. British Journal of Psychiatry, 171, 9-14.

Wagner, G. & Newcombe, H. B. (1970). Record linkage. Its methodology and application in medical data processing. Methods of Information in Medicine, 9(2), 121-38.

Waien, S. A. (1997). Linking large administrative databases: a method for conducting emergency medical services cohort studies using existing data. Academic Emergency Medicine, 4(11), 1087-95.

Walz, C., Strickland, O. & Lenz, E. (1991). Measurement in nursing research. Philadelphia: Davis.

Ware, J. E. & Sherbourne, C. D. (1992). The MOS 36-item short-form health status survey (SF-36). Medical Care, 30(6), 473-483.

Wasylenki, D. A. (1994). The cost of schizophrenia. Canadian Journal of Psychiatry, 39(9), 65-69.

Weinstein, R. M. (1994). Goffman's Asylums and the total institution model of mental hospitals. Psychiatry, 57(4), 348-67.

West, P. & Isenberg, M. (1997). Instrument development: the Mental Health-Related Self-Care Agency Scale. Archives of Psychiatric Nursing, 11(3), 126-32.

Wiersma, D. (1986). Psychological impairments and social disabilities; on the applicability of the ICIDH to psychiatry. International Rehabilitation Medicine, 8(1), 3-7.

Wiersma, D., DeJong, A. & Ormel, J. (1988). The Groningen social disabilities schedule: development, relationship with ICIDH, and psychometric properties. International Journal of Rehabilitation Research, 11(3), 213-24.

Wiersma, D., Nienhuis, F. J., Giel, R., de Jong, A. & Slooff, C. J. (1996). Assessment of the need for care 15 years after onset of a Dutch cohort of patients with schizophrenia, and an international comparison. Social Psychiatry and Psychiatric Epidemiology, 31(3/4), 114-21.

Wiersma, D. (1996). Measuring social disabilities in mental health. Social Psychiatry and Psychiatric Epidemiology , 31(3-4), 101-8.

Wiersma, D., Giel, R., de Jong, A., Nienhuis, F. J. & Slooff, C. J. (1996). Assessment of the need for care 15 years after onset of a Dutch cohort of patients with schizophrenia, and an international comparison. Journal of Social Psychiatry and Psychiatric Epidemiology, 31(3-4), 114-21.

Wiersma, D., Nienhuis, F. J., Giel, R. & Slooff, C. J. (1998). Stability and change in needs of patients with schizophrenic disorders: a 15- and 17-year follow-up from first onset of psychosis, and a comparison between 'objective' and 'subjective' assessments of needs for care. Social Psychiatry and Psychiatric Epidemiology , 33(2), 49-56.

Wiersma, D., Nienhuis, F. J., Slooff, C. J. & Giel, R. (1998). Natural course of schizophrenic disorders: a 15-year follow-up of a Dutch incidence cohort. Schizophrenia Bulletin, 24(1), 75-85.

Wilhelm, S., McNally, R. J., Baer, L. & Florin, I. (1997). Autobiographical memory in obsessive-compulsive disorder. British Journal of Clinical Psychiatry, 36(1), 21-31.

Williams, R. & Dickson, R. A. (1995). Economics of schizophrenia. Canadian Journal of Psychiatry, 40(7), 60-70.

Wing, J. K. (1995). Concepts of schizophrenia. In S. R. Hirsch & D. R. Weinberger (Eds.). Schizophrenia (pp. 3-14). Oxford: Blackwell Sciences Limited.

Wing, J. E., Curtis, R. H. & Beevor, A. S. (1996). HoNOS: Health of the nation outcome scales. London: Royal College of Psychiatrists.

Wing, J. K., Sartorius, N. & Ustun, T. B. (1998). Measurement and classification in psychiatry. In J. K. Wing, N. Sartorius, & T. B. Ustun (Eds.), Diagnosis and clinical measurement in psychiatry (pp. 1-11). Cambridge: Cambridge University Press.

Wood, P. H. N. (1989). Measuring the consequences of illness. World Health Statistics Quarterly, 42, 115-121.

World Health Organisation (1973). The international pilot study of schizophrenia. Geneva, Switzerland: Author.

World Health Organisation (1979). Schizophrenia: An international follow-up study. New York: Wiley.

World Health Organisation (1980). International classification of impairments, disabilities and handicaps. Geneva, Switzerland: Author.

World Health Organisation. (1988). WHO Psychiatric disability assessment schedule. Geneva, Switzerland: Author.

World Health Organisation (1992). SCAN: Schedules for clinical assessment in neuropsychiatry. Geneva, Switzerland: Author.

World Health Organisation (1992). The ICD-10 classification of mental and behavioural disorders: clinical descriptions and diagnostic guidelines. Geneva, Switzerland: Author.

Zeller, R. A. & Carmines, E. G. (1980). Measurement in the social sciences: The link between theory and data.(pp. 77-100). New York: Cambridge University Press.

Zhan, L. 91992). Quality of life: conceptual and measurement issues. Journal of Advanced Nursing, 17(7), 795-800.

Appendix A

Converting the SAS Dataset into an SPSS Dataset (data & format)

In SAS run the following code:

```
options ls=80 ps=60
libname dd 'c:\sas\who\datasets';
proc format cntlout=formats;
run;

libname tranfile sasv5xpt 'c:\sas\who\datasets\dasiu.dat';
libname dd 'c:\sas\who\datasets';
proc copy in=dd out=tranfile
select dasiu;
run;

libname tranfile sasv5xpt 'c:\sas\who\datasets\dasiufmt.dat';
libname dd 'c:\sas\who\datasets';
proc copy in=dd out=tranfile
select dasiu;
run;
```

In SPSS run the following code

```
get sas data='c:\sas\who\datasets\dasiu.dat'dset (dasiu)
/formats='c:\sas\who\datasets\dasiufmt.dat' fset (dasiu.fmt)
execute
```

Appendix B

Expert panel

Name	Qualifications	Position	Area of expertise		
			Clinical	Research	Theory
Assen JABLENSKY	MD DMSc	Professor of Psychiatry	X	X	X
Eamon SHANLEY	RN, MA, Ph.D	Professor of Mental Health Nursing	X		X
David CASTLE	MD MSc	Senior Lecturer	X	X	
John COOPER	MD Ph.D	Emeritus Professor	X	X	X
Farooq AHMAD	BM Ch.B	Consultant Psychiatrist	X		
Aaron GROVES	MB BS	Deputy Chief Psychiatrist for Western Australia	X		
Aleksander JANCA	MD	Associate Professor of Psychiatry	X	X	X
Peter WYNN OWEN	BM ChB (Birm.)	Director of Postgraduate Training, RANZCP	X		X

Appendix C

Dummy Data Included in the First Draft

A corrected version of draft one is given in Appendix I.

1.Self care

0. Not a problem
1. Some minimal loss in self care (e.g. dirty clothes,
2. Obviously lacking but not dangerously so (e.g. rotten teeth or nails seriously overgrown)
3. Seriously lacking (e.g. lice in hair or never washes, unshaven, hair unkempt or dangerously malnourished or malodorous)
4. Total lack of any self care activity – totally dependent on others

Catch target: The example "malodourous" given in stem 3 is a descriptor for stem 2 and not an example of "seriously lacking" self care (the stem 3 grouping descriptor).

2. Underactivity

0. Not a problem
1. Aimless or futile activity for hours at a time but whilst noteworthy not of concern to patient or others
2. Markedly reduced activity – does little beyond basic routine of eating, sleeping, dressing. Spends extended periods (hours) doing nothing
3. Very restricted activity – most of day spent inactive. No interests or daily involvement. Passively obeys direct instructions but if not supervised returns to inactive state.
4. Total lack of any daily activity. Does absolutely nothing – needs to be strongly coerced to carry out even the most basic activities

Catch target: The item anchor title "Underactivity" is incongruent with the item content which refers to "Purposeful Activity"

4. Social Withdrawal

0. Not a problem
1. Taciturn and solitary in social situations but can be engaged in a limited manner
2. Narrow range of social contacts (only one or two friends) and even with these individuals only limited interaction takes place
3. No social contact at all – totally unresponsive to any social interaction and avoids all social situations all of the time
4. Actively avoids social contact most of the time. No friends. Solitary lifestyle with only minimal contact with agency staff eg. Community nurses, DSS officers.

Catch trial: Stem descriptors 3 and 4 are transposed such that stem 3 represented less severe social withdrawal than stem 4, contrary to the scaling system

10. Possess matches

0. Not a problem
1. Needs to be checked occasionally and reminded about the possibility of risks in handling inflammables
2. Unsupervised during daytime but at risk at night where inflammables need to be locked away (e.g.. Would smoke in bed, use lighter to find way about in the dark)
3. Can be trusted to cook or smoke or be in possession of matches but only under general supervision
4. Can never be given responsibility for matches or lighter without being a serious risk. Needs constant direct supervision smoking, cooking or any activity that is a fire risk

Catch trial: Anchor heading "Possess matches" is overly specific and does not reflect the content of the item as set out in the stem descriptors

Appendix D

Blinded Performance Table of Experts

	Failure to identify dummy variable			
Expert	Item 1	Item 2	Item 4	Item 11
1				
2	X			X
3				
4		X		
5				
6				
7	X			

Appendix E

Explained Variance of Each Factor for Each Dataset

Table 1

Factor Loadings (Varimax Normalised) dasiu

	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6	Factor 7	Factor 8	Factor 9	Factor 10	Factor 11	Factor 12
CARE11	-0.03446	-0.03115	0.027048	0.801357	0.034649	0.052371	0.008096	-0.02848	-0.03384	0.135324	-0.07042	0.007681
UNACT12	-0.10772	-0.01912	0.085324	0.689711	0.024304	0.065911	0.018689	0.198819	0.016498	0.229336	-0.18618	0.16554
SLOW13	-0.06884	0.048541	-0.01967	0.686024	0.007091	0.034655	-0.01874	0.087058	-0.01308	0.246212	-0.19798	0.135559
SWITH14	0.016488	-0.01289	0.040171	0.766468	0.027868	0.10374	0.020975	0.063756	0.103475	0.131569	0.046492	-0.05218
HSHO21	0.344333	0.157417	0.188748	0.227306	0.043365	0.69515	0.043772	-0.09638	-0.04633	0.081595	-0.12726	0.072785
AFFECT22	0.05797	0.096609	0.912921	0.086504	0.047324	0.065799	-0.00808	-0.01475	-0.01157	0.02388	0.034125	-0.07626
SEX23	-0.02154	0.086384	0.864575	0.035119	0.008432	0.102842	0.00133	-0.02937	0.07504	0.030216	-0.02082	-0.02793
CHILD24	0.08123	0.001889	0.790115	0.052696	0.024506	0.085803	-0.00304	0.011326	-0.1448	0.04235	-0.07817	0.063914
HETSEX25	0.257985	0.088607	-0.64806	-0.0131	0.052332	0.516802	0.042913	-0.19932	-0.01588	0.016425	-0.03176	-0.03388
FRICT26	0.409863	0.134406	0.0574	0.094296	0.014038	0.661421	-0.16458	-0.21537	-0.047	0.035566	0.002669	-0.07534
PERF27	0.194569	0.028061	-0.07201	0.18608	0.080014	0.580264	0.116583	0.484041	-0.03714	0.012367	0.247173	-0.28426
JOB28	0.136131	0.103853	-0.04788	-0.07927	0.039605	0.182273	0.068099	-0.86361	-0.04529	0.024923	0.026322	-0.07383
INFO29	0.310817	0.146854	0.138021	0.184849	-0.00546	0.712335	0.037294	-0.04825	-0.04225	0.115437	-0.11301	0.03595
CRIS210	0.073742	-0.01514	0.109892	-0.00224	-0.046	0.104352	-0.01332	-0.05389	-0.91909	0.072948	-0.00846	0.010115
MOV311	-0.96042	-0.08997	0.004556	0.058037	0.001907	-0.13435	-0.00092	0.016962	0.014836	-0.02483	0.005649	0.041382

UNDAC312	-0.97341	-0.09645	-3.3E-05	0.048812	-0.00763	-0.12936	0.001228	0.021158	0.02134	-0.045	0.002814	0.052671
OVER313	-0.97656	-0.09799	0.000457	0.041161	-0.01706	-0.12361	0.001157	0.017141	0.02207	-0.0471	0.002239	0.04565
CONVR314	-0.97544	-0.08663	-0.00174	0.0456	-0.00999	-0.13349	-0.00099	0.019792	0.018321	-0.04658	0.005506	0.04141
SWITH315	-0.97163	-0.08826	-0.00332	0.0515	-0.01303	-0.14857	-0.00014	0.024061	0.02066	-0.03533	0.008844	0.04815
LEIS316	-0.96905	-0.07919	0.000396	0.047595	-0.00439	-0.11821	-0.00054	0.01623	0.017325	-0.05537	0.00774	0.049347
IRREL317	-0.97497	-0.09611	-0.00545	0.040216	-0.0265	-0.14253	-0.00046	0.024193	0.023644	-0.03673	-0.00022	0.039336
POSTU318	-0.97869	-0.08982	-0.00238	0.038	-0.01607	-0.13263	-2.3E-05	0.021336	0.022482	-0.04513	0.003653	0.042272
THRT319	-0.97755	-0.09426	-0.00285	0.03858	-0.01157	-0.13121	0.000658	0.022931	0.022509	-0.04921	-0.00258	0.046345
BED3110	-0.97201	-0.08705	-0.00219	0.037279	-0.00952	-0.12574	0.001464	0.017697	0.021897	-0.03577	0.007332	0.040651
APP3111	-0.97859	-0.09436	-0.00153	0.03969	-0.01081	-0.13012	0.000248	0.018363	0.022909	-0.04663	0.000363	0.035384
MEAL3112	-0.98001	-0.09237	-0.00257	0.036392	-0.01071	-0.12815	0.001137	0.020026	0.022172	-0.04447	-0.00279	0.036192
WORK321	-0.96251	-0.09527	0.008968	-0.00186	-0.02395	-0.0963	0.003544	0.012408	0.015099	-0.05625	0.010842	0.026407
MATCH322	-0.97102	-0.1028	-0.01618	0.019266	-0.01067	-0.12641	0.001118	0.024778	0.018952	-0.05299	0.009916	0.034907
VISIT323	-0.97775	-0.10312	0.00184	0.022164	-0.01702	-0.11069	0.002619	0.019176	0.022243	-0.06856	-0.00285	0.041534
GOOUT324	-0.91826	-0.083	-0.01353	0.008216	-0.04418	-0.12836	-0.00021	0.023779	0.017956	-0.04542	0.021658	-0.02697
SCISS325	-0.95827	-0.10296	-0.01405	0.011014	-0.01637	-0.10432	0.00224	0.026736	0.014132	-0.06372	0.011741	0.029633
MONEY326	-0.97173	-0.09916	0.001026	0.01055	0.002928	-0.10921	0.001782	0.024894	0.019199	-0.06026	0.02047	0.051317
OUTSD327	-0.90271	-0.08621	-0.00627	-0.03849	-0.02338	-0.13395	0.012905	0.031969	0.010112	-0.05563	0.004246	-0.03224
DISCH328	-0.95276	-0.106	-0.00376	0.020773	-0.03511	-0.13976	0.002068	0.035983	0.031407	-0.06092	-0.01414	0.018023
LOCK329	-0.88099	0.022218	0.013219	-0.05514	0.039281	-0.0101	0.003106	-0.00866	0.02005	0.080936	0.072836	0.00554
HOUSE331	-0.92161	-0.10257	-0.00739	-0.00058	-0.05159	-0.08067	0.005426	0.021389	0.015837	-0.08507	0.000507	0.037872
OUTWD332	-0.70211	0.031348	0.027381	-0.20488	-0.0871	0.093993	0.011969	0.04419	-0.01181	-0.08851	-0.0087	0.074517
OT333	-0.92108	-0.12682	0.006281	-0.0582	-0.07053	-0.03159	0.005697	0.03999	0.01982	-0.07447	-0.00997	0.038096
SOCTH334	-0.87734	-0.04503	0.015103	-0.03926	-0.04424	0.045039	0.007633	0.011606	0.002158	-0.13003	-0.01824	0.031128
BEVIS341	-0.86347	-0.08653	-0.0169	0.017182	-0.07218	-0.21418	-0.00708	0.034259	0.019373	0.017214	-0.00321	-0.14068
HOME342	-0.80505	-0.04722	-0.05444	0.035942	-0.08234	-0.28518	-0.00718	0.063925	0.010176	-0.01869	0.009857	-0.14207
PAROL343	-0.68004	0.142316	-0.01806	-0.11407	-0.03694	-0.14147	-8.4E-05	0.04172	-0.02225	0.038103	0.074397	-0.20528
SUPER344	-0.91958	-0.09884	0.006484	0.057893	-0.04724	-0.0792	-0.00405	-0.00213	0.023013	-0.00366	0.015314	0.079259
SKILL411	0.0437	0.006082	0.020295	0.074504	0.267593	0.075787	0.030002	-0.04996	0.029429	0.123593	-0.75221	-0.13054
INTER412	0.031603	0.071396	0.035351	0.11449	0.207577	0.099938	0.008862	0.033628	0.012913	-0.0237	-0.81529	-0.02939

ASSET413	0.005142	-0.04489	-0.00932	0.063685	0.591071	-0.05281	0.086149	0.079386	-0.03494	-0.06234	-0.38841	0.041184
FACT414	0.078281	0.008956	0.021982	-0.05105	0.80391	0.037846	-0.14463	-0.05103	0.013421	-0.01721	-0.14389	0.007068
CONFD415	-0.02847	-0.0383	0.110486	-0.01827	0.089694	-0.00178	-0.19258	0.002871	0.044358	0.093256	-0.41746	-0.50077
ENVIR421	0.014767	-0.00011	0.062456	0.030908	0.833654	0.122495	0.071299	-0.04258	0.070844	0.018878	-0.09427	-0.04231
HDCAP422	0.106927	0.128268	0.042179	0.204267	0.520628	-0.0977	0.057338	-0.03609	-0.00447	0.044529	-0.25384	-0.16096
LIAB423	0.096034	0.041734	-0.01155	-0.03505	0.719594	0.08002	-0.14152	-0.01661	-0.03588	0.110768	0.082548	0.028835
KEYFG431	0.188228	0.338089	-0.23569	-0.14632	0.090192	0.235977	0.019363	-0.11265	-0.037	0.14434	-0.13161	-0.2176
HOURS432	0.041527	0.785603	-0.00362	-0.02226	0.051631	0.057363	0.027305	-0.00333	-0.01496	0.050793	0.001136	0.010851
EE433	0.243763	0.827552	0.090522	-0.03104	-0.00086	0.146994	-0.0237	-0.0829	-0.00461	0.143058	0.011711	-0.05764
KCONT434	0.248593	0.838991	0.092639	-0.01868	0.028843	0.128724	0.005003	0.017221	-0.00637	0.117632	-0.06675	-0.05007
REJ435	0.287081	0.808588	0.095727	0.033018	-0.03034	0.144358	0.003564	-0.0744	-0.01115	0.161995	0.024855	-0.05314
PRIV436	0.188995	0.671601	0.114128	0.081651	0.034944	0.171536	0.005732	-0.08778	0.003809	0.263492	-0.01715	0.019935
SEEK441	0.019876	-0.01029	0.02269	-0.0303	0.056145	-0.00383	-0.97755	0.040359	-0.00185	-0.02314	0.005092	-0.01475
RECE442	0.012678	-0.01729	0.014164	-0.02817	0.056802	-0.00476	-0.97887	0.042221	-0.00726	-0.01886	0.004126	-0.0122
JOB443	0.130353	0.129691	0.008557	0.003642	0.015987	0.098181	0.04523	-0.08517	-0.03301	0.19073	-0.07755	-0.72425
GLOB5	-0.00174	0.003417	0.064105	0.742798	0.015529	0.037503	0.029279	0.032016	-0.0321	-0.01344	0.104697	-0.15974
Explained variance	26.4969	3.908994	5.568065	3.274974	2.711323	5.051235	2.110715	2.150972	1.850668	3.261184	1.990102	1.348922
Proportion of total	0.353292	0.05212	0.074241	0.043666	0.036151	0.06735	0.028143	0.02868	0.024676	0.043482	0.026535	0.017986

Table 2

Factor Loadings (Varimax Normalised) dasfu1

	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6	Factor 7	Factor 8	Factor 9	Factor 10	Factor 11	Factor 12
CARE11	0.090132	0.357512	-0.08182	-0.11607	0.04424	-0.01952	0.064405	0.04832	0.042364	0.819537	0.053583	0.113964
UNACT12	0.081033	0.304859	-0.1403	-0.20408	0.012832	0.01343	0.091832	0.030736	0.126876	0.763085	0.063897	0.072183
SLOW13	0.089561	0.411041	-0.06093	-0.07683	0.063065	-0.04605	-0.02165	-0.00115	-0.05634	0.744044	0.076187	0.185277
SWITH14	0.058523	0.329193	-0.06067	-0.07281	0.09478	-0.03443	0.123865	0.006458	0.030703	0.761229	0.109515	0.102465
HSHO21	0.273961	0.201583	-0.13072	-0.03133	0.193295	-0.06092	0.749632	0.035918	0.030936	0.42169	0.063003	-0.04632
AFFECT22	0.0825	0.073907	-0.90509	-0.02893	0.069231	0.010535	0.009502	0.08384	0.079357	0.134784	0.073711	-0.05588
SEX23	0.066717	0.063817	-0.85557	-0.10818	0.10758	-0.01075	-0.0198	0.187839	0.018981	0.133972	0.014666	-0.09921
CHILD24	0.07482	0.047601	-0.82408	0.079804	0.003995	-0.00223	0.143802	-0.0692	-0.03002	0.083463	0.040234	0.030569
HETSEX25	0.083354	0.108505	0.71828	-0.25222	0.025926	-0.01837	0.147316	0.14359	-0.29379	0.124672	0.168511	0.005843
FRICT2	0.250692	0.255251	-0.03302	-0.04122	0.124448	-0.04158	0.058997	0.002433	-0.01337	0.432831	0.727955	-0.00448
PERF27	0.071731	0.028074	0.027947	-0.83553	0.035765	-0.03416	0.012028	0.006605	0.234179	0.287982	0.046433	0.039218
JOB28	0.02808	0.081268	0.132219	0.216281	0.072142	0.014896	-0.03659	-0.01075	-0.87184	0.029155	0.026818	-0.05883
INFO29	0.278588	0.236724	-0.08796	-0.15073	0.150426	-0.04575	0.136704	0.104725	-0.05907	0.608102	0.347811	0.109887
CRIS210	0.022401	-0.00552	-0.09607	0.117712	0.045338	-0.02376	-0.03029	0.865645	0.030456	0.14446	0.060899	-0.04688
MOV311	-0.95224	-0.02144	0.021708	-0.01305	-0.05887	0.00495	-0.03025	0.021807	-0.00231	-0.03397	0.004841	-0.00832
UNDAC312	-0.99069	-0.02603	0.023378	0.020017	-0.03369	0.005455	-0.01659	0.002164	0.012937	-0.03871	-0.01873	-0.01903
OVER313	-0.95955	-0.01348	0.034308	0.015099	-0.03714	0.005499	-0.01993	0.004166	0.006848	-0.05428	-0.01121	-0.03145
CONVR314	-0.99239	-0.01745	0.023578	0.022624	-0.04018	0.004993	-0.01581	0.002494	0.014199	-0.03579	-0.01956	-0.01418
SWITH315	-0.99041	-0.03029	0.022833	0.020804	-0.04058	0.004789	-0.00932	0.003639	0.009695	-0.03458	-0.00992	-0.0115
LEIS31	-0.98964	-0.01899	0.023964	0.020362	-0.03418	0.005633	-0.01698	0.002576	0.011871	-0.03706	-0.01577	-0.01776
IRREL317	-0.98982	-0.01943	0.023953	0.020511	-0.04054	0.00486	-0.01265	0.006754	0.012454	-0.05277	-0.00659	-0.02719
POSTU318	-0.99146	-0.01822	0.024755	0.021347	-0.0426	0.004958	-0.01533	0.005938	0.013415	-0.04703	-0.01084	-0.02157

THRT319	-0.9916	-0.01928	0.024282	0.022651	-0.04648	0.005026	-0.01707	0.00648	0.012996	-0.04162	-0.01131	-0.01704
BED3110	-0.9693	-0.02111	0.013676	0.003547	-0.06298	0.004414	-0.01848	-0.03935	0.025255	-0.04729	-0.03782	-0.00954
APP3111	-0.99047	-0.0198	0.024369	0.019179	-0.03795	0.005757	-0.02202	0.005332	0.008061	-0.04495	-0.0137	-0.02462
MEAL3112	-0.99243	-0.01701	0.024891	0.022116	-0.04371	0.00495	-0.01736	0.005493	0.014646	-0.0465	-0.01434	-0.02119
WORK321	-0.95106	-0.01567	0.019926	0.017499	-0.07641	0.001416	0.010038	0.011638	0.007803	-0.06446	0.020355	-0.01957
MATCH322	-0.99234	-0.01533	0.024922	0.024122	-0.04133	0.004842	-0.02261	0.003832	0.017147	-0.04527	-0.02253	-0.01996
VISIT323	-0.96842	-0.01461	0.014064	0.010649	-0.06674	0.003625	-0.02903	-0.0393	0.02676	-0.04291	-0.04998	-0.00281
GOOUT324	-0.95727	-0.01723	0.014596	0.013196	-0.08022	0.002741	0.018128	0.012829	0.003451	-0.05432	0.005312	-0.00816
SCISS325	-0.99082	-0.0151	0.018554	0.017643	-0.03575	0.004816	-0.02732	0.006093	0.007361	-0.04581	-0.02751	-0.01991
MONEY32	-0.94934	-0.01632	0.016305	0.01307	-0.07501	0.004468	0.010861	0.017438	0.003055	-0.0533	0.010798	-0.02257
OUTSD327	-0.90038	-0.02028	0.021721	-0.01049	-0.0805	0.001858	-0.03863	-0.03326	0.001316	-0.04616	-0.07694	-0.00703
DISCH328	-0.93529	-0.02265	0.005661	-0.00133	-0.09794	0.003222	0.000586	-0.03248	0.011357	-0.05364	-0.0291	0.000647
LOCK329	-0.83124	-0.05337	0.019812	-0.02626	0.041729	0.006774	-0.12185	-0.04314	-0.05148	-0.01033	-0.125	-0.04615
HOUSE331	-0.94681	-0.03398	0.02002	0.017607	-0.03369	-0.01659	0.001678	0.007705	0.004452	-0.0758	0.026149	-0.03238
OUTWD332	-0.84675	-0.02285	0.034589	-0.00468	0.062576	-0.00992	-0.13072	-0.00174	-0.07138	0.014322	-0.08425	-0.04822
OT333	-0.96624	-0.05472	0.022965	0.026163	-0.03809	-0.015	-0.05841	-0.01041	-0.01554	-9.6E-05	-0.05071	0.004667
SOCTH334	-0.82173	0.021776	0.034353	0.075573	0.057197	-0.01038	-0.17305	-0.00076	-0.00838	0.087567	-0.14235	0.037958
BEVIS341	-0.92008	-0.02434	0.016484	0.022684	-0.04599	-0.01159	-0.05725	-0.04183	0.03299	-0.06923	-0.06344	-0.0316
HOME342	-0.9062	-0.02815	-0.00795	0.025947	-0.13151	0.000292	0.042559	-0.02926	0.046507	-0.0667	-0.04395	0.003105
PAROL343	-0.7334	-0.04712	0.00143	0.029593	-0.02311	-0.02069	-0.05815	-0.03498	-0.00939	-0.04715	-0.09452	-0.02169
SUPER344	-0.96068	-0.02164	0.033164	0.017328	-0.01793	0.006872	-0.0001	0.015336	0.020583	-0.05738	-0.00641	-0.03352
SKILL411	0.019486	0.84434	-0.01685	0.013839	0.039996	-0.03512	0.049759	-0.00769	-0.05836	0.207462	0.056774	0.189885
INTER412	0.015153	0.86748	-0.02364	0.076493	0.051692	-0.0147	0.032354	-0.04266	0.000402	0.104753	0.023537	0.119246
ASSET413	0.006191	0.843995	-0.04586	0.002943	0.038456	-0.03016	0.045478	-0.04054	-0.05696	0.162789	0.022461	0.070546
FACT414	-0.01133	0.864564	0.001461	-0.02261	0.113547	0.009215	-0.00042	-0.02472	-0.0548	0.0789	0.084232	0.018677
CONFD415	-0.0126	0.6584	-0.10574	0.152476	0.179673	0.040342	0.093868	0.005254	0.111797	0.020856	0.057812	0.017043
ENVIR421	-0.0009	0.761083	-0.04715	-0.10355	0.140208	0.009104	0.047755	0.068398	-0.04721	0.258472	0.053322	0.009343
HDCAP422	0.021757	0.822406	-0.01427	-0.05833	0.143507	0.020105	0.011661	0.065982	-0.06685	0.236723	0.010691	0.097263
LIAB423	0.050661	0.724829	-0.01154	0.003572	0.203514	0.021557	-0.02899	0.031053	0.028576	0.231244	0.167819	0.043758
KEYFG431	0.047152	0.470095	0.316928	-0.11835	0.235107	-0.03523	-0.01583	-0.02715	-0.04614	0.032184	0.088569	0.041575

HOURS432	0.035657	0.157787	-0.11836	-0.28152	0.646303	-0.02008	-0.08261	-0.00506	-0.05171	0.026054	0.094173	0.108973
EE433	0.123049	0.227929	-0.01164	0.030642	0.856231	-0.0697	0.06749	0.023799	-0.0347	0.016966	-0.05203	0.036967
KCONT434	0.127947	0.203556	-0.07126	0.047926	0.853487	-0.06915	0.048925	0.002681	-0.02999	0.049556	0.028728	0.076264
REJ435	0.132635	0.20264	-0.03425	0.013206	0.825919	0.020486	0.093389	0.027801	-0.06997	0.118068	-0.00363	0.007619
PRIV43	0.159732	0.304966	-0.01394	-0.0188	0.639595	0.07099	0.197499	0.035739	0.004412	0.171623	0.214559	0.005812
SEEK441	0.004318	0.015041	0.013006	-0.06003	0.004993	-0.94444	0.034541	-0.02302	0.042768	0.024288	-0.01079	-0.04188
RECE442	-0.00413	-0.01666	0.005894	0.001528	0.071729	-0.94846	-0.03087	0.000689	0.013367	0.029563	0.022719	-0.0222
JOB443	0.044257	0.557294	0.056875	-0.37279	0.098395	-0.00328	0.084763	0.043884	-0.23618	0.193392	-0.13909	0.006997
GLOB5	0.081144	0.466007	-0.1235	-0.3873	0.128727	-0.00725	0.114792	0.069574	0.056664	0.395396	0.076403	0.051378
Explained variance	26.53763	7.727984	5.720438	2.311087	3.564432	1.846489	1.565515	1.735516	1.907006	4.850553	1.721423	1.9949
Proportion of total	0.353835	0.10304	0.076273	0.030814	0.047526	0.02462	0.020874	0.02314	0.025427	0.064674	0.022952	0.026599

Table 3

Factor Loadings (Varimax Normalised) dasfu2

	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6	Factor 7	Factor 8	Factor 9	Factor 10	Factor 11	Factor 12	Factor 13
CARE11	-0.00828	0.121355	0.064108	0.281758	0.013176	-0.04996	0.030463	0.106937	0.777974	0.040052	-0.01436	0.032855	0.130338
UNACT12	-0.05393	0.082066	0.133706	0.186552	-0.01369	-0.14116	0.052485	0.115235	0.794615	0.008036	0.004608	0.137664	0.033939
SLOW13	0.009943	0.160578	0.024895	0.262881	0.057308	-0.03197	-0.19479	0.201142	0.689028	0.035703	-0.05374	-0.10348	-0.16416
SWITH14	-0.00097	0.062313	0.087989	0.212704	0.03052	0.032994	0.094637	0.108434	0.809739	-0.02352	0.025329	0.061344	0.108117
HSHO21	-0.12647	0.007474	0.194127	0.08739	0.153924	0.063922	0.088937	0.468018	0.413611	0.059427	0.121912	0.143517	0.406025
AFFECT22	-0.04739	0.010791	0.797593	0.074565	0.097378	-0.05062	0.273649	0.043207	0.045395	-0.00418	-0.0233	0.072237	0.157405
SEX23	0.01686	0.04364	0.847852	0.02833	0.061476	0.01508	0.00757	0.103441	0.114582	-0.01335	0.025888	0.060123	-0.00616
CHILD24	-0.02408	-0.00378	0.814123	0.032638	0.013862	-0.01677	-0.04072	0.136626	0.129071	0.031196	0.034167	-0.06815	0.040879
HETSEX25	-0.02457	-0.00616	-0.68049	-0.06375	0.084754	0.195732	-0.19479	0.23109	0.13466	0.014231	0.009412	0.116712	-0.05475
FRICT26	-0.14321	0.066219	0.044275	-0.02222	0.031478	-0.08352	0.178243	0.718804	0.183602	0.036946	0.011022	0.023918	0.06221
PERF27	0.000501	-0.00603	0.001003	0.103763	0.028807	-0.31267	0.153247	0.09067	0.279544	-0.00851	0.017819	0.753805	0.006902
JOB28	0.0174	-0.01344	-0.13367	-0.07402	0.005208	0.809445	0.102342	-0.01266	0.049859	0.003255	0.049958	-0.19873	0.026855
INFO29	-0.0192	0.016903	0.033306	0.070659	0.042259	0.011653	-0.01763	0.609976	0.509925	-0.00102	0.106719	0.083092	-0.11448
CRIS210	0.001302	-0.06518	0.030954	-0.07564	-0.01836	0.09494	0.875104	0.080933	0.111876	0.014822	0.035299	-0.00418	0.013652
MOV311	0.675811	0.106877	-0.02333	0.083711	0.013926	0.02687	0.179527	-0.24193	0.139946	-0.15909	0.066531	0.050166	0.363129
UNDAC312	0.623052	0.093216	-0.01797	0.074589	0.038783	-0.02847	0.157743	-0.18656	0.044197	-0.24085	0.059846	0.081291	0.397917
OVER313	0.957057	0.039922	-0.0099	-0.03515	-0.0067	0.013697	-0.02521	-0.01183	-0.02077	-0.08894	-0.02371	0.021258	-0.01008
CONVR314	0.958715	-0.07913	-0.01026	-0.02679	-0.03587	0.007586	-0.03278	0.007556	-0.00828	-0.14748	-0.05783	-0.00109	0.0458
SWITH315	0.968328	-0.05719	-0.00942	-0.02314	-0.05791	0.014198	-0.02934	0.032909	-0.01021	-0.0929	-0.05749	0.008475	0.035457
LEIS316	0.882669	0.021162	-0.01428	-0.03562	-0.00539	0.008242	-0.02906	-0.02082	-0.01507	-0.13584	0.026605	0.018007	-0.03688
IRREL317	0.959029	0.047392	-0.01284	-0.03337	-0.00767	0.013613	-0.02255	-0.01036	-0.01922	-0.0834	-0.01715	0.022632	-0.02118
POSTU318	0.962927	0.029068	-0.01414	-0.02879	-0.01523	0.010107	-0.02572	-0.0238	-0.01413	-0.08648	-0.02796	0.016199	-0.01822
THRT319	0.972993	-0.05751	-0.01014	-0.02713	-0.04206	0.007067	-0.0297	0.014213	-0.01397	-0.10677	-0.05688	0.005755	0.02976

BED3110	0.957065	-0.0917	-0.00441	-0.02921	-0.06454	0.01593	-0.03033	0.041539	-0.01019	-0.10305	-0.07022	0.005075	0.055041
APP3111	0.96923	-0.07863	-0.00543	-0.02448	-0.05593	0.004796	-0.0272	0.018423	-0.01171	-0.08452	-0.06844	-0.00096	0.047794
MEAL3112	0.962256	-0.07877	-0.01072	-0.02314	-0.03385	0.004517	-0.03184	-0.00641	-0.01155	-0.13495	-0.06053	-0.00562	0.041654
WORK321	0.961803	-0.07395	-0.00153	-0.01071	-0.02995	-0.00147	-0.00965	-0.02098	-0.01449	-0.03537	-0.02629	-0.02472	0.047088
MATCH322	0.963696	-0.08609	0.002314	-0.01397	-0.04559	0.004644	-0.00639	0.020203	-0.01666	0.014576	-0.04122	-0.01439	0.06012
VISIT323	0.941198	0.0304	-0.00401	0.009392	-0.026	-0.01047	0.01356	-0.08006	-0.01336	0.170384	0.003501	-0.024	-0.07228
GOOUT324	0.793413	0.09864	-0.00666	0.019712	-0.02423	-0.01903	0.024279	-0.15577	0.005109	0.119485	-0.00571	-0.03859	-0.1217
SCISS325	0.926865	0.069187	-0.00353	-0.02314	-0.03577	0.009397	0.001299	-0.01673	-0.02518	0.005889	0.013478	-0.00284	-0.01007
MONEY326	0.882148	0.158535	-0.00646	0.00183	-0.01786	-0.01603	0.0293	-0.07055	-0.02904	0.119459	0.041727	-0.02495	-0.10731
OUTSD327	0.819547	0.109032	-0.00367	0.054146	-0.01461	-0.00191	0.044496	-0.06725	0.030508	0.219368	-0.05176	-0.03717	-0.11078
DISCH328	0.859515	0.117092	-0.00685	0.013741	0.00909	-0.01579	0.034258	-0.12606	-0.00316	0.166482	0.033118	-0.03304	-0.14812
LOCK329	0.781235	0.08375	0.021319	-0.05951	-0.00131	0.064017	0.00697	0.131023	-0.04966	0.279758	0.056832	0.0789	0.052833
HOUSE331	0.961762	0.033744	-0.00242	-0.01103	-0.03676	0.014122	0.003599	-0.01708	-0.02051	0.08281	0.000945	-0.00197	-0.01121
OUTWD332	0.663151	-0.02162	0.01047	0.036053	0.013848	0.024897	0.015605	0.049768	-0.00189	0.636182	0.027139	0.052181	0.001148
OT333	0.782141	-0.01249	0.019662	0.00632	0.011254	0.009277	0.026865	-0.00751	-0.01423	0.477807	0.014534	0.018229	-0.04162
SOCTH334	0.755307	-0.0088	0.026624	-0.00577	0.017289	0.005203	0.002839	-0.02825	-0.03589	0.484346	0.083329	0.028255	0.010279
BEVIS341	0.89164	0.02452	-0.01915	-0.0119	-0.06887	-0.03057	0.020703	0.003616	-0.02405	0.030607	-0.05921	-0.01905	0.003526
HOME342	0.850141	-0.10788	-0.01485	0.019119	-0.05431	-0.0465	0.009125	-0.03389	-0.02018	0.18729	-0.01672	-0.02577	-0.01665
PAROL343	0.762908	-0.12037	0.023043	-0.01368	-0.05984	-0.00209	-0.03086	0.014518	-0.00817	0.274316	-0.05141	-0.00831	0.092308
SUPER344	0.886244	-0.06191	0.004028	-0.01586	-0.00153	-0.00207	-0.01718	-0.02584	-0.01616	-0.1201	0.023041	-0.02498	0.040613
SKILL411	0.0264	0.763845	-0.01993	0.096095	0.052401	-0.06402	-0.02163	0.091022	0.058608	0.04817	0.032031	-0.08888	0.106746
INTER412	0.008926	0.703183	0.018089	0.056021	0.10294	-0.0449	-0.02948	0.066609	0.000629	0.152053	-0.02305	-0.08626	0.142515
ASSET413	-0.00722	0.770258	-0.03944	0.041125	0.024049	-0.06208	-0.01568	-0.0075	0.002526	0.001619	0.033462	0.015065	0.125702
FACT414	0.032495	0.72187	-0.04248	0.020048	0.036474	0.0071	-0.0425	-0.10346	0.008858	0.020015	0.094815	-0.04235	0.136152
CONFD415	-0.03793	0.53823	0.04976	-0.00703	0.149908	0.102115	0.033944	-0.01923	0.058894	-0.06286	-0.02217	0.063155	-0.09389
ENVIR421	0.01041	0.64618	0.041936	0.044608	0.080908	0.066835	-0.00477	0.005561	0.13654	-0.05171	0.072497	0.174624	-0.22735
HDCAP422	0.00498	0.766198	0.019065	0.000933	0.053252	0.010462	-0.04506	0.051319	0.033547	-0.04473	0.087042	0.045151	-0.12857
LIAB423	0.010307	0.592232	0.044967	0.027887	0.085684	0.000941	0.031127	0.159974	0.132613	-0.13646	0.266898	0.069874	-0.35621
KEYFG431	-0.06093	0.160525	-0.26621	0.017138	0.333231	0.107347	0.054876	0.071276	-0.04186	-0.0511	0.240278	-0.02186	-0.01208
HOURS432	-0.07772	0.089654	0.031922	0.114768	0.684444	-0.07026	-0.19629	0.014951	0.07314	0.035737	-0.0498	0.152338	0.077832

EE433	-0.0623	0.118885	0.027868	0.053116	0.886266	0.02089	-0.01054	0.008735	-0.01268	0.030695	0.018156	0.032993	0.03466
KCONT434	-0.05925	0.083419	0.067153	-0.01968	0.853608	-0.02513	0.071912	-0.01593	0.002622	0.014548	0.096775	-0.06716	0.029004
REJ435	-0.01266	0.0711	0.056928	0.007216	0.825949	0.028825	0.021175	0.067497	0.069222	0.001618	0.129868	0.034509	-0.06164
PRIV436	-0.10263	0.152637	0.09399	0.007622	0.628294	0.090024	0.032611	0.109104	0.063212	-0.07354	0.269119	-0.04015	-0.05466
SEEK441	-0.09018	0.201038	0.051102	0.043718	0.249041	0.004363	-0.00427	0.104955	0.047039	0.0138	0.894035	0.018203	0.020961
RECE442	-0.07855	0.186469	0.052849	0.057623	0.249846	0.004599	0.007019	0.119314	0.060552	0.035879	0.894946	0.003528	0.011923
JOB443	0.03242	0.129644	-0.03382	-0.00993	0.03142	0.310355	-0.093	0.168874	0.005076	0.150581	0.059507	0.535748	0.142574
GLOB5	-0.08334	-0.00537	0.230667	0.091517	0.15301	0.01465	0.064449	0.033193	0.593048	-0.16656	0.204135	0.298847	-0.23234
Explained variance	22.54541	4.283153	5.359157	3.581975	3.53491	1.888435	1.966411	2.696753	3.69875	1.557993	2.064553	2.046757	1.180323
Proportion of total	0.300606	0.057109	0.071455	0.04776	0.047132	0.025179	0.026219	0.035957	0.049317	0.020773	0.027527	0.02729	0.015738

Appendix F

Eigenvalues per Factor for Each Dataset

Table 1

Eigenvalues dasiu

Factor	Eigenvalue	% total	Cumulative. eigenvalue	Cumulative percentage
1	7.988277	10.65104	7.988277	10.65104
2	6.609998	8.813331	14.59828	19.46437
3	2.875553	3.834071	17.47383	23.29844
4	2.477555	3.303407	19.95138	26.60185
5	2.330208	3.106944	22.28159	29.70879
6	2.149259	2.865679	24.43085	32.57447
7	2.081441	2.775255	26.51229	35.34972
8	1.959038	2.612051	28.47133	37.96177
9	1.922388	2.563185	30.39372	40.52496
10	1.773075	2.3641	32.16679	42.88906
11	1.637494	2.183326	33.80429	45.07239
12	1.591664	2.122218	35.39595	47.1946
13	1.563753	2.085004	36.95971	49.27961
14	1.453677	1.938237	38.41338	51.21784
15	1.401784	1.869045	39.81517	53.08689
16	1.382041	1.842721	41.19721	54.92961
17	1.329185	1.772247	42.52639	56.70186
18	1.268323	1.691097	43.79472	58.39295
19	1.220173	1.626897	45.01489	60.01985
20	1.169771	1.559694	46.18466	61.57954
21	1.128461	1.504614	47.31312	63.08416
22	1.080572	1.440763	48.39369	64.52492
23	1.075027	1.433369	49.46872	65.95829
24	1.028564	1.371419	50.49728	67.32971
25	1.013503	1.351338	51.51079	68.68105

Table 2

Eigenvalues dasful

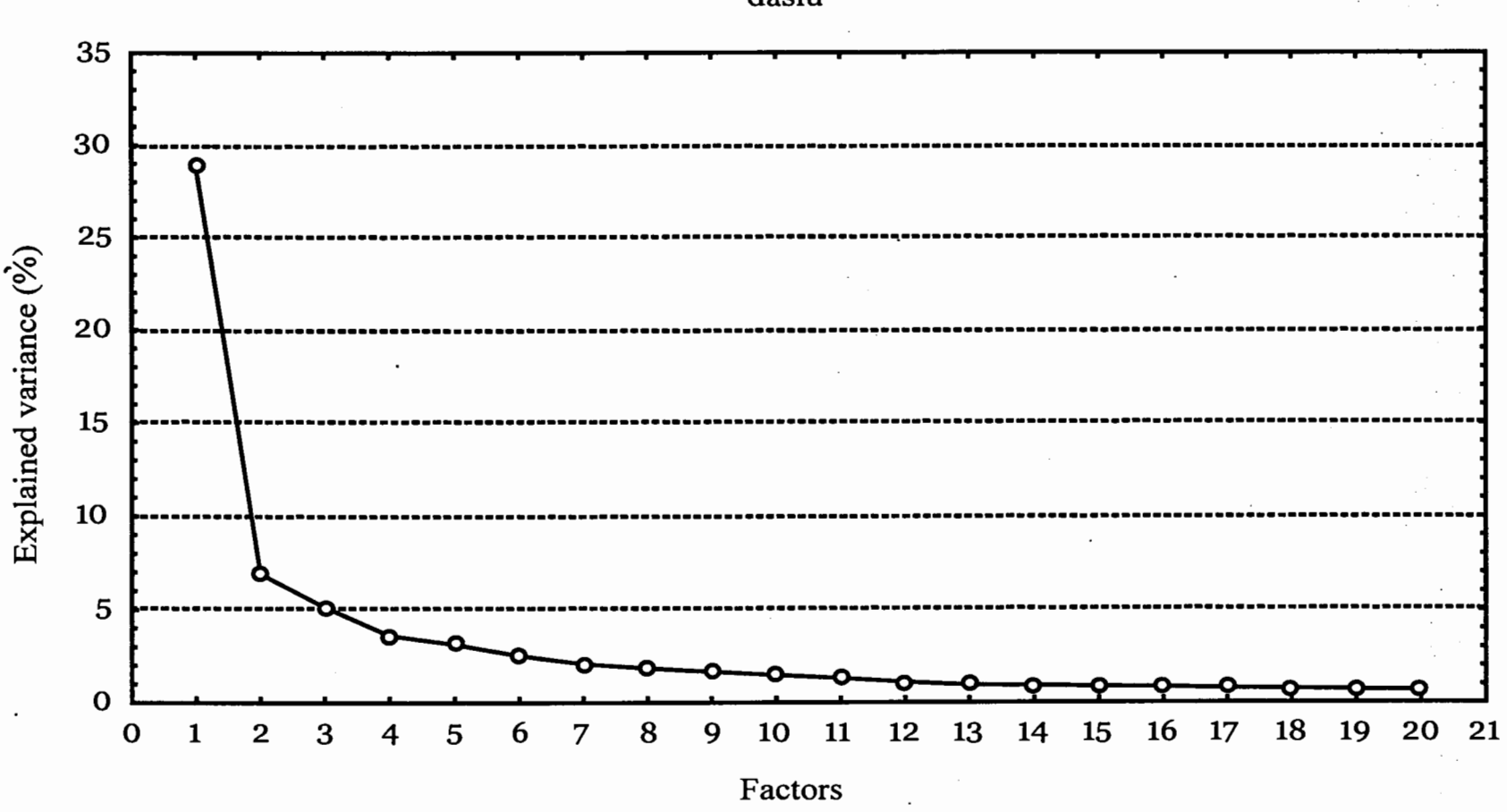
Factor	Eigenvalue	% total	Cumulative. eigenvalue	Cumulative percentage
1	8.712471	11.77361	8.712471	11.77361
2	6.437498	8.699321	15.14997	20.47293
3	4.858051	6.564934	20.00802	27.03786
4	3.256383	4.400518	23.2644	31.43838
5	3.019002	4.079732	26.2834	35.51811
6	2.836431	3.833015	29.11984	39.35113
7	2.665805	3.602439	31.78564	42.95357
8	2.211856	2.988995	33.9975	45.94256
9	2.107262	2.847652	36.10476	48.79022
10	1.969038	2.660862	38.0738	51.45108
11	1.717341	2.320731	39.79114	53.77181
12	1.638982	2.214841	41.43012	55.98665
13	1.534634	2.073829	42.96475	58.06048
14	1.481683	2.002274	44.44644	60.06275
15	1.396962	1.887787	45.8434	61.95054
16	1.342101	1.813649	47.1855	63.76419
17	1.297748	1.753714	48.48325	65.5179
18	1.199506	1.620954	49.68275	67.13886
19	1.181564	1.596708	50.86432	68.73556
20	1.163884	1.572816	52.0282	70.30838
21	1.048064	1.416303	53.07627	71.72468

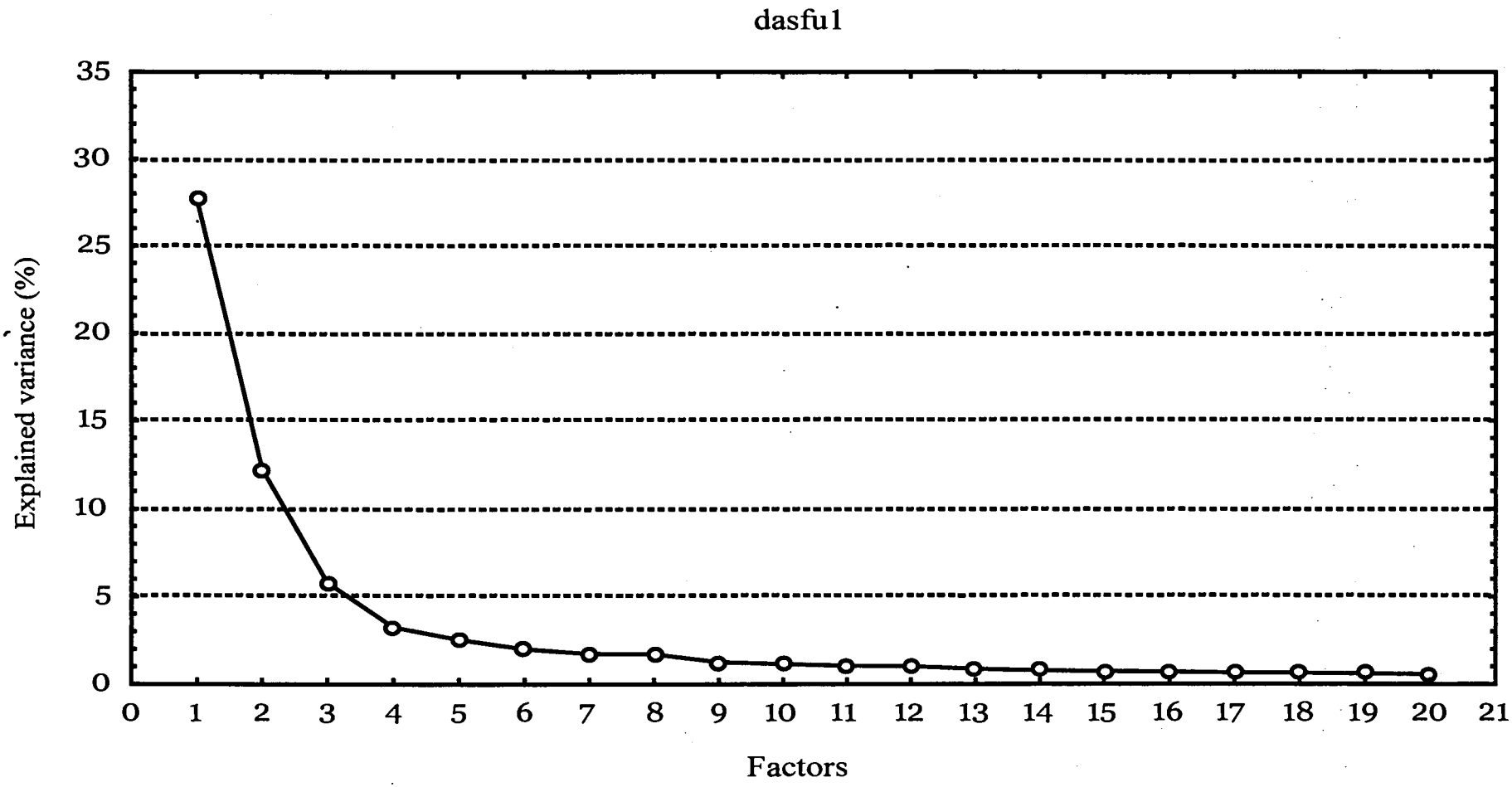
Table 3

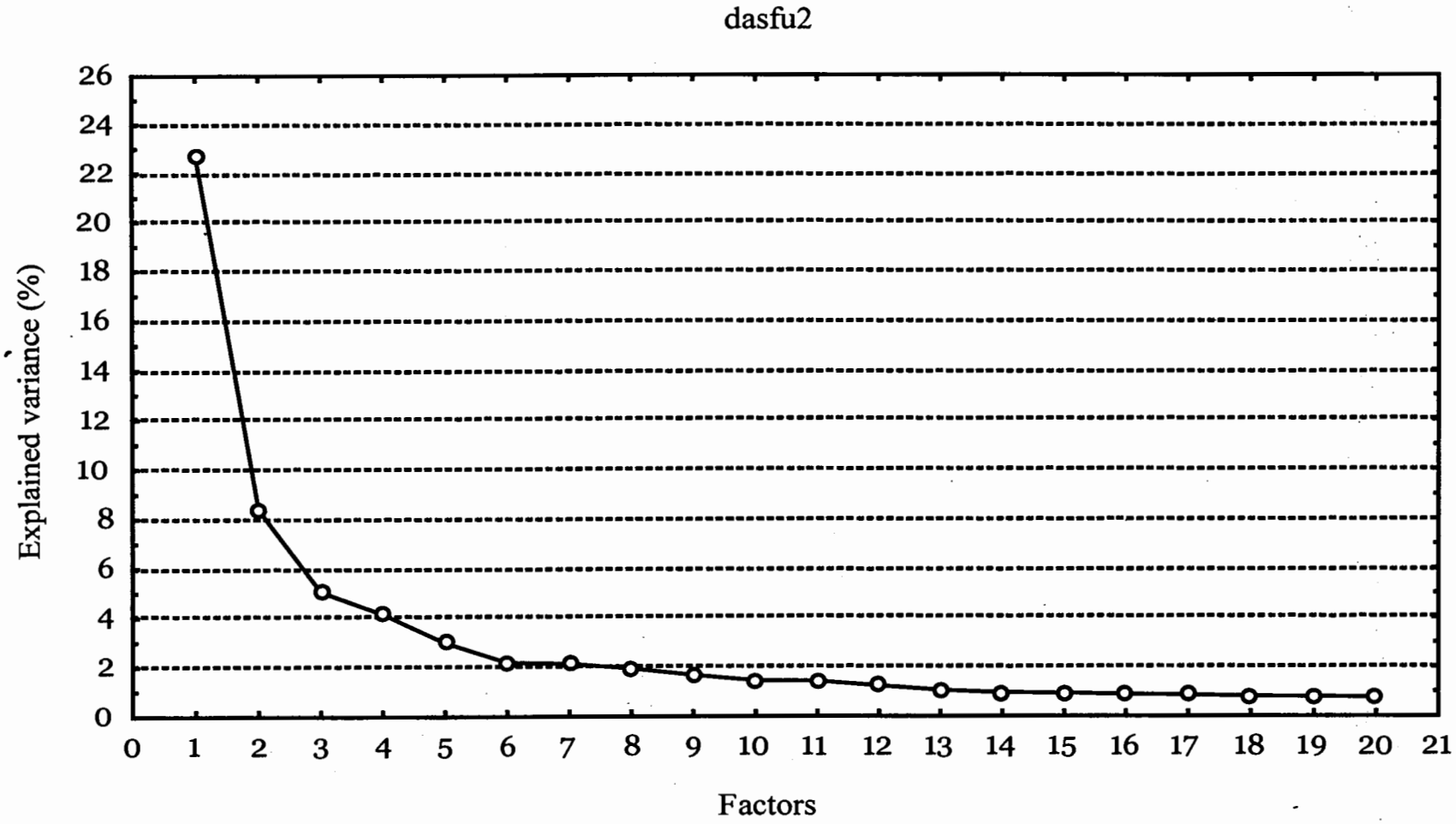
Eigenvalues dasfu2

Factor	Eigenvalue	% total	Cumulative. eigenvalue	Cumulative percentage
1	7.951049	10.6014	7.951049	10.6014
2	7.188721	9.584961	15.13977	20.18636
3	2.798664	3.731551	17.93843	23.91791
4	2.744993	3.65999	20.68343	27.5779
5	2.523464	3.364619	23.20689	30.94252
6	2.414356	3.219141	25.62125	34.16166
7	2.03974	2.719653	27.66099	36.88132
8	1.888407	2.517876	29.54939	39.39919
9	1.759697	2.346263	31.30909	41.74545
10	1.663897	2.218529	32.97299	43.96398
11	1.635643	2.180857	34.60863	46.14484
12	1.606608	2.142144	36.21524	48.28699
13	1.483496	1.977994	37.69873	50.26498
14	1.365272	1.820363	39.06401	52.08534
15	1.308505	1.744674	40.37251	53.83002
16	1.253944	1.671925	41.62646	55.50194
17	1.234358	1.645811	42.86081	57.14775
18	1.217667	1.623556	44.07848	58.77131
19	1.157328	1.543104	45.23581	60.31441
20	1.101898	1.469198	46.33771	61.78361
21	1.083341	1.444455	47.42105	63.22807
22	1.077913	1.437218	48.49896	64.66528
23	1.043609	1.391479	49.54257	66.05676

Appendix G: Scree Plots of the Three Datasets




dasful
Explained variance (%)
35
30
25
20
15
10
5
0
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21
Factors


dasfu2
Explained variance (%)
26
24
22
20
18
16
14
12
10
8
6
4
2
0
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21
Factors

Appendix H

WHO-DAS Variables and Descriptors

Variable	Descriptor
CARE11	Patient self-care during past month
UNACT12	Underactivity during past month
SLOW13	Slowness during past month
SWITH14	Social withdrawal during past month
HSHO21	Household activities during past month
AFFECT22	Affective relationship during past month
SEX23	Sexual relationship during past month
CHILD24	Interest and child-care
HETSEX25	Heterosexual relationships
FRICT26	Friction in interpersonal relationships
PERF27	Work performance
JOB28	Interest in getting job or back to work
INFO29	Interest and information
CRIS210	Behaviour in emergencies or crises
MOV311	Slowness of movement – hospital behaviour
UNDAC312	Underactivity – hospital behaviour
OVER313	Overactivity – hospital behaviour
CONVR314	Conversation – hospital behaviour
SWITH315	Social withdrawal – hospital behaviour
LEIS316	Leisure interests – hospital behaviour
IRREL317	Irrelevant or incomprehensible talk – hospital behaviour
POSTU318	Posturing and mannerisms – hospital behaviour
THRT319	Threatening or violent behaviour – hospital behaviour
BED3110	Tendency to remain in or return to bed – hospital behaviour
APP3111	Personal appearance – hospital behaviour
MEAL3112	Behaviour at meal time – hospital behaviour
WORK321	Work – nurses opinion
MATCH322	Possess matches – nurses opinion
VISIT323	Visit relatives or friends – nurses opinion
GOOUT324	Go out with opposite sex – nurses opinion
SCISS325	Scissors or razor blades – nurses opinion
MONEY326	Look after money – nurses opinion
OUTSD327	Work outside – nurses opinion
DISCH328	Discharge – nurses opinion
LOCK329	Need locked ward – nurses opinion
HOUSE331	Housekeeping
OUTWD332	Work outside ward
OT333	Occupational therapy
SOCTH334	Social therapy
BEVIS341	Being visited
HOME342	Visiting home

PAROL343	Parole
SUPER344	Supervision
SKILL411	Activity to improve skills
INTER412	Special interests
ASSET413	Above average assets
FACT414	Other favourable factors
CONFD415	Confiding relationship
ENVIR421	Environmental disadvantage
HDCAP422	Specific liabilities: handicap
LIAB423	Specific liabilities: other factors
KEYFG431	Home atmosphere: key figure
HOURS432	Hours face-to-face
EE433	Emotional involvement of key figure
KCONT434	Key figure: control and demand
REJ435	Home atmosphere: rejection
PRIV436	Home atmosphere: privacy
SEEK441	Outside support: seeking help
RECE442	Outside support: receiving help
JOB443	Outside support: job situation
GLOB5	Global evaluation

Appendix I

First Draft

DAS-Revised

Rating period

PS: Last four weeks ending [...../...../.....]

or

LB: from [...../...../.....] to [...../...../.....] Number of years [ ]

Informant: [ ]

Rater: [ ]

Instructions

The scale should always be completed using all available sources unless otherwise specified. Where there is obvious disagreement between sources separate coding sheets should be completed for each informant.

Ratings: In addition to the ratings **0-4** two additional ratings may be given for any item:

8: The rater is unsure whether the item is ratable even after adequate examination. It is not the rating for not applicable.

9: Not appropriate to make a rating because the examination (eg because of refusal) or available supplementary records are *incomplete* or *item not applicable.*

Rating period

Present State (**PS**): Rates the 4 weeks *prior to rating date.* All PS ratings are to be from the date the rating is made.

Lifetime Before (**LB**). Rates the period *preceding* the period rated for present state.

DAS-R

This instrument should only be completed by trained persons (nurses, psychiatrists, psychiatric social workers etc.).

Model: Psychiatric disability can be broken into two distinct domains. The first, called Personal Disability, is analogous to the Activities of Daily Living and is basically a self care dimension. The second domain, Role Disability, is a disruption in Social Role performance. Social Roles are characterised by the need for a reciprocal role relationship involving another individual who also has a social role.

NOTE: The range of behaviour incorporated within each of the items in Section A varies. The definition of each item has been set at a level that offers the best predictive power based on the data analysis used in the development of this instrument. As the instrument undergoes further development it is anticipated that this level of definition will increase. The specific nature of an item relates to its predictive and not descriptive power.

1.Self care

0. Not a problem
1. Some minimal loss in self care (e.g. dirty clothes,
2. Obviously lacking but not dangerously so (e.g. rotten teeth or malodorous or nails seriously overgrown)
3. Seriously lacking (e.g. lice in hair or never washes, unshaven, hair unkempt or dangerously malnourished)
4. Total lack of any self care activity – totally dependent on others

PS ☐ LB ☐

2. Purposeful Activity

0. Not a problem
1. Aimless or futile activity for hours at a time but whilst noteworthy not of concern to patient or others
2. Markedly reduced activity – does little beyond basic routine of eating, sleeping, dressing. Spends extended periods (hours) doing nothing
3. Very restricted activity – most of day spent inactive. No interests or daily involvement. Passively obeys direct instructions but if not supervised returns to inactive state.
4. Total lack of any daily activity. Does absolutely nothing – needs to be strongly coerced to carry out even the most basic activities

PS ☐ LB ☐

3. Slowness

0. Not a problem
1. Somewhat slow in movement/response – lacking normal vigour, apathetic
2. Slow such that this interferes with many daily routines and is/has been present much of the time
3. Markedly slow such that interferes with most daily routine and is present most of the time. Speed of action unresponsive to intervention.
4. Absence of movement for hours at a time – incompatible with anything but basic functions

PS ☐ LB ☐

4. Social Withdrawal

0. Not a problem
1. Taciturn and solitary in social situations but can be engaged in a limited manner
2. Narrow range of social contacts (only one or two friends) and even with these individuals only limited interaction takes place
3. Actively avoids social contact most of the time. No friends. Solitary lifestyle with only minimal contact with agency staff eg. Community nurses, DSS officers.
4. No social contact at all – totally unresponsive to any social interaction and avoids all social situations all of the time

PS ☐ LB ☐

5. Participation in family/communal activities

0. Not a problem
1. Participates in family/communal life but in a detached, aloof, unemotional manner
2. Inadequately participates – incompetent and/or irritating and/or deliberately disruptive
3. Participates but greatly disrupts family life because of intolerable behaviour (e.g. personal habits are found grossly offensive by others)
4. Totally alienated from all family/communal activities –totally non participatory

PS ☐ LB ☐

6. Affective relationship with spouse/partner

0. Not a problem
1. Somewhat remote and/or uncaring
2. Substantially poor relationship. General lack of communication. Emotionally cool
3. No emotional involvement with partner. Grossly inadequate participation in relationship.
4. Patient hostile and (potentially) very dangerous all of the time . Partner is at serious risk of violence. Partner is afraid of patient all of the time. Partner wishes separation or may have already left

PS ☐ LB ☐

7. Sexual relation with spouse/partner

0. Not a problem
1. Sexual activity has occurred but either the patient or partner find their sexual relationship unsatisfactory or lacking
2. Sexual interest and initiative have deviated (e.g.. increased or decreased) from a previously established pattern to a degree that the partner finds concerning
3. Avoidance of any intimacy cues or overt hostility at cues/initiation of sexual activity by the partner
4. Persistent and complete lack of sexual interest with spouse/partner

PS ☐ LB ☐

8. Conversation (restricted)

0. Not a problem
1. Reduced flow, content or rate but manages to converse with prompting
2. Minimal conversation but responsive to others. Does not initiate conversation
3. Markedly reduced conversation – only converses in response to persistent efforts, otherwise not responsive. Conversations generally limited to monosyllabic or unintelligible replies
4. Totally mute and unresponsive

PS ☐ LB ☐

9. Leisure interests

0. Not a problem
1. Participates in leisure activities but reports lack of interest and/or pleasure
2. Can be persuaded to watch TV, join in games for a while at least but tends towards passivity.
3. Almost total lack of interests but may watch some TV or listen to some radio with persuasion but without any apparent active interest. Has no leisure interests. Doesn't participate in structured activities
4. Totally uninterested in local or world events; never reads anything. Does not watch TV or listen to radio. Does not participate in any leisure activity

PS ☐ LB ☐

10. Hazardous Behaviour

0. Not a problem
1. Needs to be checked occasionally and reminded about the possibility of risks in handling inflammables
2. Unsupervised during daytime but at risk at night where inflammables need to be locked away (e.g.. Would smoke in bed, use lighter to find way about in the dark)
3. Can be trusted to cook or smoke or be in possession of matches but only under general supervision
4. Can never be given responsibility for matches or lighter without being a serious risk. Needs constant direct supervision smoking, cooking or any activity that is a fire risk

PS ☐ LB ☐

11. Need for supervision for security reasons

0. Not a problem
1. Can be in community with occasional contact with supervising authoroties (eg community nurse, probation officer)
2. Can be in community but needs a least daily contact and assessment
3. Needs a semi secure environment where there is 24 hour supervision
4. Direct supervision all of the time – needs to be in a locked facility because of serious risk to self or others

PS ☐ LB ☐

Appendix J

Second Draft

DAS-Revised - second draft

Subject	
Informant	

	PS	LB
Please tick (✓) as appropriate		

	Please complete for non-standard rating periods		If to present date leave blank	No of weeks
Rating period (dates)		*to*		

Date of rating	
Rater	

Instructions:

This instrument should only be completed by trained mental health professionals (nurses, psychiatrists, social workers etc)

Complete all sections using all available sources of information unless specified otherwise.

If there is irreconcilable disagreement between sources then separate coding booklets should be completed separating the data from the differing informants.

Ratings:

Ratings follow the general format of

0: Satisfactory or not present
1: Present but not problematic
2: Marked (obvious) problem
3: Severe problem
4: Maximal problem

In addition the following to ratings can be given for any item

8: The rater is unsure the item is ratable *even after adequate inquiry*. This is <u>not</u> the rating for not applicable.
9: Not appropriate to make a rating because of incomplete information (eg subject refuses to answer questions) *or* item is <u>not</u> <u>applicable</u>.

1:Self care

This item refers to basic, self initiated actions carried out in order to maintain and/or promote reasonable physical health. This would include washing/bathing, maintaining an adequate dietary intake, basic physical care such as nail-care, dental care, provision and selection of appropriate clean clothing.

Do not include oddities of appearance such as strange hairstyles, odd dress sense, un-ironed clothing or unusual dietary habits.

Only rate on the basis of those self care factors which potentially diminish or impede optimum physical health.

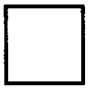
0. **Satisfactory self care.**
1. **Some minimal non-problematic loss in self care (e.g. dirty clothes or unkempt hair, unshaven).**
2. **Markedly lacking but not dangerously so (e.g. rotten teeth or chronically malodorous or nails seriously overgrown, or hair matted).**
3. **Seriously lacking (e.g. lice in hair or never washes, or nails seriously overgrown or dangerously malnourished).**
4. **Total lack of any self care activity – totally dependent on others.**

2: Purposeful activity

This item refers to the degree of logical, goal-directed activity. The rater should consider behaviour which is intentional, associated with planned independent actions and carried out to a reasonable degree. There will usually be an articulable or (potentially) observable end-point to the activity.

Rate solely on the basis of goal attainment, not effort.

0. **Satisfactory.**
1. **Minimal problem. Aimless or futile activity for hours at a time but whilst noteworthy, not of concern to the individual or others.**
2. **Markedly reduced activity – does little beyond basic routine of eating, sleeping, dressing. Spends extended periods (hours) doing nothing purposeful.**
3. **Severely restricted activity – most of day spent inactive. No interests or daily involvement. Passively obeys direct instructions but if not supervised returns to inactive state.**
4. **Total lack of any daily activity. Does absolutely nothing – needs to be strongly coerced to carry out even the most basic activities.**



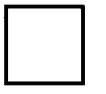
3: Objectively slowed movement

This item rates the overall impact of slowness associated with diminished motor activity *for any reason* which interferes with the individual's everyday life.

Do not rate solely on the basis of individual self report – it is important *not to* rate this item based on the individual's subjective experience of slowness.

Do not rate on the basis of intentionality – it does not matter if the interference is due to intentional slowness.

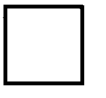
0. Not a problem
1. Minimally slow in movement/response – apathetic and/or lacking normal vigour and/or hesitant. Noticeable but not problematic.
2. Marked slowness such that this interferes with many daily routines and is/has been present much of the time.
3. Severely slow such that interferes with most daily routine and is present most of the time. Speed of action unresponsive to intervention.
4. Absence of movement for hours at a time – incompatible with anything but basic functions during this time.



4: Social activity

This item rates diminished social activity *for whatever reason.* It is not a rating of social competence or of the subjective experience of social aversion.

This item should only be rated on the *objective* degree of social engagement.



0. Satisfactory.
1. Minimal participation in social situations. Taciturn and solitary but can be engaged in a limited manner.
2. Marked problem. Narrow range of social contacts (only one or two acquaintances) and even with these individuals only limited interaction takes place.
3. Severely withdrawn. Actively avoids social contact most of the time. No friends. Solitary lifestyle with only minimal contact with agency staff eg. Community nurses, welfare officers.
4. No social contact at all – totally unresponsive to any social interaction and does not engage in any social activity.

5: Participation in family/communal activities

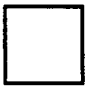
This item rates both the degree and quality of the individual's interaction with family/friends and defined community.

The defined community is those members of the individual's local community that the individual would be expected to interact with in normal daily activities, including persons in statutory positions such a law enforcement officers, elected officials as well as others in the locale such as shopkeepers.

Friends includes the individual's friends, family friends and any family partners (brother's wife for example).

Do not include the relationship between the individual and his/her treating team.

Do not rate family function/dysfunction, other than that which can be attributed to the behaviour or actions of the individual.



0. **Satisfactory.**
1. **Participates in family/communal life but in a detached, aloof, unemotional manner. Whilst noticeable is not a problem.**
2. **Markedly inadequate participation – incompetent and/or irritating and/or embarrassing.**
3. **Participation severely disrupts family life because of intolerable and/or deliberately disruptive behaviour(e.g. the individuals personal habits are found grossly offensive by others).**
4. **Totally alienated from all family/communal activities – totally non participatory.**

6: Affective relationship with spouse/partner

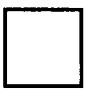
This is a rating of the individual's affective relationship with his/her primary life partner .

This item should be rated on the basis of *subjective* report (primarily by the individual but also where possible the partner).

Ratings should not be inferred from the individual's or partner's action alone.

Do not consider the opinions of those external to the individual/partner in making this rating.

It is important to emphasise that it is the individual's and not the partner's affective participation that is being rated.



0. **Satisfactory.**
1. **Somewhat remote and/or uncaringbut not to a degree that is problematic**
2. **Markedly poor relationship. General lack of communication. Emotionally ambivalent.**
3. **Little emotional involvement with partner or alternatively the relationship is characterised by severe emotional instability. Severely inadequate participation in relationship. Partner is afraid of individual some of the time.**
4. **No participation at all. Individual hostile and (potentially) very dangerous all of the time . Partner is at serious risk of violence. Partner is afraid of individual all of the time. Partner wishes separation or may have already left.**

7: Sexual relations with spouse/partner

In this item, "sexual relations" refers to whatever the individual and partner feel it is for them, and makes no attempt to externally define adequate sexual relations.

Rate on the basis of the adequacy of the sexual relationship.

Do not modify rating on the basis of inferred or understandable cause such as impotence induced by prescribed medications.

It is important that the adequacy of the relationship is rated as defined by both partners.

It cannot be rated in the absence of a the individual's and/or partner's report (score item 9).

0. **Satisfactory.**
1. **Sexual activity has occurred but either the individual or partner find their sexual relationship lacking but not to a problematic degree**
2. **Sexual interest and initiative have deviated (e.g.. increased or decreased) from a previously established pattern to a marked degree.**
3. **Severe problem. Avoidance of any intimacy cues or overt hostility at cues/initiation of sexual activity by the partner.**
4. **Persistent and complete lack of sexual interest with spouse/partner.**

8: Restricted verbal communication

This item rates any reduction in the individual's verbal communication (flow, content, rate) in familiar circumstances.

Do not rate the *quality* of the conversational content or poor *non-verbal* communication in this item.

0. Not present.
1. Minimally reduced flow, content or rate but responsive to others. Not problemtaic.
2. Markedly reduced conversation but manages to converse with prompting. Does not initiate conversation.
3. Severely reduced conversation – only converses in response to persistent efforts, otherwise not responsive. Conversations generally limited to monosyllabic or unintelligible replies.
4. Totally mute and unresponsive.

9: Leisure interests

This item rates the individual's active participation in non-obligatory, self-initiated, structured activity external to occupation. This may include hobbies, sports, as well as primarily intellectual pursuits. It is the self-initiated pattern of engagement that defines the rating of this item.

0. Satisfactory.
1. Participates in leisure activities but reports some lack of interest and/or pleasure. Not regarded as a problem.
2. Interests are markedly reduced but can be persuaded to watch TV, join in games for a while at least but tends towards passivity.
3. Severely reduced leisure activity. Almost total lack of interests but may watch some TV or listen to some radio with persuasion but without any apparent active interest. Does not participate in any structured leisure activities.
4. Totally uninterested in local or world events; never reads anything (if literate). Does not watch TV or listen to radio. Does not participate in any leisure activity.

10: Hazardous behaviour

(The very specific nature of the anchors for this item are a result of the data analysis performed in the development of this instrument and the highly predictive nature of these items. See NOTE on the coversheet

This item rates whether or not the individual can safely handle matches or similar items such as a cigarette lighter.

It does not matter whether the individual actually has access or use for such items.

The rating is made on the basis of the raters opinion given all available information.

The question raters should ask themselves is, "Should this person be allowed to have matches (or similar) and under what circumstances?".

0. **No supervision required.**
1. **Needs to be checked occasionally and reminded about the possibility of risks in handling inflammables. Not regarded as a problem.**
2. **Marked risk. Unsupervised during daytime but at risk at night where inflammables need to be locked away (e.g.. Would smoke in bed, use lighter to find way about in the dark).**
3. **Severe risk. Can be trusted to cook or smoke or be in possession of matches but only under general supervision.**
4. **Can never be given responsibility for matches or lighter without being a serious risk. Needs constant direct supervision smoking, cooking or any activity that is a fire risk.**

11: Need for supervision for security reasons

The rater may find this item difficult to rate.

It is a rating of the supervision judged necessary to minimise the risk of violence (self or externally directed) balanced against the use of the least restrictive alternative.

This item rates both environmental restriction considered necessary and frequency of supervisory contact. The item should be rated on the basis of the rater's judgment of the individual's supervisory needs and not the actual circumstance that the individual is in (should they differ).

0. **No supervision required.**
1. **Minimal risk. Can be in community with occasional (weekly or greater) contact with supervising authorities (eg community nurse, probation officer). Not problematic.**
2. **Marked risk to self or others. Can be in community but needs a least daily contact and assessment.**
3. **Needs a semi secure environment where there is 24 hour supervision because of severe risk.**
4. **Direct supervision all of the time – needs to be in a locked facility because of serious risk to self or others.**

Appendix K:CVI Ratings by Expert Panel Members

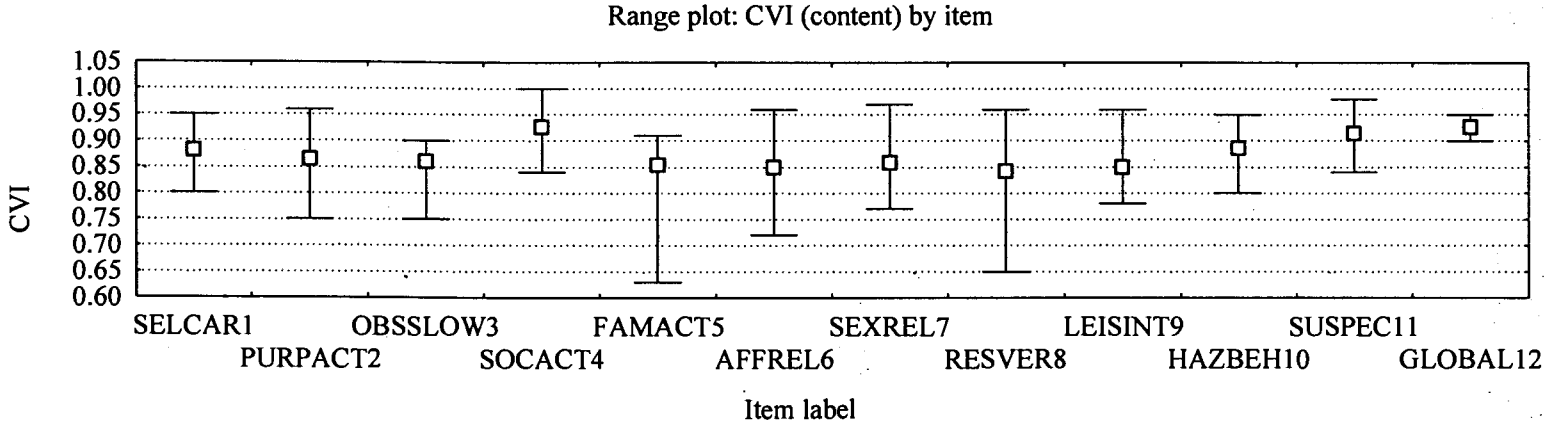
Expert	SELCAR1		PURPACT2		OBSSLOW3		SOCACT4		FAMACT5		AFFREL6	
	CONTENT	SCALING	CONTENT	SCALING	CONTENT	SCALING	CONTENT	SCALING	CONTENT	SCALING	CONTENT	SCALING
1	0.90	0.91	0.90	0.90	0.90	0.69	1.00	1.00	0.63	0.67	0.72	0.70
2	0.95	0.90	0.80	0.90	0.85	0.90	0.95	0.95	0.90	0.76	0.80	0.75
3	0.80	0.80	0.75	0.80	0.75	0.75	0.90	0.90	0.90	0.90	0.85	0.90
4	0.88	0.90	0.80	0.82	0.90	0.88	1.00	0.98	0.88	0.78	0.92	0.96
5	0.85	0.85	0.94	0.83	0.85	0.85	0.93	0.83	0.87	0.92	0.79	0.90
6	0.85	0.90	0.90	0.85	0.87	0.85	0.86	0.90	0.91	0.96	0.96	0.90
7	0.94	0.86	0.96	0.87	0.90	0.90	0.84	0.94	0.89	0.84	0.90	0.96
Mean	**0.881**	**0.874**	**0.864**	**0.853**	**0.860**	**0.831**	**0.926**	**0.929**	**0.854**	**0.833**	**0.849**	**0.867**

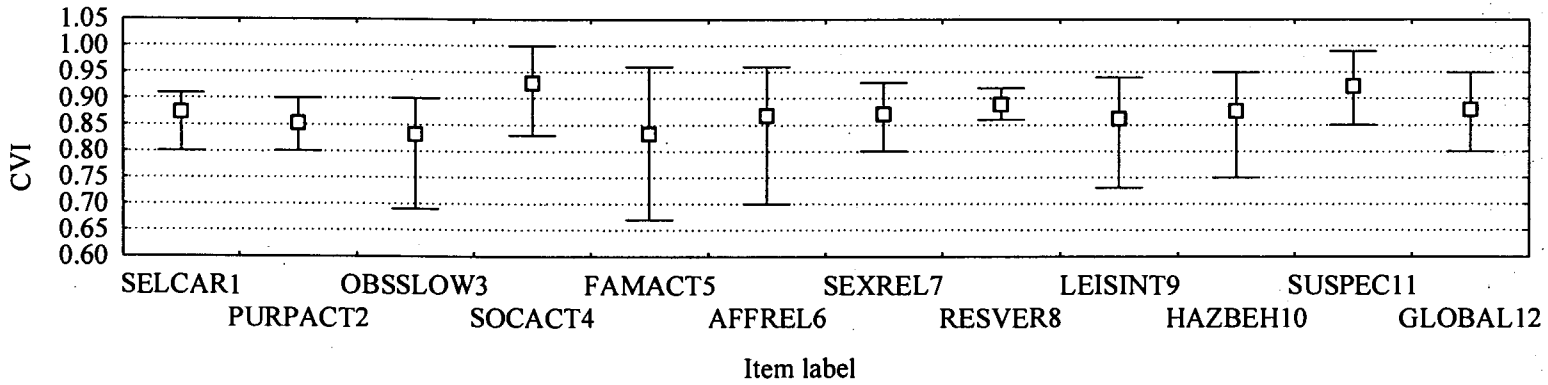
Expert	SEXREL7		RESVER8		LEISINT9		HAZBEH10		SUPSEC11		GLOBAL12	
	CONTENT	SCALING	CONTENT	SCALING	CONTENT	SCALING	CONTENT	SCALING	CONTENT	SCALING	CONTENT	SCALING
1	0.82	0.92	0.65	0.88	0.85	0.73	0.90	0.95	0.92	0.88	0.94	0.80
2	0.80	0.80	0.75	0.90	0.80	0.75	0.90	0.90	0.95	0.85	0.90	0.80
3	0.85	0.85	0.85	0.90	0.90	0.90	0.95	0.95	0.95	0.90	0.90	0.95
4	0.96	0.90	0.88	0.92	0.78	0.94	0.80	0.80	0.98	0.99	0.95	0.94
5	0.77	0.81	0.84	0.89	0.79	0.89	0.84	0.75	0.94	0.95	0.95	0.90
6	0.84	0.89	0.96	0.87	0.86	0.90	0.86	0.88	0.90	0.96	0.90	0.85
7	0.97	0.93	0.96	0.86	0.96	0.92	0.95	0.90	0.84	0.94	0.94	0.92
Mean	**0.859**	**0.871**	**0.841**	**0.889**	**0.849**	**0.861**	**0.886**	**0.876**	**0.914**	**0.924**	**0.926**	**0.880**

Factor	Descriptor
SELCAR1	Self care
PURPACT2	Purposeful activity
OBSSLOW3	Objectively slowed movement
SOCACT4	Social activity
FAMACT5	Participation in family/communal activities
AFFREL6	Affective relationship with partner/spouse

Factor	Descriptor
SEXREL7	Sexual relationship with spouse/partner
RESVER8	Restricted verbal communication
LEISINT9	Leisure interests
HAZBEH10	Hazardous behaviour
SUPSEC11	Need for supervision for security reasons
GLOBAL12	Global rating of CVI by experts

Appendix L: Rangeplot of CVI per Item





Appendix M

Draft 2 with CVI Scales Included

DAS-Revised

Subject	
Informant	

	PS	LB
Please tick (✓) as appropriate		

	Please complete for non-standard rating periods		If to present date leave blank	No of weeks
Rating period (dates)		*to*		

Date of rating	
Rater	

Instructions:

This instrument should only be completed by trained mental health professionals (nurses, psychiatrists, social workers etc)

Complete all sections using all available sources of information unless specified otherwise.

If there is irreconcilable disagreement between sources then separate coding booklets should be completed separating the data from the differing informants.

Ratings:

Ratings follow the general format of

0: Satisfactory or not present
1: Present but not problematic
2: Marked (obvious) problem
3: Severe problem
4: Maximal problem

In addition the following to ratings can be given for any item

8: The rater is unsure the item is ratable *even after adequate inquiry*. This is <u>not</u> the rating for not applicable.
9: Not appropriate to make a rating because of incomplete information (eg subject refuses to answer questions) *or* item is <u>not applicable</u>.

1:Self care

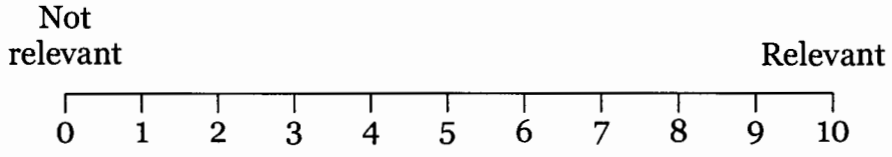
This item refers to basic, self initiated actions carried out in order to maintain and/or promote reasonable physical health. This would include washing/bathing, maintaining an adequate dietary intake, basic physical care such as nail-care, dental care, provision and selection of appropriate clean clothing.

Do not include oddities of appearance such as strange hairstyles, odd dress sense, un-ironed clothing or unusual dietary habits.

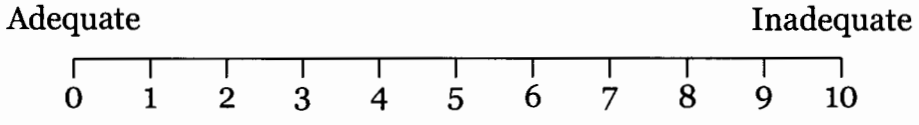
Only rate on the basis of those self care factors which potentially diminish or impede optimum physical health.

0. **Satisfactory self care.**
1. **Some minimal non-problematic loss in self care (e.g. dirty clothes or unkempt hair, unshaven).**
2. **Markedly lacking but not dangerously so (e.g. rotten teeth or chronically malodorous or nails seriously overgrown, or hair matted).**
3. **Seriously lacking (e.g. lice in hair or never washes, or nails seriously overgrown or dangerously malnourished).**
4. **Total lack of any self care activity – totally dependent on others.**

Content



Scaling



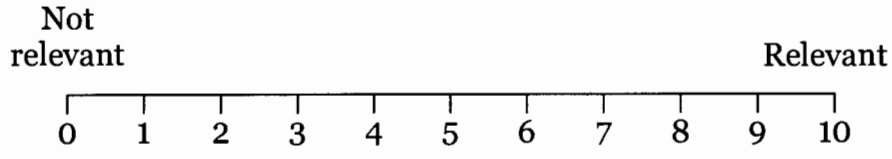
2: Purposeful activity

This item refers to the degree of logical, goal-directed activity. The rater should consider behaviour which is intentional, associated with planned independent actions and carried out to a reasonable degree. There will usually be an articulable or (potentially) observable end-point to the activity.

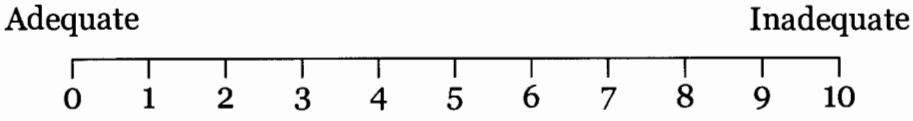
Rate solely on the basis of goal attainment, not effort.

0. **Satisfactory.**
1. **Minimal problem. Aimless or futile activity for hours at a time but whilst noteworthy, not of concern to the individual or others.**
2. **Markedly reduced activity – does little beyond basic routine of eating, sleeping, dressing. Spends extended periods (hours) doing nothing purposeful.**
3. **Severely restricted activity – most of day spent inactive. No interests or daily involvement. Passively obeys direct instructions but if not supervised returns to inactive state.**
4. **Total lack of any daily activity. Does absolutely nothing – needs to be strongly coerced to carry out even the most basic activities.**

Content



Scaling



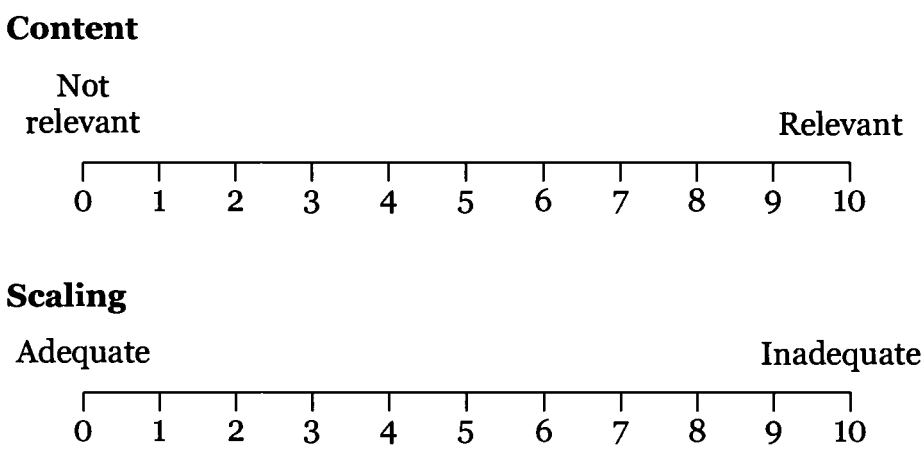
3: Objectively slowed movement

This item rates the overall impact of slowness associated with diminished motor activity *for any reason* which interferes with the individual's everyday life.

Do not rate solely on the basis of individual self report – it is important *not to* rate this item based on the individual's subjective experience of slowness.

Do not rate on the basis of intentionality – it does not matter if the interference is due to intentional slowness.

0. **Not a problem**
1. **Minimally slow in movement/response – apathetic and/or lacking normal vigour and/or hesitant. Noticeable but not problematic.**
2. **Marked slowness such that this interferes with many daily routines and is/has been present much of the time.**
3. **Severely slow such that interferes with most daily routine and is present most of the time. Speed of action unresponsive to intervention.**
4. **Absence of movement for hours at a time – incompatible with anything but basic functions during this time.**

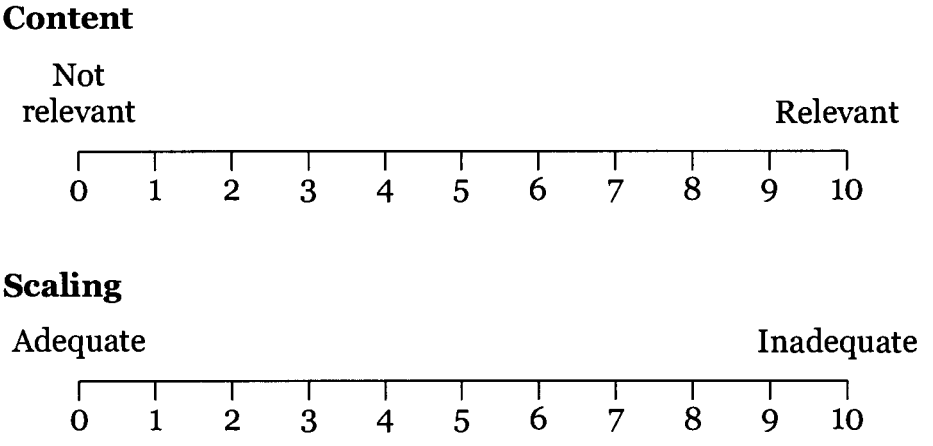


4: Social activity

This item rates diminished social activity *for whatever reason*. It is not a rating of social competence or of the subjective experience of social aversion.

This item should only be rated on the *objective* degree of social engagement.

0. **Satisfactory.**
1. **Minimal participation in social situations. Taciturn and solitary but can be engaged in a limited manner.**
2. **Marked problem. Narrow range of social contacts (only one or two acquaintances) and even with these individuals only limited interaction takes place.**
3. **Severely withdrawn. Actively avoids social contact most of the time. No friends. Solitary lifestyle with only minimal contact with agency staff eg. Community nurses, welfare officers.**
4. **No social contact at all – totally unresponsive to any social interaction and does not engage in any social activity.**



5: Participation in family/communal activities

This item rates both the degree and quality of the individual's interaction with family/friends and defined community.

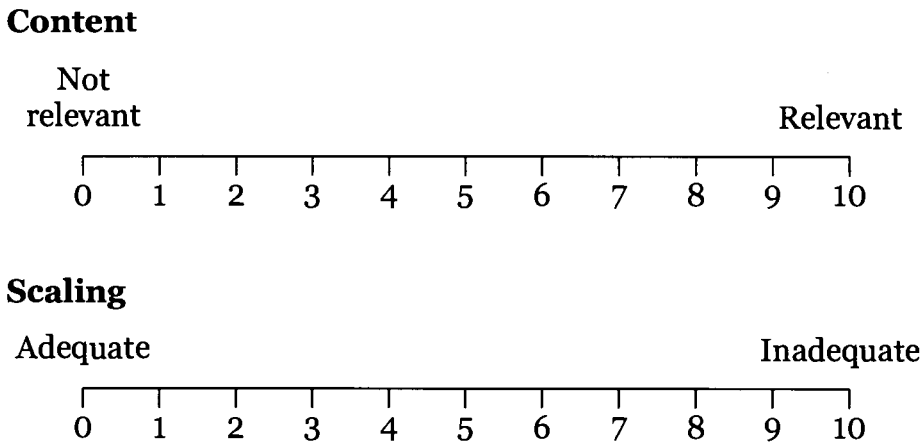
The defined community is those members of the individual's local community that the individual would be expected to interact with in normal daily activities, including persons in statutory positions such a law enforcement officers, elected officials as well as others in the locale such as shopkeepers.

Friends includes the individual's friends, family friends and any family partners (brother's wife for example).

Do not include the relationship between the individual and his/her treating team.

Do not rate family function/dysfunction, other than that which can be attributed to the behaviour or actions of the individual.

0. **Satisfactory.**
1. **Participates in family/communal life but in a detached, aloof, unemotional manner. Whilst noticeable is not a problem.**
2. **Markedly inadequate participation – incompetent and/or irritating and/or embarrassing.**
3. **Participation severely disrupts family life because of intolerable and/or deliberately disruptive behaviour(e.g. the individuals personal habits are found grossly offensive by others).**
4. **Totally alienated from all family/communal activities – totally non participatory.**



6: Affective relationship with spouse/partner

This is a rating of the individual's affective relationship with his/her primary life partner .

This item should be rated on the basis of *subjective* report (primarily by the individual but also where possible the partner).

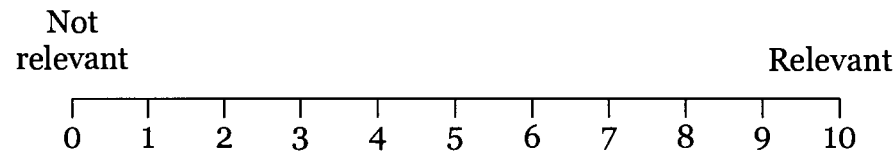
Ratings should not be inferred from the individual's or partner's action alone.

Do not consider the opinions of those external to the individual/partner in making this rating.

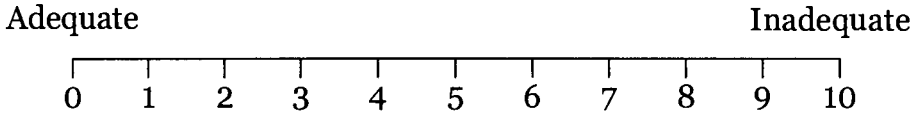
It is important to emphasise that it is the individual's and not the partner's affective participation that is being rated.

0. **Satisfactory.**
1. **Somewhat remote and/or uncaringbut not to a degree that is problematic**
2. **Markedly poor relationship. General lack of communication. Emotionally ambivalent.**
3. **Little emotional involvement with partner or alternatively the relationship is characterised by severe emotional instability. Severely inadequate participation in relationship. Partner is afraid of individual some of the time.**
4. **No participation at all. Individual hostile and (potentially) very dangerous all of the time . Partner is at serious risk of violence. Partner is afraid of individual all of the time. Partner wishes separation or may have already left.**

Content



Scaling



7: Sexual relations with spouse/partner

In this item, "sexual relations" refers to whatever the individual and partner feel it is for them, and makes no attempt to externally define adequate sexual relations.

Rate on the basis of the adequacy of the sexual relationship.

Do not modify rating on the basis of inferred or understandable cause such as impotence induced by prescribed medications.

It is important that the adequacy of the relationship is rated as defined by both partners.

It cannot be rated in the absence of a the individual's and/or partner's report (score item 9).

0. **Satisfactory.**
1. **Sexual activity has occurred but either the individual or partner find their sexual relationship lacking but not to a problematic degree**
2. **Sexual interest and initiative have deviated (e.g.. increased or decreased) from a previously established pattern to a marked degree.**
3. **Severe problem. Avoidance of any intimacy cues or overt hostility at cues/initiation of sexual activity by the partner.**
4. **Persistent and complete lack of sexual interest with spouse/partner.**

Content

Not relevant — Relevant

0 1 2 3 4 5 6 7 8 9 10

Scaling

Adequate — Inadequate

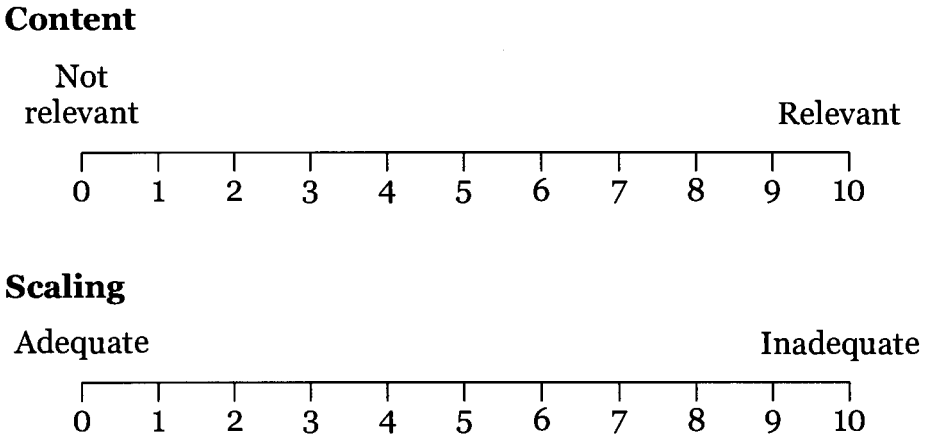
0 1 2 3 4 5 6 7 8 9 10

8: Restricted verbal communication

This item rates any reduction in the individual's verbal communication (flow, content, rate) in familiar circumstances.

Do not rate the *quality* of the conversational content or poor *non-verbal* communication in this item.

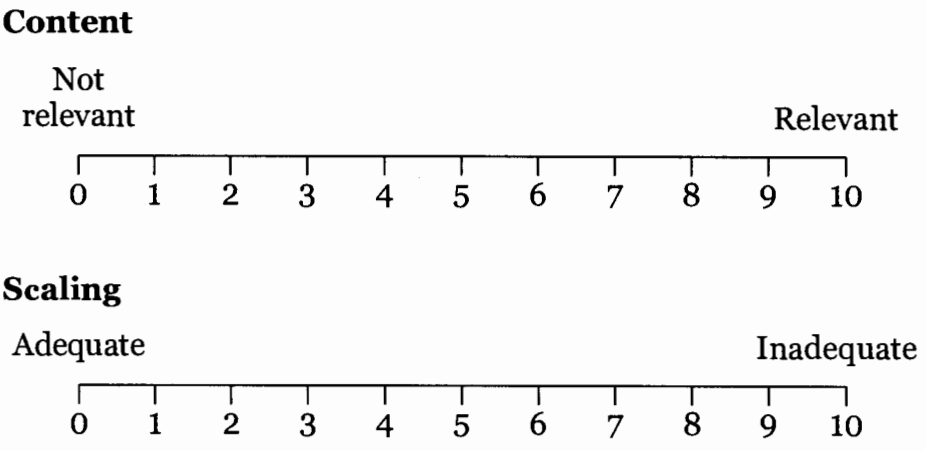
0. Not present.
1. Minimally reduced flow, content or rate but responsive to others. Not problemtaic.
2. Markedly reduced conversation but manages to converse with prompting. Does not initiate conversation.
3. Severely reduced conversation – only converses in response to persistent efforts, otherwise not responsive. Conversations generally limited to monosyllabic or unintelligible replies.
4. Totally mute and unresponsive.



9: Leisure interests

This item rates the individual's active participation in non-obligatory, self-initiated, structured activity external to occupation. This may include hobbies, sports, as well as primarily intellectual pursuits. It is the self-initiated pattern of engagement that defines the rating of this item.

0. **Satisfactory.**
1. **Participates in leisure activities but reports some lack of interest and/or pleasure. Not regarded as a problem.**
2. **Interests are markedly reduced but can be persuaded to watch TV, join in games for a while at least but tends towards passivity.**
3. **Severely reduced leisure activity. Almost total lack of interests but may watch some TV or listen to some radio with persuasion but without any apparent active interest. Does not participate in any structured leisure activities.**
4. **Totally uninterested in local or world events; never reads anything (if literate). Does not watch TV or listen to radio. Does not participate in any leisure activity.**



10: Hazardous behaviour

(The very specific nature of the anchors for this item are a result of the data analysis performed in the development of this instrument and the highly predictive nature of these items. See NOTE on the coversheet

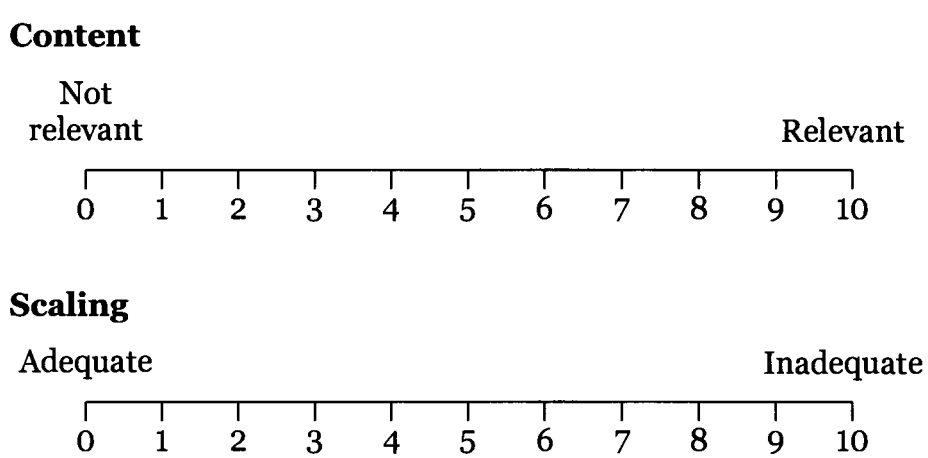
This item rates whether or not the individual can safely handle matches or similar items such as a cigarette lighter.

It does not matter whether the individual actually has access or use for such items.

The rating is made on the basis of the raters opinion given all available information.

The question raters should ask themselves is, "Should this person be allowed to have matches (or similar) and under what circumstances?".

0. **No supervision required.**
1. **Needs to be checked occasionally and reminded about the possibility of risks in handling inflammables. Not regarded as a problem.**
2. **Marked risk. Unsupervised during daytime but at risk at night where inflammables need to be locked away (e.g.. Would smoke in bed, use lighter to find way about in the dark).**
3. **Severe risk. Can be trusted to cook or smoke or be in possession of matches but only under general supervision.**
4. **Can never be given responsibility for matches or lighter without being a serious risk. Needs constant direct supervision smoking, cooking or any activity that is a fire risk.**



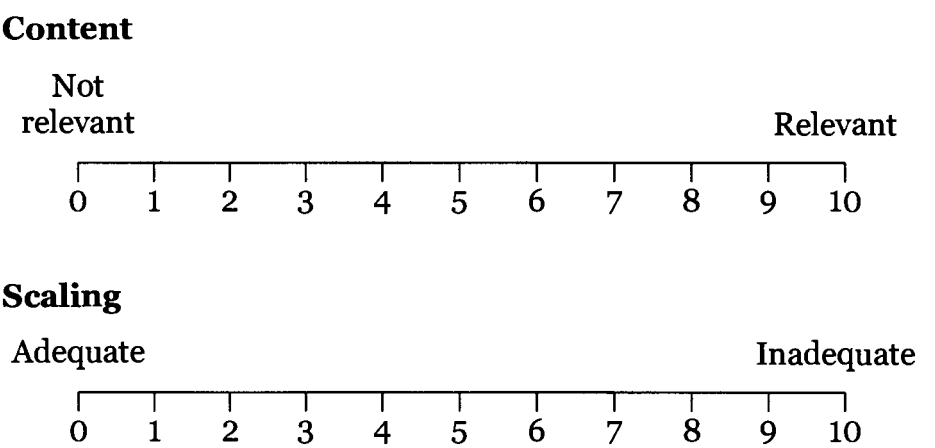
11: Need for supervision for security reasons

The rater may find this item difficult to rate.

It is a rating of the supervision judged necessary to minimise the risk of violence (self or externally directed) balanced against the use of the least restrictive alternative.

This item rates both environmental restriction considered necessary and frequency of supervisory contact. The item should be rated on the basis of the rater's judgment of the individual's supervisory needs and not the actual circumstance that the individual is in (should they differ).

0. **No supervision required.**
1. **Minimal risk. Can be in community with occasional (weekly or greater) contact with supervising authorities (eg community nurse, probation officer). Not problematic.**
2. **Marked risk to self or others. Can be in community but needs a least daily contact and assessment.**
3. **Needs a semi secure environment where there is 24 hour supervision because of severe risk.**
4. **Direct supervision all of the time – needs to be in a locked facility because of serious risk to self or others.**



Appendix N

Final Draft

DAS-Revised - second draft

Subject	
Informant	

	PS	LB
Please tick (✓) as appropriate		

	Please complete for non-standard rating periods		If to present date leave blank	No of weeks
Rating period (dates)		*to*		

Date of rating	
Rater	

Instructions:

This instrument should only be completed by trained mental health professionals (nurses, psychiatrists, social workers etc)

Complete all sections using all available sources of information unless specified otherwise.

If there is irreconcilable disagreement between sources then separate coding booklets should be completed separating the data from the differing informants.

Ratings:

Ratings follow the general format of

0: Satisfactory or not present
1: Present but not problematic
2: Mild (obvious) problem
3: Moderate problem
4: Severe problem

In addition the following to ratings can be given for any item

8: The rater is unsure the item is ratable *even after adequate inquiry*. This is <u>not</u> the rating for not applicable.
9: Not appropriate to make a rating because of incomplete information (eg subject refuses to answer questions) *or* item is <u>not applicable</u>.

1:Self care

This item refers to basic, self initiated actions carried out in order to maintain and/or promote reasonable physical health. This would include washing/bathing, maintaining an adequate dietary intake, basic physical care such as nail-care, dental care, provision and selection of appropriate clean clothing.

Do not include oddities of appearance such as strange hairstyles, odd dress sense, un-ironed clothing or special dietary habits.

Only rate on the basis of those self care factors which potentially diminish or impede optimum physical health.

0. **Satisfactory self care.**
1. **Some minimal non-problematic loss in self care (e.g. dirty clothes or unkempt hair, unshaven).**
2. **Markedly lacking but not dangerously so (e.g. rotten teeth or chronically malodorous or nails seriously overgrown, or hair matted).**
3. **Seriously lacking (e.g. lice in hair or never washes, or nails seriously overgrown or dangerously malnourished).**
4. **Total lack of any self care activity – totally dependent on others.**

2: Purposeful activity

This item refers to the degree of logical, goal-directed activity. The rater should consider behaviour which is intentional, associated with planned independent actions and carried out to a reasonable degree. There will usually be an articulable or (potentially) observable end-point to the activity.

Rate solely on the basis of goal attainment, not effort.

0. **Satisfactory.**
1. **Minimal problem. Aimless or futile activity for hours at a time but whilst noteworthy, not of concern to the individual or others.**
2. **Markedly reduced activity – does little beyond basic routine of eating, sleeping, dressing. Spends extended periods (hours) doing nothing purposeful.**
3. **Severely restricted activity – most of day spent inactive. No interests or daily involvement. Passively obeys direct instructions but if not supervised returns to inactive state.**
4. **Total lack of any daily activity. Does absolutely nothing – needs to be strongly coerced to carry out even the most basic activities.**

3: Objectively slowed movement

This item rates the overall impact of slowness associated with diminished motor activity *for any reason* which interferes with the individual's everyday life.

Do not rate solely on the basis of individual self report – it is important *not to* rate this item based on the individual's subjective experience of slowness.

Do not rate on the basis of intentionality – it does not matter if the interference is due to intentional slowness.

0. Not a problem
1. Minimally slow in movement/response – apathetic and/or lacking normal vigour and/or hesitant. Noticeable but not problematic.
2. Marked slowness such that this interferes with many daily routines and is/has been present much of the time.
3. Severely slow such that interferes with most daily routine and is present most of the time. Speed of action unresponsive to intervention.
4. Absence of movement for hours at a time – incompatible with anything but basic functions during this time.

4: Social activity

This item rates diminished social activity *for whatever reason.* It is not a rating of social competence or of the subjective experience of social aversion.

This item should only be rated on the *objective* degree of social engagement.

0. Satisfactory.
1. Minimal participation in social situations. Taciturn and solitary but can be engaged in a limited manner.
2. Marked problem. Narrow range of social contacts (only one or two acquaintances) and even with these individuals only limited interaction takes place.
3. Severely withdrawn. Actively avoids social contact most of the time. No friends. Solitary lifestyle with only minimal contact with agency staff eg. Community nurses, welfare officers etc.
4. No social contact at all – totally unresponsive to any social interaction and does not engage in any social activity.

5: Participation in family/communal activities

This item rates both the degree and quality of the individual's interaction with family/friends and defined community.

The defined community is those members of the individual's local community that the individual would be expected to interact with in normal daily activities, including persons in statutory positions such a law enforcement officers, elected officials as well as others in the locale such as shopkeepers.

Friends includes the individual's friends, family friends and any family partners (brother's wife for example).

Do not include the relationship between the individual and his/her treating team/clinicians.

Do not rate family function/dysfunction, other than that which can be attributed to the behaviour or actions of the individual.

0. **Satisfactory.**
1. **Participates in family/communal life but in a detached, aloof, unemotional manner. Whilst noticeable is not a problem.**
2. **Markedly inadequate participation – incompetent and/or irritating and/or embarrassing.**
3. **Participation severely disrupts family life because of intolerable and/or deliberately disruptive behaviour(e.g. the individuals personal habits are found grossly offensive by others).**
4. **Totally alienated from all family/communal activities – totally non participatory.**

6: Affective relationship with spouse/partner

This is a rating of the individual's affective relationship with his/her primary life partner .

This item should be rated on the basis of *subjective* report (primarily by the individual but also where possible the partner).

Ratings should not be inferred from the individual's or partner's action alone.

Do not consider the opinions of those external to the individual/partner in making this rating.

It is important to emphasise that it is the individual's and not the partner's affective participation that is being rated.

0. **Satisfactory.**
1. **Somewhat remote and/or uncaringbut not to a degree that is problematic**
2. **Markedly poor relationship. General lack of communication. Emotionally ambivalent.**
3. **Little emotional involvement with partner or alternatively the relationship is characterised by severe emotional instability. Severely inadequate participation in relationship. Partner is afraid of individual some of the time.**
4. **No participation at all. Individual hostile and (potentially) very dangerous all of the time . Partner is at serious risk of violence. Partner is afraid of individual all of the time. Partner wishes separation or may have already left.**

7: Sexual relations with spouse/partner

In this item, "sexual relations" refers to whatever the individual and partner feel it is for them, and makes no attempt to externally define adequate sexual relations.

Rate on the basis of the adequacy of the sexual relationship.

Do not modify rating on the basis of inferred or understandable cause such as impotence induced by prescribed medications.

It is important that the adequacy of the sexual relationship is rated as defined by both partners.

It cannot be rated in the absence of a the individual's and/or partner's report (score item 9).

0. **Satisfactory.**
1. **Sexual activity has occurred but either the individual or partner find their sexual relationship lacking but not to a problematic degree**
2. **Sexual interest and initiative have deviated (e.g.. increased or decreased) from a previously established pattern to a marked degree.**
3. **Severe problem. Avoidance of any intimacy cues or overt hostility at cues/initiation of sexual activity by the partner.**
4. **Persistent and complete lack of sexual interest with spouse/partner.**

8: Restricted verbal communication

This item rates any reduction in the individual's verbal communication (flow, content, rate) in familiar circumstances.

Do not rate the *quality* of the conversational content or poor *non-verbal* communication in this item.

0. Not present.
1. Minimally reduced flow, content or rate but responsive to others. Not problemtaic.
2. Markedly reduced conversation but manages to converse with prompting. Does not initiate conversation.
3. Severely reduced conversation – only converses in response to persistent efforts, otherwise not responsive. Conversations generally limited to monosyllabic or unintelligible replies.
4. Totally mute and unresponsive.

9: Leisure interests

This item rates the individual's active participation in non-obligatory, self-initiated, structured activity external to occupation. This may include hobbies, sports, as well as primarily intellectual pursuits. It is the self-initiated pattern of engagement that defines the rating of this item.

0. Satisfactory.
1. Participates in leisure activities but reports some lack of interest and/or pleasure. Not regarded as a problem.
2. Interests are markedly reduced but can be persuaded to watch TV, join in games for a while at least but tends towards passivity.
3. Severely reduced leisure activity. Almost total lack of interests but may watch some TV or listen to some radio with persuasion but without any apparent active interest. Does not participate in any structured leisure activities.
4. Totally uninterested in local or world events; never reads anything (if literate). Does not watch TV or listen to radio. Does not participate in any defined leisure activity.

10: Hazardous behaviour

This item rates whether or not the individual can safely handle matches or similar items such as a cigarette lighter.

It does not matter whether the individual actually has access or use for such items. The rater should make a judgement as if the individual would have access to such material.

The rating is made on the basis of the raters opinion given all available information.

The question raters should ask themselves is, "Should this person be allowed to have matches (or similar) and under what circumstances?".

0. **No supervision required.**
1. **Needs to be checked occasionally and reminded about the possibility of risks in handling inflammables. Not regarded as a problem.**
2. **Marked risk. Unsupervised during daytime but at risk at night where inflammables need to be locked away (e.g.. Would smoke in bed, use lighter to find way about in the dark).**
3. **Severe risk. Can be trusted to cook or smoke or be in possession of matches but only under general supervision.**
4. **Can never be given responsibility for matches or lighter without being a serious risk. Needs constant direct supervision smoking, cooking or any activity that is a fire risk.**

11: Need for supervision for security reasons

The rater may find this item difficult to rate.

It is a rating of the supervision judged necessary to minimise the risk of violence (self or externally directed) balanced against the use of the least restrictive alternative.

This item rates both environmental restriction considered necessary and frequency of supervisory contact. The item should be rated on the basis of the rater's judgment of the individual's supervisory needs and not the actual circumstance that the individual is in (should they differ).

0. **No supervision required.**
1. **Minimal risk. Can be in community with occasional (weekly or greater) contact with supervising authorities (eg community nurse, probation officer). Not problematic.**
2. **Marked risk to self or others. Can be in community but needs a least daily contact and assessment.**
3. **Needs a semi secure environment where there is 24 hour supervision because of severe risk.**
4. **Direct supervision all of the time – needs to be in a locked facility because of serious risk to self or others.**

Appendix O

Instrument Development Template