Prevalence of psychiatric morbidity in the adult population of Mamre: an empirical and methodological investigation

Sharon Rumble

A Dissertation submitted in fulfilment of the requirements for the degree of Master of Social Science

Department of Psychology
University of Cape Town

1994

ABSTRACT

A broad age-range community study was undertaken to estimate the base-line prevalence of mental disorder in the adult population of Mamre. The study further undertook to investigate a possible association between mental disorder and specified socio-demographic variables; to obtain basic information on health care utilization and attitudes to health services currently available in Mamre; and to investigate the validity of the Self Reporting Questionnaire (SRQ) (Harding et al., 1980) as a first-stage screen in a community sample in the South African context.

The general population of Mamre, a "coloured" community of approximately 5000 and presently undergoing transition from rural to urban status, was selected for study as there was an indication of a high rate of mental disorder among clinic attenders (Miller et al., 1991).

Using a cross-sectional descriptive study, prevalence was estimated in a two-stage design in which the SRQ was selected as the first-stage screen and the Present State Examination (PSE, 9th ed.) (Wing et al. 1974) as the second-stage criterion or gold standard. Both instrument have been used extensively in Africa (Parry, 1992), and both have been translated into Afrikaans.

Using power analysis to estimate first-and second-stage sample size (Shrout, 1992), a random sample of 481 respondents was selected from the total Mamre population frame, and was administered the SRQ along with additional socio-demographic questions in a first-stage questionnaire. A stratified sample of 121 respondents were then randomly selected from the two groups of SRQ high and low scorers (27.5% from SRQ positives and 12.5% from SRQ negatives) to participate in the second-stage PSE interview.

Adjusting for sample strategy, the weighted prevalence of mental disorder was estimated at 27.1% (95% confidence interval of 19.5-34.7%). This estimate was not conclusive however, as methodological weaknesses in the study, (such as the inadequate PSE inter-rater reliability study), and shortcomings of the instruments used (such as questions regarding the PSE's tendency to over- or mis-diagnose paranoid states) may have falsely increased the rate.

With a weighted sensitivity of 0.49 and specificity of 0.82 the SRQ was found to be at most moderately useful as a first-stage screen in this population. The high number of SRQ false positives (23%) was a particular factor attenuating its low false negative rate (10%).

There was no significant association between psychiatric disorder and gender, marital status, socio-economic status and alcohol use. Those reporting current illness or illness in the past month more likely to be rated positive on the PSE than those reporting physical health

($p<0.05$). For those ill during the last month (but not currently), those found to be mentally ill via the PSE were roughly three times more likely to have sought health care than those who were not mentally ill ($p<0.05$).

The findings are discussed in terms of both methodological and substantive implications.

ACKNOWLEDGEMENTS

I should like to thank a number of people and organizations.

My supervisor, Professor Leslie Swartz, whose expertise in this area of research has informed all aspect of the study and whose insight, energy and standards of excellence have helped me attain a goal I never thought possible.

My co-supervisor, Dr Charles Parry provided in-depth informed criticism at all stages of the study and gave encouragement that was always uplifting.

Dr Merrick Zwarenstein suggested useful questions for inclusion regarding utilization of primary health care services.

Professor Dirk Van Schalkwyk gave invaluable assistance with the statistical analysis of data found in Chapter 11.

Abe Koch, Gerrard Drennan, Pippa Rogers, Ronnelle Carolissen and Karen Ensink all gave generously of their time in administering PSE clinical interviews, and Colin Tredoux kindly ran the PSE-Catego programme.

Adam Davids was invaluable as a data-co-ordinator. The four lay-interviewers, Nigel Carlse, Veronica Arendse, Cedric Cupido and Randall Arendse remained enthusiastic and responsible through-out the study, often under difficult conditions. Merica Arendse and Diedree Pieterse provided useful assistance.

Helen Moffet meticulously edited a "tortuous" document. Marilyn Pether has done wonderful work typesetting the document and laying out the tables.

The Medical Research Council provided funding for the study. All views and conclusions drawn however, are my own and do not necessarily reflect those of this organization or any person who has provided me with information and guidance.

And my family, Christopher, Lauren and Olivia, who have kept me company all the way.

TABLE OF CONTENTS

TABLE OF APPENDICES[1]

GLOSSARY

(printed on green paper)

INSTRUMENTS

(printed on pink paper)

---

1 Appendices have been colour coded for ease of access and are arranged according to content and not in the order in which they occur in the text.

CALCULATIONS

(printed on grey paper)

METHODOLOGICAL DISCUSSIONS

(printed on fawn paper)

LIST OF TABLES

LISTOF FIGURES

PREFACE

Studies around the world have shown that between 17% and 24% of persons within general populations suffer from psychiatric disorders, (Parry, 1992; Reeler, 1987a). Generally, the severity of these disorders is not as extreme as that found in patients in psychiatric hospitals (Goldberg & Huxley, 1992). Nevertheless, they may be highly distressing and disruptive, both to the individual and his/her community.

Correct estimates of mental disorder are essential to the planning and evaluation of mental health services, (Parry, 1992). Given that a large number of "ordinary" persons may be at risk of suffering from mental disorder at some time in their lives, prevalence estimates of psychiatric morbidity, at community and primary health care (phc) service levels, are particularly important.

Goldberg & Huxley (1992) observe that common mental disorders, while they may not be severe in terms of symptomatology, cause enormous suffering and pain, may last a long time, and are financially costly in terms of inappropriate usage both to the individual and the phc services in general. Both *prevalence*[1] studies of the *rate*, nature and commonly associated socio-demographic variables of mental disorder at community level, as well as studies which provide information on the duration and extent of functional impairment caused by these disorders, are therefore important if effective measures are to be taken to alleviate this distress. In South Africa, no broad age-based, community study of the prevalence of psychiatric morbidity, nor the variables associated with it, has been conducted since that of Gillis et al., (1968). The need therefore, for a study that would provide base-line information pertinent to the planning and development of general phc policy, is identified.

When planning a prevalence study of mental disorder in a community, it is necessary to consider the population on which this estimate is to be based. According to Goldberg and Huxley (1992), although many people in the community suffer from mental disorder, the majority of these disorders are never seen by the mental health services. The reasons given by the authors will be discussed in more detail in a later chapter, but for now it can be said that many mentally disordered persons do not come to the attention of the mental health services, because:

(i) they do not present;

---

1 Definitions of italicized terms can be found in the Alphabetical Glossary of Technical Terms, see Appendix 1.1. Appendices are colour coded for ease of access and are arranged according to subject context and not in the order in which they occur in the text.

(ii) present, but are not identified as psychiatric cases;

(iii) are identified, but not referred to mental health services;

(iv) present to alternative services, such as an indigenous healer, who may or may not refer the person to the phc services; or

(iv) are unable to present, especially in developing countries, because they do not have access to either primary or mental health care facilities.

In order to ensure that the total spectrum of psychiatric morbidity of a community is measured, it is important that persons from all groups listed above form part of the morbidity estimate. It has been argued that this can be achieved only if the research study is implemented at community level (Goldberg & Huxley, 1980; Pan & Goldberg, 1990).

The relatively high rate of psychiatric morbidity in general populations gives credence to the importance of psychiatric epidemiology studies, and the need to develop appropriate psychiatric epidemiological instruments. As will be seen from the literature (reviewed in a later chapter), over the past two decades, epidemiological psychiatric measuring instruments have become increasingly refined and specifically orientated to community studies. There is, however, still controversy about their robustness and usefulness. This is no less an issue in South Africa, where the problem is compounded by the possible effect of using tests designed on and for one culture, to measure morbidity in another.

The foregoing discussion identifies two areas of enquiry, which this study undertakes to address:

(i) the need to conduct an epidemiological study to estimate the prevalence of psychiatric morbidity in a general population sample, in a community in South Africa;

(ii) the need to investigate the robustness and appropriateness of psychiatric epidemiological instruments, (chiefly, selected brief screening instruments), when used in the South African context.

CHAPTER OUTLINE

There are various ways in which the following research study may be presented. Although a stream-lined and goal- orientated approach may present the aims, methodology and findings of the research with more clarity, given the difficulties encountered during the course of the study, and the effect these had on the final research design and results, it was considered more

useful and informative to adopt a more discursive method in which the influences of methodological issues on the study are reported historically.

Chapter 1 considers the problems surrounding a universally acceptable **definition of mental disorder** appropriate to the description, not only of the more severe, but also the less severe disorders commonly encountered in the community. Advantages and disadvantages of the **categorical and dimensional models**, as appropriate means of measuring mental disorder, are also discussed. It is concluded that, in spite of its obvious limitations, the categorical model most nearly meets the specific goal of "counting cases" necessary to the satisfactory attainment of the aims of the present research.

Chapter 2 sets out the reasons for the necessity of conducting **psychiatric epidemiological studies at community level**, when attempting to measure and describe the total spectrum of mental morbidity. The historical development of these studies is also reviewed here.

Using Goldberg and Huxley's model (1980) to describe some of the possible pathways to psychiatric care from community level to psychiatric institutionalization, **Chapter 3** discusses the **prevalence** of mental disorder found at different levels within society, and reviews prevalence rates of disorder in both western and non-western communities, and in South Africa in particular. The effect on prevalence estimates when instruments designed for one culture are used to measure disorder in another, is also considered.

Chapter 4 considers the possibility of identifying a **socio-demographic profile** associated with the level of mental disorder encountered in the community. It is concluded that, although there have been some convincing research findings with regard to commonly investigated variables, the results do not indicate an incontestable profile.

Chapter 5 states the **rationale** for the present study and sets out the factors which led to the selection of the community of Mamre as the most suitable site for the research project. In this chapter the **initial aims** of the research, to estimate the prevalence of mental disorder in a general population study, and to investigate the efficiency of a first-stage screening instrument, are specified.

Chapter 6 discusses the **research design** and general methodological issues, including the selection of the **General Health Questionnaire**, version 28 (GHQ-28) (Goldberg & Hillier, 1979) as the first-stage questionnaire; the **Present State Examination** (PSE) (Wing et al., 1974) as the second-stage criterion interview; the translation of these instruments; and the selection and training of personnel.

Chapter 7 gives an account of the **first pilot** study and the events giving rise to questions related to the appropriateness of both first- and second-stage instruments, and discusses the inclusion of a second screening instrument, the **Self Reporting Questionnaire** (SRQ)

(Harding et al., 1980) as well as two additional criterion measures, the **Psychiatric Assessment Schedule** (PAS) (Dean et al., 1983) and **"suggested clinical diagnosis"** in a **second pilot study**. The latter is described in **Chapter 8**.

Results of this study led to the reformulation of the initial aims, which are reviewed in **Chapter 9**. In terms of the **reformulated aims**, primary and secondary objectives of the research study were identified. Primary objectives repeated the initial aim but replaced the GHQ-28 with the SRQ as the primary first-stage screen. Secondary aims were oriented to obtaining exploratory information regarding the efficiency of alternative first- and second-stage instruments, and the effect on estimated prevalence when different instruments were used simultaneously in the same population. To this end, the GHQ-28, the CGHQ-28 (an alternative method of scoring the GHQ-28), the PAS and "suggested clinical diagnosis" were retained as additional first- and second- stage measures.

Chapter 10 reports on the **main** study, and includes a review of the calculation of optimal size for first- and second- stage samples, the administration of first- and second-stage instruments, and problems encountered in the course of their administration.

Chapter 11, the results indicated an overall **prevalence rate of 27%** are discussed. The determination of the prevalence of specific diagnoses was problematic, however, as there was question regarding the mis-diagnoses of cases of paranoid schizophrenia and other paranoid states by the PSE-Catego, which is discussed further in **Chapter 12**. However, after examination of all PSE score sheets by an external psychologist, and in discussion with the rating psychologists it was estimated that approximately **95% of cases were either depression or anxiety related**. No statistically significant association was found between mental disorder and specified socio-demographic variables, although there was a significant **trend** between PSE positives and rate of consultation, for those who reported being ill in the last month. The weighted *sensitivity* of the SRQ was unacceptably low **(0.49)** although weighted *specificity* was within the higher range **(0.82)**. The SRQ correctly identified 67% of PSE cases, and its false negative rate was under 10%; however this was attenuated by its high false positive rate (49% of SRQ high scorers).

Aside from the PSE-Catego's possible **inappropriate diagnoses of paranoid states**, **Chapter 12** also discusses the strengths and weakness inherent in both the methodology used in the study and in the instruments themselves. It is concluded that, in spite of possible mis-diagnoses of paranoid states, the high prevalence rate of mental disorder estimated in the community is realistic. It is also concluded that the SRQ was moderately useful as a first-stage instrument in terms of the goals of the present research.

Chapter 13 concludes with a set of recommendations regarding **improvements to methodology**, (specifically in relation to the importance of a comprehensive inter-rater

reliability study); the possibility of **modifying** existing psychiatric first- and second-stage epidemiological instruments to ensure their validity in the context in which they are to be used; and the need to consider the implications of the high prevalence of mental disorder for the future **pattern of services and for training** needs of mental health professionals in the community.

1 DEFINING MENTAL DISORDER

PROBLEMS WITH A UNIVERSALLY ACCEPTABLE DEFINITION OF MENTAL DISORDER

A great deal has been written about what constitutes *psychiatric caseness* (Wing et al., 1981) and there is still controversy surrounding proposed definitions. This dissertation does not intend to enter this debate, but a basic understanding of how mental disorder is defined is important to inform the selection of psychiatric epidemiological measuring instruments, and to assess prevalence estimates of mental disorder found in similar community studies.

A fundamental problem for psychiatry is that there is no observable or measurable physical representation of mental illness, (Wing et al., 1974). Its presence must be inferred from observable behaviour, such as suffering and distress, impairment in functioning, impeded development, number and severity of symptoms, and in some instances, bizarre conduct. Mental illness is often thought of, and used, as if it were a construct which exists in reality in some concrete form. However, mental disorder is a "concept", an abstraction, liable at any time to be adjusted or discarded, (Goldberg & Huxley, 1980; Shepherd, et al., 1986; Wing, 1983).

Most prevalence studies of mental disorder have defined psychiatric caseness in terms of a critical number of a particular constellation of symptoms, which have continued for more than a critical time (Goldberg & Huxley, 1992).

Counting symptoms "fits" with the popular categorical model which allows mental disorders to be thought of as discrete entities with distinct causes, course and treatments.

Determining caseness in terms of number and severity of symptoms alone may however, be a limited definition of mental disorder. This approach does not take account of important factors such as vulnerability, severity of stressor, and environmental and social supports or deprivations (Goldberg & Huxley, 1992). It has long been observed that persons sharing the same symptom pattern do not necessarily experience the same outcome. Some deteriorate to the point where they can no longer function in the community and must be admitted to psychiatric hospitals, while others continue to function effectively within the community. If psychiatric caseness were dependent solely on the number and severity of symptoms, differences in response to the same set of symptoms could not be reasonable explained. Goldberg and Huxley (1992) suggest that a dimensional approach to understanding mental disorder, in which the socio-biological context of the individual is considered, may offer a plausible explanation for adaptive or maladaptive responses to similar sets of stressors.

CATEGORICAL MODEL VERSUS DIMENSIONAL MODEL OF MENTAL ILLNESS

Psychiatric epidemiological studies have shown that mental illness is not homogeneous (Goldberg & Huxley, 1980,1992; Pan & Goldberg, 1990; Vazquez-Barquero et al., 1988). While there is evidence of a common denominator of symptoms of depression, anxiety, somatic disturbance, fatigue and irritability (Goldberg & Huxley, 1980), which appears to underlie nearly all forms of mental disorder, disorders found in the community are generally less severe than those found in psychiatric hospitals (Goldberg & Huxley, 1992). The categorical model however, assumes that mental disorder can be understood in terms of discrete categories. While there is evidence that this model may be appropriate in measuring the more severe mental disorders, such as organic mental disorders and schizophrenia, there is question regarding its suitability as a measure of the common mental disorders found in the community, such as depressive and anxiety related disorders, (Goldberg & Huxley, 1992; Van den Brink, et al., 1989). A version of Goldberg and Huxley's (1992) table of common mental disorders found in the community, and their accompanying diagnostic classifications according to two categorical diagnostic systems, the International Classification of Disorders, 10th edition (ICD-10), (World Health Organization, 1973) and the Diagnostic and Statistical Manual-111R, (DSM-111R), (American Psychiatric Association, 1987), is reproduced in Table 1.1.

Table 1.1 Diagnostic classification of some common mental disorders, showing current International and United States diagnostic codes, (adapted from Goldberg & Huxley, 1992, Table 1.2, p.10)

World Health Organization: International Classification of Disease 10th Ed (ICD-10) (1973)	American Psychiatric Association: Diagnostic & Statistical Manual-111R (DSM111R) (1987)
Depression:	
296.2 Psychotic depression	296.2x Major depression, episode
296.3 Manic depressive psychosis	296.3 Major depression, recurrent
300.4 Depressive neurosis	300.4 Dysthymia
Anxiety related disorders:	
300.0-3 Neurosis	300.29 Simple phobia
300.0-3 Neurosis	300.01 Panic disorder
300.0-3 Neurosis	300.02 General anxiety disorder
300.0-3 Neurosis	300.30 Obsessive compulsive disorder
300.5-9 Neurosis	309.90 Adjustment disorder

Problems with the Categorical Model

Goldberg and Huxley (1992) suggest that the categorical model is an inadequate description of mental illness in the community for the following reasons:

(i) Mental disorders found in the community are not discrete entities. There is evidence to suggest that there is a significant overlap between anxiety and depression, and that combined syndromes are frequently encountered in the general population, (Eaton & Ritter, 1988).

(ii) Mental disorders in the community are continuously distributed in the population, (Angst & Dobler-Mikola, 1984). This means that there is no definitive point on some underlying axis, which distinguishes normal from abnormal behaviour. Any decision regarding threshold is therefore arbitrary and dependant upon the purposes of the research (Wing et al., 1981). Mental illness in the community is best understood as an evolving process, existing along a continuum, so that an individual has either more or less of a given symptom at any one time. In setting a threshold, a distinction is made between normality and borderline psychiatric morbidity. This involves imposing a

dichotomy upon a continuum of severity (Goldberg & Huxley, 1980). In the general population, where normal and abnormal reactions to life events frequently overlap, a definitive threshold between psychiatric caseness and non-caseness is inappropriate.

(iii) There are over a dozen different diagnostic interview schedules, each differing from the others in terms of coverage, threshold levels and time framing (Williams, et al., 1980). The use of different criteria to define psychiatric caseness means that when using different diagnostic systems (for example the ICD or DSM111) within the same population, it is possible to identify different persons as cases, with different diagnoses (Surtees et al., 1983). This problem is applicable when identifying severe mental disorders, but is especially problematic when addressing less severe disorders.

Although steps have been taken in recent years to redress this problem (see for example, the Composite International Diagnostic Interview, (CIDI) (Robins et al., 1988) it seems clear that what gets counted in epidemiological surveys is very much a function of the way mental disorder is defined (Eisenberg, 1991, in Goldberg & Huxley, 1992). Most diagnostic systems are based on the characteristics of patients who present for medical treatment, identified as mentally disturbed by the health care services, and subsequently referred to professional mental health care. These persons may present similar symptom patterns to many others in the community, who do not present, or who are "missed" by their phc service. The operational criteria for psychiatric caseness identify cases which are generally more severe, but which do not differ in kind from "missed" cases (Eisenberg, 1991, in Goldberg & Huxley, 1992). Goldberg and Huxley (1992) observe that those who survey mental disorder within communities "...are taking a concept of mental disorder which has been found to be meaningful among those seeking specialist care, and in effect asking how many people with similar symptoms are there in the community?" (Goldberg & Huxley, 1992, p. 7).

Proposed Dimensional Model

Goldberg and Huxley (1992) propose an alternative multi-dimensional model more appropriate to the measurement of mental disorder in the community. This model allows for a more comprehensive picture of common mental disorders, through the assessment of an individual's position on various bio-social dimensions, such as the individual's genetic make-up, parental influence, personality, social adversity and support. Whilst this model has obvious advantages over the categorical model, it is still at the conceptual stage and has not yet been operationalised.

Reasons for selecting the Categorical Model in this study

The choice of model depends on the purpose of the research (Goldberg & Huxley, 1992). While diagnostic systems based on the categorical model may not be the most appropriate systems to describe and measure mental disorder in the community, their practical usefulness in counting cases in prevalence studies has been well documented, (Bebbington et al., 1981; Cheng, 1988; Hodiamont et al., 1987; Orley & Wing, 1979; Vazquez-Barquero et al., 1987). According to Tarnopolsky et al., (1979), psychiatric evaluation in terms of this model remains the primary means for making judgements for the purpose of case identification. It is also the ultimate criterion against which other measures of identifying psychological disorder can be validated.

Furthermore, most of the current diagnostic systems, used in both clinical and community studies, are based on the categorical model, for example, the ICD8 (1973) and subsequent ICD editions, and the DSM-111R (1987). These systems are readily available and their validity well established, within the confines of their limitations, (Van den Brink, et al., 1989). Use of these instruments allows for the comparison of prevalence rates between studies. Above all, the categorical model is useful because classification facilitates communication about disorders. It assists in the "bookkeeping" of health administration, in carrying out clinical work, and producing homogeneous groups for research, (Goldberg & Huxley, 1992).

These advantages argue in favour of selecting measuring and diagnostic psychiatric instruments based on the categorical model for use in the present study. For the purposes of this dissertation, mental health status is therefore assessed in terms of number and severity of symptoms, and psychiatric caseness is determined according to diagnostic systems which describe mental illness in terms of discrete entities, such as the ICD-8 and the DSM111-R. Results, however, need to be assessed in light of the dimensional model's broader perspective of mental disorder, and in terms of the categorical model's limitations.

As mental disorder is not homogeneous, and cannot be encompassed by a single definition, based solely on severely distressed patients, the following chapter suggests that studies aimed at measuring the full spectrum of morbidity, must be directed at community level, where both severe and less severe mental disorders may be encountered.

2 RATIONALE FOR A COMMUNITY BASED STUDY

This chapter discusses the rationale for the decision to conduct a community based study. The approach falls within the ambit of psychiatric epidemiology, and refers to the investigation of psychiatric morbidity within the total community, including both persons currently identified as mentally ill and those not identified as such. The need for a community based study is examined from two points of view.

(i) Firstly, if many persons with psychiatric morbidity never come to the attention of the phc services or mental health facilities (Goldberg & Huxley, 1992), then prevalence estimates based on clinic attenders or psychiatric in-patients alone do not truly reflect the total psychiatric morbidity of a given community. Goldberg and Huxley's (1980) framework to describe a possible pathway to psychiatric care is used to illustrate this point.

(ii) Secondly, it is suggested that by embracing public health awareness policy, and informing epidemiological psychiatric instrument development relevant to the measurement of mental disorder in the community, the psychiatric epidemiological approach is most appropriate to the attainment of the goals of the present research.

GOLDBERG AND HUXLEY'S PROPOSED MODEL OF ONE POSSIBLE PATHWAY TO ATTAINING PSYCHIATRIC CARE

Rationale for the selection of Goldberg and Huxley's model

There are several different pathways leading to the identification of psychiatric caseness and the attainment of mental health treatment. Goldberg and Huxley's (1980) model was selected as a basis for discussion in the present study because it presents a convincing argument in favour of adopting a community based approach. Specifically, it addresses the problem of excluding persons who do not consult their phc services, or who are not identified as psychiatric cases by those services, from total morbidity estimates. This model is however, medically centred and biased in favour of central mental health services. It is thus especially relevant in communities where these services are readily available. In many developing countries, including South Africa, where these services are scarce or non-existent, the pathway to psychiatric care is particularly tortuous (Miller, et al., 1991; Vazquez-Barquero et al., 1990) and may not follow Goldberg and Huxley's (1980) model. Non-medical routes are also possible, for example an individual may consult an indigenous healer, herbalist,

pharmacist, pastor, or other resource, before being referred to the mental health services (Gater et al., 1991).

The five levels and four filters

Table 2.1 indicates an individual's position relative to health services in terms of the five levels and four filters proposed by Goldberg and Huxley's (1980) model. Also listed in Table 2.1 are the estimated prevalence rates for mental disorder at each level. Discussion regarding these rates is found in Chapter 3, p. 23).

Table 2.1 Five levels and four filters, with estimates of prevalence rates at each level. (Table adapted from Goldberg and Huxley (1992), p. 4).

Level 1	The community, including all persons who would meet the criteria of psychiatric caseness, should they be assessed by a psychiatric interview
	26%-31.5%
.....	1st filter
	(Illness behaviour)
Level 2	Total mental morbidity-attenders in primary care, irrespective of whether or not their mental disorder is detected
	23%/
.....	2nd filter
	(Ability to detect disorder)
Level 3	Mental disorders identified by doctors (conspicuous psychiatric Morbidity)
	10.2%
.....	3rd filter
	(Referred to mental illness services)
Level 4	Total morbidity-mental illness services, ie, patients attending psychiatrists, psychologists and social-workers both in and out-patient clinics and psychiatrists' offices
	2.3%
.....	4th filter
	(Admission to psychiatric beds)
Level 5	Patients admitted to psychiatric hospitals
	0.5%

This table shows that each of the five levels is separated by four filters. At each filter, different factors operate either to promote or retard the individual's progress from one level to the next. These factors include socio-cultural context, prevailing theoretical concepts of mental illness and health, socio-demographic variables, vulnerability, availability of hospital beds, etc. The relevance of this model to the present research design, lies in its ability to demonstrate that factors other than symptoms are important in defining who is "sick" and what constitutes mental illness. The factors which operate at the first and second filters are particularly pertinent to the decision to conduct a community based study and are discussed below.

First Filter: Illness behaviour

Various factors operate in determining whether an individual perceives him/herself as emotionally distressed, and seeks confirmation of the illness experience from phc services. These include: the current state of medical knowledge and prevailing attitudes towards health and illness, the availability of and ease of access to health services, the socio-demographic characteristics and health beliefs of the individual, ability to pay for care, symptom disability and the individual's evaluation of the meaning of such perceptions (Goldberg & Huxley, 1992). Once the individual adopts the "sick role" he/she passes through the first filter. Alternatively, illness may not be experienced as such, or may be denied, and treatment delayed.

Goldberg and Huxley's model however, may not be appropriate in all cultures. In many non-western cultures for example, spirit possession by ancestors is regarded as normal. Persons in the community experiencing such symptoms would not necessarily interpret their behaviour as needing medical attention. However, if they were assessed by a standard psychiatric research interview, it is likely that some could be given a psychiatric diagnosis (Gater et al., 1991).

Second filter: Ability to detect disorder

While there is evidence that many persons with psychological disorder in the community do present to their phc services (Goldberg & Huxley, 1992), there is also evidence that they may do so for reasons other than their psychological symptoms (Goldberg & Huxley, 1992).

Psychological symptoms are often presented somatically. The detection of the psychopathological component of the illness experience is largely dependant on the training and sensitivity of the phc professional to underlying psychological factors. If mental illness is not recognised, the passage to level 3 is blocked, and the individual becomes part of the hidden morbidity of the community.

Conclusions regarding the need to assess mental disorder at level 1, the community

From the above discussion it is evident that many persons who would meet the criteria for psychiatric caseness remain undetected in the community. Prevalence of mental disorder estimated on data obtained only from levels 2, 3, 4 or 5 is therefore inadequate as a description of the full spectrum of morbidity in the community. Such estimates do not include those who do not interpret their discomfort as illness and therefore are never seen by the phc services, and those who do consult, but whose underlying psychiatric morbidity is missed because they present with somatic symptoms.

General population psychiatric epidemiological surveys are designed to discriminate across the whole spectrum of psychiatric morbidity, including those persons whose mental disorder is not conspicuous at either levels 1 or 2. Such studies are therefore essential when estimating prevalence of mental disorder in the community.

RELEVANCE OF THE PSYCHIATRIC EPIDEMIOLOGICAL APPROACH TO THE PRESENT STUDY

Psychiatric epidemiology

Psychiatric epidemiology attempts to understand and measure mental health status from a community-centred point of view, a more comprehensive approach than studying known psychiatric patients. With the view that the understanding of mental disorder can be enhanced through the study of environmental issues, health in general and the distribution of both mental health and ill-health in populations, this approach attempts to include all aspects relevant to mental health. Moreover, psychiatric epidemiology focuses on identifying factors both within the environment and the individual which either increase or decrease vulnerability to disorder. Knowledge of these factors informs programmes aimed at *preventing* disorder before it occurs. As such, psychiatric epidemiology goes well beyond determining *incidence*, prevalence and the distribution of disorder.

Historical review of psychiatric epidemiology, and the role of public health awareness policy

The following section gives a selective overview of the historical development of the psychiatric epidemiological approach and refers to public health awareness policies that both prompted and informed its most recent developments. The review is presented as motivation for the selection of research design and psychiatric instruments used in the present study.

Chapter 2: Rationale for a community based study

Since the time of Ancient Greece, epidemiological methods have been applied to the investigation of physical infectious diseases. During the Victorian era in Europe, these methods were widened to embrace the investigation of non-infectious disorders (Parry, 1992). Psychiatric epidemiology emerged in its modern form in the USA during the 1920s and 1930s, largely as the result of the initiatives of social scientists (Grob, 1986, in Parry, 1992). Early studies tended to focus on hospital populations or outpatients at university psychiatric clinics and relied on key informants and agency records to supply the information that would enable them to identify cases (Dohrenwend, 1990). In general, diagnostic methods were not given, sample sizes were small, and there was little control over the case finding process (Bland, 1988, in Parry, 1992).

From the mid-1930s onwards a number of factors emerged which motivated surveys of the general population in psychiatric research. The social reform movement of this period proposed that individual behaviour and psychological well-being could not be understood in isolation from broader social issues, such as poverty and over-crowding (Heller et al., 1977). This move towards community studies was accompanied by a greater reliance on direct interviews with subjects. There were, however, still shortcomings, most important of which was that psychiatry was still person-centred, and did not have the conceptual framework, or technical tools, to understand and measure problems in terms of the person-environment interaction.

After the second World War, there was a tremendous expansion in psychiatric nomenclatures. This followed a major conceptual re-orientation in clinical psychology, prompted by the belief that traditional forms of psychological therapy lacked in optimal social utility. New helping modalities came about through concern with environmental issues, health in general, prevention rather than cure, and the realization that the distribution of disease in a population is fundamental to the understanding of disease aetiology. This approach had an increasingly environmental focus, and was based on the concepts of prevention, competence and adaptation (Heller et al., 1977). New areas for research and practice were thus opened up. In studies such as the Midtown Manhattan Study (Srole et al., 1962, in Parry, 1992) and the Stirling County Study (Leighton et al., 1963, in Parry, 1992), interviews became more standardized and explicit data was collected. Case identification however, depended on the psychiatrist's evaluation of the protocols, even when interviews were conducted by psychologists. This method assumed that psychiatric diagnosis between psychiatrists was reliable, an assumption which has since proved to be inaccurate (Van den Brink et al., 1989). The lack of any standardized psychiatric diagnostic system during this period meant that all diagnoses were subjective to a degree, depending on the cultural, social and theoretical background of the individual psychiatrist. In general, classification was made in terms of caseness and impairment, rather than in terms of diagnostic types. Even where these were attempted, reliability of diagnosis was questionable as interviewing procedures were typically

not made explicit (Parry, 1992). Each study tended to pioneer its own unique methods and procedures for counting cases, with the result that they were inadequate in terms of content, criterion and construct validity (Dohrenwend 1990). As there was no common definition of mental illness, nor a common research design, comparison between studies was meaningless. This impeded the usefulness of both the research itself and psychiatric epidemiology as a whole. In a review of prevalence studies of psychiatric morbidity in general populations world wide, Dohrenwend and Dohrenwend, (1982) found rates ranging from 8.7% to 20.3%. The extent of this range was seen as a reflection not so much of a difference in the actual prevalence rates, as a difference in criteria specified in the large number of different nomenclatures and measuring instruments of this period (Dohrenwend, 1990; Surtees et al., 1983).

Psychiatric studies undertaken in the mid- 1970s and 1980s, sometimes referred to as third generation psychiatric epidemiological studies (Dohrenwend, 1990), were characterized by an emphasis on standardization and increasingly refined diagnostic criteria, for example, the PSE. During this time, the first diagnostic systems emerged in which the criteria were sufficiently specific enough for questions to be matched to them, such as, Research Diagnostic Criteria, (RDC), (Feighner et al., 1972), and the Diagnostic and Statistical Manual of Mental Disorders, (DSM111), (American Association of Psychiatry, 1980). As a result, modern psychiatric epidemiologists now have at their disposal interviews which are closely linked to diagnostic criteria and standard nomenclatures (Parry, 1992).

While these instruments represent a significant improvement over earlier measuring tools, in the semi-structured interviews, for example, the PSE, the Schedule for Affective Disorders and Schizophrenia, (SADS), (Endicott & Spitzer 1978), and the Psychiatric Assessment Schedule, (PAS), (Dean et al., 1983), there is still an element of reliance on the experience and skill of the clinician to reduce measurement error (Wing, 1983). Fully structured instruments, such as, the Centre for Epidemiological Studies-Depression Scale, (CES-DS), (Radloff 1977), the Psychiatric Epidemiology Research Interview, (PERI), (Dohrenwend et al., 1986), and the Diagnostic Interview Schedule, (DIS), (Robins et al., 1981), were also designed specifically for use in psychiatric epidemiological studies. Both semi-structured and structured interviews may be criticized on the grounds that very few validity studies have been done on these instruments. It is assumed that their validity derives from their development with psychiatric patients, and from their use in clinical research (Dowenrend, 1990). The possibility that these instruments may be less sensitive to the level of mental disorder commonly found in the community, means that their use in community studies, while an improvement on earlier instruments, may fall short of the ideal. Whatever their shortcomings however, they represent the best objective measurement of psychiatric morbidity currently available. They are therefore used as gold standards against which other case-finding instruments, which need to be validated in a given context, are measured.

In the 1970s and 1980s, the need to describe mental disorders in the community, and to measure large numbers of persons economically, led to the development of screening questionnaires, such as, the General Health Questionnaire, (GHQ), (Goldberg, 1972) and the Self Reporting Questionnaire, (SRQ), (Harding et al., 1980). Using lay-interviewers, these questionnaires were able to measure large numbers of people. This was a considerable advantage over the costly method of having a psychologist or psychiatrist administer a clinical interview to all persons in the sample.

While these instruments are brief and easily administered, there is evidence to suggest that they do not converge closely with diagnoses based on clinical interviews (Dohrenwend, 1990). They therefore cannot be used as diagnostic instruments. Lack of agreement between screening and clinical interviews is not unexpected, given that diagnostic systems and clinical interviews agree only poorly with one another at all levels, but especially at community level. Moreover, screening instruments tend to overestimate the number of true cases and are therefore, not recommended for use on their own in single one-stage designs (Parry, 1992).

In the studies of the 1970s and mid-1980s, a single one-stage design was traditionally used, in which all participants were administered a clinical interview (Orley & Wing, 1979). This method was generally costly and wasteful of time, finance and manpower. Currently, a multi-stage research design, using third-generation instruments is recommended. This involves using the short inexpensive screen to survey a large sample of the population. Potential cases are then given the longer and more costly follow-up clinical interview by a psychiatric clinician. Only those persons identified as potential cases, and possibly a percentage of persons identified as potential non-cases, receive the full-scale clinical interview. Semi-structured interviews and standardized screening questionnaires, when used together, in a two-stage case finding method, complement and cross-check each other. This capitalizes on the ability of the screening instrument to provide reliable measurement over the full range of important dimensions of psychopathology, and on the ability of the clinical examination to provide reliable diagnoses in groups, where the symptomatology involved is not rare (Lewis & Williams, 1989). The design allows standardized systems of evaluating psychiatric symptoms to be incorporated into the study of large populations with a high degree of reliability (Vazquez-Barquero et al., 1987). It is also appropriate in situations where the validity and reliability of the screening instrument itself is under investigation. Here the criterion serves as a gold standard in the identification of true cases. These designs have frequently been used in studies to estimate prevalence in the community (Banks, 1983b; Benjamin et al., 1982; Cheng, 1985; Goldberg, 1972; Goldberg et al., 1976; Goldberg & Hillier, 1979; Marie & Williams, 1985; Rabins & Brooks, 1981; Tarnopolsky et al., 1979; Vazques-Barquero, et al., 1986). The advantage of the two-stage design lies largely in its cost-effectiveness (Bebbington et al., 1981; Cheng, 1985; Dohrenwend & Dohrenwend, 1982; Duncan-Jones & Henderson, 1978; Williams et al., 1980). However, in certain

circumstances, the two-stage design may not be the most cost-effective option (Shrout 1992). The two-stage design should only be used when it can be shown that the screen is significantly cheaper (in terms of cost, time and efficacy) than the clinical interview.

Developments in standardized psychiatric interviews, diagnostic systems and research design have together resulted in the improvement of measuring and describing mental disorder in the community.

Disadvantages and advantages of psychiatric epidemiology in the measuring of the entire spectrum of mental health/ill health

Disadvantages

The definition of psychiatric caseness in the community remains a problem in psychiatric epidemiological studies. In spite of the fact that these studies focus on the community, where the level of disorder is expected to be less severe than that found either in clinics or hospitals (Goldberg & Huxley, 1992), the concept of psychiatric caseness tends to remain grounded in concepts derived from hospital patients (Williams et al., 1980). This disadvantage has already been discussed in Chapter 1, where the rationale for using measuring instruments based on the categorical model in the present study, was given.

Advantages

Aside from contributing to the understanding, description and measurement of the total spectrum of mental disorder and health, the psychiatric epidemiological approach enables the researcher to:

(i) enumerate as accurately as possible the prevalence of mental disorder in a given population, within the limitations of current measuring instruments;

(ii) expand the clinical picture of mental disorder by not confining attention solely to those who have sought a particular type of care;

(iii) study complex questions, (such as the relationship between social class and psychological disorder) which cannot be studied with laboratory or experimental methods (D'Auanno & Price 1989, in Helman, 1990);

(iv) access the community while disturbing the usual behaviour of that community as little as possible (D'Auanno & Price in Helman, 1990); and

(v) provide relevant information for the planning and provision of services.

CONCLUSIONS REGARDING THE DECISION TO CONDUCT A COMMUNITY BASED PREVALENCE STUDY

Given that there are persons in the community with unidentified mental disorder, and given the availability of psychiatric epidemiological instruments and case-identifying designs, capable of identifying and measuring the full spectrum of morbidity, it was concluded that a community-based study was both appropriate and realistic to attain the goals of the present research.

3 REVIEW OF PREVALENCE RATES OF MENTAL DISORDER IN COMMUNITY STUDIES IN DIFFERENT CONTEXTS

This chapter reviews prevalence rates of mental disorder estimated in several community studies world wide. Prevalence rates in western developed communities are compared with those found in non-western, or developing communities, and differences in rates found between community and hospital populations are examined. The appropriateness of using psychiatric epidemiological instruments designed in one culture for use in another, and the effect of their use on estimated prevalence rates, is also discussed.

A COMPARISON OF PREVALENCE RATES FOUND IN ELEVEN COMMUNITY STUDIES WORLD-WIDE

Goldberg and Huxley (1992) have compiled a table of a number of recent community-based surveys conducted world-wide. All these studies used direct assessment of community respondents by standardized research interview, and showed a range of morbidity from 9% to 24.1%. Their table is reproduced in Table 3.1 and adapted to include four community-based prevalence studies conducted in Africa.

Table 3.1. Prevalence rates for random samples of the general population for all psychiatric illness in the past month; recent surveys based on direct assessment by standardized research interview. (Adapted from Table 2.2, Goldberg & Huxley, 1992, p. 18, to include four African studies, (Parry, 1992).

Investigator	Place	Method	Sample size	Prevalence %
Hodiamont et al. 1987	Holland	GHQ/PSE	486	7.3
Duncan-Jones and Henderson 1987	Australia	GHQ/PSE	157	9.0
Bebbington et al. 1981	England	PSE	800	10.9
Gillis et al. 1968	South Africa	Structured Questionnaire	500	11.8-23.0
Vazquez-Barquero et al. 1987	Spain	PSE	425	14.7
Hollifield et al. 1990	Lesotho	DIS	356	22.8
Ben Arie et al. 1983*	South Africa	PSE/MMPI[1]	139	23.7
Vazquez-Barquero et al. 1981	Spain	CIS[2]	415	23.9
Orley and Wing 1979	Uganda	PSE	191	24.1
Cheng 1988	Taiwan	CHQ3[3]/CIS	489	26.2
Bester et al. 1991*	South Africa	SRQ	400	39.0
Overall rates				17.4-18.5

* The age group in these studies was restricted, to the 65+ year group

\# The findings of Bester et al., 's study are not included in the overall rate, as a standard psychiatric research instrument was not used.

1 Minnesota Multiphasis Personality Inventory (MMPI) Hathaway, S., & MuKinley, J., (1951)

2 Clinical Interview Schedule (CIS) Goldberg, D., Cooper, B., Eastwood, M., Kedword, M., & Shepherd, H. (1970)

3 Chinese General Health Questionnaire, (CHQ), Cheng, 1985

From this table, it can be seen that the overall prevalence rate is between 17.4% and 18.6%. These rates are compatible with prevalence rates suggested by Parry (1992) and Reeler (1987a). The lowest rate, 7.3%, was found in Hodiamont et al.,'s 1987 study, and the highest, 26.2%, in Cheng's 1988 study. In general it would appear that prevalence rates are higher in studies conducted in developing and rural communities. Following Goldberg and Huxley's (1992) method, a mean prevalence rate reported for these community studies indicates a range of between 11.8% and 24.1%, with an overall rate of between 20.1% and 22.0% in developing countries, compared with an overall rate of 13.4% found in more developed communities. These rates suggest that psychiatric morbidity in non-western cultures may be of the same order as, or even higher than that found in more developed countries (Ben-Tovim, 1987; Harding et al., 1980; Reeler, 1987a). Statements concerning suspected higher prevalence rates in developing non-western communities should however, be treated with caution. In Ben-Arie et al., (1983) and Bester et al. 1991, the age range was restricted to the elderly ie. 65 years old. As there is evidence that persons over 65 years are likely to experience a high rate of mental distress (Gillis, 1991), inclusion of these results in the overall prevalence rate may have falsely raised this estimate. Of the six studies in developing communities, three were conducted in rural districts (Bester et al., 1983; Hollifield et al., 1990; Orley & Wing, 1979). In rural communities in developing countries, the scarcity of phc services may result in the failure to initiate or implement preventative measures. This in turn could explain the high prevalence rate found in these communities (Goldberg & Huxley, 1992). Notwithstanding these factors, prevalence estimates of mental disorder in communities around the world appear to fall within a similar range, regardless of cultural background.

Prevalence estimates found in community studies conducted in South Africa

No broad age-range community studies, using standardized research interviews, have been conducted in South Africa since that of Gillis et al., (1968). Studies which are restricted either in terms of their instrument selection, sample size or choice of population, include those of Ben-Arie et al., (1983) and Bester et al., (1991). Prevalence estimates from all three studies should be regarded only as a rough guideline. In two of the studies, (Ben-Arie et al., 1983; Bester et al., 1991), the age range was restricted, so that neither sample was randomly selected from the total population; nor was the estimated prevalence in either sample weighted back to reflect prevalence in the respective general populations. In Bester et al.,'s (1991) study, a screening instrument which was not designed to be a psychiatric case finding instrument, (the SRQ), was used to assess prevalence in a one-stage design. Given that screen positive estimates are likely to be higher than the true positives, these rates are nevertheless consistent with the highest range (24.1%) found in other developing communities (Freeman, 1990a; Miller et al., 1991).

South Africa has a heterogeneous population which, until recently, was segregated by law into distinct racial (ethnic) groups. The different racial groups have generally experienced different socio-economic status, with many "coloured" and "black" people living in depressed socio-economic conditions. Evidence suggests that persons from informal settlements and the lower income groups, are likely to experience a high level of mental distress (Goldberg & Huxley, 1992). It seems safe to assume that the prevalence of mental disorder in these groups may be higher than that found in more advantaged groups. However, evidence in support of this view is not consistent. A study by Lambley (1973) found no great difference in anxiety between different population groups, and a three year community survey by Dick et al., (1978a) found mental disorder to be a major community health issue for **all** groups in South Africa, regardless of socio-economic status. The differences in prevalence rate of mental disorder between different socio-economic groups cannot be satisfactorily resolved by examination of prevalence rates recorded in the table 3.1. Although Bester et al.,'s (1991) study was conducted on an urban/rural population and indicated a higher prevalence rate than the other two studies conducted in urban communities, the use of a screening instrument in this study as a case-finding tool does not lend validity to the findings. The methodological problems associated with the definition of psychiatric caseness, and the difficulties of using instruments which may be inappropriate, (which is discussed below), do not permit any global statements about which group has more or less disorder. The inability to make any firm statement regarding the prevalence of mental disorder either in advantaged or disadvantaged socio-economic groups, highlights the need for community based prevalence studies of mental disorder in South Africa.

COMPARISON OF PREVALENCE RATES FOUND IN THE COMMUNITY AND IN CLINIC AND HOSPITAL SETTINGS

Goldberg and Huxley, (1992) compiled a table of prevalence rates of mental disorder at the five levels specified in their 1980 study. (see Chapter 2, Table 2.1, p. 12) Examination of this table shows that prevalence of mental disorder at community level (26%-31.5%) is greater than at any other level and only sightly lower at level 2 (23%). The minimal difference in prevalence rates between levels 1 and 2 suggests that the majority of those with mild disorders **do** contact their phc services (Goldberg & Huxley, 1980). However, although many of these patients consult their phc services, they may do so for reasons other than their psychological symptoms, or at times when they are not in fact symptomatic (Goldberg & Huxley, 1992).

Prevalence rates at levels 3, 4, and 5 do not necessarily include all persons contained within the lower levels (Goldberg & Huxley, 1992). Many of those diagnosed by general practitioners (level 3), as well as a small group of those referred for specialist mental illness services (level 4), do not meet research criteria for a diagnosis (Goldberg & Huxley, 1992).

High prevalence of disorder at level 1 lends support to the argument for the need for direct assessment at community level, put forward in Chapter 2. Assessment at levels 4 and 5, and even at level 3, give prevalence rates which are significantly skewed in favour of the more severely disturbed person.

THE APPROPRIATENESS OF USING PSYCHIATRIC INSTRUMENTS DESIGNED IN ONE CULTURE FOR USE IN ANOTHER

The problem of defining and measuring mental disorder in general, and in the community in particular, has been discussed in earlier chapters. For practical reasons, and for the purposes of the present research, instruments based on the categorical model were selected for use in this study. However, it has also been acknowledged that instruments based on this model, are capable of yielding only moderate concordance in terms of psychiatric caseness (Goldberg & Huxley, 1992; Surtees et al., 1986). Table 3.1 shows that the PSE is the most commonly used research interview in community studies internationally. Studies using the PSE, report rates which are generally lower than those using either assessments based on the DSM-111 system (for example, SADS and the DIS) or on the Clinical Interview Schedule, (CIS), (Goldberg et al., 1970). This suggests that the use of different diagnostic systems may yield different prevalence rates in the same population (Dean et al., 1983). Further examination of Table 3.1, however, shows that in general, rates are within a comparable range. The question of whether it is appropriate to use psychiatric instruments designed on hospital in-patients to assess mental disorder at community level, was addressed in Chapter 1. Here it was argued that, while these instruments may be more appropriate to the description of the severely disturbed person, they are nevertheless, the best measuring instruments currently available for use in community surveys.

The question arises regarding the appropriateness of using instruments designed in western countries, in non-western countries. This debate revolves around the issue of whether the features of mental illness are universally experienced and interpreted, or whether they are specific to the cultures in which they occur.

Universalists within psychiatry form part of the Western biomedical tradition which conceives science, medicine and illness as concerned with empirically derived "objects" which reflect what is "real" and in "nature" (Fabrega, 1989). As such, mental disorders may be viewed as discrete entities with specific causes, course and treatments. These concepts are seen to be objective, neutral and value-free, and consequently have the potential to be universally applied. Cultural relativists, on the other hand, argue that disease is not an entity, but rather an explanation, and that, "... culture does considerably more than shape illness as an experience - it shapes the very way illness is conceived" (Kleinman, 1977 p. 3). Cultural relativists contend that in order to understand the experience of mental illness within the

context in which it occurs, it is necessary that the researcher understand the way in which culture influences the perception, classification, process of labelling, explanation, expression of symptoms, course and the decisions regarding the treatment of sickness (Kleinman, 1977).

Opponents of the universal approach suggest that psychiatry is itself part of the western cultural tradition; consequently, any definition of mental illness represents a cultural category "constructed by psychiatrists in the west to yield a homogeneous group of patients" (Kleinman, 1977, p. 3). To apply this definition of mental illness to persons in other cultures is to impose western cultural categories as if they were culture free. The danger of the universalist approach lies in the potential to miss aspects of a phenomenon viewed as important by people in other cultures, and of imposing conclusions based on concepts which may be valid in one culture, but which may be foreign when used in another (Brislin, 1986).

If the existence of a universal psychological abnormality can not be presumed, it is necessary to take into account the whole context of a particular experience and its personal meaning (Littlewood, 1990). This approach may be particularly relevant to those mental disorders which have a preponderance of neurotic or somatic symptoms and which seem to be unique clusters of symptoms and behaviour changes, for example, koro in Chinese communities, windigo found in the Ojibwa people of north Ontario, and brief reactive psychosis which occurs in South African "black" communities. These experiences make sense only within a particular context and within a particular culture, and appear to have no equivalent in other societies (Helman, 1990). If taken to extremes however, this "emic" view applied to the description of mental disorders in general, may not get beyond describing culture-bound syndromes and constructing inter cultural explanations. Relativists may thus be unable to contribute very much to cross-cultural understanding and assessment of mental disorder in non-western cultures.

The current standing of the universalist versus cultural relativist debate proposes that, while certain psychiatric concepts of mental illness may have a cultural basis, the illness behaviours associated with them, may have universalistic meaning, although these behaviours may require a cultural interpretation (Swartz, 1987, in Gaines, 1992).

The use of western nosology to define and measure severe mental disorders in non-western cultures

Most studies in non-western cultures have used psychiatric instruments designed in western countries, notably the USA and UK. While these instruments may have reliability and validity in the societies in which they were developed, there is no guarantee that they are appropriate when applied to persons from other cultures (Parry, 1992).

There is evidence, however, of certain universals in abnormal behaviour, especially extreme disturbances in conduct, thought or affect (Helman, 1990). While there is wide variation in their form and distribution, it seems that categories of illness developed in the west, and based on the categorical model, are probably usable for describing the more severe emotional disorders encountered in non-western societies, for example, schizophrenia, manic depressive psychosis and disorders arising from organic brain disease (Ferrando, 1988; Helman, 1990). The success of the International Pilot Study of Schizophrenia, IPSS, (1973) in which the PSE (Wing, 1974) reliably diagnosed acute schizophrenia (as defined in the west) in nine non-western cultures, lends support to the use of western nosology for diagnosing the major psychoses, particularly schizophrenia. On the other hand, even where there is evidence of universals in abnormal behaviour, presentations are usually influenced by the local culture in which they occur (Ferrando, 1988; Goldberg & Huxley, 1992; Helman, 1990).

The use of western nosology to define and measure common mental disorders in non-western cultures

Depressive illness and anxiety-related disorders are the disorders most likely to be encountered in the community (Goldberg and Huxley, 1992). In a review of psychiatric epidemiological studies of mental disorder in non-western communities, Parry (1992) confirmed that there is a higher prevalence of neurotic over psychotic disorders in these populations, with a large percentage of the neurotic population suffering from anxiety, mood (depression), and somatoform disorder. In general, these disorders are less severe, less distinctive, and more likely to be shaped by the culture in which they occur. It was noted in Chapter 1, that psychiatric instruments based on the categorical model may not be the most useful in measuring common mental disorders (Goldberg and Huxley, 1992). On the other hand, there is evidence of a common core of symptoms underlying nearly all forms of mental illness which persists regardless of culture (Cheng, 1985; Marie & Williams, 1985; Medina-Mora et al., 1983). If psychiatric instruments are shown to be sensitive to these aspects of mental disorder, it can be argued that such instruments adequately measure mental disorders in the community, regardless of cultural context. The similarity in prevalence rates shown in Table 3.1 gives credence to the argument that psychiatric instruments, developed in western cultures, do in fact measure a dimension of mental illness found in common mental disorders, regardless of the cultural context in which they are assessed (Cheng, 1985). Goldberg and Williams (1988) comment that neither "ethnic" nor cultural factors appear to affect the identification of mental disorders found in the community. Notwithstanding these arguments, it is important to bear in mind when selecting psychiatric case-finding instruments and assessing results, that "...these instruments will find what is universal and systematically miss what does not fit their tight parameters" (Kleinmann, 1977, p. 9).

Threats to validity of western psychiatric instruments when used in non-western cultures

Parry (1992) identifies two major threats to the construct/concept validity of western instruments when applied in non-western cultures. The first arises from cultural differences, and the second from the translation process.

1. Cultural problems which affect the validity of psychiatric instruments when applied cross-culturally:

   (i) Confusion between culturally distinctive behaviour and psychopathologial manifestations; for example, hallucinations and delusions in persons from western cultures are usually considered to be symptomatic of psychoses, yet in a number of African settings, hallucinations may occur among normal people as a result of encounters with ancestors who are believed to live in the same psychic world (Parry, 1992).

   (ii) Problems arising from the way in which the specific content of a psychiatric test may influence responses culturally. Many rural persons in Africa have no conventional concept of time, yet the subjective perception of time is a factor in determining a number of psychiatric conditions in many of the psychiatric questionnaires such as the PSE (Gillis et al., 1982).

   (iii) Problems arising from the format of psychiatric tests which may influence responses. The interrogative style of many clinical interviews may be foreign to the expectations and practises of those who converse in a more circuitous way (Gillis et al., 1982).

2. Translation problems which affect the validity of psychiatric instruments when applied in cross-cultural settings:

   Languages differ in the extent to which they allow for similar expression of inner distress (Swartz et al., 1985). This is particularly true of emotional problems, such as depression, irritability and anxiety. There is no guarantee that instruments developed in one cultural setting, operationalizing definitions of certain concepts characteristic of that culture, will find concepts, or those same operational definitions in another culture (Brislin, 1986; Chan, 1985).

Fortunately, there are measures which reduce the negative affects of the problems identified above, and these are discussed on p. 28.

Use in South Africa of psychiatric instruments based on western nosology and designed for use in western cultures

Issues which are pertinent to the use of western psychiatric instruments in non-western countries are equally relevant in South Africa, which has an heterogeneous population. In addition to having different socio-economic status, the various "ethnic" groups do not share the same cultural heritage. South Africa is a country in the process of transition. While "black", and to a lesser extent, "coloured" communities in South Africa have distinct cultural heritages, including interpretation of mental illness and pathway to mental care, these traditions constantly interact with the western model of mental health. Thus, while western definitions of psychiatric morbidity may be inappropriate in some contexts, they may be appropriate in others. In the past researchers of the "black" population of South Africa adopted the "noble savage" view of an African personality as "natural" and close to nature. This perspective negated the existence of the intellect, and postulated a view of "blacks" at a lower stage of evolutionary progression (Cheetham & Griffiths, 1980, in Swartz, 1987). Present trends in South African transcultural psychiatry accept that cultural differences exist, but at the same time, attempt to maintain an integrative perspective (Swartz, 1987).

The overwhelming majority of formally qualified mental health practitioners (ie, psychiatrists and clinical psychologists) in South Africa are "white". The interaction between interviewer and interviewee may therefore be significantly influenced by the power dynamics which prevail in a society which, for many years, has entrenched racial differences (Swartz, 1987).

Recommended measures to enhance the appropriateness of western psychiatric instruments when used in non-western developing communities

When selecting psychiatric epidemiological instruments it is important to consider not only their reliability and validity, but also their use in similar studies. This allows for a comparison of rates between studies. Parry (1992) recommends that certain measures be taken to enhance the appropriateness of these instruments when used in cultures for which they were not designed. These measures include:

(i) Those administering the psychiatric test must be conversant with abnormal behaviour patterns of the culture in which the test is to be used, as well as the socio-cultural context within which such behaviour occurs. Furthermore, they must be familiar with the manner in which mental illness is conceived, expressed and classified in the given context.

(ii) Culture-specific instruments must be developed where necessary. Adding questions about specific cultural phenomena, to existing instruments, or modifying questions to fit the cultural context should be considered.

(iii) Where appropriate, indigenous persons should be used as interviewers and interpreters.

In the translation process itself, measures may also be taken to insure that western instruments have a greater relevance to the culture in which they are to be used.

These include:

(i) Translators must have a knowledge of the subtle cultural norms and idiomatic expressions of the community.

(ii) Translated items must convey the **intent** which lies behind the question, rather than the content (Gough, 1978, in Brislin, 1986). Items may need intuitive rather than linguistic conversion. This can be achieved if translators have an intimate knowledge of the community and the underlying intent of the questionnaire, and if the translation process itself allows for the decentering from the original language in which the questionnaire is written.

A fuller account of the translation process may be found in Chapter 6, p. 59.

Conclusions regarding the use of psychiatric instruments designed in western cultures in non-western cultures and in South Africa in particular

Notwithstanding the possibility that western nosology may be usable in non-western cultures, use of imported psychiatric instruments may not be the best way of describing emotional disorder in non-western settings. This is especially applicable if nosology is to lead to an understanding of the illness and an approach to treatment (Ferrando, 1988). Western nosolgy may miss important aspects of illness and misunderstand the meaning of the illness in its cultural totality. If the ultimate aim of studying emotional disorder in non-western cultures is to benefit the people who are studied, understanding disorder only from the perspective of a western psychiatric framework, may not be useful, justifiable or even the most economical way of proceeding (Ferrando, 1988). These reservations need to be borne in mind when assessing results and making recommendations for future research in this area.

The development of culture-specific instruments would be the ideal solution in meeting the problems identified above. In South Africa, however both the urgent need for information and the cost involved in the design of such instruments, mean that compromises need to be made. The "sensitive" use of existing culture-general instruments constitutes one part of a broader, more complex strategy to understand psychopathology in this country.

4 REVIEW OF COMMONLY RESEARCHED SOCIO-DEMOGRAPHIC VARIABLES ASSOCIATED WITH MENTAL DISORDER

It has been suggested in previous chapters that accurate prevalence estimates of all levels of mental disorder in the community are essential to the planning and evaluation of mental health services in the community (Parry, 1992). Research that provides prevalence data alone has limited practical usefulness, however, as it says nothing about the variables which either increase or decrease an individual's vulnerability to distress. In order to plan and implement preventative and treatment programmes pertinent to specific communities, investigation of socio-demographic variables, likely to be associated with mental disorder, is important. In the present study therefore, assessment of prevalence was combined with an exploratory investigation of specific socio-demographic variables. Exploratory investigation was aimed at identifying possible risk factors, such as gender, age and employment status, relevant to mental disorder in the community. As the methods of data collection regarding these variables were such that would yield only correlation data between a demographic profile and mental disorder, no definitive conclusion regarding the association was anticipated. Nevertheless it was hoped that findings would point the way to future investigations of risk factors within the community.

COMMONLY RESEARCHED SOCIO-DEMOGRAPHIC VARIABLES AND THEIR ASSOCIATION WITH MENTAL DISORDER

Gender, age, marital status, education, employment, socio-economic status and physical health are the variables most commonly researched in relation to mental disorder. A review of the research findings regarding these variable is given below. Results show that in general, findings have not been consistent (Bebbington et al,. 1991; Chan, 1985; Cheng, 1988; Dube, 1970; Duncan-Jones & Henderson, 1978; Goldberg & Huxley, 1980, 1992; Vazques-Barquero et al., 1986). Conflicting results have meant that an indisputable socio-demographic profile invariably related to the level of mental illness found in the community, has not been identified. In the present study, it was therefore considered appropriate to investigate the variables most commonly associated with disorder.

Gender

Although several studies have shown a higher rate of mental disorder in women than in men (Finlay-Jones & Burville, 1978; Vazques-Barquero et al., 1986) the findings are not uncontested (Cheng, 1988; Henderson et al., 1979; Hodiamont et al., 1987; Newman et al.,

1988). When social adjustment for both sexes is comparable, the sex difference disappears altogether (Finlay-Jones & Burville, 1978). This suggests that studies showing a sex difference in rate of psychiatric morbidity are likely to be indicating a difference which has its origin in factors such as debilitating socio-economic circumstances rather than a biological difference between men and women (Goldberg & Huxley, 1992). Further to this argument, gender studies on physically healthy men and women indicate a sex ratio of mental disorder which is almost equal (Vazques-Barquero et al., 1987). However, there does appear to be a higher rate of disorder in women with physical illness (Goldberg & Huxley, 1992). Vazques-Barquero and his colleagues (1987) suggest that this difference may be understood in terms of the higher rate of consultation found in women, their willingness to acknowledge illness, to present with psychological symptoms, and to remember psychological symptoms more readily than men.

Age

Surveys regarding the association of disorder with age are generally inconclusive, (Bebbington et al., 1991). While some studies have found that disorder peaks in men and women between 30 and 40 years (Cheng, 1985; Finlay-Jones & Burville, 1978), others have found that it peaks only in men between 50 and 60 years (Hodiamont et al., 1987). Still other studies show no significant difference in rates of disorder between any of the age groups (Orley & Wing, 1979). Given this conflicting evidence it is inappropriate to make any confident predictions about the effect of age on disorder (Bebbington et al., 1991).

Marital status

According to Finlay-Jones and Burville (1978) disorder does not appear to be affected by marital status. This observation has however been contested, and other studies suggest that while divorced and separated men have about the same morbidity as single men, they have less morbidity than widowers, and married men appear to have the lowest morbidity of all (Bebbington et al., 1991). For both divorced and separated women there appears to be an increase in disorder which is also found in widows and widowers (Goldberg & Huxley, 1992).

Employment status

Although there is evidence that unemployment may be associated with depression, and may have a substantial effect on self-reported physical health status, somatization, and anxiety (Brown and Harris, 1978), generally, findings in this area are inconsistent. Some studies suggest that morbidity is higher for both unemployed males and females over 30 years

(Finlay-Jones & Ekhardt, 1984), while other studies suggest that this rate may not apply to unemployed females (Finlay-Jones & Burville, 1978; Hodiamont et al., 1987).

For women employment may, in certain circumstances, be associated with increased levels of disorder (Rosenfield 1989, in Goldberg & Huxley, 1992). For some women, a job means trading a situation of little control for another with low control (Rosenfield 1987, in Goldberg & Huxley 1992). For working women with young families, employment can impose even more demands (Bebbington et al., 1991). Cleary and Mechanic (1983) found higher rates of distress in employed married men and women than in unemployed housewives. On the other hand, studies in non-western societies, where most women share in the economic responsibility of the family, found that employed women had a relatively low rate of disorder (Cheng & Williams, 1986; Dube, 1970).

Social class

Since people in lower social class are likely to encounter more adverse and fewer supportive experiences and to be less in control of their environments, one can expect a higher rate of disorder in the lower social classes (Bebbington et al., 1991). Most research supports this view, indicating a higher rate of mental disorder especially in the semi-skilled and unskilled section of the population (Finlay-Jones & Ekhardt, 1984; Hodiamont et al., 1987). However, in a study conducted by Finlay-Jones and Ekhardt (1984) the highest rate of disorder was found in clerical workers. In surveying a number of studies investigating an association between mental disorder and socio-demographic variables, Bebbington et al., (1991) conclude that social-class may not be a significant factor in determining mental disorder.

Race

Generally studies have shown a higher rate of disorder in "blacks" than "whites", (Goldberg, et al., 1976). It has been argued that this difference persists regardless of social class, so that lower-status "blacks" are more vulnerable to the impact of economic problems than lower-status "whites" (Kessler & Neighbours 1986). These results have been contested however in other studies which have shown that when social class is controlled, the difference in rate of disorder between races disappears (Hirschfeld & Cross 1982). What seems clear, is the importance of recognizing that life events may have different impacts on different cultures (Goldberg & Huxley, 1992).

Physical environment

Adverse living conditions, such as over-crowding, have been shown to be associated with mental disorder (Goldberg & Huxley, 1992). Studies in urban areas, where these conditions

may be more prevalent, have found that the rate of mental disorder is indeed high (Hodiamont et al., 1987; Vazques-Barquero et al., 1986). On the other hand, studies in rural areas, where these conditions are likely to be less prevalent, have also found a high prevalence of mental disorder (De Jong et al., 1986; Miller et al., 1991). The lack of readily available phc services in these areas may however, be an explanation for the high rate.

Physical illness

Research findings indicate that psychological illness is more likely, in both sexes, when physical ill health is present, particularly certain chronic conditions (Weyerer, 1990). There is however, much ambiguity about the extent and nature of the association, and disentangling cause from effect has proved to be problematic (Goldberg & Huxley, 1992). According to Aneshensel et al., (1984, p. 109), "...there may be a possible connection between physical morbidity and depression which is self-perpetuating and mutually reinforcing, and in which both biological and social mechanisms are involved".

Goldberg and Huxley (1992) suggest that only a small minority of mental disorders are secondary to physical disease in community and primary care settings; for example, certain cases of depression are caused by particular drug therapies used in patients with hypertension (Goldberg & Huxley, 1992). The Department of Community Health, Cape Town (1990) established that there is a high prevalence of hypertension in Mamre, the community selected for study in the present research (see Chapter 5, p. 35). Investigation of a possible association between psychiatric caseness and persons using depression-inducing drug treatments for hypertension, is therefore relevant when assessing results of the present study.

Prevalence estimates and their interpretation are also affected by the relationship between consultation with phc services and the presence of mental disorder. While this is particularly pertinent when estimating prevalence of disorder in clinic settings, information concerning an association between consultation and mental disorder is also important to the general understanding of mental disorder in the community. In addition, this provides useful data for the planning of health care services. Discussion regarding the compilation of questions to test this relationship is found in Chapter 6, p. 58.

According to Goldberg and Huxley (1992) there is a direct relation between degree of psychological distress and the probability of consulting with phc services for both men and women. Women however, are more likely than men to consult at all levels of distress, although there is no evidence to suggest that there is an interaction between sex, psychological disorder and probability of consulting health care services (Goldberg & Huxley, 1992).

CONCLUSION

Discussion in Chapters 1 to 4 has focused on factors in favour of conducting a community based study (such as the high rate of mental disorder at community level); issues which are problematic to such a study (such as psychiatric case definition); and factors which would increase the usefulness of the study (such as the need to identify possible risk factors in the community). Issues which motivated the research study, and its participation in a programme to improve the general health status of the community of Mamre, are discussed in the following chapter, which also sets out the initial aims of the present research.

5 INITIAL AIMS

This chapter includes:

(i) a review of the general goals, specified by the over-arching Mamre Community Health Project, (MCHP), of which the present research forms a part;

(ii) a presentation of the initial aims of the present study; and

(iii) a schematic representation of the factors which led to the decision to conduct a second pilot study and to the reformulation of the initial aims.

The current research study forms part of a much larger research project. In 1986, the Department of Community Health (DCH), University of Cape Town, in collaboration with the Centre for Epidemiological Research in Southern Africa (CERSA) began a long-term community study aimed at upgrading the general health status of a selected community in South Africa. This project came to be known as the Mamre Community Health Project (MCHP), and the motivation for it reflects international trends in public health awareness policy: ie, to develop and enhance the role of the Public Health Sector in health care. This approach recommends active interaction between the selected community and phc services, both in the assessment of community mental health requirements, and in the implementation of preventative or treatment programmes. By encouraging community involvement in all stages of research, planning and implementation, it is hoped that phc services will become more relevant to the communities they serve. Furthermore, this approach empowers communities to participate in meeting their own health requirements.

THE DECISION TO CONDUCT THE PRESENT RESEARCH STUDY IN THE COMMUNITY OF MAMRE

Mamre, an Afrikaans "coloured" community, 50 km outside of Cape Town, was selected as the most suitable community in which to situate the present study. Several factors made this community particularly appropriate:

(i) Mamre has a fairly stable working-class population, which sees itself as historically part of a settled community. Conducting a study in a stable environment is clearly easier than in less settled communities. Results can also be more readily interpreted and applied.

(ii) Mamre is a community undergoing transition from rural to urban status, with infrastructures such as sewage and electricity currently being supplied to some of the homes (Katzenellenbogen et al., 1988). Communities experiencing transition are particularly interesting, as factors commonly occurring in industrialized communities, and which may be associated with disorder, may more readily be assessed.

(iii) Mamre's locality makes it accessible to researchers from Cape Town.

In line with the commitment to stimulate community participation, consent for the MCHP was sought from the Mamre Village Management Board, prior to the beginning of the study in 1986. Next, a project steering committee, consisting of Mamre residents and representatives from University of Cape Town, CERSA and the Divisional Council of the Cape, (DCC), was elected (Katzenellenbogen et al.,, 1988). Their agenda was to review and approve future research projects in the community.

In 1986, a general health survey was conducted in Mamre to provide base-line information on the community's health status (Hoffman et al., 1988). Initial findings suggested that 2.9% of the population were suffering from "nerves", a broad non-specific term implying psychological distress (Katzenellenbogen et al., 1988). On the basis of these findings, the Department of Psychology at University of Cape Town was approached to further investigate the nature and extent of mental disorder in the community.

An initial study by University of Cape Town's Psychology Department was conducted among clinic attenders at phc services in Mamre (Miller et al., 1991), including those persons attending a GP in private practice, a district surgeon and a day hospital. The GHQ-28 was used as the first stage screen, and a random 20% of GHQ-28 high scorers were administered an unstructured clinical interview. Estimates from this study need to be treated with caution because of the use of an unstructured clinical interview, the lack of power analysis to calculate optimal second-stage sample size, and the use of a screening instrument to estimate prevalence. Nevertheless, estimates based on persons scoring above 4/5 on the GHQ-28, indicated a high prevalence rate of 45.28%, suggesting that nearly half of the clinic attenders were psychologically distressed (Miller et al., 1991). These results indicated that there might be a high rate of mental disorder in the community, which was not being detected, even at clinic level (Miller et al., 1991). The suspected high rate of disorder in the clinic population, the urgent need for data relevant to the mental health status of the community as a whole, and the inappropriateness of using clinical data to estimate prevalence of mental disorder in the community, pointed to the need to conduct a broad age-range community based study in Mamre.

Participation in the Mamre Research Project had two important general implications for the planning and implementation of the present study. In specifying that research should include the participation of the community in the planning and implementation of the study, and that the overall goal should be to uplift the health status of the community, the objectives of the Mamre Community Health Project informed not only the specific aims of the present research, but also the manner in which the research study was conducted. Aims were therefore considered in terms of their benefit to the community, and not only in terms of academic insight. Where possible, community participation was encouraged rather than passive acceptance.

INITIAL AIMS

(i) to conduct a community-based prevalence study of psychiatric morbidity in Mamre, using a two-stage research design method of case identification, in which the GHQ-28 was used as the first-stage screen and the PSE as the second-stage criterion measure;

(ii) to assess any association between mental disorder and selected socio-demographic variables, ie, gender, age, marital status, education, employment, socio-economic status, support systems, alcohol consumption and substance abuse;

(iii) to obtain information on health care utilization; specifically, to assess a possible association between physical health status and mental disorder, and mental health status and rate of consultation with phc services. Also to obtain information relating to attitudes to health services currently available in Mamre; and

(iv) to assess the validity of the GHQ-28 as a first-stage screen for mental disorder when used with the PSE as the second-stage criterion, in a community sample in South Africa.

Attention is drawn to the fact that these were **initial** aims. In the course of the first pilot study, events occurred which necessitated their reformulation (see Chapter 7, p. 68). Chapters 6, 7, 8, and 9 will show how and why the reformulation took place, In order to help orient the reader to the structure of the argument in those chapters, however, the current chapter ends with a schematic overview of the processes of reformulation of the aims which is given in Table 5.1.

Table 5.1 Diagrammatic overview of Chapter 5 and Chapters 7 to 9: factors leading to the reformulation of initial aims

INITIAL AIMS (Chapter 5)

1. To estimate prevalence of mental disorder in the community of Mamre.
2. To validate the GHQ-28 as a first-stage screen in a two-stage research design, in which the PSE is used as the second-stage gold standard.
3. To assess any possible association between mental disorder and specified socio-demographic variables.
4. To obtain basic information on health care utilization and attitudes to health services currently available in Mamre.

FIRST PILOT STUDY (Chapter 7)

Methods:	
Design:	Cross-sectional descriptive study.
Instruments:	A two-stage research design in which the GHQ-28 is used as the first-stage criterion measure.
Sample:	35 adults randomly selected from the total Mamre population frame.
Results:	Unacceptably low validity coefficients on the GHQ-28
Conclusions:	Need to conduct second pilot study in which alternative first- and second-stage measures are included.

SECOND PILOT STUDY (Chapter 8)

Methods:	
Design:	Cross-sectional descriptive study.
Instruments:	Both GHQ-28 and SRQ are used in a cross-over design as first-stage screens. The PSE is combined with the PAS in a single second-stage criterion measures along with the "suggested clinical diagnosis" of the clinician administering the interview.
Sample:	35 adults randomly selected from the total Mamre population frame.
Results:	Highly satisfactory validity coefficients for both the SRQ and the GHQ-28 when used with all three criterion measures but especially when used with the PSE.
Conclusion:	Results good enough to proceed with main study. In the main study, the SRQ to be used as the first-stage screen and the PSE and the second-stage criterion. The GHQ-28, the PAS and clinician's "suggested clinical diagnosis" to be included in the main study to provide data on the appropriateness of psychiatric instruments to assess mental disorder in the community, in the South African context.

REFORMULATED AIMS (Chapter 9)

	Overall aim:	To conduct a cross-sectional descriptive study and at the same time, to ensure a design which will yield data sufficient for an analytic assessment.
1.	Primary aims:	
	(i)	To estimate the prevalence of mental disorder in the community of Mamre.
	(ii)	To validate the SRQ as a first-stage screen in a two-stage research design, in which the PSE is used as the second-stage gold standard.
	(iii)	To assess and possible association between mental disorder and specified socio-demographic variables.
	(iv)	To obtain basic information on health care utilization and attitudes to health services currently available in Mamre.
2.	Secondary aims:	
	(i)	To investigate the appropriateness of the PSE as a case-identifying instrument in a community study in South Africa.
	(ii)	To assess the effect on prevalence rate of mental disorder when three different psychiatric measures are used, and
	(iii)	To investigate any possible changes in the validity coefficients of the SRQ when the SRQ is used with any on the three criterion measures, the PSE, the PAS and "suggested clinical diagnosis".

6 OVERALL DESIGN AND GENERAL METHODOLOGY

The study's overall design and general methodology are reviewed in the chapter. The section on general methodology includes discussion on:

(i) the selection of a random sample from the adult population of Mamre;

(ii) the selection of first- and second- stage instruments in a two-stage case-finding design;

(iii) the compilation of additional questions to be included in the first-stage questionnaire;

(iv) the translation of instruments;

(v) the selection and training of personnel; and

(vi) ethical considerations.

The overall research design and general methodology remain the same for the first and second pilots as well as the main study.

RESEARCH DESIGN

The design was a cross-sectional descriptive study in which a two-stage case identifying method was selected to estimate the prevalence of mental disorder in the community.

The two-stage case-finding method used here is based on the assumption that the second-stage criterion measure is error free, and therefore an appropriate gold standard against which the validity of the screen may be assessed. Although the correctness of this assumption will be discussed later in this chapter, it is given here as an intrinsic part of the research design.

GENERAL METHODOLOGY

SAMPLE

Description

The appropriate sampling frame was identified as the adult population of Mamre.

The size and demographic composition of Mamre's adult community had been appraised by the Department of Community Health in a general population survey in 1986 (Hoffman et al., 1988). A later study, up-dating the 1986 Mamre census, found little change in the base-line demographic information (Azevedo & Gilbert, 1991). Accordingly, the 1986 Mamre census, containing a comprehensive list of names and addresses of Mamre residents, was used as a population sampling frame for the present research. As age is a continuous variable on the computer population data-base, it was automatically updated to the 6 year interval.

In terms of the 1986 census:

(i) the ratio of men to women in the community was very nearly one to one (49% males and 51% females);

(ii) the age distribution was typical of a developing community, with a broad 15-24 year age base tapering off steeply to the 70+ age group (Hoffman et al., 1988);

(iii) there were more women than men in the 55-64 year group;

(iv) the majority of adults had secondary education up to the level of std. 6;

(v) 65% of adults were employed. By 1992 however, the economic down-turn in the South African economy had resulted in a reversal in the employed/unemployed figures in Mamre, so that approximately 64% were unemployed and only 36% were employed; and

(vi) of those employed, 75% were in the semi-skilled and unskilled employment categories.

Random sampling

Using the SAS computerized random selection programme, samples for both pilot studies and the main study were randomly selected. A randomly selected back-up list of names

was also drawn to be used where selected respondents were not available, or declined to be interviewed.

Decision not to stratify sample

It was decided not to stratify the sample given the inconclusive results regarding a possible association between mental disorder and commonly researched socio-demographic variables (see Chapter 4). Support for this decision with respect to gender was found in Miller et al's., (1991) clinic study in the same community, in which an almost equal rate of men and women were found to be suffering from mental distress.

Calculation of sample size

Calculation of sample size for the main study is described in Appendix 3.1).

INSTRUMENTS

Broad advantages of selecting a two-stage method of case identification have already been enumerated in Chapter 2, p. 17. This method of case detection was felt to be the most appropriate in the present research for the following reasons:

(i) Shrout (1992) recommends it as the "optimal" design in general population prevalence studies where neither the sensitivity, nor the specificity of the first-stage screen is known in the context in which it is to be used; and

(ii) in estimates calculated prior to beginning the study, the cost of administering the screen to the total sample, in terms of rands and hours, was likely to be approximately 0.025 that of the administration of the criterion to the total sample. (see Appendix 3.1, Table 3.1.1)

The General Health Questionnaire, version 28, (GHQ-28), (Goldberg & Hillier, 1979) was selected as the first-stage screen and the Present State Examination, PSE (Wing et al., 1974) as the second-stage criterion.

First-stage screening instrument: the General Health Questionnaire, GHQ-28.

There are several versions of the GHQ. As the principles which govern the rationale for its design, its strengths and weakness apply to all versions, it is appropriate to begin with a

discussion of the GHQ in general. This is followed by a description of the GHQ-28, and the reasons for its selection as the first-stage screen in this study.

Rationale for the construction of the GHQ

The GHQ (Goldberg, 1972) was designed specifically to take account of the transient nature of mental distress in clinic and community populations. At the same time, it seeks to measure a core of symptoms which are believed to be common to most mental disorders, both the severe and less severe, such as depression, anxiety, somatic disturbance, fatigue and irritability (Goldberg & Huxley, 1980). By identifying and measuring those factors which represent the lowest common denominator of mental disorder, the screen is designed to be used at all levels of psychiatric morbidity.

The GHQ focuses on signs and symptoms of mental disorder which are usually distressing, and signify a break in the normal continuity of behaviour (Goldberg, 1972). As such, the questionnaire has two fundamental concerns:

(i) the inability of the respondent to carry out his/her usual daily functions; and

(ii) the appearance of new phenomena of a distressing nature (Goldberg & Williams, 1988).

Respondents are asked to compare the extent to which they experience each item in the **present** with the extent to which they **usually** experience it (Goldberg, 1972). This privileges symptoms over personality traits, which are of longer duration (Goldberg & Williams, 1988).

The GHQ was also designed to address the need to screen large numbers of persons economically and in the minimum amount of time.

Item selection

Initial items for the GHQ-60 were selected from earlier investigations on a non-hospitalized sample of 2460 persons, by Gurin, Veroff and Feld (1960, 1962, in Goldberg, 1972); from key items in the Cornell Medical Inventory (Brodman et al., 1949, in Goldberg, 1972); and from input from a body of psychiatric clinicians who had considerable experience with psychiatric patients.

The criteria behind item selection were :

(i) that all items should be applicable to the general population, i.e., to those defined as mentally healthy as well as unhealthy;

(ii) that the emphasis should be on changing aspects of behaviour; and

(iii) that only items measuring observable behaviours should be included (Goldberg, 1972).

The final version of the GHQ was calibrated on 3 different groups: severely disturbed hospital patients, moderately disturbed hospital patients, and psychologically healthy persons in the general population.

The original GHQ comprises 60 items. Items are both positively and negatively phrased, for example, "Have you recently felt on the whole you were doing things well?", and "Have you recently been thinking of yourself as a worthless person?" It has been suggested that positively and negatively worded items discriminate differently, with positively worded items assessing more transient states, and negatively worded items those that are more stable (Duncan-Jones, 1979). Positively worded items also tend to discriminate less accurately than negatively worded items (Elton et al., 1988).

Setting the threshold on the GHQ

The GHQ does not have a fixed threshold which distinguishes normal from abnormal behaviour. Mental illness is considered to be continuously distributed in the population (Angst & Dobler-Mikola, 1984; Goldberg & Huxley, 1992) so that the imposition of a threshold places a dichotomy upon a continuum of severity (Goldberg & Huxley, 1980). In prevalence studies where a distinction between psychiatric caseness and non-caseness is necessary, setting a threshold may be practically useful. However, it should be regarded as arbitrary. Setting a threshold appropriate to the particular goals of a research project, is determined in a pilot study conducted with a sample of the population to be researched. This involves a comparison of GHQ scores with an independent psychiatric instrument used as the gold standard. The comparison computes the best validity coefficients at all possible thresholds, by applying Receiver Operating Characteristics (ROC) (Metz, 1978). (see Appendix 1.1). The optimal threshold in a particular context is usually a compromise between the screen's high sensitivity and low false positive rate, at all possible threshold levels (Riegleman, 1981).

In a review of 72 validity studies of the GHQ, Goldberg & Williams (1988) found that the threshold for the original 60 item version as well as the 12 item version was stable at a cut-off of 4/5. While that of the 28 item version was fairly unstable, ranging from 1 to 13, with a modal value of 4/5.

Response Scale of the GHQ

The GHQ differs from most other screening instruments in its four-point response scale. To each question the respondent is required to select a response from a four-point scale

ranging from, "less than usual" to "much more than usual". This format focuses the respondent's attention on his/her current physical or mental health status.

Scoring procedure of the GHQ

The GHQ employs a distinctive scoring procedure. Scoring may be carried out by means of two different scoring methods known as the GHQ and the CGHQ.

(i) The GHQ condenses the 4 point response scale into a binomial measurement in which a zero score is assigned to the first two response categories, and a score of 1 is assigned to the last two:

0, 0, 1, 1. Only the last two response categories are treated as indicators of pathological deviation from the respondent's usual pattern of behaviour, even in the case of negatively worded items.

(ii) The CGHQ however, retains the original scoring method for all positive items, but applies a scoring system of 0, 1, 1, 1, to all negatively worded items. Thus a response of "no more than usual" records the presence of a symptom. This is particularly useful when measuring populations with a high rate of long-standing illness, where the response, "same as usual", is likely to obscure chronic illness (Goodchild & Duncan-Jones, 1985).

The GHQ score considers only the number of symptoms, and therefore can only yield an **area** score as opposed to the more conventional Likert 0, 1, 2, 3, method, which gives a composite measure indicating both area and intensity. Goldberg (1972) argues that the GHQ's overall count of symptoms gives results which are as accurate as, if not better than, the more complex Likert method.

The GHQ method of scoring has several advantages:

(i) it is simple;

(ii) it eliminates the error of central tendency;

(iii) it eliminates "end" and "middle" users, since a score in either the first or second column, signifies non-possession of the item, and a score in either the third or fourth column signifies possession;

(iv) it reduces bias associated with binomial response scales, in other words, "yes" and "no" sayers, by the random allocation of items indicating illness to the left and right of the questionnaire; and

(vi) it eliminates the problem of between row scoring, in other words, the subjective value given by the respondent to the interval between 1 and 2, or between 3 and 4, which is typical of the Likert method.

Reliability of the GHQ

Reliability refers to the ability of an instrument to measure some quality in the respondent in a consistent way (Goldberg, 1972). This is important, as imperfect reliability attenuates correlation coefficients between the GHQ and criterion measure (Goldberg & Williams, 1988).

The two most common methods of assessing reliability are test-retest, and split-half reliability testing. The test-retest method poses methodological problems when applied to all versions of the GHQ, while the split-half method is problematic when applied to the GHQ-28 in which positively and negatively worded items are not equally weighted. The test-retest method is difficult to apply as the GHQ is designed to measure a highly variable quality; present or current functioning, which may remit spontaneously within a very short time, or remit as a result of the administration of the test itself may (Goldberg, 1972). Any change at re-test therefore, may reflect remission, but may equally reflect unreliability of the screen itself. Nevertheless, some studies have used this method to assess reliability of various versions of the GHQ, and have found test-retest estimates to range from 0.51 to 0.90 (Vieweg & Hedlund, 1983).

In a split-half reliability study of the GHQ-60 (Robins & Price, 1982, in McDowell & Newell, 1987) a correlation of 0.90 was found which is comparable to the range of 0.90 to 0.95 found in other split-half reliability studies for the GHQ-12 and GHQ-30 (McDowell & Newell (1987).

According to Shrout et al., (1981, in Wing et al., 1981), if a screening test is to be evaluated in terms of sensitivity and specificity, its reliability should not be a major criterion, as long as it is moderate. Although sensitivity is negatively affected when reliability is imperfect, the effect on specificity appears to be negligible (Robins & Price, 1982, in McDowell & Newell, 1987). The overall effect of imperfect reliability on the validity coefficients of the screen is therefore relatively small (Shrout & Fleiss, 1981, in Goldberg & Williams, 1988).

Validity of the GHQ

Three measures which indicate the validity of a test include; construct, content and criterion or concurrent validity.

Construct validity refers to the extent to which a test measures a particular theoretical construct, in this case, mental disorder, and may be assessed by principal component

analysis (Elton et al., 1988; Neale & Liebert, 1980). Such studies have reported a large general factor found in the GHQ with a pattern of item loading, ranging from a kappa correlation of 0.60 to 0.85, across different versions of the questionnaire (Duncan-Jones, 1979).

Content validity assesses whether the content of questionnaire items is representative of the variable being measured (Neale & Libert, 1980). Construction of the GHQ virtually guarantees its content validity; and the calibration study of the GHQ-60 has demonstrated the ability of items to discriminate between groups of respondents ranging from normal to severely disturbed (Goldberg, 1972).

Most studies of the GHQ have concerned themselves with its **criterion** validity. These assess the GHQ's ability to yield scores which are comparable to scores on an external measurement or criterion which is assumed to be error-free (Goldberg & Williams, 1988). The major draw-back with this method is that there is no independent or error-free criterion measure which can be shown to be error free (Tarnopolsky et al., 1979; Williams & Tarnopolsky, 1980).

Two measures used to assess the criterion validity of the GHQ are those of sensitivity and specificity. (see Appendix 1.1). In a review of 43 validity studies, Goldberg & Williams 1988 found the variance weighted mean (VWM) sensitivity for the GHQ to be 0.76 (asymmetric 95% confidence limits 0.74 and 0.78) and the variance weighted mean specificity to be 0.85 (95% confidence limits 0.84 and 0.86).

Two other studies, however, have found the GHQ to be an inadequate measure of psychiatric morbidity (Benjamin et al., 1982; Tarnopoplsky et al., 1979). Assessing the performance of the GHQ in terms of its false positive and false negative rates, these studies found that the GHQ-60 and GHQ-30 had an *overall misclassification rate* of between 25% and 50% of psychiatric cases. The low validity coefficients found in these two studies may however, be the result of the unacceptably long time-lag between first and second stage interviews (Cheng, 1985).

The GHQ has been extensively tested among different groups of sample populations, in different cultural settings and against a variety of different criterion measures (Goldberg & Williams, 1988; McDowell & Newell 1987; Skuse, 1984). Almost without exception, its validity coefficients have been satisfactory with a kappa correlation between itself and the criterion measure ranging between 0.45 and 0.83. The median correlation coefficient of the GHQ-28 is 0.76 (Goldberg & Williams, 1988).

In general, there is overwhelming evidence to suggest that all versions of the GHQ are feasible and are useful as screening instruments for psychiatric morbidity. Furthermore,

the GHQ has a higher degree of validity and consistency of validity coefficients than most rival instruments (Boardman 1987; Cheng 1988; Chong & Wilkinson 1989; Duncan-Jones et al., 1979; Finlay-Jones & Murphy 1979; Fontensesi et al., 1985; Hodiamont et al., 1987; Krause et al., 1990; Lobo et al., 1986; Mcdowell & Newell, 1983; Medina-Mora et al., 1983; Pan & Goldberg 1990; Robins 1985; Skuse & Williams 1984; Vazques-Barquero et al., 1990).

Factors which may threaten the validity of the GHQ

Goldberg and Williams (1988) list several factors as possible threats to the validity of the GHQ. These include:

(i) **Time delay between first and second stage interviews**

As the GHQ measures changeable mental states, it is important that the second-stage clinical interview follows the first-stage screen within the shortest possible time (Goldberg, 1972). Cheng (1988) found that validity coefficients, especially sensitivity, were significantly improved when the clinical interview followed the screen within one week. The longer the time interval the more likely it is that false positives will be recorded, as some conditions may remit in the interval between the interviews.

(ii) **Effects of language and culture**

The possible effects of language and culture on psychiatric instruments have been discussed in Chapter 3, p. 27. The GHQ has been translated into over thirty languages and used in a number of cultural settings (Goldberg & Williams, 1988). In a meta-analysis of these studies, Williams et al. (1987) found a VWM sensitivity of 78% in non-English studies, and 74% in English studies, and a VWM specificity of 85% and 86% respectively. These results suggest that cultural factors play only a small part in the identification of a core of underlying symptoms found in most psychiatric disorders (Chan, 1985; Cheng, 1985, 1988; Goldberg & Williams 1988). All the above validity coefficients however, were estimated against clinical interviews which were themselves designed in western countries. This may have positively influenced the extent of the agreement.

(iii) **The effects of long-standing disorders**

As discussed before, the design of the GHQ response format means that long-standing disorders may be missed (Goodchild & Duncan-Jones, 1985). As chronic disorders are usually found to represent about half of those contributing to point prevalence estimates, in both general population and

primary care settings, this may have an impact on the validity of the GHQ (Goldberg & Williams, 1988). In populations with a high rate of physical illness, the alternative CGHQ method of scoring is recommended (Goodchild & Duncan-Jones, 1985).

(iv) **Community versus clinic studies**
Validity studies of the GHQ indicate that sensitivity may be lower in community as opposed to clinic or hospital samples. Specificity on the other hand, does not appear to be affected by the context in which the study occurs (Goldberg & Williams, 1988).

(v) **The effect on the validity of the GHQ when different criterion measures are used**
In a meta-analysis of 22 studies in which the CIS was used as the criterion interview, and 12 studies in which the PSE was used, Williams et al. (1987) found that the VWM sensitivities of the GHQ were within 1% of each other and the VWM specificities within 3% of one another. These results suggest that the choice of criterion interview does not significantly affect the validity of the GHQ.

Advantages of the GHQ

The GHQ has several features which make it suitable as a first-stage screen:

(i) it is easy to administer and can be administered by lay-interviewers;

(ii) interviewers are not required to make any subjective assessment about the respondent;

(iii) large numbers of people can be interviewed economically;

(iv) it is acceptable to both mentally healthy and mentally disordered persons, and therefore suitable for use in community studies;

(v) its response format and scoring procedure may yield superior validity coefficients to those of other screens (Goldberg & Williams, 1988);

(vi) items are scored positively only if the behaviour is considered by the **respondent** to be a departure from his/her usual behaviour. An individual's perception of himself/herself as unable to maintain daily functioning, may be the best discriminator of psychiatric illness (Goldberg, 1972); and

(vii) it has been widely used internationally, in different cultural contexts and has been found to be valid (Goldberg & Williams, 1988).

Limitations of the GHQ

The GHQ does, however, have three important limitations:

(i) it is not intended to detect personality disorders, or long-standing functional psychoses, such as schizophrenia or psychotic depression. One study however, has indicated that the GHQ may be sensitive to psychotic disorder (Lobo et al., 1986),

(ii) it detects disorder of less than two weeks duration. This may affect the screen's validity coefficients, given that most second-stage criterion instruments, (such as the PSE and the SADS) and diagnostic systems (such as the ICD and the DSM-111), require symptoms to have been present for at least two weeks or longer; and

(iii) questions relating to substance abuse are not included.

Selection of the GHQ-28 as the first-stage screen

Rationale for selecting the GHQ-28 as the first-stage screen

The GHQ-28 was selected for the following reasons:

(i) it is short, and therefore appropriate in a semi-urbanised community such as Mamre, where the level of literacy may be expected to be fairly low. Brevity was an advantage, as the questionnaire was to be verbally administered by lay-interviewers from the community;

(ii) in spite of the reduced number of items, the GHQ-28 out-performs the GHQ-30 and GHQ-60 with regard to sensitivity. In a comparison of VWM sensitivity and specificity from 43 validity studies, Goldberg and Williams (1988) found the GHQ-28 VWM sensitivity to be 84 % - higher than either the GHQ-30 (VWM 74 %) or the GHQ-60 (VWM 78%). Specificity of the GHQ-28 was equal to that of the GHQ-30 (VWM 82 %). Although the GHQ-60's specificity was higher (VWM 87%), this has to be considered in terms of its greater length;

(iii) an Afrikaans version of the GHQ-28 has been calibrated on a "coloured" in-patient sample in South Africa (Isaacs, 1990);

(iv) it has been validated in over 55 validation studies (Goldberg & Williams, 1988);

(v) it has been translated into over 36 languages with good results; and

(vi) it had already been used in a study of mental disorder in a clinic sample in Mamre (Miller et al., 1991). As one of the future goals of the Mamre Community Health Project is to investigate a possible association between consultation/non-consultation and mental disorder, by comparing the GHQ scores of clinic and community respondents, it was necessary that the same screening instrument be used in both studies.

Description of the GHQ-28

A "scaled" version of the GHQ, the GHQ-28, was designed to increase the amount of variance accounted for by symptoms of depression, anxiety, somatic disturbance, fatigue and irritability (Goldberg & Hillier, 1979). Four sub-scales were identified, using principal components analysis on the entire 60 item version:

A scale, somatic symptoms;

B scale, anxiety and sleep disturbance;

C scale, social dysfunction; and

D scale, depression (Goldberg & Hillier, 1979).

The four dimensions underlying the GHQ-28 have been found to be relatively consistent across samples (McDowell & Newell, 1983). This indicates the existence of a core of psychopathological symptoms which cuts across diagnoses, and which appears to be common to nearly all forms of mental pathology in all cultures (Duncan-Jones et al., 1990; Eaton et al., 1988; Goldberg, 1979; Goldberg & Williams, 1988; Leon, 1972).

While these subscales measure different aspects of mental pathology, they are not independent of one another (Goldberg & Williams, 1988). Correlation between the B scale, (anxiety and sleep disturbance) and the total score suggests that anxiety may be a core phenomenon underlying the common syndromes of psychiatric disorder (Goldberg & Hillier, 1979).

Reliability of the GHQ-28

Problems in establishing reliability of the GHQ and particularly the GHQ-28, have already been discussed. As the GHQ-28 was constructed on the same principles as the GHQ-60, it may be assumed that its test-retest reliability has similar satisfactory reliability.

Validity of the GHQ-28.

Comparison between validity coefficients of the various versions of the GHQ is shown in Table 6.1.

Table 6.1 Comparison of the variance-weighted mean (VWM) *validity coefficients*[1] of the GHQ-28 with the other versions of the GHQ found in 43 validity studies. (Table adapted from Goldberg & Williams, 1988, p.54)

GHQ version	Sensitivity %	Specificity %
GHQ-12	89	80
GHQ-28	84	82
GHQ-30	74	82
GHQ-60	78	87
All	78	85

In a review of 12 validity studies, in which the GHQ-28 was used with either the PSE or the CIS the VWM sensitivity ranged between 0.44 and 1, and VWM specificity between 0.74 and 0.93 (Banks, 1983; Bridges & Goldberg, 1986; Elton et al., 1988; Goldberg & Hillier, 1979; Krause et al., 1990; Lobo, 1986; Mann et al., 1983; Medina-Mora et al., 1983; Rabins & Brooks, 1981). A stronger association was found between the GHQ-28 and the PSE than between the PSE and either the GHQ-12 or GHQ-30 (Banks, 1983). The GHQ-28 may therefore be almost as sensitive in discriminating cases from non-cases as the GHQ-60 (Lobo et al., 1986; Medina-Mora et al., 1983).

1 See Glossary of Technical Terms, Appendix 1.1.

Second-stage criterion measure: The Present State Examination (PSE)

Rationale for the construction of the PSE

The PSE is a semi-structured clinical interview (Wing et al., 1974). It was designed in response to the need for a psychiatric measuring instrument that would standardize the collection of information, concerning certain broad classes of psychiatric conditions, according to clinical consensus (Luria, 1979). The aims were to describe clinical phenomena clearly, precisely and reliably, and to facilitate the investigation of diagnostic rules and practices (Wing, 1983). The PSE was not intended to serve as, or to be equated with a clinical diagnosis (Wing, 1983).

Description of the PSE

The PSE was designed on and for use with clinical samples. Essentially, it is a check-list of 140 questions which systematically cover all the psychiatric phenomena likely to be considered during a mental state examination (Grayson et al., 1987), excluding tests for cognitive impairment. The PSE assesses the presence or absence of symptoms within the general area of the functional psychoses and neuroses, through a process of cross-examination used by a clinician. The emphasis is on current mental state, and the interview is predominantly concerned with symptoms experienced during the previous month.

For most items, a form of questioning is suggested. However, the interviewer is free at all times to depart from the format and pursue his/her own line of enquiry. Several measures are employed to increase standardization:

(i) questions are structured;

(ii) probes are suggested if the interviewer requires further clarification;

(iii) a system of optimal cut-off points allows the examiner to skip certain questions if he/she believes that they are irrelevant, particularly in cases where the respondent is clearly non-pathological;

(iv) guidelines as to how each item should be scored are given;

(v) an accompanying glossary provides definitions of symptoms, as well as a technique for clinical examination and instructions for rating the presence and severity of each symptom.

PSE symptoms can be clustered into syndromes, giving a more manageable profile, and syndrome scores can be summarized to give various kinds of sub-totals as well as a total score (Wing, 1983).

A computer programme, the Catego, (Taylor & Gourlay 1974, in Wing et al., 1974) was designed specifically for use with the PSE. The programme incorporates diagnostic rules which give a diagnosis according to the ICD8, for patients who have enough symptoms to be regarded as psychiatrically ill (Wing, 1983). The Catego, when applied to a PSE symptom-profile, indicates an Index of Definition, (ID), which allocates each profile to one of 8 levels of confidence that a specific psychiatric disorder is present or not present. Levels range from no symptoms (ID-1), through threshold disorders (ID-5), to definite cases (ID6-8). The ID assigned to each case is arrived at by taking into account the number, severity and the specificity of the signs and symptoms recorded (Dean et al., 1983). An individual who registers ID5 or higher, is considered to have a diagnosable condition and can be assigned a Catego class and an ICD tentative diagnosis (Wing 1983).

Reliability of the PSE

The PSE's questionnaire format, suggested rating scale, and glossary, together with the computer programme, are aimed at increasing the standardization of the interview (Wing, 1983). The glossary in particular attempts to ensure comparable ratings even when made by psychologists and psychiatrists from different schools of thought. Use of the Catego programme means that once PSE ratings are made, the rest of the classification process is automatic, with no room for further personal judgement, apart from the specification of rules incorporated in the Catego computer programme (Grayson et al., 1987; Orley & Wing, 1979).

In spite of these measures, the PSE should be considered a semi-standardized interview. Within the interview process, there is still much room for subjective judgement on the part of the interviewer (Cohen & Cohen, 1984; Wing, 1983). It is left to the examiner's discretion to decide whether the wording of any question should be altered, whether the order of the questions be changed, whether or not to rate a symptom present in spite of the respondent's reply, and which questions ought to be omitted. As Wing (1983, p. 117) notes, "The importance of the skill and experience of the clinician should not be under estimated". The inter-rater reliability of the PSE may therefore not be taken for granted.

Commonly used methods of assessing inter-rater reliability on the PSE include:

(i) a comparison between different interviewers' ratings of the same subjects;

(ii) the re-rating by experienced administrators of recorded field work interviews;

(iii) the rating of the same subject by interviewers and more experienced administrators (Cicchetti, 1976; Cooper et al., 1977).

Several inter-rater reliability studies have yielded satisfactory results. Kendall, et al., (1968)1[2] (in Wing et al., 1974) found the value of kappa for all items to be 0.77. Similar results were found in the WHO International Pilot Study of Schizophrenia, IPSS (1973). Rodgers & Mann (1986) found the inter-rater reliability between lay-interviewers and clinicians to be satisfactory when the PSE-short form was used in a community study.

In an inter-item reliability study (IPSS, WHO, 1973), the median inter-class correlation coefficient was 0.77. Items on depression were considerably more reliable than others, particularly items on anxiety (Rodgers & Mann, 1986).

As long as properly designed programmes of learning procedures and reliability studies are adhered to, the PSE can fulfil the aim of providing an acceptable degree of reliability and repeatability at all stages of the diagnostic process (Wing et al., 1974).

Assessment of local reliability is important given questions regarding the use of psychiatric instruments developed in western cultures when used in non-western cultures. In the course of reviewing the literature for the present study, no evidence was found of any such study in South Africa, though there are informal reports of satisfactory reliability (Gillis et al., 1982).

Validity of the PSE

Assessing the validity of the PSE is a complex issue. It is assumed that the PSE effectively identifies cases and differentiates among major classes of psychopathology, because it has been developed by psychiatrists with considerable experience with clinical patients (Dohrenwhend, 1990). The best means currently available to assess validity of instruments such as the PSE, is to compare their assessment with clinical consensus (Rodgers & Mann, 1986), and to assess the convergence across multiple instruments (Parry, 1992). Assessing validity in terms of clinical consensus, however may be a rather circular argument when one is attempting to assess the validity of an instrument designed to minimize the effects of interviewer bias. Among validity studies already conducted are Rodgers & Mann (1986) and Wing et al., (1978). In addition, Gillis et al., (1982) and Orley & Wing (1979), though not validity studies per se, found the PSE to be a satisfactory measure of psychiatric caseness. In general, these studies suggest that the PSE is suitable as a case-finding instrument.

---

2 US-UK Diagnostic Project, 1968

Other factors which may influence the validity of the PSE include:

(i) the reliability of ratings of the presence and severity of reported symptoms; and

(ii) the interviewer's ability to elicit accurate reports from interviewees (Rodgers & Mann, 1986).

Limitations of the PSE

(i) There are large gaps in the PSE.
It does not deal in any detail with organic psychoses, with hysterical or subcultural conditions, or with disorders which require thorough historical detail, such as mental retardation, bipolar affective disorder, antisocial personality, chronic disabilities, or substance abuse (Grayson et al., 1990);

(ii) The criteria that vague or contradictory definitions be excluded in the construction of the questionnaire lead to the exclusion of a number of important symptoms;

(iii) Definition of psychiatric caseness, according to the PSE, may not be as finite as it seems. Although the ID takes into account the number, severity and the specificity of the signs and symptoms recorded, and allocates caseness accordingly, the level at which caseness is defined is not fixed. Depending upon the context in which the PSE is used, the level may need to be raised or lowered (Dean et al., 1983; Cohen & Cohen, 1984);

(iv) The PSE was designed to measure mental disorder in clinic patients (Wing et al., 1974). However, the level of mental disorder found in the community may be less severe and less stable than that found in clinic and hospital patients (Goldberg & Huxley, 1980, 1992). This being so, there are questions regarding the appropriateness of the PSE as a measure of mental disorder in community studies, where it is required to provide reliable measures across the whole range of various dimensions of psychopathology, particularly depressive and anxiety related disorders (Rodgers & Mann, 1986; Swartz, 1987). Dohrenwend (1990) found a striking contrast in internal consistency reliabilities of the scales for depressive disorders, between psychiatric patients and non-patients. He concluded that as far as the depressive disorders are concerned, the PSE does not measure the same thing in the general population sample as it does in samples of psychiatric patients.

(v) PSE items were originally chosen by a small number of psychiatrists from a school of thought which might be termed "Western European", in its origins, (Wing, 1983). The dangers of using these items to describe and identify mental distress in non-western cultures has already been discussed (see Chapter 1). According to Swartz (1987, p. 6), the PSE sets out to "tap universals but has internal inconsistencies with regard to cultural phenomena". Problems which may arise when using it in non-western societies include:

(i) Difficulties in translating certain items, especially those relating to depression;

(ii) the inclusion of items which require respondents to have knowledge or experience of variables which may be commonplace in western societies, but not in non-western communities, such as weighing scales to assess weight gain or loss; and

(iii) various interpretative factors relating to cultural beliefs, such as spirit possession, which may also affect assessment (Gillis et al., 1982).

Reasons for selecting the PSE as the second-stage criterion

In spite of the PSE's limitations, it was selected as the second-stage criterion for the following reasons:

(i) the PSE has been widely used as a criterion measure with the GHQ and also the GHQ-28, in several studies world wide (Banks, 1983; Finlay-Jones, 1979; Hodiamont et al., 1987; Rabins & Brooks, 1981; Vazques-Barquero et al., 1986, 1987, 1988). The correlation coefficients between it and the GHQ have been satisfactory, ranging between 0.53-0.83 (Goldberg & Williams, 1988);

(ii) although not validity studies per se, the studies of Gilis et al., (1982) and Orley and Wing (1979) found the PSE to be a satisfactory measure of psychiatric disorder in the community;

(iii) an Afrikaans version of the PSE was available and its usefulness established in studies with South African clinic patients (Gillis et al., 1982);

(iv) a psychiatric nurse trained in the PSE by a principal researcher in the International Pilot Study of Schizophrenia, and with considerable experience in its administration, was prepared to participate in the present

research, and assist with the training of four interviewers who had successfully completed the clinical requirements for the MA (Clinical Psychology) degree; and

(v) alternative clinical interviews, such as PAS and the CIS, were not available in Cape Town at the time of the study. Furthermore, the selection of either of these would require training, and also reliability and validity testing in the South African context before they could be used. Financial constraints also motivated against their selection.

The inclusion of additional questions in the first-stage interview

Questions were designed for inclusion in the first-stage interview which would yield data relevant to the investigation of a possible association between mental disorder and socio-demographic variables. Variables under investigation included gender, age, marital status, education, employment and socio-economic status, physical health, and alcohol and substance abuse (Bebbington et al., 1991; Chan, 1985; Cheng, 1988; Dube, 1970; Duncan-Jones & Henderson, 1978; Hoffman et al., 1988; Goldberg & Huxley, 1980, 1992; Vazques-Barquero et al., 1986; Yach & Kuhn, 1989).

Additional questions were also included to ascertain the strength of the relationship between psychological distress and rate of primary health consultation, as well as utilization of informal health services.

All additional questions were designed in collaboration with the Departments of Psychology and Community Health at University of Cape Town, CERSA, and the National Urbanization and Health Research Programme (MRC). Those engaged in this process had considerable experience in the design of questionnaires for the purposes of obtaining base-line socio-demographic information, particularly in Mamre (Hoffman et al., 1988; Katzenellenbogen et al., 1988; Miller et al., 1991). The CAGE (Mayfield et al., 1974) alcohol abuse questionnaire was embedded in a series of questions on general liquid consumption. Questions regarding alcohol intake were interspersed with less provocative questions on the advice of a consultant from the South African National Council on Alcoholism (SANCA). It was hoped that this would allay respondents' suspicions and dispose them to answer questions on alcohol and substance abuse.

The additional questions were included with the screen in a single questionnaire, which was administered in the first-stage interview. The benefits derived from this extra data were considered to outweigh the disadvantage of the longer time taken to administer the first-stage questionnaire.

TRANSLATION OF INSTRUMENTS

Translation of the first-stage questionnaire: the GHQ-28 and additional questions.

Mamre is an Afrikaans speaking community. The GHQ-28 and additional questions therefore had to be translated into the local vernacular. An Afrikaans version of the PSE had reportedly been informally validated in a previous study (Gillis et al., 1982).

The Translation Process

Measures that could be taken to improve the validity and relevance of translated instruments were discussed in Chapter 3, p. 27. Essentially, these emphasize the importance of a familiarity with the subtle cultural norms and idiomatic expressions of the community in which the questionnaire is to be used, as well as the translator's ability to communicate the **intent** rather than the content of the instrument (Brislin, 1986; Chan, 1985).

Outline of the translation procedure

Step One: Translation
Three bilingual persons, with psychological training, and familiar with the general purpose of the GHQ-28, were appointed to translate the questionnaire into Afrikaans. Two of the translators were familiar both with Mamre and the Afrikaans dialect spoken in its community;

Step Two: Back-Translation
The translated questionnaire was sent to an Afrikaans speaking psychology clinician and to a school teacher in Mamre, for back-translation and comments on appropriate language usage.

Step Three: Reassessment of translation
Reassessment of the Afrikaans version was carried out during the training of the four lay-interviewers selected from the community. A number of items were thus reworded in terms of more appropriate idiomatic expression.

Step Four: Paraphrase
A resident of Mamre paraphrased the reassessed version. This version was then administered to a community member by a person who had completed the clinical requirements for the MA (Clinical Psychology) degree to assess its appropriateness and clarity.

Step Five: The pilot study

The final Afrikaans version of the questionnaire was administered in the first pilot study, to test its acceptability and validity in the context of this community.

SELECTION AND TRAINING OF PERSONNEL

Four lay-interviewers were selected to administer the first-stage questionnaire. A data-co-ordinator was appointed to manage the lay-interviewers and co-ordinate the study "in situ"; and five clinicians were appointed to administer the second-stage clinical interview.

First-stage lay-interviewers

Appointment of four lay-interviewers

Research suggests that when a community is involved in the planning and implementation of community programmes, research results are more likely to be meaningful (Katzenellenbogen et al., 1988). It was therefore decided to involve the community as far as possible in the present research project by appointing lay-interviewers and a data-coordinator from the community. This was not an easy option given that Mamre is a relatively small and close-knit community. Previous studies in this community (Katzenellenbogen et al., 1988) reported fairly wide-spread concern among residents that information might not be handled confidentially. This was felt to be a particularly sensitive issues when community members participated as part of the information gathering team. In spite of this very real concern, it was decided to appoint lay-interviewers from within the community, because of the benefits to participants in terms of training and financial reimbursement.

Three males and one female were appointed on the recommendation of a senior community worker and resident of Mamre. All were familiar with general interviewing procedures, as they had participated in other research programmes in the community. None of these studies however, had been of a psychological nature. Lay-interviewers were paid R10.00 for the completion of each interview and R5.00 for an unsuccessful interview, where it could be shown that three attempts had been made to contact the respondent.

Training of the four lay-interviewers

Training took place over a four-day period, in which the interviewers studied the questionnaire and suggested useful corrections to the idiomatic expression of several items. During this process, interviewers familiarized themselves with the content and the underlying meaning of the questionnaire. Training was also given on:

(i) how to introduce the questionnaire as a general health document;

(ii) how to administer the questionnaire, especially with regard to items which a respondent may not understand;

(iii) how to record responses fully and accurately; and

(iv) how to administer the questionnaire in a professional manner, which would not influence the respondents's replies. (While this is an important issue, it is less crucial in the present study, as the GHQ is a self-report questionnaire and does not require the interviewer to interpret or make judgements).

During training, and particularly during role-playing sessions, it became clear that the interviewers were having some difficulty with the wording of the response format of the translated GHQ-28. It was decided to give a typed copy of the responses to each respondent, so that they could simultaneously read and hear the response options.

Appointment of a data-co-ordinator

The size and location of the study required a person to be available in Mamre. A Mamre ex-school teacher was appointed as a data-co-ordinator, and trained over a period of 7 days. In general, his tasks were to guide and manage the lay-interviewers and to co-ordinate the study on a daily basis, including weekends. The data-co-ordinator was also required to:

(i) liaise between the research co-ordinator and the lay-interviewers;

(ii) check all completed questionnaires for non-completion or irregularities; and

(iii) monitor the lay-interviewers on a random basis, to ensure that interviewing techniques were maintained at a consistently satisfactory level of performance. (see Appendix 2.4 for Monitoring Form).

Second-stage clinical interviewers

Four persons who had completed the clinical requirements for the MA (Clinical Psychology) degree and a psychiatric nurse were selected to administer the second-stage clinical interviews. As mentioned earlier, the psychiatric nurse had extensive experience and expertise with the Afrikaans version of the PSE.

The other four clinicians were given two days of concentrated training by the psychiatric nurse on the PSE. Although this training was limited, lack of financial resources and time

did not permit a more thorough training period. Moreover, review of the literature indicated that even this limited training was more than was usually offered in most South African studies (Freeman, 1990; Thom et al., 1993). Voluntary clinic in-patients were used as subjects and were reimbursed for their participation in the training programme. A small inter-rater reliability study on the PSE was planned for the pilot study. Clinical interviewers were paid R50.00 per clinical interview, regardless of whether or not the respondent was rated as positive or asymptomatic.

ETHICAL CONSIDERATIONS

In line with the overall objectives of the Mamre Community Health Project, measures were taken to ensure that the study was acceptable to the community, and involved as little intrusion as possible.

No planned intervention

The acquisition of base line information regarding the mental health status of the community was the primary goal, and no intervention in the community was planned.

Community participation

The project relied on community participation at all levels. While the study was motivated at the request of the Department of Community Health, every effort was made to obtain the community's approval and acceptance prior to its commencement. This was achieved through consultation with members of the steering committee and key figures in the community, such as community workers and the local priest. Notices advertising the purpose and agenda of the study, were also placed in public places such as the church, clinic, and cafes.

Interview confidentiality

Every effort was made in training lay-interviewers from the community, to stress the importance of confidentiality and the professional nature of the study. Lay-interviewers were required to sign a form, in the presence of each respondent, assuring the respondent of the confidentiality of the interview.

Scoring of the first-stage questionnaire was completed by the project co-ordinator so that scoring procedures and results were not made known to either first- or second-stage personnel. As both high and low scorers participated in the follow-up interview, identification of high scorers by the community was not possible.

Voluntary participation in the study

Respondents were informed that participation in both the first- and second-stages of the study was voluntary. Participants in the second-stage clinical interview were paid R20.00.

Follow-up clinical interviews for second-stage respondents requiring counselling

Persons identified by the second-stage clinical interviewer as needing mental health care were given the opportunity to consult with the appropriate official if they so desired it. Mamre is staffed with a clinical psychologist, who lends support and guidance to the social workers, who are based in Atlantis 7 km away, as well as to school nurses, the psychiatric nurse, and other local health care workers operating in Mamre.

Report back to the community

At all stages of the study, those who had participated were up-dated on its progress. The project co-ordinator undertook to report the findings back to the community at the study's end.

The aims and methods used in the first pilot study, as well as the results which led to the second pilot study, and to the inclusion of additional psychiatric instruments are considered in the following chapter.

7 FIRST PILOT STUDY

AIMS OF THE FIRST PILOT STUDY

(i) To assess the appropriateness of the translated version of the first-stage questionnaire;

(ii) To establish the optimal threshold of the GHQ-28 in this context, using *Receiver Operating Characteristic* (ROC) Analysis (Metz, 1978);

(iii) To assess the inter-rater reliability of second-stage clinical interviewers. (Assessment of inter-rater reliability of first-stage lay-interviewers was considered unnecessary, as the GHQ-28 is fully standardized. Random monitoring of lay-interviewers by the data-co-ordinator was felt to be an adequate motivating and controlling factor).

SAMPLE

Forty-four adults over the age of 15 years were randomly selected from the same sampling frame as that planned for the main study. Of those selected, six could not be contacted, three said they were too busy and one was not interested in participating in the study. Thirty-four respondents thus took part in the first-stage questionnaire.

METHOD

Administration of the first-stage questionnaire and second- stage clinical interview

Each lay-interviewer was assigned a randomly selected list of eleven names and corresponding addresses. The first-stage questionnaire was administered to respondents in their homes. Interviewing took place over a three-day period which included a weekend, so that people who worked and lived outside of Mamre during the week, but who considered themselves Mamre residents, could be contacted. Administration of the questionnaire took approximately twenty minutes.

All participants in the pilot received both the first- and second-stage interviews, regardless of their GHQ-28 score. Although all 34 respondents agreed to the follow-up clinical interview, only 21 arrived at the Mamre clinic for the second-stage interview. The clinical interview was conducted within 5 days of the first-stage questionnaire. Second-stage clinical interviewers were blind to respondent's scores on the GHQ-28. As many of the respondents

were symptom free, many of the PSE cut-offs could be used, thus decreasing interviewing time. This saving was attenuated however, by the increased time taken to complete the interview by the four interviewers who had completed the clinical requirements for the MA (Clinical Psychology) degree but who were new to the PSE. On average, each clinical interview took approximately 25 to 45 minutes to complete.

There was a drop-out rate of 38% among respondents who failed to attend the second-stage clinical interview. Of the 13 who dropped out of the study, 6 were male, the average age was 22 years, and 4 were high scorers on the GHQ-28. No demographic variables among these respondents, distinguished them from those who did participate in the clinical interview. It was therefore concluded that the subsample of 21 who received the PSE was no different from the 34 who were administered the GHQ-28.

RESULTS

Sample representativeness

Sample representativeness could only be assessed imprecisely due to the smallness of sample size. However, as the sample was randomly drawn from the total adult population frame, it was likely to be representative. Nevertheless, there was an over-representation of women in the 65 to 74 year age group, which was 7% higher than in the general health population study conducted in 1986 (Hoffman et al., 1986).

Assessment of the suitability of the translated version of the first-stage questionnaire, including the GHQ-28 and additional questions

Lay-interviewers reported no difficulty in the administration of the questionnaire. The anticipated difficulty with the wording of the GHQ-28 response format did not materialize. Some questions from the rest of the schedule were found to be inapplicable, and others were insufficiently specific. Questions relating to tertiary education were mostly irrelevant as the majority of Mamre residents had not received schooling beyond matric level. Questions concerning the days on which alcohol was most likely to be consumed were regarded as too general, as there appeared to be a difference in the amount of alcohol consumed during the week and over weekends.

Optimal threshold for the GHQ-28 when used in this context

A single threshold yields only a limited amount of information about the performance of the measuring instrument (Goldberg & Williams, 1988). It does not allow for the assessment of the discriminating ability of the instrument across the total spectrum of morbidity, nor does it allow for the assessment of the effect on the validity coefficients of varying the threshold score (Goldberg & Williams, 1988). These questions can be addressed, however, using the

Receiver Operating Characteristic Analysis, (ROC), (Metz, 1978). Accordingly, the optimal threshold on the GHQ-28 in the context of the present study, was determined by estimating GHQ-28's best validity coefficients for **all possible thresholds**. Both methods of scoring the GHQ-28 (GHQ and CGHQ method) were used. Table 7.1 gives the *validity coefficients*[1] of the GHQ-28 and CGHQ-28 for all possible thresholds, with the PSE-ID5 as the gold standard.

Table 7.1 Validity coefficients for the GHQ-28 at all possible thresholds, using both the GHQ and CGHQ methods of scoring and with the PSE-ID5 as the gold standard.

GHQ method of scoring (0,0,1,1)				
Threshold	**Sensitivity**	**Specificity**	**False Positive**	**False Negative**
3/4	**0.57**	**0.50**	**0.50**	**0.43**
4/5	0.43	0.71	0.29	0.57
5/6	0.29	0.71	0.29	0.71
6/7	0.29	0.79	0.21	0.71
6/7	0.14	0.86	0.14	0.86
CGHQ method of scoring (0,1,1,1)				
Threshold	**Sensitivity**	**Specificity**	**False Positive**	**False Negative**
3/4	0.71	0.36	0.64	0.29
4/5	0.57	0.50	0.50	0.43
5/6	**0.57**	**0.64**	**0.36**	**0.43**
6/7	0.43	0.64	0.36	0.57
7/8	0.43	0.86	0.14	0.57
8/9	0.29	1.00	-	0.71

Table 7.1 shows that the optimal threshold (*sensitivity*, 0.57, *specificity*, 0.50) for the GHQ-28 in this context was 3/4, while the optimal threshold (sensitivity, 0.57, specificity, 0.64) of the CGHQ-28 was 5/6. These thresholds represented the point where both sensitivity and specificity were optimal, and the false negative rate lowest. They also represented a

---

1 See Glossary of Technical Terms, Appendix 1.1.

compromise between high sensitivity and low false positive rate at all possible threshold levels.

At a threshold of 3/4, the GHQ-28 identified 52% of respondents as potential cases, and at a threshold of 5/6, the CGHQ-28 identified 42% as potential cases. The PSE-ID5, acting as the gold standard, identified 33% of respondents as psychiatric cases. Table 7.2 illustrates the method used to estimate the sensitivity and specificity of the GHQ-28 at the 3/4 threshold, using the GHQ method of scoring, and with the PSE as the gold standard.

Table 7.2 Validity coefficients of the GHQ-28 at a threshold of 3/4, using the GHQ method of scoring, and the PSE-ID5 as the gold standard.

GHQ-28	**PSE-ID5**		**Total**
	non-case	**case**	
Above 3/4	7 false positive	4 true positive	11
Below 3/4	7 true negative	3 false negative	10
Total	14	7	21

Sensitivity: true positives/true positives + false negatives
= 0.57

Specificity: true negatives correctly identified/false positives + true negatives correctly identified
= 0.50

Inter-rater reliability study on the PSE

Inter-rater reliability of PSE interviewers was assessed by having one interviewer rate a GHQ-28 high scoring candidate, in the presence of the four other clinical interviewers. Each interviewer recorded his/her rating of the candidate, according to the scoring procedure set down by the PSE. At the end of the interview, each interviewer was given the opportunity to ask the respondent any questions he/she felt relevant, but which had not already been asked. Other interviewers were not permitted to alter their own ratings in response to the answer given. There was total agreement among the interviewers in terms of case/non-caseness, as well as in terms of diagnostic category. A statistic for the extent of this agreement was not calculated, given the small sample size, (n=1). In spite of these results, it was recognised that

the inter-rater reliability study was inadequate. Once again, time and financial constraints did not permit a more thorough investigation.

ANALYSIS OF RESULTS

In terms of the specified aims of the pilot study:

(i) the acceptability and appropriateness of the Afrikaans version of both the GHQ-28 and the additional questions were established;

(ii) an optimal threshold of 3/4 on the GHQ-28 and 5/6 on the CGHQ-28 was estimated. Although the 3/4 threshold was lower than that recommended by Goldberg & Williams (1988) for clinic based studies, it was considered acceptable, as the project was a community study; and

(iii) the inter-rater reliability study, although very limited in scope, suggested that agreement between clinical interviewers regarding the rating of respondents and assignment to diagnostic category, was satisfactory.

The validity coefficients for both the GHQ-28 and CGHQ-28 were, however, unsatisfactory. With a sensitivity of 0.57, and a false negative rate of 43%, the coefficients represented only above chance rates! There were five possible explanations for the low validity coefficients:

(i) small sample size;

(ii) over-representation of the 65-74 year age group. A higher rate of mental disorder than usual has frequently been found in this age group (Goldberg & Williams, 1988). Of the seven cases identified by the PSE, three were over 65;

(iii) the possibility that the GHQ-28 was an inefficient screen for the identification of potential cases in the context of this community sample. Although none of the lay-interviewers reported any difficulty with its administration, the fairly complicated response format of the GHQ-28 may have been misunderstood by respondents. This may have influenced respondents' scores, and consequently the relationship between the GHQ-28 and the PSE. Of the seven cases identified by the PSE, the GHQ-28 incorrectly identified four as non-cases, and the CGHQ-28 three as non-cases. Assuming the PSE was error free, these cases would have been missed in a study in which the GHQ-28 was used alone;

(iv) the possibility that the PSE was mis- or over- diagnosing cases, especially of paranoid schizophrenia. Of the seven cases identified by the PSE, three were given a tentative diagnosis of paranoid schizophrenia. (Questions regarding the PSE's mis- or over- diagnosis of paranoid schizophrenia are discussed in Chapter 12, p. 33 and in Appendix 4.2.

(v) the GHQ-28's high false positive rate (50%), may indicate that the PSE was insensitive to the level of mental disorder commonly found in communities, for example, depression and anxiety (Sashidharan, et al., 1990). This being so, the GHQ-28 and the PSE were measuring different levels of mental disorder. This suggests that the PSE may be an inappropriate gold standard for the validation of the GHQ-28 in a community study.

CONCLUSION

The disparity between the GHQ-28 and the PSE in identifying cases, and the corresponding inadequate sensitivity and specificity of the GHQ-28, brought into question the efficacy of **both** the GHQ-28 as a first-stage screen, and the PSE as a second-stage gold standard, when measuring mental disorder within a general population in the South African context. However, the manner in which the two instruments acted as a cross-check on one another illustrated the advantages of a two-stage method of case detection.

In an attempt to address the above problems, it was decided to conduct a second pilot study. In this study, the Self Reporting Questionnaire (SRQ) (Harding et al., 1980) which has an alternative scoring method, was included in the first stage questionnaire, and the Psychiatric Assessment Schedule (PAS) (Dean et al., 1983 was included in a single clinical interview, along with the PSE, as the second stage criterion measure. Second-stage clinical interviewers were also asked to record their "suggested clinical diagnosis" at the end of each clinical interview. It was anticipated that further information would become available regarding the performance and appropriateness of both screening and criterion measures in this and similar contexts. Details of the second pilot study appear in the following chapter.

8 SECOND PILOT STUDY

This chapter briefly describes the aims of the second pilot study and reviews the two additional first- and second-stage instruments, the SRQ and PAS. The decision to select the SRQ as the chief first-stage screen and the PSE as the primary second-stage criterion in the main study, is also discussed, as is the rationale for the inclusion of **all** second pilot instruments in the main study.

AIMS

(i) to compare the validity coefficients of the GHQ-28 and the SRQ when used with all three criterion measures, the PSE, the PAS, and the clinicians' "suggested diagnosis";

(ii) to estimate the optimal thresholds on the GHQ-28 and SRQ in this context when used with all three criterion measures; and

(iii) to assess the appropriateness and acceptability of the combined PSE/PAS clinical interview.

ADDITIONAL INSTRUMENTS FOR INCLUSION IN THE STUDY

First-stage screen: The Self Reporting Questionnaire, SRQ

The SRQ was selected for the following reasons:

(i) it was designed specifically for use in non-western cultures;

(ii) it was designed to be administered by lay-interviewers with the minimum amount of training;

(iii) its response format is simpler than that of the GHQ-28, and it is thought to be more accessible to respondents in developing populations (Harding et al., 1980);

(iv) it has been used successfully in several epidemiological studies in non-western and African cultures (Parry, 1992); and

(v) it was used in a similar prevalence study in Mexico, in which the GHQ-28 was found to have low validity coefficients in the pilot (Marie & Williams, 1984).

Rationale for the construction of the SRQ

The SRQ was designed by Harding et al., (1980) to be used in a baseline study to develop and evaluate low-cost methods of mental health care in developing countries (WHO collaborative study, 1980). Several studies had identified the need for a psychiatric screening questionnaire capable of detecting mental disorder in general populations. They noted that in developing communities, those with mental distress frequently present with somatic complaints (Harding et al., 1980), as mental illness is often regarded as degrading (Dhadphale et al., 1982). Furthermore, emotional stress is often accompanied by chronic physical symptoms (Dhadphale et al., 1982); and the high prevalence of serious physical conditions, such as malaria and infectious diseases, means that provision for mental distress usually takes second place (Harding et al., 1980). Where phc services **are** available, they generally have to rely on community health workers with only basic training and limited educational backgrounds.

Aims of the SRQ

The SRQ was designed to be:

(i) used in phc clinic samples in developing countries;

(ii) sensitive to the level of mental distress commonly found in the community;

(iii) able to detect psychological symptoms obscured or accompanied by physical symptoms; and

(iv) administered by community health workers, or even lay-interviewers, with minimum training.

Description of the SRQ

The SRQ is a self-report questionnaire which may either be completed by the respondent him/herself, or administered by community health workers or lay-interviewers. The full version of the SRQ consists of 25 items. Items 1 to 20 cover somatic complaints, sleep disturbance, anxiety and depression, while items 21 to 25 screen for psychotic morbidity and epilepsy.

Item selection

Items were selected by means of a consensual process from four pre-existing instruments: the Patient Self-Report Symptom Form (PASSR), (Climent & Plutchick, 1979, in Harding et al., 1980), the PGI Health Questionnaire NZ (Verma & Wig, 1977, in Harding et al., 1980), the GHQ (Goldberg, 1972) and the PSE (Wing, et al. 1974). Items for inclusion were selected on the basis of ease of translation and cultural relevance. The five psychotic items were based on items in Foulds' Symptom Sign Inventory, (Foulds & Hope, 1968).

Setting the threshold on the SRQ

Potential mental health/ill-health is conceptualized in terms of the dimensional model, so that respondents are screened into either group according to a predetermined threshold. The threshold on the SRQ is not fixed, and the appropriate cut-off needs to be estimated in the context in which the test is administered. The SRQ's threshold appears, however, to be fairly stable at 7/8 (Harding et al., 1980; Marie & Williams, 1985; Sen, 1987; Skuse & Williams, 1984).

Response scale and scoring procedure on the SRQ

The SRQ has a simple yes/no response format designed specifically to be accessible to less literate non-western populations (Harding et al., 1980).

A "yes" response is scored as 1, and a "no" response as 0. The summed score indicates the number of symptoms present, but does not provide a severity estimate. A "yes" response to any of the last five questions on psychotic states necessitates a follow-up clinical interview.

Reliability of the SRQ

Reliability of the SRQ has been established in several studies, (Harding et al., 1980). Given the SRQ's simple yes/no format it was decided that an inter-rater reliability study was not necessary in the present study.

Validity of the SRQ

The validity of the SRQ has been assessed in over ten studies in developing countries, including Brazil (Marie & Williams, 1985), Colombia (Climent & Plutchick, 1979, in Harding et al., 1980), Ethiopia (Kortman & Ten Horn, 1988); India (Ibrahim et al., 1980, in Harding et al., 1980), Senegal (Diop et al,, 1982), Kenya (Dhadphale et al., 1982), India (Sen, 1987), Guinea-Bissau (De Jong et al., 1986), and South Africa (Freeman, 1990a; Thom et al., 1993). Criterion instruments used in a two-stage design with the SRQ include the PSE and the CIS. Validity studies based on clinic samples showed a sensitivity between 0.78 and 0.80 and a specificity between 0.78 and 0.81 (Dhadphale et al., 1982; Harding et al., 1980; Kortman & Ten Horn, 1988; Marie & Williams, 1985; Sen et al., 1987). However, criticism has been levelled at both Harding et al's., 1980 study and Dhadphale et al's., 1983 study (Marie & Williams 1985). In the former, it was found that the criterion measure was applied to a highly skewed sub-sample and the resulting validity coefficients were not shown to be adjusted to represent the population as a whole (Marie & Williams, 1985). In Dhadphale's 1982 study, known patients were compared with known normals. This not only minimized misclassification, but may also have resulted in overestimation of the efficacy of the test (Marie & Williams, 1985).

Comparison between the GHQ-28 and the SRQ

Essentially the SRQ and the GHQ-28 are similar in purpose and design. Their similarities and differences are discussed below.

Similarities:

(i) Screen for a common core of symptoms believed to underlie both the severe and less severe mental disorders;

(ii) Items 1-28 on the GHQ-28 and 1-20 on the SRQ are similar in content;

(iii) Mental disorder is conceived as being distributed in varying degrees of severity among the general population (Goldberg & Williams, 1988), therefore there is no fixed threshold;

(iv) Yield an area score, which does not indicate the severity of the disorder;

(v) No theoretical assumptions are made about the nature of the diagnostic hierarchy within the classes of psychiatric illness. Neither screen should therefore be used to identify cases in any diagnostic sense;

(vi) Have been used with different criterion measures, including the PSE and CIS, in a number of studies in different contexts. Similarity in range of the validity coefficients suggests that the choice of criterion interview does not significantly affect the validity of either screen (Williams et al., 1987);

(vii) Measure changeable mental states requiring the administration of the second-stage clinical interview within 7-9 days (Cheng, 1988);

(viii) Sensitivity is lower in community samples than in clinic or hospital samples (Goldberg & Williams, 1988);

(ix) Acceptable to respondents who are symptom free;

(x) Translated into several languages and used in a wide number of cultural settings with good results (Cheng, 1985; Goldberg & Williams, 1988; Marie & Williams, 1985 Medina-Mora et al., 1983; Vazquez-Barquero et al., 1986); and

(xi) Inexpensive and can be administered by lay-interviewers after a brief period of simple training.

Differences:

(i) SRQ includes five items pertaining to psychotic states which are not covered in the GHQ-28;

(ii) SRQ questions may be more acceptable to respondents in non-western cultures (Sen, 1987);

(iii) The GHQ-28's response format includes both negatively and positively worded items to indicate health in some instances and ill-health in others. This may reduce errors due to overall agreement set (Goldberg, 1991), but may not be as easy to administer as the SRQ's simple yes/no response format (Sen, 1987);

(iv) The GHQ-28's method of scoring, 0,0,1,1, means that only items which the respondent rates as differing from his/her usual behaviour, are given a positive score. While this method has several advantages, it may miss chronic cases of distress.

Second-stage criterion measure: Psychiatric Assessment Schedule, PAS

The PAS was selected for the following reasons:

(i) the results of the first pilot suggested that the PSE might be relatively insensitive to the level of mental disorder commonly encountered in the community. The PAS puts special emphasis on issues related to depression and anxiety, the mental disorders most usually found in the community;

(ii) scores on the PAS yield both ICD8 and DSM111R tentative diagnoses;

(iii) as the PAS includes the first 40 items on the PSE it was possible to incorporate the two interviews into a single clinical interview and to score both instruments on a single PSE modified scoring sheet. This was achieved by interspersing PSE items with PAS items not covered by the former. (see Appendix 2.6). The advantages of combining the PSE and PAS into a single clinical interview include:

(i) the simultaneous administration of two different semi-structured interviews;

(ii) the use of the PSE-Catego computer programme to score both interviews;

(iii) the use of two diagnostic systems, the ICD8 and the DSM111R, to interpret the scores; and

(iv) the opportunity to assess both severe and less severe levels of mental disorder in the same population at the same time.

Rationale for the construction of the PAS

In spite of efforts to design increasingly standardized psychiatric instruments, there remains a disturbing rate of incongruence between both rate of diagnosis and case identification (Dean et al., 1983). In order to test the extent of this incongruence, Dean and her colleagues designed a study whereby the same sample of women were interviewed using a single semi-structured clinical interview, the PAS (1983). This was constructed so that the data could be interpreted in terms of four different diagnostic systems: the PSE-Catego/ICD, 8th ED.; the Research Diagnostic Criteria (RDC) (Spitzer et al., 1978), which was later replaced by the DSM111R; the Bedford Scale (Brown & Harris, 1978); and the Feighner criteria (Feighner et al., 1972).

Description

The PAS is made up from the first 40 items of the shortened PSE, plus 39 additional items from the Schedule for Affective Disorders and Schizophrenia, (SADS), (Endicott & Spitzer, 1978). The first 40 items from the PSE are reproduced in the PAS unchanged, so that the PSE-Catego computer programme can still yield a PSE Catego class and a diagnosis according to the ICD8 (Dean et al., 1983). Where no equivalent on the PSE could be found, or where questions were insufficiently detailed, additional items from the SADS were selected. Among the items supplemented were questions relating to increased appetite, weight gain, and sleep disturbance; preoccupation with thoughts of death; panic attacks; illness duration; extent of impaired functioning; personality disorders associated with mood disturbances; time since subject last felt like his/her normal self, and the six month period prior to interview.

Reliability

As the PAS is a composite of the PSE and SADS its reliability is not expected to be different from that of either of the aforementioned instruments (Orley & Wing, 1979; Rodgers & Mann, 1986; Wing, et al., 1974; 1978).

Validity

The PAS's validity is also assumed to be comparable to that of the PSE, and to a lesser extent, the SADS.

Limitations of the PAS

Suggested limitations of the PSE also apply to the PAS. It does not enable diagnoses of alcoholism, anorexia nervosa, obsessive compulsive disorder, phobic disorder, or any organic state to be made (Dean et al., 1983). The PAS is also unable to detect psychotic disorders including schizophrenia.

Similarities and Differences between the PSE and PAS

Similarities and differences between the two instruments are listed below.

Similarities:

(i) Include items on health, worry and tension, autonomic anxiety, thinking, concentration, depressed mood, appetite, sleep, libido, irritability, behaviour affect and speech;

(ii) Can be used with the PSE Catego computer programme; and

(iii) Provide diagnoses in terms of the ICD.

Differences:

(i) The PAS does not cover symptoms related to expansive mood and ideation, obsession, derealisation and depersonalization, perceptual disorders, hallucinations, delusions of control and psychotic symptoms;

(ii) Additional questions on the PAS focus attention on issues relating to depressive and anxiety related disorders;

(iii) The PAS may be used with three diagnostic systems: the ICD, DSM111R, and the Bedford Scale; whereas the PSE may only be used with the ICD.

"suggested clinical diagnosis"

Although psychiatric epidemiology has moved towards increasing instrument standardization, a computer diagnosis such as the PSE-Catego cannot, and is not intended to replace a clinical diagnosis (Wing, 1983). The decision that second-stage clinical interviewers should record their own "suggested clinical diagnosis" at the end of the PSE/PAS score sheet provided an opportunity to investigate a number of psychiatric diagnostic algorithms simultaneously, at no extra cost to the respondents or the study. This approach was used by Ben-Arie et al., (1983) in a study of elderly "coloured" persons in Cape Town, and was argued to be appropriate to studies using the PSE in populations on which it was not standardized.

Translation and training

All additional instruments were translated, using the same procedures used in the translation of the GHQ-28 (see Chapter 6, p. 59)

Lay-interviewers were given a one-day training session on the SRQ. The five clinical interviewers received one day's additional training on the combined PSE/PAS interview. This was given by a senior psychology clinician who has had extensive experience with the PSE in the South African context, but who was not part of the study.

No inter-rater reliability study on the combined PSE/PAS was conducted, and its absence constituted a serious weakness in the present research design. Its impact on the results is discussed in a later chapter.

SAMPLE

Using the same sampling procedures as in the first pilot study, 33 adults were randomly selected from the Mamre population frame. Four persons could not be contacted, and a total of 29 respondents thus received the first-stage questionnaire.

METHOD

Administration of the GHQ-28 and SRQ was varied to reduce the possibility of an interference factor. Accordingly, the GHQ-28 was administered first to 15 respondents and the SRQ first to the other 14 respondents. The questionnaire took approximately 15 minutes to administer, nearly 5 minutes more than the time taken to administer the GHQ-28 alone. Method of administration and monitoring of lay-interviewers remained the same as in the first pilot.

Within five days of the screen, 21 respondents received the combined PSE/PAS clinical interview. Once again, clinicians were blind to respondents' scores on the screens. Twenty eight percent of those who had agreed to participate in the clinical interview, dropped out of the study. Of these, four were male, the average age was 25 years, and three were high scorers on both the GHQ-28 and SRQ. Two of the male respondents said that they usually drank heavily over the weekend, and this may have accounted for their absence at the clinical interview. No other factor was found to explain non-attendance at the follow-up interview. It was therefore concluded that the subsample of 21, who received the second-stage interview, was not obviously different from the 29 who were administered the GHQ-28 and SRQ.

RESULTS

Sample representativeness

Smallness of sample size prevented any conclusive statement regarding its representativeness. Once again, the random sampling technique employed increased the likelihood of the sample being representative of the general population. The sample closely paralleled the gender ratio of 1 to 1 found in the 1986 Department of Community Health population census. However, both the 16 to 19 year and 35 to 54 year age groups were over-represented.

Assessment regarding the inclusion of additional first and second-stage instruments

Although the inclusion of a second screen increased the time taken to administer the interview, lay-interviewers reported that with only two exceptions, respondents were amenable to answering both screens, one directly after the other.

Clinical interviewers found that the incorporation of the PSE with the PAS did not significantly increase the time taken to administer the clinical interview.

Validity coefficients for all possible thresholds for both the GHQ-28 (GHQ and CGHQ methods of scoring) and SRQ, when used with the three criterion measures in this context

ROC analysis was used to determine the optimal thresholds on the SRQ, GHQ-28 and CGHQ-28. When used with any one of the criterion measures, a threshold of 3/4 was found to be optimal on the GHQ-28, and a threshold of 7/8 for both the SRQ and the GHQ-28.

Of the 29 persons screened, 11 (38%) were high scorers on the GHQ-28, 20 (69%) were high on the CGHQ-28 and 10 (34%) were high scorers on the SRQ.

The PSE-Catego identified five respondents as possible psychiatric cases. Of these, three were identified as positive both by the PAS and "suggested clinical diagnosis".

Table 8.1 shows the validity coefficients for the two screens (including both methods of scoring the GHQ-28) at their optimal thresholds, and also when used with the three criteria measures.

Table 8.1 Validity coefficients at optimal threshold for GHQ-28 (both methods of scoring, GHQ and CGHQ) and SRQ, using the PSE, PAS and "suggested clinical diagnosis" as criterion measures.

Criterion: PSE-Catego						
Instrument	**Cut-off**	**Sensitivity**	**Specfitivity**	**False negative**	**False positive**	***Positive predictive value***[1]
GHQ	3/4	1.0	0.81	-	0.19	1.00
CGHQ	7/8	0.60	0.81	0.40	0.19	0.84
SRQ	7/8	1.0	0.88	-	0.12	1.00
Criterion: PAS-DSMR111R						
Instrument	**Cut-off**	**Sensitivity**	**Specfitivity**	**False negative**	**False positive**	***Positive predictive value***
GHQ	3/4	1.0	0.72	-	0.28	1.00
CGHQ	7/8	0.60	0.81	0.40	0.19	0.84
SRQ	7/8	1.0	0.89	-	0.11	1.00
Criterion: "suggested clinical diagnosis"						
Instrument	**Cut-off**	**Sensitivity**	**Specfitivity**	**False negative**	**False positive**	***Positive predictive value***
GHQ	3/4	1.0	0.72	-	0.28	1.00
CGHQ	7/8	1.0	0.39	-	0.61	1.00
SRQ	7/8	1.0	0.78	-	0.22	1.00

ANALYSIS OF RESULTS

Comparison of validity coefficients of the GHQ-28, CGHQ-28, and the SRQ

Table 8.1 shows that validity coefficients for the two screens were nearly equivalent, regardless of the second-stage criterion measures used,. The sensitivity of the CGHQ-28

---

1 See Glossary of Technical Terms, Appendix 1.1

however, was lower than either that of the GHQ-28 or the SRQ when the PSE or the PAS was the criterion measure. Sensitivity for both the SRQ and GHQ-28 was 1.0 when any one of the criterion measures was used. However, the SRQ had a slightly better specificity than the GHQ-28, (0.88-v-0.81 when used with the PSE-Catego, and 0.89-v-0.72 when used with the PAS-DSM111R), suggesting that it was likely to identify fewer false positives than the GHQ-28.

It was concluded that no decision regarding the superiority of one screen over the other in this context, could be made. It was therefore decided to retain both screens in the main study. However, as the SRQ had slightly better validity coefficients and a lower false positive rate, it was also decided that the administration of the screens should be unvaried. This meant that the SRQ would be administered first in all cases. The SRQ was therefore elected as the **primary** screening instrument to be used in estimating the prevalence of mental disorder. Thus all respondents scoring above the SRQ threshold of 7/8 would be considered high scorers, regardless of their GHQ-28 score.

In spite of the CGHQ's lower validity coefficients, it was also decided to retain both methods of scoring the GHQ-28. This decision was motivated by strongly reported evidence that long-standing disorders are more likely to be detected when using the CGHQ method of scoring (Goodchild & Duncan-Jones, 1985).

Concordance between the three different case-finding instruments

All three cases identified by the PAS-DSM111R and "suggested clinical diagnoses", were also rated positive by the PSE-Catego. The two cases identified by the PSE-Catego (ICD8) only, were both given a tentative diagnosis of paranoid psychosis. Once again identification and diagnoses of these cases as paranoid psychosis by the PSE, prompted questions regarding the PSE's mis- or over-diagnoses of certain syndromes when used as a case-identifying instrument in a general population sample. (see Chapter 12, p. 133 and Appendix 4.2).

While the SRQ's specificity and false positive rate were slightly better when the PAS-DSM111R was the criterion, overall there was nothing to suggest that any one of the criterion measures was more appropriate in this context. Similarity in the validity coefficients of both screens suggested that with the exception of the psychotic conditions (which the PAS is not designed to measure), the PSE and the PAS were in fact measuring the same level of mental disorder. It was nevertheless decided to retain both the PAS and "suggested clinical diagnosis" in the main study, in order to maximize information on caseness, and to provide data on psychiatric instruments in the South African context. Inclusion of all three criterion measures neither interfered with the goals of the research, nor had any negative impact on the research design or methodology. However, it was decided that the PSE would remain the

primary second-stage criterion for estimating prevalence, and the gold standard in the validation study of the SRQ. This decision was based on the following reasons:

(i) in spite of the SRQ's slightly higher specificity and lower false positive rate when the PAS-DSM111R was the criterion, the difference in the validity coefficients was not great enough to conclude that the PAS-DSM111R was a superior criterion measure to that of the PSE-Catego in this context;

(ii) taken together, the validity coefficients of both the SRQ and the CGHQ-28 were slightly better when the PSE was the criterion measure;

(iii) the PSE has been widely used as a second-stage validating instrument internationally, both with the SRQ and the GHQ (Banks, 1983; De Jong et al., 1986; Finlay-Jones & Murphy, 1979; Freeman, 1990; Harding et al., 1980; Hodiamont et al., 1987; Rabins & Brook, 1981; Vazquez-Barquero et al., 1986, 1987, 1988; Thom et al., 1993);

(iv) the PSE has been used in studies on the African continent and in South Africa in particular (Ben-Arie et al., 1983; De Jong et al., 1986; Freeman, 1990; Orley & Wing, 1979; Thom et al., 1993).

CONCLUSION

Results of the second pilot were an improvement on those obtained in the first study. Several factors may have contributed to this improvement:

(i) variation due to random sampling between the first and second pilot samples;

(ii) additional training of second-stage clinical interviewers prior to the second pilot, which may have led to more sensitive and careful scoring of questions in the non-psychotic range;

(iii) increased second-pilot sample representativeness; or

(iv) the existence of an unidentified interference variable.

In light of the issues discussed above, the initial aims of the research were reformulated to include investigation of psychiatric first and second-stage instruments.

9 REFORMULATED AIMS

Essentially, the initial aims identified in Chapter 5 remained the **primary** objectives of the research study. However, the results of the first and second pilot studies identified further areas of investigation, which could be accommodated as **secondary** objectives within the planned research design. The areas of enquiry pertained to the investigation of certain first- and second-stage psychiatric epidemiology instruments when used in a community sample in South Africa. In assessing the appropriateness of these instruments the following issues were considered:

(i) possible differences in the validity coefficients of the SRQ, GHQ-28 and CGHQ-28 when the PSE was used as the gold standard;

(ii) possible differences in the *validity coefficients* of the SRQ when used with any one of the criterion measures: the PSE, PAS or "suggested clinical diagnosis"; and

(iii) a comparison of prevalence estimates obtained simultaneously in the same population, for all three methods of case identification.

It was recognised that these secondary aims were investigative and that findings would require more specific follow-up research.

REFORMULATED AIMS

Primary objectives

The primary objectives remained stable (see Chapter 5, p. 37) with the exception of objective (ii). In this objective the SRQ replaced the GHQ-28 as the first-stage screening instrument selected for validation in a community sample in South Africa.

Secondary objectives

(i) to investigate two different screening instruments, the SRQ, and the GHQ-28 (both methods of scoring), in terms of their ability to identify mental disorder as defined by the PSE, in a general population sample in the South African context;

(ii) to enquire into the effects on prevalence rate when three different systems of case identification are employed, ie, the PSE-Catego, the PAS-DSM111R and the clinical interviewers' "suggested clinical diagnosis"; and

(iii) to assess any changes in the SRQ's validity coefficients when used with any one of the above systems of case identification.

Taken together, the primary and secondary objectives motivated a research design which included both substantive and exploratory research in a single study. It was anticipated that the combination of these two approaches would result in findings which would be useful both to the planning of relevant health care services and to future research.

10 MAIN STUDY

This chapter describes the main study which took place over a five-week period during the winter of 1992. Several issues customarily discussed in the Method section of the main study, such as research design, sample selection, instrument selection and translation, and recruitment and training of personnel, have already been described in previous chapters. These issues are therefore not reviewed again but, for purposes of clarity reference to the appropriate Chapter in which they are discussed is given.

SAMPLE

Sample Description

(see Chapter 6, p. 41.)

Calculation of sample size

Full details on the calculation of first- and second-stage sample size are given in Appendix 3.1. Based on equations taken from Shrout (1992), it was estimated that a sample of 656 was an efficient first-stage sample size, which would both yield a small standard error of estimate and still be within the study's budget, and that a second-stage sample of 111 was appropriate.

Random selection of first-stage sample

A random sample of 700 adults was selected from the total adult population of Mamre, (see Chapter 6, p. 41). In addition, a further 200 names and addresses were selected as part of a back-up list. The back-up list was to be used where the original respondent could not be contacted, or refused to participate in the interview. Replacement respondents were selected to be similar to the original respondent in terms of gender, age and location.

Random selection of second-stage sample

A stratified random sample of 53, (27.5%) respondents from the SRQ screened positives and 57, (12.5%) respondents from the SRQ screened negatives (see Appendix 3.1 for rationale), were selected to participate in the second-stage clinical interview. As a given proportion of the sample was randomly selected form either high or low SRQ scoring groups, it was a stratified sample. This meant that prevalence estimates and SRQ validity coefficients had to

be weighted back to the original sample at the end of the study, to ensure the results were representative of the community as a whole (Goldberg & Williams, 1988).

METHOD

Instruments

(see Chapter 6, pp. 50 and 53, Chapter 8, pp. 70 and 74)

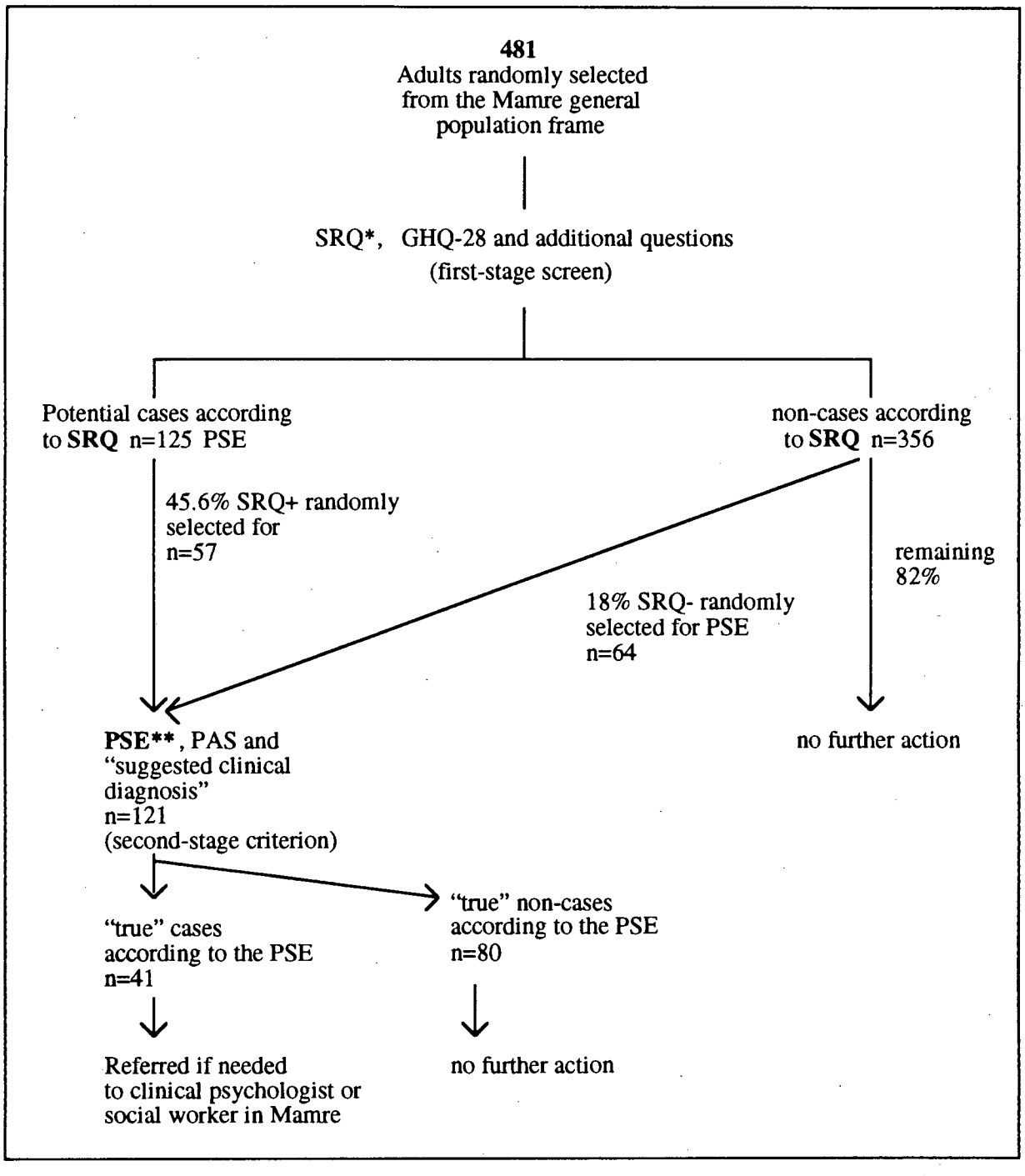
First- and second-stage personnel

(see Chapter 6, p. 60)

Administration of first- and second-stage instruments

As procedures in the main study were fairly complex, a diagrammatic representation is provided in Fig. 10.1 to clarify the methods used to select and detect potential and definite psychiatric cases. Sample size and proportions shown in this diagram reflect the numbers actually used in the main study and not those on which sample size was originally estimated (see Appendix 3.1, Fig. 3.1.2). Changes to both estimated first-stage sample size and the proportion of screened positives and negatives required to receive the second-stage interview are discussed later in this chapter (see p. 89). However, in the discussion which follows Fig. 10.1, the original estimates are used as the administration of first- and second-stage interviews were based on these estimates in the first two weeks of the study.

Figure 10.1 Diagrammatic representation of procedures used in the main study to select and detect potential and definite cases. (Actual numbers used in the main study shown and not those originally estimated).



* SRQ (threshold 7/8) used as the main screen.
** PSE-ID5 used as the main criterion measure

Out of a total of 656 first-stage respondents, 164 were randomly selected each week to receive the first-stage questionnaire. This weekly estimate was based on the need to meet the required first-stage sample size of 656 within a 4 week period, and to enable the second-stage clinical interview to be administered within 9 days of the screen.

This meant that over a period of 4 weeks, each of the 4 lay-interviewers received a list of 41 names and addresses weekly. Respondents' names were arranged according to street and area, for easier administration of interviews. Lists were randomly distributed to the interviewers to prevent any one interviewer from only interviewing respondents in his/her own residential area. If a respondent could not be contacted after three attempts, refused to participate, or was incapable of giving the interview, he/she was replaced with a respondent from the back-up list.

First-stage interviews were administered in respondents' homes. Participation was voluntary, and interview confidentiality was vouchsafed in the confidentiality form. Interviewing took place weekly over a five day period, including a weekend. On the sixth day of each week, all completed interviews were handed to the data-co-ordinator for checking. Any uncompleted form was returned to the interviewer to be completed if possible. On the seventh day, interviewers received the new week's list of names and were given feed-back on their interviewing skills, as monitored by the data-co-ordinator. Problems encountered during the previous 6 days were discussed and payment for completed interviews was made.

All completed questionnaires were scored on the sixth day of each week, by the research co-ordinator. Scoring procedures and results were not communicated to the lay-interviewers, data co-ordinator or the clinical interviewers.

Once first-stage questionnaires had been scored, a weekly proportion of SRQ high and low scores was randomly selected for the 2 groups of SRQ respondents. Letters addressed to all second-stage respondents, informing them of the time and date of the interview, were given to the lay-interviewers to deliver. On delivery, lay-interviewers were instructed to ensure that respondents knew the directions to the clinic where the interview was to take place, had transport to the interview (if not this was provided), and were satisfied with the scheduled appointment time. On the day prior to the clinical interview, both interviewers and the data-co-ordinator attempted to contact all selected respondents to confirm their interview time but more specifically, to motivate their attendance.

Administration of the second-stage clinical interview

An estimated total of 110 respondents were required to participate in the clinical interview. This figure was based on estimates calculated for a first-stage sample of 656, in which it was concluded that 27.5% of SRQ high scorers and 12.5% of SRQ low scorers were required to

participate in the second-stage interview (see Appendix 3.1, Fig. 3.1.2). On a weekly basis this meant that 13 SRQ high scorers and 14 SRQ low scorers were randomly selected from the SRQ high and low groups, and administered the clinical interviews.

Given the 25% to 38% drop-out rate in this section of the study, found in both first- and second- pilot studies, a back-up list was compiled of respondents from both SRQ high and low scoring groups. Replacement respondents were selected on the basis of their SRQ score and not matched to the non-attending candidate in terms of gender, age or marital status.

The combined PSE/PAS interview was conducted on four consecutive Saturdays. The clinical interviewers were blind to respondents' SRQ scores. With 27 clinical interviews per week, and 5 clinical interviewers, each interviewer was scheduled to see between 5 and 6 respondents each Saturday. On average, each interview was estimated to last approximately 30 to 45 minutes.

An equal number of respondents was assigned randomly to each clinician. However, some clinicians, especially the more experienced, took a shorter time to conduct the interview, and accordingly conducted more interviews than others. Also because assignment was random, some interviewers had more low-scorers than others, and accordingly these interviews took less time to complete. It was therefore decided that respondents should be seen by any one of the clinical interviewers as soon as an interviewer was free.

The second-stage interview was co-ordinated by the data-co-ordinator. He ensured that respondents were seen by one of the clinical interviewers as soon as an interview space became available; he followed-up respondents who did not arrive to ascertain the reason for their absence, (often respondents had simply forgotten); and arranged to replace those persons who failed to attend from the back-up list. This part of the study was greatly facilitated by the presence of the data co-ordinator who was familiar with the community and knew where and how to contact respondents in a situation where many of the candidates had no telephone. The data-co-ordinator was also in charge of the payment of R20.00 to all candidates at the end of each PSE/PAS interview.

The modified PSE scoring sheet (see Appendix 2.7) was completed by the clinical interviewers at each interview. These forms were collected by the project co-ordinator at the end of each week. Forms were then entered into the PSE-Catego data-base programme and tentative diagnoses made according to the ICD8 or, in the case of the PAS, scored manually, in terms of the "Key to the revised PAS", the DSM-111R, (Gask, 1988). "suggested clinical diagnoses" were also recorded.

Any respondent whom the clinical interviewer felt required immediate clinical attention was referred to the appropriate health care official working in Mamre, usually the clinical psychologist or social worker.

Problems encountered in the main study

Overall, no serious problems were encountered in the main study. However, certain issues did arise which require consideration as they may have influenced the estimated prevalence and SRQ validity coefficients. These issues are briefly reviewed below.

(i) Reduced first-stage sample size

After the first two weeks it became evident that lay-interviewers were unable to meet their target number of 164 weekly interviews. This meant that, by the end of the study, the total number of first-stage interviews would fall short of the required first-stage sample size of 656. Based on the number of first-stage interviews conducted in the first two weeks, it was estimated that the best that could be expected was a weekly rate of 120 interviews, resulting in a total first-stage sample size of 480. In effect, this meant a reduction of 176 in the original estimate.

With a sample of 656, the confidence interval at 95% was estimated at 6% and the standard error of estimate of prevalence at 0.0009 (see Appendix 3.1, Table 3.1.2). Reducing the sample to 480, the 95% confidence interval increased to 7% and the standard error of estimate of prevalence to 0.00123. This meant a loss of approximately 1%, in terms of the confidence that could be placed in the results. Alternatively, the study could be prolonged for an extra 2 weeks to attain the original estimated sample size. Before making a decision, several factors had to be considered.

(i) The extent to which the 6% confidence interval estimated on a sample of 656, was itself a compromise between financial and time constraints, and optimal sample size. And the extent to which a further increase in confidence interval could be tolerated.

(ii) The purpose/s which the results of the study were to serve. In studies where results are to be used as the basis for making costly financial decisions, such as the appointment of a resident clinical psychologist to Mamre, it is necessary that the confidence interval and standard error of estimate be as small as possible. In studies, such as the present, where results are to be used to inform

further research, a less severe confidence interval and standard error of estimate can be tolerated.

(iii) The extent to which the study could sustain extra expenditure required to meet a prolonged interviewing period

It was reasoned that the present research was initiated to gain base-line information of the mental health status of the adult population of Mamre and, as such the study was investigatory. Results could not therefore be considered conclusive and would need to be confirmed by further psychiatric epidemiological research in South Africa. Financial constraints of the study were also pressing. As the increase in the confidence interval was only slight, from 6% to 7%, it was decided expedient to accept a reduction in first-stage sample size.

A smaller first-stage sample meant not only a smaller second-stage sample but also a reduction in the number of SRQ high and low scorers required to participate in the second-stage interview. The re-estimated optimal second-stage sample was 82, and the number of SRQ high scorers and low scorers 40 and 42 respectively. Clinical interviewers had however, achieved their weekly target number of 27 clinical interviews, based on a first-stage sample of 656, during the first 2 weeks. It was decided therefore to continue to select a weekly number of 13 SRQ high scorers and 14 SRQ low scorers, as opposed to 10 SRQ high and 11 SRQ low scorers for the clinical interview.

Although not a problem as such, a further factor was identified which had a bearing on second-stage sample size. Several original respondents who did not keep their appoints were replaced from the back-up list. During the course of the day however, some of these respondents had arrived for the interview. As they could not ethically be turned away, the interview was given. This meant that the second-stage sample was increased from 111 to 121.

(ii) Falsification of interview information

At the end of the second week, one of the lay-interviewers was suspected by the data-co-ordinator to be falsifying information on some of the questionnaires. Familiarity with the community alerted the data-co-ordinator to the inaccuracy and inconsistency of some of the reported responses. Closer examination of the particular interviewer's forms also showed a singularly large number of "non-consents" to the second-stage clinical interview. Thirty-eight out of a total of 73 "non-consents" for the 2 weeks were from interviews

conducted by this interviewer. The data-co-ordinator and the project co-ordinator were able to contact 15 of those who had been recorded as "non-consent". Of these, 11 reported they had not been interviewed.

It was assumed that all respondents who had consented to the second-stage interview had been interviewed, and this was confirmed in a random check of "consent" forms. The large number of false "non-consents" may have had a bias limiting effect on results and therefore on the validity of the research findings. However, as the majority of "non-consents" were seen by one interviewer, and represented a cross-section of the sample, it was argued that the effect on the overall sample would not be biased. The lay-interviewer concerned undertook that no further irregularities would occur, and henceforth all interviews were randomly checked. It was concluded that no further falsification took place.

(iii) Duration of the study

It was planned that the study, including the pilot, would be completed in 6 weeks. However, the need for a second pilot study, and the intervention of 2 public holidays, meant there was an increase of an additional 5 weeks. For 2 of these weeks, the entire study came to a halt as a result of public holidays, and this in itself, necessitated refresher courses for lay-interviewers. Prolongation of the study, resulted in a decrease in motivation in two of the lay-interviewers in the last week and a half of the study and manifested itself in a decrease in the number of interviews given. This factor is considered in relation to the Mamre Community Health Project's aim to develop the skills of the community in Chapter 12, p 141.

(iv) Length of questionnaire

Ideally, to obtain the benefits of a brief screening questionnaire, particularly when used in a developing community, the SRQ should have been administered on its own. However, for reasons already discussed (Chapter 9, p. 83) it was decided to include both the GHQ-28 and additional questions in the first-stage questionnaire. This meant that the entire first-stage interview took approximately 20 minutes, opposed to the 5 minutes required to complete the SRQ alone. While respondents showed no evidence of irritation at the length of the questionnaire, lay-interviewers found its length exhausting, especially when administered over a protracted time.

Results are reported in the following chapter, and the impact of the problems identified above on these results and on recommendations for future research are discussed in Chapters 12 and 13.

11 RESULTS

SAMPLE REPRESENTATIVENESS

As research findings regarding a possible association between mental disorder and commonly researched socio-demographic variables are controversial, the first-stage sample was not stratified. Nevertheless, it was considered prudent to assess the degree to which the unstratified randomly selected first- and second-stage samples were representative of the adult population of Mamre, given the ongoing debate regarding this association. Representativeness was assessed in terms of the concordance between the rates for specific demographic variables found in the 1986 Mamre census and in the present research. Table 11.1 compares the rates obtained in the 1986 census and those in the first- and second-stage samples for gender, age, marital and employment status, education and socio-economic status.

Table 11.1 Comparison of the composition of first- and second-stage samples with the 1986 Mamre census, in terms of specified socio-demographic variables. (All values given as a percentage)

Variable	1986 Total Population n=3032 %	1991 First-stage sample n=481 %	1991 Second-stage sample n=121 %
Gender			
male	49	43	41
female	51	57	59
Age			
15-19	18	13	9
20-24	16	12	15
25-34	24	23	19
35-44	18	20	24
45-54	10	13	12
55-64	8	11	13
65-70	3	4	2
70+	3	4	6
Marital status			
single	46	45	42
married	47	47	51
divorced	1	1	-
widowed	6	7	7
Employment			
unemployed	35	58	69
employed	65	42	31
Socio-economic status			
professional	5	5	2
white collar	7	16	19
skilled manual	12	9	3
routine manual	45	45	49
unskilled manual	31	25	27
Educational status			
nil	3	1	2
primary education	42	34	31
secondary education	55	65	67
post-school training	4	10	7

Slight variations evident in the second-stage sample may have been the result of random sampling error, or the non-proportional sampling scheme employed. The latter meant that a higher number of SRQ screened positives (45.6%) than screened negatives (18%) were included in the second-stage interview (see Chapter, 10, Fig. 10.1).

Two fairly small differences between the 1986 and 1991 samples were identified. These included:

(i) the over-representation of the 55 to 64 and 70+ year age groups and the under-representation of the 15 to 19 year age group in the 1991 sample; and

(ii) the increased number of unemployed in 1991, compared to the number of unemployed in 1986.

Although not undisputed, evidence suggests that the 55 to 64 and 70+ year age groups, as well as unemployed men, may be associated with a higher rate of mental disorder (Goldberg & Huxley, 1992; Finlay-Jones & Ekhardt, 1984). The sample could therefore be yielding an inflated prevalence figure. The unemployment rate increased enormously in Mamre between the 1986 census and this study, owing to layoffs in industries in nearby Atlantis. Investigation of the prevalence of disorder in terms of the above categories indicated that there was no significant association either between age and mental disorder, or between unemployment in men and mental disorder in the present sample. It was concluded that the variation in the second-stage sample arising from these groups, did not affect rates of disorder detected.

Both first and second-stage samples appear to be adequately representative of the population from which they were drawn.

A study conducted in Mamre in 1993 after the data for the present study were collected (Salmon, 1993) indicated that the population frame used in the present study may not have been comprehensive. A particular area, populated by persons of a very low status socio-economic group, (referred to as the "wood gatherers") may have been only partially represented. The community reported that alcohol abuse was a serious problem in this group. Its exclusion therefore, may have affected both sample representativeness, and estimated prevalence rates.

RESPONSE RATE

Non-consent to first-stage interview

78 persons refused to participate in the study. Of these: 14 said they were not interested; 21 said they were too busy; 10 were under the influence of alcohol; and

33 gave no reason.

Data was not available to indicate if those who refused to participate belonged to any particular socio-demographic group.

A further 21 first-stage questionnaires were inadequately completed and had to be scrapped. Sample representativenss was not effected by the withdrawal of these respondents as no association was found either between incompleted questionnaires and lay-interviewers administering the screen, or between incompleted questionnaires and respondents.

Non-consent to the second-stage interview

Of the 481 persons screened, 70 (15%) refused to participate in the second-stage clinical interview. Of these: 23 were working; 18 were going away; 15 were busy; were not well enough; 7 were not interested; and 4 gave no reason.

As the clinical interview took place on a Saturday, some non-consent respondents may have been unavailable due to working over-time or playing sport. Non-consent respondents' SRQ scores were not consistently different from the general population (Chi-square = 1.758, $p > 0.10$), nor were they distinguishable in terms of any socio-demographic variable. It was therefore concluded that their non-participation in the clinical interview did not have a biasing effect on results.

Failure to attend the second-stage clinical interview

29% of respondents failed to attend the second-stage clinical interview. 57% of non-attenders were male and 63% had low SRQ scores. Where possible, respondents were contacted to establish the reasons for their failure to attend the clinical interview. Reasons given included:

- forgot appointment;

- unable to find the interview locality;

- something else had come up.

No single factor which accounted for the second-stage drop-out rate was identified.

Random selection of replacement respondents for second-stage interview from the back-up list

35 replacement names were randomly selected from a back-up list of high and low SRQ scorers on the day of interview. Replacement respondents were not matched with those they were selected to replace, other than in their SRQ score. This decision was based on the fact that the original sample was not stratified in terms of specific variables.

As no distinguishing factors were found associated with second-stage non-attenders, and as there was sufficient reason not to match replacement respondents with original second-stage participants, (other than in terms of their SRQ score), it was concluded that prevalence rates were affected neither by non-attenders, nor by the random substitution of respondents from the respective high and low SRQ groups.

Important issues which need to be considered when reporting results

Before reporting prevalence estimates, issues underlying the research study, and which qualify the findings, need to be reported. These issues are discussed more fully in the following chapter and in Appendix 4.1, but are briefly mentioned here, as they influence the order and manner in which results are presented.

(i) Symptom 58 on the PSE was erroneously rated positive on one of the PSE score sheets by one of the clinical interviewers. The effects of this error on prevalence estimates and on SRQ validity coefficients are assessed in Appendix 4.1.

(ii) Tenenbein's (1970) statistical equation (see Appendix 3.4), which was used in this study to estimate prevalence, reflects the efficiency of both the first and second-stage instruments. It follows that prevalence estimates of mental disorder therefore **depend on the validity of the two-stage case-identifying instruments** used to detect psychiatric morbidity. Evidence of the PSE's efficiency as a case-finding instrument in the community has been cited in research by Gillis et al., (1982) and Orley and Wing (1979) (see Chapter 6, p. 55). However, in spite of the SRQ's high validity coefficients in the second pilot (sensitivity 1.00, specificity 0.88) the small size of the sample meant that its validity in the context of the present research, is still a matter for investigation.

This being so, it is clear that the simultaneous achievement of two of the primary goals of the present research ie,

(i) to estimate prevalence of mental disorder in the adult population of Mamre; and

(ii) to assess the validity of the SRQ as an efficient first-stage instrument in the context of a South African community sample,

depends on the SRQ having satisfactory validity in the present context.

Results pertaining to the validity of the SRQ are therefore reported before those estimating prevalence rate.

VALIDITY OF THE SRQ AS A FIRST-STAGE SCREEN IN A COMMUNITY STUDY IN MAMRE

Calculation of SRQ validity coefficients was based on a comparison between scores obtained by the SRQ and the PSE on the 121 respondents who participated in both the first- and second-stage interviews. A two-by-two table showing the scores on the two instruments is given in Table 11.2.

Table 11.2 Two-by-two table of the relationship between scores on the SRQ and the PSE, for the second-stage sample, n=121.

SRQ (7/8 threshold)	**PSE-ID5**		
	Low non-case	**High** case	**Total**
Low	52 true negative	12 false negative	64
High	28 false positive	29 true positive	57
Total	80	41	121

As a larger proportion of SRQ screened positives (46.5%) than SRQ screened negatives (18%) were selected to participate in the second-stage interview, it was necessary to weight back the estimated validity coefficients to the original first-stage sample.

Of the 41 cases identified by the PSE-ID5, 29 (51% of the SRQ positives) were SRQ positive. Of the 80 non-cases identified by the PSE-ID5, 52 (81% of the SRQ negatives) were low scorers on the SRQ. Accordingly, sensitivity was estimated at **0.71**, and specificity at **0.65**. However, when weighted back to the original sample, sensitivity dropped to **0.49**, while specificity rose to **0.82**. The drop in weighted sensitivity and increase in specificity may be explained by the high false positive rate (49%), and low false negative rate (19%) of the SRQ in this study. Calculations of both unweighted and weighted validity coefficients of the SRQ are found in Appendix 3.3.

Unweighted SRQ validity coefficients were within the range of sensitivity and specificity found in other similar studies (Dhadphale et al., 1982; Harding et al., 1980; Kortman & Ten Horn, 1988; Marie & Williams, 1985; Sen et al., 1987). However, while the weighted specificity was well within the range reported in the above studies, the weighted sensitivity of the SRQ was considerably below the level necessary to justify the use of the SRQ as a valid first-stage screening instrument in this context. Further discussion on the appropriateness and efficiency of the SRQ as a first-stage screen in this context, may be found in Chapter 12, p. 127.

The SRQ was able to correctly screen 24% of true positives and 43% of true negatives, which, taken together, meant that the SRQ correctly identified **67%** of PSE cases and non-cases. Furthermore, the SRQ identified only 12 (10%) false negatives. The relationship between the SRQ and PSE was significant (Chi-square 13, $p< 0.001$). The relationship between the PSE and SRQ was attenuated however, by the large number of SRQ false positives, 28 (23%, or 49% of SRQ high scorers). Examination of the **false negatives** with **true positives** indicated that there was no difference between the two groups in terms of gender, age or employment status. However, false negatives were slightly more likely than true positives, to be single or widowed (Chi-square 5.6, $p< 0.06$), and more likely to have reported being currently healthy (Chi-square 6.7, $p< 0.010$). Investigation of the **false positives** with **true negatives** showed no association between false positives and demographic variables, or reported physical health status. The lack of an association between false positives and physical health status is contrary to the findings of Finlay-Jones and Murphy (1979) which observed that false positives on the GHQ tend to be associated with severe or chronic physical illness.

PREVALENCE OF MENTAL DISORDER IN THE ADULT POPULATION OF MAMRE

Just as the validity coefficients of the SRQ had to be weighted to reflect the total first-stage sample, so too prevalence had to be weighted to reflect the prevalence of disorder in the population as a whole. Weighted prevalence estimates were calculated according to Tenenbein's statistical equation (see Appendix 3.4).

PREVALENCE OF DIAGNOSTIC CATEGORIES OF MENTAL DISORDER ACCORDING TO THE PSE-CATEGO, (TENTATIVE DIAGNOSIS ICD8), FOUND IN THE ADULT POPULATION OF MAMRE

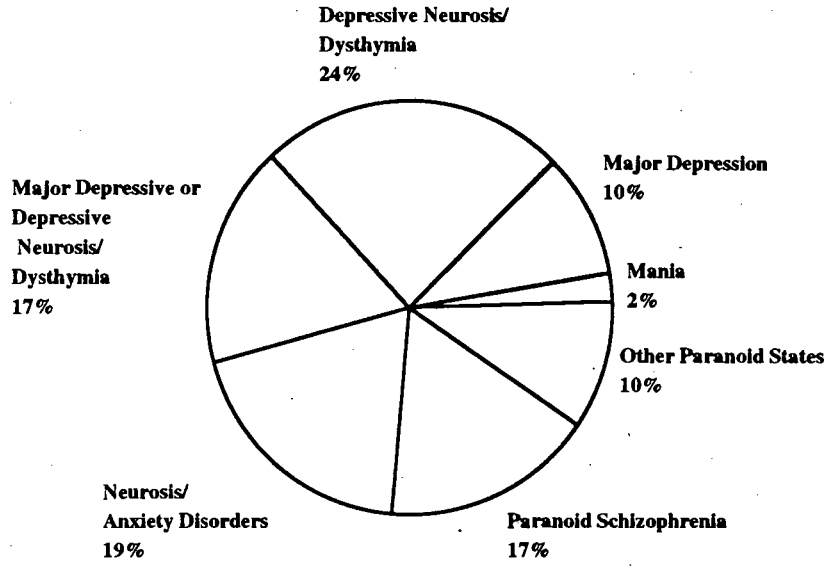
ICD8 tentative diagnosis were determined according to the PSE's Catego computer programme. For two of the diagnostic categories, ie, depressive neuroses and neuroses, the comparable DSM111R diagnosis, dysthymia and anxiety disorders, are given in the present study in addition to the ICD8, for ease of comparison with PAS findings.

According to the PSE-Catego (ICD8), of the 41 identified cases, the following tentative diagnoses were made :

7 paranoid schizophrenia, 1 mania, 4 paranoid states, 4 major depression, 10 depressive neuroses, (dysthymia), 7 major depression or depressive neuroses (dysthymia), and 8 neuroses, comparable to anxiety disorders on the DSM111R.

The distribution of categories of disorder in the 41 cases given a PSE, ICD8 tentative diagnoses are shown in the pie-chart, Fig. 11.1.

Fig. 11.1 PSE-Catego tentative diagnosis, ICD8, n=41 (% rounded to whole numbers)



The detailed weighted results, reflecting the expected prevalence of these disorders in the general adult population of Mamre, and indicating Catego class and ICD8 tentative diagnosis, are given in Table 11.3. Of the 7 cases given a PSE-Catego tentative diagnoses of paranoid schizophrenia, 1 respondent was inaccurately diagnosed as a result of an error in rating symptom 58 (see Appendix 4.1). This meant that the weighted prevalence rate for paranoid schizophrenia shown in Table 11.3 is slightly higher than if this case had not been included in the estimate. However, as the weighted prevalence of paranoid schizophrenia was found to be minimally affected by this error, no adjustment was made to the original results.

Table 11.3 Prevalence of PSE-Catego classes (with ICD8 and two DSM111R equivalents) in the adult population of Mamre.

Tentative ICD8 Diagnosis	**Code**	**Catego class**	**Unweighted prevalence 2nd stage sample** n=121 (%)	**Weighted prevalence adult population** n=3032 (%)*
Paranoid schizophrenia	295.3	S	6	5
Mania	296.1	M	1	0.1
Other paranoid states	297.9	P	3	3
Major depression	296.2	D	3	2
Depressive neurosis or Dysthymia on the DSM111R	300.4	N	8	6
Major depression/ Depressive neurosis or Dysthymia on the DSM111R	296.2/ 300.4	R	6	5
Neuroses or Anxiety disorders on the DSM111R	300.0-3	A	7	6
Total %			34	27.1*

* Weighted to represent the prevalence of specific diagnostic categories in the total adult population, n=3032

The majority of PSE cases were given a tentative diagnoses of either depressive (51.2%) or anxiety related disorders (19.5%). These results appear to confirm Goldberg and Huxley's (1992) observation that these disorders comprise the greatest proportion of mental disorders

found in the community. They also compare with rates found in other similar community-based studies (Duncan-Jones & Henderson, 1978; Hodiamont et al., 1987; Orley & Wing, 1979).

Examination of Table 11.3 suggests that the rate of paranoid schizophrenia in particular, is very high for a general population sample, and therefore requires further investigation. (see Appendix 4.2).

None of the psychiatric cases on the patient list of the psychiatric community sister working in Mamre during the year previous to the study (1990), happened to be included in the second-stage sample. It was therefore not possible to assess the extent to which the PSE, as a case finding instrument in the second-stage of the study, was able to identify known psychiatric cases.

INVESTIGATION OF POSSIBLE ASSOCIATION BETWEEN PSYCHIATRIC CASENESS AND SPECIFIED SOCIO-DEMOGRAPHIC VARIABLES

Initial examination of one-way frequency tables indicated certain socio-demographic areas for further enquiry. In particular these areas included, gender, marital and employment status, work type (or other occupation if not employed in the formal sector), alcohol abuse, physical health status and rate of consultation. Data relevant to these areas was analyzed using the SAS computer programme. P values less than 0.5 were considered significant.

Results were based on the 121 respondents who participated in the second-stage clinical interview, and were weighted back to the original sample; they are therefore applicable to the population as a whole.

With the exception of physical health status, no statistically significant associations were found between psychiatric morbidity and the demographic variables specified, and this held even when results were analysed by gender. Results are reported nevertheless, as they may be informative in terms of providing qualitative information relevant to the provision of phc services. (see Appendix 3.2).

In terms of identifying a distinguishing demographic profile associated with mental disorder, no undisputed factor/s could be identified. In general, results indicated that:

(i) prevalence rates of mental distress were slightly higher among women than men. The difference was not significant, however, as there was an overlap in their confidence intervals. In the present study, similarity in prevalence rate between the sexes may reflect the inherent stresses of life-style in Mamre;

(ii) there was no association between presence of mental disorder and age. This may have been a result of the small sample size, which did not allow for in-depth cross-tabulations;

(iii) prevalence of mental disorder was comparable in the single and married groups, but was considerably higher among widows. Although this association accords with other findings (Goldberg and Huxley, 1992), it has to be considered in the light of the very small number of respondents in this class, n=8.

Examination of the data in terms of gender, showed no difference between the rate of mental disorder in single men and women and married women. However, there was a lower rate in married men, which is consistent with findings from other studies (Cox et al., 1987). Samples were too small to permit analysis in terms of gender for the categories of separated, divorced and widowed persons;

(iv) rate of mental disorder was nearly equivalent in both the employed and unemployed, including unemployed males. Unemployed females however, had a lower rate of morbidity than employed females. This supports the argument that employment may not be a factor protecting against mental disorder in women (Cleary & Mechanic, 1983), probably due to the increased stress of occupying dual roles;

(v) prevalence rates were not calculated for all occupational classes, as there were not enough observations in each of the various groups. Prevalence rates were highest among routine non-manual and semi-skilled manual and unskilled manual workers;.

Among the unemployed, prevalence was highest for pensioners and lowest for housewives. The lower rate for housewives is consistent with the lower rate found among unemployed females reported above;

(vi) in an earlier study (Katzenellenbogen et al., 1988), Mamre residents reported that alcohol was a serious health problem in the community. It was recognised at the outset of the study that obtaining accurate information on alcohol consumption might be difficult (Kinney & Leaton, 1978). Although questions related to alcohol abuse were couched in a battery of questions concerning general liquid intake, results indicated that this "ploy" had not been successful. Only 3% of respondents reported that they "drank" or had a problem with

alcohol. This rate was clearly inconsistent with the community's perception of alcohol prevalence.

INVESTIGATION OF POSSIBLE ASSOCIATION BETWEEN PSYCHIATRIC CASENESS AND PHYSICAL HEALTH STATUS, AND RATE OF CONSULTATION WITH PHC SERVICES

Association between physical health status and mental disorder

A significant **trend** of 95% probability was discerned between physical health status and prevalence of mental disorder. This suggested that mental disorder was likely to be lowest among those reporting physical health, and highest among those reporting current physical ill health. This trend held when the data were reanalyzed in terms of gender. In all instances, prevalence of mental disorder was higher in women than in men. These findings are consistent with the results from other studies, which show that psychological illness is between 1.5 to 3 times more likely when physical ill health is present (Weyerer, 1990)

The direction of the relationship between mental disorder and physical illness is complex and research in this area has been controversial (Schulberg et al., 1987 a, b, in Goldberg & Huxley, 1992). Information has not been collected in such a way to make a clear indication of whether mental disorder preceded physical ill health, or vice-a-versa.

Types of illness reported ranged widely. Colds and influenza, back pain, hypertension, stomach pain, and headaches, were the most commonly reported illnesses. The planned investigation of an association between hypertension and depression was not possible, as only 2% of those reporting physical illness, claimed to be suffering from high blood pressure.

Association between mental disorder and rate of consultation

Just over a third (n=47) of the 121 second-stage respondents reported being currently ill (n=29) or having been ill during the last month (n=18). Of these, 30 (70%) sought treatment. The most common site of service was a private doctor, followed by consultation with a doctor at the day hospital.

There was no significant difference in the health care utilization behaviour between PSE cases and non-cases. Thus, people who scored high on the PSE and reported current illness or illness within the past month, were no more likely to seek health care than people who scored low on the PSE: Rate ratio, 1.28 (95% CI 0.81-2.03). However, this rate ratio was not homogeneous when stratified by **timing** of illness. For those who reported current illness, the rate ratio for health-care utilization in PSE positives, as compared to PSE negatives, was 0.81

(95% CI 0.5-1.32). For respondents reporting illness within the previous month (but not currently), the rate ratio of PSE positives to PSE negatives was higher: 3.14 (95% CI 1.14-8.64). The difference in rate ratios for health care seeking behaviour of those currently ill (0.81), and those ill during the last month, but not currently ill (3.14), could be explained by a delay in seeking health care by those currently ill, to some time in the weeks following the administration of the PSE. For those ill during the last month (but not currently ill), those found to be mentally ill according to the PSE were roughly three times more likely to have sought health care than those who were not mentally ill.

Inspection of PSE high scorers and those reporting current illness or illness in the past month, indicated that the rate ratio for health care consultation among these two groups was not significant (1.25 (95% CI 0.9-2.58)). Timing of illness therefore, does not appear to affect rate of consultation in those who score high on the PSE.

Attitudes to health care services

Health care services available in Mamre in 1991 included: a district surgeon who visited Mamre two or three times a week; three general practitioners who visited Mamre on alternate days for one session per day, each lasting between two and three hours; the Mamre clinic, staffed by senior professional nurses and a staff nurse; a psychiatrist from a psychiatric hospital in Cape Town, who visited the community once every three months; a psychiatric sister and staff nurse from the same institution, who attended the community once a month, and a half-time psychologist from the Department. of Psychology, University of Cape Town.

Results indicated that 70% of persons who consulted, consulted with the GP in Mamre, or the doctor at the day hospital in Atlantis. 95% reported that they were satisfied with the care they received. Of the 29 who consulted with either a GP, surgeon or district sister in Mamre, 54% were given a medical examination.

Overall rate of satisfaction with the phc service has to be considered in the light of the context of the research. As respondents were aware that the study was conducted under the auspices of the University of Cape Town, and was linked to the Mamre Community Health Project, it is possible that they did not feel at liberty to criticize the phc services. In-depth interviews conducted in Mamre (Rogers, 1993) suggest that satisfaction with the services may be far more equivocal than the results of the present research suggest.

Reasons given for not consulting with the health care services in Mamre included:

- services too expensive (11%);,
- lack of transport, (reported by 2 respondents); and

- unavailability of the required service (reported by one respondent, who was suffering from stomach pain).

DIFFERENCES IN PREVALENCE RATE WHEN THE SRQ IS USED WITH THREE DIFFERENT CRITERION MEASURES

Weighted prevalence rates of mental disorder, measured in the same population at the same time, and using the same first-stage screening instrument, the SRQ, but three different criterion measures, are given in Table 11.4. This shows that prevalence was lowest when the SRQ was used with "suggested clinical diagnosis" (17%), and highest when the PSE was the second-stage criterion (27%).

Table 11.4 Weighted prevalence estimates of mental disorder in the adult population of Mamre, when the PSE-Catego, tentative diagnosis, ICD8, the PAS-DSM111R, and clinicians's "suggested diagnosis" were used as second-stage criterion measures, n=121

Criterion	Prevalence	Standard error of estimate (se)	Lower confidence interval	Upper confidence interval
PSE-Catego tentative diagnosis, ICD8	0.27	0.04	0.19	0.35
PAS-DSM111R	0.23	0.04	0.16	0.30
"suggested clinical diagnosis"	0.17	0.03	0.10	0.23

When either the PAS-DSM111R or "suggested clinical diagnosis" were used, prevalence rates were within the range found in other community based studies, in which standard psychiatric instruments were used: 7.0-26% (Bland, 1988, in Parry, 1992; Cheng, 1988; Duncan-Jones & Henderson, 1978; Hodiamont et al., 1987; Orley & Wing, 1979; Vazquez-Barquero et al., 1987). With the exception of the "suggested clinical diagnosis", prevalence estimates were however, at the upper end of the range of the estimate of prevalence in community studies (Cheng, 1988; Duncan-Jones & Henderson, 1978; Hodiamont et al., 1987; Orley & Wing, 1979; Vazquez-Barquero et al., 1987). This supports other reports that prevalence of mental

disorder in developing communities may be as high, if not higher, than that found in western industrialized countries (Ben-Arie et al., 1983; Ben-Tovim, 1987; Cheng, 1988; Gillis et al., 1968; Harding et al., 1980; Hollifield et al., 1990; Orley & Wing, 1979; Reeler, 1987a; Vazquez-Barquero et al., 1987). (see Chapter 3, p. 22).

The PSE is generally regarded by professionals in the field of psychiatry as a fairly conservative psychiatric instrument (Goldberg & Huxley, 1992; Wing, 1983). Compared to community studies using the DSM-111R as the criterion, the PSE-Catego tends to yield lower prevalence rates (Goldberg and Huxley, 1992). In the present study, however, prevalence of mental disorder was highest when the PSE was used, and lowest when the "suggested clinical diagnosis" was used. The higher rate recorded by the PSE needs to be considered in light of the data collection methods used. Factors which may have inflated the prevalence estimate when the PSE was the criterion measure are examined in Chapter 12, p. 130.

CONCORDANCE BETWEEN THE THREE CRITERION MEASURES: THE PSE-CATEGO, THE PAS-DSM111R AND "SUGGESTED CLINICAL DIAGNOSIS" AND PSYCHIATRIC CASENESS

In spite of efforts to increase standardization, both in clinical interviews (PSE and PAS), and in diagnostic systems (ICD and DSM111), agreement between instruments regarding psychiatric caseness is poor, and even worse in the case of individual categories (Dean et al., 1983; Dohrenwend, 1990; Goldberg & Huxley, 1992; Grayson et al., 1990; Van den Brink et al., 1989). Disagreement is exacerbated when these instruments are used in community studies where the level of mental disorder is both less severe and less stable (Goldberg & Huxley, 1992). The Venn diagram (Fig 11.2) illustrates the extent of agreement between the three criterion measures in terms of psychiatric caseness. (The respondent inaccurately rated as a case is excluded from the number of PSE cases in the diagram).

Fig 11.2 Venn diagram showing the extent of agreement regarding psychiatric caseness between the three criterion measures, the PSE-Catego, the PAS-DSM111R and the "suggested clinical diagnosis", n=49

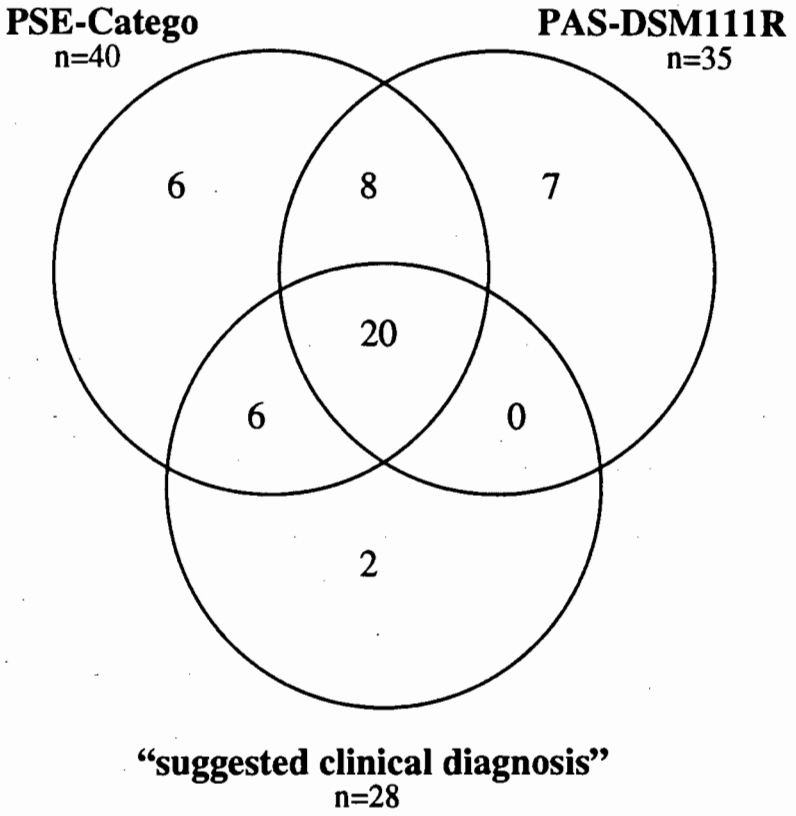


Fig 11.2 shows agreement by all three criterion measures on 20 cases. Investigation of a possible distinguishing demographic profile among the 20 cases did not indicate any significant association between cases and gender, age, marital or employment status. A significant association (Chi-square = 16, $p< 0.0010$) was found between cases and physical health status. Of the 20 cases, 38% reported currently illness, 17% reported illness within the last month, and only 6% reported current health. This suggests that criterion measures were most likely to agree on psychiatric caseness when assessing respondents who reported physical **ill health**, and least likely to agree on caseness when assessing respondents who reported being **healthy**. Although this finding must be treated with caution because of the small sample size (n=20), the conclusion held when those who were <u>not</u> identified as cases by all three criterion measures were assessed (ie, of the fifteen cases not identified as cases by all three criterion measures, only two reported being currently ill or ill in the past month).

The extent of the association between each of the criterion measures was estimated using the Chi-square statistic, and was found to be significant for all three criterion measures. (A more

stringent value for p, 0.1, was used to allow for the repeated comparison of the Chi-square statistic).

(i) PSE and PAS, Chi-square 84, $p < 0.001$;

(ii) PSE and "suggested clinical diagnosis", Chi-square, 113 $p < 0.001$;

(iii) PAS and "suggested clinical diagnosis", Chi-square, 78 $p < 0.001$.

Although agreement on psychiatric caseness between the criterion measures was only moderate (between 49% to 55%), the association reached significance because of the fairly high level of agreement between the instruments regarding non-caseness (77% to 79%).

Concordance between the three criterion measures and diagnostic categories

Table 11.5 shows the weighted prevalence of specified diagnostic categories of mental disorder as classified according to the PSE-Catego tentative diagnosis, ICD8, the PAS-DSM111R and "suggested clinical diagnosis". (The original weighted estimate for the PSE-Catego's tentative diagnosis of paranoid schizophrenia is reported in this table, as it was concluded that the inclusion of the one respondent inaccurately assigned a tentative diagnosis of paranoid schizophrenia did not affect the original estimate of weighted prevalence (see Appendix 3.4)).

Table 11.5 Weighted prevalence of diagnostic categories of psychiatric disorder in the adult population of Mamre, n=481: three alternative diagnostic systems applied (PSE-Catego and tentative diagnosis ICD8, PAS-DSM111R and "suggested clinical diagnosis".

Diagnostic system and assigned diagnosis		**Frequency** n=121	**Unweighted prevalence** (%)	**Weighted prevalence total pop.** (%)
PSE-Catego, tentative diagnosis ICD8				
Paranoid schizophrenia	295.3	7	6	5
Mania	296.1	1	1	0.1
Other paranoid states	297.9	4	3	3
Major depression	296.2	4	3	2
Depressive neurosis/ Dysthymia on DSM111R	300.4	10	8	6
Major depression/ Depressive neurosis/ Dysthymia on DSM111R	296.2/ 300.4	7	6	5
Neurosis/ Anxiety disorders	300.0-3	8	7	6
PAS-DSM111-R				
Major depression		22	18	13
Dysthymic disorder		1	0.8	0.5
Agoraphobia		11	9	9
Social phobia		1	0.8	0.5
"suggested clinical diagnosis"				
Major depression		9	7	4
Dysthymic disorder		3	2	6
Generalized anxiety		12	10	7
Subcultural delusions		1	0.8	0.01
Substance abuse		1	0.8	0.01
Alcohol abuse		1	0.8	0.01

There was acceptable agreement among the three criterion measures with regard to diagnostic categories, in terms of depressive and anxiety related disorders. Depressive disorders made up between 43% and 68% of those respondents identified as psychiatric cases, and anxiety disorders between 20% and 58%.

COMPARISON OF SRQ VALIDITY COEFFICIENTS WHEN THREE DIFFERENT CRITERION MEASURES, THE PSE, PAS AND "SUGGESTED CLINICAL DIAGNOSIS" WERE USED AS THE GOLD STANDARD

The investigation of possible differences in the validity coefficients of the SRQ, when used with different criterion measures in the same sample population, was one of the secondary goals identified in the reformulated aims. Unweighted and weighted validity coefficients of the SRQ, when used with the PSE-ID5, the PAS-DSM111R and "suggested clinical diagnosis" are given in Table 11.6. The SRQ's positive predicative values (PPV) and overall misclassification rates (OMR) were calculated for the different prevalence rates, as estimated by the three criterion measures: (estimated prevalence PSE-ID5 : 27%, PAS-DSM111R: 23%; and "suggested clinical diagnosis": 17%).

Table 11.6 Comparison of the unweighted and weighted validity coefficients of the SRQ when used with the PSE-ID5, the PAS-DSM111R and "suggested clinical diagnosis" in the same community-based population.

SRQ, threshold 7/8

Criterion	Sensitivity	Weighted sensitivity	Specificity	Weighted specificity	Positive Predictive Value	Overall Misclass-ification Rate
PSE-ID5	0.71	0.49	0.65	0.82	0.50	0.27
PAS/DSM111R	0.71	0.50	0.63	0.81	0.45	0.26
"suggested clinical diagnosis"	0.74	0.44	0.60	0.79	0.36	0.24

Table 11.6 shows that regardless of the criterion measure used, SRQ weighted and unweighted validity coefficients are comparable. This suggests that the use of alternative criterion measures in this population does not result in improved SRQ validity coefficients.

Comparison of SRQ validity coefficients with validity coefficients of the GHQ-28 and CGHQ-28 when used with the PSE-ID5, the PAS-DSM111R and "suggested clinical diagnosis"

The SRQ's low weighted sensitivity raises the question of the appropriateness of the SRQ as a first-stage screen in the context of the present research. A more appropriate screen is investigated by comparing the validity coefficients of the SRQ with those obtained for the GHQ-28 and CGHQ-28, when used with the three criterion measures. Table 11.7 shows unweighted and weighted sensitivities and specificities, false negative rates (FNR), positive predictive values (PPV) and overall misclassification rates (OMR) for the SRQ, GHQ-28 and CGHQ-28 for all three criterion measures.

Table 11.7 Comparison of the weighted validity coefficients, of the SRQ, GHQ-28 and CGHQ-28, when used with three criterion measures, the PSE-ID5, the PAS-DSM111R and "suggested clinical diagnosis".

Criterion: PSE-ID5						
Instrument	**Cut/off**	**Sensitivity**	**Specificity**	**False Negative Rate** %	**Positive Predictive Value** (prevalence 0.27)	**Overall Misclassification Rate** (prevalence 0.27)
SRQ	7/8	0.49	0.82	9	0.50	0.27
GHQ-28	3/4	0.45	0.75	8	0.40	0.33
CGHQ-28*	7/8	0.49	0.82	5	0.50	0.27
*CGHQ denotes the alternative 0,1,1,1, method of scoring						
Criterion: PAS-DSM111R						
Instrument	**Cut/off**	**Sensitivity**	**Specificity**	**False Negative Rate** %	**Positive Predictive Value** (prevalence 0.23)	**Overall Misclassification Rate** (prevalence 0.23)
SRQ	7/8	0.50	0.81	11	0.45	0.26
GHQ-28	3/4	0.50	0.75	7	0.43	0.23
CGHQ-28	7/8	0.56	0.82	4	0.53	0.24
Criterion: "suggested clinical diagnosis"						
Instrument	**Cut/off**	**Sensitivity**	**Specificity**	**False Negative Rate** %	**Positive Predictive Value** (prevalence 0.17)	**Overall Misclassification Rate** (prevalence 0.17)
SRQ	7/8	0.44	0.79	12	0.36	0.24
GHQ-28	3/4	0.54	0.75	9	0.31	0.29
CGHQ-28	7/8	0.65	0.81	7	0.41	0.22

Examination of sensitivity values indicates that sensitivity is best when the CGHQ-28 is used with "suggested clinical diagnosis". Specificity is comparable for both screens (including the CGHQ-28 method of scoring) when used with any one of the three criterion measures. In general, the CGHQ-28 gives better validity coefficients than the SRQ, when used with any of the three criterion measures. The CGHQ-28 also has a lower rate of false negatives (FNR) (4% to 6%) than either the SRQ (9% to 12%) or the GHQ-28 (7% to 9%). The screens' low false negative rate is particularly important in a studies such as the present one, where the primary goal is to estimate base-line prevalence of mental disorder in a community sample. A low false negative rate suggests that the screen is sufficiently "sensitive" to permit the identification of the majority of potential psychiatric cases.

The weighted PPV for both the SRQ, GHQ-28 and CGHQ-28 is very low, (0.31 to 0.41), and is compounded by a fairly high OMR (0.22 to 0.29). Although the OMRs are only slightly higher than those found in studies with similar prevalence rates (0.21 to 0.26) (Bellantuono et al., 1987; Cheng, 1986; Tarnopolsky et al., 1979), the PPVs are well below the range found in these same studies (0.60 to 0.80). However, they are comparable to values found in Mann et al's. study (1983), (0.48).

CONCORDANCE REGARDING POTENTIAL CASES IDENTIFIED BY THE SRQ GHQ-28 AND CGHQ-28 IN THE SECOND-STAGE SAMPLE, n=121

Concordance between high scorers on the SRQ, GHQ-28 and CGHQ-28 is shown in the Venn diagram, Fig 11.3.

Fig. 11.3 Venn diagram showing the extent of agreement regarding high scorers on the SRQ, GHQ-28 and CGHQ-28, n=121

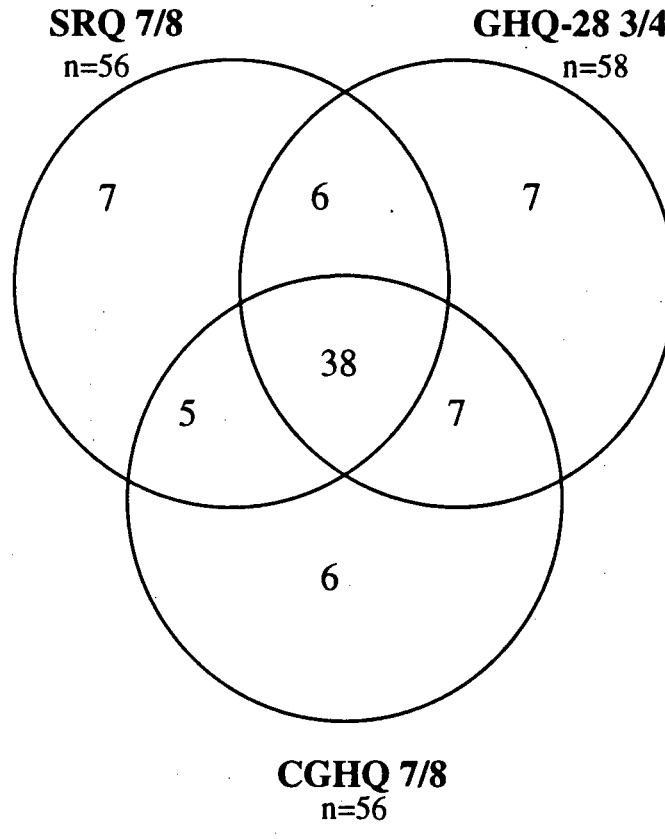


Fig 11.3 shows that of the 121 second-stage respondents, the SRQ, GHQ-28 and CGHQ-28 concurred on 38 potential cases. The Chi-statistic of association was significant between all screens ($p<0.1$):

(i) Chi-square value between SRQ and GHQ-28, 377 $p< 0.001$;

(ii) Chi-square value between SRQ and CGHQ-28, 5570 $p< 0.001$; and

(iii) Chi-square value between GHQ-28 and CGHQ-28, 710 $p< 0.001$.

The strength of the association reflected by the Chi-square statistic, however, represents the extent of agreement between the screens regarding low scorers, rather than their agreement regarding high scorers.

Examination of respondents who were high scorers on both screens in terms of gender, age, marital and employment status did not indicate any distinguishing demographic profile. However, there was a significant association between high scorers on all screens, and reported physical health status. The Chi-square statistic for those who reported current illness or illness within the past month was significant (7.75, $p< 0.0054$), suggesting that respondents who reported being physically ill were most likely to be rated positive on all screens. The association between high scorers on both screens and reported physical illness is consistent with the finding that the PSE, PAS and "suggested clinical diagnosis" were more likely to agree on psychiatric caseness when assessing respondents who reported physical ill health. However, as those cases on which there was disagreement between the screens were not assessed on this variable, this finding is not conclusive.

Diagnostic categories of mental disorder most commonly identified in respondents rated potential cases by the SRQ GHQ-28 and CGHQ-28, and confirmed as true cases by the PSE-ID5, n=17

Seventeen high scorers on the SRQ, GHQ-28 and CGHQ-28 were later confirmed as cases by the PSE-ID5. 50% of screen positives were given the PSE-Catego tentative diagnoses of major depression or depressive related disorder. 22% were given a tentative diagnosis of anxiety related disorder. Results may suggest that the SRQ, GHQ-28 and CGHQ-28 are more likely to agree with the PSE-Catego in terms of caseness when assessing respondents with depressive or anxiety related disorders. However, no conclusive statement regarding this association can be made due to the small sample size.

Examination of these 17 cases found no distinguishing demographic profile. However, there was a strong association (Chi-square 13, $p< 0.001$) between reported physical health status and agreement on caseness by the above instruments. This may suggest that persons reporting current physical ill health, or illness in the past month, may be more likely to be identified as potentially positive on the screens, and to be confirmed as true positive by the PSE-ID5.

CONCLUSION

Results are assessed in the following chapter in light of both the strengths and weaknesses inherent in the research methodology and in the first- and second-stage instruments selected.

12 DISCUSSION

It has been argued that prevalence estimates are dependent on the validity of the two-stage case identifying method. This being so, it is appropriate that the validity of the SRQ and PSE in the context of the present research is discussed before the estimated prevalence of mental disorder in the adult population of Mamre is assessed.

METHODOLOGICAL ISSUES

THE VALIDITY OF THE SRQ AS A FIRST-STAGE SCREEN IN THE CONTEXT OF THE PRESENT RESEARCH (Primary objective (v))

Criterion validity of the SRQ

The unweighted sensitivity and specificity of the SRQ were estimated at 0.71 and 0.65 respectively. When these coefficients were weighted to reflect their values in the original sample, however, sensitivity dropped to 0.49, whereas specificity increased to 0.82. Although the weighted specificity was comparable with the top end of the range of reported specificity in other similar studies (0.78 to 0.81), the weighted sensitivity was considerably lower (0.78 to 0.80) (Dhadphale et al., 1983; Harding et al., 1980; Kortman & Ten Horn, 1988; Marie & Williams, 1985; Sen et al., 1987). Differences between the weighted and unweighted validity coefficients reflect the SRQ's high false positive rate (23%) and the low false negative rate (10%) in this sample. However, comparison of the weighted sensitivity with sensitivity reported in the above studies may be inappropriate. The sensitivity and specificity quoted in these studies were obtained in the measurement of clinic populations. Robins (1985) observes that sensitivity, in particular, may be lower in community as compared to clinic studies. As sensitivity and specificity are dependant upon the ratio of screen high to screen low scores in the sample (Benjamin et al., 1982; Chan & Chan, 1983; Tarnopolsky et al., 1979), it follows that when there is an equal ratio of screen high to low scorers, the chances of obtaining satisfactory sensitivity are favourable (Tarnopolsky et al., 1979). Samples drawn from the community are more likely to have a lower proportion of potential cases than non-cases. It may therefore be expected that the ratio of screen positives to screen negatives will be disproportionate. According to Mann et al., (1983) sensitivity may drop by as much as 24% if the proportion of high scorers in the sample drops from 50% to 22%. In the present study, unweighted sensitivity was estimated on a second-stage sample of 121 respondents, in which there was an almost equal proportion of high (45%) to low (53%) scorers. Weighted sensitivity, however, was adjusted to the original first-stage sample

of 481 respondents, in which the proportion of SRQ positives (26%) to negatives (74%) was far from equivalent. This disproportionate representation of high to low scorers in the original sample may have contributed to the considerable drop in sensitivity once it was weighted back. It has also been argued that as long as the proportion of high and low scorers in the original sample approximates the proportion of **true** positives and **true** negatives, validity coefficients should not be adversely affected (Chan & Chan, 1983). Use of power analysis to estimate optimal first- and second-stage sample size meant that the proportion of high and low scorers in the original sample did approximate the proportion of true positives and true negatives in the second-stage sample. It would therefore seem that disproportionate representation may not be an appropriate explanation for the SRQ's low sensitivity in this case. Nevertheless, in the face of evidence of lower sensitivity than found in other community studies using similar screening instruments (Benjamin et al., 1982; Chan, 1985; Cheng, 1985; Hobbs et al., 1984; Kortman, 1990; Pan & Goldberg, 1990), the possibility that disproportionate representation of high and low scorers between the original and second-stage samples may have contributed to the SRQ's lower sensitivity should not be discounted.

There is also evidence to suggest that not all reported validity coefficients have been weighted to reflect the original sample, and may therefore be artificially high (Marie and Williams, 1985). In Harding et al.,'s (1980) study for example, while the PSE was applied to a highly skewed sub-sample of respondents (90% were high SRQ scorers and only 10% were low scorers), questions arise as to whether sensitivity was weighted back to the original sample (Marie & Williams, 1985). Alteration in the weighted sensitivity in the present study demonstrates that the weighting back process may considerably influence sensitivity, as it takes into account not only the PPV and NPV of the test, but also the proportion of screen high and low scorers in the original sample. Hobbs et al., (1984),and his colleagues reported an unweighted sensitivity and specificity of 0.72 and 0.87 respectively, but a weighted sensitivity and specificity of 0.55 and 0.72.

These arguments offer a plausible explanation for the lower sensitivity found in the present study. However, there are other equally sound possible explanations for the SRQ's low sensitivity, which highlight the shortcomings of the present research. These are discussed below.

(i) Questions relating to the definition of psychiatric caseness

In chapters 1 and 2 it has been argued that the task of defining mental disorder, and particularly mental disorder commonly encountered at community level, is problematic (Goldberg & Huxley, 1992), given that concepts of psychiatric caseness remain grounded in concepts derived from hospitalized patients (Finlay-Jones & Murphy, 1979). Furthermore, the distinction between normality and psychiatric morbidity involves imposing an arbitrary dichotomy

on a continuum of severity. Although the threshold is estimated in terms of an informed understanding of the context, methodology and instrument application of the research, it nevertheless reflects assumptions regarding what constitutes mental disorder. In the absence of a comprehensive and universally accepted definition of psychiatric caseness, it is reasonable to suggest that the poor validity coefficients found in this study are, in part, a product of the inadequacy of the currently held definition of psychiatric caseness, particularly when this is applied to the level of disorder found within communities.

(ii) Questions relating to the appropriateness of the PSE as a criterion measure in a two-stage research design, in the context of a community study

This issue is explored in greater detail later in this chapter. Suffice to say, the validation of the screen is completely dependent on the psychiatric assessment, which is assumed to be correct (Surtees, et al., 1986). The PSE was selected as a valid second-stage criterion measure of psychiatric morbidity, in spite of evidence suggesting that it might be inappropriate as a gold standard when used to validate a screen in a community survey (Finlay-Jones & Murphy, 1979). (see Chapter 6, p. 56). As the PSE was developed from studies of psychiatric patients, it tends to rate only symptoms that are "clinically severe", and to miss those which are less stable and less well defined (Sen et al., 1987). Use of the PSE with screening instruments such as the SRQ is therefore likely to produce comparatively low validity coefficients (Sen et al., 1987).

Six cases identified by the PAS-DSM111R, and two cases identified by the "suggested clinical diagnosis" were not identified by the PSE-Catego. Examination of these cases showed that all six PAS-DSM111R cases were given a diagnosis of agoraphobia, while the two rated by the "suggested clinical diagnosis" were given diagnoses of substance abuse and alcohol abuse respectively. As the PSE-Catego cannot detect the above disorders, it is not surprising that it was unable to identify these cases. However, it is interesting to note that none of the cases of agoraphobia identified by the PAS-DSM111R were identified by the "suggested clinical diagnosis". This may suggest that the agoraphobia algorithm may be too lenient. Furthermore, all cases given a diagnosis of depression or anxiety related disorder by the PAS-DSM111R and "suggested clinical diagnosis" were also rated positive by the PSE-Catego. Together, these factors suggest that in this study at least, the PSE was no less sensitive to disorder found in the community than either of the other two criterion measures. Therefore, there appears sufficient reason to conclude that

the SRQ's low sensitivity was not necessarily due to the insensitivity of the PSE to the type of disorder commonly encountered in the community.

(iii) Questions relating to the adequacy of the PSE inter-rater reliability study, conducted in the first pilot

Poor sensitivity may also have resulted from an inadequate PSE inter-rater reliability study conducted in the first pilot. As has been noted, the inaccurate rating of symptom 58 by one of the second-stage interviewers resulted in the mis-assignment of psychiatric caseness in the case of one respondent. The PSE is a semi-structured interview, which means that the interviewer is required to judge whether or not reported symptoms are present, and how far to go with additional probes in order to elicit sufficient information for rating purposes. Rigorous training on the instrument is necessary to provide sufficient knowledge to enable the interviewer to interpret the data in a clinically meaningful way (Rodgers & Mann, 1986). Although training on the PSE was provided on two occasions, by persons with considerable experience in its administration, it may nevertheless have been inadequate in terms of both length of training period, and supervised practise interviewing. Furthermore, the inter-rater reliability study conducted in the first pilot, was simplistic to the extent that only one respondent was rated by all PSE clinicians. While clinicians were blind to the respondent's score on the SRQ, the fact that there was only one respondent may well have led them to conclude that he/she was a high scorer on the screen. This in turn, may have had a biasing influence on their ratings on the PSE. At the time the study was undertaken, financial constraints did not permit a more extensive inter-rater reliability assessment. However, lack of good inter-rater reliability on the PSE may have introduced systematic bias in the rating of individual, symptom abnormalities, and hence classification of case/non-case. Poor inter-rater reliability may therefore also have impacted negatively on the sensitivity of the SRQ. The limitations of the inter-rater reliability study remain a severe shortcoming in the present research.

(iv) Questions relating to the ability of the SRQ to discriminate between psychiatrically ill and non-ill people in a community study

If one concludes that the PSE was not insensitive to the type of mental disorder found in the adult population of Mamre, an alternative explanation for the SRQ's low sensitivity in this study, might be the large number of SRQ false positives. The high rate suggests that the SRQ itself might be inefficient in discriminating between psychiatrically ill and non-ill people in community

samples. Kortman (1990) comments that the validity of the SRQ may be attenuated by its sensitivity to help-seeking behaviour among clinic attenders, even in the absence of any mental illness. Although this observation referred to persons attending phc services, it may also serve as a reasonable explanation for the high SRQ false positive rate found in the Mamre sample. There was considerable social and economic stress in Mamre at the time the study was conducted (Katzenellenbogen et al., 1988), yet the community did not have easy access to the full range of phc services. It is therefore feasible that respondents experiencing stress might have been inclined to answer SRQ questions positively, even in the absence of mental distress, if they anticipated that positive responses might lead to health care benefits.

(v) Questions relating to the appropriateness of the translated version of the SRQ

The importance of the translation process to the validity of the SRQ in the present context, was recognised from the outset. Every effort was therefore made to ensure that the "intent" behind items was communicated in translation. However, the possibility remains that the SRQ might have lost its discriminatory power in translation due to cultural differences. Unfortunately, sufficient data relevant to the exploration of this possibility was not obtained. The high positive response to items such as "Are you unable to play a useful part in life ?" (93%) and "Are you a much more important person than most people think?" (62%) may reflect community rather than individual thinking. In a similar study in Ethiopia, Kortman (1990) commented that concepts used in the SRQ might be too western to be transposed unchanged to non-western cultures, and might therefore require fairly extensive modification in these contexts. Fabrega (1987) has also suggested that the SRQ might be less universally applicable in different cultures than was supposed by the WHO. It may be that the discrepancy between the original and translated versions of the SRQ may have constituted another source of information bias, which in turn may have contributed to the low sensitivity found in this study.

(vi) Questions relating to the SRQ's inability to rate as positive persons who are true positives on the PSE, but who report being physically healthy

Examination of true positives on the PSE-Catego indicated that persons reporting **physical ill health**, both currently or in the past month, were likely to be correctly rated as positive on the SRQ. This finding is consistent with Kleinman's observation (1988) that, for some cultures, the physical presentation of symptoms (somatization) may be a meaningful way of expressing psychological or emotional distress. The association between

positive SRQ scores and physical health status may therefore indicate that the SRQ is sensitive to underlying psychological disorder, even when only somatic symptoms are presented. In the present study, true positives who reported **current physical healthy**, on the other hand, were less likely to be given a high score on the SRQ. It is possible that self perceptions of physical health may mask underlying emotional distress, resulting in the SRQ's inability to correctly rate mental disorder in those who reported current physical health in this population. It seems therefore, that this may have contributed to lower sensitivity in the SRQ.

EXTENT TO WHICH THE SRQ MEASURES THE SEVERITY OF PSYCHIATRIC DISTURBANCE

The measure of association between scores on the SRQ and the PSE is a means of assessing how well the SRQ measures the severity of psychiatric disturbance, or whether the same cases are identified by the two instruments to the same extent. In the present study, the SRQ identified 67% of the PSE-Catego cases and non-cases, yielding a Chi-square coefficient of 13, $p< 0.001$. The association between the two instruments was attenuated, however, by the high number of SRQ false positives. Possible associations between commonly researched variables and false positive rate have been suggested by researchers in similar studies (Finlay-Jones & Murphy, 1979; Lobo et al., 1986; Mann et al., 1983; Skuse & Williams, 1984) and these are considered below in the context of the present research.

Association between SRQ false positives and gender, social class and physical health status

Finlay-Jones and Murphy (1979) and Lobo et al., (1986) observe that false positives tend to be either respondents with a physical illness and/or women from working-class backgrounds, who are overburdened with heavy domestic duties and responsibilities, and who therefore tend to respond positively in the absence of either symptoms or other sources of distress. Examination of SRQ false positives in the present study, however, indicated that false positives did not differ from true negatives in terms of gender, socio-economic or reported physical health status. The absence of any association suggests that the false positive rate may be due to other factors.

False positive rate and spontaneous remission of mental disorder in the interval between the administration of the screen and the second-stage clinical interview

The level of mental illness found in communities has been described as typically unstable, ill-defined, and frequently of a transient nature (Goldberg & Huxley, 1992; Lobo et al., 1986). Distress may spontaneously remit during the interval between the administration of the first and second-stage interviews, or may remit as a result of the administration of the screen itself. This possibility was recognised in the present research and the second-stage interview was consequently administered within nine days of the screen. The effect of remission may be assessed by re-administering the screen at the second-stage interview. Unfortunately, this was not done. No conclusive statement can therefore be made regarding possible remission rates or effect of remission on the validity coefficients of the SRQ. It is recommended that future studies, using similar instruments in a two-stage research design, should take measures to assess possible spontaneous remission.

The SRQ's tendency to over-identify respondents as positive in a community sample

If the PSE was a valid measurement of psychiatric caseness in the present study, then it must be concluded that the SRQ tended to over identify positives. This bias has also been found in other similar studies in which the SRQ and the GHQ have been used (Finlay-Jones & Murphy, 1979; Kortman, 1990; Skuse & Williams, 1984). According to Kortman (1990) the over identification of screen positives may be one of the inherent limitations of self-report measures in community studies.

Conclusions regarding high rate of SRQ false positives in the present research

Notwithstanding these limitations, the high rate of false positives has to be considered in the light of the aim to establish a base-line estimate of the prevalence of mental disorder in the adult population of Mamre. In terms of this aim, it was more important that the screen did not miss potential cases. A high false positive rate was therefore preferable to a high false negative rate, given that the latter might have led to an underestimate of the prevalence of disorder, and to the irretrievable loss of true positives. On the other hand, the high false positive rate means that prevalence estimates found in this population must be treated with caution, and that the comparison of estimated prevalence rate of disorder in Mamre with other populations can only be tentative.

COMPARISON OF THE VALIDITY COEFFICIENTS OF THE SRQ WITH THOSE OF THE GHQ-28 AND CGHQ-28 (Secondary objective (i))

The following discussion evaluate the appropriateness of the decision to use the SRQ as the primary first-stage screen in the main study. Efficiency is evaluated in terms of the validity coefficients for both SRQ, GHQ-28, and CGHQ-28 when used with the three different criterion measures.

The weighted validity coefficients for the SRQ and the CGHQ-28 were equivalent when the PSE was used as the criterion measure. However, when the PAS or "suggested clinical diagnosis" were the criterion measures, the CGHQ-28 had superior validity coefficients, PPV and OMR, as well as a lower false positive rate, than either the SRQ or the GHQ-28. Furthermore, when the CGHQ-28 was used with "suggested clinical diagnosis", its weighted validity coefficients, unlike that of the SRQ, were within the range of both weighted and unweighted sensitivity and specificity found in other community based studies (Benjamin et al., 1982; Chan, 1985; Cheng, 1985; Hobbs et al., 1985; Kortman, 1990; Pan & Goldberg, 1990). (The disadvantages of using an unstructured and non-standardised measure such as the "suggested clinical diagnosis" as a criterion measure are discussed in Chapter 2, p 15). The results suggest that the CGHQ-28 may yield better validity coefficients, and may therefore, be a more appropriate screening instrument than the SRQ in a two-stage design in the context of a community study. Its superior validity coefficients may be understood in terms of its four point scoring system and its method of scoring all negatively worded items as indicative of problems, which allows it to be sensitive to respondents with chronic illness (see Chapter 6, p. 45).

CONCLUSIONS REGARDING THE STRENGTHS AND WEAKNESS OF THE SRQ AS A FIRST-STAGE SCREEN IN A COMMUNITY STUDY

The above discussion identifies both weakness inherent in the methodology of the research, and weakness inherent in the SRQ as a first-stage screen for psychiatric morbidity in a community survey.

Weaknesses in research methodology which might have negatively affected the validity coefficients of the SRQ

(i) Inadequacies in the PSE inter-rater reliability study: validity of the SRQ was dependent not only on the validity, but also on the reliability of the PSE. Given that the five clinical interviewers may not have been rating respondents in the same way, the SRQ's validity may have been adversely affected. (see Appendix 4.1; Erroneous rating of symptom 58 on the PSE).

(ii) Inadequacies in the translation process, which may not have taken sufficient cognizance of possible cultural differences in the Mamre community: these shortcomings may have diminished the SRQ's ability to discriminate effectively between disordered and mentally healthy persons.

(iii) Disproportionate representation of SRQ low to high scorers in the original sample, compared to that in the second-stage sample may also have been a factor in the SRQ's lower weighted sensitivity.

Of these the inadequacy of the PSE inter-rater reliability study is by far the most serious shortcoming.

Weakness inherent in the SRQ which may have accounted for its low sensitivity in the present study

(i) Questions regarding the SRQ's ability to distinguish psychiatrically ill from psychiatrically non-ill respondents in a community study: most validity studies on the SRQ have been conducted with clinic populations. There is little evidence to suggest that the SRQ is equally effective as a screening instrument in community samples. Some researchers have questioned whether respondents in community conditions will behave in an identical manner, regardless of how they are identified, or of the circumstances in which the questionnaire is presented (Benjamin et al., 1982);

(ii) Questions regarding bias resulting from the SRQ's response format which permits both "yes" and "no" sayers: this means that respondents with a tendency to give "yes" responses to all or most questions, are likely to give rise to a number of false positives;

(iii) Questions regarding the SRQ's ability to identify the chronically distressed person;

(iv) Questions regarding the SRQ's ability to identify those with mental disorder, who perceive themselves as being physically well; and

(v) Questions regarding the SRQ's sensitivity to help-seeking behaviour among community respondents who are burdened with heavy socio-economic responsibilities, even in the absence of any mental illness.

CONCLUSIONS REGARDING THE USE OF THE SRQ IN THE PRESENT RESEARCH

Before a decision can be made regarding the appropriateness of the SRQ as a first-stage screen for psychiatric morbidity in a community study in South Africa, its weaknesses and strengths have to be considered in terms of the goals of the research.

In terms of estimating base-line prevalence of mental disorder in the Mamre community, and estimating the validity of the SRQ as a first-stage screen when used with the PSE in a South African community sample, there is evidence that the SRQ may be moderately useful for the following reasons:

(i) Its low false negative rate meant that few true positives were missed. This is a crucial aspect of a first-stage screen, when attempting to estimate prevalence in a population where little is known about the mental health status of the community;

(ii) The SRQ was able to correctly identify 67% of respondents as either true positives or true negatives when the PSE was used as the gold standard;

(iii) While the SRQ over identified false positives, it only correctly identified 42% of true positives. Although weighted sensitivity was far from optimal, it compares with normal clinic practice in developing and more developed countries. These studies have shown that between 50% and 75% of psychiatric illness remains undetected, or is confused with somatic disorders (Goldberg & Huxley, 1980).

Although not specifically stated as a goal of the present research, one further concern regarding the usefulness of the SRQ needs to be considered. It was hoped that in establishing the validity of the SRQ, it would be possible to provide phc workers (who have limited psychiatric training) with an efficient instrument that would help them discriminate between psychiatrically ill and psychiatrically non-ill persons. Although the number of false negatives (10% of the second-stage sample, or 29.3% of all true negatives) suggests that the SRQ might be moderately useful in terms of the above objective, its low PPV (0.50) and relatively high OMR (0.27) argue against its usefulness in predicting the probability that someone with a high score will be found to have the disorder at a subsequent clinical examination.

The estimation of prevalence of mental disorder in the adult population, and the assessment of the effect on prevalence estimates when three different systems of case identification were employed, are examined in the following section in light of the SRQ's moderate validity.

ESTIMATED PREVALENCE OF MENTAL DISORDER IN THE ADULT POPULATION OF MAMRE, (Primary objective (i))

The weighted prevalence of mental disorder (27%) in Mamre confirms the view that mental disorder at community level may not only be high at community level (Goldberg & Huxley, 1992), but may even be higher than that found in the communities of developed countries (Ben-Tovim, 1987; Harding et al., 1980; Reeler, 1987a).

Although the estimate calculated in this study is high, it is in keeping with the high rate (45%) found in the methodologically less sophisticated Mamre clinic study conducted in the same population eight months previous to the beginning of the present research (Miller et al., 1991). It is also consistent with the upper range found by Gillis et al., 1968 (23%), and with that found by Ben-Arie et al., (1983) (23.7%) in a community sample of "coloured" elderly in Cape Town, although there is literature to suggest that rates may be higher in the elderly (Gillis et al, 1991). While there is evidence that the first filter between the community and the phc services may not be as permeable as first reported (Goldberg & Huxley, 1980), the apparently high rates of disorder in both community and clinic samples in Mamre appears to indicate that in Mamre at least, the first filter may be fairly permeable.

Comparison with the prevalence rates found in similar studies in non-western or developing communities indicates that the estimated rate (27%) in this community exceeds the overall rate found in the majority of these studies (20.1-22%) by 5% (Cheng, 1988; Hollifield et al., 1990; Orley & Wing, 1979; Thom et al., 1993; Vasquez-Barquero et al., 1981, 1987). It is therefore important to examine the validity of the estimated prevalence in Mamre in the light of possible variables which may have contributed to its inflation.

Questions regarding the validity of the SRQ as a first-stage screen

It was concluded earlier in the chapter that the SRQ was moderately useful as a screen in a two-stage research design. However, as weighted prevalence is estimated according to a statistical formula which takes the number of screen false positives into account, the presence of a high SRQ false positive rate in the present study is likely to have spuriously inflated the estimated prevalence (Wing et al., 1977). No definitive statement regarding prevalence of mental disorder in this population can therefore be made.

Questions regarding the inadequacy of the PSE inter-rater reliability study

The inadequacy of the inter-rater reliability study has already been discussed. Its limitations are an important aspect of the research. Any question of inter-rater unreliability, puts in doubt conclusions regarding estimated prevalence. Furthermore, the fact that one of the raters was mistakenly rating symptom 58, not only highlights the paucity of the PSE training

programme in this research study, but also questions conclusions regarding the estimated prevalence of cases of severe mental disorder in the community.

Questions regarding the validity of the PSE as a second-stage criterion in the context of community research in a developing country

(i) PSE identification of known psychiatric patients in Mamre

It was ascertained that none of the psychiatric cases in Mamre known to the psychiatric nurse in 1991, happened to be selected for the second-stage clinical interview. The random sampling technique employed to select respondents for the second-stage interview, however, gives an equal chance to all respondents to be selected. It cannot be assumed, therefore, that the rate of PSE cases would have been different had known cases been selected through random sampling.

(ii) Questions regarding the PSE's tendency to under-diagnose psychiatric disorder in a community study

It has been argued that the PSE's criteria for caseness may be too stringent when applied to a community study, where the level of mental disorder may be difficult to defined (Ormel et al., 1990; Vasquez-Barquero et al., 1986). According to Goldberg and Huxley (1992), if second-stage criteria are too stringent, it is likely that many borderline cases will be missed, and prevalence consequently lowered. This also means that a high rate of what would be incorrectly termed as false positives would be recorded on the screen. With the exception of three cases, all those who attained borderline scores on the SRQ (45%), but who were not identified as cases by the PSE-Catego, were not rated as positive by either the PAS-DSM111R or "suggested clinical diagnosis". This suggests that the PSE is no more likely to miss borderline cases than the other two criterion measures. It can therefore also be argued that the PSE ID-5 may be an appropriate threshold for psychiatric caseness in the context of this research study.

Of the three criterion measures, the PSE yielded the highest prevalence estimate. This, plus the fact that the PSE did not appear to be missing borderline cases, suggests that the PSE was not underrating psychiatric disorder. On the contrary, it may very well have been over estimating disorder.

(iii) Questions regarding the PSE's tendency to over-diagnose psychiatric disorder in a community sample

When the PAS-DSM111R and "suggested clinical diagnosis" were used as criterion measures, the estimated prevalence of mental disorder was 24% and 17% respectively. These rates fall within the range of prevalence estimated in other community studies in developing countries (Bland, 1988, in Parry, 1992; Cheng, 1988; Orley & Wing, 1979; Vazquez-Barquero et al., 1987). An examination of those cases identified by the PSE-Catego alone was therefore considered appropriate, and may be found in Table 12.1.

Table 12.1 Examination by independent clinical psychologist of seven cases identified as positive by the PSE-Catego but not identified by the PAS-DSM111R or "suggested clinical diagnosis"

Case number	Explanation of results on PSE-Catego	Probable diagnosis
225	Symptom 58 incorrectly rated positive, resulting in incorrect identification of respondent as a psychiatric case. PSE-Catego tentative diagnosis: Paranoid schizophrenia	Probable non-case
36	External psychologist and interviewing psychologist concurred that anxiety symptoms occurred in response to a specific life-experience. PSE-Catego tentative diagnosis: Neurosis	Probable non-case
300	Symptoms of delusion explainable in terms of culturally acceptable experience in this population. PSE-Catego tentative diagnosis: Paranoid schizophrenia	Probable non-case
253	Sufficient symptoms for positive diagnosis. PSE-Catego tentative diagnosis: Major depression	Major depression
262	Sufficient symptoms for positive diagnosis. PSE-Catego tentative diagnosis: Major depression	Major depression
92	Sufficient symptoms for positive diagnosis. PSE-Catego tentative diagnosis: Other paranoid states.	Dysthymia
358	Sufficient symptoms for positive diagnosis. PSE-Catego tentative diagnosis: Major depression or Dysthymia	Dysthymia confirmed

Inspection of these cases was carried out by a clinician with considerable experience with the PSE, but who had not been involved in its administration in the present study. Table 12.1 shows that one of the cases, Case 225, was mis-assigned psychiatric caseness as a result of the inaccurate rating of symptom 58 by one of the second-stage interviewers (see Appendix 4.1). The remaining six cases were examined in terms of their SRQ score and

specific demographic variables. No single variable was identified which might have accounted for their diagnoses on the PSE-Catego. It was concluded that with the exception of two cases, there was evidence of sufficient symptomatology for the other four to have been correctly diagnosed as positive by the PSE-Catego (Appendix 4.3, cases 36 and 300). The two cases considered to be non-cases by the external psychologist and interviewing clinician were assigned a PSE-Catego tentative diagnosis of paranoid schizophrenia, on the basis of symptoms which should have been more appropriately understood either in terms of culturally acceptable experiences in this population, or as a response to specific life experiences.

It is interesting to note that if the three probable non-cases were excluded from the prevalence rate, the weighted prevalence of mental disorder would fall to 24%, which is comparable with the weighted prevalence as estimated by the PAS-DSM111R.

In the two probable non-cases, the PSE-ID was 5. Examination of cases which were identified as positive by all three criterion measures showed that 17 out of the 20 cases (90%) had an ID of 6 or above. It may therefore be concluded that in this population, an ID of 5 may have contributed to the number of cases identified by the PSE, but not by the PAS or "suggested clinical diagnosis".

However, as only two cases diagnosed as positive by the PSE-Catego were considered probable non-cases, it may be argued that the PAS-DSM111R and "suggested clinical diagnosis" were under-rating the extent of mental disorder in this population. Although there is some question regarding the appropriateness of some of the PSE-Catego tentative diagnostic categories (as will become clearer later in the chapter), of the six cases identified by the PSE-Catego alone, it was the considered opinion of the external psychologist that 4 were **definite** cases. It may be that the PAS-DSM111R's inability to identify these 4 cases was due to its more restricted range of symptoms than that of the PSE-Catego. Regarding the inability of the "suggested clinical diagnosis" to identify these four cases, it must be recognised that this is not a structured system, and therefore there is no guarantee that it was either valid or reliable in this study. There was no attempt to establish inter-rater reliability among the raters on this criterion, and the criterion functioned more as a local check on the PSE, as suggested by Ben-Arie et al. (1983).

Of the five second-stage clinical interviewers, the psychiatric nurse had considerable experience in the administration of psychiatric interviews. Experienced clinicians tend to rate more conservatively than those with less

experience (Sen et al., 1987). In spite of finding that the psychiatric nurse was responsible for the incorrectly rated symptom 58, the above argument still seems to apply in this context. In two of the cases rated positive by the PSE-Catego, and confirmed as positive by the independent clinician, a "suggested clinical diagnosis" of asymptomatic was given by the psychiatric nurse. The possibility therefore exists that like the PAS-DSM111R, "suggested clinical diagnosis" may also be missing cases.

While the above discussion suggests that the PSE-Catego is neither under- nor over-diagnosing mental disorder in the present community study, there is insufficient evidence to conclude that the PAS-DSM111R and "suggested clinical diagnosis" are under-diagnosing mental disorder in the sample.

(iv) Questions regarding the PSE's tendency to mis-diagnose paranoid schizophrenia and other paranoid psychoses in this community study

The high rate of both paranoid schizophrenia (5%) and other paranoid states (3%), as diagnosed by the PSE-Catego (but not by the two other criterion measures), while consistent with the rate of psychosis found in a community study conducted in Guinea Bissau (De Jong et al., 1986) is considerably higher than that found in other community studies. The need for further investigation of the validity of these diagnoses is therefore identified. (Examination of the 11 cases assigned these tentative diagnoses is given in Appendix 4.3).

Disagreement regarding the classification of cases into specific diagnostic categories by different diagnostic systems has been well documented (Dean et al., 1983; Van den Brink et al., 1989). The argument that the PSE-Catego may be mis-diagnosing certain cases is not directed at this discrepancy. What is examined, is the possibility that paranoid schizophrenia, or other paranoid states, are being unjustifiably suggested as an appropriate diagnosis in this population, as a result of the PSE-Catego's inability to modify positive ratings on specific symptoms. This applies specifically to symptoms regarding cultural appropriateness and organic impairment. The PSE-Catego is unable to take into account symptoms of organic impairment or symptoms that can be understood in terms of subcultural experiences. In these instances, the PSE-Catego erroneously assigns a tentative diagnosis of either paranoid schizophrenia, or other paranoid states. (Fuller discussion of this serious shortcoming is found in Appendix 4.2). Evidence that the PSE-Catego may be over-diagnosing paranoid and psychotic disorders in the community, although convincing, can be considered only tentative. Methods used to

reassess these cases were exploratory, and were therefore unable to yield conclusive results.

(v) Questions regarding the PSE's tendency to under-diagnose non-psychotic disorder in the community

It has been argued that the PAS may be a more appropriate psychiatric instrument to use as a second-stage criterion in a community study than the PSE because of the former's lower threshold, particularly for non-psychotic disorders (Van den Brink et al., 1989). (see Chapter 8, p. 74). It has also been suggested that the PSE may be insensitive to these disorders, especially social phobias, and obsessive and compulsive disorders (Ormel et al., 1990), while having a strong bias towards depressive and anxiety related disorders (Van den Brink et al., 1989). Of the 11 cases of agoraphobia diagnosed by the PAS-DSM111R, 2 were given a PSE-Catego tentative diagnosis of psychotic disorder, and 3 were diagnosed as depressive or anxiety related disorders. The remaining 6 cases of agoraphobia, and the one case of PAS-DSM111R social phobia, were not rated as cases by either the PSE-Catego, or the "suggested clinical diagnosis". While the PSE-Catego is unable to diagnose agoraphobia, the strong possibility exists that the PAS-DSM111R itself may be over-diagnosing agoraphobia. On the other hand, the prevalence of these disorders in this community may not be as great as expected.

(vi) The PSE-Catego's inability to identify either alcohol or substance abuse

Gillis et al., (1968) identified alcohol abuse as a serious problem among "coloured" persons in South Africa. This was and also reported by community residents to be one of the most common problems in Mamre (Katzenellenbogen et al., 1988). Although only 21% of respondents reported abusing alcohol during the weekends in the first-stage questionnaire, it may be anticipated that alcohol abuse might be a prominent feature of cases of mental disorder in Mamre. As the PSE-Catego is not designed to rate alcohol abuse, this could not be verified in this study. However, one of the cases initially given a tentative diagnosis of paranoid schizophrenia, was assessed as dementia by both the interviewing psychologist and the external psychologist, on the basis that the symptoms of hallucination and delusion could best be understood in terms of the subject's organic impairment and abuse of alcohol (see Case 62, Appendix 4.3). Of the 27 cases identified by the "suggested clinical diagnosis", only two (7%) were diagnosed as abusing alcohol. Neither of these cases was rated positive by the PSE-Catego. How and

whether the prevalence rate would have been affected if cases of alcohol abuse were identified, remains a matter of speculation, but could be elevated.

Interestingly enough, the psychiatric nurse commented on the apparent lack of respondents with symptoms of alcohol abuse. It has already been suggested that the sample may not have been fully representative of the total population in this respect, as one section of the population, thought by the community to abuse alcohol, may not been fully represented in the first-stage sample (Salmon, 1993). Although alcohol was given as the reason for non-participation in the study in only 10 instances,, it may have been one of the undisclosed reasons for non-participation.

The PSE-Catego is similarly unable to rate drug abuse. Only 1 case was identified as a substance abuser by the rating psychologist, which is contrary to the findings of Katzenellenbogen et al.,'s. 1988 study, in which drug abuse was perceived by the community as an important health problem.. This may suggest the sample may not have been representative of the community in this respect.

(vii) Questions relating to the PSE-Catego's bias towards depressive and anxiety related disorders

Most studies have confirmed that depression and anxiety make up the greatest proportion on mental disorder within communities (De Jong et al., 1986; Duncan-Jones & Henderson, 1978; Goldberg & Huxley, 1992; Harding et al., 1980; Hodiamont et al., 1987; Hoeper et al., 1979; Lobo et al., 1986; Orley & Wing, 1979; Sashidharan et al., 1988; Surtees et al., 1986). The extent of agreement between the three criterion measures regarding these disorders, suggests that the PSE-Catego weighted prevalence estimates (19%) for depression (13%) and anxiety (6%) in the adult population of Mamre, are appropriate. Similar rates were also found in the studies listed above.

CONCLUSIONS REGARDING THE VALIDITY OF THE PSE

Definitive statements regarding the PSE's validity as a second-stage criterion measure in the context of the community under study, are problematic, as methods used to assess the instrument's appropriateness were investigative. Nevertheless, areas for further enquiry were identified. These include:

(i) the PSE-Catego's tendency to mis-diagnose paranoid

(ii) its inability to distinguish between true delusions and delusions which are sub-cultural in nature;

(iii) its inability to rate organic impairment; and

(iv) its inability to rate substance abuse.

These limitations, along with the inaccurate rating of symptom 58 by one of the raters, mean that conclusions regarding the prevalence estimate of particular psychiatric diagnoses, (especially of the paranoid disorders) within the community, have to be treated with caution. In the original calculation, the weighted prevalence of severe mental disorders (including paranoid schizophrenia, other paranoid states, and psychotic depression) was estimated at 16.1%. If the cases regarded as non-cases by the external psychologist are excluded from the calculation, and the remaining cases are reassigned to the alternate suggested diagnostic classes, the prevalence estimate of severe mental disorders would fall to 10.1%. This is comparable with estimates of rates of severe mental disorder in other community studies (Goldberg & Huxley, 1992). Furthermore, the weighted prevalence of paranoid schizophrenia and other paranoid states would then fall to zero.

In spite of these limitations, there is insufficient evidence to suggest that the PSE is either under- or over-diagnosing psychiatric morbidity in the community. While there is strong evidence that it may be mis-diagnosing paranoid schizophrenia and other paranoid psychosis, there is no evidence to imply that it fails to identify social phobias or obsessive and compulsive disorders. Significant agreement between the three criterion measures regarding depressive and anxiety related disorders, confirms that the PSE may be effective in identifying the level of disorder most commonly encountered in communities (Goldberg & Huxley, 1992). Therefore, it can be concluded that the use of the PSE as the second-stage criterion measure in the present research, did not spuriously inflate the prevalence estimate.

SUBSTANTIVE ISSUES

SOCIO-POLITICAL FACTORS IN SOUTH AFRICA WHICH MAY HAVE CONTRIBUTED TO THE APPARENTLY HIGH PREVALENCE RATE OF MENTAL DISORDER IN MAMRE

Until recently, South African laws have enforced racial segregation (apartheid) among different ethnic groups. Experiences of discrimination and alienation among Afro-Caribbean immigrants in Britain have been found to be highly threatening and may play a part in precipitating mental illness, particularly psychotic disorder, in vulnerable individuals (Littlewood, 1990). In South Africa, the "coloured" population has experienced

discrimination and alienation, not only as a result of the apartheid system instituted by the ruling "white" minority, but also from the oppressed "black" majority, resulting from its racial position in "no man's land". Several researchers have found higher rates of psychiatric disorder in the "coloured" population when compared to other groups (Dick et al., 1978). This seems to lend support to the argument that adverse political and social systems may contribute to mental illness. However, the assumption of a linear association between oppressive political systems and mental disorder has been criticized by Swartz (1987) who argues that the identification of a simplistic linear relationship fails to take mechanisms that communities may use to deal with oppression into account, and may even lead, in areas of mental health, to arguments in defence of the apartheid system.

Nevertheless, there can be no dispute that the apartheid system has played a crucial part in the life experience of every South African. Although the effects of apartheid on mental health may not be discernible as a simple causal relationship, attempts to "ignore this factor in any theory of mental health and illness is likely to render the theory decontextualzied and unsatisfactory" (Swartz, 1987 p. 26). Given that the "coloured" population of Mamre has lived through years of apartheid, it is reasonable to interpret the apparently high prevalence rate of mental disorder found in the present study, within the context of this experience.

LACK OF PRIMARY HEALTH CARE FACILITIES IN THE COMMUNITY

Where few or no facilities exist, it is likely that mental disorder will go undetected (De Jong et al., 1986). Up until 1991, Mamre was visited by a psychiatric sister only once a month, and a psychiatrist only once every three months. Furthermore, those persons who were detected as psychiatric cases, usually of a psychotic nature, had to be referred to a psychiatric hospital 80 kms away. As public transport from Mamre to the city is both unreliable and infrequent, the opportunity for psychiatric care was tenuous. The situation was even more problematic for those suffering from non-psychotic disorders, where the facilities for detection or referral were virtually non-existent. As a result, many persons with mental disorder remained untreated in the community, and their presence may partly account for the high prevalence rate found in this study.

It may also be that the apparently high prevalence of disordered persons in the community is the result of either the inability or unwillingness of phc practitioners to acknowledge psychological distress. When practitioners are in the situation of having no resources or referral methods, they may not identify psychiatric disorder because they may believe that nothing can be done about it (Miller et al., 1991). In this situation, GPs tend to treat the physical symptoms where these are presented, instead.

LITTLE KNOWLEDGE OF BIOMEDICAL CONCEPTS OF MENTAL DISORDER AMONG COMMUNITY RESIDENTS

Similarly, absence of phc services may also mean that the community itself lacks knowledge about mental disorder, in terms of the dominant biomedical definition and interpretation. The way both the individual and the community interpret mental disorder depends on the understanding of deviant/unusual behaviour within that particular community. Where communities are not fully socialized into biomedical definitions of mental disorder, behaviour that is considered deviant by the community may be interpreted in terms of culture specific structures, such as spirit possession, witchcraft, and so forth. It has already been suggested that it is not uncommon for Mamre residents to explain certain life experiences in terms of bewitchment, "getoor" (Miller, 1989). Such behaviour is likely to be interpreted and dealt with in a way that increases the probability that it will not come to the attention of the phc services.

A pertinent observation by one of the senior members of the Mamre Community Health Project was that since the inception of various phc facilities in Mamre, community residents showed a tendency to report symptoms relative to the types of services provided. This would suggest that illness may not only be understood as something which develops first within the individual and which must thereafter find appropriate structures for its identification and treatment, but that the provision of services may themselves play a part in the development, interpretation and experiences of certain disorders (Swartz, 1993).

TOLERANCE OF MENTAL DISORDER IN THE COMMUNITY

It is also possible that in the absence of mental care facilities, the community has been forced to find means of dealing among themselves with members of the community who are mentally ill. This may have led to a tolerance for mental disorders, especially those which are neither violent nor bizarre. This in turn would allow patients to remain undetected and untreated in the community.

THE ROLE OF THE CHURCH IN MAMRE

90% of the community identifies itself as members of the Moravian church. Mamre was first established as a Moravian mission station in the 1808, and the church remains an important influence in the village and plays a central role in its everyday life (Katzenellenbogen et al., 1988). The role of the church in affecting the utilization of phc services and its impact on the management of mental disorders among the general population, is a matter for further investigation.

CONCLUDING COMMENTS ON PREVALENCE

CONCLUSIONS REGARDING THE PREVALENCE ESTIMATE OF MENTAL DISORDER AMONG THE ADULT POPULATION OF MAMRE

In summary, there appear to be several convincing factors which could have resulted in the high prevalence estimate of mental disorder. The most important of these include:

(i) the SRQ's high false positive rate;

(ii) the inadequate PSE inter-rater reliability study;

(iii) socio-political factors;

(iv) documented evidence of a higher rate of mental disorder among the "coloured" population compared to other population groups;

(v) non-exposure to the dominant biomedical definition and interpretation of mental illness;

(vi) lack of phc care facilities, either to identify these disorders, or to offer treatment;

(vii) tolerance within the community for deviant non-violent behaviour; and

(viii) a high level of psychopathology.

It is therefore concluded that while the estimate of disorder may be higher than that found in other similar studies, it was probably realistic within the context of the present research. Furthermore, the evidence of a high rate of both the more severe and less severe mental disorders in the community seems sound. Nevertheless, issues such as the moderate validity of the screen, and the inadequacy of the PSE inter-rater reliability study, mean that the estimate should not be regarded as conclusive.

ASSOCIATION BETWEEN PREVALENCE OF MENTAL DISORDER AND COMMONLY RESEARCHED SOCIO-DEMOGRAPHIC VARIABLES (Primary objective (ii))

Differences in prevalence rates were noted on the following variables: gender, marital status, employment status and alcohol use. However, due to the small numbers in the various cells, these associations failed to reach statistical significance. It was therefore not possible to

identify a distinguishable socio-demographic profile associated with the prevalence of mental disorder in the community.

ASSOCIATION BETWEEN PREVALENCE OF MENTAL DISORDER AND PHYSICAL HEALTH STATUS (Primary objective (iii))

As in other studies, a significant trend was found between prevalence of mental disorder and reported ill health, both current and within the past month (Weyerer 1990). The data necessary to reach conclusions on the direction of the relationship are not available, but this association confirms the view that many persons with mental disorder experience accompanying physical illnesses (Goldberg & Huxley, 1980). As attenders at phc clinics frequently present with somatic complaints in cases of psychological distress (Good & Good, 1982), it is important that phc physicians are trained to be sensitive to possible underlying emotional disorder (Goldberg & Huxley, 1980). Although the issue of phc practitioners' ability to detect mental disorder did not fall within the present research, it is interesting to note that in a study conducted by Miller et al., (1991) on clinic attenders in Mamre, only 44.8% of those scoring positively on the GHQ-28, were considered by their attending phc practitioner to have a psychological or stress-related disorder. This suggests that although there was a relationship between the practitioners' and the GHQ-28 assessments of psychological factors, the practitioners tended to underestimate the degree of psychological involvement, relative to the GHQ-28 (Miller et al., 1991). Similar results have also been obtained in a study in Kangwane by Freeman et al., (1991), in which less than one third of the mental health problems identified by the researchers were identified by the clinic staff.

DISORDER AND RATE OF CONSULTATION (Primary objective (iii))

In spite of a significant association between mental disorder and reported physical ill-health, no significant relationship was found between PSE high scorers reporting physical ill health and rate of consultation. The failure to show a significant relationship between physical health status, PSE positive score and rate of consultation, is contrary to the findings of several other studies (Finlay-Jones & Burville, 1978; Vazquez-Barquero et al., 1987). A relationship may not have emerged in the present study due to the possibility that respondents with mental disorder may consult at other times of the year when they are not in fact symptomatic (Goldberg & Huxley, 1992). Furthermore, the first-stage questionnaire was inadequate in the way phrasing of questions relating to this area of enquiry . Only those respondents who reported being currently ill or ill in the past month had to complete the section relating to phc consultation. The results, therefore, do not reflect the phc utilization of those respondents who may have been PSE high scorers, but who reported currently good health or health within the last month.

The finding that approximately three times as many persons with high PSE scores who reported illness within the **last month** as those who had PSE low scores and reported physical illness did consult, (although there was no significant relation between timing of illness and rate of consultation among PSE high scorers alone), suggests that the timing of illness in relation to rate of consultation may be important. The possibility suggested by this association, that mental health interventions in the community by lay- carers may selectively reduce inappropriate phc utilization, identifies this as an important area for future investigation.

Investigation of type of PSE-Catego tentative psychiatric diagnoses and health care utilization, indicated no obvious association between type of diagnosis and rate of consultation.

Further research is planned which will examine health care utilization among clinic and community samples in Mamre. It is envisaged that this enquiry will enhance our understanding of the illness experience, by providing insight into those factors which both increase the likelihood of consultation as well as those factors which appear to offer protection or support to those persons with symptoms of mental disorder but who do not consult.

UNDERTAKING TO DEVELOP THE COMMUNITIES' SKILLS

The training of lay-interviewers to administer the first-stage screen was in line with the Mamre Community Health Project's commitment to develop the skills of the community. However, fatigue and loss of motivation experienced by the lay-interviewers as a result of the prolonged study, and the distances that had to be covered, often in unseasonable weather without transport, has to be considered in the light of this goal. While neither the length of the study nor the long distances were found to have an effect on the quality of the data collected, and therefore did not impact on results, these are issues which need consideration in future studies of this nature.

CONCLUSION

Recommendations related to both methodology and substantive issues identified in the above discussion, and their relation to the detection and measurement of psychiatric morbidity, and to the future planning of relevant health care services, are examined in the following chapter.

13 RECOMMENDATIONS

Methodological shortcomings and inadequacies in the use of the SRQ and PSE in the present study do not permit a definitive estimate of prevalence of mental disorder in Mamre. Nevertheless, there is sufficient consensus among the various second-stage instruments, and between the results obtained in the present study and the previous clinic study in Mamre (Miller et al., 1990), to conclude that the rate of mental disorder in the adult population of Mamre is in the average to high range. Until further research is undertaken, however, there can be no way of knowing how many of the psychiatrically disordered will consult with their phc services, and will be identified as psychiatric cases should they do so. Since the inception of the Mamre Community Health Project in 1986, important steps have been taken to improve the availability of relevant primary mental health care services within the community. While these services have been useful, the possible high prevalence of mental disorder in the community indicates that there may be a sizable gap between its needs, and the provision of facilities which allow for detection, referral and treatment. Recommendations pertaining to these issues, as well as proposals regarding suggested improvements in general methodology are discussed below.

ISSUES RELATED TO THE DETECTION, REFERRAL AND TREATMENT OF MENTAL DISORDERS COMMONLY ENCOUNTERED IN THE ADULT POPULATION OF MAMRE

Detection

Participation of community health workers in the community

Given the suspected high prevalence of mental disorder at community level, there is reason to believe that community health workers currently working in Mamre, may be the personnel best placed to detect mental distress. In building reliable and close relationships with community leaders such as teachers and church officials, community workers are able to make contact with mentally disordered persons who might otherwise never come to the attention of the phc services. A trusting relationship between the community worker and the individual is also important for the successful treatment of those with non-psychotic disorders, who do not require hospitalization, but who require restitution of their functioning within the community.

Methods of detection by community health workers

Community workers need to be provided with competent screening techniques, in which they are thoroughly trained, if they are to be able to effectively detect mental disorder at community level. Immediate requirements therefore include:

(i) Provision of a modified screening instrument relevant to the Mamre context.

In spite of the SRQ's concordance (67%) with the PSE-ID5 in the present study, its low weighted sensitivity (0.49) and PPV (0.50) does not identify it as an effective tool for the identification of potential cases by phc workers in this community. While the validity coefficients of the CGHQ-28 were marginally better than those of the SRQ, particularly when the PAS was the criterion measure, there is no strong evidence to suggests that it is a preferable alternative to the SRQ in the present context. The design of a new screening instrument appropriate to the Mamre context is one option that needs to be seriously considered. On the other hand, the urgent need for accurate data to inform management planning, and the scarcity of available funding, suggests that consideration also be given to modifying existing screening instruments such as the SRQ and GHQ.

If screen modification is planned, and the aim is to restore functioning, and to classify only in so far as this contributes to better health care, it may be useful to make changes to the screen based on concepts employed in the dimensional model. It has been argued that this model may provide a more meaningful approach to the understanding and measurement of common mental disorders (Goldberg & Huxley, 1992). In view of the estimated high prevalence of both the severe and less severe mental disorders in the community, modifications should focus on:

(a) issues relevant to depression and anxiety, organic disorder, paranoid schizophrenia and alcohol abuse;

(b) issues relevant to subcultural delusions specific to the South African context;

(c) issues relevant to socio-demographic factors which may either increase or decrease the individual's vulnerability to mental disorder; and

(d) issues relevant to the planning of appropriate phc services, such as prior use of services, attitudes towards services, and availability of

different service modalities within the community.

(ii) Training of community health workers.

Community health workers need to be trained to develop their understanding of mental disorder in terms of the dimensional model. As this model conceives of individuals as moving between different dimensions in response to environmental stresses (rather than fixed in a two-dimensional position), it increases the worker's sensitivity to the various factors which contribute to an individual's vulnerability or resilience to mental disorder.

The potentially high rate of psychotic disorder in the community means that it is also important that community workers be trained to identify psychotic symptoms, and to discriminate between those persons who require hospital treatment and those who can be managed in the home. At the same time, the estimated high rate of depression and anxiety in Mamre suggests that workers also need to be trained to recognise the psychosocial and psychological symptoms relevant to these disorders, since these are often presented as physical complaints. Any such training should build on and enhance skills already held by community health workers. This mode of training is in fact already in use in Mamre (Swartz, 1993).

Referral and treatment facilities

Although the weighted prevalence of depression and anxiety related disorders was high (20%), many of these disorders are likely to remit spontaneously over time, unless maladaptive coping mechanisms are learnt (Goldberg & Huxley, 1992). Of the second-stage respondents who requested further psychiatric counselling, several no longer felt the need for counselling when later contacted by the clinical psychologist working in Mamre. It is possible that the administration of the PSE alone may have resulted in the spontaneous remission of disorder in a number of cases. Goldberg and Huxley (1992) recommend that in order to decrease inappropriate utilization of phc services, and to restore functioning as quickly as possible, community health workers should be trained to identify and manage this level of disorder. The suggested training programme includes the use of specially designed treatment packages, such as anxiety management techniques, and methods of bereavement counselling, to be used with certain targeted groups. Eight community workers in Mamre have already been taught basic counselling skills by a clinical psychologist employed by the Department of Psychology, at the University of Cape Town, and certain preventative programmes have been initiated over the past two years by the same Department. Counselling skills specific to the management of depression and anxiety, may be of particular

relevance in the community, especially if counselling is directed at groups known to be vulnerable. The use of community workers to deal with such disorders also frees more highly trained professional mental health personnel to attend to the more serious mental disorders and to offer consultation services. In some respects, this recommendation has already been followed in place in Mamre. During the past two years, the psychologist in Mamre has established a role as a "liaison" psychologist to health workers. In this role, she has given guidance and support to health workers in their management of patients, rather than becoming involved in direct consultation with the individual patients themselves.

Perhaps most important of all, community health workers need to be trained to change attitudes towards mental disorder in the community and to demonstrate that distressed persons **can** be helped.

However, psychotic disorders often require specialist mental health services. While it is recommended that phc workers be trained to identify these disorders, it is also recommended that adequate provision be made for persons with psychotic disorders to receive effective mental health care. Goldberg & Huxley (1992) comment that it is inaccurate to think of mental disorder as a homogeneous entity, so that all resources are funnelled into primary care. The spontaneous remission rate of disorders commonly found in the community is high, while those referred to mental health services are much more likely to have special treatment needs. Notwithstanding the need for cost-effective specialized mental health care services, it is recommended that phc is increasingly integrated with these specialized services.

RECOMMENDED IMPROVEMENTS IN GENERAL METHODOLOGY TO INCREASE THE RELIABILITY, VALIDITY AND USEFULNESS OF RESULTS

Measures regarding the SRQ

The advisability of modifying or completely replacing the SRQ as a psychiatric screening instrument, has already been discussed. If the SRQ is to be used in future studies in Mamre, certain precautions and procedures need to be considered:

(i) Translation of the SRQ

Particular attention needs to be given to the possible rephrasing of some SRQ items, in order to take into consideration the cultural differences between other communities and the specific context in which the SRQ is to be used. Bilingual persons from the community elected for study, need to be appointed as early as the initial translation phase, to collaborate with personnel trained in psychology assigned to translate the SRQ.

(ii) SRQ high false positive rate

Before the SRQ can be used with confidence, its high false positive rate in this study needs to be further investigated in a similar research context. Increased standardization of the instrument (WHO, 1994) means that the number of false positives in future studies are likely to be reduced.

(iii) Training of lay-interviewers on the SRQ

Although the SRQ is a structured instrument with a binomial response format, an element of judgement is required by the lay-interviewers in order to detect the presence or absence of symptoms. For example, lay-interviewers must be able to judge whether to give a positive or negative score to answers such as "yes - once in a while", or "yes - some time back". This weakness has been recognized, and is addressed in the WHO's user's guide to the SRQ (1994) in which measures to increase the instrument's standardization are given. It is important that this aspect of the questionnaire is given sufficient emphasis during training.

Measures regarding the PSE

It was concluded that the PSE was neither more nor less appropriate as a second-stage criterion measure in this context, than either the PAS or "suggested clinical diagnoses". Based on this conclusion, it is an open question as to whether other case-finding instruments such as the Diagnostic Interview Schedule (DIS) (Robins et al., 1981) or the Clinical Interview Schedule (CIS-R) (Lewis & Pelosi, 1990) would be more effective criterion measures in a similar context. While the DIS has the advantage of being a structured psychiatric interview, which can be administered by trained lay-interviewers and which can obtain information on a much broader range of diagnoses, its length, and the problems associated with the use of non-psychiatrically trained personnel in psychiatric epidemiology research, are severe limitations (Parry, 1992). While the CIS-R was designed to be sensitive to the level of mental disorder commonly encountered in the community, its suitability for community research in South Africa has yet to be investigated. Commenting on an earlier version of the CIS-R, the CIS (Goldberg et al., 1970), Cleary et al., (1982) observed that the sensitivities of the GHQ were within 1% of each other when either the PSE or the CIS was used as the criterion measure. Given the substantial body of professionals in the psychiatric field in South Africa who are familiar with the PSE and its administration, it is recommended that, rather than investing resources in the exploration of alternative psychiatric clinical instruments whose appropriateness is no more evident than the PSE, these resources would be best spent on the following measures:

(i) Investigation of the PSE's validity in the context of South African community research

Given the lack of an indisputable definition of mental disorder, this may not be an easy task. Nevertheless, if the PSE is to be used as a criterion measure, either in validity studies or on its own in a single-stage design, it should itself be validated in the context in which it is to be used. While the PSE-9 has shown itself to be a valid instrument of mental disorder in hospital and clinic samples, there is still uncertainty regarding its validity when used in the community, in the South African context. The PSE-10 (SCAN) (Wing et al., 1990) currently in development, which covers a broader area (including questions on alcohol abuse) may be a more appropriate instrument.

(ii) Provision of thorough training in the administration of the PSE

Thorough training in the PSE is essential. Wing et al., (1977) recommend a period of at least one week's training on the PSE to be given by an experienced clinician who has been trained specifically to instruct personnel on the PSE. Training should also include practice interviews on hospital and community patients with video feed-back of interviews.

(iii) The implementation of effective and comprehensive inter-rater reliability studies

The use of power analysis to estimate optimal first and second-stage sample size (Shrout, 1992) did provide a cost-effective estimate of the **minimum** number of persons required to receive both the first and second-stage interviews, and also the **minimum** sample size from which prevalence could be estimated. However, when cross-tabulations between cells were required to establish a possible distinctive demographic profile of persons at risk of mental disorder, the sample size was grossly inadequate. This was a severe limitation of the research. Although there was some suggestion of an association between mental disorder and certain demographic variables, the numbers were too small for the association to reach significance. This aspect of the research needs careful consideration when estimating sample size.

IMPLEMENTATION

The implementation of the findings of this study is subject to the approval of the Mamre steering committee. It is envisaged that the data obtained will be used as:

(i) base-line information when planning future research studies relevant to the implementation of phc services in the community; and

(ii) to assist in the motivation, guidance, and training of community health workers, in their efforts to identify and assist those experiencing mental distress.

It was also the intention of this study that any person identified by the PSE as a psychiatric case who wished to have further contact with the psychiatric services, would have the opportunity to consult with the clinical psychologist working in Mamre under the auspices of the Mamre Community Health Project.

CONCLUSION

Overall, the study achieved what it set out to do. Base-line prevalence of mental disorder in the adult population was estimated, a trend was discerned in those who were mentally distressed and reported physical ill health, and their rate of consultation, and methodological issues pertaining to both the SRQ and the PSE in this context were investigated. While it may be argued that none of the data obtained provided conclusive results or statements, the cost of the research may be justified by the nature of the issues identified for further enquiry. No previous psychiatric two-stage design has been used in a community-based study in South Africa. Similar studies which have been conducted, such as Freeman (1990), Thom et al. (1991), and Miller et al. (1991), have either used inadequate research designs or have assumed that the instruments selected are appropriate measures of psychiatric morbidity in the South African context. Notwithstanding the insufficient PSE inter-rater reliability study, the present research has demonstrated that there may be serious methodological problems with the use of single screening instruments in South Africa. These may have yielded false estimates of psychiatric morbidity in the populations in which they were used. In terms of psychiatric epidemiological instruments appropriate for community studies in South Africa, future research needs to be directed in the following two ways: i) the validation of existing psychiatric clinical interviews such as the PSE, whether they are to be used as second-stage criterion measures or on their own; and ii) the modification of existing screens, or even the design of a screen specific to the South African context. With regard to the development of an instrument specific to the South African context, the first draft of the South African Mental Health Indicators (SAMHI) (Parry, 1994, personal communication) has been designed. This instrument is somewhere between a screening and a second-stage instrument, and the final version will still have to be validated before it can be considered more useful than existing instruments in the South African context. However, the expense and effort of such studies suggest that a nationally orchestrated plan for multi-site investigations would allow for the best use of available resources. The lessons of the current study - both positive

and negative- could help inform planning for a broader-based national psychiatric epidemiological effort.

Issues which have been identified in the present research need to be investigated further in similar community studies using a two-stage method of case-detection.

REFERENCES

REFERENCES

American Psychiatric Association. (1980). Diagnostic and statistical manual of mental disorders, 1st ed., DSM-111. Washington DC.

American Psychiatric Association. (1987). Diagnostic and statistical manual of mental disorders, 3rd ed., DSM-111R. Washington DC.

Aneshensel, C., Fredricks, R., & Huba, G. (1984). Depression and physical illness: a multi-wave, non-recursive causal model. Journal of Health and Social Behaviour, 25, 350-371.

Angst, J., & Dobler-Mikola, A. (1984). The Zurich Study. 11. The continuum from normal to pathological depressive mood swings. European Archives of Psychiatry and Neurological Science, 234, 21-29.

Antony, J., Folstein, M., Romanoski, A., Van Korft, M., Nestadt, G., Chahal, R., Merchant, A., Brown, H., Shapiro, S., Kramer, M., & Gruenberg, E. (1985). Comparison of the lay Diagnostic Interview Schedule and a standard psychiatric diagnosis. Archives of General Psychiatry, 42, 667-675.

Azevedo, C., & Gilbert, A. (1991). A follow-up study of persons identified in the base-line study in Mamre. Unpublished honours paper. Department of Community Health, UCT.

Banks, M. (1983). Validation of GHQ-28 in young community sample. Psychological Medicine, 13, 349-353.

Bebbington, P., Dean, C., Der, G., Hurry, J., & Tennant, C. (1991). Gender, parity and the prevalence of minor affective disorder. British Journal of Psychiatry, 158, 40-45.

Bebbington, P., Hurry, J., Tennant, C., Sturt, E., & Wing, J. (1981). Epidemiology of mental disorders in Camberwell. Psychological Medicine, 11, 561-579.

Bellantuono, C., Fiorio, R., Williams, P., & Cortina, P. (1987). Psychiatric morbidity in an Italian general practice. Psychological Medicine, 17, 243-247.

Ben-Arie, O., Swartz, L., Teggin, A., & Elk, R. (1983). The coloured elderly in Cape Town - A psychological, psychiatric and medical community survey: Part II. Prevalence of psychiatric disorders. South African Medical Journal, 64, 1056-1061.

Benjamin, S., Decalmer, P., & Haran, D. (1982). Community screening for mental illness: A validity study of the GHQ. British Journal of Psychiatry, 140, 174-180.

Ben-Tovim, D. (1987). Development psychiatry: Mental health and primary health care in Botswana. London: Tavistock Publications.

Bester, F., Weich, D., & Gagiano. C. (1991). Mental disorders of elderly blacks in the Orange Free State. The Journal of Age Related Disorders, 3, 13-16.

Boardman, A. (1987). The GHQ and the detection of emotional disorders by general health practitioners. British Journal of Psychiatry, 151, 373-381.

Bridges, K., & Goldberg, D. (1986). The validation of the GHQ-28 and the M.M.S.E. in neurological in-patients. British Journal of Psychiatry, 148, 548-553.

Brislin, R. (1986). The wording and translation of research instruments. In W.J. Lonner & J.W. Berry (Eds), Field-methods in cross-cultural research (pp. 139-164). Berkeley: Sage Publications.

Brown, G. & Harris, T. (1978). Social origins of depression. London: Tavistock.

Chan, D. (1985). The Chinese version of the General Health Questionnaire: Does language make a difference? Psychological Medicine, 15, 147-155.

Chan, D., & Chan, L. (1983). Reliability, validity and the structure of the GHQ in a Chinese context. Psychological Medicine, 13, 363-371.

Cheng, T.A. (1985). A pilot study of mental disorders in Taiwan. Psychological Medicine, 15, 195-203.

Cheng, T.A. (1988). A community study of minor psychiatric morbidity in Taiwan. Psychological Medicine, 18, 953-968.

Cheng, T., & Williams, P. (1986). The design and development of a screening questionnaire (CHQ) for use in community studies of mental disorders in Taiwan. Psychological Medicine, 16, 415-422.

Chong, M., & Wilkinson, G. (1989). A validation study of the GHQ. Psychological Medicine, 19, 495-505.

Cicchetti, D. (1976). Assessing inter-rater reliability rating scales: Resolving some basic issues. British Journal of Psychiatry, 129, 452-456.

Cleary, P., Goldberg, I., Kessler, L., & Nyez, G. (1982). Screening for mental disorder using primary care patients. Usefulness of GHQ. Archives of General Psychiatry, 39, 837-40.

Cleary, P., & Mechanic, D. (1983). Sex differences in psychological distress among married people. Journal of Health and Social Behaviour, 23, 159-169.

Climent, C., Diop, B., Harding, T., Ibrahim, H., Ladrido-Ignacio, L., & Wig, N. (1980). Mental health in primary care. WHO Chronicle, 34, 231-236.

Cohen, P., & Cohen, J. (1984). The clinician's illusion. Archives of General Psychiatry, 14, 1178-1182.

Cooper, J., Copeland, J., Brown, G., Harris, T., & Gourlay, A. (1977). Further studies on interviewer training and interrater reliability of the Present State Examination (PSE). Psychological Medicine, 7, 517-523.

Cox, B., Blaxter, M., Buckle, A., Fenner, N., Golding, J., Gore, M., Huppert, F., Nickson, J., Roth, M., Stark, J., Wadsworth, M., & Wichelow, M. (1987). The health and lifestyle survey. Cambridge: Health Promotion Research Trust.

Dean, C., Surtees, P., & Sashidharan, S. (1983). Comparison of research diagnostic systems in an Edinburgh community sample. British Journal of Psychiatry, 142, 247-256.

De Jong, J., De Klein, G., & Ten Horn, S. (1986). A baseline study on mental disorders in Guinea-Bissau. British Journal of Psychiatry, 148, 27-32.

Dhadphale, M., & Ellison, R., Griffin, L. (1982). Frequency of mental disorder among outpatients at a rural district hospital in Kenya. The Central African Journal of Medicine, 28, 85-89.

Dhadphale, R., & Griffin, L. (1983). The frequency of psychiatric disorders among patients attending semi-urban and rural general out-patient clinics in Kenya. British Journal of Psychiatry, 142, 379-383.

Dick, B., Spencer, I., Watermeyer, G., Bourne, D., Wolff, E., & Moyle, G. (1978). Chronic illness in non-institutionalized persons. Part 1. Prevalence and epidemiology. South African Medical Journal, 53, 892-904.

Dick, B., Spencer, I., Watermeyer, G., Bourne, D., Wolff, E., & Moyle, G. (1978). Chronic illness in non-institutionalized persons. Part II. Who cares?. South African Medical Journal, 53, 928-937.

Diop, B., Collignon, R., Gueye, M., & Harding, T. (1982). Diagnosis and symptoms of mental disorder in a rural area of Senegal. African Journal of Medicine and Science, 11, 95-103.

Dohrenwend, B. (1990). Problems of validity in field studies revisited. Psychological Medicine, 20, 195-208.

Dohrenwend, B.P., & Dohrenwend, B.S. (1982). Perspectives on the past and future of psychiatric epidemiology. American Journal of Public Health, 72, 1271-1279.

Dohrenwend, B., Levav, I., & Shrout, P. (1986). Screening scales from the Psychiatric Epidemiology Research Interview (PERI). In M. Weissman, J. Myers, & C. Ross (Eds), In community surveys of psychiatric disorders. New Brunswick: University Press.

Dube, K. (1970). A study of prevalence and biosocial variables in mental illness in a rural and an urban community in Uttat Pradesh - India. Acta Psychiatrica Scandinavia, 46, 327-359.

Duncan-Jones, P. (1979). Validity and uses of the GHQ. British Journal of Psychiatry, 135, 382.

Duncan-Jones, P., & Henderson, P. (1978). The use of a two-stage design in a prevalence survey. Social Psychiatry, 13, 231-237.

Duncan-Jones, P., Fergusson, D., Ormel, J., & Horwood, L. (1990). A model of stability and change in minor psychiatric symptoms: Results from three longitudinal studies. Psychological Medicine, Monograph Supplement, 18.

Eaton, W., & Ritter, C. (1988). Distinguishing anxiety and depression with field survey data. Psychological Medicine, 18, 155-166.

Elton, M., Patton, G., Weyerer, S., Diallina, M., & Fichter, M. (1988). A comparative investigation of the principal component structure of the 28 item version of the General Health Questionnaire (GHQ). Acta Psychiatric Scandinavia, 77, 124-132.

Endicott, J., & Spitzer, R. (1978). A diagnostic interview: The schedules of affective disorders and schizophrenia. Archives of General Psychiatry, 35, 837-844.

Fabrega, H. (1989). Cultural relativism and psychiatric illness. The Journal of Nervous and Mental Disease, 177, p. 415-425.

Feighner, J., Robins, E., Guze, S., Woodruff, R., Winokur, G., & Munoz, R. (1972). Diagnostic criteria for use in psychiatric research. Archives of General Psychiatry, 26, 57-63.

Ferrando, S. (1988). Race and culture in psychiatry. London & New York: Tavistock/Routledge.

Finlay-Jones, R., & Burvill, P. (1978). Contrasting demographic patterns of minor psychiatric morbidity in general practice and the community. Psychological Medicine, 8, 455-466.

Finlay-Jones, R., & Ekhardt, B. (1984). Psychiatric disorder among young unemployed. Australian and New Zealand Journal of Psychiatry, 15, 265-270.

Finlay-Jones, R., & Murphy, E. (1979). Severity of psychiatric disorder and the 30-item General Health Questionnaire. British Journal of Psychiatry, 134, 609-616.

Fontensesi, F., Gobetti, C., Zimmerman-Tansella, Ch., & Tansella, M. (1985). Validation of the Italian version of the GHQ in a general practice setting. Psychological Medicine, 15, 411-416.

Foulds, G., & Hope, K. (1968). Manual of the symptom sign inventory (SSI). University of London Press: London.

Freeman, M. (1990). The need for improved mental health care in South Africa. The Centre for the Study of Health Policy. University of Witwatersrand, Johannesburg.

Freeman, M. (1991). An evaluation of mental health services in the South-Eastern Transvaal. The Centre for the Study of Health Policy. University of Witwatersrand, Johannesburg.

Gaines, A. (1992). Ethnopsychiatry: The cultural construction of professional and folk psychiatries. State University of New York Press. New York.

Gask, L. (1988). Key to the modified PAS. Institute of Psychiatry, Manchester.

Gater, R., De Almeidae Sousa, R., Caraveo, J., Chandrasekar, C., Dhadphale M., Golding D., al Kathiri, A., de Llano, G., Mubbashar, M., Silhan, K., Thong, D., Torres, F. & Sartorius, N. (1991). The parthways to psychiatric care: a cross-cultural study. Psychological Medicine.

Gillis, L.S. (1991). Depression in elderly black persons. Paper presented at the 10th Epidemiological Conference. Cape Town, South Africa.

Gillis, L., Elk, R., Ben-Arie, O., & Teggin, A. (1982). The present state examination: Experiences with Xhosa speaking psychiatric patients. British Journal of Psychiatry, 141, 143-147.

Gillis, L., Lewis, J., & Slabbert, M. (1968). Psychiatric disorder amongst the coloured people of the Cape Peninsula. British Journal of Psychiatry, 114, 1575-1587.

Goldberg, D. (1972). The detection of psychiatric illness by questionnaire. Maudsley Monograph No. 21, London: Oxford University Press.

Goldberg, D., Cooper, B., Eastwood, M., Kedword, M., & Shepherd, H. (1970). A standard psychiatric interview suitable for use in community surveys. British Journal of Preventative and Social Medicine, 24, 18-26.

Goldberg, D., & Hillier, V. (1979). A scaled version of the General Health Questionnaire. Psychological Medicine, 9, 139-145.

Goldberg, D., & Huxley, P. (1980). Mental illness in the community. London: Tavistock Publications.

Goldberg, D., & Huxley, P. (1992). Common mental disorders. A bio-social model. London & New York: Tavistock/Routledge.

Goldberg, D., Kay, C., & Thompson, L. (1976). Psychiatric morbidity in general practice and the community. Psychological Medicine, 6, 565-569.

Goldberg, D., & Williams, P. (1988). A user's guide to the general health questionnaire. Windsor: NFER-Nelson.

Goodchild, M., & Duncan-Jones, P. (1985). Chronicity and the General Health Questionnaire. British Journal of Psychiatry, 146, 55-61.

Grayson, D., Bridges, K., Cook, D., & Goldberg, D. (1990). The validity of diagnostic systems for common mental disorders: A comparison between the PSE-CATEGO and the DSMIII systems. Psychological Medicine, 20, 209-218.

Grayson, D., Bridges, K., Duncan-Jones, P., & Goldberg, D. (1987). The relationship between symptoms and diagnoses of minor psychiatric disorder in general practice. Psychological Medicine, 17, 933-943.

Harding, T., Climent, C., Diop, M., Giel, R., Murthy, R., Suleiman, M., & Wig, N. (1983). The WHO collaborative study on strategies for extending mental health care, II: The development of new research methods. American Journal of Psychiatry, 140, 1474-1430.

Harding, T., De Arango, M., Baltazar, J., Climent, C., Ibrahim, H., Ladrido-Ignacio, L., Murthy, R., & Wig, N. (1980). Mental disorders in primary health care: A study of their frequency and diagnosis in four developing countries. Psychological Medicine, 10, 231-241.

Heller, K., Price, R., Reinharz, S., Riger, S., Wandersman, A., & D'Aunno, T. (1977). Psychology and community change: Challenges of the future. Pacific Grove, California: Brooks/Cole Publishing.

Helman, C. (1990). Culture, health and illness. An introduction for health professionals. UK: Butterworth-Heinemann.

Helzer, J., Robins, L., McEvoy, L., Spitznage, E., Stoltzamn, R., Farmer, A., Brokington, I. (1985). A comparison of clinical and Diagnostic Interview Schedule diagnoses: Physician reexamination of lay-interviewed cases in the general population. Archives of General Psychiatry, 42, 657-666.

Henderson, S., Duncan-Jones, P., Byrne, D., & Adcock, S. (1979). Psychiatric disorder in Canberra: A standardized study of prevalence. Acta Psychiatrica Scandinavia, 60, 355-374.

Hirschfeld, R., & Cross, C. (1982). Epidemiology of affective disorders. Psychosocial risk factors. Archives of General Psychiatry, 3935, 46-50.

Hobbs, P., Ballinger, C., Greenwood, C., Martin, B., & McClure, A. (1984). Factor analysis and validation of the General Health Questionnaire in men: A general practice survey. British Journal of Psychiatry, 144, 270-275.

Hodiamont, P., Peer, N., & Syben, N. (1987). Epidemiological aspects of psychiatric disorder in a Dutch health centre. Psychological Medicine, 17, 495-505.

Hoeper, E., Nycz, G., Cleary, P., Regier, D., & Goldberg, D. (1979). Estimating prevalence of RDC mental disorders in primary medical care. International Journal of Mental Health, 8(2), 6-15.

Hoffman, M., Yach, D., Katzenellenbogen, J., Pick, W., & Klopper, M. (1988). Mamre community health project - rationale and methods. South African Medical Journal, 74, 323-367.

Hollifield, M., Katon, W., Spain, D., & Pule, L. (1990). Anxiety and depression in a village in Lesotho, Africa: A comparison with the United States. British Journal of Psychiatry, 156, 343-350.

Isaacs, C. (1990). A calibration study of the Afrikaans version of the GHQ-28 with coloured in-patients. Unpublished manuscript.

Katzenellenbogen, J., Pick, W., Hoffman, M., & Weir, G. (1988). Community participation in the Mamre community health project. South African Medical Journal, 74, 335-338.

Kessler, R., & Neighobours, H. (1986). A new perspective on the relationships among race, social class and psychological distress. Journal of Health and Social Behaviour, 27, 107-115.

Kinney, J., & Leaton, G. (1978). Handbook of alcohol information. St. Louis: Mosby & Company.

Kleinman, A. (1977). Depression, somatization and the "new cross-cultural psychiatry". Social Science & Medicine, 11, 1-10.

Kleinman, A. (1988). Rethinking psychiatry. From cultural category to personal experience. London: The Free Press.

Kortman, F. (1990). Psychiatric case finding in Ethiopia: Shortcomings of the Self-Report Questionnaire. Culture, Medicine & Psychiatry, 14, 381-391.

Kortman, F., & Ten Horn, S. (1988). Comprehension and motivation in response to a psychiatric screening instrument: Validity of the SRQ in Ethiopia. British Journal of Psychiatry, 153, 95-101.

Krause, I., Rosser, R., Khiani, M., & Lotaij, N. (1990). Psychiatric morbidity among Punjabi medical patients in England measured by GHQ-28. Psychological Medicine, 20, 711-719.

Kreitman, N. (1975). In P. Sainsbury & N Kreitman (Eds), Methods of psychiatric research: and introduction for clinical psychologists. London: Oxford University Press.

Lambley, P. (1973). Psychology and socio-political reality: Apartheid psychology and its links with trends in humanistic psychology and behaviour theory. International Journal of Psychology, 8, 73-79.

Leon, C. (1972). Psychiatry in Latin America. British Journal of Psychiatry, 121, 121-136.

Lewis, G., & Pelosi, A. (1990). Manual of the revised clinical interview schedule (CIS-R). London: Institute of Psychiatry.

Lewis, G., & Williams, P. (1989). Clinical judgement ad the standardized interview in psychiatry. Psychological Medicine, 19, 971-979.

Littlewood, R. (1990). From categories to contexts: A decade of the new cross-cultural psychiatry. British Journal of Psychiatry, 156, 308-327.

Lobo, A., Pérez Echeverria, M., & Artal, J. (1986). Validity of the scaled version of GHQ-28 in a Spanish population. Psychological Medicine, 16, 135-140.

Luria, R. (1979). Reliability and descriptive validity of PSE syndromes. Archives of General Psychiatry, 36, 1187-1195.

Mann, A., Wakeling, K., Monck, E., Dobbs, R., & Szmukler, G. (1983). Screening for abnormal eating attitudes and psychological morbidity in an unselected population of 15 year old schoolgirls. Psychological Medicine, 13, 573-580.

Marie, J., & Williams, P. (1984). Minor psychiatric disorder in primary care in Brazil: A pilot study. Psychological Medicine, 14, 223-227.

Marie, J., & Williams, P. (1985). A comparison of the validity of two psychiatric screening questionnaires (GHQ-12 and SRQ-20) in Brazil, using Receiver Operating Characteristic (ROC) analysis. Psychological Medicine, 15, 651-659.

Marie, J., & Williams, P. (1985). A validity study of a psychiatric screening questionnaire (SRQ-20) in primary care in the city of Sao Paulo. British Journal of Psychiatry, 148, 23-26

Mayfield, D., McLeod, G., Hall, P. (1974). Validation of a New Alcoholism Screening Instrument. American Journal of Psychiatry, 131, 1221-1223.

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

Medina-Mora, M., Padilla, G., Campillo-Serrano, D., Mas, C., Ezban, M., Caraveo, J., & Corona, J. (1983). The factor structure of the GHQ: A scaled version for hospital general practice service in Mexico. Psychological Medicine, 13, 355-361.

Metz, C. (1978). Basic principles of ROC analysis. Seminars in Nuclear Medicine, 8, 283-298.

Miller, T. (1989). Unpublished data.

Miller, T., Swartz, L., & Rumble, S. (1991, July). Psychological issues in primary health care in Mamre: An exploratory study. Paper presented at the 10th Epidemiological Conference, Cape Town, South Africa.

Neale, J., & Liebert, R. (1980). Science and behavior. An introduction to methods of research, 2nd edition. New Jersey: Prentice-Hall Inc.

Newman, S., Bland, R., & Orn, H. (1988). The General Health Questionnaire as a screening instrument in a community survey. Psychological Medicine, (submitted).

Orley, J., & Wing, J. (1979). Psychiatric disorder in two African villages. Archives of General Psychiatry, 36, 513-520.

Ormel, J., & Giel, R. (1990). Medical effects of non-recognition of affective disorders. In N. Sartorius, D. Goldberg, G. de Girolamo, J., Cosae Silva, Y., LeCrubier, & H.U. Witchen (Eds), Primary care: Psychological disorders in general medical settings, Bern: Huber-Hagrete.

Pan, P., & Goldberg, D. (1990). A comparison of the validity of GHQ-12 and CHQ-12 in Chinese primary care patients in Manchester. Psychological Medicine, 20, 931-940.

Parry, C. (1992). Psychiatric epidemiology in Africa: methodological considerations. Paper presented at the 11th annual conference of the Epidemiological Association of Southern Africa, Johannesburg, August, 1992.

Rabins, P., & Brooks, B. (1981). Emotional disturbance in multiple sclerosis patients: validity of the General Health Questionnaire. Psychiatric Medicine, 11, 425-427.

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

Reeler, A. (1987). Psychological disorders in Africa. Part III: Service delivery. Central African Journal of Medicine, 33, 37-41.

Riegelman, R.K. (1981). Studying a study and testing a test. Boston: Little, Brown & Company.

Robins, L. (1985). Epidemiology: Reflections on testing the validity of psychiatric interviews. Archives of General Psychiatry, 42, 918-924.

Robins, E., Gruze, S., Woodruff, R., Windkut, G., & Munoz, R. (1972). Diagnostic criteria for use in psychiatric research. Archives of General Psychiatry, 26, 57-63.

Robins, L., Wing, J., Wittchen, H., Helzer, J., Babor, T., Burke, J., Farmer, A., Jabelenski A., Pickens, R., Regier, D., Sratorius, N., & Towle, L. (1988). The Composite International Diagnostic Interview. Archives of General Psychiatry, 45, 1069-1077

Rodgers, B., & Mann, S. (1986). The reliability and validity of PSE assessments by lay interviewers: A national population survey. Psychological Medicine, 16, 689-700.

Rogers, P. (1993). Illness experience in primary health care in Mamre. (Unpublished masters thesis, UCT).

Salmon, J. (1993). Prevalence of closed-angle glaucoma and open-angle glaucoma in Mamre, Western Cape. Archives of Ophthalmology, 3, 1263-1269.

Sashidharan, S., Surtees, P., Kreitman, N., Ingham, J., & Miller, P. (1988). Hospital-treated and general population morbidity from affective disorders. British Journal of Psychiatry, 152, 499-505.

Sashhidharan, P., Surtees, P., Kreitman, N., Ingham, J., & Miller, P. (1990). Affective disorders among women in the general population and among those referred to psychiatrists. Clinical features and demographic correlates. British Journal of Psychiatry, 157, 828-834.

Sen, B., Wilkinson, G., & Marie, J. (1987). Psychiatric morbidity in primary health care. A two-stage screening procedure in developing countries: Choice of instruments and cost-effectiveness. British Journal of Psychiatry, 151, 33-38.

Shepherd, M., Wilkinson, G., & Williams, P. (Eds), (1986). Mental illness in primary care settings. London: Tavistock.

Shrout, P. (1992). Statistical design of screening procedures. To be published in C. Attkisson & J. Zich (Eds), Screening for depression in primary care. New York: Routledge, Chapman & Hall, Inc.

Skuse, D., & Williams, P. (1984). Screening for psychiatric disorder in general practice. Psychological Medicine, 14, 365-377.

Spitzer, R., Endicott, J., & Robins, E. (1978). Research Diagnostic Criteria: rationale and reliability. Archives of General Psychiatry, 35, 773-782.

Surtees, P., Dean, C., Ingham, J., Kreitman, N.,MacMiller, P., & Sashidharan, S. (1983). Psychiatric disorders in women from an Edinburgh community: Association with demographic factors. British Journal of Psychiatry, 142, 238-246.

Surtees, P., & Sashidharan, S. (1986). Psychiatric morbidity in two matched community samples: A comparison of rates and risks in Edinburgh and St. Louis. Journal of Affective Disorders, 10, 101-113.

Swartz, L. (1985). Issues for cross-cultural psychiatric research in South Africa. Culture, Medicine & Psychiatry, 9, 59-74.

Swartz, L. (1986). Transcultural psychiatry in South Africa: Part I. Transcultural Psychiatric Research Review, 23, 273-303.

Swartz, L. (1987). Transcultural psychiatry in South Africa: II. Cross-cultural issues in mental health practice. Transcultural Psychiatric Research Review, 24, 5-30.

Swartz. L. (1993). Clinical Psychiatry in Mamre Community Health Project: Issues at the interface between research and service provision. Presentation given at the MRC, November, 1993.

Swartz, L., Ben-Arie, O., & Teggin, A. (1985). Subcultural delusions and hallucinations. Comments on the present state examination in a multi-cultural context. British Journal of Psychiatry, 146, 391-394.

Tarnopolsky, A., Hard, D., McLean, E., Roberts, H., & Wiggins, R.D. (1979). Validity and uses of a screening questionnaire (GHQ) in the community. British Journal of Psychiatry, 134, 508-515.

Tenenbein, A. (1970). A double sampling scheme for estimating from binomial data with misclassification. Journal of the American Statistical Association, 65, 331, p. 1350-1361.

Thom, R., Zwi, R., & Reinach, S. (1993). The prevalence of psychiatric disorders at a primary care clinic in Soweto, Johannesburg. South African Medical Journal, 83, 653-655.

Van den Brink, W., Koeter, M., Ormer, J., Dijkstra, W., Giel, R., Slooff, C., & Wonifarth, T. (1989). Psychiatric diagnosis in an outpatient population. A comparative study of PSE-Catego and DSM-111. Archives of General Psychiatry, 46, 369-372.

Vazquez-Barquero, J., & Diez-Manrique, J. (1986). Two stage design in a community survey. British Journal of Psychiatry, 149, 88-97.

References

Vazquez-Barquero, J., Diez-Manrique, J., Pera, C., Aldama, J., Samaniego-Rodriguez, C., Mendez Arango J., & Mirapeix, C. (1987). A community mental health survey in Cantabria: A general description of morbidity. Psychological Medicine, 17, 207-242.

Vazquez-Barquero, J., Munoz, P., & Jauregui, V. (1981). The interaction between physical illness and neurotic morbidity in the community. British Journal of Psychiatry, 139, 328-335.

Vazquez-Barquero, J., Wilkinson, G., Williams, P., Diez-Manrique, J., & Péna, C. (1990). Mental health and medical consultation in primary care settings. Psychological Medicine, 20, 681-694.

Vazquez-Barquero, J.K., Williams, P., Diez-Manrique, J.F., Lequerica, J., & Arenal, A. (1988). The factor structure of the GHQ-60 in a community sample. Psychological Medicine, 18, 211-218.

Vieweg, B., & Hedlund, J. (1983). GHQ. A comprehensive review. Journal of Operational Psychology, 14, 74-81.

Weyerer, S. (1990). Relationships between physical and psychological disorders. In N. Sartorius, D Goldberg, G de Girolamo, J. Costa e Silva, Y. Le Crubier, & H. Wittchen (Eds.), Psychological disorders in general medical settings. Bern: Huber-Hogrefe.

Williams, P., Goldberg, D., & Marie, J. (1987). The validity of the General Health Questionnaire. Unpublished research report.

Williams, P., Tarnopolsky, A., & Hand, D. (1980). Case definition and case identification in psychiatric epidemiology: Review and assessment. Psychological Medicine, 10, 101-114.

Williams, P., Tarnopolsky, A., Hand, D., & Shepherd, M. (1986). Minor psychiatric morbidity and general practice consultations: the West London survey. Psychological Medicine, Monograph Supplements, 9.

Wing, J. (1983). Use and misuse of the PSE. British Journal of Psychiatry, 143, 111-117.

Wing, J., Barbor, T., Burgha, T., Burke, T., Cooper, J., Giel, R., Jabolensky, A., Regier, D., & Sartorius, N. (1991). SCAN: Schedule for clinical associations in neuro-psychiatry. Archives of General Psychiatry, 47, 589-593.

Wing, J.K., Bebbington, P., & Robins, L.N. (1981). What is a case? UK: Grant McIntyre Ltd.

Wing, J., Cooper, J., & Sartorius, N. (1974). The measurement and classification of psychiatric symptoms. London: Cambridge University Press.

Wing, J., Mann, S., Leff, J., & Nixon, J. (1978). The concept of a "case" in psychiatric population surveys. Psychological Medicine, 8, 203-217.

Wing, J.K., Nixon, J.M., Mann, S.A., & Leff, J. (1977). Reliability of the PSE used in a population study. Psychological Medicine, 9, 505-546.

World Health Organization (1973). International Classification of Disorder, 8th ed. Geneva: World Health Organisation.

World Health Organization (1978). International Study of Schizophrenia (IPPS).

World Health Organization (1988), 10th ed. International Classification of Disease. Geneva: World Health Organisation.

World Health Organization (1994). A user's guide to the Self Reporting Questionnaire (SRQ). Geneva: Switzerland.

Yach, D., & Kuhn, L. (1989). Use of epidemiology in psychiatric research. Paper presented at the Psychiatric Congress, Johannesburg, South Africa.

BIBLIOGRAPHY

BIBLIOGRAPHY[1]

Abramson, J. (1974). Survey methods in community medicine. London: Churchill Livingstone.

Ahlbom, A., & Novell, S. (1984). Introduction to modern epidemiology. USA: Resources Inc.

Alderson, H. (1976). An introduction to epidemiology. London: MacMillan Press.

Corney, R. (1984). The mental and physical health of clients referred to social workers in a local authority department and a general practice attachment scheme. Psychological Medicine, 14, 137-144.

Corser, C., & Philip, A. (1978). Emotional disturbance in newly registered general practice patients. British Journal of Psychiatry, 132, 172-176.

Diamond, E., & Lilienfield, A. (1962). Effects of errors in classification and diagnosis in various types of epidemiological studies. American Journal of Public Health, 1137-1144.

Drennan, G., Levett, A., & Swartz, L. (1991). Hidden dimensions of power and resistance in the translation process: A South African study. Culture, Medicine & Psychiatry, 15, 361-381.

Duncan-Jones, P., Grayson, D., & Moran, P. (1986). The utility of latent trait models in psychiatric epidemiology. Psychological Medicine, 16, 391-405.

Eaton, W., Holzer, C., Von Korff, M., Anthony, J., Helzer, J., George, L., Burnam, A., Boyd, J., Kessler, L., & Locke, B. (1984). The design of the epidemiologic catchment area surveys. Archives of General Psychiatry, 41, 942-948.

Erdreich, L., & Lee, E., (1981). Use of relative operating characteristic analysis in epidemiology. American Journal of Epidemiology, 114, 649-662.

---

1 Bibliography of material consulted in the preparation of a thesis is not standard in the reference section. However, as part of the purpose of the dissertation is to make information regarding method and material available to the South African researcher, a bibliography of material not used in the text but relevant to psychiatric epidemiology, has been included.

Fleiss, J., Gurland, B., & Cooper, J. (1971). Some contributions to the measurement of psychopathology. British Journal of Psychiatry, 119, 647-656.

Fowler, F. Jr., & Mangione, T. (1990). Standardized survey interviewing. Minimizing interviewer-related error. (Applied social research methods.) Series Vol. 18, London: Sage Publications.

Gillis, L.S. (1984). The diagnosis of psychiatric illness in different cultures. Continuing Medical Education, 2, 5-9.

Goldberg, D., & Bridges, K. (1988). Somatic presentations of psychiatric illness in primary care settings. Journal of Psychometric Research, 32, 137-144.

Goldberg, D., Rickels, C., Downing, R., & Herbacher, P. (1976). A comparison of two psychiatric screening tests. British Journal of Psychiatry, 129, 61-67.

Hanley, J., & McNeil, B. (1982). The meaning and use of the area under a Receiver Operating Characteristic (ROC) curve. Radiology, 143, 29-36.

Huppert, F., Gore, M., & Elliot, B. (1988). The value of an improved scoring system (CGHQ) for the General Health Questionnaire in a representative community sample. Psychological Medicine, 18, 1001-1006.

Klerman, G. (1985). Diagnosis of psychiatric disorders in epidemiologic field studies. Archives of General Psychiatry, 42, 723-724.

Knuiman, M., & Burville, P. (1984). A statistical modelling approach to community prevalence data. Psychological Medicine, 14, 167-173.

Lilienfeld, A., & Lilienfeld, D. (1980). Foundations of epidemiology, 2nd ed. Oxford: Oxford University Press.

Masella, A., Sartorius, N., Jablensky, A., & Fenton, F. (1985). Cross-cultural studies of depressive disorders: an overview. In A. Kleinman & B. Good (Eds), culture and depression (pp. 299-324). Berkeley: University of California Press.

Mausner, J., & Kramer, S. (1985). Epidemiology: An introduction to text. USA: Saunders & Co.

Myers, J., Weissman, M., Tischler, G., Holzer, III, C., Leafe, P., Orvaschet, H., Anthony, J., Boyd, J., Burke, J., Kramer, M., & Stoltzman, R. (1984). Six-month prevalence of psychiatric disorders in three communities. Archives of General Psychiatry, 41, 959-967.

Newman, S.C., Shrout, P.E., & Bland, R.C. (1990). The efficiency of two-phase designs in prevalence surveys of mental disorders. Psychological Medicine, 20, 183-193.

Odejide, A., Oyewunmi, L., & Ohaeri, J. (1989). Psychiatry in Africa: An overview. American Journal of Psychiatry, 146, 708-716.

Perugi, G., Musetti, L., Simonini, E., Piagentini, F., Cassano, G., & Alciskal, H. (1990). Gender-mediated clinical features of depressive illness. The importance of temperamental differences. British Journal of Psychiatry, 157, 835-841.

Regier, D., Boyd, J., Burke, J., Rae, D., Myers, J., Kramer, M., Robins, L., George, L., Karno, M., & Locke, A. (1988). One-month prevalence of mental disorders in the United States. Archives of General Psychiatry, 45, 977-985.

Robins, L., Helzer, J., Croughan, J. (1981). National Institute of Mental Health Diagnostic Interview Schedule: Its history, characteristics and validity. Archives of General Psychiatry, 38, 381-389.

Rose, G., & Barker, D. (1986). Epidemiology for the uninitiated (2nd ed.). UK: Tavistock.

Rothman, K. (1982). Modern epidemiology. Toronto: Little, Brown & Co.

Sashidharan, S. (1985). Definitions of psychiatric syndromes - comparison in hospital and general population. British Journal of Psychiatry, 47, 547-557.

Schaeffer, R., Mendenhall, W., & L. Ott. (1986). Elementary survey sampling. Boston: Duxbury Press.

Schlotzhauer, S., & Littel, R. (1987). SAS system for elementary statistical analysis. Cary, NC: SAS Institute Inc.

Schmid, W., Bronisch, T., & Von Zerssen, D. A comparative study of PSE/Catego and Dasika: Two psychiatric computer diagnostic systems. British Journal of Psychiatry, 141, 292-295.

Spitznagel, E., & Helzer, J. (1985). A proposed solution to the base rate problem in the Kappa (K) statistic. Archives of General Psychiatry, 42, 725-728.

Stansfeld, S., Gallacher, J., Sharp, D., & Yarnell, J. (1991). Social factors and minor psychiatric disorder in middle-aged men: A validation study and a population survey. Psychological Medicine, 21, 157-167.

Stewart, G.T. (1972). (Ed.), Trends in Epidemiology. USA: Thomas.

Surtees, P., & McC. Miller, P. (1990). The internal general health questionnaire. British Journal of Psychiatry, 157, 679-686.

Surtees, P., Sashidharan, S., & Dean, C. (1985). Affective disorders amongst women in the general population: A longitudinal study. British Journal of Psychiatry, 143, 176-186.

Swanson-Fisher, B., & Martin, C.. (1981). Standardized interviews in psychiatry: Issues of reliability. British Journal of Psychiatry, 139, 133-143.

Weissman, M., & Myers, J. (1978). Affective disorders in a US urban community. The use of research diagnostic criteria in an epidemiological survey. Archives of General Psychiatry, 35, 705-712.

Williams, P., Tarnopolsky, A., Hand, D., & Shepherd, M. (1986). Minor psychiatric morbidity and general practice consultations: the West London survey. Psychological Medicine, Monograph Supplements, 9.

GLOSSARY

APPENDIX 1.1 ALPHABETICAL GLOSSARY OF TECHNICAL TERMS

APPENDIX 1.1

ALPHABETICAL GLOSSARY OF TECHNICAL TERMS

APPENDIX 1.1 ALPHABETICAL GLOSSARY OF TECHNICAL TERMS

Incidence (see prevalence)

Prevalence

Prevalence and incidence are two measures relevant to psychiatric epidemiology. Incidence measures the number of new occurrences of a disorder in a population within a specified **interval**, while prevalence refers to the total number of cases of disorder within a population at a **specific point in time**. Prevalence is a more inclusive measure than incidence, and is also simpler to calculate. However, it is also more difficult to interpret as it reflects both the incidence and the duration of a disorder (Helman, 1990). For the purposes of the present research, it was decided that prevalence was a more appropriate measure, as it provides relevant data concerning the accurate estimate of mental health status and the demand for treatment at a given point in time (Eaton et al., 1984).

Prevention

One of the concerns of psychiatric epidemiology as broadly understood is to improve the health and well-being of all citizens, not only those with established and recognized disorder. The emphasis is thus on the prevention of the onset of disease, rather than on finding cures once disease has occurred (D'Auanno & Price, 1984, in Helman, 1990). Moreover, the aim is to push the level of detection and intervention back to the precursors and risk factors of mental pathology, in an attempt to detect morbidity in apparently asymptomatic individuals (Mausner & Kramer, 1985). Early identification is important, as it spotlights high-risk groups, such as the unemployed, who may then be targeted for intervention programmes. Programmes which promote psychological competency and reinforce adaptive behaviours, are aimed at assisting those identified as vulnerable in resisting disorder (Heller et al., 1977). Prevention programmes are sometimes difficult to operationalise as the links between mental disorder and precipitating factors are not always apparent (D'Auanno & Price, 1984, in Helman, 1990). Nevertheless, they are consistent with attempts to assess the individual-environment interaction, and promote mental health in general.

Psychiatric caseness/non-caseness

Psychiatric epidemiologists are faced with the major problem of defining psychiatric caseness. This is because:

(i) there is no definitive definition of mental illness;

(ii) the categorical model on which most diagnostic systems and psychiatric epidemiological instruments are based, may not be appropriate to the measurement and description of mental disorders found in the community; and

(iii) disease behaves as a continuously distributed variable within the community, therefore the imposition of an arbitrary threshold between mental health and disorder may not be appropriate.

While the dimensional model (which requires no decision regarding psychiatric caseness or non-caseness) may be the most appropriate measure of mental disorder in the community, it may not be practical in estimating prevalence. "Disease in the population is nearly always quantitative and not qualitative" (Goldberg et al., 1976, p. 36).

In psychiatric epidemiological prevalence studies, the investigator is concerned with whether something is wrong, and not with what is wrong. The accurate measurement of psychiatric caseness is therefore among the most important aspects of the research. One of the major difficulties in psychiatric epidemiological prevalence studies, is the measuring of large numbers of respondents with less severe, non-specific symptoms (Williams et al., 1980). In community based prevalence studies, there is considerable argument regarding precisely what criteria to adopt to designate psychiatric caseness, and what threshold to impose on an hypothetical axis. If either criteria or threshold are inappropriate, prevalence estimates are likely to be inaccurate.

Rate

Prevalence is most commonly estimated in terms of a rate. This is used to indicate how many occurrences of disorder per given unit may occur within a given time. Rates have the advantage that they can be expressed simply and compared with one another by elementary statistical procedures (Kreitman, 1975). The principal difficulty with the concept of rate is that the implication that members of the study population can be unequivocally differentiated into cases and non-cases (Kreitman, 1975). For certain disorders, particularly the more severe mental disorders, clear-cut distinctions are feasible, but for the less severe mental disorders, the issue is less definitive. A rate is meaningful only if it is interpreted in conjunction with the definition of case which the investigator has used (Kreitman, 1985).

Receiver Operating Characteristic (ROC) Analysis (Metz, 1978)

ROC allows the discriminating ability of an instrument to be assessed across the total spectrum of morbidity. It also allows for the assessment of the effect on an instruments

validity coefficients by varying the threshold score. A ROC curve is obtained by plotting sensitivity against false positive rate for all possible thresholds of the screening instrument. Typically the ROC curve is convex and is located in the area above the leading diagonal, which represents chance performance. The better the screening test, the closer the ROC curve will be to the boundary of the square, and the further away form the leading diagonal. Inspection of the ROC curve makes visible the compromise between best sensitivity and specificity, and allows for easy selection of the optimal threshold most effective in meeting the requirements of the study in hand (Goldberg & Williams, 1988).

Validity Coefficients

Five validity coefficients including sensitivity, specificity, overall misclassification rate, positive predictive value and negative predictive value indicate a test's efficiency. Fig. A1.1 graphically shows these coefficients and briefly defines then. Fuller explanations of each coefficient follow.

Fig. 1.1.1 Validity Coefficients

Test score	**Psychiatric interview**	
	non-case	**case**
Above threshold	a false positives	b true cases correctly identified
Below threshold	c true normals correctly identified	d false negatives

Specificity:
proportion of true normals correctly identified $= c/(a+c)$

Sensitivity:
proportion of true cases correctly identified $= b/(b+d)$

Overall Misclassification Rate:
proportion of respondents correctly identified $= (a+d)/(a+b+c+d)$

Positive Predictive Value:
proportion of high scorers who are cases $= b/(a+b)$

Negative Predictive Value:
proportion of low scorers who are not cases $= c/(c+d)$

Overall Misclassification Rate (OMR)

OMR refers to the number of false positives and false negatives a test identifies, and therefore is a measure of the cost of the test. The OMR is a overall summary of the performance of a test.

Positive Predictive Value (PPV) and Negative Predictive Value (NPV)

PPV indicates the expected efficiency of the two-stage case identifying procedure and suggests the number of respondents with a high score who will be found to be a case at subsequent interview. NPV indicates the extent to which the test succeeds in predicting the probability of a low score being a non-case. Together PPV and NPV are the most practically useful values to the researcher for they indicate the extent to which the test succeeds in predicting psychiatric cases in a given population. Predictive values vary with the prevalence of psychiatric morbidity. The higher the prevalence the greater the positive predictive value (Sen 1987). PPV and NPV also vary with changes in the threshold of the test. This means that in population surveys, the threshold should be set fairly low to maximize the proportion of test positives.

Sensitivity and Specificity

Sensitivity and specificity indicate the test's ability to discriminate between those with potential disorder and those who are potentially disorder-free (McDowell & Newell, 1987). Sensitivity refers to the test's ability to correctly identify true cases in the sample, while specificity refers to the test's ability to correctly identify those persons who are true negatives. Together, with the proportion of false positives, these measures represent the overall misclassification rate, or cost of the test. Sensitivity and specificity are considered the most appropriate statistics in estimating the validity of a test, as they are regarded as independent of base rate. In other words, they remain the same regardless of prevalence rate. They are regarded as inherent characteristics of the test, which do not change when the test is applied to different populations (Goldberg & Williams, 1988; Riegleman, 1981). However, invariance of sensitivity and specificity has not been found across all samples (Anthony et al., 1985; Helzer et al., 1985; Robins, 1985). Robins (1985, p. 5) observes that "... neither sensitivity nor specificity are independent of the base rate of signs and symptoms of the disorder in the population, or the strictness of the diagnostic definition, or the study's time frame".

If the study's goal is the best estimate of population prevalence, and specificity is good, it is better to have **moderate** rather than high sensitivity (Robins, 1985). This applies in situations where prevalence is expected to be low. In such cases, a high sensitivity will yield only a few false negatives, which will be insufficient to cancel out the false positives. This could result in a falsely high prevalence estimate.

In general, sensitivity is expected to be slightly lower in community samples than in clinic samples (Chan & Chan, 1983). In community studies, it is probable that the ratio of high to low scorers will depart from equality, to a large degree, and this results in reduced sensitivity

(Benjamin et al., 1982; Tarnopolsky, 1979). Specificity, however, does not appear to be significantly affected.

INSTRUMENTS

APPENDIX 2.1 ENGLISH VERSION OF THE GHQ-28

APPENDIX 2.2 ENGLISH VERSION OF THE SRQ

APPENDIX 2.3 FIRST-STAGE QUESTIONNAIRE (including Afrikaans version of the GHQ-28 and SRQ)[1]

APPENDIX 2.4 LAY-INTERVIEWERS' MONITORING FORM

APPENDIX 2.5 CONFIDENTIALITY FORM

APPENDIX 2.6 ENGLISH/AFRIKAANS VERSION OF PAS QUESTIONS NOT INCLUDED IN THE PSE, BUT CONTAINED IN THE COMBINED PSE/PAS CLINICAL INTERVIEW

APPENDIX 2.7 SCORING SHEET FOR THE PSE, PAS AND "SUGGESTED CLINICAL DIAGNOSIS"

APPENDIX 2.8 KEY TO PAS (MODIFIED)

---

1 SRQ and GHQ-28 were printed on blue and yellow paper respectively in the main study for ease of access.

APPENDIX 2.1

ENGLISH VERSION OF THE GENERAL HEALTH QUESTIONNAIRE-28

APPENDIX 2.1 ENGLISH VERSION OF THE GENERAL HEALTH QUESTIONNAIRE-28

We should like to know if you have had any medical complaints; and how your health has been in general, *over the past few weeks*. Please answer ALL the questions on the following pages simply by underlining the answer which you think most nearly applies to you. Remember that we want to know about present and recent complaints, not those that you had in the past.
It is important that you try to answer ALL the questions.
Thank you very much for your cooperation.

HAVE YOU RECENTLY:

A1.	Been feeling perfectly well and in good health?	Better than usual	Same as usual	Worse than usual	Much worse than usual	
A2.	Been feeling in need of a good tonic?	Not at all	No more than usual	Rather more than usual	Much more than usual	
A3.	Been feeling run down and out of sorts?	Not at all	No more than usual	Rather more than usual	Much more than usual	
A4.	Felt that you are ill?	Not at all	No more than usual	Rather more than usual	Much more than usual	
A5.	Been getting any pains in your head?	Not at all	No more than usual	Rather more than usual	Much more than usual	
A6.	Been getting a feeling of tightness or pressure in your head?	Not at all	No more than usual	Rather more than usual	Much more than usual	
A7.	Been having hot or cold spells?	Not at all	No more than usual	Rather more than usual	Much more than usual	
B1.	Lost much sleep over worry?	Not at all	No more than usual	Rather more than usual	Much more than usual	
B2.	Had difficulty in staying asleep once you are off?	Not at all	No more than usual	Rather more than usual	Much more than usual	
B3.	Felt constantly under strain?	Not at all	No more than usual	Rather more than usual	Much more than usual	
B4.	Been getting edgy and bad-tempered?	Not at all	No more than usual	Rather more than usual	Much more than usual	
B5.	Been getting scared or panicky for no good reason?	Not at all	No more than usual	Rather more than usual	Much more than usual	
B6.	Found everything getting on top of you?	Not at all	No more than usual	Rather more than usual	Much more than usual	
B7.	Been feeling nervous and strung-up all the time?	Not at all	No more than usual	Rather more than usual	Much more than usual	
C1.	Been managing to keep yourself busy and occupied?	More so than usual	Same as usual	Rather less than usual	Much less than usual	
C2.	Been taking longer over the things you do?	Quicker than usual	Same as usual	Longer than usual	Much longer than usual	
C3.	Felt on the whole you were doing things well?	Better than usual	About the same	Less well than usual	Much less well	
C4.	Been satisfied with the way you've carried out your task?	More satisfied	About same as usual	Less satisfied than usual	Much less satisfied	
C5.	Felt that you are playing a useful part in things?	More so than usual	Same as usual	Less useful than usual	Much less useful	
C6.	Felt capable of making decisions about things?	More so than usual	Same as usual	Less so than usual	Much less capable	
C7.	Been able to enjoy your normal day-to-day activities?	More so than usual	Same as usual	Less so than usual	Much less than usual	
D1.	Been thinking of yourself as a worthless person?	Not at all	No more than usual	Rather more than usual	Much more than usual	
D2.	Felt that life is entirely hopeless?	Not at all	No more than usual	Rather more than usual	Much more than usual	
D3.	Felt that life is not worth living?	Not at all	No more than usual	Rather more than usual	Much more than usual	
D4.	Thought of the possibility that you might make away with yourself?	Definitely not	I don't think so	Has crossed my mind	Definitely have	
D5.	Found at times you couldn't do anything because your nerves were too bad?	Not at all	No more than usual	Rather more than usual	Much more than usual	
D6.	Found yourself wishing you were dead and away from it all?	Not at all	No more than usual	Rather more than usual	Much more than usual	
D7.	Found that the idea of taking your own life kept coming into your mind?	Definitely not	I don't think so	Has crossed my mind	Definitely has	

Goldberg, D., & Hillier, V. (1979)

APPENDIX 2.2

ENGLISH VERSION OF THE SELF REPORTING QUESTIONNAIRE

APPENDIX 2.2 ENGLISH VERSION OF THE SELF REPORTING QUESTIONNAIRE

1.	Do you often have headaches?	Yes	No
2.	Is your appetite poor?	Yes	No
3.	Do you sleep badly?	Yes	No
4.	Are you easily frightened?	Yes	No
5.	Do your hands shake?	Yes	No
6.	Do you feel nervous, tense or worried?	Yes	No
7.	Is your digestion poor?	Yes	No
8.	Do you have trouble thinking clearly?	Yes	No
9.	Do you feel unhappy?	Yes	No
10.	Do you cry more than usual?	Yes	No
11.	Do you find it difficult to enjoy your daily activities?	Yes	No
12.	Do you find it difficult to make decisions?	Yes	No
13.	Is your daily work suffering?	Yes	No
14.	Are you unable to play a useful partin life?	Yes	No
15.	Have you lost interest in things?	Yes	No
16.	Do you feel that you are a worthless person?	Yes	No
17.	Has the thought of ending your life been in your mind?	Yes	No
18.	Do you feel tired all the time?	Yes	No
19.	Do you have uncomfortable feelings in your stomach?	Yes	No
20.	Are you easily tired?	Yes	No
21.	Do your feel that somebody has been trying to harm you in some way?	Yes	No
22.	Are you a much more important person than most people think?	Yes	No
23.	Have you noticed any interference or anything else unusual with your thinking?	Yes	No
24.	Do you ever hear voices without knowing where they come from or which other people cannot hear?	Yes	No
25.	Have you ever had any fits, convulsions or falls to the ground, with movements of arms and legs, biting of the tongue or loss of consciousness?	Yes	No

(Harding, T., De Arango, M., Baltazar, J., Climent, C., Ibrahim, H., Ladrido-Ignacio, L., Murthy, R., & Wig, N. (1980)

APPENDIX 2.3

FIRST-STAGE QUESTIONNAIRE

INCLUDING THE AFRIKAANS VERSIONS OF THE SRQ AND GHQ-28[1] AND ADDITIONAL QUESTIONS ON SOCIO-DEMOGRAPHIC VARIABLES, SUBSTANCE ABUSE AND HEALTH CARE UTILISATION

---

1 SRQ and GHQ-28 were printed on blue and yellow paper in the study for ease of access.

APPENDIX 2.3 FIRST-STAGE QUESTIONNAIRE

Kode van Respondent: | | | |

Universiteit van Kaapstad

Algemene Gesondheidstudie Volwasse Bevolking, Mamre, 1992

1. Naam van onderhoudvoerder: | | |

_____

2. Kode van Ondervraer:

3. Datum van onderhoud:

D	D	M	M	Y	Y

4. Tyd van onderhoud: _____

5. My naam is _____.
Ek voer onderhoude vir die Universiteit van Kaapstad wat graag die algemene gesondheidStoestand van die volwasse inwoners van Mamre wil bestudeer. Die inligting verkry uit hierdie studie sal gebruik word om die gesondheidsbehoeftes van ons gemeenskap vas te stel. Waar moontlik, sal programme geloods word om hierdie behoeftes aan te spreek. Ek wil dit graag duidelik stel dat alles wat u in hierdie onderhoud se, vertroulik is. u vraelys sal direk aan die Universiteit se navorsingspan oorhandig word, wat dit sal ontleed. Ons stel nie belang in individuele antwoorde nie. Al die vraelyste sal saam gebruik word om die gesondheidstoestand en die behoeftes van die gemeenskap van Mamre, vas te stel. U antwoorde sal vertroulik bly, en sal slegs vir akademiese doeleindes gebruik word.

Die onderhoud sal omtrent 15 minute duur. As dit nie nou vir u gerieflik is om die vrae te beantwoord nie, kan ons 'n afspraak maak vir 'n meer geriefliker tyd binne die volgende paar dae.

6. Volgende afspraak:

datum _____

tyd _____

7. Is u bereid om aan die studie deel te neem?

JA	NEE

8. Handtekening van respondent

_____

* * * *

9. Indien u NEE antwoord, gee asseblief die rede.

_____

_____

_____

10. Manlik / Vroulik

M	V

11. Ouderdom:

12. Besoek aan huishouding:

Datum _____

1. | | | | | | |
2. | | | | | | |
3. | | | | | | |

Redes waarom u onsuksesvol was.

1. _____
2. _____
3. _____

Eerstens dankie vir u tyd.

Is daar 'n stil plek waar ons die onderhoud kan voer?.

JA	NEE

Hier is my geskrewe belofte van vertroulikheid:

As u ten enige tyd gedurende die onderhoud nie seker is wat gevra word nie, kan u vir my raadpleeg.

A. Algemene Inligting

Ek wil graag begin deur u 'n paar besonderhede te vra aangaande uself.

Kode van Respondent: | | | |

1. Wat is u volle naam?

_____

2. Wat is u straat adres?

_____

_____

3. Wat is u telefoonnommer?

_____

4. **Kan u my die naam, telefoonnommer, of adres van 'n buurman/buurvrou of vriend gee, waar ek 'n boodskap vir u kan los as u nie tuis is nie?**

_____

_____

5. Manlik / Vroulik

1. Manlik — 1.

2. Vroulik — 2.

6. Wanneer was u gebore?

D	D	M	M	Y	Y

7. Is u:

1. ongetroud
2. getroud
3. ongetroud, maar lewe saam met iemand
4. getroud, maar lewe apart van u man/vrou
5. geskei
6. 'n weduwee of wewenaar

8. Hoeveel biologiese (eie) kinders het u gehad, so wel as aangenome kinders, en pleegkinders?

_____

9. Werk u vir geld of aalmoese, b.v. klere of kos ?

JA	NEE

(i) Indien NEE

Is u

1. 'n tuisteskepper
2. student
3. ongeskik (kan nie werk nie)
4. werkloos (kan werk)
5. pensioentrekker
6. ander (spesifiseer)

_____

Soek u vir werk?

JA	NEE

Indien JA;
Hoe lank is u werkloos?

_____ _____ _____
jaar maande weke

(ii) Indien JA;

Watter soort werk doen u tans?

_____

_____

Waar werk u?
In 'n:

1. fabriek	1
2. winkel	2
3. skool	3
4. hospitaal	4
5. huis	5
6. groente tuin	6
7. raad (council)	7
ander, spesifiseer	8

_____

10. Kry u 'n toelaag?

JA	NEE

Indien JA:
Watter tipe toelaag kry u?

1. 'n ouderdomspensioen	1
2.n ongeskiktheidstoelaag	2
3. 'n oorlogspensioen	3
4. kindersorgtoelaag	4
5. 'n spesiale sorgtoelaag	5
6. ander, spesifiseer	6

_____

11. Wat is die hoogste standerd wat u op skool geslaag het?

1. Het glad nie skool gegaan nie

_____

2. Die hoogste standerd

_____

12. Het u enige kursusse geloop of opleiding ondergaan sedert u die skool verlaat het b.v. NTS, verpleging tik?

JA	NEE

Indien JA,

Spesifiseer, _____

13. Het u aan enige akademiese instituut studeer sedert u die skool verlaat het?

Indien JA, waar _____

Wat was u hoogste kwalifikasie?

_____

14. Hoeveel kamers is daar in u huis, uitsluitend die kombuis, badkamer, toilet en stoorkamer?

_____

15. Wat is die voorletters, ouderdomme, en verwantskap van almal wat van Maandag tot Vrydag in die huis slaap?

Voorletters/Ouderdom/Verwantskap/

1 _____

2 _____

3 _____

4 _____

5 _____

6 _____

7 _____

8 _____

9 _____

16. Is u 'n lid van enige kerk of sosiale groep b.v. rugbyklub, kaartspel, koor?

JA	NEE

Spesifiseer _____

_____

17. Op wie kan u regtig staatmaak om u uit 'n krisissituasie te help?

	Voorletters	Verwantskap
1		
2		
3		
4		
5		
6		
7		
8		
9		

kode: _____

B. Algemene Gesondheid

Die volgende vrae is dieselfde soos die wat u nounet beantwoord het maar hierdie keer hoef u net JA of NEE te antwoord.

1.	Kry u gereeld kopseer?	JA	NEE	
2.	Is u eetlus swak?	JA	NEE	
3.	Slaap u sleg?	JA	NEE	
4.	Word u maklik bang?	JA	NEE	
5.	Bewe u hande?	JA	NEE	
6.	Voel u senuweeagtig, gespanne of bekommerd?	JA	NEE	
7.	Is u spysvertering swak?	JA	NEE	
8.	Het u 'n probleem om helder te dink?	JA	NEE	
9.	Voel u ongelukkig?	JA	NEE	
10.	Huil u meer as gewoonlik?	JA	NEE	
11.	Vind u dit moeilik om u daaglikse take te geniet?	JA	NEE	
12.	Vind u dit moeilik om besluite te neem?	JA	NEE	
13.	Word u daaglikse werk geaffekteer?	JA	NEE	
14.	Is u in staat om 'n waardevolle rol in die lewe te speel ?	JA	NEE	
15.	Het u belangstelling in dinge verloor?	JA	NEE	
16.	Voel u dat u 'n waardelose persoon is?	JA	NEE	
17.	Het die gedagte/idee om u eie lewe te neem al by u opgekom?	JA	NEE	
18.	Voel u gedurig moeg?	JA	NEE	

19.	Het u ongemaklike gevoelens in u maag?	JA / NEE	
20.	Raak u gou moeg?	JA / NEE	
21.	Voel u dat iemand op een of ander manier u kwaad wil aandoen?	JA / NEE	
22.	Is u 'n veel meer belangrike persoon as wat meeste mense dink?	JA / NEE	
23.	Het u enige versteuring of enigiets ongewoons omtrent u gedagtes opgemerk?	JA / NEE	
24.	Hoor u ooit stemme sonder om te weet waar hulle vandaan kom of wat ander mense nie kan hoor nie?	JA / NEE	
25.	Het u al ooit enige stuiptrekkings gehad of op die grond neergeval, met gepaardgaande beweging van die arm en bene en byt van die tong of bewusteloosheid?	JA / NEE	

kode: _____

B. Algemene Gesondheidstoestand

In die tweede deel van hierdie vraelys, wil ons graag weet of u enige mediese klagtes gehad het, en hoe u gesondheid in die algemeen, oor die laaste paar weke gewees het. Beantwoord asseblief al die vrae deur slegs die antwoord te kies wat u dink naastenby op u van toepassing is. Onthou dat ons van gesondheidsprobleme wat u op die tydstip ervaar wil hoor en nie die wat u die afgelope paar jaar ervaar het nie.

Dit is belangrik dat u moet probeer om al die vrae te beantwoord.

Het u onlangs:

A1 Heeltemal goed gevoel en in goeie gesondheid verkeer?

Beter as gewoonlik,
Dieselfde as gewoonlik,
Slegter as gewoonlik,
Baie slegter as gewoonlik.

A2 Gevoel dat u 'n tonikum nodig het?

Glad nie,
Nie meer as gewoonlik nie,
Meer as gewoonlik,
Baie meer as gewoonlik.

A3 Aanhoudend afgemat en moeg gevoel?

Glad nie
Nie meer as gewoonlik nie,
Meer as gewoonlik,
Baie meer as gewoonlik.

A4 Gevoel dat u siek is?

Glad nie,
Nie meer as gewoonlik nie,
Meer as gewoonlik,
Baie meer as gewoonlik

A5 Enige pyne in u kop gekry?

Glad nie,
Nie meer as gewoonlik nie,
Meer as gewoonlik,
Baie meer as gewoonlik.

A6 'n Gevoel van styfheid of druk in u kop gekry?

Glad nie,
Nie meer as gewoonlik nie,
Meer as gewoonlik,
Baie meer as gewoonlik.

A7 Warm of koue aanvalle gekry?

Glad nie,
Nie meer as gewoonlik nie,
Meer as gewoonlik,
Baie meer as gewoonlik.

B1 Baie slaap verloor as gevolg van bekommernis?

Glad nie,
Nie meer as gewoonlik nie,
Meer as gewoonlik,
Baie meer as gewoonlik.

B2 Moeilikheid ervaar om aan die slaap te bly as u eers begin slaap het?

Glad nie,
Nie meer as gewoonlik nie,
Meer as gewoonlik,
Baie meer as gewoonlik.

B3 Al die tyd onder druk gevoel?

Glad nie,
Nie meer as gewoonlik nie,
Meer as gewoonlik,
Baie meer as gewoonlik.

B4 Humeurig begin voel?

Glad nie,
Nie meer as gewoonlik nie,
Meer as gewoonlik,
Baie meer as gewoonlik.

B5 Begin bang en paniekerig word vir geen rede?

Glad nie,
Nie meer as gewoonlik nie,
Meer as gewoonlik,
Baie meer as gewoonlik.

B6 Gevoel asof alles te veel word vir u?

Glad nie,
Nie meer as gewoonlik nie,
Meer as gewoonlik,
Baie meer as gewoonlik.

B7 Die hele tyd senuweeagtig gevoel?

Glad nie,
Nie meer as gewoonlik nie,
Meer as gewoonlik,
Baie meer as gewoonlik.

C1 Daarin geslaag om u besig te hou?

Baie meer as gewoonlik,
Dieselfde as gewoonlik ,
Nie meer as gewoonlik nie,
Glad nie.

C2 Gevoel dat u langer neem om dinge te doen?

Gouer as gewoonlik,
Dieselfde as gewoonlik,
Langer as gewoonlik,
Baie langer as gewoonlik.

C3 Oor die algemeen gevoel dat u dinge goed doen?

Beter as gewoonlik,
Dieselfde as gewoonlik ,
Nie so goed soos gewoonlik,
Baie slegter as gewoonlik.

C4 Tevrede gevoel met die manier waarop u take verrig?

Meer as gewoonlik,
Dieselfde soos gewoonlik,
Minder as gewoonlik,
Baie minder as gewoonlik.

C5 Gevoel dat u 'n betekenisvolle rol speel in dinge wat u doen?

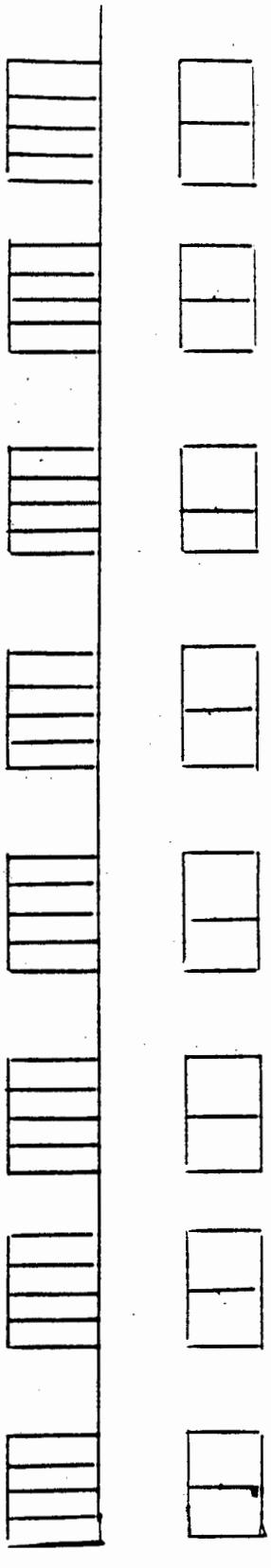
Baie meer as gewoonlik,
Dieselfde as gewoonlik,
Minder as gewoonlik,
Baie minder as gewoonlik.

C6 Bekwaam gevoel om besluite te neem?

Baie meer as gewoonlik,
Dieselfde as gewoonlik ,
Minder as gewoonlik,
Baie minder as gewoonlik.

C7 In staat gevoel om u alledaagse aktiwiteite te geniet?

Glad nie,
Nie meer as gewoonlik nie,
Meer as gewoonlik,
Baie meer as gewoonlik



D1 Van uself gedink as 'n nuttelose persoon?

Glad nie,
Nie meer as gewoonlik nie,
Meer as gewoonlik,
Baie meer as gewoonlik.

D2 Gevoel dat die lewe heeltemal hopeloos is?

Glad nie,
Nie meer as gewoonlik nie,
Meer as gewoonlik,
Baie meer as gewoonlik.

D3 Gevoel dat dit nie werd is om te lewe nie?

Glad nie,
Nie meer as gewoonlik nie,
Meer as gewoonlik,
Baie meer as gewoonlik.

D4 Gedink dat u miskien nie meer wil lewe nie?

Glad nie,
Ek dink nie so nie,
Die gedagte het al by my opgekom,
Ek het beslis daaraan gedink.

D5 Gevind dat u niks kon doen nie omdat u senuwees so sleg gevoel het?

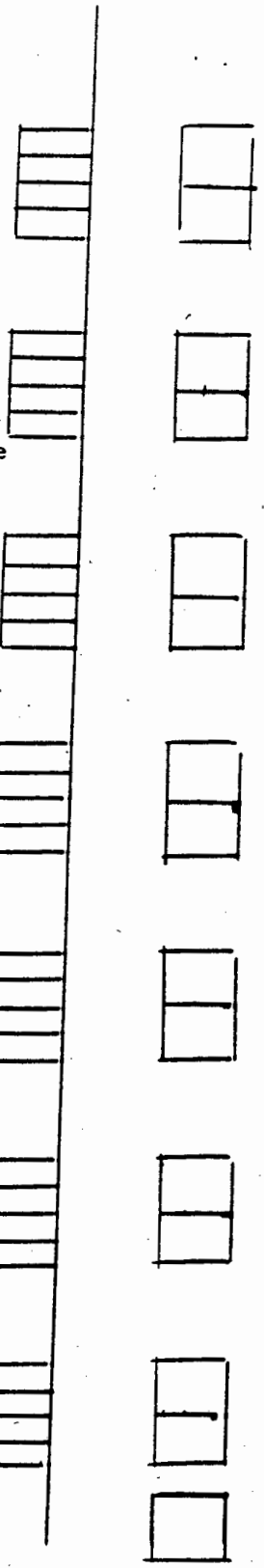
Glad nie,
Nie meer as gewoonlik nie,
Meer as gewoonlik,
Baie meer as gewoonlik.

D6 Gevoel dat u wens u was dood en weg was van alles?

Glad nie,
Nie meer as gewoonlik nie,
Meer as gewoonlik,
Baie meer as gewoonlik.

D7 Gevoel dat die idee om u eie lewe te neem aanhoudend in u gedagte was?

Glad nie,
Ek dink nie so nie,
Die gedagte het al by my opgekom,
Ek het beslis daaraan gedink.



C. Gesondheidsorgdienste

Die volgende paar vrae het betrekking op die beskikbaarheid van die huidige gesondheidsorgdienste en hoe u voel dit in u behoeftes voorsien.

1 Is u nou siek? | JA | NEE | ☐

Indien JA:
Wat is u probleem?

_____

Indien NEE?
Was u siek in die

1 laaste twee dae? | 1 | ☐
2 laaste week? | 2
3 laaste twee weke? | 3
4 laaste maand? | 4

Wat was die probleme?(5)

_____

(Indien NEE, aan al die bogenoende, gaan na afdeling D).

Indien JA aan enige een van die bogenoende dan:

2. Is/was u in staat om u gewone dagtake uit te voer? | JA | NEE | ☐

3. Het u uself behandel vir hierdie siekte? | JA | NEE | ☐

Indien JA, spesifiseer

_____

Deur wie is u behandel?

1. 'n familielid | 1
2. 'n buurman/buurvrou | 2
3. 'n kruiedoktor | 3
4. 'n apteker | 4
5 'n distriksverpleegster | 5

6'n hospitaalverpleegster	6
7. 'n privaatdokter	7
8. 'n dokter by 'n daghospitaal of gewone hospitaal	8
9. 'n distriksgeneesheer	9
10. 'n dokter/verpleegster by die werksplek	10
11. 'n priester	11
12. 'n selfhulp groep	12
13. Ander, (spesifiseer)	13

_____

4. Indien u nie behandel was nie hoekom nie?

1. Siekte of besering te gering	1
2. 'n Ou kwaal	2
3. Te veel onkoste	3
4. Te besig	4
5. Geen diens beskikbaar op daardie tydstip nie	5
6. Geen vervoer nie	6
7. Ander, (spesifiseer)	7

_____

_____

5. Indien u deur 'n dokter, kliniek of hospitaal behandel is, was u medies ondersoek?

JA	NEE

6. Was daar enigiets wat u anders sou verkies met hierdie besoek aan die dokter, kliniek of hospitaal en is daar iets wat u sou byvoeg?

JA	NEE

Spesifiseer: _____

_____

7. Het 'n mediese fonds vir u mediese sorg betaal?

JA	NEE

D. Vloeistofinname

Laastens wil ons graag weet wat die verwantskap tussen u gesondheid en u vloeistofinname is. Die vrae is gerig om inligting in te win ten opsigte van die gemeenskap se gesondheids-behoeftes en nie om individuele gedrag te veroordeel nie.

1. Hoeveel glase water drink u per dag?

Geen	

2. Hoeveel koppies tee drink u per dag?

Geen	

3. Hoeveel koppies koffie drink u per dag?

Geen	

4. Hoeveel glase of bottels bier, wyn of sterk drank drink u per dag ?

Geen	

5. Hoeveel glase, of bottels bier, wyn of sterk drank drink u oor naweke?

Geen	

(Indien 2 of minder glase, gaan na vraag 9).

6. Het u al ooit gevoel dat u minder alkohol behoort te drink?

JA	NEE

7. Het mense u al aanstoot gegee wanneer hulle u drinkery kritiseer?

JA	NEE

8. Het u al ooit soggens net na u opgestaan het 'n drankie geneem om u senuwees te kalmeer of om van 'n babelaas ontslae te raak?

JA	NEE

9. Hoeveel sigarette rook u per dag?

Geen	

10. Is daar iets anders bv. tablette, of ander medikasie wat u gebruik om u meer ontspanne te laat voel?

JA	NEE

Spesifiseer _____

_____

E. Slotsom

Hierdie vraelys is deel van 'n groter studie wat vereis dat 'n deel van die mense wat die eerste vraelys voltooi, deelneem aan 'n opvolgonderhoud. Die tweede onderhoud is 'n kliniese onderhoud van omtrent 30 minute en sal by die kliniek in Johannesstraat, plaasvind, op Saterdae.

Deelnemers sal R20,00 betaal word.

Sal u bereid wees om deel te neem aan die tweede helfte van die studie?

JA	NEE

Watter tyd sal u die beste pas?

Sat v.m. ☐

Sat n.m. ☐

Indien u NEE antwoord gee asseblief die redes waarom u nie aan die tweede onderhoud wil deelneem nie.

Baie dankie vir die tyd wat u aan my kon afstaan.

APPENDIX 2.4

LAY-INTERVIEWERS' MONITORING FORM

APPENDIX 2.4 MONITORING FORM FOR LAY-INTERVIEWERS

Interviewer's code: _____
Date of interview: _____
Time of interview: _____
Respondent's name: _____

1. Did the interviewer introduce him or herself? _____ yes/no

2. Did the interviewer give the name of the institution which is conducting the study? _____ yes/no

3. Did the interviewer state the aim of the study? _____ yes/no

4. Was the form of confidentiality signed, dated and left with the respondent? _____ yes/no

5. Were all questions read exactly as worded in the questionnaire? _____ yes/no

6. Were all questions read out in the correct order? _____ yes/no

7. Were all answers recorded exactly as given? _____ yes/no

8. Did the interviewer influence the way in which the respondent answered any of the questions? _____ yes/no

   In what way did the interviewer influence the respondent?
   _____
   _____

9. If the respondent did not respond in the appropriate response form what did the interviewer do? _____
   _____
   _____

10. Was the questionnaire completed? _____ yes/no

11. Did the interviewer read slowly and clearly? _____ yes/no

Rating of Interview

1	poor
2	fair
3	good
4	very good
5	excellent

What were the interviewer's good points in the interview?

_____

_____

What were the interviewer's weak points in the interview?

_____

_____

Signed _____

Date._____

APPENDIX 2.5

CONFIDENTIALITY FORM

APPENDIX 2.5 CONFIDENTIALITY FORM

UNIVERSITIET VAN KAAPSTAD

Algemene Gesondheidstudie, Volwasse Bevolking,
Mamre, 1992

Belofte van Vertroulikheid

Ek, _____ belowe dat enigiets

wat gedurend die onderhoud vir my gese word, vertroulik sal wees en met niemand, behalwe die navorsingspan van die studie, bespeek sal word nie.

Geteken: _____
ONDERHOUDER

Datum: _____

APPENDIX 2.6

ENGLISH AND AFRIKAANS VERSIONS OF QUESTIONS ON THE PAS NOT INCLUDED IN THE PSE, BUT CONTAINED IN THE COMBINED PSE/PAS CLINICAL INTERVIEW

APPENDIX 2.6 ENGLISH AND AFRIKAANS VERSIONS OF QUESTIONS IN THE PAS NOT FOUND IN THE PSE (9TH ED.), BUT INCLUDED IN THE COMBINED PSE/PAS CLINICAL INTERVIEW

Both English and Afrikaans versions of the PSE (8th edition) are available in South Africa. Only those PAS questions not found in the PSE, but included in this study in the combined PSE/PAS clinical interview, are given below. Questions are numbered in the order in which they occur in the combined PSE/PAS interview. Some items pertain to additional methods of rating a particular symptom identified in the PSE, and are not additional question per se.

In reference to **PANIC DISORDERS** (Question 14 section 3 on the PSE)

24. RATE SECONDARY FEAR OF DYING OR LOSING CONTROL
    0 = Not present
    1 = Definitely present

25. How rapidly to these symptoms appear after the attack begins?
    Hou gou na die paniek aanalle begin heirdie simptome?
    RATE SPEED AT WHICH SYMPTOMS DEVELOP-if during at least some of the attacks at least 4 of the symptoms given in the check list of PSE question 14 developed suddenly and increased in intensity within 10 minutes of the beginning of the first symptom rate as "1"
    If this is not the case rate as "0"

26. RATE PANIC ATTACKS WITH AUTONOMIC SYMPTOMS: AND LEADING TO ACTION.
    Rate here if panic or intolerable anxiety leads to some action to end it, eg, leaving a bus, phoning husband at work, going to see a neighbour.
    1 = 1 to 4 panic attacks leading to action during month.
    2 = 5 or more

27. Did you have any panic attacks which did not lead to any action?
    Het u enige paniek aanvalle gehad wat nie tot enige aksis aanleiding gegee het?
    How many of both kinds of attacks have you been having?
    How veel van beide hierdie tipe aanvalle het u al gehad?
    RATE FREQUENCY OF PANIC ATTACKS
    N.B. maximum frequency at any time during the episode of illness not just in last 4 weeks.
    0 = less than 4 in 4 week period.
    1 = 4 or more attacks in a 4 week period.

28. Were you nervous much of the time between attacks?
    Was u angstig meeste van die tyd tussen aanvalle?
    RATE ANXIETY BETWEEN ATTACKS
    0 = Anxiety present between attacks to a marked degree.
    1 = Comparative freedom from anxiety between attacks.

29. If you only had 1-2 attacks did you find that you were inconstant fear of it happening again?
    As u net 1 of 2 aanvalle gehad het, vind u dat u gedurig bang is dat die weer sal gebeur?

RATE FEAR OF HAVING A PANIC ATTACK (present at any time since onset)
0 = No fear experienced.
1 = Fear experienced but not persistent or for less than a period of one month.
2 = Fear of a panic attack occurring present persistently for at least one month.

30. What have the attacks been like in the last month?
Hou was die aanvalle gedurende die afgelope maand?
RATE CURRENT SEVERITY OF PANIC ATTACKS
0 = Mild. During the past month either all attacks have been fewer than 4 symptoms or there has been no more than one panic attack
1 = Moderate. During the past month attacks have been intermediate between "mild" and "severe".
2 = Severe. During the past month there have been at least 8 attacks.

In reference to **SITUATIONAL ANXIETY** (questions 15 and 16 section 3 on the PSE)

32. How many of the above situations do you experience anxiety in? (or would if exposed to them).
In hoeveel van bogenoemde situasies ondevind u angs? (Of sal u angstig word as u daaraan biootgestel word?)
RATE NUMBER OF SITUATIONS WHICH TRIGGER ANXIETY
0 = None or 1
1 = 2 or more

34. What do you think about the way that you react in these situations?
Wat dink u van die mannier how u reageer in hierdie situasies?
RATE ATTITUDE TO SOCIAL PHOBIA
If recognises fear to be excessive and/or unreasonable rate as "1". If does not rate as "0".

In reference to **SPECIFIC PHOBIAS** (question 18 section 3 on the PSE)

37. RATE CURRENT SEVERITY OF AGORAPHOBIC AVOIDANCE (if box 36=1 or 2)
0 = Mild. Some avoidance (or endurance with distress) but relatively normal life-style.
1 = Moderate. Avoidance results in constricted life-style, eg, the person is able to leave house alone, but cannot go far unaccompanied.
2 = Severe. Avoidance results in being nearly or completely house bound or unable to leave the house unaccompanied.

DEPRESSED MOOD section 5 on PSE

45. When a person gets depressed he may think about dying or suicide. Have you?
Wanneer 'n persoon depressief (teneergedruk) raak, kan hy dink aan selfmoord. Het u?
RATE THOUGHTS ABOUT DEATH
1 = Frequent thoughts about death (would be better off dead) or thoughts of suicide without plans

APPETITE, SLEEP RETARDATION LIBIOR section 7 on PSE

58. Have you had an increase in appetite?
Het u 'n toename in eetlus gehad?

RATE INCREASE IN APPETITE
0 = No increase or slight increase
1 = Mild to moderate increase
2 = Hungry all the time

59. Have you gained weight over the last three months?
Het u gewig opgetel gedurende die afgelope 3 maande?
RATE WEIGHT GAIN
0 = No weight gain or only regained lost weight.
1 = Doubtful or up to 5lbs.
2 = 5lb.

60. Have you had trouble sleeping?
Het u probleme gehad om to slaap?
RATE SLEEP DIFFICULTY
0 = No difficulty or occasional difficulty.
1 = Mild to moderate difficulty -often or usually has significant difficulty.
2 = Severe, almost always has great difficulty.

61. Are you sleeping longer or more than usual?
Slaap u langer of meer as gewoonlik?
RATE HYPERSOMNIA
0 = Normal sleep or occasionally sleeps more than usual.
1 = Frequently sleeps at least 1 hour more than usual.
2 = Frequently sleeps 2-4 hours more than usual:

MAJOR DEPRESSIVE EPISODE (section 9 on the PSE)

74. You've told me about a number of symptoms you have been experiencing, can you tell me which of the following you have experienced almost everyday for at least two weeks?
    1. Depressed mood (43)
    2. Loss of interest of pleasure (including loss of libido) (41, 42, 65)
    3. Weight loss, weight gain, or marked increase or decrease in appetite (56, 57, 58, 59)
    4. Insomnia or hypersomnia (60, 51, 62, 64)
    5. Fatigue or loss of energy (16, 23)
    6. Worthlessness or excessive inappropriate guilt (51, 55)
    7. Diminished ability to think or concentrate (38, 39)
    8. Recurrent thoughts of death, recurrent suicidal ideation without a specific plan or a suicide attempt or a specific plan for committing suicide (45, 46)
    9. Psychomotor agitation or retardation (observable by others, not merely subjective feelings of restlessness or being slowed down) (69, 70)

    U het my vertel van 'n aantal simptome wat u ondervind het, kan u my se wat van die volgende het u ondervind omtrent elke dag gedurende die afgelope 2 weke?
    1. Teneergedruktheid
    2. Verlies aan belangstelling of plesier (insluitende verlies aan seksdrang)
    3. Gewig verlies, gewig toename, of merkabre toename of afname in eetlus
    4. Slaaploosheid of te veel slaap (hipersomnia)
    5. Moegheid of verlies aan energie
    6. Waardeloosheid of oormatige of nie toepastelike skild gevoel
    7. Afname in vermoe om te dink of konsentreer

8. Herhaaldelik dink aan dood, herhaaldelik selfmoord gedagtes sonder spesifieke plan of selfmoord daad, or 'n spesifieke plan om selfmoord to pleeg
9. Psigomotoriese agitasie of vertraging (waarneembaar deur ander, nie net subjektiewe gevoellens of voorkoms of vertraging)

COUNT NUMBER OF SYMPTOMS PRESENT ALMOST EVERYDAY FOR AT LEAST TWO WEEKS

GENERAL ANXIETY DISORDER (section 9 on the PSE)

75. How many of these symptoms do you often experience when you are felling anxious?

Motor tension

1. Trembling, twitching or feeling shaky
2. Muscle tension, aches or soreness
3. Restlessness
4. Easy fatiguability

Autonomic hyperactivity

5. Shortness of breath or smothering sensations
6. Palpitations
7. Sweating or cold clammy hands
8. Dry mouth
9. Dizziness or light-headedness
10. Nausea, diarrhoea or other abdominal distress
11. Flushes or chills
12. Frequent urination
13. Trouble swallowing or "lump in throat"

Vigilance and scanning

14. Feeling keyed up or on edge
15. Exaggerated startle response
16. Difficulty concentrating or mind going bland
17. Trouble falling or staying asleep
18 Irritability

Hoeveel van hierdie simptome ondervind u baie keer wanneer u angstig voel?

Motoriese spanning

1. Bewerasie, rukkings or bewerige gevoel
2. Spier spanning, pyn of seergevoel
3. Rustelоosheid
4. maklike uitputting

Autonomiese hiperaktiwiteit

5. Kort asem of versmorende gevoel
6. Hartkloppens
7. Sweetagtige of aangeklamde hande
8. Drow mond
9. Dronk in die kop of lighoofdig
10. Naarheid, diarrhoea (maagwerking0 of ander abdominale (buik) ongemak
11. Warm of koue gloede
12. Herhaaldelike urienering
13. Probleme om te sluk of knop in die keel

Waarneming

14. Opgewerk of gespanne
15. Veregerde skrik reaksie

16. Moeilike konsentrasie of gedagtes wat leeg raak
17. Probleme om aan die slaap te raak of mon te bly slaap
18. Irriteerbaarheid

COUNT NUMBER OF SYMPTOMS

76. DURATION OF SYMPTOMS

0 = No illness or less than 1 week
1 = Week or more, but less than 2 weeks
2 = 2 weeks or more, but less than 4 weeks
3 = 4 weeks or more
4 = 3 months or more
5 = 6 months or more

77. When did you last feel like your normal self for 2 months or more?
Hoe lank gelede het u soos u normale self gevoel vir 2 maande or meer?
1 = less than 2 years ago
2 = 2 years ago or more than 2 years ago
3 = Always like this

IMPAIRMENT OF FUNCTIONING (section 9 on the PSE)

78. (If depression or anxiety or panic present)
During this time you have been depressed (or panicky or anxious):
did you seek help from someone?
did anyone suggest you seek help?
did you take any medication?
did you act any differently with people, family or at work?
Gedurende die tyd wat u teneergedruk gevoel het (of paniekerig of angstig)
Het u huop van iemand gevra?
Het iemand voorgestel dat u hulp moet soek?
Het u enige medikasie geneem?
Het u anders opgetree teenoor mense, familie of by die werk?
If yes to one of above score 1 in box
If no to all score 0

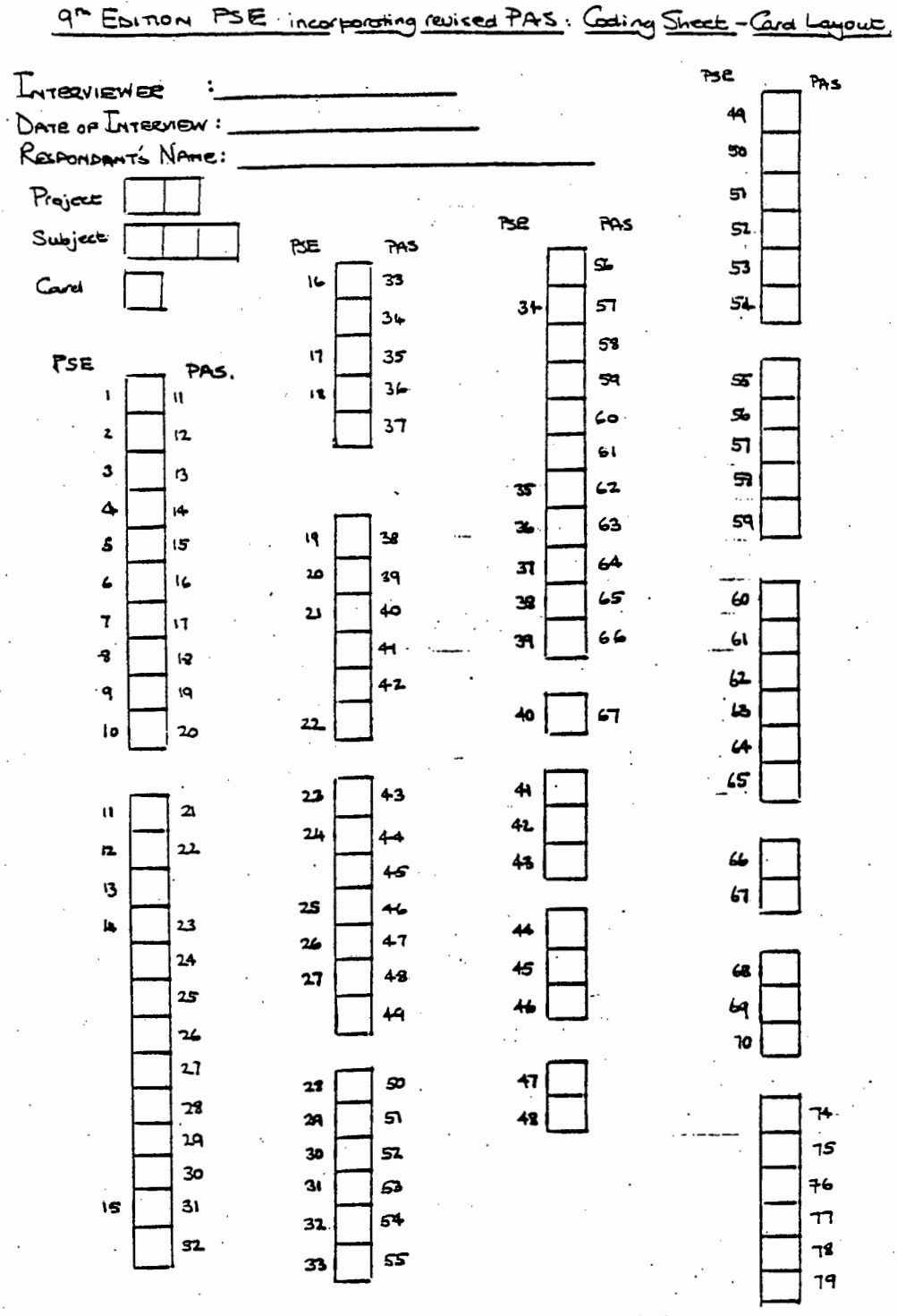
DELUSIONS AND HALLUCINATIONS (section 15 on the PSE)

79. Now I should like to ask you some questions which we ask of everybody.
Has your imagination been playing tricks on you in any way?
Do you ever seem to hear noises or voices when there is no-one about and nothing else to explain it?
Is that true of visions or other unusual experiences that other people have?
Nou wil ek u graag vrae vra wat ons vir aimal vra.
Het you verbeelding snaakse dinge met hou gedoem in enige manier?
Het dit al voorgekom dat u geluide of stemme hoor as daar miemand daar is nie, of niks naders om dit to verduidelik nie?
Is dit waar van visioene of ander buitengewone ondervindings wat ander mense het?
If any reply is YES score as 1

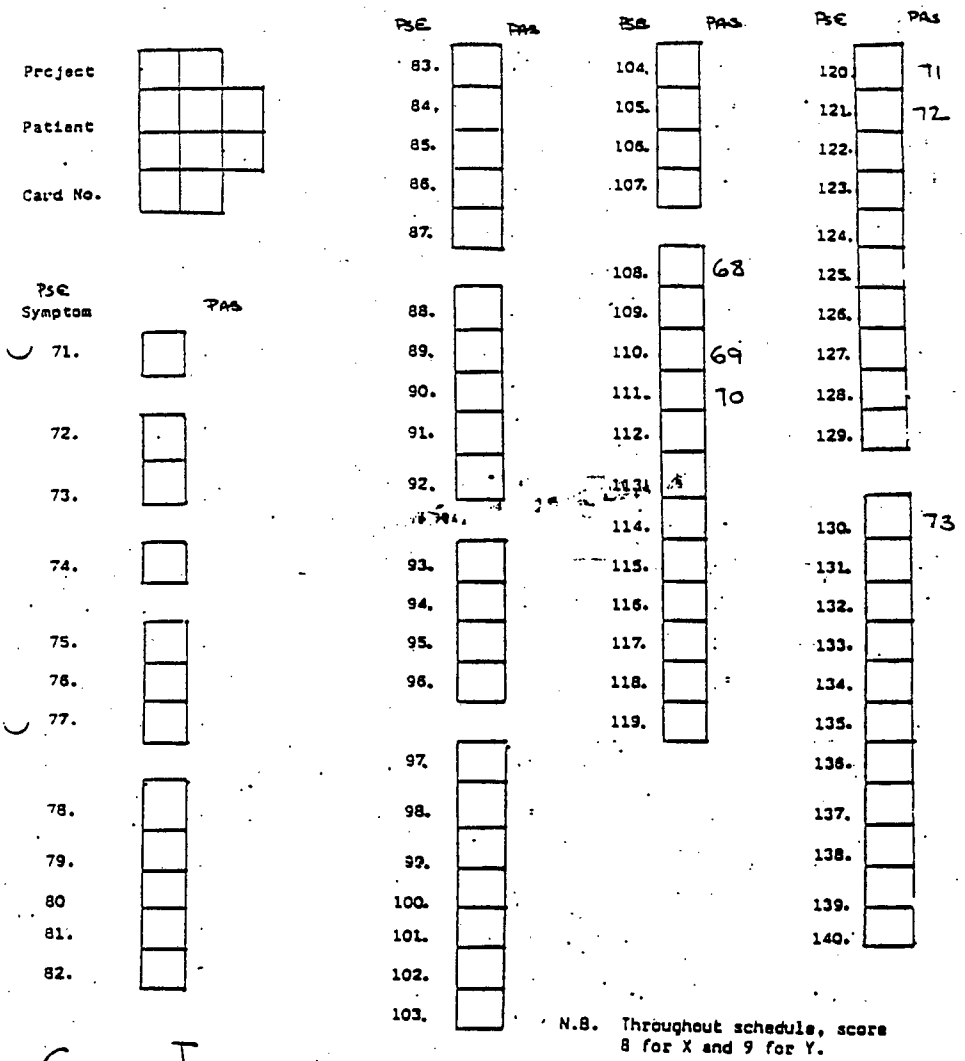
APPENDIX 2.7

SCORING SHEET FOR COMBINED PSE, PAS AND "SUGGESTED CLINICAL DIAGNOSIS"

APPENDIX 2.7 SCORING SHEET FOR COMBINED PSE, PAS AND SUGGESTED CLINICAL DIAGNOSIS



9 th Edition of P.S.E. : Coding sheet - Card Layout



N.B. Throughout schedule, score 8 for X and 9 for Y.

CLINICAL IMPRESSION: _____

_____

_____

SUGGESTED DIAGNOSES (DSM III.): _____

APPENDIX 2.8

KEY TO THE PAS (MODIFIED)

APPENDIX 2.8 KEY TO PAS (MODIFIED)

KEY TO THE PAS (modified)

DSM-IIIR

Major Depressive Episode

5 OR MORE IN BOX 74

Dysthymic Disorder

1 or 2 in box 43 and
2 or more in box 77

+ at least 2 of: 1 or 2 in box 56/57
60/61/62/64
16/63
51
38/39
44

Generalised Anxiety Disorder

1 in box 13 and
5 or more in box 76 and
6 or more in box 75

Panic Disorder

Either 1 in box 27 or
2 in box 29

and 2 in box 23
and 1 in box 25

Panic disorder may be + or - agoraphobia

Specify severity box 30

Agoraphobia

1 or 2 in box 31
Consider whether Panic Disorder is present
Specify severity box 37

Social Phobia

1 or 2 in box 33
1 in box 34

CALCULATIONS

APPENDIX 3.1 CALCULATION OF FIRST- AND SECOND-STAGE SAMPLE SIZE

APPENDIX 3.2 WEIGHTED PREVALENCE OF MENTAL DISORDER REFLECTED FOR SPECIFIED SOCIO-DEMOGRAPHIC GROUPS, PHYSICAL HEALTH STATUS AND RATE OF CONSULTATION

APPENDIX 3.3 CALCULATION OF UNWEIGHTED AND WEIGHTED VALIDITY COEFFICIENTS OF THE SRQ

APPENDIX 3.4 CALCULATION OF WEIGHTED PREVALENCE OF MENTAL DISORDER IN THE ADULT POPULATION OF MAMRE

APPENDIX 3.1

CALCULATION OF FIRST- AND SECOND-STAGE SAMPLE SIZE

APPENDIX 3.1 CALCULATION OF FIRST- AND SECOND-STAGE SAMPLE SIZE

To estimate prevalence of disorder in a general population a two-stage case-identifying method was selected. This method involves the administration of a short and inexpensive first-stage interview followed by a more costly and comprehensive second-stage clinical interview (Shrout, 1992). (The merits of this method of case detection are discussed in Chapters 2, p, 17). Two samples are involved in this design:

(i) the relatively large random sample selected to receive the first-stage questionnaire; and

(ii) the smaller stratified sample selected to receive the second-stage clinical interview.

The statistical equations for the calculation of required sample size, both for the first- and second- stages are identified by Shrout (1992) and are given later in this appendix. In community prevalence studies, where the sensitivity of the screen is unknown, this equation estimates second-stage sample size based on a random selection of a proportion of both screen positives and negatives. Calculations therefore include:

(i) the estimation of size of both first- and second-stage samples, and

(ii) the estimation of the proportion of SRQ high and low scorers required to receive the PSE interview.

In the calculations which follow, the estimation of the proportion of SRQ positives and SRQ negatives required to participate in the second-stage interview is calculated first, as it is on this estimate that the second equation, to estimate first-stage sample size, is based.

Calculation of second-stage sample size based on the estimated proportion of SRQ high and low scorers required to participate in second-stage clinical interview. Estimates based on Shrout's statistical equations (1992)

(i) Equation 1: proportion of SRQ high scorers required to participate in second-stage interview

$$f_1 = \sqrt{\frac{\lambda_1(1-\lambda_1) \qquad C_s}{(\lambda_1-\lambda_2)^2 \pi(1-\pi)C_d}}$$

(ii) Equation 2: Proportion of SRQ low scorers required to participate in second-stage interview

$$f_2 = \sqrt{\frac{\lambda_2(1-\lambda_2) \quad C_s}{(\lambda_1 - \lambda_2)^2 \pi(1-\pi) C_d}}$$

Where:

π_1 = proportion of population screened positive

π_2 = proportion of population screened negative

λ_1 = true positives (++)

λ_2 = true negatives (--)

C_d = cost of diagnostic procedure (PSE)

C_s = cost of screen (SRQ)

C_s / C_d = ratio of screening to diagnostic costs (SRQ:PSE)

f_1 = fraction of screened positive to participate in second-stage clinical interview

f_2 = fraction of screened negative to participate in second-stage clinical interview

Assumptions on which the values used in the above equations were based.

Values used in the above equations were based on assumptions which had to be made in advance of the study. These derived from findings from similar studies and are reviewed below (Parry, 1992).

(i) Assumption 1.

Based on prevalence estimates of mental disorder in developing countries in Africa, it was assumed that prevalence would be in the region of 22% (Ben-Arie et al., 1991; Hollifield et al., 1990; Kortman, 1990).

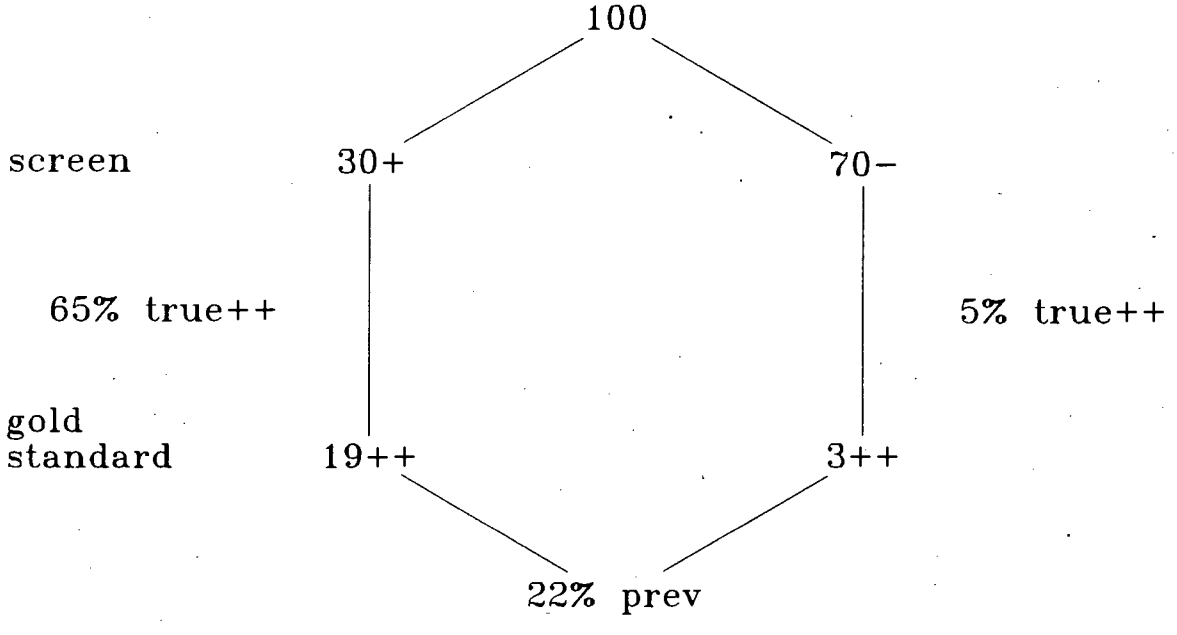
(ii) Assumption 2.

Based on the findings of similar research studies in developing countries, it was estimated that the SRQ would identify 29% of respondents as positive and 71% as negative (without a second-stage criterion, screening instruments tend to over estimate the number of true cases (Parry, 1992). It was also assumed that of those screened positive, 65% would be "true" positive, according to the

PSE, and 5% screened negative would also be "true" positive (De Jong et al., 1986; Harding et al., 1980).

Fig. 3.1.1 shows the number of screened positives and screened negatives estimated to yield a prevalence of 22%, assuming that 65% of screen positives and 5% of screen negatives will be true positive, n=100. Values used in this diagram have been rounded either up, in the case of estimated screen positives (+) and negatives(-), or down, in the case of number of true positives (++), to avoid working with fractions.

Fig. 3.1.1 Diagrammatic representation of number of screened positives and negatives and number of true positives estimated to yield a prevalence of 22%.



(iii) Assumption 3.

The cost of administering and scoring the screen was assumed to be very much lower than that of the criterion. According to Shrout (1992) the two-stage method of case detection is efficient if it can be shown that the first-stage screen is considerably more economic than the criterion.

The estimated cost ratio of the SRQ to the PSE (Cs/Cd) in the present study was 2:80 or 0.025. Estimates used in this calculation are shown in Table 3.1.1.

Table 3.1.1 Estimated cost ratio of administration and scoring of SRQ to PSE (Cs/Cd), n=100. (Figures rounded to whole numbers).

Instrument	**Administration costs**			**Scoring costs**			**Total time-cost per administration**
	Estimated cost of single administration (E_a)	Estimate of single administration *(Et)*	Estimated hourly cost per administration *(Eat)* $= E_a \times Et$	Estimated hourly cost of scoring (R/hour) (E_c)	Estimated time to score single administration *(Es)*	Estimated cost per scoring of single administration *(Ecs)* $= E_c \times E_s$	$Eat + E_{cs}$
PSE	R75	1 hour	R75 x 1 = 75	R10	20 min = 1/3hour	R10 x 1/3 = R3.30	R75 + R3.30 ~ R80
SRQ	R10	10 min = 1/6 hour	R10 x 1/6 = R1.67	R10	1 min = 1/60 hour	R10 x 1/60 = R0.17	R1.67 + R0.17 ~ R2

Using the values estimated above for n=100, the following numerical values were used in Shrout's statistical equations (see p. 1) to estimate proportion of screen positives and negatives to participate in the second-stage interview.

Where:

π_1 = *0.30* (proportion of population screened positive, n=100)

π_2 = *0.70* (proportion of population screened negative, n=100)

λ_1 = *0.65* (++) (positive on screen and criterion, ie, true positive)

$(1-\lambda_1)$ = 0.35 (positive on screen, but negative on criterion, ie, false positive)

$(\lambda_1-\lambda_2)^2 = (0.65-0.05)^2$

$= (0.06)(0.06)$

$= 0.36$

$\pi_1(1-\pi_1) = 0.30\,(1-0.30)$

$= 0.21$

λ_2 = *0.05* (negative on the screen but positive on criterion, ie, false negative)

$(1-\lambda_2)$ = *0.95* (--) (negative on screen and criterion, ie, true negative)

Cs/Cd = *0.025* (cost ratio of screen to criterion)

Therefore:

$$f_1 = \sqrt{\frac{\lambda_1(1-\lambda_1)}{(1-\lambda_2)^2\pi(1-\pi)}\frac{C_s}{C_d}}$$

$$f_1 = \frac{(0.65)(0.35)(0.025)}{(0.36)(0.21)} = 0.275$$

and:

$$f_2 = \sqrt{\frac{\lambda_2(1-\lambda_2)}{(\lambda_1-\lambda_2)^2\pi(1-\pi)}\frac{C_s}{C_d}}$$

$$f_2 = \frac{(0.05)(0.95)(0.025)}{(0.36)(0.21)} = 0.125$$

Accordingly, the estimated proportion of SRQ high scorers, randomly selected from the group of SRQ positives, to participate in the second-stage clinical interview was 27.5%, and the estimated proportion of SRQ low scorers was 12.5%

Calculation of first-stage sample size

Based on the above estimates for f_1 and f_2 the most efficient first-stage sample size was calculated that would yield a small standard error of estimate (se) of prevalence (p) at 95% confidence, and still be within the given budget of the present study. Shrout's (1992) statistical equation was used to estimate first-stage sample size. Calculation is shown below.

$$V(\hat{p}) = \frac{1}{N}\left[\frac{\pi\lambda_1(1-\lambda_1)}{f_1} + \frac{(1-\pi)\lambda_2(1-\lambda_2)}{f_2} + \pi(1-\pi)(\lambda_1-\lambda_2)^2\right]$$

Where $V(\hat{p})$ = standard error of estimate of prevalence

$$V(\hat{p}) = \frac{1}{N}\frac{(0.3)(0.65)(0.35)}{(0.275)} + \frac{(0.7)(0.05)(0.95)}{(0.125)} + \frac{(0.3)(0.7)(0.36)}{}$$

$$V(\hat{p}) = \frac{1}{N}0.5907$$

$$\therefore N = 0.5907 / V(\hat{p})$$

$$= \frac{0.5907}{0.009}$$

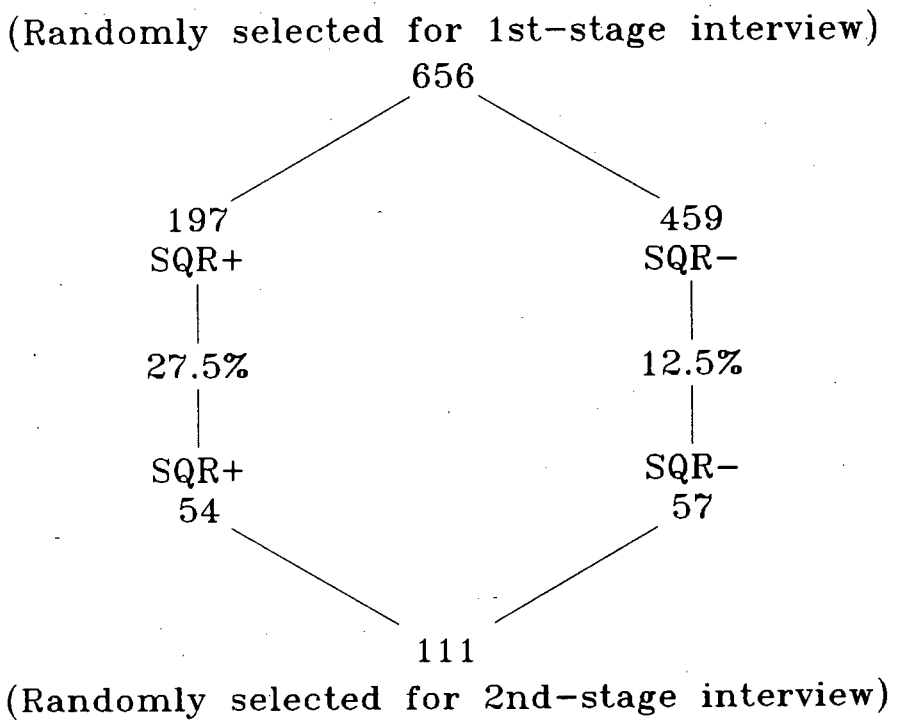
$$= 656$$

It was therefore estimated that a first-stage sample of 656 was optimal in terms both of standard error of estimate and budget constraints. Table 3.1.2 shows alternative first- stage sample sizes based on the same proportions of f_1 and f_2, but giving three different standard error of estimates of p. This table shows that a sample of 1476 yields the smallest standard error of estimate of p. However, within the limitations of the budget, a first-stage sample of 656 which yields only a slightly higher se of 0.0009, was considered both realistic and tolerable. It was therefore concluded that a first-stage sample of 656, yielding a se of 0.0009 and with a 95% confidence interval (CI) of 6% was appropriate in the present context. Fig. 3.1.2 diagrammatically shows both estimated first- and second-stage sample size and the proportion of screen positives and negatives randomly selected to participate in the second-stage interview.

Table 3.1.2 Different standard error of estimates for different first-stage sample sizes, but for the same proportion of screen positives and screen negatives elected to participate in the second-stage interview, and for a given budget.

Confidence Interval (CI)	Standard deviation	Standard error (se)	First-stage sample size
4%	2%	0.0004	1476
5%	2.5%	0.000625	945
6%	3%	0.0009	656

Fig. 3.1.2 Estimated size for first- and second-stage samples, in which 27.5% of SRQ positives and 12.5% of SRQ negatives are randomly selected to participate in the second-stage clinical interview



Cross-check of estimates used in the calculation of sample size

Assumption used in the calculations of sample size were based on community prevalence studies conducted in other countries (Parry, 1992). To ascertain that these assumptions were appropriate to the estimation of sample size in the present research, these estimates were reassessed in terms of second pilot study results.

(i) In the second pilot a prevalence of 22% was assumed. The PSE identified 5 cases indicating an estimated prevalence of 24%. This was comparable to the assumed prevalence of 22%.

(ii) In the second pilot it was assumed that of those screened positive, 65% would be true positive. The SRQ identified 8 of the 21 respondents screened as positive. Of these 5 were true positive on the PSE, indicating that 63% of those identified as positive by the screen were true positive. This figure compared favourably with the assumption that 65% of respondents would be positive on the screen.

(iii) In the pilot, out of the 16 true negatives identified by the PSE, the screen correctly identified 13, or 81%. Although not as good as the anticipated negative predictive value, of 96%, this was considered sufficiency close to the original estimate.

(iv) Finally in the pilot, the screen did not identify any false negatives, although it was assumed it would yield a 5% false negative rate.

These findings were sufficiently close to the original estimates, to accept the assumptions and estimates used in the calculation of sample size.

APPENDIX 3.2

WEIGHTED PREVALENCE OF MENTAL DISORDER REFLECTED FOR SPECIFIED SOCIO-DEMOGRAPHIC GROUPS, PHYSICAL HEALTH STATUS AND RATE OF CONSULTATION (Criterion:PSE)

APPENDIX 3.2 WEIGHTED PREVALENCE ON MENTAL DISORDER FOR SPECIFIED SOCIO-DEMOGRAPHIC GROUPS, PHYSICAL HEALTH STATUS AND RATE OF CONSULTATION (Criterion: PSE)(p=0.5)

Table 3.2.1

Test	Prevalence*	Standard error of estimate	Lower confidence interval	Upper confidence interval	Number in sample
Entire sample	0.27	0.39	0.19	0.35	121
Gender					
Male	0.21	0.06	0.10	0.31	50
Female	0.32	0.05	0.21	0.42	71
Marital Status					
Single	0.26	0.06	0.14	0.38	51
Married	0.24	0.05	0.14	0.34	62
Widowed	0.52	0.16	0.21	0.83	8
Employment					
Employed	0.28	0.07	0.15	0.42	37
Unemployed	0.27	0.05	0.18	0.37	82
Type of Work					
White collar	0.16	0.06	0.03	0.28	7
Skill manual	0.30	0.10	0.10	0.51	18
Routine manual	0.25	0.13	0.02	0.51	10
Unemployed					
Student	0.31	0.09	0.13	0.50	22
Housewife	0.10	0.10	0.08	0.29	10
Looking for work	0.25	0.07	0.10	0.40	27
Pensioner	0.46	0.12	0.22	0.69	16
Health Status					
Healthy	0.20	0.05	0.11	0.30	70
Sick past month	0.41	0.10	0.21	0.60	18
Sick now	0.44	0.08	0.27	0.60	29
Treatment if currently ill or ill in the last month					
Self treated	0.26	0.09	0.07	0.44	16
Sought help	0.46	0.08	0.31	0.62	30
If ill and treated, treated by whom					
Family	0.25	0.20	0.14	0.64	4
Private GP	0.45	0.09	0.27	0.64	23
GP Day hospital	0.68	0.11	0.47	0.89	3
Alcohol consumed					
yes	0.50	0.25	0.01	1.00	4
no	0.26	0.04	0.19	0.34	26

*Figures adjusted to reflect total sample, n= 481.

WEIGHTED PREVALENCE FOR SOCIO-DEMOGRAPHIC VARIABLES, PHYSICAL HEALTH STATUS AND CONSULTATION BY GENDER (Criterion, PSE) (p=0.5)

Table 3.2.2

Test	Prevalence*	Standard error of estimate	Lower confidence interval	Upper confidence interval
Entire sample	0.27	0.04	0.19	0.35
Marital Status				
Male single	0.24	0.08	0.09	0.39
Female single	0.29	0.10	0.09	0.49
Male married	0.11	0.06	0.00	0.22
Female married	0.29	0.06	0.17	0.42
Employment Status				
Male employed	0.17	0.08	0.02	0.32
Female employed	0.41	0.12	0.18	0.64
Male unemployed	0.23	0.08	0.08	0.38
Female unemployed	0.30	0.06	0.18	0.42
Health Status				
Male healthy	0.16	0.06	0.04	0.29
Female healthy	0.24	0.07	0.10	0.38
Male sick past month	0.41	0.16	0.09	0.72
Female sick past month	0.44	0.11	0.21	0.66
Male sick now	0.32	0.14	0.04	0.60
Female sick now	0.48	0.10	0.29	0.67
Treatment				
Male self treated	0.21	0.20	0.18	0.60
Female self treated	0.28	0.11	0.06	0.50
Male sought treatment	0.32	0.10	0.11	0.52
Female sought treatment	0.53	0.10	0.33	0.73
Alcohol consumed				
Male yes	0.50	0.25	.01	0.99
Male no	0.18	0.05	0.08	0.29
Female no	0.32	0.05	0.21	0.42

* Figures adjusted to reflect total sample, n=481

APPENDIX 3.3

CALCULATION OF VALIDITY COEFFICIENTS OF THE SRQ

APPENDIX 3.3 CALCULATION OF SRQ UNWEIGHTED SPECIFICITY AND SENSITIVITY, POSITIVE PREDICTIVE VALUE AND NEGATIVE PREDICTIVE VALUE WITH THE PSE-ID5 AS THE GOLD STANDARD

Calculation of the weighted SRQ validity coefficients were based on procedures recommended by Goldberg and Williams (1988).

Table 3.3.1 Scores on the PSE and SRQ for second-stage interview.

SRQ score threshold 7/8 n=121	**PSE-ID5** n=121		Proportion screened + and screened - in first-stage sample n=481
	non-case	case	
Above threshold	**28** false positives (a)	**29** true cases correctly identified (b)	**0.2604** proportion of high SRQs (pHS)
Below threshold	**52** true normals correctly identified (c)	**12** false negatives (d)	**0.7417** proportion of low SRQs (pLS)
Total	a + c **80**	b + d **41**	

Where:

Specificity:

proportion of true normals correctly identified $= \frac{c}{(a+c)}$

Sensitivity:

proportion of true cases correctly identified $= \frac{b}{(b+d)}$

Overall Misclassification rate:

proportion of respondents incorrectly identified $= \frac{a+d}{(a+b+c+d)}$

PPV:

proportion of high scores who are cases (ppv) $= \frac{b}{(a+b)}$

NPV:

proportion of low scorers who are not cases (npv) $= \frac{c}{(c+d)}$

(i) Calculation of unweighted **specificity:**

$$\frac{52}{(28+52)} = 0.65$$

(ii) Calculation of unweighted **sensitivity**:

$$\frac{29}{(12+29)} = 0.71$$

(iii) Calculation of **PPV**:

$$\frac{29}{(28+29)} = 0.5088$$

(iv) Calculation of **NPV**:

$$\frac{52}{(52+12)} = 0.8125$$

CALCULATION OF SRQ WEIGHTED SPECIFICITY AND SENSITIVITY

(i) Proportion of high scores (pHS) in the original sample, (n=481);

$$pHS = \frac{125}{481}$$

$$= \boldsymbol{0.2604}$$

(ii) Proportion of low scores (pLS) in the original sample, (n=481),

$$pLS = \frac{356}{481}$$

$$= \boldsymbol{0.7401}$$

therefore,

(iii) Calculation of **weighted specificity**:

$$\frac{npv \cdot pLS}{(1-ppv)\cdot pHS + npv \cdot pLS}$$

$$= \frac{(0.8125)(0.7401)}{(1-0.5088)(0.2604)+(0.8125)(0.7401)}$$

$$= \boldsymbol{0.8246}$$

(iv) Calculation of **weighted sensitivity**:

$$\frac{ppv \cdot pHS}{(ppv \cdot pHS)+(1-npv)\cdot pLS}$$

$$= \frac{(0.5088)(0.2604)}{(0.5088)(0.2604)+(1-0.8125)(0.7401)}$$

$$= \boldsymbol{0.4879}$$

APPENDIX 3.4

CALCULATION OF WEIGHTED PREVALENCE OF MENTAL DISORDER IN THE ADULT POPULATION OF MAMRE

APPENDIX 3.4 CALCULATION OF WEIGHTED PREVALENCE OF MENTAL DISORDER IN THE ADULT POPULATION OF MAMRE

Table 3.4.1 Two-by-two table of PSE-ID5 second-stage cases and non-cases, and SRQ positives and negatives for both first and second-stage samples

PSE-ID5	SRQ 7/8				TOTAL	
	Screened Negative		Screened Positive			
	356 (74%)*** (Y)		**125** (26%)**** (X)		**481** (N)	
	Given the PSE	Did not participate in PSE	Given the PSE	Did not participate in PSE	Given the PSE	Did not participate in PSE
	64 (18%)* *(n.o)*	**292**	**57** (45.6%)** *(n.i)*	**68**	**121** *(n)*	**360**
Case	**52** *(noo)*	-	**28** *(noi)*	-	**80** *(no)*	-
Non Case	**12** *(nio)*	-	**29** *(nii)*	-	**41** *(ni)*	-

* Out of the total number of screened negatives in first-stage sample, (n=356), 18% were selected for PSE interview

** Out of the total number of screened positives in first-stage sample, (n=125), 45.6% were selected for PSE interview

*** 74% of the total sample, (n=481) were screened negative

**** 26% of the total sample, (n=481) were screened positive

Statistical equation for the calculation of weighted prevalence in a two-stage case finding design (Tenenbein, 1970)

$$\hat{p} = \frac{nii}{n.i}\left(\frac{n.i+X}{N}\right) + \frac{nio}{n.o}\left(\frac{n.o+Y}{N}\right)$$

Where:

nii = false positive on the screen

n.i = true positive on the screen

X = total number of screened positives

nio = true negative on the screen

n.o = false negative on the screen

Y = total number of screened negatives

N = total number in first-stage sample

Therefore weighted prevalence is estimated at:

$$\left(\frac{29}{57}\right)\left(\frac{125}{481}\right) + \left(\frac{12}{64}\right)\left(\frac{356}{481}\right)$$

$$= \mathbf{0.27099}$$

Where:

$q = 1 - \hat{p} = 1\text{-}0.2709 = \mathbf{0.729}$ (% of healthy people)

$\hat{p}q = (0.2709)(0.729) = \mathbf{0.1976}$

$$\hat{\theta} = \frac{\frac{nio}{n.o} - \left[\frac{n.o+Y}{N}\right]}{\hat{p}} = \frac{0.1388}{0.2710} = \mathbf{0.5122}$$

$$\left(1-\hat{\theta}\right) = \mathbf{0.4875}$$

$$\hat{\varnothing} = \frac{\left(\frac{noi}{n.i}\right)\left(\frac{n.i+X}{N}\right)}{\left(1-\hat{p}\right)} = \frac{0.1277}{0.729} = \mathbf{0.1751}$$

$$\pi = \hat{p}\left(1-\hat{\theta}\right)+q\hat{\varnothing} = (0.27099)(0.4878)+(0.729)(0.1751) = \mathbf{0.2598}$$

$$1-\pi = \mathbf{0.7402}$$

$$K = \frac{\hat{p}q(1-\hat{\theta}-\hat{\varnothing})^2}{\pi(1-\pi)} = \frac{(0.271)(0.729)(1-0.5122-0.1751)^2}{0.1923} = \mathbf{0.1005}$$

$$1\text{-}K = \mathbf{0.8995}$$

$$V = \frac{\hat{p}q}{n}(1-K)+\frac{\hat{p}q}{N}K = \frac{0.1976}{121}(0.8995)+\frac{0.1976}{481}(0.1005) = 0.001469+0.00004129 = \mathbf{0.001510}$$

$$se = \mathbf{0.03886}$$

95% CI for $\hat{p}$ is $\hat{p} \pm (1.96)(se)$

ie. $0.271 \pm (1.96)(0.03886)$

0.271 ± 0.076

$0.271 \pm 0.195 \rightarrow 0.347$ *or* $19\% \rightarrow 34\%$

Conclusion

Weighted prevalence of mental disorder in the adult population of Mamre was estimated at **27%**, (se **4**), with 95% CI (**19-34**).

METHODOLOGICAL DISCUSSION

APPENDIX 4.1 ERRONEOUS SCORING OF SYMPTOM 58 ON THE PSE IN THE CASE OF ONE RESPONDENT

APPENDIX 4.2 ELEVEN CASES GIVEN A PSE-CATEGO TENTATIVE DIAGNOSIS OF PARANOID SCHIZOPHRENIA OR OTHER PARANOID STATES

APPENDIX 4.3 REVIEW OF THE ELEVEN CASES GIVEN A PSE-CATEGO TENTATIVE DIAGNOSIS OF PARANOID SCHIZOPHRENIA OR OTHER PARANOID STATES

APPENDIX 4.1

ERRONEOUS SCORING ON SYMPTOM 58 ON THE PSE IN THE CASE OF ONE RESPONDENT

APPENDIX 4.1 ERRONEOUS SCORING OF SYMPTOM 58 ON THE PSE

The high prevalence rate of mental disorder in the population, particularly of paranoid schizophrenia and other paranoid states, together with discrepancies in prevalence rates when the three criterion measures, (the PSE, PAS or "suggested clinical diagnosis") were used, prompted closer examination of all second-stage PSE score sheets. An independent external clinical psychologist, with considerable experience in the use and administration of the PSE, in both clinic and community studies, examined the score sheets of all second-stage respondents. Inspection of these score sheets was therefore conducted **after** results had been calculated.

It was found that one of the second-stage clinical interviewers had erroneously rated symptom 58 on the PSE. This had resulted in the inaccurate positive identification of one of the second-stage respondents. Estimated prevalence rates when the PSE was the second-stage criterion, were accordingly assessed to determine the possible effect this error may have had. As only one respondent had been inaccurately identified as a psychiatric case, it was concluded that there had been only minimal effect on results.

Inaccurate rating of PSE symptom 58

Examination of all PSE score sheets by the external psychologist indicated that symptom 58 (thought block/withdrawal) had been inaccurately rated positive in three cases (see Cases: 49, 225 and 310, Appendix 4.3). This had resulted in a PSE-Catego, ICD8 tentative diagnosis of paranoid schizophrenia. Further examination showed that all three cases had been rated by the same rater. These score sheets were checked with the interviewer in question and it was ascertained that the rater had been scoring symptom 58 positive **in the presence of anxiety**. The PSE-Glossary of Definitions specifically states that this symptom should not be rated present if there is any evidence of anxiety (Wing et al., 1974). The external psychologist and second-stage interviewer concurred that in all three instances, a tentative diagnoses of paranoid schizophrenia was inappropriate. While in two of the cases, the positive rating of symptom 58 could be understood in terms of the respondents' underlying organic impairment, the incorrect rating of symptom 58 had resulted in one respondent being incorrectly identified as a psychiatric case. All remaining score sheets were subsequently checked, and it was found that no other errors had been made with this item.

Effect of the inaccurate rating of symptom 58 on the prevalence estimate and validity coefficients of the SRQ

The effect of the inaccurate rating of symptom 58 had to be considered in terms of the estimated prevalence rate; the prevalence of specific diagnostic categories (particularly paranoid schizophrenia); and the sensitivity and specificity of the SRQ. As an initial investigatory step, the unweighted sensitivity and specificity of the SRQ were re-calculated. No difference between the original and re-calculated unweighted specificity was found, although the unweighted sensitivity minimally increased by 0.02. It was concluded that the mis-identification of one respondent had only a minimal effect on the estimated prevalence and validity coefficients of the SRQ. The effect on estimated prevalence of diagnostic categories is discussed in the following section.

Effect of the inaccurate rating of symptom 58 on the prevalence rate of paranoid schizophrenia in the population

As symptom 58 is part of the Nuclear Syndrome, a positive score on this symptom automatically results in a PSE-Catego, ICD8 tentative diagnosis of paranoid schizophrenia. Of the seven cases given this tentative diagnosis, only one respondent was assigned this diagnosis as a result of the inaccurate rating of symptom 58. Of the remaining six cases, symptom 58 was appropriately rated positive either in the presence of organic impairment or could be explained in terms of a subcultural delusion. This means that while the prevalence rate of paranoid schizophrenia reported in Chapter 11 may be inaccurate, the inaccuracy lies not in the idiosyncratic rating of one of the interviewers, but in the possibility that the PSE-Catego was mis-diagnosing paranoid conditions in this population. It was therefore decided that the original estimated weighted prevalence of paranoid schizophrenia should be reported in the results. Questions regarding the mis- or over-diagnoses of paranoid schizophrenia by the PSE-Catego, and its appropriateness as a criterion measure in a community sample, are discussed in Appendix 4.2.

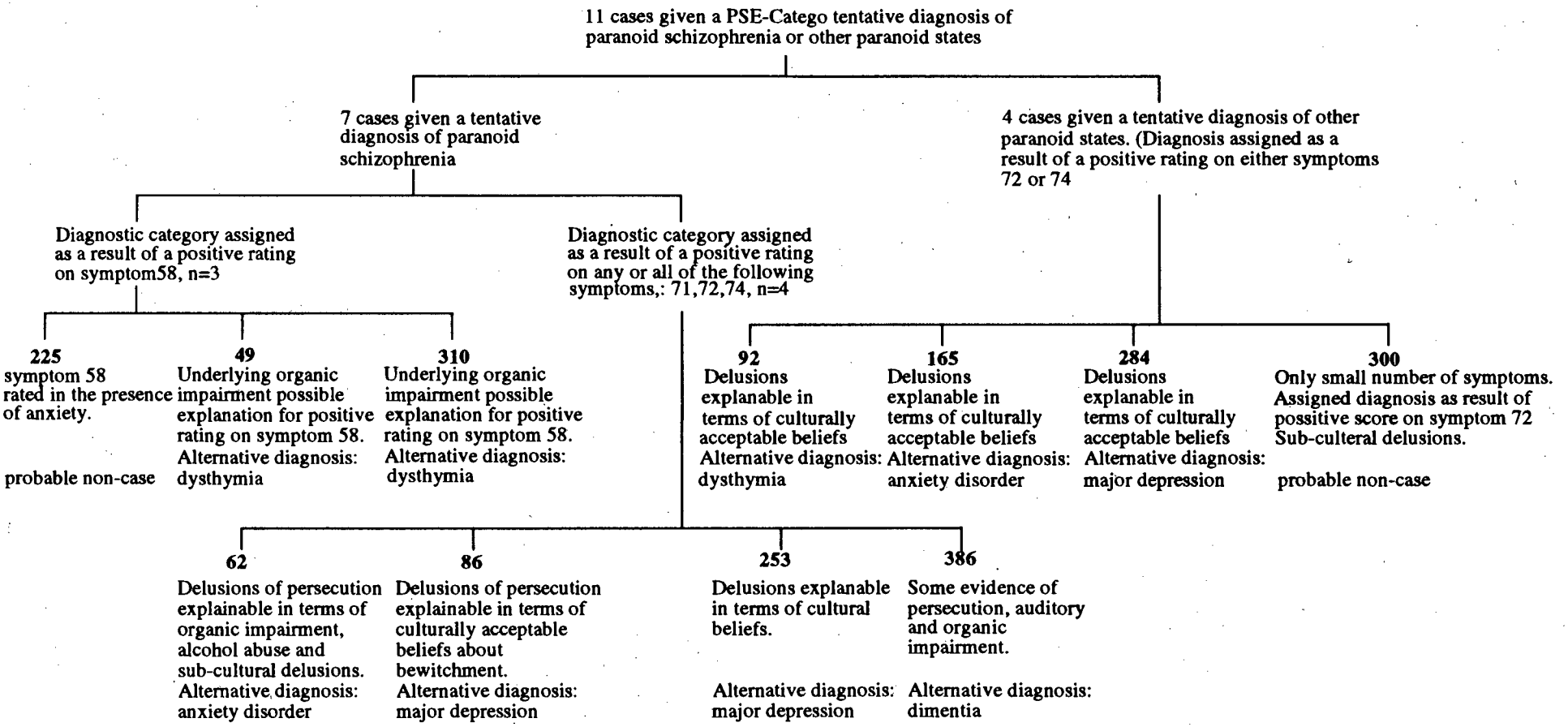
APPENDIX 4.2

ELEVEN CASES GIVEN A PSE-CATEGO TENTATIVE DIAGNOSIS OF PARANOID SCHIZOPHRENIA OR OTHER PARANOID STATES

APPENDIX 4.2 ELEVEN CASES GIVEN A PSE-CATEGO TENTATIVE DIAGNOSIS OF PARANOID SCHIZOPHRENIA OR OTHER PARANOID STATES

The PSE score sheets of the eleven cases assigned PSE-Catego tentative diagnoses of paranoid schizophrenia or other paranoid states were examined by the external clinician, and also re-examined by the five second-stage interviewers. The concerns regarding the validity of the PSE diagnosis were not communicated to the five clinicians. Interviewers were simply asked to confirm or refute their original "suggested diagnoses". While two of the cases were thought to have been incorrectly identified as psychiatric cases (see cases 225 and 300, Appendix 4.3), nine were considered to have been correctly identified as psychiatric cases, but to have been incorrectly classified as paranoid. The eleven cases are shown in Table 4.2.1, which also indicates the role played by symptoms 58, 71, 72 and 74 in the assignment of the respective diagnostic categories, and the external psychologist's alternative suggested diagnoses.

Table 4.2.1. Respondents assigned a PSE-Catego tentative diagnosis of paranoid schizophrenia or other paranoid states



Of the eleven cases, three were given a tentative diagnoses of paranoid schizophrenia on the basis of **symptom 58** (thought block/withdrawal); (see cases 49, 225 and 310, Appendix 4.3) and four were given a similar diagnoses on the basis of **symptoms 71** (delusions of control), **72** (delusions of reference) and **74** (delusions of persecution) (see cases 62, 86, 253 and 386, Appendix 4.3). Four other cases (See cases 92, 165, 284 and 300, Appendix 4.3) were also given tentative diagnoses of other paranoid states, as a result of their positive score on any of the above symptoms.

All three cases rated positive on symptom 58 were rated by the same interviewer. Their PSE score sheets were checked with the rater in question, and it was ascertained that in one case (case 225), the rater had been rating symptom 58 positive in the presence of anxiety (see Appendix 4.3). This resulted in the mis-identification of the respondent as a psychiatric case by the PSE-Catego. Organic impairment may have been responsible for the positive rating of symptom 58 in cases 49 and 310. As the PSE-Catego is unable to rate organic impairment, the positive scoring of symptom 58 resulted in the automatic erroneous assignment of a tentative diagnoses of paranoid schizophrenia. The external psychologist considered the diagnoses to be inaccurate in both cases, and suggested that in the absence of full information about the organic state, dysthymia was a more appropriate and conservative diagnosis in both cases.

With regard to the remaining eight cases, the tentative diagnoses of paranoid schizophrenia and other paranoid states had been made by the PSE-Catego on the basis of the presence of delusions, especially of a persecutory or hallucinatory nature (symptoms 61, 63 and 64). In this particular population, these symptoms may be understood in terms of culture specific experiences, and therefore may be subculturally appropriate. In Case 86, for example, the respondent was convinced that the washerwoman she employed was jealous of her and had "bewitched" her so that her leg would not heal. Although previous studies in Mamre have suggested that spirit possession is not a common explanatory model, residents nevertheless do speak of being "getoor" (bewitched) (Miller, 1989 unpublished paper). Aside from the fact that terms such as "delusion", "culture" and "subculture" are not explicitly defined in the PSE, (which has a meticulous glossary for most other terms (Swartz et al., 1985)), a positive rating on **symptom 83** (subcultural delusions) is not taken into account in the PSE-Catego computer programme. This means that in the compilation of syndromes, the Catego computer programme is unable to accommodate delusions which may be subculturally appropriate. Failure to modify ratings on symptoms of delusion and hallucination in the Catego programme appears to have resulted in both mis- and over-diagnosis of paranoid schizophrenia and other paranoid states in this population.

Table 4.2.1 also shows that of the eleven cases rated positive by the PSE-Catego, nine were confirmed cases according to the external psychologist and rating interviewers. However,

there was no evidence in any of the nine cases, to warrant a tentative diagnoses of either paranoid schizophrenia or other paranoid states. Alternative suggested diagnoses included dementia, major depression, dysthymia and anxiety.

APPENDIX 4.3

REVIEW OF ELEVEN CASES GIVEN A PSE-CATEGO TENTATIVE DIAGNOSIS OF PARANOID SCHIZOPHRENIA OR OTHER PARANOID STATES

APPENDIX 4.3 REVIEW OF THE ELEVEN CASES GIVEN A PSE-CATEGO TENTATIVE DIAGNOSIS OF PARANOID SCHIZOPHRENIA OR OTHER PARANOID STATES

CASES IDENTIFIED BY THE PSE-CATEGO BUT CONSIDERED TO BE PROBABLE NON-CASES BY THE EXTERNAL PSYCHOLOGIST EXAMINING THE PSE SCORE SHEETS

Case 225

This respondent was assigned a PSE-Catego tentative diagnosis of paranoid schizophrenia on the basis of symptoms of mild physical illness (symptoms 1, 2, 5), worry (symptom 4), irritability (symptom 40) and thought block (symptom 58). Examination of the respondent's PSE score sheet revealed that symptom 58 had been erroneously rated positive *in the presence of anxiety*. This rating error resulted in both a mis-identification of psychiatric caseness, and mis-diagnosis of paranoid schizophrenia. Both the external psychologist and the rating interviewer concurred that the respondent was likely to be a non-case.

Case 300

This respondent was rated positive on symptoms of irritability (symptom 40), delusions of reference (symptom 72), and slowness of speech (symptom 130). In spite of the small number of symptoms, the respondent was given a tentative diagnosis of other paranoid states, on the basis of symptom 72. It was concluded however, that the subject's delusions were explicable in terms of the subcultural context of the community, and that the respondent was a probable non-case.

Case 36

A PSE-Catego tentative diagnosis of neurosis was assigned, based on the presence of symptoms of worry (symptom 4), muscular tension (symptom 7), depressed mood (symptom 23), hopelessness (symptom 24), and observed anxiety (symptom 120). However, it was also noted by the rating interviewer that these symptoms reflected an "anniversary mood" which occurred in response to a specific life-experience (death of the subject's father). As the PSE-Catego has a one month time-frame, it was unable to distinguish between emotional distress that was specific to a particular event at a particular time of the year, and generalized anxiety.

Given that the respondent was able function effectively in the community, it was agreed that the subject's symptoms were not severe enough for a diagnosis at this stage. It was concluded that the respondent was likely to be a non-case.

CASES GIVEN A PSE-CATEGO TENTATIVE DIAGNOSIS OF PARANOID SCHIZOPHRENIA AS A RESULT OF THEIR POSITIVE RATING ON SYMPTOM 58

Case 49

The respondent was rated positive on symptoms of thought block (symptom 58), delusional mood (symptom 49) and organic memory impairment (symptom 103). While both symptoms 58 and 49 were explicable in terms of the underlying organic impairment, the PSE-Catego assigned a tentative diagnoses of paranoid schizophrenia, as the assessment of organic factors is beyond its scope. According to the Glossary of Definitions (Wing et al. 1974) the presence of symptom 58, which is one of the symptoms of the Nuclear Syndrome, results in the automatic assignment of the tentative diagnoses of paranoid schizophrenia. Given that the delusions were not of a persecutory nature, and occurred only after dreaming, this diagnosis was considered particularly inappropriate. Dysthymia was suggested as an alternative diagnosis.

Case 310

The rating psychologist observed that this respondent was likely to be suffering from organic impairment of memory (symptom 103) and that inefficient thinking (symptom 19), depressed mood (symptom 23), social withdrawal (symptom 28), delayed sleep (symptom 35), early morning waking (symptom 37) and thought withdrawal (symptom 58) were consistent with the underlying organic disorder. Unable to rate organic impairment, the PSE-Catego inappropriately assigned a tentative diagnoses of paranoid schizophrenia based on the presence of symptom 58. Dysthymia was suggested as an alternative diagnosis in the absence of information about organic factors in this case.

Appendix 4.3

CASES GIVEN A PSE-CATEGO TENTATIVE DIAGNOSIS OF PARANOID SCHIZOPHRENIA AS A RESULT OF THEIR POSITIVE RATING ON ANY OR ALL OF THE FOLLOWING SYMPTOMS, 71, 72 AND 74

Case 62

Symptoms of anxiety (symptoms 11, 12, 15), auditory hallucinations (symptoms 60 and 65), delusions (symptoms 71, 74, 92, 95, 96 and 97) and organic impairment (symptom 103) were considered part of the respondent's underlying organic impairment. However, as the PSE-Catego is unable to rate this disorder, it inappropriately assigned a tentative diagnosis of paranoid schizophrenia based on the presence of symptom 71. Symptom 71 is one of the symptoms of the Nuclear Syndrome and, like symptom 58, results in the automatic assignment of a tentative diagnoses of paranoid schizophrenia. Beliefs about being controlled (symptom 71), and grandiose hallucinations and delusions, are explicable in terms of subcultural factors of the population. Although symptom 83 (sub-cultural delusions) was rated positive, it was not taken into account in the compilation of syndromes by the Catego computer programme. Therefore, there is no indication that the delusions recorded, might be acceptable explanations of behaviour within the given community. In addition, the rating interviewer noted that the respondent appeared to be abusing alcohol. The above symptoms are also explicable in terms of alcohol abuse. The PSE-Catego however, is not designed to rate this syndrome. Anxiety disorder was the suggested alternative diagnosis.

Cases 86 and 253

While symptoms of delusions of a persecutory nature (71, 79, 93, 95 and 96) were evident in these respondents, it was considered that these could be explained in terms of culturally accepted beliefs about bewitchment ("getoor"). Within the cultural heritage of Mamre, bewitchment or "getoor" is an acceptable explanation for various life experiences (Miller, 1989). Given that symptom 83 is not considered in the compilation of PSE-Catego syndromes, an inappropriate diagnosis of paranoid schizophrenia was made based on the presence of symptom 71. Major depression was considered a more appropriate diagnosis.

CASES GIVEN A PSE-CATEGO TENTATIVE DIAGNOSIS OF OTHER PARANOID STATES AS A RESULT OF THEIR POSITIVE RATING ON EITHER SYMPTOM 72 OR 74

Cases 92, 165 and 284

In all the above cases, symptoms of delusions of a persecutory nature (symptoms 71, 79, 93, 95 and 96) were evident. However, there was agreement by both the external and rating psychologist that these delusions could be explained in terms of culturally accepted beliefs about bewitchment in this community (symptom 83). Failure to modify ratings on symptoms of delusion, in terms of symptom 83 in the Catego programme, appears to have resulted in the inaccurate diagnoses of other paranoid states in this population. Alternate suggested diagnoses for the three cases were dysthymia, anxiety disorder and major depression respectively.