

Measuring Consumer Outcomes In Mental Health

Field Testing of Selected Measures of Consumer Outcome
in Mental Health

Terry Stedman
Peter Yellowlees
Graham Mellsop
Robert Clarke
Suzanne Drake

University of Queensland
St. Lucia, Qld.

October 1997

© Commonwealth of Australia

ISBN 0 642 28347 8

This work is copyright. It may be reproduced in whole or in part for study or training purposes subject to the inclusion of an acknowledgement of the source and no commercial usage or sale. Reproduction for purposes other than those indicated above requires the written permission of the Department of Health and Family Services, GPO Box 9848, Canberra, ACT, 2601.

Additional copies of this paper are available free of charge and can be obtained by writing to:

Quality and Effectiveness Section
Mental Health Branch
Department of Health and Family Services
MDP 37
GPO Box 9848
Canberra ACT 2601
Facsimile: (02) 6289 7703

Authors

Dr Terry Stedman M.B.B.S. F.R.A.N.Z.C.P. Director, Clinical Services, Wolston Park Hospital; Clinical Director of the Barrett Supra Regional Acute Service; Senior Clinical Lecturer, Department of Psychiatry, University of Queensland.

Professor Peter M Yellowlees M.D. B.Sc., M.B.B.S., F.R.A.N.Z.C.P., M.R.C.Psych., M.A.Ps.S., M.R.A.C.M.A.; Head, Department of Psychiatry, University of Queensland; Director, Valley Integrated Adult Mental Health Service, R.B.H. District.

Professor Graham Mellsop M.B., Ch.B., D.P.M., M.D., F.R.A.N.Z.C.P., M.R.C.Psych. Clinical Services Director, Mental Health, Health Waikato.

Robert Clarke B.A.(Hons.), Senior Research Officer, Department of Psychiatry, University of Queensland.

Suzanne Drake B.A. M.Litt.(Psych.), Research Assistant, Department of Psychiatry, University of Queensland.

Research Assistants

Samantha Jenkins B.A.(Hons.) M.C.P.

Betty Cooper Dip.App.Sc(O.T.), M.Qual.O.T.,

Grad.DipO.H.&S.

Peter Newcombe B.A.(Hons.), B.Ed., Dip.T.

Consumer Consultant

Tony Hennessy, Consultant Evaluator, Valley Integrated Adult Mental Health Service, R.B.H. District.

Acknowledgments

Dr. Tony Baglioni Ph.D.

Toni Bell B. Health Sc., R.N., R.P.N., R.M.

Greg Brugman Dip.T, B.Ed., B.A.,

P.G.Dip.Psych.

Dr Larry Evans M.B.Ch.B., M.B.B.S.,

D.P.M., F.R.C.Psych, F.R.A.N.Z.C.P.

Jennifer Rose B.A.(Hons.)

James Whittred B.Sc., B.Bus.Mgt.

Wolston Park Hospital, Wacol, Qld.

Dr David Bartholmeuz M.B.B.S.

Mark Boschen B.A.(Hons.)

Nicole Clement B.A.(Hons.)

Dr Fred Delooze M.B.B.S.

Dr. Paul Maruff Ph.D.

Nigel Pearn B.A.(Hons.)

Joy Welham B.A.(Hons.) M.A.P.S.

Tess Withers R.P.N., R.G.N.

Commonwealth Liaison

Dermot Casey

Therese Manson

Short Title

Consumer Outcome Measures Project.

Proposed Reference

Stedman, T., Yellowlees, P., Mellsop, G., Clarke, R., and Drake, S. (1997). *Measuring Consumer Outcomes in Mental Health*. Canberra, ACT: Department of Health and Family Services.

FOREWORD

MEASURING CONSUMER OUTCOMES IN MENTAL HEALTH

Consumer Outcomes Project Advisory Group to the Mental Health Branch

Commonwealth Department of Health and Family Services

Context:

Consumers, families, clinicians, services and governments all assess the outcomes of mental health services. Often this is an informal process, with limited discussion of the findings. Many now want to do it better and to see the results used appropriately. This is part of a wide movement in health care and other areas of public life in Australia and other countries. It is linked with concerns that services are accountable, are engaged in a constant endeavour to improve the quality of their work, and use the available resources in the best and fairest way possible.

This includes the drive to ensure that evidence influences practice in an appropriate way. In mental health care, all of the above groups are especially concerned that the effects of discrimination are not perpetuated, and that the nature of mental illnesses, their broad effects, and the need for continuing care to maintain as well as to improve quality of life is recognised and included in calculation of resource provision.

The project reported here is part of the work commissioned by the Commonwealth to address two key objectives of the first National Mental Health Strategy relating to outcomes assessment (see p1). The first phase of the work (Andrews et al, *The Measurement of Consumer Outcome in Mental Health*, 1994) provided a clear review of the field and recommendations for further testing of six outcome measures. Three of the recommended measures use self reports from consumers, and three are based on clinician assessments. This was the starting point for the present project. The formation of the Project Advisory Group allowed a group of people with a range of expertise and representative views to engage in dialogue with the project group who had successfully tendered to undertake the work. This process allowed consumers, a carer, service providers in several specialist mental health and primary health care settings, and experts in measurement, evaluation and information systems to learn from and influence the project. The project group provided remarkable and consistent leadership in the project, as well as a willingness to be flexible.

The project has a clear set of findings which form the basis for a series of recommendations for the next phase of the work. The recommendations take account of the pressing need to make practical advances in the field in the face of competing pressures for health care and social resources, and the equally important need to avoid simplification and stultification in the assessment of outcomes. Both of these sets of pressures point in the same direction: the need to integrate this area with developments in quality management and evidence-based practice in mental health care, and the introduction of local information systems.

Findings:

- 1 It is both ethical and practical to evaluate mental health services. Evaluation is an obligation and while questions about what is being measured and how it might be interpreted should be subject to critical debate, good evaluation which seeks to improve the quality and effectiveness of services remains the aim.

- 2 Mental health services staff and consumers responded to the suggestion to undertake outcome assessment in a constructive and generally enthusiastic manner. The implication is that staff and consumers and families will see advantages and become engaged as partners in the process if the appropriate setting and resources are available. Whether the same applies to general practice settings is less clear.
- 3 Consumer and carer involvement helped significantly to shape the project and interpret the findings. A carer representative was invited part of the way through the project, and was able thereby to have less influence than he would have desired on the form of the project.
- 4 The reasons for assessing outcomes will influence the evaluation process and the types of measures used. Also the context in which measurement occurs is different depending on personal characteristics and the social setting of the consumer, and the nature of the illness and associated effects. There will for instance be differences in the concerns of a young person with a severe psychotic illness and a middle-aged woman with depression. The process for assessing outcomes needs to reflect these differences.
- 5 Service provider and consumer and family perspectives are different and require different measures, but each is vital for a full evaluation. Each perspective can be assessed in a variety of ways, including the use of focus groups, or individual stories, as well as scales. Group outcomes can also be assessed using demographic variables such as employment and housing types.
- 6 Measurement of consumer outcomes is only part of quality management and an evidence-based approach to service development. Other aspects such as consumer satisfaction, service standards, treatment goals and consumer empowerment are equally important.
- 7 The scales tested which assess consumer outcomes on various dimensions of symptoms and disability are feasible for use in the settings and groups included in the project. Simple global assessments of change, and visual analogue scales as also used in the project may have a place for certain purposes.
- 8 Scales of other types especially quality of life measurements were not included in the project. The development of useful generic quality of life assessments and other scales is occurring at an increasing pace, and additional instruments will continue to emerge as suitable for trial and use in mental health outcomes assessment, alone or in combination.

Recommendations: (see also p97 – 102)

- 1 The National Mental Health Policy included two key objectives relating to consumer outcomes (see p 4-5 & 110): the institution of reviews of service outcomes as a regular component of service delivery; and encouraging the development of national outcome standards and systems for assessing whether services are meeting these standards. The work of this project has stemmed from these objectives. The recommendation made in the light of the project is to modify the second of these objectives to encourage Commonwealth, State and Territory governments to provide an appropriate environment, resources and incentives for services to support outcomes assessments. The choice of a single national consumer outcomes measure should not be the primary aim at this time.

The States and Territories or services may decide to take different approaches to tackling this aim. Some could make an overall strategic decision about routine outcomes assessment. Others could encourage a variety of service responses, including the design of specific evaluation studies for particular consumer groups. In each case the fundamental requirements are (1) a clear question or set of questions; (2) the development of useful information systems, working through barriers to collection, and engagement of the staff and consumers and families in the work required (the best ways to do these things are important questions); and (3) monitoring the progress of the work. That in turn requires an open analysis and discussion of the range of findings.

- 2 The involvement of consumers and carers in planning outcome assessment studies is vital;

as is ensuring that the consumer voice and the family viewpoint are heard in the choice of instruments and assessment techniques. States and Territories and services are urged to consider the value of a combination of quantitative and qualitative techniques in this work. Each can be rigorous, and the possibilities exist for each to be used routinely or for particular times or purposes, and cover the total service or for particular populations. For instance a scale such as the HoNOS could be used over a selected period in a mental health service, and in addition focus groups of consumers and families convened.

- 3 The development and monitoring of the field should be linked explicitly with use of the national mental health service standards, with quality management and with work on evidence-based practice.
- 4 States, Territories and services may now choose from a variety of instruments detailed in the report, and it is recommended that information on the development of new instruments and on assessment methods and interpretation of findings and development of the field continue to be made available nationally.
- 5 The direct use of consumer outcome measures as a guide to resource allocation is premature and requires further study, such as would be encouraged in the approach outlined. The data so far do not allow us to conclude that the measures tested are a gold standard for consumer outcomes.
- 6 The Commonwealth, States and Territories should encourage research into service effectiveness, and the development of outcomes measurement, information systems and evaluation methods in service settings.
- 7 This approach should include a number of **strategies**:

Training and education:

- with a focus on the attitudes of staff towards consumers and families. One of the important findings of the study was that using outcome measures provided a basis for a dialogue between clinician and consumer;
- with a focus on enabling consumers to be more involved in the treatment process by using and understanding outcome measures;
- of undergraduates in clinical training courses about their role in treatment and working with consumers and families; and
- with a focus on using outcome measures in practice as part of clinician continuing education stressing the benefits to both clinician and consumer.

Consultation:

- with various disciplines by promoting the use of outcome measures;
- with consumers, clinicians, service providers and funders to encourage the development of agreed processes and use of information which is seen as mutually helpful rather than threatening.

Multi-disciplinary approach:

- examining the approach used by other areas of human services, (eg disability services) will be informative;
- incorporating the provision for collecting and using outcome measures into service procedures and information systems development (a service culture of quality and assessment).

Research:

- the use of additional outcome measures such as quality of life measures, and assessment of specific groups of consumers in different settings;
- to increase understanding of how to measure change and how to interpret the findings in different situations.

Other groups:

Carer as well as consumer representatives need to be included fully in the planning and implementing of outcome assessments. Further work needs to be undertaken at a national level before recommendations are made about outcome assessment in child and adolescent mental health services, and for Aboriginal mental health services.

The Consumer Outcomes Project Advisory Group (COPAG) believes that this project is part of an international effort, and is on the leading edge of this effort. It is also an effort shared with many fields of health care, some of which are engaged in very similar debates.

These recommendations set a new direction. It is important to take further steps right away, to provide resources and incentives or imperatives for services to assess outcomes, and to ensure that the progress and findings are monitored and discussed at the national level as well as locally. Coordination of the work and open debate about the findings are essential for progress and will allow development of the methods for outcomes assessment and services evaluation. The development of these methods is as important in the view of the COPAG as development of the measures.

Professor Helen Herrman
Chair
Consumer Outcomes Project
Advisory Group

CONTENTS

Executive Summary: <i>Measuring Consumer Outcomes in Mental Health</i>	1
Introduction	1
Methodology	1
Results	2
Conclusions and Recommendations	4
1.0 Introduction	7
1.1 Background to the Project	7
1.1.1 Report on Stage One: Measurement of Consumer Outcome in Mental Health (1994)	7
1.1.2 Developments in the Field of Consumer Outcomes Measurement in Mental Health in Australia	8
1.1.3 The Present Study	10
1.2 Outcomes Measurement	11
1.2.1 The Purpose of Outcomes Measurement	11
1.2.2 Choosing an Appropriate Measure of Consumer Outcomes in Mental Health	13
1.2.3 Issues in the Assessment of Change	16
1.3 Summary	21
1.4 Selected Measures of Consumer Outcomes	22
1.4.1 Behaviour and Symptom Identification Scale (BASIS-32)	22
1.4.2 Mental Health Inventory (MHI)	26
1.4.3 The Medical Outcomes Study Short Form 36 (SF-36)	26
1.4.4 The Health of the Nation Outcome Scales (HoNOS)	27
1.4.5 Life Skills Profile (LSP)	28
1.4.6 The Role Functioning Scale (RFS)	29
1.4.7 Summary of Issues	29
1.5 Computers and Consumer Outcomes in Mental Health	30
2.0 Aims and Research Questions	31
2.1 Aims	31
2.2 Research Rationale and Hypotheses	31
2.2.1 Feasibility	31
2.2.2 The Selected Measures as Measures of Mental Health Status	33
2.2.3 The Selected Measures as Measures of Change in Mental Health Status	35
2.2.4 The Effect of Using Computers to Collect Outcomes Data	36

CONTENTS

3.0	Method	37
3.1	Design	37
3.2	Sample	37
3.3	Procedure	42
3.4	Questionnaires	43
3.5	Focus Groups	47
3.6	Consultations with Service Provider Peak Bodies	48
4.0	Results	49
4.1	Participant Issues	49
4.1.1	Demographic, Comorbidity and Service Utilisation Data	49
4.1.2	Comparison of MHI and SF-36 Data with Normative Samples	51
4.1.3	Differences Due to Participant Attrition	51
4.1.4	Scores on the Selected Measures	52
4.1.5	Summary	52
4.2	Feasibility Analyses	53
4.2.1	Utility Ratings of Consumer Measures	53
4.2.2	Report of the Consumer Focus Groups	53
4.2.3	Report for the Consumer Support Group Focus Group	62
4.2.4	Utility Ratings of Service Provider Measures	64
4.2.5	Report of the Service Providers' Focus Groups	65
4.2.6	Findings from the Peak Service Provider Association Submissions	73
4.2.7	Discussion of Results of Feasibility Analyses	74
4.3	The Selected Measures as Measures of Mental Health Status	77
4.3.1	Validity	77
4.3.1.1	Associations between Scores for the Selected Measures	77
4.3.1.2	Analyses of Differences in Mental Health Status across Diagnostic Groups and Practice Settings	78
4.3.2	Reliability	81
4.3.2.1	Test-Retest Reliabilities	81
4.3.2.2	Internal Consistency	81
4.3.2.3	Inter-Rater Reliabilities of the Service Provider Measures	81
4.4	Selected Measures as Measures of Change in Mental Health Status	82
4.4.2	Differences in Changes in Scores for the Selected Measures from Time 1 to Time 2 across Change Groups	83
4.4.3	Change in Scores across Diagnostic Groups and Practice Settings	84
4.4.3.1	Differences in Global Change Ratings across Diagnostic Groups and Practice Settings	84
4.4.3.2	Differences in Scores for the Selected Measures across Diagnostic Groups and Practice Settings	85
4.4.4	Correlational Analyses	86
4.4.4.1	Association between Change Scores for the Selected Measures	86
4.4.4.2	Association between Change Scores on the Selected Measures and Global Change Ratings	86
4.4.4.3	Internal Consistency of Change Scores	87

CONTENTS

4.5	Effect of Computer Administration	87
4.6	Discussion of Results of Analyses of the Selected Measures as Measures of Mental Health Status and Change in Mental Health Status	88
4.7	General Discussion	90
4.8	Some Practical Issues	92
5.0	<i>Conclusions and Recommendations</i>	97
	General Recommendations	97
	Recommendations in Relation to the Six Selected Measures	101
	What Does it All Mean?	101
6.0	<i>References</i>	103
	Appendix 1: Executive Summary-Measurement of Consumer Outcome (Andrews et al., 1994)	109
	Appendix 2: Consumer Consultant's Report on Focus Groups Process, Mr Tony Hennessey	111
	Appendix 3: Tables of Results	115
	Appendix 4: Figures of Results	145
	Appendix 5: Selected Measures	173
	Appendix 6: Additional Measures Used in the Study	211

Executive Summary:

Measuring Consumer Outcomes in Mental Health

Introduction

In recent years efforts have been undertaken to improve our knowledge of the outcomes for consumers of Australian mental health services. Much of this work has been promoted and encouraged in the context of the National Mental Health Strategy that has as two of its key objectives:

- To undertake regular reviews of outcomes of services provided to people with serious mental health problems and mental health disorders as a central component of the delivery of mental health services; and,
- To encourage the development of national outcome standards for mental health services and systems for assessing whether services are meeting these standards.

A review of available approaches to outcomes assessment has been completed and six measures have been identified as potentially useful for the routine measurement of consumer outcomes in Australian mental health services (Andrews et al., 1994): the Behaviour and Symptom Identification Scale (BASIS-32), the Mental Health Inventory (MHI), the Medical Outcomes Study 36-Item Short-Form Survey (SF-36), the Health of the Nation Outcome Scales (HoNOS), the Life Skills Profile (LSP), and the Role Functioning Scale (RFS).

The present study set out to:

- Assess the level to which the selected measures are feasible (applicable, acceptable, practical), valid, reliable and sensitive to change, for three selected mental disorders (*schizophrenia, affective disorders, and anxiety disorders*), and in three clinical practice settings (*public psychiatry, private psychiatry, and general practice*);
- Subsequently identify a measure or measures which may be suitable for use as a national measure of consumer outcomes to be pilot tested at 20-40 pilot implementation sites; and,
- Articulate the potential difficulties in the implementation of the measures at the pilot sites, and to identify strategies to overcome these.

A review of the background to the study is included, and outcomes measurement is discussed in the context of the decision-making processes relevant to mental health services. Issues relating to the choice of a suitable measure, and the assessment of change (validity, reliability, measurement bias and artefact, and the significance and attribution of change), are also discussed.

Methodology

Sixty-five service providers representing a range of professional groups (*private psychiatry, public psychiatry, and general practice*) nominated 183 consumers who agreed to participate in the study, and who met the ICD-10 primary care criteria for one of the identified *diagnostic groups* (*schizophrenia, affective disorders, and anxiety disorders*). Data was collected from consumers in three stages (Time 1, Time 1(Retest), and Time 2). Time 1 data was collected from consumers at the relevant clinical setting at a convenient time. Time 1(Retest) data was collected an average of 24 hours after the initial assessment. Time 2 data was collected approximately three months later. Data was collected from service providers in two stages corresponding to Time 1 and Time 2 for consumers.

Consumers completed two of the three selected consumer measures (i.e., BASIS-32, MHI, or SF-36). Service providers completed two of the three selected service provider measures (i.e., HoNOS, LSP, or RFS). Consumers were also invited to complete adapted versions of the selected service provider measures as a method of investigating inter-rater reliability for the HoNOS, LSP, and RFS, respectively. Consumers and service providers each completed additional measures to assess the severity of illness of the consumer (visual analogue scales and modified clinical global impressions scales), global change in condition over a three month period, and to assess the utility of the selected measures (utility questionnaires).

Eight focus groups were convened to address issues related to the implementation of procedures for the assessment of consumer outcomes in mental health. As well, six peak organizations representing service providers were invited to provide submissions in relation to the validity and potential utility of the selected measures.

Results

Feasibility: In relation to the specific measures, the majority of consumers considered the MHI to have more relevant questions and to be a more useful measure than either the BASIS-32 or SF-36. It was also considered to be a *better* measure than the SF-36. In the latter case consumers' comments suggested that the MHI was more relevant for assessing mental health issues. However, comments received in relation to each questionnaire indicated a number of concerns including the time period covered, and the format and language used, by the measures.

Service providers in the *public psychiatry setting* rated the LSP significantly higher in relation to a scale score representing usefulness, effectiveness and relevance, than either the HoNOS or RFS. In the *general practice* or *private psychiatry settings* no single measure was rated significantly more highly than the others. As well, *general practitioners'* average ratings for the measures were significantly higher than those in the *public psychiatry setting*, which in turn were significantly higher than those in the *private psychiatry setting*.

The focus group data, along with comments provided by participants in relation to the measures, indicated that consumers and service providers generally supported the concept of implementing routine assessment of consumer outcomes in mental health. This is consistent with the findings of Andrews et al. (1994). The concerns raised by consumers and service providers were in relation to how such a process would actually work, and the purposes that such an exercise would serve.

Implementation issues identified by consumers and service providers included the protection of consumer confidentiality, the ability of consumers to complete the measures when they are unwell, the negative impact on consumers that completing such questions may produce, and the attribution of change in mental health symptoms to the effects of intervention.

Other factors to emerge included the need for: simplicity, brevity and comprehensiveness in the assessment process; improved assessment of general social functioning and quality of life; measures which meet accepted criteria of scientific rigor; standardised training and instruction in the use of the measures. The theme that emerged most strongly from these results is that the *measure* is not the issue: more attention needs to be given to the *processes* of outcomes assessment.

It was pointed out that where a measure did detect change in a person's condition, the attribution of this change to any single factor, including intervention, may be problematic. Service providers and consumers, alike, were concerned about this issue. Consumers and service providers suggested, for example, that people may respond in a biased manner should resource or service delivery implications be attendant upon the results.

The resource implications of implementing systems for the routine assessment of outcomes were discussed at length with issues such as the time necessary to complete the measures, the site of completion, and the reimbursement for this process, being of particular interest and concern. Confidentiality and anonymity were also highlighted as being of particular importance. Other practical issues discussed included the proposed methods of administration, and the storage and the accessibility of the obtained information.

Finally, an issue raised throughout the focus group consultations was the relationship between consumer outcomes and the quality of service provision. Consumers suggested that if the data were to be used to evaluate service effectiveness then the questions of the selected measure should address such issues in a direct manner. Some participants argued that there is a relationship between clinical outcomes and issues of service satisfaction and quality of service provision. Consequently, the questions included in the selected measures of outcomes were considered inappropriate to the evaluation of service provision. Others, however, stressed that their personal outcomes were separate to their opinion of the service they received. In either case, emphasis was placed on the need for

consumers to be asked whether they felt that the service being provided to them was effective in meeting their needs.

Many of these views were also expressed in three submissions received from service provider peak organisations.

There was a general consensus that the use of routine measures of consumer outcomes was feasible and potentially useful. Concerns which were raised highlighted issues that need to be addressed as part of any process for implementing such measures, rather than as obstacles to the process.

The Selected Measures as Measures of Health Status: The results of correlational analyses were generally consistent with predictions. Moderate to strong associations were observed between the scores of the selected consumer measures (BASIS-32, MHI, and SF-36). The weakest associations were between the *impulsive/addictive behaviour* and *psychosis* scores of the BASIS-32 and the scores of the other consumer measures; and between the 'physical health' scores of the SF-36 and the scores of the other selected consumer measures. Given the constraints in the design of the study further research is required to investigate the construct validity of the sub-scale scores and to confirm the factor structures of the measures for a large sample of people with mental health problems.

While the associations between the LSP and RFS scores were strong there was greater variability in correlations between the HoNOS scores and those of the other selected service provider measures reflecting the more diverse range of dimensions covered by the HoNOS. The test-retest reliability and internal consistency data for each measure was similar to data reported previously.

There was poor convergence between consumers' and service providers' assessments. This highlights that the assessment of mental health status is not straightforward, and any assessment process requires the contribution of the relevant stakeholders, wherever possible. A simple reading of the results of the consumer data may lead one to believe that the people with *affective disorders*, and receiving services in the *private psychiatry setting*, had greater levels of illness in comparison to the *schizophrenia* and *public psychiatry* groups; whereas a simple reading of some of the service provider data might support the inverse conclusion. These results were consistent with expectations given the distribution of *diagnostic groups across practice settings*: a large proportion of people in the *affective disorders group* were receiving services in a private acute inpatient setting; whereas most people with *schizophrenia* were receiving services in a public outpatient service. Differences between consumers' and service providers' assessments also reflect the different contents of the measures (the consumer measures generally assessed how the consumer was feeling, whereas the service provider measures generally assessed functional ability), and may also reflect different processes for completing such measures on the part of consumers and service providers.

The Selected Measures as Measures of Change: Change in mental health condition was recorded on Global Change Ratings. Almost 81% of consumers indicated some change in their mental health (60% better; 21% worse). Approximately 84% of service providers indicated that their consumers had changed (67% better; 17% worse). Ratings did not differ significantly across *diagnostic groups* or *practice settings*. On the basis of this data, three *change groups* were identified and the associations between these groups and change recorded on the selected measure scores were investigated.

There was evidence of convergence between the scores of the selected measures and the three *change groups*. The strongest results were obtained for the MHI and SF-36 data. The HoNOS and RFS also demonstrated interesting results. On each of these scales we observed significant interaction effects between *time of administration* and *change groups* which appeared to indicate that people who stated that they had experienced, or observed, a global improvement in their own, or their consumers', mental health over the three month period, recorded scores on the selected measures which were consistent with this. On the other hand, people who indicated no global change recorded similar scores at Time 1 as at Time 2. In the case of the data collected from the MHI, people who had reported a global deterioration in their mental health also recorded significant 'declines' on a number of scale scores for this measure.

Using the simple change scores (Δ =Time 2 score-Time 1 score), significant associations were observed between the *Global Change Ratings* (consumer and service provider, respectively) and the composite scores of the selected measures, except the RFS. Change for the better was associated

with 'positive' changes on the BASIS-32 Average, MHI Index, MCS (SF-36), HoNOS and LSP *total* scores; while changes for the worse were associated with 'negative' changes on those scores.

Significant multivariate effects were observed for the sub-scale change scores of the MHI, mental health scores of the SF-36, HoNOS and LSP.

Correlational analyses indicated reasonable amounts of internal consistency for the change scores (except the HoNOS and RFS); significant associations among change scores for consumer and service provider measures, respectively; and relatively low convergence between consumer and service provider change scores.

The assessment of change used in this study was limited by the quasi-experimental nature of the study; the fact that the design did not include the analyses of treatment effects; the high rate of attrition; the possibility of floor effects in the data at Time 1; and the fact that only two waves of data were collected. However, the data provides some evidence of the validity and reliability of the six measures as measures of change.

The use of computers to collect outcomes information did not appear to influence the results in any way.

The study faced a number of challenges in its implementation and in the analyses of the data obtained and these are discussed at some length in the context of their implications for practical application of routine outcomes assessment.

Conclusions and Recommendations

The main conclusion drawn from these results is that greater attention needs to be given to the *processes* for the routine measurement of consumer outcomes in clinical practice. The feasibility of implementing routine measures of outcomes will be closely related to the perceived worth of the measures to those completing them for the purposes of making care-related decisions. Moreover outcomes data needs to be developed as an integral component of any effective clinical information system.

The following general recommendations are proposed:

- The Commonwealth and State Governments should aim to develop an environment which encourages and supports routine outcomes assessment in accordance with National Standards for Mental Health Services and guidelines for clinical practice. The choice of a national consumer outcomes measure should not be the primary goal at this stage.
- Measures of outcomes should be considered primarily as tools for improving the quality of services. Outcomes assessment is not the only measurement exercise that a modern service requires. Alternative uses of outcomes data such as resource decision-making will, however, require more specific and detailed examinations of the suitability of these measures and the data obtained from them.
- Outcomes assessment should be considered as a goal in itself and needs to be differentiated from other information gathering tasks and developments.
- There is a need for more applied research into service effectiveness and the dimensions of consumer outcomes.

The focus for Phase 2 of the current research program should shift from 'measures' to 'methods' and concentrate on the issue of *how* routine outcomes assessment will actually work. This will entail addressing many of the issues raised through the focus groups, as well as:

- Consultation: consulting with service providers and consumers to determine their exact needs and how routine outcomes assessment can fit into existing service frameworks. Determining the appropriate measures for the situation. Dealing with issues of what outcomes to measure, ethical issues (e.g., confidentiality), the attribution of change, etc.
- Education: clearly defining the purpose of the process and providing a conceptual map of how outcomes data will be of direct benefit to all stakeholders (consumers, service providers, government, etc.) in the process.

- Training: providing service providers with clear training, guidelines and support for collecting outcomes data. Assisting consumers in any way they require.
- Analyses and Reporting of Outcomes Data: provide support in the analyses of the data obtained and the design of the methods for reporting back to service providers and consumers the outcomes of the assessments.
- Ongoing Support, Feedback, and Review: assisting service providers and consumers over time; following-up and reviewing the progress of the system; testing the validity and reliability of the data obtained over time and in relation to other indices of outcomes.

We simply do not know enough about the utility of these measures, nor enough about the proposed uses of a 'national' measure, to conclusively recommend any of the six selected measures as a national measure for the routine assessment of consumer outcomes in mental health. More work needs to be undertaken to investigate the utility of these measures, as well as alternative measures and systems for outcomes assessment. However, as the implementation of routine outcomes measurement requires at some stage a choice of a measure or measures, the following recommendations are presented with regard to the six selected measures:

1. The six selected measures trialed in this study should be considered as a pool of measures which have the potential to facilitate the assessment of consumer outcomes in various situations depending on the needs of consumers and the focus of service provision.
2. Services should be encouraged to implement routine assessment of consumer outcomes. Where services already have a preferred measure or approach as part of their system, unless change is clearly warranted, they should continue with this and review it in due course. Where the needs of a service do not suggest a particular approach or instrument, it is recommended that the utility and applicability of the six selected measures should be considered, and in particular that,
 - a) The MHI should be considered for trialing for an interim period as a suitable self-report measure for the routine assessment of consumer outcomes in appropriate situations.
 - b) The HoNOS should be considered for trialing for an interim period as a service provider-rated measure of outcomes across all settings.
 - c) In situations where high levels of disability are expected, the LSP or RFS may be acceptable alternatives.
3. The measures should be chosen on the basis of the identified information needs of services and consumers in accordance with the principles stated above.
4. The implementation of any measure should be subject to review following a specified period of time to assess the adequacy of the measure. The implementation of any measure should involve consultation with consumers or their representatives, and the suitable training of staff, as described above.

The recommendations concerning the MHI and HoNOS should not be taken to suggest that the other measures are not useful. Clearly the data collected in this study indicate that the other measures are valid and reliable and may be useful in different circumstances. For instance, the SF-36 may be very useful when issues of physical health are clearly relevant (e.g., in general practice); and many consumers and service providers considered the BASIS-32 and RFS to be equally reliable and valid. Indeed all of the measures have previously been recommended as useful measures of mental health outcomes in various settings. As well, the strong associations between the scores of the measures suggest a certain amount of equivalence in data collected by the measures.

The findings of this study have demonstrated:

- That there is more to the implementation of measures of outcomes than simply choosing and administering questionnaires.
- That further research is needed to examine how useful such measures are to clinical and resource-allocation decision-making processes.
- That there should be a shift towards discussion of the needs of local and regional services; and, of how data collected at the grass-roots level can be meaningfully translated into information that is relevant and useful to higher levels of management in mental health.

1.0 Introduction

1.1 Background to the Project

The assessment of clinical outcomes is a central concern for mental health service providers, consumers, and policy makers in Australia. The call for health providers to improve service quality and to produce objective evidence of accountability and effectiveness is consistent with international trends (Ellwood, 1988; Relman, 1988; Sederer & Dickey, 1996; Walters, Kirkby, Marks, et al., 1996). Active participants in this debate include governments, public and private medical organisations, funding agencies, researchers, as well as consumers and the general public.

In 1992 the Commonwealth, State and Territory Health Ministers agreed on a National Mental Health Strategy and Policy that set out an approach to improve Australia's mental health services. Consequently, positive consumer outcomes have become the first priority of both mental health policy and service delivery (National Mental Health Policy, 1992).

The National Mental Health Strategy has as two of its key objectives:

- To undertake regular reviews of outcomes of services provided to people with serious mental health problems and mental health disorders, as a central component of the delivery of mental health services; and
- To encourage the development of national outcomes standards for mental health services, and systems for assessing whether services are meeting these standards.

To implement these objectives a two-stage research program was devised by the National Mental Health Information Strategy Committee of the Australian Health Ministers' Advisory Council (AHMAC) National Mental Health Working Group. The aims of this research program are to investigate and propose appropriate methods and systems for the routine assessment of consumer outcomes in mental health.

1.1.1 Report on Stage One: Measurement of Consumer Outcome in Mental Health (1994)

Professor Gavin Andrews, Dr Lorna Peters and Dr Maree Teeson completed stage 1 of this research program. The consultancy was engaged:

- To review existing measures of consumer outcomes in mental health;
- To consult with service provider, consumer and carer organisations in order to assess current attitudes to outcomes measurement; and,
- To develop a methodology for arriving at a national outcomes measure (a copy of the Executive Summary of Stage 1 is included in Appendix 1 of this report).

In the report, *Measurement of Consumer Outcome in Mental Health* (Andrews, Peters & Teeson, 1994), six measures are identified as potentially useful for the routine measurement of consumer outcomes in Australian mental health services. These measures are: the Behaviour and Symptom Identification Scale (BASIS-32), the Mental Health Inventory (MHI), the Medical Outcomes Study (MOS) 36-Item Short-Form Survey (SF-36), the Health of the Nation Outcome Scales (HoNOS), the Life Skills Profile (LSP) and the Role Functioning Scale (RFS).

In their consultations with service provider, consumer and carer organisations, Andrews et al. (1994) observed the following:

- Of the four areas of assessment considered, 'disability' and 'quality of life' were regarded as the most important, followed by 'satisfaction with service' and 'symptoms';
- Interviews between consumers and service providers were considered the most important method for collecting outcomes data;
- The assessment of service efficacy was considered the most important use of outcomes data, followed by informing clinicians/consumers of progress and then service funding; and,
- About half of the services surveyed indicated a current use of outcomes measurements.

The authors stated, "...the general attitude to outcome measurement among service providers and

consumer groups was positive in terms of areas, methods, and uses of outcome data” (Andrews et al., 1994, p.39).

Two of the recommendations from this study were:

- That the six selected measures be field tested to determine the relative feasibility (applicability, acceptability and practicality) of using the instruments for the routine assessment of consumer outcomes; as well as to assess the validity, reliability and sensitivity to change of the instruments; and
- That nationwide piloting (over 20-40 sites) of the measure(s) selected from the Field Testing Phase be implemented.

These recommendations form the basis of Stage 2 of the consumer outcomes research program.

1.1.2 Developments in the Field of Consumer Outcomes Measurement in Mental Health in Australia

In the time since the report by Andrews et al. (1994) the field of consumer outcomes measurement has developed significantly in Australia. Indeed the present project has occurred simultaneously with a number of other research and policy initiatives.

National Standards for Mental Health Services (1997) contains a number of standards that specifically address the assessment and review processes of mental health services. In particular, Standard 11.3 states that consumers and their carers should receive a “...comprehensive, timely and accurate assessment and a regular review of progress” (p.31). The criteria for applying and meeting this standard are included in Table A1.

In recent years the Australian Health Outcomes Clearing House (AHOCH) has been established to disseminate information regarding outcomes assessment in general medical and mental health services.

On a project level, particularly in promoting local area projects, State Governments are becoming more active. For instance, the State Government of Victoria has been conducting trials to assess the utility and feasibility of the HoNOS as a statewide measure of consumer outcomes.

At the Commonwealth level, the Mental Health Classification and Services Cost (MH-CASC) project is investigating methods to create a costing and classification system for mental health services based on case attributes.

In summary, there is quite a deal of activity in relation to the assessment of consumer outcomes in mental health in this country. Regarding specific measures of outcomes there is a significant amount of research concerning the Life Skills Profile, Health of the Nation Outcome Scale, and MOS Short Form (SF-36) being collected at present. The present study is a further contribution to this area of research.

Table A1: National Service Standards-Section 11.3.

<i>Consumers and their carers receive a comprehensive, timely and accurate assessment and a regular review of progress.</i>
Criteria
<ul style="list-style-type: none">• Assessments are conducted by appropriately qualified and experienced mental health professionals.• Wherever possible, the assessment is conducted in a setting chosen by the consumer. The choice of setting is negotiated by the consumer and the MHS and considers the safety of those people involved.• The MHS has a procedure for appropriately following up people who decline to participate in an assessment.• The MHS has a system for commencing and recording assessment during the consumer's first contact with the service.• The assessment process is comprehensive and, with the consumer's informed consent, includes the consumer's carers (including children), other service providers and other people nominated by the consumer.• The assessment is conducted using accepted methods and tools.• The MHS has documented protocols and procedures describing the assessment process.• The assessment is recorded in an individualised clinical record in a timely and accurate manner.• There is opportunity for the assessment to be conducted in the preferred language of the consumer and their carers.• Staff are aware of, and sensitive to, cultural and language issues which may affect the assessment.• Diagnosis is made using internationally accepted medical standards by an appropriately qualified and experienced mental health professional.• Where a diagnosis is made, the consumer and carers (with the consumer's informed consent) are provided with information on the diagnosis, options for treatment and possible prognoses.• Wherever possible, the MHS conducts face-to-face assessments but may use telephone and video technologies where this is not possible due to distance or the consumer's preference.
Review
<ul style="list-style-type: none">• The MHS ensures that the assessment is continually reviewed throughout the consumer's contact with the service.• Staff of the MHS involved in providing assessment undergo specific training in assessment and receive supervision from a more experienced colleague.• New assessments are subjected to a clinical review process by the MHS.• All active consumers, whether voluntary or involuntary, are reviewed at least every three months. The review should be multidisciplinary, conducted with peers and more experienced colleagues and recorded in the individual clinical record.• A review of the consumer is additionally conducted when:<ul style="list-style-type: none">* the consumer declines treatment and support* the consumer requests a review* the consumer injures themselves or another person* the consumer receives involuntary treatment* there has been no contact between the consumer and the MHS for three months* the consumer is going to exit the MHS* monitoring of consumer outcomes (satisfaction with service, measure of quality of life, measure of functioning) indicates a sustained decline.• The MHS has a system for the routine monitoring of staff case loads in terms of number and mix of cases, frequency of contact and outcomes of care.• Documented policies and procedures exist and are used to achieve the above criteria.• The MHS monitors its performance in regard to the above criteria and utilizes data collected to improve performance as part of a quality improvement process.

1.1.3 The Present Study

In December 1995 the consultants for the present study were commissioned to perform the field testing of the six measures of consumer outcomes selected by Andrews et al. (1994).

The study had two goals:

- To provide advice to the AHMAC National Mental Health Working Group as to which of the six selected measures identified by Andrews et al. (1994) would be most suitable for the routine measurement of consumer outcomes in Australian mental health services; and,
- To provide advice to the AHMAC National Mental Health Working Group as to relevant issues with regard to the implementation of such measures in everyday clinical practice.

More specifically the project aimed:

- To assess the level to which the six selected measures of outcomes are *feasible (that is, applicable, acceptable, and practical), valid, reliable and sensitive to change*, in three selected *mental disorders (schizophrenia, affective disorder and anxiety disorder)* and in three clinical *practice settings (public sector, private sector and general practice)*;
- To subsequently *identify a measure or measures* which may be suitable for use as a *national measure of consumer outcomes* to be pilot tested at 20-40 pilot implementation sites; and
- To *articulate the potential difficulties in the implementation of the measurements* at the pilot sites, and to *identify strategies* to overcome these.

Phase 2 of Stage 2 of the Commonwealth's research program on the measurement of mental health outcomes will aim to:

- Test the preferred measure(s) at various sites covering a range of geographical locations and service types.
- Assess the applicability and requirements for national implementation of the measurement of consumer outcomes in routine clinical use.
- Develop a training package and engage a trainer(s) to educate service providers at pilot sites regarding the value and utility of measures of consumer outcomes.
- Produce a training package for use on a national level of implementation.
- Organise the collection and analysis of consumer outcomes data from the pilot testing and report on a suitable research methodology for the collection and analysis of consumer outcomes data.

The purpose of the present study was not to address the issue of *why* measure consumer outcomes. This has been comprehensively dealt with previously as part of the Andrews et al. (1994) report and by other investigators (e.g., Ellwood, 1988; Jenkins, 1990; Mirin & Namerow, 1991; Sederer & Dickey, 1996). The principle question underpinning the current study is: *How* to routinely measure consumer outcomes in Australian mental health services in an appropriate manner? Until recently much of the literature concerning outcomes assessment has not dealt directly with the issue of implementing systems of assessment. This has started to change over the last couple of years (Smith, Fischer, Nordquist et al., 1997). With this has come an appreciation of the complexity of outcomes assessment and of the potential pitfalls that may be encountered if a range of issues are not adequately addressed. The present study can be seen as part of a growing literature with respect to the practical application of routine outcomes assessment.

To set this study in its proper context, and to discuss the criteria for judging the appropriateness of a measure, it is necessary to briefly address the issue of the purposes and motivations for measuring consumer outcomes in mental health.

1.2 Outcomes Measurement

1.2.1 The Purpose of Outcomes Measurement

Clearly, one of the major problems for the development of mental health policy in countries around the world is the lack of complete and accurate data on which to base decisions...What is also clear is that mental health problems are one of the major public health problems facing countries. (Kemp, 1993, p.1).

Mental health services and planners face the ubiquitous problem of a lack of detailed information concerning vital elements of service delivery in mental health. The introduction of measures for the routine assessment of consumer outcomes is a practical response to this problem.

What is an 'Outcome'?

The Australian Health Ministers Advisory Council in 1992 defined 'consumer outcome' as "...the effect on a patient's health status that is attributable to an intervention" (Andrews et al., 1994, p.12). To talk about consumer outcomes measurement involves two criteria being met. First, there is the assessment of change within an individual, or groups of individuals. Second this change must be attributed to the effects of a health intervention or treatment. However, the issue of the attribution of change to treatment effects is not a simple and straightforward process, as will be discussed below.

The simplistic approach to assessing outcomes is to routinely use an acceptable and standardised outcomes measure. The concepts of 'consumer outcomes' and 'consumer outcomes measure' are, however, quite separate (Kraemer & Telch, 1992). An 'outcome' denotes a characteristic of a person. An 'outcomes measure' is "...a protocol to obtain a number or classification from or about the patient that is indicative of the 'outcome'...Because the quality of outcome measures is determined by how closely an *outcome measure* reflects the *outcome* of interest, it is critical to maintain clear distinction (sic) between the two terms" (Kraemer & Telch, 1992, p.86). In other words, one must not confuse a person's rating on a questionnaire (the measure of outcomes) with what the individual or group has actually experienced (their personal outcomes). The degree to which the two correspond will depend on the precision and accuracy of the measure used.

Goals of Outcomes Measurement

Historically, the goals of outcomes research have been to investigate treatment efficacy (at individual and group levels), to identify groups for whom interventions are most effective, to describe service provider and consumer characteristics related to different outcomes, and to identify processes responsible for favourable outcomes (Eisen & Dickey, 1996). In recent times there has been a shift away from a 'research' paradigm with practitioners and service managers now being encouraged to undertake their own outcomes research. "Implicit in this shift is the idea that although not all clinical practitioners can be outcome researchers, all can (and perhaps should) conduct outcome assessments with their clients in their clinical practices" (Eisen & Dickey, 1996, p.181).

Outcomes research is also being increasingly called upon in the economic analyses of service and treatment effectiveness. That is, it is used to link clinical outcomes with the costs incurred by consumers and service providers (Mirin and Namerow, 1991; Sederer, Dickey & Hermann, 1992).

Fundamentally, the primary goal of outcomes measurement and research is to facilitate decision-making processes (Benjamin, Perfetto & Greene, 1995). Adapting a definition first proposed by Ellwood (1988), outcomes assessment can be defined as *the collection and analyses of information from consumers and service providers designed to assist consumers, providers and policy makers to make rational care-related choices based on a better insight into the likely effect of such choices on the lives of consumers.*

Sutherland & Till (1993), in discussing the assessment of quality of life, identify three levels of decision-making in health services: micro (clinical), meso (agency, institutional or regional), and macro/meta (governmental) levels (see Table A2). At a clinical or micro-level "...the object of decision-making is individual patient benefit" (p.298). Decision-making involves consumers and service providers.

Those factors which hold the greatest relevance to treatment goals are those which are consequently targeted for outcomes assessments.

At a meso-level the major feature is the group perspective. It involves decisions about treatments and resource allocation with regard to the health care of defined groups of consumers "...where the goal is to offer the most effective interventions to persons using the services of particular institutions, agencies or regional networks" (Sutherland & Till, 1993, p.300). Macro-level decision-making is concerned with policy and resource allocation at governmental and population levels (Sutherland & Till, 1993, p.301).

Intersecting each level of this model are two factors that influence decision-making processes (see Table A2): clinical needs and resource needs. Outcomes assessment has been proposed by various authors as one way in which to address (see Sederer & Dickey, 1996):

- a) the need to improve *clinical* decision making processes, and thereby the quality of services to individual consumers; and
- b) the need to improve *resource allocation* decision making processes to ensure that service quality is maintained by the appropriate and equitable targeting of resources at each level of service delivery.

Table A2: Model of Levels and Dimensions of Decision Making for Mental Health Services (Adapted from Sutherland & Till, 1993)

Level	Target	Dimensions of Decision-Making	
		Clinical	Resource Allocation
Micro	⇒ Individuals ⇒	e.g., "What treatment goals are necessary to meet the needs of this individual?"	e.g., "What resources are needed to meet the goals of treatment for this person?"
Meso	⇒ Groups ⇒	e.g., "What are the best treatment approaches for addressing the needs of this group of people?"	e.g., "What resources are needed to meet the goals of treatment for this group?"
Macro	⇒ Population ⇒	e.g., "What are the best treatment approaches for meeting the needs of this population of people?"	e.g., "What resources are needed to meet the goals of treatment for this population?"

Clinical Decision Making

“Clearly the consumers of mental health services are the primary focus of any examination of treatment outcomes” (Mirin & Namerow, 1991, p.2). The primary benefit of collecting and utilising outcomes assessments should be the improvement in the care and services that are provided to consumers (Walters et al., 1996).

Measures of outcomes have the potential to benefit clinicians and consumers directly in a number of ways:

- For clinicians, measures of outcomes are an obvious method of objectively assessing the worth of treatment interventions for their consumers. Service providers, “...physicians and others who provide care need outcome data to help guide their efforts...Reliable outcome data are an essential first step in making treatment decisions that incorporate advances in our scientific understanding about the etiology and treatment of mental illness” (Mirin & Namerow, 1991, p.4).
- For consumers, measures of outcomes provide a point of feedback for themselves, and their carers, to evaluate the value of the treatments that they receive.

Despite the other uses that outcomes data may be put to, the clinical utility of the measures of outcomes will be the primary concern of consumers and clinicians alike. The ease with which measures of outcomes can be implemented as part of routine service practice, and the extent to which they are taken seriously, will likely be determined by the degree to which consumers and service providers value their clinical utility and consider them to be relevant (Smith et al., 1997; Thompson, 1992).

Resource Allocation Decision Making

Much of the literature discussing measures of outcomes has either been prefaced by, or set clearly in the context of, a discussion of the need to improve service quality and efficiency, and of maximising the cost effectiveness of the mental health dollar (e.g., Sederer & Dickey, 1996; Wells, 1995). Measures of consumer outcomes are proposed as one method by which service managers, policy makers and funding bodies (public and private) can evaluate service provision at local, regional, and national levels, and make decisions as to the appropriateness of service delivery models and the allocation of scarce resources.

Just how useful measures of outcomes can be in facilitating economic decision-making, is at this stage a moot point. Resource allocation, however, has a wider definition than simply dollars and cents. And a number of processes have been identified in which outcomes data may be used to facilitate decision-making at the local and regional service level:

- The identification and matching of service consumers with respect to specific service types;
- The assessment of staffing and funding needs in light of consumers’ service needs; and
- The assessment of the effectiveness of specific treatment modalities given the needs of local and regional communities (Sonnanburg, 1996).

1.2.2 Choosing an Appropriate Measure of Consumer Outcomes in Mental Health

The choice of a measure of outcomes will depend first and foremost on the decisions that are to be taken on the basis of the information obtained from the measure. Clearly the first question in choosing any measure is ‘What outcomes do you wish to measure?’ Mental health services provide a range of services to people with various mental health conditions. Obviously the outcomes desired in each case will necessarily vary. Furthermore the measure of outcomes that is chosen must be able to provide information that will assist the service provider and consumer to make a decision as to whether or not the consumer has experienced a *change* on the dimension of interest.

These considerations are implicit in the criteria described by Andrews et al. (1994) who highlighted a number of dimensions for assessing the suitability of a measure of consumer outcomes (Andrews et al., 1994, p.p.29-33):

- *The measure must be applicable.* The measure should address dimensions that are important to consumers (symptoms, disability and consumer satisfaction) as well as being useful for clinicians in formulating and conducting treatment. The measure should also provide information of a nature, and in a format, that facilitates the management of services.
- *The measure must be acceptable.* The measure should be brief, and the purpose, wording, and interpretation should be clear (i.e., it should be user-friendly).
- *The measure must be practical.* Issues of practicality relate to: the burden imposed upon consumers and service providers in terms of time, costs, training in the use of the measure, and level of skill required in the scoring and interpretation of the data.
- *The measure must be valid.* The measure should have sound psychometric properties and should measure what it is supposed to measure.
- *The measure must be reliable.* The measure should (within acceptable limits) provide the same results when given to the same person on two occasions or by two different people.
- *The measure must be sensitive to change.* The measure should be able to indicate whether a *clinically* significant change has occurred for a consumer over consecutive administrations of the measure. This is the definitive property of a measure of consumer outcomes. A measure may provide information with respect to mental health status, but it is the extent to which it assesses meaningful change in a person's condition that it can be called an *outcomes measure*.

Andrews et al. (1994) concluded: "In summary, measures which are suitable for use in routine clinical practice are likely to be brief, low cost, multidimensional measures which require minimal training in their administration, scoring, and interpretation, but which are sufficiently reliable, valid and sensitive to change to indicate the outcome of the therapeutic intervention" (p.33).

The criteria proposed by Andrews et al. (1994) are similar to those proposed by the United States National Institute of Mental Health (NIMH). The NIMH described eleven criteria of the 'ideal outcome measure' (Ciarlo, Edwards, Kiresuk et al., 1986; Green & Gracely, 1987). These are included in Table A3.

Perusing the criteria developed by Andrews et al. (1994) and Ciarlo et al. (1986), and considering the different levels of decision-making which need to be addressed within mental health services, one might legitimately ask, does such a measure exist? Can any one measure or questionnaire address the diverse range of needs of people with mental illness, as well as the differing needs of clinicians and service managers? Clearly there are those who believe that it is not possible (e.g., Brown, McCartney & Bell, 1995).

Furthermore value judgments about what should be considered as 'outcomes' will necessarily influence the choice of measures of outcomes. As Cummins & Baxter (1994) argue: "Choice of outcome variables also defines what the program should be aiming to achieve and, therefore, what are the most important outcomes for the recipients of the services. In this way, evaluators make a value judgment on behalf of the people who are being served" (p.22). Because such information will be used to provide feedback to policy makers and planners, who will in turn determine the shape of future services, measures must be used "...which are most relevant to the service recipients".

Table A3: Criteria for an Outcomes Measure-National Institute of Mental Health (Adapted from Green & Gracely, 1987; Ciarlo et al., 1986)

1. The measure (or set of measures) should be relevant and appropriate to the client group(s) whose treatment is being studied (i.e., it should cover the most important and frequently observed symptoms, problems, goals, or other domains of change for the group(s)).
2. The measure(s) should involve simple methodology and procedures that can be implemented uniformly by a majority of service facilities, using accessible and well-defined training materials and instructions. The scores from a measure should have clear and objective referents ('meanings') that are consistent across consumers, to ensure interpretability of scores as well as changes in scores.
3. The scores of the measures should have clear and objective referents.
4. The measure(s) should reflect the perspectives of all relevant participants in the treatment process.
5. Measure(s) that provide information regarding the means or processes by which treatments may produce positive effects are preferred to those that do not.
6. The measure(s) should meet the minimal criteria of psychometric adequacy, including: reliability, validity, sensitivity to treatment change; freedom from respondent bias, and nonreactivity to extraneous situational factors that may exist (including physical setting, client expectations, staff behaviour, accountability pressures, etc.).
7. The measurement materials and implementation procedures should be relatively inexpensive.
8. The content and the presentation of results should be understandable to a wide audience, including consumers, Public Servants, and the General Public, as well as mental health professionals.
9. A measure's scores should provide easy feedback to various audiences, and readily interpretable without extensive statistical skill.
10. Measure(s) that are useful in clinical service functions (diagnosis, treatment planning, case review) are preferred to help facilitate acceptance and implementation of the outcome measurement effort.
11. The measure(s) used should be compatible with a wide range of theories of psychopathology and the goals and procedures of various treatment approaches.

1.2.3 Issues in the Assessment of Change

The first goal of any health service is to facilitate an improvement in the health of the consumers of the service (Streiner & Norman, 1989). The effectiveness of a health service is determined by the ability of the service to assist consumers to make positive changes. Furthermore, the ability to detect meaningful change within an individual, and within groups of individuals, over time is the defining property of a measure of outcomes (Guyatt, Walter, & Norman, 1987).

In order to assess change one must use methods that can accurately measure:

- The mental health status of an individual or group of individuals at a specific point in time, in particular, prior to treatment intervention; and
- The change in mental health status of an individual or group of individuals during, and following the termination of, treatment.

The simple approach to measuring change is to assess a person on a specific dimension (e.g., symptoms of depression) at one point in time, and then again at a second point in time. The difference between the two scores can be called a 'difference score', 'change score', or 'gain score' (Rogosa, Brandt & Zimowski, 1982). If there is no difference between the scores it may be concluded that there has been no change on the dimension measured. Likewise, if there is a difference between the two scores, it may be concluded that there has been a change (either for better or worse). In the latter case any 'change' still needs further investigation in the light of the *reliability*, *significance*, and potential *cause(s)* of this change. Such considerations relate more specifically to the issues of: the subject of measurement; the sources of information; the validity and reliability of the measure of outcomes as a measure of change; bias and artifact in the assessment process; the *statistical* and *clinical* significance of change (i.e., What does a change score really mean?); and, the attribution of change (i.e., What caused or contributed to the change in this person?).

What to measure?

There is a large body of literature on the dimensions that should be included in outcomes assessment (see Chandler, Meisel, Hu et al., 1996; Clark & Friedman, 1983; Jerrell & Ridgeley, 1995; Meltzer, 1995; Mirin, Gossett and Grob, 1991; Sederer & Dickey, 1996). Various approaches adopt a condition-specific format (e.g., Smith, Ross & Rost, 1996a & 1996b) or utilise more generic approaches (such as that recommended by Andrews et al., 1994). Ultimately the choice of what to measure will depend upon the decisions that are to be made on the basis of the information obtained from the measure. At the micro-level of decision-making the measures will be specific to the goals of treatment as designed in accordance with the needs of the consumer. At the meso- and macro-level, measures will be more generic in order to summarise outcomes on a limited range of dimensions across large groups of individuals.

If routine questionnaires are introduced across settings compromises must be reached. On one hand, the implementation of any measure to be used on a routine basis will require the acceptance and support of service providers and consumers. Moreover, the implementation of any measure will depend on the *utility* of the measure: that is, the likelihood that the results of the measurement will influence the course of treatment (Thompson, 1992). However, as Andrews et al. (1994) acknowledged, the need for data across large groups of people necessitates the use of generic measures to speed collection, analysis, reporting etc., and to allow comparison between different groups of consumers.

Sources of Information

It is now widely recognised that the assessment of treatment outcomes should involve all relevant stakeholders. It is an accepted philosophy that consumers, and where appropriate their carers, should be involved in all stages of treatment decision-making including assessment and review (Anthony & Nemeck, 1984; Eisen, Grob & Dill, 1991; National Standards for Mental Health Services, 1996).

While such a principle makes good sense, issues arise with regard to the level of agreement between stakeholders on the degree and importance of change in mental health status. In a recent review of quality of life research, Stedman (1996) reported that agreement between doctors' and their patients' judgments regarding the patients' quality of life were generally poor. In relation to mental health, low levels of agreement between service providers' and consumers' assessments have also been observed in regard to:

- Assessments of symptom distress and global functioning (Clark & Friedman, 1983; Piersma & Boes, 1995).
- Assessments of treatment outcomes (Conte, Plutchik, Picard et al., 1988; Sullivan & Grubea, 1991).
- Assessments of depression (White, White, & Razani 1984).
- Assessment of comorbidity amongst substance abusers (Ross, Swinson, Larkin & Doumani, 1994).
- Social skills assessment (Monti, Corriveau & Currna, 1982).
- Judgments of the importance of different aspects of treatment (Perreault, Rogers, Leichner & Sabourin, 1996).

Factors that may influence the level of agreement between service providers' and consumers' judgments include: diagnosis (Perreault et al., 1996), level of disability (Corrigan, Buican, & McCracken, 1996), severity of illness of consumers (White et al., 1984) as well as consumers' level of insight into their condition (Dickerson, Boronow, Ringel & Parente, 1997); and initial reaction to treatment (Priebe & Gruyters, 1994). These dimensions focus mainly on the consumer. Relatively little research is to be found on the dimensions of service provider characteristics that may influence agreement between service providers and consumers, although the work on clinician effectiveness may eventually shed light on this (e.g., Blatt, Sanislow & Zuroff, 1996; Seligman, 1995). It is likely that factors such as the time spent with the consumer, the comprehensiveness and degree of standardization of the service providers' assessment, and the effort to understand the meaning of consumers' unique individual circumstances, are obvious factors that may influence the level of concordance between ratings.

At the same time it is important to concede that lack of agreement between service providers and consumers need not always be framed negatively. Instead it may simply suggest that both parties approach the assessment process from different perspectives, experiences and knowledge bases, and thus bring a "multidimensional" aspect to the assessment process.

The Validity and Reliability of an Outcomes Measure as a Measure of Change

When attempting to assess whether change has occurred it is important to know whether the method used provides information that is meaningful (i.e., *valid*) and free of error (i.e., *reliable*).

Validity: Validity is defined as the degree to which a scale measures what it is supposed to measure. Validity is not simply a property that resides in a questionnaire or measure. Rather it is a function of the *purposes* to which the questionnaire or measure is to be put, and the *inferences* that are to be made from the information derived from the measure (American Psychological Association, 1985; Cronbach, 1984; Messick, 1995; Pedhazur & Pedhazur-Schmelkin, 1991). The *Standards for Educational and Psychological Testing* of the American Psychological Association (1995) state that validity refers to the "...appropriateness, meaningfulness, and usefulness of the specific inferences made from test scores. Test validation is the process of accumulating evidence to support such inferences" (p.9).

A number of categories of validity are usually identified, including, face, content, criterion, and construct validity (see Pedhazur & Pedhazur-Schmelkin, 1991).

Two other categories of validity include *convergent* and *divergent* validity. These two terms refer to the expectation that scales that purport to measure similar constructs should provide similar results when their results are compared (convergent validity); whereas scales that purport dissimilar constructs should provide dissimilar results when the results are compared (divergent validity).

Sensitivity to change is closely related to the concept of validity. The sensitivity to change, or

responsiveness of a measure, is defined as its ability to detect clinically significant changes in a person's condition over time (Guyatt et al., 1987; Guyatt, Deyo, Charlson et al., 1989). Essentially the sensitivity to change of a measure can be considered as the validity and reliability of the measure as a *measure of change*.

Reliability: Reliability is closely related to validity, and essentially refers to the degree to which "...test scores are free from errors of measurement" (American Psychological Association, 1985, p.19). Three categories of information can be used to establish the reliability of a measure: test-retest reliability, inter-rater reliability, and internal consistency data.

The *test-retest method* of assessing reliability involves administering a measure on two separate occasions to the same people. As long as there is no reason to expect any change in the manner that the people respond to the items, it would be expected that the scores from the two occasions would display a high level of correspondence, or *correlation*. The *inter-rater method* of assessing reliability involves using two or more independent raters to assess the same phenomenon or event. The degree to which the raters agree can be used to evaluate the reliability of inferences being made from the data obtained from the measure. The *internal consistency* of a measure refers to the degree to which the items or questions of an item 'hold together'. That is, the degree to which items that are combined to create a composite score (e.g., 5 items that are used to create a Depression sub-scale) measure the construct consistently.

Bias and Artifact in the Assessment Process

In any measurement process there is the potential for extraneous factors to influence the outcomes of the measurement. Such factors are described as *bias* or *artifact* and relate to the social psychological processes surrounding the completion of forms, decision making etc., as well as the statistical procedures for analysing the data obtained from the measure. Some of the more readily identified biases are described in Table A4.

There are a number of measurement artifacts that can impede the interpretability of change data. These include:

- *Regression to the mean*: A statistical artifact whereby the scores for people at Time 2 tend to become less extreme and closer to the average for the group than at Time 1, irrespective of the potential effect of any intervention (Campbell, 1996).
- *Spontaneous remission*: Some disorders improve over time without any intervention at all.
- *Inert placebo*: Simply being involved in a treatment program may produce positive effects for a person irrespective of the efficacy of that program.

Table A4: Potential Sources of Measurement Biases (Streiner & Norman, 1989, p.54)

- *Differing Perspectives.* "...the people who develop a scale, those who use it in their work, and the ones who are asked to fill it out, all approach scales from different perspectives, and with differing amounts of information about the scale".
- *Social Desirability and Faking Good.* Most people want to be seen in a good light. When asked questions, especially those which relate to deviant behaviour, there is a tendency for people either unconsciously (social desirability) or consciously (faking good) to create a positive impression.
- *Deviation and Faking Bad.* These are simply the opposite of social desirability and faking good. They may occur when the objective of the questionnaire respondents (unconsciously or consciously, respectively) is to obtain a service to fulfill their needs.
- *Yea-saying and Acquiescence.* This is the tendency to respond positively or in the affirmative to any question presented, leading to contradictory responses.
- *End-aversion, Positive Skew and Halo.* Scales that are scored on a continuum (e.g., Visual Analogue Scales and Likert Scales) are susceptible to a number of biases. End-aversion refers to a reluctance amongst some people to use extreme choices or categories of a scale, which is probably due to a difficulty in making absolute judgments like "always", "never" etc.. The *halo effect* describes the situation in which judgments made about aspects of an individual's behaviour are influenced by the rater's overall impression of the person.
- *Framing.* This bias refers to the fact that the manner in which a question is posed, or framed, will influence the way a person responds to it.

Finally, the fact that these measures are intended to be completed by or for people with mental disorders or illnesses raises special issues, including:

- *Cognitive deficits of consumers:* Many consumers of mental health services experience severe cognitive deficits which impair their judgment. This fact in itself can impinge upon the reliability of data collected on standardised self-report measures, especially if care is not taken to take account of the individual needs of the consumer (Corrigan et al., 1996).
- *Vulnerability of consumers.* As a social group people with mental illness experience a high degree of discrimination. Needless to say this may engender a degree of cynicism and reluctance to use questionnaires which may be perceived to be irrelevant, or to be mechanistically labeling the person.

Assessing the Significance of Change

If a person is assessed on two occasions, and a 'change' in scores has been observed, how *significant* is this change? Two types of significance are discussed in relation to the assessment of change: *statistical* and *clinical significance*.

Statistical Significance: Treatment effects are typically assessed and reported on the basis of statistical comparisons of ratings on some relevant measure or scale across two or more time periods (Keppel, 1982; Jacobson & Traux, 1991). A test of statistical significance tells us whether the differences between two or more groups of data are likely to be due to chance.

Tests of statistical significance, however, can be influenced by different factors, including the sample size and the variability within the sample with respect to responses to the test questions (Deyo & Patrick, 1995). Furthermore, statistical techniques deal with *groups* of people and compare differences across *groups* of people. The variability amongst people within the group is treated as a source of error. Statistical analyses across groups do not really tell us about the characteristics of the individuals within these groups (Jacobson & Traux, 1991).

Statistical Measurement of Change: Conceptually the easiest way of measuring change is to assess a person at one point in time (Time 1) and then again at a later point in time (Time 2) and then to

subtract the Time 1 score from the Time 2 score. The magnitude of the change can then be examined for its statistical significance. However, writers such as Cronbach & Furby (1970), Lord & Novick (1968), and Nunnally (1978) argue that simple change scores are unreliable as true measures of change. The reason for this is that change scores are "...systematically related to any random error of measurement" (Cronbach & Furby, 1970, p.68). Alternative methods involving analyses of variance and covariance, and regression techniques, are recommended in lieu of change scores.

Nevertheless, there is a growing body of literature supporting the use of 'change scores' as valid indicators of change (Rogosa et al., 1982; Rogosa & Willet, 1983; Williams & Zimmerman, 1996). The argument of these authors is that the statistical reasoning traditionally adopted to refute the use of change scores employs unrealistic case scenarios. It has also been argued that the use of analytical methods such as those suggested by Cronbach & Furby are unreliable in quasi-experimental studies (Streiner & Norman, 1989; Weisberg, 1979).

Clinical Significance: Clinical significance refers to the extent to which treatment meets the goals and expectations of all parties involved with the treatment, including consumer, clinician, carer, etc.. Jacobson & Traux (1991) provide a non-exclusive list of criteria defining clinical significance, which includes:

- The level of change that is recognized as meaningful by peers and significant others;
- The elimination of the presenting problem;
- The achievement of normal or 'high end-state' levels of functioning by the end of treatment; and
- Changes that reduce the risk of health problems.

Clinical significance is defined in relation to the consumer's expectations and the goals of treatment, as well as the clinician's expertise and knowledge of the consumer and their mental illness.

The Attribution of Change

Assuming that the conceptual difficulties relating to the significance of change have been overcome, and it can be reliably concluded that a person has changed, can change be attributed to the effects of the treatment provided to that person? The answer is 'not necessarily'. The attribution of change is complicated by the fact that there are so many forces influencing people in their day-to-day lives (Sonnanburg, 1996). Treatments and assessments do not occur in a vacuum. As has been highlighted above, there are a number of factors that can contribute to improvements on test scores that need not be related to treatment effects. When considering the case of change for individuals or groups, these factors need to be taken into account.

The data collected from measures of outcomes, at best, can tell us if there has been a change in an individual's mental health condition. But they cannot, necessarily, tell us what caused this change (Raphael, Hugh & Stewart, 1997). The fact that the observed change coincided with an intervention could simply be due to chance or any of the other factors mentioned above. The possibility of misinterpreting information about change in a person or group is always present. Sonnanburg (1996) suggests the following strategies to minimise such errors:

- Use multiple assessment methods (i.e., more than one questionnaire);
- Use multidimensional methods of assessment;
- If possible use more than one rater;
- Examine the perspective of all who are effected by the treatment (consumer, clinician, carer);
- Use qualitative data to assess the consumers' and clinicians' opinions concerning potential change and the meaning of change scores;
- Investigate other factors in the consumer's life, other than treatment, that may be contributing to change; and
- Use multiple assessments over time. Simply taking before and after ratings is not enough. Assessments during treatment and follow-up assessments are recommended.

What this indicates is that single measures of consumer outcomes, on their own, are probably not sufficient to provide the appropriate amount of information required to facilitate the kinds of decision-making processes required by mental health services.

1.3 Summary

The overview presented here is meant to illustrate the diverse range of issues that need to be taken into account when implementing systems of routine outcomes assessment, and interpreting the information collected from measures of outcome.

Obviously it is beyond any single piece of research to comprehensively answer the questions that are raised in this review. The aims of the present study are to provide a broad response to what are very broad and general questions, in order to facilitate Phase 2 of the consumer outcomes research project. Issues relating to the assessment of change, the appropriate methods for assessing outcomes, and the processes required to ensure the feasibility of implementation of any measure, are topics that will be of long-term interest and concern.

Ultimately the utility of any measure will be determined by the nature of the information it provides and we turn now to an examination of the six selected measures.

1.4 Selected Measures of Consumer Outcomes

Andrews et al. (1994) identified six measures of consumer outcomes that may be of potential value as routine measures. The scales are: the Behaviour and Symptom Identification Scale (BASIS-32), the Medical Outcomes Study 36-Item Short-Form Survey (SF-36), the Mental Health Inventory (MHI), the Health of the Nation Outcome Scales (HoNOS), the Role Functioning Scale (RFS) and the Life Skills Profile (LSP). Three of the scales are designed to be completed by service providers (HoNOS, RFS, LSP) and three are self-report measures designed to be completed by consumers (BASIS, SF-36, MHI). Tables A5 and A6 provide brief descriptions of each of the measures (copies of the measures are included in Appendix 5).

1.4.1 Behaviour and Symptom Identification Scale (BASIS-32)

The authors of the BASIS-32 scale describe it as "...an empirically derived, patient-oriented measure of symptoms and behavioural distress designed for use in outcome assessment and developed specifically for a psychiatric inpatient population" (Eisen, Dill & Grob, 1994, p.246).

The authors (Eisen et al., 1994) also state that the BASIS-32 possesses four features which distinguish it from other measures of outcomes. These include:

- That it was empirically derived from consumers' perspectives;
- That it was developed with an acutely ill psychiatric inpatient sample;
- That it includes the major psychiatric symptoms and functioning difficulties in one measure; and,
- That it combines individualised and standardised approaches to patient assessment.

The BASIS-32 contains 32 items designed to investigate the extent to which a person has been experiencing difficulties on a range of dimensions. The period of assessment is the preceding week. The respondent answers the question by marking a rating of: 0 (no difficulty), 1 (a little difficulty), 2 (moderate difficulty), 3 (quite a bit of difficulty), and 4 (extreme difficulty). The BASIS-32 contains five sub-scales: *relation to self and others*, *daily living and role functioning*, *depression and anxiety*, *impulsive and addictive behaviour*, and *psychosis*. As well, a single average score for the whole questionnaire is computed.

Assessments can be undertaken at intake and at specified intervals during or following treatment. The BASIS-32 is a flexible questionnaire to administer. At least four techniques have been used: structured interviews (either with a rater present or by telephone), or self-report (either on-site or through mail-out). Eisen (1995) suggests that comparable participation rates are obtained from self-administered versus interview-administered methods.

Acceptable test-retest reliability, discriminant validity and concurrent validity for at least three of the sub-scales have been reported. Significant change in all measures at six months follow-up suggests some sensitivity to change.

There are a number of concerns relating to the use of the BASIS-32. The first of these is that the sub-scales were derived from the results of factor analyses and, therefore, may or may not truly represent the construct nominated. Also, the study population from which this data was obtained included patients from a private inpatient psychiatric setting, more than 80% of whom had a DSM-III-R Primary Axis I diagnosis of affective disorder or substance abuse. Only 9% of the consumers had a diagnosis of schizophrenia or a non-affective psychoses and only 2% had a diagnoses of anxiety disorder. In addition to this the authors (Eisen et al., 1994) reported some difficulties in validating the psychoses sub-scale.

Table A5: Selected Measures of Consumer Outcomes in Mental Health (Adapted from Andrews, G, Peters, L., and Teeson, M. (1994).)

<i>Title</i>	<i>Time</i>	<i>Description</i>	<i>Use</i>	<i>Application</i>	<i>Reliability</i>	<i>Validity</i>	<i>Sensitivity to Change</i>
Behaviour and Symptom Identification Scale (Eisen et al, 1994)	20-30 mins.	Interviewer-administered, self-administered questionnaire. 32 items with 5 sub-scale: relation to self and others, daily living and role functioning, depression and anxiety, impulsive and addictive behaviour, psychosis. Uses 5-point scale. Cost unknown, nil training.	Mental health	clinical research	item-adequate test-retest-adequate	content-adequate construct-adequate criterion-adequate	yes
Health of the Nation Outcome Scales (Wing, Curtis & Beevor 1996)	15-30 mins.	Clinician-rate scale; 12 sub-scales: aggression, self-harm, alcohol and drugs, memory/orientation, physical problems, mood disturbance, hallucinations and delusions, other mental, social relationships, social environment (housing and finance), overall severity. Nil cost, training required.	Mental health	clinical	item-adequate inter-rater-adequate	nil	yes
Life Skills Profile (Rosen et al, 1989; Parker et al, 1991)	20-25 mins (est)	Questionnaire completed by clinician or family member. 39 items, assessing general functioning over past three months, using 5 scales: self-care, non-turbulence, social contact, communication, responsibility. Uses 4-point scale. Costs \$1 per copy. No training required, but rater must be familiar with patient.	Mental health	clinical research	item-adequate test-retest excellent inter-rater adequate	content-adequate construct-adequate criterion-adequate	yes
Mental Health Inventory (Veit & Ware, 1983)	10-15 mins	Self-rated questionnaire; 38 items with 5 sub-scales: anxiety, depression, behavioural/emotional control. General positive affect, emotional ties. Rated on 5 and 6 point scales; covers past month. Other versions: MHI-18, same sub-scales as MHI-38; MHI-5, 1 item for each of the 5 sub-scales. Cost unknown, nil training.	Mental health General population	survey	item-excellent test-retest adequate	content-adequate criterion-adequate	yes
Role Functioning Scale (Goodman et al, 1993)		Clinician-rated scale. 4 single scales: working, independent living and self-care, immediate social network relationships, extended social network relationships. 7-point rating scale: scale summed to give a global index. Cost unknown, training required.	Mental health	research	item-excellent test-retest-adequate inter-rater-adequate	construct-adequate criterion-adequate	no
SF-36 (Ware & Sherbourne, 1992; McHorney et al, 1993)	5-10 mins	Self-rated questionnaire (can also be given to the informant or administered as an interview). 8 multi-item scales: physical functioning, physical and emotional role limitations, bodily pain, mental health, social functioning, vitality, general health perceptions, reported health transition. 4 week or 1 week time frame. Copyright conditions govern use, minimal training required.	Depression general health	clinical research survey	item-adequate test-retest-adequate	content-adequate construct-excellent criterion-adequate	yes

Table A6: Selected Measures of Consumer Outcomes in Mental Health

<i>Title</i>	<i>Score and Scales</i>	<i>Calculations</i>	<i>Other Details</i>
Behaviour and Symptom Identification Scale (BASIS-32) (Eisen et al, 1994)	<ul style="list-style-type: none"> • Relation to Self and Others • Depression and Anxiety • Daily living/role functioning • Impulsive/addictive behaviour • Psychosis • BASIS-32 Average 	<p>Average</p> <ul style="list-style-type: none"> • Items 7, 8, 10-12, 14 & 15. • Items 6, 9, 17, 18, 19,& 20. • Items 1, (2, 3, 4'), 5, 13, 16, 21, & 32. • Items 25, 26, 28, 29, 30, & 31. • Items 22, 23, 24, & 27. • Items 1-32. 	<ul style="list-style-type: none"> • 4 sub-scales and the BASIS-32 Average are computed by averaging the ratings for component items using the non-missing items as denominator. • Average Range: 0-4. • Higher scores obtained indicate a higher level of problems for the construct named by the item, scale or index.
Medical Outcomes Study (MOS) Short Form (SF-36) (Ware & Sherbourne, 1992; McHorney et al, 1993)	<ul style="list-style-type: none"> • Physical Component Summary (PCS) Measure • Mental Component Summary (MCS) Measure • Physical Functioning (PF) • Role Physical (RP) • Bodily Pain (BP) • General Health (GH) • Vitality (VT) • Social Functioning (SF) • Role Emotional (RE) • Mental Health (MH) 	<p>Recoding and addition of:</p> <ul style="list-style-type: none"> • PF, RP, BP, & GH • VT, SF, RE, & MH • 3a-3j • 4a-4d • 7 & 8 • 1, 11a-11d • 9a, e, g, & I • 6 & 10 • 5a-5c • 9b-9d, 9f & 9h 	<ul style="list-style-type: none"> • Possible score range: <ul style="list-style-type: none"> 21-75 14-70 10-30 4-8 2-12 5-25 4-24 2-10 3-6 5-30 • Higher scores obtained indicate a higher level of the construct named by the item, scale or index.
Mental Health Inventory (MHI) (Veit & Ware, 1983)	<ul style="list-style-type: none"> • Anxiety • Depression • Behavioural/Emotional Control • General Positive Affect • Emotional Ties • Psychological Distress • Psychological Well-Being 31,34,37 	<p>Recoding and addition of</p> <ul style="list-style-type: none"> • 3,11,13,15,25,29,32,33,& 35 • 9,19,30,&36 • 8,14,16,18,20,21,24,27,& 28 • 4-7,12,17,26,31,34,& 37 • 10&23 • 2,3,8,9,11,13-16,18-21, 24,25,27-30, 32,33,35, 36, 38 • 1, 4-7,10,12,17,22,23,26, 	<ul style="list-style-type: none"> • Possible score range: <ul style="list-style-type: none"> 9-54 4-23 9-53 10-60 2-12 24-142 14-84
Mental Health Index	<ul style="list-style-type: none"> • All items 		<ul style="list-style-type: none"> • Possible score range: <ul style="list-style-type: none"> 38-226 5-30 • Higher scores obtained indicate a higher level of the construct named by the item, scale or index.

Table A6 (continued): Selected Measures of Consumer Outcomes in Mental Health

<p>Health of the Nation Outcome Scales (HoNOS) (Wing, Curtis & Beevor 1996)</p>	<ul style="list-style-type: none"> • Overactive, aggressive, disruptive behaviour • Non-accidental self-injury • Problem drinking or drug-taking • Cognitive problems • Physical illness or disability problems • Problems with hallucinations & delusions • Problems with depressed mood • Other mental and behavioural problems • Problems with relationships • Problems with activities of daily living • Problems with living conditions • Problems with occupation and activities • Behaviour (1-3) • Impairment (4,5) • Symptoms (6-8) • Social Skills (9-12) • Total (1-12) 	<ul style="list-style-type: none"> • Individual item scores • Four sub-scale scores calculated by addition of items shown, respectively • Total score addition of individual items (1-12) 	<ul style="list-style-type: none"> • Higher scores obtained indicate a higher level of the construct named by the item, scale or index (i.e, greater mental health problem).
<p>Life Skills Profile (LSP) (Rosen et al, 1989; Parker et al, 1991)</p>	<ul style="list-style-type: none"> • Self-care • Non-turbulence • Social contact • Communication • Responsibility • Total 	<p>Addition of items:</p> <ul style="list-style-type: none"> • 10, 12-16, 23, 24, 26 & 30 • 5, 6, 25, 27-29, 32, 34-38 • 3, 4, 20-22 & 39 • 1, 2, 7-9 & 11 • 17-19, 31 & 33 • sum of all items 	<ul style="list-style-type: none"> • Items for each scale are summed, and a total score obtained by summing scale scores. • Higher scores obtained indicate a higher level of the construct named by the item, scale or index (i.e., greater social functioning and ability).
<p>Role Functioning Scale (RFS) (Goodman et al, 1993)</p>	<ul style="list-style-type: none"> • Working • Independent Living and Self Care • Immediate Social Network Relationships • Extended Social Network Relationships • Global Role Functioning Index 	<ul style="list-style-type: none"> • Raw Score • “ • “ • “ • Addition of 1-4 	<ul style="list-style-type: none"> • Range R1-R4: 1-7 • Range GRFI: 4-28 • Key to interpretation of GRFI scores: <ul style="list-style-type: none"> 4 Severely limited 5-8 Markedly Limited 9-12 Limited 13-16 Marginal 17-20 Moderate 21-24 Adequate 24-28 Optimal • Higher scores obtained indicate a higher level of the construct named by the item, scale or index (i.e., greater social functioning and ability).

The authors also highlight a number of other limitations in the original validation sample, namely the exclusion of consumers who were considered too psychotic, confused, or otherwise unable to be appropriately interviewed, and the possibility of response bias, including denial or the desire to appear well (Eisen et al., 1994). Despite these issues the BASIS-32 has in recent years attracted much attention as a measure of consumer outcomes.

1.4.2 Mental Health Inventory (MHI)

The Mental Health Inventory (MHI) was originally intended as a measure of mental health for a general population (Davies, Sherbourne, Peterson & Ware, 1988; Veit & Ware, 1983). The Rand research group developed it alongside the SF-36 with John Ware as a principal architect. Indeed a number of questions have been taken directly from the MHI to make up the *mental health* sub-scale of the SF-36.

The measure contains 38 items. Each item includes a description of a particular symptom or state of mind, and the respondent indicates on a scale the degree to which they have experienced this in the past month. All of the scales, except two, are scored on a six point scale. A number of scores are derived from the MHI. These include a global Mental Health *Index* score; *psychological distress* and *well-being* scores; and six sub-scale scores representing *anxiety*, *depression*, *loss of behavioural/emotional control*, *general positive affect*, *emotional ties*, and *life satisfaction*. The Mental Health Inventory can be completed either as a self-report measure or as part of an interview.

As with the SF-36, there have been analyses of the properties of the MHI as part of the RAND HIE project. The measure has been factor analyzed using exploratory and confirmatory factor analytic methods, on a very large sample of the general population ($N=5089$), across a number of sites in the United States (Veit & Ware, 1983). The factor structure has displayed robustness across cultures (Liang, Wu, Krause et al., 1992).

The MHI has been used in a wide variety of studies to assess mental health. These have included: the mental health of consumers infected with human immunodeficiency virus (HIV) (Grummon, Rigby, Orr et al., 1994; Tsevat, Solzan, Kuntz et al., 1996); the mental health of consumers with ovarian cancer (Kornblith, Thaler, Wong et al., 1995; Portenoy, Kornblith, Wong et al., 1994) and breast cancer (Kornblith, Hollis, Zuckerman et al., 1993; Siedman, Portenoy, Yao et al., 1995); the effects of dietary fat intervention on the mental health of women (Bowen, Kestin, McTiernan et al., 1995); as a screening tool for depression amongst the elderly (Dorfman, Lubben, Mayer-Oakes et al., 1995) and as a general measure of mental illness for primary care consumers (Berwick, Murphy, Goldman et al., 1991; Weinstein, Berwick, Goldman et al., 1989;). The MHI has also been recommended as a suitable tool for the assessment of adolescent mental health (Ostroff, Woolverton, Berry & Lesko, 1996).

Despite the large amount of research involving the MHI, its status as an outcomes measure for people with serious mental health problems requires further investigation. In most cases, it has been used in non-psychiatric populations and for the assessment of group differences.

1.4.3 The Medical Outcomes Study Short Form 36 (SF-36)

The SF-36 scale was developed as a generic measure of health status (Ware, 1996; Ware, Snow, Kosinski & Gandek, 1993). It was originally developed as part of the Medical Outcomes Study (MOS), a large project carried out by the RAND Corporation in the United States of America during the 1980s. The purpose of the study was to develop practical tools for monitoring consumer outcomes and their determinants in routine practice (Stevenson, 1995).

In contrast to the BASIS-32, the scale items contained in the SF-36 were chosen by the authors (Ware & Sherbourne, 1992) on the basis of those concepts most commonly used in other scales measuring consumer outcomes. These include: *physical health*; *role limitations due to physical health problems*; *bodily pain*; *social functioning*; *general mental health* (psychological distress and psychological well-being); *role limitations due to emotional problems*; *vitality* (energy/fatigue); and *general health perceptions*.

The questionnaire includes 36 items in total, organised into 11 sections. Items from the various sections are used to calculate the eight scale scores. These may be displayed as a profile of health status. In turn, the scale scores are also used to calculate a *physical component score (PCS)* and a *mental component score (MCS)*.

The SF-36 is a short measure that is easy to administer in a number of modalities (e.g., face-to-face, telephone interview, self-report etc.). Among its considerable strengths are its wide use in large groups with a range of illnesses; its use in the general population; and, the large volume of research that has been undertaken using this measure. The SF-36 also has a very comprehensive and useful manual for administrative and interpretation purposes.

The SF-36 has been used widely with people suffering from affective disorders in community settings (Hayes, Wells, Sherbourne et al., 1995; McHorney, Ware & Raczek, 1993). Use amongst people suffering from anxiety disorders and schizophrenia is not extensive.

The Australian Bureau of Statistics included the SF-36 in the 1995 National Health Survey. Data from this survey is now available (Stevenson, 1995). This information will provide a sound basis for comparing both the burdens caused by mental and physical illness, and the results of treatment. This may be helpful in supporting the National Mental Health Plan's objective to mainstream mental health services.

Commensurate with its extensive use, there have been comments made in regard to the shortcomings of the SF-36 (Sansoni, 1995). In summary, it has been suggested that some questions are poorly constructed, and that the development of the measure has been excessively reliant on psychometric considerations at the expense of relevance to consumers. The outstanding issues in relation to this instrument are its reliability, acceptability and relevance in non-affective disordered and inpatient groups, and its clinical relevance for individual consumers and clinicians.

1.4.4 The Health of the Nation Outcome Scales (HoNOS)

The Health of the Nation Outcome Scales (HoNOS) (Wing, Curtis & Beevor, 1996) was developed for use as a routine measure of consumer outcomes in mental health services in the United Kingdom. As in Australia, the Government of the United Kingdom has set targets for the improvement of the health and functioning of people with mental illness. 'Health and social gain' for people with mental illness incorporate:

- An improvement in the mental, physical and/or social functioning, over that which one would expect without intervention; and
- The maintenance of an 'optimal functioning state' by preventing, slowing and/or mitigating deterioration.

The HoNOS is the method by which to assess whether these goals are being achieved. As such, the HoNOS is a relatively new questionnaire. Pilot testing commenced in 1993 and the measure is undergoing development in the U.K. Currently a self-report version and a version for children and adolescents are under development.

The HoNOS contains 12 scales. Each scale measures a type of problem commonly presented by consumers of mental health services. A completed HoNOS sheet provides a profile of 12 severity ratings and a total score. The 12 scale items address problems resulting from *overactive, aggressive, disruptive or agitated behaviour; suicidal thoughts or behaviour, non-accidental self-injury; problem drinking or drug taking (substance use); cognitive problems* involving memory, orientation, understanding; problems associated with *physical illness or disability*; problems associated with *hallucinations and delusions; depressed mood; other mental and behavioural problems*; problems with making supportive *social relationships*; problems with *activities of daily living (overall disability)*; problems with *living conditions*; and problems with *occupation and activities*.

Four sub-scale scores are also computed: *behaviour, impairment, symptoms* and *social skills*. A *total* score is also computed through the addition of the twelve item scores. Ratings are made on a five point scale ranging from 0 (no problem within the period rated) to 4 (severe to very severe). Service providers complete the HoNOS using all relevant and useful information that they would

gather through their normal procedures (i.e., case notes, team meetings, interviews with consumers, carers, etc.). The authors state clearly that training is required to ensure that accurate assessments are made with the HoNOS and they have developed a training package to introduce new users to the measure.

More information is required on the psychometric properties of this measure, especially within an Australian sample. There are currently a number of studies and field trials underway (e.g., Goldney, Fisher, Walmsley et al., 1996). The Victorian Government is considering using the HoNOS as a statewide measure of consumer outcomes, and the Commonwealth is using the HoNOS in its MH-CASC Project. No doubt, given the enthusiasm with which the HoNOS has been greeted by evaluators in this country and in the U.K., a substantial amount of information is likely to become available about this measure in the near future.

1.4.5 Life Skills Profile (LSP)

The Life Skills Profile (LSP) was developed to measure function and disability in people with schizophrenia and chronic mental disabilities and addresses itself particularly towards community-based services (Parker, Rosen, Emdur & Hadzi-Pavlovic, 1991; Rosen, Hadzi-Pavlovic & Parker, 1989). The measure has already been adopted by a number of services in Australia as a useful assessment tool. It has also been used in a number of studies to assess service delivery and consumer outcomes (Parker & Hadzi-Pavlovic, 1995; Trauer, Duckmanton & Chiu, 1995). Although it has been developed in Australia, versions of the questionnaire are used in other countries (e.g., Fernandez de Larrinoa Palacios et al., 1992).

The authors of the LSP note a number of minimum requirements that this measure should meet. These are:

- That it focus on aspects of function that affect survival and adaptation in the community;
- That it assess broad and relevant constructs rather than fine details of behaviour;
- That it be able to be completed by both professional and nonprofessional raters;
- That it be capable of ready administration without any need for special training;
- That scores be derived easily to maximise its utility for service users; and
- That it meet appropriate standards of reliability and validity.

The Life Skills Profile is a 39-item measure. Items assess general levels of psychosocial functioning and disability. The questionnaire provides five sub-scale scores and a *total* score. Scores are derived through the summation of the scale items. The five sub-scales are *self-care*, *non-turbulence*, *social contact*, *communication*, and *responsibility*. Higher scores represent more positive functioning. The LSP is intended to be completed by service providers (e.g., case manager), or a person who knows the individual in question well (e.g., a family member or carer).

The LSP has been used as an outcomes measure in various studies (Hambridge & Rosen, 1994). It appears to demonstrate sensitivity to changes in the functioning of people who have participated in a community based intervention program. In a recent study, Trauer et al. (1995) concluded that the internal consistency of the scale was good, but the inter-rater reliabilities were only marginally acceptable. They also recommended that the factor structure be fine-tuned and, to this end, described an alternative five-factor model. Familiarity (a function of the duration of contact with the consumer) emerged as a variable that had a significant influence on a number of the scale scores.

The items included in the LSP are highly relevant to the seriously mentally ill in public mental health settings and it appears to be a very good measure for the purposes for which it was designed. There is good reliability and validity data from Australian studies and very similar results from a Spanish population. There is also considerable local experience with this scale and it has been used as an outcomes measure in several Australian mental health services. Moreover, there is some evidence of its sensitivity to change.

However, the Life Skills Profile is designed to measure the needs of people who already have a clearly defined disability. More specifically, the LSP has become the generally accepted measure for public or community mental health services assessing the mental health outcomes of people with chronic schizophrenia. The period of assessment is problematic for people suffering from acute

episodes of mental illness rather than a chronic illness. Further information is required on the applicability, acceptability, reliability and sensitivity to change of the LSP for people with affective and anxiety disorders, as well as people who receive treatment from private psychiatric services or general practitioners.

1.4.6 The Role Functioning Scale (RFS)

The Role Functioning Scale was designed to be a statewide outcomes assessment measure of community mental health programs in the state of Georgia in the United States (Goodman, Sewell, Cooley & Leavitt, 1993; McPheeters, 1984; Newman, 1980). However, apart from some recent work (Brekke, 1992) this scale has not been extensively reported in the literature. Despite this, Andrews et al. (1994) describe the RFS as including most of the dimensions which are likely to be relevant to consumers of mental health care. In this regard they are supported by Ciarlo et al. (1986).

The RFS includes four sub-scales. Each is scored on a seven-point scale. The four scales are: *working productivity* (the consumer's functioning in the most appropriate expected role of homemaker, student or wage earner); *independent living, self-care* (the consumer's management of household, eating, sleeping, hygiene care); *immediate social network relationships* (the consumer's transactions as a friend, spouse or member of a primary group); and *extended social network relationships* (the consumer's transactions in the neighbourhood, community church, clubs, agencies, recreational and other community resources). A *Global Role Functioning Index (GRFI)* score is derived by the simple summation of the four sub-scale scores.

Further investigation is warranted on the applicability, acceptability, reliability and sensitivity to change, for a range of consumer groups and across different clinical settings of the RFS. As explained by McPheeters (1984), soon after its development there was a change in mental health policy in Georgia and the RFS received less attention than previously because it was felt to lack clinical utility, validity and reliability. It is, however, accurate to point out that the RFS has come to be seen as an outcomes measure primarily for people with schizophrenia and chronic mental disability (Brekke, 1992).

1.4.7 Summary of Issues

Each of the measures addresses a specific range of dimensions of mental health in a different manner. Each was designed with different purposes in mind. Some of the measures have obviously attracted a greater deal of attention in regard to outcomes assessment, most notably the HoNOS, LSP, BASIS-32 and SF-36. Furthermore the measures differ in terms of the amount of research that has been undertaken into the validity and reliability of the information derived from them.

In summary there were three key areas that needed to be addressed in relation to the selected measures, including:

- Whether, and to what extent, these measures assess mental health status, and the various aspects of mental health status that are of relevance to service providers and consumers in making rational care-related decisions;
- Whether, and to what extent, these measures report meaningful changes in the mental health status of consumers;
- Those issues which consumers and service providers consider relevant to the effective implementation and utilization of measures for the routine assessment of consumer outcomes in mental health.

1.5 Computers and Consumer Outcomes in Mental Health

It was decided to collect as much data as possible using computer-formatted questionnaires. Should a measure, or a system, of routine outcomes measurement be introduced into mental health services, then computers represent a potential labour-saving and economically efficient tool for the collection of data (Butcher, 1994; Marks, 1996; Yokley, Coleman & Yates, 1990). On the other hand the question of whether the use of such technology has a differential effect on the validity and reliability of the data obtained needs to be carefully considered.

Computers are being increasingly utilized as assessment tools and new applications are being created regularly (Kirkby & Lambert, 1996; Kobak, 1996). At least half of the measures used in this study are already available in computer format: the SF-36 (under evaluation), the Life Skills Profile, and the Health of the Nation Outcomes Scale (HoNOS) (Marks, 1996). Factors that need to be considered before using computers for such a purpose include:

- Whether the introduction of such technology impedes self-disclosure (Skinner & Allen, 1983; Weisband & Keisler, 1996);
- The reactions by consumers and service providers to computer forms (Bagley & Genuis, 1991; Locke & Gilbert, 1995; Murrelle, Ainsworth, Bulger et al., 1992; Russell, Peace & Mellsop, 1986; Sawyer, Sarris & Quigley, 1993; Sawyer, Sarris, Quigley et al., 1990; Spinhoven, Labbe & Rombouts, (1993);
- Issues of comparability between computer-administered forms and pen-and-paper versions in a study that used both; and
- Service provider access to computers.

The present study *was not* a trial of the efficacy of computer administration of the selected outcomes measures, nor does it attempt to address all of the issues discussed above. Only two factors influenced the choice of administration format for each participant: personal preference and practicality.

2.0 Aims and Research Questions

2.1 Aims

Six measures have previously been identified as potentially useful for the routine assessment of consumer outcomes in mental health services (Andrews et al., 1994). Tables A5 and A6 provide descriptions of these measures. The aims of this project were:

- To assess the level to which the six selected outcomes measures are *feasible (that is, applicable, acceptable and practical), valid, reliable, and sensitive to change*, in three selected *mental disorders (schizophrenia, affective disorder and anxiety disorder)* and in three *clinical practice settings (public sector, private sector and general practice)*;
- To subsequently *identify a measure or measures* which may be suitable for use as a *national measure of consumer outcomes* to be pilot tested at 20-40 pilot implementation sites; and,
- To *articulate the potential difficulties in the implementation of the measurements* at the pilot sites, and to *identify strategies* to overcome these.

2.2 Research Rationale and Hypotheses

The research questions and hypotheses addressed in this study (see Table A7 for summary) are organised under three headings: feasibility issues; the validity and reliability of the selected measures as *measures of mental health status*; and, the validity and reliability of the selected measures as *measures of change in mental health status*.

2.2.1 Feasibility

Andrews et al. (1994) identified three dimensions of feasibility: applicability, acceptability, and practicality. The 'applicability' of a measure was defined as the degree to which a measure addresses dimensions of importance to the consumer (symptoms, disability and consumer satisfaction); is useful for service providers in formulating and conducting treatment (a useful measure is one which is multi-dimensional, yet also provides a summary of information in the form of a single score); and allows for the aggregation of data in a meaningful way to meet the purposes of service management.

The 'acceptability' of a measure describes the ease with which a consumer or clinician can use a particular measure (i.e., user-friendliness). Practicality relates to: cost of implementation; training requirements; and the complexity of scoring, reporting and interpreting the data (Andrews et al., 1994, p.p. 30-31).

Feasibility in this study was approached in two ways: in relation to the use of the six selected measures, and in relation to implementing any system of routine outcomes assessment.

In addressing the feasibility of the six selected measures a short range of dimensions were chosen for investigation which correspond to some of the aspects of applicability and acceptability described by Andrews et al. (1994). (Issues relating to practicality were primarily addressed through the focus groups consultations; see below). For consumers we were interested in the following questions in relation to each measure:

- Was the language of the measures difficult to understand?
- Did the questions address issues of importance and relevance to the consumer?
- Was the measure considered useful as a means for assessing treatment progress?

What other issues did consumers identify which might impact upon the feasibility of implementing this measure?

Table A7: Overview of Research Questions and Hypotheses

<p><i>Feasibility</i></p> <p>Feasibility of using the six selected measures (from Utility Questionnaires):</p> <ul style="list-style-type: none"> • How was each measure responded to in relation to the identified individual dimensions of feasibility and across all dimensions (i.e., the ease with which the language of the measures could be understood; the importance and relevance of the questions used, to the consumer; the relevance and effectiveness of the measure for service providers; and the usefulness of the measure as a means for assessing treatment progress)? • How was each measure evaluated in comparison to the other measures? • Were there differences in the evaluations of the measures according to the diagnosis of the consumer or the practice setting? • What related issues were raised by participants which could potentially impact upon the feasibility of implementing any of the measures? <p>Feasibility of implementing routine outcomes assessment (from focus group data):</p> <ul style="list-style-type: none"> • The potential benefits and impediments to outcomes assessment? • Specific feasibility issues? • The structure and content required of outcomes measures? and, • Other related issues?
<p><i>The Selected Measures as Measures of Mental Health Status</i></p>
<p><i>Validity</i></p> <ul style="list-style-type: none"> • That significant relationships will be observed between the scores of the measures that purport to assess similar dimensions of mental health. • That any differences observed between the sub-groups of participants on the selected measures will be consistent across the measures to the degree that the measures are related. • That differences in scores for the selected measures across diagnostic groups with those differences observed on alternative measures of mental health.
<p><i>Reliability</i></p> <ul style="list-style-type: none"> • That test-retest coefficients calculated for the consumer measures will be comparable to those reported in the research literature. • That internal consistency coefficients calculated for the measures will be comparable to those reported in the research literature. • That there will be significant associations between scores derived from the service provider measures and scores derived from adaptations of these same measure which have been completed by consumers.
<p><i>The Selected Measures as Measures of Change in Mental Health Status</i></p>
<p><i>Validity</i></p> <ul style="list-style-type: none"> • That significant relationships will be observed between the 'change scores' of those measures which purport to assess similar dimensions of mental health outcomes. • That significant relationships will be observed between participants' direct assessments of global change in their mental health over a two to three month period, and the scores derived from the selected measures.
<p><i>Reliability</i></p> <ul style="list-style-type: none"> • That strong internal consistency will be observed for the change scores of the measures.

For service providers, a similar set of questions was investigated, viz.:

- Was the language difficult to understand?
- Were the questions relevant to measuring outcomes?
- Was the measure an effective means of measuring outcomes?
- Was the measure useful?
- What other issues did service providers identify which might impact upon the feasibility of the measure?

Furthermore questions relating to the relevance and effectiveness of the measures were framed in terms of the needs of the service providers' immediate clients (i.e., those who were being assessed as part of the study), the needs of other people with a similar diagnosis and the needs of all people with mental illness.

To answer these questions, all participants were invited to complete a 'Utility Questionnaire' following the administration of each of the selected measures (see Method). Quantitative and qualitative response formats were utilised in collecting the relevant data. No specific hypotheses were tested in this regard. Rather the aim was to investigate:

- How each measure was responded to on the individual dimensions and across all dimensions described above (represented by a Global Utility Score; see Method);
- How each measure was evaluated in comparison to the other measures;
- Whether there were differences in the evaluations of the measures according to the diagnosis of the consumer or the practice setting; and
- The range of related issues raised by participants that could potentially impact upon the feasibility of implementing any of these measures.

In addition to asking study participants their views on the selected measures, six (6) peak service provider organisations were invited to provide submissions about the selected measures in respect of to the *face validity* of the measures and the anticipated usefulness of the measures. The aim of this was to introduce another perspective into the discussion of outcomes assessment.

With regard to investigating general issues relating to the feasibility of routine outcomes assessment, eight focus groups were convened (four for consumers and four for service providers). Again no specific hypotheses were tested. Rather the aim was to provide a forum for consumers and service providers to discuss the implementation of standardised measures for the assessment of consumer outcomes. The principal question explored by the focus groups was: "What is the feasibility of routinely using standardised measures to measure the outcomes for consumers of mental health services?". In each group participants were presented with a short series of questions to initiate and encourage discussion (see Method). The foci for analyses and reporting of data were:

- The potential benefits and impediments to outcomes assessment;
- Specific feasibility issues;
- The structure and content required of outcomes measures; and
- Other related issues.

2.2.2 The Selected Measures as Measures of Mental Health Status

In order for a measure to provide accurate information regarding mental health outcomes it must first be shown to provide valid and reliable information regarding a person's mental health status. The six selected measures have each been examined to varying degrees in terms of their properties in providing accurate and meaningful information about consumers' mental health status (see above). Our first aim in this regard was to collect data from a heterogeneous Australian sample and compare common validity and reliability indices with those that have been previously reported. The next aim was to look at how the measures reported the mental health status for various groups of consumers, and to investigate whether differences observed on one measure were also observed on other measures. As well we compared mental health data from the consumers' perspective and that of their service provider, and sought to determine those issues that might influence the validity and reliability of any decision-making processes that would be based on information derived from these measures.

Validity

Do the selected measures provide meaningful information about consumers' mental health status? Each of the measures purports to assess different dimensions of mental health status. In many cases they assess similar dimensions of mental health and disability. We would expect the measures to provide fairly similar information regarding the mental health status of consumers to the degree that the measures appear related. And moreover, that they would describe any differences observed between sub-groups (as defined by *diagnostic group* and *practice setting*) in similar ways. On the other hand, even where the dimensions of mental health assessed appear to be the same there may be only a small convergence in score depending on whether the information is collected from the consumer or the service provider, as was discussed in section 1.2.3.

The issue of validity was approached in two ways. The first was to investigate the association between scores derived from the measures. Analyses were undertaken to determine the degree of convergence and divergence between the scores derived from the consumer and service provider measures, respectively. As well we examined the degree of convergence between ratings collected from the consumers and service providers.

The second approach was to look closely at the characteristics of the sample of consumers and service providers, and to describe any differences in mental health status ratings across different groups. As described below, a very heterogeneous group of people participated in the study. While controlling for *diagnostic group* and *practice setting*, in general, the researchers had little control over the final composition of the group. We did have certain expectations given foreknowledge of the settings and likely participants. However, the final analyses depended on whoever volunteered for the study. The approach was to look at patterns of differences in mental health status between sub-groups of the sample defined by *diagnostic group* or *practice setting* as reported by the selected measures. To the extent that the measures converge we would expect similar findings on each measure.

As an additional check, we also used alternative methods (that is, *visual analogue scales* and *modified clinical global impressions scales*) to describe the mental health status of the sub-groups and compared these findings with those obtained from the selected measures. We expected that any observed differences in mental health across *diagnostic groups* and *practice settings* reported on the alternative measures, should correspond to those observed on the scores of the selected measures. We also compared the results obtained on the selected consumer measures to those obtained on the service provider measures as an additional method of determining the level of correspondence between the two sets of measures.

The hypotheses tested in relation to the validity of the measures as measures of mental health status were:

- That significant relationships will be observed between the scores of the measures that purport to assess similar dimensions of mental health.
- That any differences observed between the sub-groups of participants on the selected measures will be consistent across the measures to the degree that the measures are related.
- That differences in scores for the selected measures across diagnostic groups will be consistent with those differences observed on alternative measures of mental health status.

Reliability

Do the selected measures provide reliable information about consumers' mental health status? As with the validity analyses we would expect that the measures would demonstrate reasonable levels of reliability in light of previous results. The reliability of the measures as measures of mental health status was assessed using: test-retest methods (for consumer measures only); internal consistency (all measures); and inter-rater reliability (service provider measures only).

The hypotheses tested in relation to the reliability of the measures as measures of health status were:

- That test-retest coefficients calculated for the consumer measures will be significant and comparable to those reported in the research literature.

- That internal consistency coefficients calculated for the measures will be high and comparable to those reported in the research literature.
- That there will be significant associations between scores derived from the service provider measures and scores derived from adaptations of these same measures which have been completed by consumers.

2.2.3 The Selected Measures as Measures of Change in Mental Health Status

As described above, one of the chief purposes of using a standardized measure of outcomes is to improve the ability of services to accurately and reliably detect clinically significant change in a person's mental health over time (Guyatt et al., 1989; Guyatt et al., 1987). This is usually termed the *sensitivity to change* of a measure, although it shares similarities to the definitions of validity and reliability. In this study we address 'sensitivity to change' as the validity and reliability of the measures as *measures of change*.

Validity

Usually change is investigated by comparing the effects of a specified intervention of known utility with that of a control condition. However, the present study was not designed to investigate treatment effects, but rather the applicability of certain measures, to measure change. The sample of people who participated was heterogeneous, including people with different diagnoses (including comorbidity), illness severity and histories of illness, etc.. The sample also represented a range of intervention modalities and stages of intervention progress. Moreover the voluntary nature of the exercise meant that there was little control over who was actually re-assessed after two or three months.

In assessing change we examined the convergence between *global ratings of change* and changes in scores for the measures, using repeated measures, analyses of variance and regression techniques. We also examined the relationship between simple change scores calculated from the measures (Score for Time 2 - Score for Time 1) to investigate the degree to which changes on one measure related to changes observed on other measures. As well the relationships between changes observed on the measures and participants' assessments of *general* or *global* change in mental health (i.e., for better or worse, or not at all) were examined.

Both procedures represent an assessment of the validity of the measures as measures of change in mental health status. Change on one measure should correlate significantly with change on another measure to the degree to which the two measures assess similar aspect of mental health status. Moreover, there should be a relationship between a person's global assessment of change in their mental health over a three month period, and change as recorded on the selected measures.

Reliability

Examining the consistency of change in scale scores within each measure assessed the reliability of the measures as measures of change in mental health status. That is, if the items of a measure are supposed to be related, then change on one item should be consistent with change on related items. Calculating internal consistency coefficients for the scales and the relevant sub-scales using the raw change scores assessed the consistency of change across the items of the measures.

The following hypotheses were tested in relation to the validity and reliability of the selected measures as measures of change in mental health status:

Validity

- That significant relationships will be observed between participants' direct assessments of global change in their mental health over a two to three month period, and the scores derived from the selected measures.
- That significant relationships will be observed between the 'change scores' of those measures which purport to assess similar dimensions of mental health outcomes.

Reliability

- That strong internal consistency will be observed for the change scores of the measures.

2.2.4 The Effect of Using Computers to Collect Outcomes Data

As stated above, it was not one of the primary aims of this study to examine the efficacy of using computers to collect outcomes data. However, it was decided to utilise computer administrations and to ascertain whether this had any effect on the nature of the data obtained. Data derived from pen-and-paper versions of the measures was compared to data collected with computer versions to determine whether there were any systematic differences. Upon completion of the computer versions, participants were asked about their attitude towards using computers for such a purpose. Furthermore, as part of their submission the Peak Service Provider Organizations were asked to comment upon the use of computers in the outcomes assessment process. In accordance with the literature reviewed above it was hypothesised:

- That the use of computer versions of the selected measures will not unduly influence the results obtained.
- That participants who use the computer versions of the measures will not indicate negative responses to the use of computers to collect outcomes data.

3.0 Method

3.1 Design

Data was collected from consumers in three stages (Time 1, Time 1(Retest) and Time 2) during the period of May to December 1996. Time 1 data was collected at the relevant clinical setting at a time appropriate for the consumers. Time 1(Retest) data was collected an average of 24 hours (mean=1 day) after the initial assessment. Time 2 data was collected approximately three months after the initial assessment (mean for consumers=96 days) (see Table A8).

Table A8: Sequence for Administration of Selected Outcomes Measures and Other Instruments for the Consumer Outcomes Study

<i>SERVICE PROVIDER</i>	<i>CONSUMER</i>
1. Initial administration of selected service provider rated measures.	1a. Initial administration of selected consumer rated measures. 1b. Re-administration of selected consumer rated measures (1-2 days)
2. Follow-up administration of selected service provider rated measures (2-3 months)	2. Follow up administration of selected consumer rated measures (2-3 months)

Data was collected from service providers, over the same time period as above, in two stages. Time 1 data was collected at the relevant clinical setting at a time negotiated with the research assistant. Follow-up data was collected approximately 3 months after this initial assessment (mean for service providers=89 days). Every effort was made to collect the service provider assessments within 1-2 weeks (mean at Time 1=5 days) of the consumer assessments.

Consumers completed two of the three selected consumer measures (i.e., BASIS-32, MHI or SF-36). Service providers completed two of the three selected service provider measures (i.e., HoNOS, LSP or RFS). A sequence for the selection of the measures was devised. This sequence controlled the combination and ordering of the questionnaires. There were eighteen possible combinations and selection of the specific measures for any particular participant was dependent only on their rank in the order of recruitment and the correspondence of this to the administration sequence. Consumers were also invited to complete adapted versions of the measures completed by their service providers (see below).

All participants received \$20.00 for each session of involvement in the study (including the focus groups).

3.2 Sample

Service Providers: Sixty-five service providers were recruited from three clinical settings: (*general practice* (n=20), *private psychiatry* (n=16) and *public psychiatry* (n=29)), and completed a total of 176 assessments. Table A9 displays information concerning the range of professional groups represented (see also Figure 1, Appendix 4). All service providers were from the Brisbane and Ipswich metropolitan areas. General practitioners were recruited through a community-based group practice coordinated by a university medical school, and through a general practitioners' special interest group in psychiatric issues.

Table A9: Distribution of Professional Groups by Number of Assessments and Practice Setting

Professional Group	Number	Number of Assessments	Number of Professional Groups Across Practice Settings		
			Private	Public	General Practice
General Practitioners	20	62			20
Psychiatrists (Private)	10	28	10		
Psychiatrists (Public)	2	5		2	
Psychologists	2	8	1	1	
Social Workers	6	9		6	
Nurses	15	42	5	10	
Occupational Therapists	5	12		5	
Social Work Associates	4	8		4	
Other (i.e., Resident Medical Officer)	1	2		1	
Total	65	176	16	29	20

Table A10a: Distribution of Consumers across Diagnostic Groupings and Practice Settings

Diagnostic Groups	Practice Settings			Total
	Private	Public	General Practice	
Anxiety	15	4	20	39
Affective	29	17	32	78
Schizophrenia	8	47	11	66
Total	52	68	63	183

Table A10b: Distribution of Consumers who Remained and Dropped out of the Study across Diagnostic Groupings and Practice Settings

	Remained	Dropped Out	Total
Diagnostic Groups			
Anxiety Disorders	27	12	39
Affective Disorders	50	28	78
Schizophrenia	36	30	66
Practice Settings			
Private Psychiatry	20	32	52
Public Psychiatry	40	28	68
General Practice	53	10	63
Total	113	70	183

Table A11: Consumer Demographic Data

Age (in years)	Mean	42
	SD	13
	Minimum	17
	Maximum	87
Gender	Male	85 (46%)
	Female	98 (54%)
Australia as Birth Place(%)	Yes	81
	No	19
Aboriginal or Torres Strait Islanders (%)	Yes	1
	No	99
Other Birth Place (%)	New Zealand	6
	United Kingdom	9
	Asia	1
	Europe	2
	Pacific Islands	-
	Americas	-
	Africa	1
Total	19	
Marital Status (%)	Married	35
	De Facto	6
	Divorced	15
	Widowed	4
	Never Married	40
Employment (%)	Employed	23
	Unemployed	55
	Full-time Student	2
	Homemaker	20
Duration of Absence from the Workplace (% of total sample)	<1 month	1
	1-3 months	2
	3-6 months	5
	6-12 months	4
	1-2 years	6
	>2 years	45

Service providers in the *private psychiatry* setting were primarily recruited from two private psychiatric hospitals. Four psychiatrists in private community practice also took part in the study. As Table A10a reflects, service providers in the *private psychiatry settings* were generally providing inpatient services to people with anxiety and affective disorders.

Service providers in the *public psychiatry setting* were recruited from two large community mental health clinics providing community based case management and outpatient services, and from an acute inpatient ward of a large public hospital.

Service providers were approached in person and provided written and verbal information about the study. Their involvement in the study was contingent upon the consent of the consumers who they nominated for participation. The decline in service provider numbers at Time 2 was primarily due to consumers choosing not to continue with the study.

Consumers: Service providers nominated up to four consumers who displayed symptomatology consistent with one of three *diagnostic groupings* in accordance with ICD-10 primary care criteria (World Health Organization, 1992): that is, *anxiety disorders*, *affective disorders* or *schizophrenia*. It was decided to allow service providers to assess a maximum of four consumers, as the research team was conscious of imposing upon the service providers' time. In some circumstances certain service providers did assess more than four of their consumers. On average the service provider to consumer ratio was approximately 1:3. Table 10a and Figure 2 describe the distribution of consumers across *diagnostic groups* and *practice settings* at Time 1. Table 10b describes the distribution across diagnostic groups and practice settings of consumers who remained in the study and those who dropped out after Time 1. The sampling procedures resulted in an uneven distribution of participants across *diagnostic groups* and *practice settings*.

One hundred and eighty-three consumers participated at Time 1 (see Table A10a). Of the initial participants, 158 (86%) participated at Time 1 (Retest). One hundred and thirteen (61%) of the initial participants participated at Time 2. Socio-demographic characteristics, including age, gender and employment status of these consumers are included in Table A11. Further descriptions of the characteristics of participants are included in Section 4.1.

The consumers were generally middle-aged and of Anglo-Celtic ethnic background with most born in Australia. Of those born outside Australia, the United Kingdom and New Zealand were the most frequent birthplaces. Approximately one third of the consumers were married, but approximately 40% indicated never having been married. Seventy-five percent of the sample were currently unemployed or classified themselves as 'homemakers'. Seventy percent of these people indicated being out of employment for over two years.

Representativeness of the Sample: As Andrews et al (1994) state, at any time approximately 25-30% of the Australian population will meet the diagnostic criteria for a mental disorder, many people having more than one disorder: approximately 10% for an affective disorder, 15% for an anxiety disorder, 10% for a substance abuse disorder, and 0.5% for schizophrenia. Of the three percent of the population who could be classified as having a 'serious mental disorder' 10% will have schizophrenia, 44% will have an affective disorder, 25% will have an anxiety disorder, and 21% will have a substance use disorder. Of these people 15% will receive the service of a general practitioner, 30% will receive private psychiatric services, 41% will receive psychiatric services from public facilities, and 13% will receive service from facilities specializing in addictive disorders (see Tables A12 and A13).

Table A12: Prevalence of Mental Disorders in Australia by Severity (Source: Andrews et al., 1994; figures are based on estimates of potential consumers within the Australian population)

Disorder	12 month total	Serious mental disorder ^a	Chronic mental disorder ^b	Mild and transient disorder ^c
Schizophrenia	0.5	0.5	0	0
Any Affective Disorder	9.5	2.1	2.2	5.2
Any Anxiety Disorder	12.6	1.2	2.2	9.2
Any Substance Use Disorder ^d	9.5	1.0	2.2	6.3
Total %	26.8	2.9	5.0	18.8
Total Number (Australia)	4.7M	0.5M	0.9M	3.4M

Notes: ^a 'Serious mental disorder' follows the US National Advisory Mental Health Council (1993) guidelines and includes all schizophrenia and related disorders, all bipolar disorder, 20% of major depressive disorder, and 20% of panic disorder and OCS, together with 10% of social phobia (that comorbid with avoidant personality disorder) and 10% of substance abuse disorder (principally drug dependence). At this level these disorders are chronic and disabling, frequently lead to hospitalization and require treatment by a specialist mental health or addictive service, or by a very experienced general practitioner.

^b 'Chronic mental disorders' are present, like the 'serious mental disorders', throughout the 12 month period and are associated with disability (i.e., GAF ratings of less than 70). They include 25% of all affective disorders, 17% of all anxiety disorders, and 21% of the substance abuse disorders. Substance dependence is treated by specialist drug and alcohol services in all Australian states but general practitioners have an important role to play in treating substance abuse. General practitioners should be competent to manage the other chronic mental disorders.

^c 'Mild and transient disorder'; 70% of all mental disorders are either so mild as not to handicap or else they remit without treatment within the year. There are many human services outside the funded health system who already attend to such patients (Stress Management Programs, Counselors, Psychologists and Clergy; self help groups and programs such as AA, other non-government organisations) but general practitioners and specialist mental health services are often required to see such patients. They should either advise appropriately or refer to non-medical community services.

^d As persons with one disorder tend to suffer another and the more severe the disorder the greater this possibility, the totals have been discounted for this comorbidity; by multiplying the total by 0.6 in respect to serious mental disorders, 0.75 in respect to chronic mental disorders, and by 0.9 in respect to mild and transient mental disorders.

Table 13: Relation between individuals with severe, chronic, mild or transient mental disorders, and the number (and %) of patients in Australia in these categories being seen by various sectors of the medical workforce for their disorders (all numbers as 000's). (Source: Andrews et al., 1994)

	Severe	Chronic	Mild or Transient	Total: any diagnosis
Total Numbers (000's)	510	880	3290	4680
Public mental health service	94 (41%)	33 (6%)	—	127 (9%)
Private psychiatrists	69 (30%)	92 (17%)	—	161 (12%)
General practitioners	35 (15%)	402 (72%)	560	997 (74%)
Addictive services	30 (13%)	30 (5%)	3	63 (5%)
Total patients seen	228	557	563	1348
Per cent being seen	45%	63%	17%	29%

The study sample was not strictly representative of the population of people receiving and providing mental health services in Australia. For instance, in this study there was a greater number of people with *schizophrenia* than *anxiety*, and a greater number of participants representing the *public* sector as opposed to the *private* sector. Recruitment of all participants was on a volunteer basis, and there was a strict time limit to the study, which placed constraints upon the sampling procedures. It must be kept in mind that this is largely exploratory research that sought to gather and analyse the opinions of consumers and service providers in relation to the measurement of consumer outcome. As will be discussed in the focus group methodology, having a group of people who feel that they are well enough to participate in such a study and provide forthright views on the matter of routine outcomes assessment is extremely useful. Including people from such disparate backgrounds, representing a range of experiences in relation to mental health services, further enhances the range of information gathered. The information gathered is likely to be of a nature to inform future efforts to implement outcomes assessments as part of routine clinical practice.

3.3 Procedure

Service Providers Time 1: At Time 1 service providers were approached by a researcher and invited to participate in the study. Verbal and written information about the study was provided. Service providers were invited to nominate up to four consumers whom they believed would be interested, and capable of participating, in the study, and who met the ICD-10 primary care diagnostic criteria for one of the specified *diagnostic groups: anxiety disorders, affective disorders or schizophrenia*. Upon obtaining the consent of the consumers, service providers were provided with instructions pertaining to the completion of the measures and a research assistant was available for further consultation.

Service Providers Time 2: Three months after completion of Time 1 measures, service providers were provided with follow-up measures to complete and return. Sets of measures included identical outcomes measures to those completed at Time 1, a modified information sheet, questions 2 and 3 of the Modified Clinical Global Impressions Scale, and a modified version of the outcomes measure utility questionnaire.

Consumers Time 1: At Time 1 nominated consumers were approached and invited to participate in the study. Verbal and written information about the study was provided and informed consent to participate was obtained. Consumers were then provided with instructions pertaining to the completion of the measures. Sets of measures included two of three selected measures of outcomes, an information sheet, visual analogue scales and utility questionnaires. Depending on the capability and wishes of the consumer, assistance was available to aid in the completion of the measures.

Consumers were also invited to complete adapted versions of the measures completed by their

service providers (see below). One hundred and eleven consumers agreed to do so. Table A14 displays the number of consumers who completed each of the measures and the distribution in completed questionnaires across diagnostic groups and practice settings.

Table A14: Distribution of Consumer-Versions of the Service Provider Measures

Diagnostic Groups	Private Psychiatry	Public Psychiatry	General Practice	Total
Anxiety	12	2	9	23
Affective	26	7	19	52
Schizophrenia	6	25	5	36
Total	44	34	33	111

Consumers Time 1(Retest): Within three days ($M= 1$ day) of completion of Time 1 measures, consumers were invited to complete the Time 1(Retest) measures. Sets of measures were identical to those completed at Time 1, with the exception that the Demographic Survey was omitted. Once again, these measures were generally completed in the presence of the research assistant. The exceptions were a small number of participants who were unable to attend for follow-up and who completed the measures at home.

Consumers Time 2: Three months following completion of Time 1 measures, all consumers were contacted and invited to complete the final set of measures. Those participants who agreed to undertake to complete the measures at Time 2 did not have access to their previous responses to the selected measures.

Method of Questionnaire Administration: Questionnaires were administered using either pen-and-paper or computer format. For the purpose of computer administration the instruments were redrafted using the Jetform computer program. Measures were presented on a laptop computer. The individual participants determined the choice of format (pen-and-paper versus computer).

3.4 Questionnaires

The questionnaires used in this study fall into a number of categories: selected measures for consumers; selected measures for service providers; other consumer measures; and other service provider measures (see Tables A5, A6 and A15). Copies of all measures used in this study are included in Appendices 5 and 6. Two categories of scores are derived from the selected consumer and service provider measures: *composite* scores (sum, total or average scores across all items of the measure) and *scale* scores representing the individual sub-scales (or a combination of sub-scale scores) of each measure.

The selected consumer and service provider measures have already been described in some detail in Section 1.4.

Given the constraints in training up to 40 service providers in the standardized use of the HoNOS, service providers were provided with the glossary of the scale and were able to consult with the research assistant.

Other Consumer Measures

Other measures administered to consumers were:

Information Sheet: This survey was completed by consumers and included questions that related to age, gender, ethnicity, language and occupation.

Visual Analogue Scales: Visual Analogue Scales (VAS) were used to assess the severity of illness for consumers. They have been used in a range of situations (Aitken, 1969; Bond, Shine & Bruce, 1995; Luria, 1975; Maruff, Wood, Currie, et al., 1994; McCormack, Horne & Sheather, 1988; Pfenning, Cohen & van der Ploeg, 1995; Priebe & Gruyters, 1994; Zealley & Aitken, 1976). The structure of

the scales used in this study was similar to that described by Cella & Perry (1986) for pen-and-paper testing, and to those used by Maruff et al. (1994) for computer-administered VAS. The program for the computer administered versions of the VAS was developed by Dr Paul Maruff and associates.

Participants were asked to rate themselves over the last seven days according to: (1) perceived mental *well-being*; (2) perceived *depression*; (3) perceived *anxiety*; and, (4) perceived *distress*. Participants were asked to respond to each question by placing a vertical mark on a 100mm horizontal line (or, in the case of the computer version, they moved a cursor along a 100 pixel line) which separated two anchor statements. For the question relating to mental *well-being* these anchor statements were: "*The worst that it has ever been*", and "*The best that it has ever been*". With regards to *depression*, *anxiety*, and *distress*, the anchor statements were "*The least depressed/ anxious/ distressed I have ever felt*", and "*The most depressed/ anxious/ distressed I have ever felt*".

The score was calculated as the number of millimetres or pixels from the left-hand anchor point to the position along the line where the participant placed their mark.

Scores ranged from 0 to 100 for the four scales and 0 to 400 for the total score. The greater the score the *more* of the relevant dimension (i.e., *well-being* or *depression*) that the consumer had experienced in the preceding seven days.

Global Change Rating (consumers): At the three-month follow-up all participants were asked, "*Compared to how you felt three months ago, how much have you changed?*" Participants responded by indicating a point on a seven-point scale ranging from 'Extremely Better' through to 'No Change' to 'Extremely Worse'. This is based on the Modified Clinical Global Impressions Scale (see below).

Table A15: Other Measures of the Consumer Outcomes Project

Category		Variable Types and Names
Demographic and Information Measures	• Consumer	<p>Eleven indices:</p> <ul style="list-style-type: none"> • Age • Gender • Born in Australia • Aboriginality • Place of Birth (Australia, Other) • Language • Marital Status • Employment • Student Status • Nature of Employment • Duration of Unemployment
	• Service Provider	<p>Eight indices:</p> <ul style="list-style-type: none"> • Professional Description <ul style="list-style-type: none"> • Consumer's Primary Diagnosis • Consumer's Secondary Diagnosis • Hospital admissions in last 6 months • Days spent in hospital in last 6 months • Drug/Alcohol Problem • Length of time as Service Provider to Consumer
Change Measures	• Consumer	<ul style="list-style-type: none"> ◇ Visual Analogue Scales <ul style="list-style-type: none"> • Total • Well Being • Depression • Anxiety • Distress ◇ Change Measure ◇ Modified Clinical Global Impressions Scale (Parts A & B)
	• Service Provider	
Utility Measures	• Consumer	<ul style="list-style-type: none"> • Utility Questionnaire • Participant Response Sheet • Utility Questionnaire • Measure of Consumer Outcomes Questionnaire

Utility Questionnaire - Consumer Version: Following the administration of each selected measure, consumers were asked a short series of questions about the measure. The four questions were:

1. Were the words used in the questions difficult to understand?
2. Did the questions ask you things that are important to you?
3. Did the questions ask you things that relate to how you feel?
4. Is this questionnaire a useful way for your doctor or specialist to find out whether you are getting better?

For questions 1-3 consumers responded by indicating: none, some, most or all. For question 4 consumers responded by indicating: not useful at all, slightly useful, reasonably useful or very useful. A Global Utility Rating was computed by summing the individual items (excluding Question 1; see Section 4.2.1). Upon completion of the second utility questionnaire participants were asked which of the two measures of outcomes they preferred, and, if the computer-administered format was used, how they felt about using a computer to provide such information.

Other Service Provider Measures

Other measures administered to service providers were:

Diagnostic Screening Sheet: Prior to referrals being followed up, service providers were asked to read the ICD-10 Primary Care Diagnostic Criteria (World Health Organization, 1992) for the relevant disorder to confirm that the consumers met such criteria. An example of a screening sheet is included in the Appendix 6. The format of the screening sheets is based upon that used in the Climate Computer Program developed by Professor Gavin Andrews and associates.

A primary care diagnostic sheet was prepared for each of the major disorders represented in this study. These included acute and transient psychotic disorder, adjustment disorder, bipolar disorder, depression, generalised anxiety disorder, obsessive compulsive disorder, panic disorder, phobic anxiety disorder, schizoaffective disorder, schizophrenia, and severe stress disorders.

Information Sheet: Information was collected from service providers relating to professional background and the consumers' treatment history.

Modified Clinical Global Impressions (CGI) Rating: This scale was designed to assess the service providers' impressions of the severity of the consumers' mental illness (at Time 1) and the degree of change in the consumers' condition (at Time 2; referred to as *Global Change Rating (service providers)*). The CGI includes seven categories ranged from 'normal, not at all ill' to 'among the most extremely ill patients' (for Time 1 and 2 ratings). The *Global Change Rating* includes seven categories ranging from 'Extremely Better' through to 'No Change' to 'Extremely Worse'. This scale is based upon the Clinical Global Impressions Scale (Guy, 1976; NIMH, 1985) which is a measure of severity of illness, and of change in severity, widely used in pharmacological trials.

Utility Questionnaire - Service Provider Version: Following the administration of each selected measure, service providers were asked a short series of questions about the measure:

1. Were the words used in the questions difficult to understand?
2. Were the questions relevant to measuring the treatment outcomes for your client?
3. Were the questions relevant to measuring the treatment outcomes for people with this particular diagnosis (i.e., that of your client), in general?
4. Were the questions relevant to measuring the treatment outcomes for all people with mental illness?
5. Is this questionnaire an effective way of measuring treatment outcomes for your client?
6. Is this questionnaire an effective way of measuring treatment outcomes for people with this particular diagnosis (i.e., that of your client), in general?
7. Is this questionnaire an effective way of measuring treatment outcomes for all people with mental illness?
8. Is this questionnaire a useful way of measuring treatment outcomes for all people with mental illness?

Each question was responded to using a four-point scale similar to that described for the consumer utility questionnaires. A Global Utility Rating was computed by summing the individual items (excluding

Question 1; see Section 4.2.4). After completion of the second utility questionnaire participants were asked which of the two measures of outcomes they preferred, and, if the computer-administered format was used, how they felt about using a computer to complete the questions.

3.5 Focus Groups

Focus groups were used to collect information that would assist the consultancy in shaping recommendations regarding the implementation of the selected measures of consumer outcomes.

The aims of the focus groups were:

1. To discuss the routine use of standardised measures for the assessment of consumer outcomes in mental health; and
2. To raise issues relating to the implementation of procedures for the assessment of consumer outcomes in mental health.

A focus group is a small group discussion that gathers together people from similar backgrounds or experiences to discuss a specific topic. The group is guided by a moderator/facilitator who assists the group participate in a 'natural' discussion. The discussion is 'focused' in that it aims to cover only a small range of issues in some detail while respecting the participants' feelings, attitudes and opinions.

A moderator is aided by a prepared question guide which outlines the major questions to be presented to the group for discussion, but is flexible and adaptable to the needs and circumstances of the groups. An observer/note-taker records the key issues and other factors that may influence the interpretation of the information. In the case of the consumer focus groups, the moderator was a Consumer Consultant, Mr Tony Hennessey.

Focus groups are developed to explore people's beliefs, attitudes and opinions on a particular issue or topic. In this regard, they can provide insights into the beliefs/opinions/ideas of the group's participants. However, they *cannot* indicate how general these opinions are in the community as the groups do not necessarily constitute a *representative sample* of the target population. Nor can the information collected be used to build up a detailed picture of specific beliefs, although they may give indications for further investigation. Focus groups have been used for a variety of purposes in a range of research designs in the social and medical sciences (see Albrecht, Johnson & Walther, 1993; Dawson, Manderson & Tallo, 1993; Epstein & Wadsworth, 1994; Gross, Fogg & Conrad, 1993; Knodel, 1993; Morgan, 1993; Pulice, McCormick & Dewees, 1995).

The advantages of focus groups are that they produce a lot of information quickly and cost effectively; allow in-depth exploration of specific issues of concern; assist in collecting information from stigmatized communities; can be undertaken without a great deal of staff training; can assist in the collection of information on attitudes/opinions not revealed in a survey questionnaire; and, allow follow-up and clarification of responses if required.

The limitations of focus groups include: information gathered cannot be generalised to a wider community; social compliance effects and other group processes may interfere with the information collected; there can be negative influences on the part of the moderator; they can be of limited value in exploring complex beliefs; and they can paint a picture of what is socially acceptable.

Focus groups are rarely composed of representative samples from the target population (Fuller, Edwards, Vorakitphokatorn & Sermsri, 1993). In the present study the 'pool' of participants came from a sample of people who had already volunteered to take part in the Consumer Outcomes Study. All participants were invited to be involved in the groups. Groups were limited to 12 participants in each. In this regard, the sample was not representative of either consumers or service providers in the wider community, nor of the original sample of participants. Groups were organised according to the practice setting from which participants were drawn.

A separate focus group convened for representatives of support groups for people with mental

illness in the Brisbane region. The meeting was a useful opportunity to explore different perspectives on the consumers' experience and reaction to the issues of outcomes assessment.

Eight discussion groups were convened: four for service providers; three for consumers; and an additional group for members of local consumer support groups who were not otherwise involved with the study. An extra group for service providers in the public psychiatry setting was also convened. Twenty-six service providers participated (*general practice* = 6; *private psychiatry* = 7; *public psychiatry* = 13). Twenty consumers participated in three groups (*general practice* = 7; *private psychiatry* = 5; *public psychiatry* = 8) and seven members of local consumer support groups participated in a fourth group.

Prior to becoming involved, all potential focus group participants were provided with an information sheet that described the aims and procedures for the meeting. A 'question line' was developed and was presented to the participants during the discussion sessions. The question line was developed as follows:

- An initial set of questions was developed by the research team;
- The consumer focus group question line was presented to the consumer Consultant for refinement;
- The question line was refined following each of the focus groups to ensure that the questions asked were relevant and effective in highlighting the range of opinions and attitudes of consumers and service providers concerning outcomes measurement.

The focus group discussions were recorded through note taking and audiocassette recording. A research assistant acted as an observer, and was responsible for the recording of the session. Cassette recordings of the proceedings were transcribed to facilitate analysis of the information. Immediately following a debrief with the focus group participants, the Moderator and Observer (and Consumer Consultant in the case of focus groups convened for consumers) met to discuss the outcomes of the session and to compile impressions and notes, as well as to discuss potential modifications to the process.

Analysis of the focus group data was an ongoing process. It involved all data that had been collected in note form and recorded by audiocassette. All members of the research team including the Consumer Consultant were involved in the analysis process. A copy of the Consumer Consultant's report on the focus group process is included in Appendix 2. It provides a useful resource for future studies that intend to use this methodology.

As was the case for their previous involvement in the study, all participants were reimbursed \$20.00 for their involvement in the focus groups.

3.6 Consultations with Service Provider Peak Bodies

Organisations representing six professional groups were invited to submit an opinion about the six selected measures with respect to the face validity and the overall utility of such measures. The six organisations were the Royal Australian and New Zealand College of Psychiatrists, the Royal Australian College of General Practitioners, the Clinical Board of the Australian Psychological Society, the Australian Association of Occupational Therapists, the Australian Association of Social Workers, and the Australian and New Zealand College of Mental Health Nurses. These organisations had previously been invited to provide submissions for the Andrews et al. (1994) study. Of these bodies, submissions were received from the Royal Australian College of General Practitioners, the Australian Association of Occupational Therapists, and the staff of the Bankstown-Lidcombe Hospital (on behalf of the Royal Australian and New Zealand College of Psychiatrists). A summary of the contents of these submissions is included in the Section 4.2.6.

4.0 Results

The results of the major analyses undertaken are summarized in Table A16. Summaries of the focus group findings can be found in Tables 2.4 and 2.8 included in Section 4.2. Tables and figures of results of all other results are included in Appendices 3 and 4, respectively.

4.1 Participant Issues

Preliminary analyses were undertaken to investigate the severity of illness of the sample of consumers. The results reported here describe the consumers in terms of demographic and service utilisation characteristics. Data for the MHI and SF-36 were compared to published norms for these measures. As well, the differences in Time 1 ratings between those people who remained in the study and those who dropped out were analysed to assess the effect of subject attrition.

4.1.1 Demographic, Comorbidity and Service Utilisation Data

Socio-demographic characteristics, including *age*, *gender*, and *employment status* of the consumers, are included in Table A11 (see section 3).

Age: Significant *age* differences were observed across *gender*, $F(1,179)=6.1$, $p<.05$, *diagnostic groups*, $F(2,178)=6.5$, $p<.01$, and *practice settings*, $F(2,178)=6.0$, $p<.01$. Male consumers ($M=39.5$, $SD=13.3$) were on average, younger than *female* consumers ($M=44.2$, $SD=12.7$).

Consumers in the *public psychiatry setting* were significantly younger ($M=38.0$, $SD=11.1$) than those in the *general practice setting* ($M=45.7$, $SD=16.1$). Consumers in the *schizophrenia group* ($M=37.5$, $SD=11.7$) were significantly younger than consumers in the *anxiety disorders group* ($M=44.0$, $SD=16.0$) and *affective disorders group* ($M=44.9$, $SD=11.8$).

Gender: Differences were observed in the distribution of *gender* across *diagnostic groups* and *practice settings* (see Table 1.3). Females outnumbered males in the *affective* and *anxiety disorders groups*, and in the *private* and *general practice settings*. Males outnumbered females in the *schizophrenia* and *public psychiatry groups*.

Employment: Seventy-five percent of the consumers indicated that they were currently unemployed or occupied as a homemaker. Of these people, 70% had been unemployed for more than two years. The highest frequencies of unemployment were recorded in the *schizophrenia* and *public psychiatry settings*. Eighty-eight percent of people with *schizophrenia* and 86% of consumers in the *public psychiatry setting* stated that they were unemployed or occupied as a homemaker.

Comorbidity: Sixty-four consumers were indicated as having a *comorbid mental health condition*. The disorders recorded included anxiety, affective, substance abuse, personality, organic and eating disorders (see Figure 3). Forty-four percent of this group was represented in the *general practice setting*, as opposed to 28% in the *public* and *private psychiatry setting*, respectively. Of the people who were indicated as having a comorbid condition, 42% had an *affective disorder* as their primary diagnosis, as opposed to 30% of those people with an *anxiety disorder* and 28% with *schizophrenia*. These results need to be considered in light of the fact that no organised assessment for diagnoses were undertaken in relation to comorbidity, and the reliability of the findings may be open to question.

Table A16: Summary of Analyses and Results

Outcome measures	Analyses Undertaken								
	Validity		Reliability			Sensitivity to Change			Feasibility
	Correlational Analyses	Differences Across Diagnoses and Practice Settings	Test-Retest Reliability	Internal Consistency	Inter-Rater Reliability	Convergnence with Change Groups	Correlation between Change Scores and Global Change Ratings	Internal Consistency	Utility Questionnare
BASIS-32	As expected. Good correlations with MHI scales, and mental helath scales of the SF-36.	Differences observed	Consistent with previous reports	Consistent with previous reports	N/A	Slight evidence	Low	Moderate High	Generally good level of acceptance among Consumers
MHI	As expected. Good correlatiom with related scales of BASIS-32. Good correlations with most SF-36 scales	Differences observed	Consistent with previous reports	Consistent with previous reports	N/A	Good	Fair	Moderate to High	Most preferred measure for Consumers.
SF-36	See above	Differences observed	Consistent with Previous results	Consistent with previous results	N/A	Good	Fair	Moderate to High	Generally the least preferred measure for Consumers. Wide preference for a scale that specifically measured "mental health"
HoNOS	Fair correlations LSP and RFS on 'social skills' scales, and all four sub-scale and total scores	Differences observed	N/A	Generally low	Significant but low correlations between Service Provider and Consumer ratings on this measure	Reasonable	Fair	Low to Moderate	Generally good level of acceptance among Service Providers
LSP	Good correlations with RFS as expected.	Differences observed	N/A	Consistent with previous results	Generally low correlations between Service Provider and Consumer ratings on this measure.	Reaasonable	Fair	Moderate to High	Higher level of acceptance in the Public Practice Setting.
RFS	See above.								
	See above	Differences observed	N/A	Acceptable for Global Role Functioning Index	Generally low correlations between Service Provider Consumer ratings on this measure	Slight evidence	Low	Moderate	Least accepted of the three measures in the Public Setting

While only thirteen consumers were considered to have a *comorbid substance abuse disorder*, thirty-three consumers were indicated as having a *drug or alcohol problem*. Across *diagnostic groups*, 13% of people with *anxiety disorders*, 21% of people with *affective disorders*, and 25% of people with *schizophrenia* were indicated as having a *drug or alcohol problem*. Across *practice settings*, 12% of people in the *private psychiatry setting*, 27% of people in the *public psychiatry setting*, and 20% of people in the *general practice setting*, were indicated as having a *drug or alcohol problem*.

Service Utilisation: Approximately 33% of the sample had been admitted to hospital at least once in the six months prior to the study for mental health problems. There was a tendency for these people to be in the *private psychiatry setting* and to have an *affective disorder*. Sixty-seven percent of the sample who were in the *private psychiatry settings* had been admitted to hospital in the preceding six months, as opposed to 29% of people in the *public psychiatry setting*, and 9% in the *general practice setting*. Forty-three percent of the sample with *affective disorders* had been admitted to hospital as opposed to 11% of those with an *anxiety disorder* and 33% of people with *schizophrenia*.

Of those people who had been admitted to hospital in the previous 6 months, people in the *private psychiatry setting* had spent the longest time in hospital with 56% spending more than three weeks in hospital as opposed to 23% and 7% of those people in the *public* and *general practice groups*, respectively. Similarly, 40% of those people with *affective disorders* who had spent time in hospital had spent three or more weeks in hospital as opposed to 21% of people with *schizophrenia* and 9% of people with *anxiety disorders*.

While there were no significant differences across *diagnostic groups*, service providers and consumers in the *general practice setting* generally had a significantly longer period of contact than was the case in the *public psychiatry setting*, and both groups had significantly longer periods of service provider-consumer contact than that observed in the *private psychiatry setting*, $F(2,170)=25.3, p<.001$. In the *general practice setting* 66% of service providers and consumers had been in contact for more than two years, as opposed to 20% and 19% in the *public* and *private psychiatry settings*, respectively.

4.1.2 Comparison of MHI and SF-36 Data with Normative Samples

Australian norms have recently been published for the eight sub-scale scores of the SF-36 (Stevenson, 1995) and there are U.S. norms for the physical and mental component scores (Ware, Kosinski & Keller, 1994). Norms based on a U.S. sample have also been published for the MHI (Davies et al., 1988). Tables B1.1 (a) and (b) and B1.2 display the Australian and U.S. norms for the SF-36 and the MHI, respectively. Independent sample *t*-tests indicated significant differences ($p<.001$) between the population means and the means for the present sample on all scores of these measures, indicative of higher levels of mental health problems in the study group.

4.1.3 Differences Due to Participant Attrition

Approximately 40 percent of the sample of consumers dropped out between Time 1 and Time 2. Analyses were undertaken to determine whether there were any differences across *age*, *gender*, *severity of illness*, service utilisation variables, and Time 1 composite scores for the selected measures, between those people who remained in the study and those who dropped out. Results indicated significant differences on only a few variables (see Table 1.4). People who dropped out tended to obtain scores that suggested higher levels of severity of illness and disability than those who remained in the study. People who remained in the study tended to have longer associations with their service provider, and to have had fewer admissions and spent fewer days in hospital in the six months preceding the study.

Those who dropped-out had significantly higher mean *CGI* ratings than those who continued with the study, suggesting greater degrees of severity of illness in the former group. Of the six selected measures, only the scores of the RFS suggested differences between the groups. Univariate and multivariate analyses of variance indicated significant differences between the groups for the *global role functioning index* scores, and the scores for the *working/productivity*, and *immediate* and *extended social network relationship scales*.

The effect of the attrition of participants was further investigated in relation to the analyses of change on the measures and is discussed in section 5.4.

4.1.4 Scores on the Selected Measures

Inspection of the descriptive statistics of the scores for the measures suggested that they were generally low in relation to the mid-point for the relevant scale. For instance, the mean score across the sample for the BASIS-32 was 1.2 ($SD=0.8$). The mean scores on the 12 item scales of the HoNOS are all below 2, and in most cases below 1. Both scales have a score range of 0-4. Equating these to the terms of the measures themselves they would suggest that on average people in this sample reported experiencing little or no difficulty on the dimensions of mental health measured by the BASIS-32, and that service providers generally considered their consumers as having relatively minor difficulty on the dimensions assessed by the HoNOS. This is a trend that can also be observed for data on the other measures. It suggests that there may be *floor effects* in the data. That is participants are, on average, reporting initial scores that are indicative of only minor levels of illness. The prospects of seeing an improvement over a short period of time, like three months, is thereby diminished.

Alternatively, rather than this being a peculiarity of the present sample of consumers, one might question whether there is something about the measures that makes them susceptible to such effects. For instance, the manner in which questions are presented or worded might invite respondents to answer in a particular manner. It could be that consumers have a general bias against making judgments in relation to their feelings such as 'all the time' 'extreme', etc. (i.e., 'end-aversion'). Moreover there may be some questions (e.g., such as those on the LSP) which some service providers may feel unable to answer and so provide a low score to give the consumer 'the benefit of the doubt'. These are issues that cannot be immediately resolved, but to which we will return in later discussions.

4.1.5 Summary

The preliminary analyses reported here present a number of general findings in relation to the sample of consumers and service providers who participated in this study:

- Over one third of the consumers were considered by their service providers to have a comorbid mental health condition. The majority of these people were in the *general practice* setting and had an *affective disorder*.
- Thirty-three people were indicated as having drug or alcohol problems with the majority belonging to the *schizophrenia group*.
- A third of the sample had been hospitalised in the previous six months for psychiatric related illness. The majority of these people were recipients of *private* services and had an *affective disorder*.
- Consumers and service providers in the *general practice* setting tended to have much longer periods of contact than those people in the other settings.
- In comparison to published norms, the present sample, represented by those people who completed the MHI and SF-36, indicated scores that were significantly different to those of the norms and indicative of severe mental health problems.
- Approximately 40 percent of the original sample, dropped out of the study between Time 1 and Time 2. Analyses indicated some differences between the groups indicative of greater severity of illness in the group of people who dropped-out, based on their Time 1 scores. Further analyses in relation to this issue are reported in section 5.4.
- Mean scores for the measures collected across the sample raise questions as to the severity of illness of the sample of consumers and the possibility of floor effects, as well as issues relating to biases that may be inherent in the measures.

The analyses reported here are by way of describing the sample of consumers who participated and thereby place the data collected from the selected measures of outcomes in an appropriate context.

4.2 Feasibility Analyses

The following set of analyses investigated issues relating to the feasibility of implementing the selected measures of outcome in routine clinical practice.

4.2.1 Utility Ratings of Consumer Measures

Following the administration of each of the selected measures (BASIS-32, MHI, SF-36), consumers were asked a series of questions (see Tables B2.1 and B2.2). In general, consumers reported little difficulty in understanding the wording of the questionnaires, and indicated that each of the measures was either *reasonably* or *very useful* as an outcomes measure. This is consistent with the fact that consumers were only asked to participate if the service provider felt that they were able to complete the measures.

The majority of consumers responded that *most* or *all* of the questions on the questionnaires addressed issues that were *important* and *relevant* to them. There was a trend for participants to indicate that *most* or *all* of the questions of the MHI addressed *relevant* issues to a greater degree than either the BASIS-32 or SF-36 (81.1.% versus 64.2% and 69.5%, respectively).

Analyses revealed a reasonable level of consistency in responses across questions 2-4 ($\alpha=0.74$). Summing the scores on these items created a General Utility (GU) score. Three way analyses of variance (ANOVA) investigated the differences in GU scores for the three measures used, and across the three *diagnostic groups* and *practice settings*. Significant differences in GU scores were observed for the three measures, $F(2,331)=3.5$, $p<.05$ (see Figure 4). Post-hoc analyses with Scheffe corrections, indicated significant difference between the mean ratings for the MHI ($M=9.6$, $SD=1.9$) and the mean ratings of the BASIS-32 ($M=8.7$, $SD=2.1$) and SF36 ($M=8.9$, $SD=2.0$), respectively: on average consumers reported that the MHI had greater 'utility' than either the BASIS-32 or the SF-36. No significant differences in GU scores were observed across *diagnostic groups* or *practice settings*.

Upon completion of two of the three selected measures (see Section 3), consumers were asked which measure they thought was *better* (see Table B2.2). A significant difference between ratings for the MHI and SF-36, $\chi^2(1)=9.0$, $p<.01$, was observed: of those consumers who used both the MHI and SF-36 ($N=46$), the majority (71%) considered the MHI a better measure. Comments provided by consumers suggested that the MHI was considered more relevant to the assessment of mental health problems.

Comments Received From Consumers in Relation to the Selected Measures

Upon completing the utility questionnaire consumers were invited to provide written comments about each of the selected measures. The range of issues raised by consumers in relation to the three selected consumer measures (see Table B2.3) highlight preferences for:

- A clear and consistent response format. (The response formats of the MHI and BASIS-32 were preferred. The changes in format for the SF-36 were considered confusing by some consumers.)
- An adequate range of alternative responses.
- Open-ended questions which allow for additional information to be provided.
- Culturally appropriate and familiar language. (Some of the language used in each measure was considered old fashioned or ambiguous.)
- A comprehensive range of questions which address both functioning and symptomatology factors.
- An appropriate time period in relation to assessment. (This point was raised in relation to each of the measures. The time period covered by the MHI and SF-36 (i.e., one month) was considered difficult given problems with recall of feelings, and the changes that can occur in a person's mental health. The one week period assessed by the BASIS-32 was considered more appropriate by some consumers.)
- Positively framed questions.
- Measures that are relevant on a personal level.

These comments were reinforced and elaborated upon during the three focus groups which were

convened for consumers. A number of other issues were raised which were consistent with information provided by service providers and these are discussed below in section 4.2.4.

4.2.2 Report of the Consumer Focus Groups

Three focus groups were convened for the consumers who participated in this study. All participants of the focus groups had previously completed at least two of the three consumer outcomes measures, prior to their involvement in the discussion groups. Most participants had completed the questionnaires on two occasions. What follows is a summary of the main issues which arose during the focus group discussions (see Table B2.4 for a summary of the main points).

What Benefits Would There Be In Implementing Routine Assessment of Consumer Outcomes?

The following points were raised as potential benefits of implementing routine assessments of consumer outcomes:

- Increased self-awareness of condition.
"...I mean it is a good idea in theory, but if you even just had the forms yourself to keep a record every month without necessarily showing them to anyone else..." (Group 3)
"...think it also highlights for yourself when you actually sit down to write, you sit down and you know, you feel like you list down, ... and you know why you're feeling the way you are in different areas..." (Group 3)
- Improved communication between mental health professionals and consumers.
"...I've found that I have difficulty trying to get across just how bad I was, so unless you are in that stage - I mean, how are you today, oh, I'm great, but the other day I was... and it's hard to get across just how really bad you felt, I feel, just in words...Perhaps it's something that you could have, like a number of them, and you fill it out regularly yourself, and then it's something that your doctor could browse through at a later date..." (Group 2)
- Provide and/or improve the structure of clinical interview sessions.
"...usually they're running late, and I understand that ... they don't have a bell on their desk that says ding, your time is up, see you next week sort of thing. But what I find really difficult is although it's your turn and you've got to go in and you've got a certain time frame and you've got to say everything you need to say, and at times I find that incredibly difficult. You know, you've got yourself dressed, you've put your make-up on, you've pinned a mask on your face to go out the door, you've held yourself together and all of a sudden you've just got to go in and he'll say how are you and you've got to unravel everything and spit it all out and I find that incredibly hard so a lot of times I write notes...Maybe if I had a questionnaire I'd fill that in..." (Group 3)
- General benefit to the community through an increased understanding of mental health.
"...And also, I think, if having the form filled out and so forth for the bureaucrats, then it could go the full circle, if the bureaucrats let the media know, and sort of say well, there is quite a lot of mental illness in the area, and it would start to get rid of the stigma of mental illness. It's still very much in the closet..." (Group 2)

Table B2.4: Summary of Issues Raised during Consumer Focus Groups**Benefits of Implementing Routine Assessment of Consumer Outcomes**

- Increase self-awareness of one's mental health.
- Improve communication between service providers and consumers.
- Provide and/or improve the structure of clinical interview sessions.
- Indirectly assist increasing knowledge within the community about mental health.
- Provide the Government and service providers with better information by which to target funding of mental health services.
- Provide a less threatening means of communicating dissatisfaction with a service. People may be more willing to make a complaint about a service in a format such as this, rather than confronting the service provider directly.

Feasibility Issues: Applicability, Acceptability and Practicality

- Unrealistic to expect that any single questionnaire could be appropriate for all diagnostic groups; the measurement of different stages of illness; for use in inpatient as well as outpatient settings;
- To use one standardized questionnaire may place limitations on consumers' individualism.
- To use one questionnaire on a routine basis may invite familiarization with the questions it contains, and thereby reduce its validity.
- Relying on one measure may minimize the effects of important personal factors.

Structure and Content

- Items should address living skills and general functioning, as well as symptomatology and quality of life.
- Open-ended questions and alternative response modalities may be useful.
- The time periods need to be addressed. It can be difficult to recollect events which have occurred throughout the assessment period. During a stipulated time frame a consumer may have experienced each of the options presented by the questionnaire response format.
- Debriefing process required following administration of the measure so that adequate feedback is provided to both the consumer and the service provider.
- Some language used in the selected measures was considered old fashioned and inappropriate.
- Why is a 'questionnaire' necessary to measure consumer outcomes? Could other processes be utilised?

Procedural Issues

- Who will be responsible for deciding whether a consumer will be required to complete this standardized measure on a routine basis?
- What will be the defining criteria regarding inclusion or exclusion in an outcomes assessment process? This encompasses factors such as minimum standards, and whether the measure should address symptomatology, daily functioning, etc.?
- Will this be a mandatory requirement for all consumers of a mental health service, regardless of service setting?
- Who will be responsible for deciding upon the time frame by which these measures of outcomes are completed? Will such a time frame be uniform, or will it vary depending on diagnosis, practice setting and severity of illness, etc.?

Other Issues

- Both consumers and service providers have to be committed to the process for it to work. This may involve increased education to specific groups of service providers, such as general practitioners.
- The site for completion of the measures, and whether assistance would be provided. Completion of the measures outside the service setting was viewed by some as a potential problem. Completing the measure at the site of service delivery could increase completion rate, provide consumers with the opportunity to complete the questionnaire without interruption, and provide a similar or more uniform environment in which to complete the forms. Other consumers, however, suggested that they were likely to be more objective if they were able to complete the questionnaires in the comfort of their home environment.

Table B2.4 (cont.): Summary of Issues Raised during Consumer Focus Groups

- Several consumers stated that they would be uncomfortable if somebody else were to complete an outcomes measure pertaining to them. A collaborative effort involving themselves and their service provider was a more desirable option. Some indicated a preference to have another person present whilst they were completing the questionnaires to help clarify confusing points, and to provide support if the exercise became distressing.
- The process of completing the questionnaires can be distressing. Not everyone would feel comfortable with having to complete a questionnaire, and it is often a very difficult process to have to choose between available responses.
- Some consumers felt that it would be unreasonable to complete a questionnaire when they are unwell. Others suggested, however, that if an attempt were made by consumers to complete the forms when they were unwell, it could provide their service provider with an increased awareness of the severity of their illness.

Confidentiality

- Issue of confidentiality raised and re-emphasized on numerous occasions. Unless anonymity and confidentiality can be guaranteed, it is unlikely that consumers would consent or comply with this process.

Purposes of outcomes measurement and the use to which the information may be put, needs to be clarified.

The following questions were raised:

- Is the information obtained from this exercise to be used primarily as a clinical tool to enhance individual patient care?
- Is the information obtained from this exercise to be used primarily as a clinical tool to assess service provision?
- Is the information obtained from this exercise to be used primarily by the Government on an economic level? This could lead to bias in the results. Concerns about losing services, budget cuts etc., may influence responses to outcomes measures. Fears were expressed that this information may be used as a form of ammunition to support restrictions in the services made available to consumers, or for comparative purposes in hospital versus community settings.

A further issue raised was that questions should address the issue of the treatment being received. It was noted that none of the questions contained in the selected questionnaires did not address the issue of service provision, and that they would therefore not provide useful information pertinent to the services themselves. Consumers felt they should be consulted in a direct manner about their opinions and concerns regarding the services they receive.

The selected questionnaires seemed to exclude environmental events and immediate life situations which could impact on an individual's response, and so may not provide a true indication of an individual's progress.

The validity of the data obtained from outcomes measures was discussed. Issues raised here included:

- The possibility that some individuals may respond in a manner which is reflective of changes that they perceive are desired by their service provider.
- Differing cognitive and emotional levels may effect question interpretation and therefore response.
- The ease of manipulation when responding to a questionnaire.

Consumers of mental health services want a voice regarding their individual treatment plan and its progress. There is a need for improved communication between services. It was stated that a collaborative approach is necessary to break down the walls between professionals and consumers of mental health services.

It was argued that some service providers do not like feedback from consumers in relation to quality of service issues. There is a need for consumers to be able to be specific in relation to their service provider and their satisfaction with service if a true measure of outcomes is to be made. Constructive criticism is a necessary component of the measuring of outcomes, that being "...what I want from a service provider versus what I didn't get".

- Improve the ability of government and service providers to target funding to mental health services.

"...I think that it is very important to provide some feedback to the bean counters, to the bureaucrats, on how to divide the cake up. Like, if they've got no information there on the degree of mental disorders and illnesses within society, they're just not going to give any dollars to it, and that is the bottom line in this day and age. That's what it is all about. Like, if it doesn't appear on a bit of paper or on a computer file somewhere, it is just forgotten about. I think that it's very good that it does provide that evidence, albeit, as everyone has said here, that it mightn't be quite right, it might vary a bit this way or that way, but at least it does provide something. So it's a step in the right direction..." (Group 2)

"...I think that the overall statistics should be available to any government agency or whatever that requires them, in my view - in the long term they've got to be of benefit to other people with the same needs and it might help people in the future to start getting treated properly..." (Group 3)

Feasibility Issues: Applicability, Acceptability and Practicality

Several factors emerged relating to the feasibility of using a single standardized questionnaire. These are discussed in relation to general issues, structure and content, and process issues.

General Issues

It may be necessary to use different questionnaires to meet the needs of people with different diagnoses.

"...well even when I was filling out the questions they were too broad..." (Group 2)

"...the questions that would be relevant say to someone suffering schizophrenia, might not be relevant to someone suffering say bipolar..." (Group 1)

"...you can't really use a uniform assessment scale because of the different varieties of mental illness..." (Group 1)

It may be necessary to use different questionnaires or different response formats to account for the different stages of illness.

"...a month ago I wouldn't have been capable to fill out the questionnaire, but a month later I was able to help..." (Group 3)

"...instead of having numbers from 1-10, maybe you could have 10 questions, and if you answer yes to the first question, you go to the second, yes to that you go to the third one. So they get more and more severe. So you aren't just saying do you suffer from anxiety, and you say no, then they think you just scored a one on the scale. Whereas if you can go through all those 10 questions answering yes, yes, yes, then you score 10. If your only half way through the questionnaire and say not so applicable, then you score a five. So they don't have to actually decide upon a number, they can decide on the question. More specific, and put less pressure on people as well, but it would also reduce the time limit of the questionnaire, because there are some issues that R....might find has nothing to do with him, so he only answers one..." (Group 1)

It may be necessary to use different questionnaires for different clinical settings.

"...I just thought because I was in hospital there should have been separate questions for me. I felt that asking me routine questions about my daily life was absolutely useless because I couldn't answer them, you know I was sitting in bed. You know my routine life was getting up taking my medication, going to a group and coming back having lunch and doing the same thing again. You know I mean I couldn't answer those questions, it was ridiculous..." (Group 3)

Use of a standardized questionnaire may place limitations on the consumer's individualism.

"...my concern is that if we're trying to be given standardized forms, we're trying to be standardized, and in psychiatry, patients are not standardized. Everyone is different. Even in little ways, everyone is different. And what you can cope with might be different to what I can cope with..." (Group 3)

"...when you look at it every individual is different. So there should be a questionnaire for every individual...They try and treat you like a group, where actually they should be treating you as an individual and your own personal self..." (Group 1)

Use of a questionnaire on a routine basis may invite familiarization with the questions it contains, and thereby reduce its validity.

"...you think, am I more a six or am I a five, and then the next time you go to do it you think, oh well, I put five down last time so I'll put five down again, aiming at the average to try and not make myself feel that bad..." (Group 1)

Structure and Content

Regarding the use of questionnaires as measures of treatment outcomes, the following points were raised in relation to the structure and content desired in such a tool:

- An outcomes measure should contain both living skills and general functioning information, as well as questions pertinent to symptomatology.

"...I found with my schizophrenia, a big aspect of my well-being is my routine for the day, and there is nothing on that type of thing in the questionnaire..." (Group 1)

"...how well you function - what can you do and how do you do it - that kind of stuff is good..." (Group 1)

- Inclusion of open-ended questions would be beneficial; some participants suggested that the range of choices were restrictive, and that as a result they felt 'boxed'.

"...I've been on it from both sides. I myself am a GP, and I've been sick for over three years and I've certainly learned a lot from being a patient as opposed to a service provider shall we say. I have a lot of concern about the forms in Psych.. Definitely because you have to get the patient to fill it out, and it depends upon the patient at the time, and my concern also is that you're getting boxed in - it's not how you say it, it's not an assessment of you, but it's you ticking boxes and rating things and my concern is that, OK you've got 30 questions and you rank them from zero to ten, and if you score over two hundred you're pretty sick and if you score under two hundred you're managing OK..." (Group 3)

- Alternative response methods, such as the Visual Analogue Scales, may have potential benefits especially by eliminating problems associated with language use.

"...I sometimes have trouble choosing boxes, it depends on the situation. Sliding scales are much better. It depends on the question. If you have 0, 1, 2, 3, happens all the time, or doesn't happen at all, well there's no 1 ½..." (Group 1)

- The time frame inherent in responding to a questionnaire may lead to difficulties in the recollection of events and feelings. In a stipulated time frame a consumer may have experienced each of the options presented to them by the response format.

"...I don't really like filling in boxes at all. But I found too that each day can be different. Like, when they say over the last fortnight or the last week or whatever it is. I mean there can be so many differences in one day alone it's like playing snakes and ladders - you know, you're up and down, and to call that precisely, you know I'm personally concerned - it just doesn't gel with me at all..." (Group 2)

"...you have a lot of forgetfulness when you're asking how things have been over the last month or two..." (Group 3)

- Some of the language used in these particular questionnaires was considered old fashioned, and inappropriate.

"...I actually found them a little condescending, old fashioned in the words..." (Group 1)

"...and the other thing that I found too - did you find that the terminology was old? You know, there were words being used that we don't use any more, you know, it was sort of as though they designed the test 25 to 30 years ago..." (Group 3)

- Questions should address the issue of the treatment being received. None of the questions contained in the questionnaires actually addressed the issue of service provision. It was suggested that consumers should be consulted in a direct manner about their opinions and concerns regarding the services they receive.

"...No, there's nothing about have you been in hospital, or what medications have you been taking, have you had psychotherapy, have you had counseling, how long have you needed to be in hospital,

how many hospitalizations have you had, so there is nothing on the individual form about your treatment...” (Group 3)

- The selected questionnaires exclude environmental events and immediate life situations which can be expected to have a major impact on an individual's response. The questionnaires therefore may not provide a *true* indication of an individual's progress based solely on mental health issues.

“...what about life situations, what about anxiety triggered off by life situations? Because to me this is just so normal.... And I don't really think that the anxiety that I experienced had anything to do with my illness, I think that it was just plain normal...” (Group 2)

Process Issues

The issue of exactly who would be required to complete the questionnaire on a routine basis, and across what time frame, was discussed:

- Who will be responsible for deciding whether a consumer will be required to complete this standardized measure on a routine basis - the service provider, the consumer or the Government?

“...Well it's obviously not suitable for all people. That's plainly obvious from the start...So I would imagine that you clear that up by allowing the medical practitioner to make some pre-assessment of whom the questionnaire should be directed to...” (Group 2)

“...As soon as I saw it I put the warning bells up - we are all individuals, we all have different body chemistry, we are all taking different drugs for our illness, some people have more anxiety than others, I have hardly any anxiety, but I have other things, like mood changes and all kind of stuff, and it really worries me - I'm worried because of what's going to happen with this...” (Group 1)

- What will be the defining criteria regarding inclusion or exclusion in such a process? This encompasses factors such as minimum standards, and whether measures should be based on symptomatology or functioning.

“...if it was going to be used, it could be used by some GP's as a way to become more aware or to assess if somebody needs some psychiatric assessments...That's where it could be used, I felt. But beyond that point, I wouldn't...” (Group 1)

- Will outcomes assessment be a mandatory requirement for all consumers of mental health services regardless of service setting?

“...we are also working on a weekly basis and I am also doing the other things that I am doing, and I'm seeing a psychiatrist as well...However, I wouldn't want someone that I didn't consider to be very well qualified to be making decisions about what was going on in my head...” (Group 1)

- Who will be responsible for deciding upon the time frame by which these measures of outcomes are completed? Will this time frame be uniform? Will it vary depending on issues such as diagnosis, practice setting and severity of illness?

“...personally, I think you've probably determined yourself how often ... I don't just mean filling out the questionnaire and handing it in to someone, but actually, actually going through the questionnaire and some of the questions and doing a self diagnosis - I suppose that everybody is sort of an individual...” (Group 3)

In order for this process to be beneficial, both consumers and the service providers have to take it seriously. This would involve increasing the education and awareness of mental health issues amongst different groups of service providers.

“...maybe I shouldn't repeat this but when I was at the doctors and I was asked to fill the form out he said to me he thought it was an absolute waste of time...Not that I'm criticizing my doctor in any way, but you're going to come across that with medical practitioners - it's going to be a waste of time and they might not fill it in ...” (Group 3)

“...basically that questionnaire...it's giving the GP's that aren't qualified to pick that (mental illness) up themselves.....which is great in one way, but not so good in another way. I still think that they should be trained...” (Group 2)

- While some people preferred that questionnaires such as these be completed at the service setting, others viewed this as a potential problem. Completing a questionnaire at the practice setting would also allow consumers the opportunity to complete the questionnaire without interruption, and this may provide a similar or more uniform environment in which to complete the forms. Several participants, however, suggested that they were likely to be more objective if they were able to complete the questionnaires in the comfort of their home environment.

"...I filled out two in the doctors surgery, you know, in fairly quick succession, and I chose - I was given the choice to come back to the surgery and do the third one, or to take it home, and I said I'd rather do it at home, because I'd feel more relaxed then..." (Group 2)

"...I would have to make sure that I made an appointment with the doctor before I walked out that door, so I wouldn't procrastinate..." (Group 2)
- Several participants stated that they would be uncomfortable if somebody else were to complete an outcomes measure pertaining to them. Others considered the idea of a collaborative effort involving themselves and their service provider a more desirable option. It was suggested that this process may serve to increase awareness between both parties as to how treatment is progressing. Some individuals preferred to have somebody present whilst they were completing the questionnaires.

"...it would have been handy to have had somebody there to go through it with me because it did have quite a few questions on medication and those sort of areas and that type of thing, and it was not an easy one to fill out..." (Group 3)
- It was suggested by some participants that it may be unreasonable for consumers to complete a questionnaire when they are unwell. Contrary to this view, others suggested that if an attempt could be made by consumers to complete the forms when they were in fact unwell, it may provide their service provider with an increased awareness of the severity of their illness.

"...I would say twice a year. But then occasionally if you get an acute state, to do one then to see whether the doctor perceives how sick you are, and how you perceive how sick you are at that time..." (Group 2)

"...I had problems filling out the one for the anxiety clinic. I was in...at the time and I really wasn't very well when I did it, and I found that I had an actual panic attack after filling it out, that's how upsetting it was..." (Group 3)

"...I think I was more upset by filling out the form because at that particular stage I was going through a lot of anxiety and a lot of panic attacks and I found that extremely difficult to fill out, and then after I'd filled it out I was more anxious because I didn't know whether I'd filled it out correctly and there were a lot of questions on it that were sort of ambiguous, and that really upset me quite a bit, because I wasn't sure whether I'd filled it out correctly or not, and it was a fairly lengthy form..." (Group 3)
- A debriefing process would be required following the administration of the questionnaire so that adequate feedback may be provided by both the consumer and the service provider. This may include discussing the meaning of results, relating new results to previous results etc..

"...I find that when I do a questionnaire it is very important to make sure that the person leaves caring, because you brought up all of those issues and sometimes when I leave....and I just need to sit down and just have a brief chat after the whole thing is finished just to know that you are alright...that they don't make a judgment of you, or that they don't care any less for you..." (Group 1)

Concerns

- It was suggested that unless anonymity and confidentiality can be guaranteed, it is unlikely that consumers would consent or comply with this process.

"...just one other concern I've got about this is the issue of anonymity, because there is just no way that your name could be linked to those results that go into a data base somewhere. This is absolutely paramount, because if that information got out to insurance companies or lawyers or something like that, you wouldn't have a leg to stand on. That is another fear that I have, because confidentiality must be absolutely water-tight..." (Group 2)

"...the information can go anywhere. But my name and the information should be separated at an early stage with people that are only directly involved with me..." (Group 1)

Participants were very concerned about the purposes of outcomes measurement and the use to which the information may be put:

- Is the information obtained from this exercise to be used primarily as a clinical tool to enhance individual patient care?

"...the benefits of this project hopefully should be to us, not the service providers. You should be able to, or I would really like to be able to go to a service provider and say, at which service you're providing gives the best outcomes for my particular illness?..." (Group 1)

- Is the information obtained from this exercise to be used primarily as a clinical tool to assess service provision?

"...so there's two functions to it then, are there? It's not only to me on outcome, but also a tool to the GP on how his performance is going... What I'm saying is that it's all very well to have this system, but you've got to look at the outcomes of it in terms of what do you do if the GP is not up to scratch? What recourse do you use?..." (Group 2)

- Is the information obtained from this exercise to be used primarily by the Government on an economic level? Such a motivation may lead to bias in the results. Some consumers stated that concerns about losing services, budget cuts etc., may influence their responses to a questionnaire. Fears were expressed that this information may be used as a form of ammunition to support restrictions in the services made available to consumers, or for comparative purposes in hospital versus community settings.

"...but isn't the whole point of the questionnaire not so much to assess your personal health level, but to provide data to the bureaucrats to be able to divide up the health dollars? And to what the outcome is, it's not actually the treatment, it's the outcome of the treatment..." (Group 2)

On several occasions it was indicated that completing questionnaires like the selected measures can be a distressing process for consumers. Not everyone would feel comfortable with having to complete a "questionnaire", and that it is often a very difficult process to have to choose one of the responses available.

"...it triggered the condition you see. Any time you saw a question that related to your particular condition, it acted like a trigger for it..." (Group 2)

"...I had problems filling out the one for the anxiety clinic, I was in ... at the time and I really wasn't very well when I did it, and I found that I had an actual panic attack after filling it out, that's how upsetting it was..." (Group 3)

The validity of the data obtained from questionnaires was also questioned. Issues raised here included:

- The possibility that some individuals may respond in a manner which is reflective of changes that they perceive are desired by their service provider.

"...well that's right, they could just tick the ones that you want. I mean, some people can be very, even though they may want help and want to get better, they can also be very manipulating..." (Group 2)

- Differing cognitive and emotional levels may effect question interpretation and therefore response.

"...but then again on the other end of the scale, you can get people who are very basic and not really intelligent, and because they don't actually understand the questionnaires, you know, they guess or they just don't know how to respond, there again, you'll get a false reading..." (Group 2)

- The ease of manipulation when responding to a questionnaire.

"...if the GP is making the marks, then he is probably putting some of his own interpretations down..." (Group 2)

"...my concern there is that often the patient knows when they need to be in hospital more than anyone else does, and ticking boxes and squares and rating things is not only stressful at the time, but also open to manipulation, I mean you can say I've learnt from the past that if I score under two hundred I don't get to go to hospital, so if I rate everything eight or nine I'll get to go to hospital and stuff. So it's fraught with that sort of possibility..." (Group 3)

4.2.3 Report for the Consumer Support Group Focus Group

The focus group held for members of local consumer support groups is discussed separately from the main report of the Consumer focus groups for several reasons. Firstly, the participants who were involved in this focus group had not participated in the field testing prior to this meeting, and had not been involved in completing the outcomes measures. Secondly, those who participated in this particular focus group were invited to do so primarily because of their involvement with the Queensland Association for Mental Health and related Support Groups. Each of these participants was involved with a Consumer Support Group. They were all consumers and have an indepth and firsthand awareness of the services and resources available to consumers of mental health services in the Brisbane Region.

Participants were provided with copies of the questionnaires approximately one week prior to this meeting in order to familiarize themselves with the type of measures being considered for routine use. What follows is a summary of the main issues which arose during this focus group discussion (see also Table 2.4).

What Benefits Would There Be In Implementing Routine Assessment of Consumer Outcomes?

Several benefits emerged in relation to the use of standardised questionnaires. These included the following:

- Routine assessment of outcomes with a standardized measure would provide consumers with an opportunity to express themselves more comprehensively. Issues which may not be addressed during a clinical session may be commented on when responding to a questionnaire. *"...the questions are a good idea for some, for others who can't object too high about what's going on, and it makes it really hard to identify. Whereas if you've written it out, yes that's something they may not have thought out. So you've identified feelings or behaviours that someone might not of if you'd asked an open question to identify those things.*
- Using a standardized measure may allow some individuals greater objectivity in reporting their mental health needs. *"...well, it's very hard to measure the quality of the services. It's not easy to measure. But they do have to measure them I believe, they do have to measure them..."*
"...what you're actually doing is [giving consumers] a voice about private treatment outcomes or public hospitals, which is fine, because no one has the time or the power to say, I need this but I didn't get it...So the system doesn't allow the patient to feedback on what would improve treatment to them..."

Feasibility Issues: Applicability, Acceptability and Practicality

Several factors emerged relating to the feasibility of using a single standardized questionnaire. These included the following:

It is unreasonable to expect that one standardized questionnaire could be applicable across the range of mental illnesses. One questionnaire could not be sensitive enough to be utilized both across differing diagnostic groups, and at the various stages of illness.

"...I'm very worried about a standardized form to be filled in, in an area with such a huge variation, and that you're not going to get things that are standardized anyway within this area..."

Using only one questionnaire may in fact minimize the effects of factors such as:

- History - including both life history and history of illness. *"...well I don't know because if you come from an impoverished background say, and you've been treated like a piece of s— the whole of your life, how do you know what self esteem is, how do you know what respect is, how do you know..."*
"...and I think that could pose huge problems for people who hear voices or who come from different ethnicity, because they could very easily be slotted into a box, and I find that quite frightening..."

- Geography - encompassing differing and varied cultural and location issues e.g., urban versus rural.

"...or in a small town where there is only one psychiatrist, it's pretty hard for that psychiatrist to drop that patient..."

"...don't you think you should take into account that there are migrants in Australia, there's the ethnicity of Australia for starters, and the indigenous population..."

Structure and Content

It was felt that more attention needs to be given to issues such as:

- Life skills, life functioning and quality of life issues.

"...is my life fulfilling to me? That, that comes nowhere near that. So you're filling in a questionnaire that someone else is going to look at and make their own decision on your lifestyle. That hasn't given you any space at all to put what you think, to put how fulfilling you think your lifestyle is..."

- Open-ended formats should be included to promote feedback from consumers. This would hopefully be considered as a component of the negotiation process between the consumers and the provider of the service.
- Questionnaires may provide a less threatening means of communicating dissatisfaction with a service. People may be more willing to make a complaint about a service in a format such as this, when compared to confronting the service provider directly.

"...I'd rather write the answer myself rather than just filling in their questions. I'd like straight out questions ..."

- It was asked, "Why do we need a 'questionnaire' to measure outcomes?" It was argued that there exists a broad range of research methodology that was not being utilized because it was considered both too labour intensive and expensive.

"...It's the process you know, and I think that that's what we need to address. And I go back to why do we need questionnaires. Isn't there another way of doing this, like a one on one interview with an objective person who can sit down and document? I mean that might be more time consuming and more labour intensive, but the outcomes are going to be far more objective, far more solid, far more concrete. ...I mean there is a lot of research methodology around but we're just not using it because...it's too labour intensive and it's also financially not productive..."

"...My whole beef about the whole thing is why do we need questionnaires when there are other ways of evaluating the service rather than by where you have four options, you know, what are they, sometimes, never, hardly ever or always. You know, I mean how can you operate on that, it's very confusing to someone who, particularly when you're in a confused state of being and your emotions are all over the place. That is just a time bomb, because you feel you're inadequate because you can't answer a basic questionnaire that's asking basic questions that are very personal to yourself for starters..."

Process Issues

- The consensus to emerge amongst those participating in this discussion was that consumers of mental health services want a voice regarding their individual treatment plan and its progress. It was also expressed that there is a need for improved communication between the services. It was stated that a collaborative approach is necessary to breakdown the walls between professionals and consumers of mental health services.

"...I think it would be valuable if consumers could have their own input into things, how their progress is going, rather than having the nurses reports saying this is how I felt and this is etc. Consumers are at a capable level of being able to communicate with their doctor and being able to say well this is how my treatment is going while you're not here, and have some input into perhaps some dialogue with their doctor in terms of negotiating a treatment plan and sticking with that treatment plan. With the doctor and the patient informing the nursing staff about this, how we have decided treatment should go for this person, and we have negotiated this, and this is what we will abide by and not what the hospital or the DSM-IV says is a way of treatment because everybody is a unique individual, and needs to be treated as such..."

"...you know there are boxes being compiled, and you know there are far too many of them, and that would break down the boxes, with being given individual treatment and assessment plans, with the patient involved to keep their cultural issues involved, you know, there's a lot to nut out between the

doctor and the patient. The doctor quite often just goes by the DSM-IV criteria, not looking at the whole of the person, but just looking at the psychiatric basis of the condition, not looking at lifestyle, not looking at living, housing situations and all those sorts of things that are important in integrating the person, to become a whole person..."

- Some participants argued that service providers do not like feedback from consumers in relation to quality of service issues. There is a need for consumers to be able to be specific in relation to their service provider, and their satisfaction with the service, if a true measure of outcomes is to be made. Again, a consensus emerged amongst the group that constructive criticism is a necessary component of the measuring of outcomes; that being "...what I want from a service provider versus what I didn't get".
"...they're supposed to be looking after you and your well being is at stake, so I would speak out..."
"...I'd feel threatened, but I'd fill it out, no worries..."
"...Ah, disagreeing with the service or the service provider, you know once again you may think a service is a failure at one point in the [treatment] process, and a little further down the track you may view the whole thing in a different way, so it's a difficult one, but I'd fill it in..."
"...but if you're going to criticise you've got to put back something in so that they know what you want, you can't just go slagging them all the time..."
"...we know what's wrong, but they, the service providers, they don't want to know, they don't like being picked on..."
- The questions included in the selected outcomes measures did not address the issue of service provision. Emphasis was placed on the need to ask straightforward questions.
"...this questionnaire is supposed to be evaluating the services, and all it is doing is evaluating the patient..."
"...some of the questions should be things like, how do you rate your service to your consumer, or your client? Do you think that the hospital you work for is giving you enough support?..."
"...it's supposed to be about patient services, but there's not one thing in here about services..."

4.2.4 Utility Ratings of Service Provider Measures

Following administration of each of the selected measures (HoNOS, LSP, RFS) service providers were asked a series of questions (see Tables B2.5 and B2.6). A substantial number of service providers experienced some difficulty with the wording of the questionnaires, viz., the HoNOS and RFS. For instance, 39% and 24% of those who completed the HoNOS and RFS, respectively, had difficulty with *some* or *most* of the wording of these measures, as compared to 11% of those who completed the LSP (see Table B2.5). This may have been due to the format in which the measures were presented. As indicated previously the HoNOS was administered in a non-standardised format. The authors of the HoNOS recommend a minimum amount of training in the use of this measure. In the present study, service providers were simply given the glossary to the HoNOS to use as a questionnaire. This may have led to confusion given the wording of the response choices.

In general, service providers indicated that some or most of the questions of each questionnaire were relevant to assessing outcomes for their specific client(s), for people with the same diagnoses as their specific client(s), and for people with mental illness in general. Similarly, service providers generally indicated that the three measures were *slightly* or *reasonably effective* as measures of treatment outcomes. There were few substantial differences in ratings across the three measures.

Analyses revealed a high level of consistency in responses across questions 2-8 ($\alpha=0.95$). Summing scores for these items created a General Utility (GU) score. Three way ANOVA investigated differences in GU scores for the three measures used, and across the three *diagnostic groups* and *practice settings*. Significant differences in GU scores were observed across the *practice settings*, $F(2,318)=30.8$, $p<.001$. Post-hoc analyses with Scheffe corrections indicated significant differences between all groups: *general practitioners* gave higher ratings ($M=19.2$, $SD=4.0$) than service providers from *public psychiatry settings* ($M=17.4$, $SD=4.4$), who, on average, gave higher ratings than service providers from *private psychiatry settings* ($M=14.8$, $SD=3.6$). These results may be due to differences in attitudes and approaches towards outcomes assessment across settings: service providers in the *private psychiatry settings* may have had more experience with mental health status measures and therefore have different expectations of such measures. On the other hand, a number of the *general*

practitioners were currently involved in a psychiatry special interest group and may have considered such resources especially useful to facilitate their services to consumers who have a mental illness. As well, many service providers in the *public psychiatry settings* were already familiar with the LSP.

There was a significant interaction between the *measures of outcomes* and *practice settings*, $F(4,318)=4.6$, $p<.01$, indicating different assessments of the utility of the measures across *settings* (see Figure 5). Additional analyses, employing a Scheffe test to control for an increase in familywise error, were undertaken to examine the simple effects of the *practice setting* factor. Significant differences were observed between the mean *GU* scores in the *public psychiatry settings*: the mean for the LSP ($M=19.7$, $SD=4.0$) was significantly higher than that of the RFS ($M=15.8$, $SD=4.2$) or HoNOS ($M=16.5$, $SD=4.0$). No significant differences were observed in the *general practice* or *private psychiatry settings*.

Upon completion of two of the three selected measures, service providers were asked which measure they thought was *better* (see Tables B2.6). Respondents indicated few preferences for the HoNOS over the LSP or RFS, however there was a significant difference in the responses to the LSP and RFS, $\chi^2(1)=7.4$, $p<.01$. Of those service providers who used both measures ($N=49$), substantially more respondents indicated that the LSP (69%) was a *better* measure than the RFS (31%).

Comments Received From Service Providers in Relation to the Selected Measures

In completing the utility questionnaire, service providers were invited to provide written comments about each of the selected measures. A range of issues were raised in relation to the three selected service provider measures (see Table B2.7) highlighted preferences for:

- An adequate range of alternative responses. (Each of the service provider measures was viewed to be restrictive in the range of responses available.)
- Positively framed questions.
- Questions of a nature which the service provider could reasonably be expected to answer accurately. (For instance some service providers indicated an inability to answer questions about their consumer(s) due to limited information.)
- Language which is clear and unambiguous.
- Specificity in the focus of the items in comparison to 'global' questions represented by some of the items on the selected measures.
- Questions which are relevant to the needs of individual consumers.
- Questions which address issues such as comorbidity, as well as culture and ethnicity.

Several process issues were also remarked on by both consumers and service providers. These related to:

- The intended purpose of an outcomes assessment process.
- The impact of consumers' illness on their ability to complete the questionnaire (i.e., wellness), and the distress that the process of completing a measure may cause to consumers.
- Whether it is proposed that the form be completed independently or with a trained interviewer present.
- The need for training of service providers in the collection, analysis and interpretation of the data.
- The need for any system of outcomes measurement to be directly relevant to the goals of treatment, and to direct the progress of an intervention.

These comments were reinforced and elaborated upon during the focus groups which were convened for service providers.

4.2.5 Report of the Service Providers' Focus Groups

All of the participants in Service Provider focus groups had completed one version of the questionnaires prior to their involvement in the discussions. What follows is a summary of the main issues which arose during the focus group discussions (see Table B2.8 for a summary).

What Benefits Would There Be In Implementing Routine Assessment of Consumer Outcomes?

- Routine measurement of outcomes would provide a standardized language for assessing clients and communicating results.

"...and with the patients, to say well this is what happened with your treatment, and you can measure change, and see change, and that's very helpful for the patients to take away with them...Can also be helpful for the GP's, because they sort of like to have some sort of outcome measure as well. So for all intents and purposes they can be very useful..." (Group 3)

- Routine measurement of outcomes would be useful in the planning of treatment programs.

"...if you ask them questions about their quality of life, you can build it into some sort of clinical part of what you're doing...on outcome..." (Group 2)
"...To me these are tools I would use to help manage my patient. So with the patients filling in the tool - I mean in the sense that if the tool is good enough - it can almost be part of their treatment in some cases..." (Group 1)

- The observation of positive changes would enhance the service providers' work satisfaction.

"...It's caused a certain amount of satisfaction for me because I think that I did it when I started, and now I'm doing it to finish, and I haven't had anyone that's gone down in this quarter or half, and that was fine for me because I'm working with these clients. But some people might argue that that is not what it's about, and that it's really a measurement of client improvement, but I think it's a measurement of my work with this client. So in that way I found it rewarding for me..." (Group 2)

Feasibility Issues: Applicability, Acceptability and Practicality

The issue of the feasibility of using *one* standardized questionnaire was raised during each of these focus groups respectively. Several factors emerged relating to this issue, these included:

- It may be unrealistic to think that one questionnaire could be used for all diagnostic groups. If one questionnaire were to be used, it was argued, the questions would be too broad, and that this would effect the sensitivity of the information obtained. The measures of outcomes which were trialed in this study were perceived to place an emphasis on chronic schizophrenia and psychosis, and therefore became exclusionary.

"...so I think that this is no new idea, and in fact it's, the way it's set up - this is no criticism of the people who set it up - the fact is that I think it's very, very difficult to get one scale that is going to cover everything, and it's quite clear that the emphasis in the ones that have been used is on chronic psychosis..." (Group 3)

Table B2.8: Summary of Issues Raised during Service Provider Focus Groups**Benefits of Implementing Routine Assessment of Consumer Outcomes**

- Standardized language for assessing clients and communicating results.
- Useful in the planning of treatment programs.
- Observation of positive changes would enhance the service providers' work satisfaction.

There is a need to develop the best measures, in accordance with accepted scientific standards.

Feasibility Issues: Applicability, Acceptability and Practicality

- Unrealistic to expect that one questionnaire could be used for measuring clinical outcomes across all diagnostic groups and practice settings. If one questionnaire were to be used, the questions would be too broad, and this would effect the sensitivity of the information obtained.
- Use of one response style may not provide enough sensitivity to individual consumer situations. To use one standardized outcomes measure on a routine basis may lead to familiarization with the questions it contained, and hence influence response.
- Use of one standardized questionnaire with non-English speaking consumers from a variety of ethnic and cultural backgrounds was raised, and it was suggested that ethnicity will effect the process.

Structure and Content

- Outcomes measures need to contain both living skills and functioning information, as well as questions pertinent to symptomatology.
- Inclusion of some open-ended questions would be useful. Present response formats may not provide enough sensitivity to individual consumer needs.

Procedural Issues

- The process should be simple, short and useful.
- Who will be responsible for deciding who will be required to complete these measures on a routine basis?
- What will be the defining criteria regarding inclusion or exclusion in an outcomes assessment process?
- Will this be a mandatory requirement for all providers of a mental health services, regardless of service setting?
- Who will be responsible for deciding upon the time frame by which these measures of outcomes are completed? The time frame inherent in the scales themselves was considered too restrictive, and this would limit the validity of the results obtained, especially for unstable clients.
- Who will be responsible for the completing the outcomes measures - the consumer, the service provider, both, or other allied health professionals?
- For this process to be of benefit to anyone (consumers, service providers, government), then the outcomes measures must be seen by service providers to be "good measures with face validity, that they will own." Service providers need to be able to see the importance of completing the measures in terms of the impact that they will have on immediate clinical practice.

While some service providers stated that they were already using standardized assessment measures on a routine basis, others stated that they had never considered doing so. While tending to be dependent on service setting, time availability was also offered as a contributing factor here.

The way in which the outcomes measures would be administered was discussed. Some felt that pen and paper administration would be burdensome and impractical in terms of storage, and suggested that a computerized format may be more appropriate. Others stated that the type of administration should be sensitive to what the consumer is comfortable with.

The need for confidentiality and the protection of the privacy of all involved in this process was stressed throughout each session.

Table B2.8 (cont.): Summary of Issues Raised during Service Provider Focus Groups

The purposes of outcomes measurement and the use to which the information may be put were raised as issues that require clarification:

- Are these measures primarily clinical tools for the enhancement of patient care? If this is in fact seen to be the primary purpose of this exercise, then it was indicated that 'ticking boxes' does not necessarily equate with good quality clinical care, and that this process will in fact take time away from clinical care.
- Is the information obtained from this exercise to be used primarily as a clinical tool to assess service provision?
- Is the information obtained from this exercise to be used primarily by the Government on an economic level? If this information was to be used on an economic basis, then interest in completing the outcomes measures would decrease. The exercise would not be seen as a process aimed at improving the welfare of consumers, but for economic and political means. The linking of outcomes measures to funding might be viewed as more acceptable if it was believed that they provided a true reflection of change, which these particular measures were not seen to be doing.

Service providers argued that the selected measures failed to address issues relevant to the area of service provision.

Concerns were raised in relation to the issue of how upsetting/distressing this process may be for some consumers.

The validity of the data obtained from questionnaires was questioned. Issues raised here included:

- The possibility that some consumers may respond in a manner which is reflective of changes that they perceive are desired by their service provider.
- Service providers' expectations may influence their ratings, for example the difficulties inherent in preventing previous experience with a client effecting present ratings.
- The ease of manipulation when responding to a questionnaire.

The issue of inter-rater reliability was addressed as a concern when using this method of measuring outcomes. Even if trained in the administration of a specific tool, measurements can often vary widely, which will then have an effect on the reliability of the information obtained.

The issue of payment to services for undertaking routine outcomes measurement was raised, and whether this would be a uniform amount across service settings or would vary dependent on present consultation rates.

Limitations on patients' time was stated to be of particular concern, and the difficulties that would arise in time management if service providers were expected to complete an outcomes measure on each of their respective consumers.

For those who had not previously used tools such as these as a routine component of their assessment process, an emphasis was placed on the need for some sort of extended training in order that these scales be used effectively.

- The appropriateness of using one questionnaire across the differing practice settings was addressed. Questions would necessarily be too broad if one questionnaire were to be used across practice settings, and this would effect the sensitivity of the information obtained.
"...You see I didn't find the scales very useful at all. I think that they were hospital based for a hospital inpatient, and I really found it had minimal usefulness to me. I mean there is no way that you could use that scale on my patients, it's not sensitive from a community point of view. It may well be sensitive from the point of view of should this patient stay in hospital another month or can they go home. I think it might well be sensitive enough for that. But I'm not sure that it's sensitive enough to tell you whether my agoraphobic is ten percent better or my panic disorder person is not having so many attacks, or that they are being controlled on their medication or something like that. I just really found that those scales were irrelevant. I found them very, very coarse, and I think that they were designed and used by people who were not community practitioners..." (Group 1)
- The use of one response format may not provide enough sensitivity to individual consumer situations.
"...the other thing that I found difficult when I was actually trying to respond to these particular questions was that they didn't seem to take into account that people change over a period of time so that at one point in time the client that I was commenting on might have been not getting on particularly well with other people but then there are other points in time where he did do that, and I found that it was very hard to sort of just answer that question as though it just had one answer because it didn't. So I felt almost pushed to select a particular response which would reflect the client over a period of time..." (Group 2)
- To use one standardized outcomes measure on a routine basis was to invite familiarization with the questions it contained.
"...there is an issue relating to clients learning to anticipate questions and answering accordingly. But this could be addressed by rotating outcome measures in some way..." (Group 2)
- The use of one standardized questionnaire with non-English speaking consumers from a variety of ethnic and cultural backgrounds was raised, and it was suggested that ethnicity will effect the process.
"...ethnicity will effect the process with respect to increasing anxiety and obtaining positively skewed data..." (Group 1)

Structure and Content

Regarding the use of questionnaires as measures of treatment outcomes, the following points were raised in relation to the structure and content desired in such a tool:

- While tools are beneficial, it was emphasized that there is a need to develop the best tools that meet scientific rigor. Many participants considered that the measures being investigated did not meet such standards. The issue of *who* defines what is the right tool was raised.
"...whether we like it or not ...the Commonwealth Government is not doing this for any other reason than cost I'm sure, so whether we're doing it with some reluctance is not really a debatable point, the bottom line is that we're going to have to do it. And if we have to do it, then clearly we'd like to develop the best possible tools for follow-up with psychiatrists - that would be fair, equitable, and I guess the whole other scientific rigor - which I have some serious doubts as to whether some of the tools we're using now do..." (Group 3)
- An outcomes measure needs to contain both living skills and functioning information, as well as questions pertinent to symptomatology.
"...these ratings don't quite make that step. What they've done is they're trying to measure function, but they are still, in a way, doing that from a symptom base...so they are not doing things like objectively assessing someone's work record, or what their social interactions are. In a way it's rather symptom based still. There are other questions. But they're still symptom based, and that's one of the weaknesses, that they'd be much more useful if they said things like 'how many times did you go out last week?' and 'how many times did you go out per week before you had treatment, and how many times have you gone out since you've had treatment?'" (Group 3)

- Inclusion of some open-ended questions so that qualitative commentary may be received. It was indicated that the present response formats did not provide enough sensitivity to individual consumer needs.

"...despite the use of standardized measures, the results often do not show how well a client is functioning, that is, the finer points..." (Group 2)

"...questionnaires should include qualitative as well as quantitative questions..." (Group 2)

"...I think that the other question that you could ask people is, 'are you better now?', and frankly we start treatment because they come in and they say they're not well, but we never ask them at the end of it, are you better, or do you think that you're better yet..." (Group 3)

Process Issues

The process should be simple, short and useful.

"...(GP's would) see it as a paper exercise if it was too long, if it took too long, and if it was useless - if they thought it was useless. But if it was short and useful, they'd do it..." (Group 1)

The issue of exactly who would be required to complete the questionnaire on a routine basis, and what the time frame for routine outcomes assessment of might be was discussed. This encompassed factors such as:

- Who will be responsible for deciding who will be required to complete this standardized measure on a routine basis - the service provider, the consumer or the Government?

"...so that would be something that incorporates not just a questionnaire but would have like guidelines for giving that questionnaire, or guidelines for assessing people?..." (Group 1)

"...how many patients with a mental illness are going to be happy about that? Are they ready to comply and to think that information on them is going to be passed on to bureaucrats? It is really going to upset quite a number of them..." (Group 1)

- What will the defining criteria be regarding inclusion or exclusion in an outcomes assessment process?

"...Can I just play the devil's advocate and say that every patient, no matter what they come in with, has a psychological issue...The reason that they are there is that they are anxious about whatever it is that they have come in for today. So I think that the issue of, who is this for, is much more basic than saying that you're schizophrenic, you fill it in..." (Group 1)

- Will this be a mandatory requirement for all providers of a mental health service, regardless of service setting - that is, private psychiatry, public mental health service or general practice setting?

"...we're pragmatic enough to realise that this is an inevitable experience, that whether we like it or not, we're going to be pushed into this..." (Group 3)

"...I wouldn't do this if it was for the bureaucrats, to me these are tools I would use to help me manage my patient..." (Group 1)

- Who will be responsible for deciding upon the time frame by which these measures of outcomes are completed? The time frame inherent in the scales themselves was considered too restrictive, and this would limit the validity of the results obtained, especially for unstable clients.

"...also the time frame. I found that very restrictive. Two weeks. I mean in an illness of depression what is two weeks, especially with a major depressive illness that should be more than six months before you are diagnosed with it..." (Group 1)

- Who will be responsible for the completing of the outcomes measures - the consumer, the service provider, both, or other allied health professionals?

"... as a clinician I don't want to spend all of my time ticking boxes, although I understand the need for clinicians to be doing some of them. On the other hand, I would like to think that all staff could be involved in this..." (Group 3)

"...and what do we do with the consumers that don't or refuse to participate, or refuse to have a measure such as this completed? Are they put in a different box that says no? Or do we just fill one out for them?..." (Group 1)

For this process to be of benefit to anyone (consumer, service provider or Government), then the outcomes measures must be seen by service providers to be 'good measures with face validity that they will own'. Service providers need to be able to see the importance of completing the measures in terms of the impact that they will have on immediate clinical practice. While some service providers stated that they were already using standardized assessment measures on a routine basis, others stated that they had never considered doing so. While tending to be dependent on service setting, time availability was also offered as a contributing factor here.

"...I think that like with anything that you do, you do it more if you think it's intrinsically meaningful, and with me if I was having to fill one of these questionnaires every say two weeks, or one month, I'd have to know that the data, that meaningful use is going to be made of it. If it's just for me to do every two weeks to review my own work, then maybe...For me it has to have meaning - I've got to believe in what I'm doing, I've got to believe in the scale, and I've got to believe that in the end, it is going to have greater returns..." (Group 2)

The way in which the outcomes measures would be administered was discussed. Some offered that pen and paper administration would be burdensome and impractical in terms of storage, and suggested that a computerized format may be more appropriate. Others stated that the type of administration should be sensitive to what the consumer is comfortable with.

"...so I'd imagine a little cubicle with a sort of a touch screen that goes into the system. It is in my computer and doesn't go anywhere. Nobody else is handling it because you have a beginning and an end with the software. That is if you have a computer and I don't. But anyway, then I don't have to worry about big brother you see because he doesn't need analysing - I already have the package and it does it for me..." (Group 1)

"...I think that it depends on the patient, some patients would be quite happy to do it on their own, and other patients would feel more appropriate to do it with someone..." (Group 1)

Concerns

Confidentiality and protecting the privacy of those involved in this process was discussed and regarded as a crucial issue.

"...If the data has identifying features on it, which it has to, because you have to have some identity otherwise it is just across the board Australia wide, and how can anybody use that? Certainly not to shuffle dollars. So whether its comparing hospitals or comparing practitioners or comparing patients, as soon as you do one, the rest are right there..." (Group 1)

The purposes of outcomes measurement and the use to which the information may be put were issues that were raised throughout most of the discussions:

- Is the information obtained from this exercise to be used primarily as a clinical tool to enhance patient care? If this is in fact seen to be the primary purpose of this exercise, then it was indicated that 'ticking boxes' does not necessarily equate with good quality clinical care, and that this process will in fact take time away from clinical care.

"...when the amount of paperwork increased because the amount of time you spend on paperwork is taking away from the actual contact time that you have, and that's what happened (accreditation).....they are now looking at other issues relating to quality of care, because I don't think that you can correlate the number of boxes that are ticked with good patient care, and the other thing that happens certainly when accreditation came in was that staff who were involved in middle management were having to spend a huge amount of time doing all of this paperwork which was actually taking time away from patients rather than curing them..." (Group 3)

"...if we start rating them then they are going to feel like sheep..." (Group 2)

"...one of the dangers that I can see in actually filling in these forms on a regular basis could be that you would be so caught up in filling in the forms that you stop looking at your patient as the person that they are...I think that you could lose your patient in the middle of the paperwork..." (Group 1)

- Is the information obtained from this exercise to be used primarily as a clinical tool to assess service provision? If this is perceived to be the primary purpose of this exercise, then the fundamental issue of attribution, that is, that change need not necessarily be associated with the intervention provided by a service, needs to be addressed. This includes factors such as changes in accommodation and the presence of other community supports. The issue of staffing numbers at the service was also discussed.

"...then you have clients like S...who knows that she will probably improve out of that particular environment, but chooses not to move out of that particular environment and therefore that particular person on that questionnaire is not going to... You know, it looks like the service is not terribly effective. But it's not really the service as such..." (Group 2)

"...I think that if the assumption is that if you make the intervention and have predicted that some of these things will change and they change, then we've got to assume that it's got something to do with the intervention... I'm saying that the intervention could be so multidimensional - how do you then say that it's due to this or that?..." (Group 3)

"...our resources at the moment are so limited, any change that we might be finding could be due to chance - evolution of the earth, illness or medication, and until the community gets adequate resources to be able to really feel that they can make a change, then the outcomes are not going to..." (Group 2)

- Is the information obtained from this exercise to be used primarily by the Government on an economic level? If this information was to be used on an economic basis, then interest in completing the outcomes measures would decrease. The exercise would not be seen as a process aimed at improving the welfare of consumers. The linking of outcomes measures to funding might be viewed as more acceptable if it was believed that the measures provided a true reflection of change, which these particular measures were not seen to be doing.

"...the Commonwealth Government is not doing this for any other reason than cost I'm sure, so whether we're doing it with some reluctance is not really a debatable point, the bottom line is that we're going to have to do it...I guess you get the message that there are all sorts of two edge swords in all of this, you know we're supplying information and then it's used against us..." (Group 3)

"...It's not a problem providing that it's right, but it's not right, and that's the point.....and even if it is right, and it doesn't suit their purposes, then they have something else imposed on us, and they've got a track record..." (Group 3)

The measures of outcomes trialed in this study failed to ask any questions relevant to the area of service provision. It was stated that the objectives of treatment involve a negotiation between individual consumers and their respective service providers, and that these measures have not yet made this step.

"...the other approach that I feel is a useful one is that when people come for treatment, is that you have them negotiate objectives with you for treatment, so that they've got realistic objectives for treatment, and you've got realistic objectives for treatment And then you can see how many of those you achieve, or what degree that you have achieved, or what degree you haven't achieved - my objective is, you know, so that I can go out shopping, or my objective is that I am able to go back to work again. And I think that that's not a bad way of doing it..." (Group 3)

"...I think that it plays a big role in how they might direct their own recovery...open them to a process of negotiation..." (Group 2)

Concerns were raised in relation to the issue of how upsetting/distressing this process may be for some consumers.

"...and what do we do with the consumers that don't or refuse to participate, or refuse to have a measure such as this completed, are they put in a different box?..." (Group 1)

The validity of the data obtained from questionnaires was questioned. Issues raised here included:

- The possibility that some consumers may respond in a manner which is reflective of changes that they perceive are desired by their service provider.

"...I could see them sitting and looking at you wondering what do you want me to say..." (Group 1)

- Service providers' expectations of the consumer may influence ratings.

"...I found it really hard to sit down and to do the questionnaire. And I was very aware of her confidentiality, because I found it really difficult not to use that information in my daily work and not to document. Because her perception of her life was so incredibly different to how I think she functions, and I think she's got this incredibly dysfunctional life with anxiety and she thinks that everything is great, so that was a really difficult situation to be in, you know, was not documenting that information or not to expand on it more purely for my own interest..." (Group 2)

- The ease of manipulation when responding to a questionnaire.
"...can foresee trying to keep hold of certain clients to manipulate results..." (Group 2)

The issue of inter-rater reliability was addressed as a concern when using this method of measuring outcomes. Even if trained in the administration of a specific tool, measurements can often vary widely, which will then have an effect on the reliability of the information obtained.

"...I measured my most disabled client in another study, and her social worker provided one as well, and there was a twenty-twenty difference in our estimates of the same day. And that was significant. But then I think that that had something to do with knowing her long term, and seeing how far she has progressed. For the new person who has only known her for a few months, she was very pessimistic and she hasn't achieved her goals and therefore this person is making no progress. Whereas I thought that her progress was incredible over a two or three year period..." (Group 2)

The issue of payment to services for undertaking routine outcomes measurement was raised, and whether this would be a uniform amount across service settings or would vary dependent on present consultation rates.

"...but maybe the government could consider though, if they wish us to do these outcome studies, some sort of payment.....The government only pays for face to face contact with the patient, we don't get anything else, and a lot of that is eroded by filling out forms and so forth..." (Group 3)
"...if it takes ten to fifteen minutes to do one of these questionnaires properly, I mean that is almost a consultation..." (Group 1)

Limitations on time was stated to be of particular concern, and the difficulties that would arise in time management if service providers were expected to complete an outcomes measure on each of their respective consumers.

"...I don't think there is any clinician who would spend that amount of time filling out forms..." (Group 3)
"...It just wouldn't be done, it just isn't possible. It would be too time consuming, and you're just not willing to do that..." (Group 1)

For those who had not previously used tools such as these as a routine component of their assessment process, an emphasis was placed on the need for some sort of extended training in order that these scales be used effectively.

"...maybe we need some more training in general mental health issues and general practice, or up-skilling..." (Group 1)
"...yeh, not just questions one to five, but probably some expertise behind how to utilize the tool for specific circumstances..." (Group 1)

4.2.6 Findings from the Peak Service Provider Association Submissions

Six professional groups were contacted and requested to comment on the *face validity* and usefulness of the six selected measures as well as the use of computers in collecting outcomes data. Submissions were received from the Royal Australian College of General Practitioners, the Australian Association of Occupational Therapists, and the staff of the Bankstown-Lidcombe Hospital (representing the Royal Australian and New Zealand College of Psychiatrists).

Members from the Royal Australian College of General Practitioners envisaged some difficulties in relation to integrating routine outcomes measurement into everyday practice, such as:

- The time commitment required to complete a selected measure, which was viewed as a potential difficulty for both service providers and for consumers;
- The provision of a suitable environment in which to complete a selected measure, given this busy and pressured setting;
- The practicality of administering a measure 'routinely' demands clarification, that is, a time frame must necessarily be defined;
- The suggestion that each disorder might need to be considered separately; and
- Sensitivity issues relating to the ability of all measures to detect mental disorders commonly encountered in General Practice, were questioned.

The College indicated support for the use of computer data collection and storage systems for

clinically useful activities such as clinical audits, however, cautioned that the introduction of such systems must be made gradual so that potentially complex issues inherent in this process may be addressed. Of the six selected measures proposed the HoNOS was viewed as the most acceptable and practical for General Practice use.

The Australian Association of Occupational Therapists (AAOT) emphasized that they prefer a measure which is discipline sensitive; that is, a measure which places appropriate emphasis on a range of occupational skills which are of particular interest to occupational therapists (e.g., self-care, life skills, social skills, productivity, leisure, etc.). While stating that each of the three measures were viewed to have both strengths and limitations, this group indicated a preference for the LSP and the RFS as most suitable to meet their needs as service providers. Of the three selected measures to be completed by consumers, the BASIS-32 was preferred, although once again it was stated that each of these respective measures had their difficulties. The AAOT also stated that it is unlikely that one single outcomes measure would be applicable to all populations and age groups. It was also considered more appropriate to involve both the consumer and the clinician in such a process, and therefore it was felt that two measures should necessarily be used in assessing consumer outcomes. In relation to the use of computers in the collection, storage and analyses of outcomes information, the following points were raised:

- The acceptability of this method to the clients whose information it will be that is stored on computers; and
- The ability of the service to provide secure computer systems that are accessible to relevant staff in terms of entry and retrieval.

The AAOT also expressed concern as to whether this information will actually be analyzed and used to provide better services for consumers. The AAOT stated that, in the opinion of their group, *"...one of the key areas that will probably need to be addressed is that staff have to be convinced to incorporate outcomes measures into routine practice and see the value of doing so"*.

The third and final Group to provide feedback was a group of staff from the Bankstown-Lidcombe Hospital (on behalf of the RANZCP). This group indicated a preference for the use of a single measure, with the HoNOS and the BASIS-32 being offered as the most highly regarded by group members. Time management issues relating to the completion of such a measure by either a service provider or a consumer was again raised as a potential difficulty.

4.2.7 Discussion of Results of Feasibility Analyses

The broad goals of this study were to provide advice as to which of the six selected measures would be most suitable for the routine measurement of consumer outcomes, and to provide advice as to issues of relevance in regard to the implementation of such a measure in everyday clinical practice. In Section 1 we argued that the primary goal of outcomes measurement is the facilitation of decision-making processes across various levels of mental health service delivery. To facilitate such decision-making a measure must be feasible, that is applicable, acceptable and practical. How, then, do the results reported above inform us as to the feasibility of using any of the selected measures, or the feasibility of implementing standardized processes, for the routine assessment of consumer outcomes in mental health?

Consumers and service providers generally supported the concept of implementing routine assessment of consumer outcomes in mental health. This is consistent with the findings of Andrews et al. (1994). The main points of concern raised by consumers and service providers were in relation to how such a process would actually work, and moreover the relevance of such processes to the everyday needs of consumers and services. The data from the focus groups suggests that the pragmatics of implementing effective systems of outcomes assessment are more complicated than simply choosing a questionnaire. The comments provided by consumers and service providers highlight important *process* issues that need to be addressed before any system of outcomes assessment can become feasible.

The utility data indicated that the majority of consumers who used it considered the MHI to have more relevant questions and to be a more useful measure than the BASIS-32, and to be a *better* measure than the SF-36. In the latter case, consumers' comments suggested that the MHI was

more relevant for assessing mental health issues. Many of the items of the SF-36 measure physical health or 'general' health. However, some consumers' comments suggested that the time period assessed by the MHI (1 month) presented difficulties for recall of feelings (a problem shared with the SF-36), and that some of the language used by the measure was either inappropriate or old-fashioned.

For service providers, the utility data indicated different patterns of responses across practice settings. Service providers rated the LSP significantly more highly than either the HoNOS or RFS in the *public psychiatry setting*. In the *general practice* or *private psychiatry settings*, no single measure was rated significantly more highly than the others. As well *general practitioners* average ratings for the measures were significantly higher than those in the *public psychiatry settings*, which in turn were significantly higher than those in the *private psychiatry settings*. A number of factors may account for these findings that are also reflected in the qualitative data. For instance, the *general practitioners* may be welcoming of resources to help meet the needs of their patients who have a mental illness or disorder. Many service providers in the *public psychiatry settings* were already familiar with the LSP. On the other hand, service providers in the *private psychiatry settings* may have expectations that such measures be short and comprehensive, yet assess the general social functioning of the individual consumer.

The comments provided by consumers and service providers also highlighted a range of implementation issues, including confidentiality, the ability of consumers to complete the measures when they are unwell, and the negative impact on consumers' that completing such questions may produce. Other factors to emerge were the need for simplicity, brevity, and comprehensiveness in the assessment process; improved assessment of general social functioning and quality of life; measures which meet accepted criteria of scientific rigor; standardised training and instruction in the use of the measures; and a focus on process issues.

The eight focus groups produced a large and diverse amount of information relating to the feasibility of implementing routine outcomes measurement in mental health services. While not necessarily constituting a representative sample of the target populations from which participants were drawn, each participant presented as aware of, and concerned with, the potential impacts that this may present for themselves and others.

In general, both consumers and service providers were supportive of the concept of routine outcomes assessment. However, a range of important pragmatic issues were raised:

- How would such a process work?
- What outcomes would be measured?
- Who would control the information collected?
- For what purposes would the information be used?
- How would decision-making processes be formally influenced by the information collected with these measures?
- How suitable are the selected measures for the assessment of mental health outcomes across diagnostic groups and in differing practice settings?
- How would confidentiality be protected?
- What sort of burden would be placed upon consumers and service providers and the clinical process?

The issue of attribution of change was also discussed. It was pointed out that where a measure did detect change in a person's condition, the attribution of this change to any single factor, including intervention, was difficult. Service providers and consumers, alike, were concerned in regard to this point. It was suggested, for example, if resource or service delivery implications were to be attendant upon the results, people may respond in a biased manner.

The resource implications of implementing systems for the routine assessment of outcomes were discussed at length with issues such as the time necessary to complete the measures, the site of completion, and the reimbursement for this process, being of particular interest and concern. Confidentiality and anonymity were also highlighted as being of particular importance. Other practical issues such as the proposed methods of administration, storage, and the accessibility of the obtained information were remarked on.

Not only were time and financial costs considered problematic, but concerns were raised in relation to the negative impact, in particular feelings of distress, for certain consumers, as a direct consequence of completing these forms. This was a factor that the research team became aware of through anecdotal reports received from service providers in the private inpatient setting. On at least two occasions consumers became very upset after completing some of the questionnaires.

Finally, the relationship between consumer outcomes and the quality of service provision arose as an important issue. Some consumers felt that the questions included in the selected outcomes measures did not address issues of service provision. Others, however, stressed that their personal outcomes were independent of their opinion of the service they received. Nevertheless, a clear need was recognised for consumers to be asked whether they felt that the service being provided to them was effective in meeting their needs.

In relation to the model of decision-making in mental health presented above (see Table A2) the information reported here can be interpreted as being concerned mostly with *micro* level decision-making; that is, involving individual consumers and service providers. However, mental health services are not provided in a vacuum and most of the participants, especially those involved in the focus groups, were aware of, and raised issues relevant to, *meso* and *macro* levels of decision-making. Moreover, they raised important questions about how information collected at the consumer-service provider level will be analysed and interpreted by service managers, policy-makers, bureaucrats, etc.

Consumers and service providers are aware that they will be the source of outcomes data, and that such data may have the potential to influence decisions in relation to service delivery. In this respect, while goodwill was expressed towards the use of such data to increase service effectiveness and knowledge of mental health, concern was expressed by both consumers and service providers in relation to the potential abuse of such information, and the burden to stakeholders involved in the process. As stated above, the most pertinent aspect of these results is the need for greater information about how the process of outcomes assessment will actually work and what it will mean for all stakeholders. Until such issues are confronted and resolved the question of the feasibility of any specific questionnaire for decision-making will remain relatively academic.

The process of including focus groups proved to be extremely valuable in collecting data in relation to the feasibility of implementing measures for the routine assessment of consumer outcomes. No doubt they will prove to be very useful tools in the processes of implementing and evaluating systems of outcomes assessment. Readers are directed to the Consumer Consultant's report on the focus group processes that includes a number of important points that may be of benefit to future efforts to utilise such groups in a consultative capacity.

Many of the points stated above were also raised in the three submissions received from the service provider groups. Relating this discussion back to the choice of a measure or measures the prevailing theme appears to be that the practicalities of outcomes measurement in Australia need to be examined more closely, but that some of these measures do possess potential to provide useful information in certain circumstances. For instance, each of the measures received support from different participants. It is clear that the introduction of any system must start somewhere, and these measures provide a means by which services can at least begin to explore the measurement of outcomes as they implement systems that address the needs and concerns of all stakeholders.

4.3 The Selected Measures as Measures of Mental Health Status

4.3.1 Validity

The analyses outlined in this section set out to:

- Examine the association between the scores recorded by the measures;
- Examine the convergence between results obtained on the selected consumer measures and those obtained on the selected service provider measures; and
- Examine the patterns of observed differences in scores across diagnostic groups and practice settings, and the degree to which findings reported on the selected measures correspond to observations from alternative measures of mental health status.

4.3.1.1 Associations between Scores for the Selected Measures

Consumer Measures: In general, the mental health composite scores of the three selected consumer measures (see Table B3.1) were highly correlated (Range: -0.79-0.86). As expected, both the MHI and BASIS-32 scores had stronger correlations with the *mental component summary* scores (MCS) of the SF-36 than with the *physical component summary* scores (PCS). The MHI Index scores, however, did correlate significantly with the PCS scores ($r=0.53$).

In general significant correlations were observed between the scale scores of each measure. The scores of the BASIS-32 correlated significantly with those of the MHI in the expected directions (see Table B3.2). They ranged from -0.31 (*impulsive/ addictive behavior* and *emotional ties*) to -0.81 (*BASIS-32 Average* and *MHI Index* scores). All correlations were significant at the 0.05 level or better. The scores for *relation to self/others*, *depression/anxiety* and *daily living/role functioning* scales of the BASIS-32 displayed stronger associations with the scale scores of the MHI than did those for *impulsive/addictive behaviour* and *psychosis* scales.

There were significant correlations between the scores of the BASIS-32 and the 'mental health' scale scores of the SF-36 (*role-emotional*, *social functioning*, *mental health*, and the MCS scores) as well as the *general health* and *vitality* scores (see Table B3.3). Again, the *relation to self/others*, *depression/anxiety* and *daily living/role functioning* scales of the BASIS-32 displayed stronger associations with the scale scores of the SF-36, than did those of the *impulsive/addictive behaviour* and *psychosis* scales.

Correlations across most scale scores of the MHI and SF-36 were statistically significant, with coefficients generally significant at the 0.01 level (see Table B3.4). The strongest associations involved the *mental health* sub-scale and the MCS scores of the SF-36. Of the five sub-scales the MHI *depression* sub-scale displayed the strongest correlations with the SF-36 *mental health* sub-scale score ($r=-0.88$).

Service Provider Measures: Significant correlations between the composite scores of the three selected service provider measures were observed (see Table B3.5). There were also significant correlations between scale scores that purport to measure related constructs, in particular those assessing daily living skills and social functioning. In general, significant relationships were observed between the LSP and RFS scores, respectively, on those HoNOS items which measure social functioning as opposed to symptomatology. There were, however, also significant associations observed on the *cognitive problems*, and *hallucinations/delusions* items, which is reflective of the relationship between psychotic symptoms and poor social functioning (see Tables B3.6-B3.8).

As would be expected the LSP and RFS scores consistently displayed significant correlations in the predicted directions (range: $r=0.38$ to 0.78). All correlations were significant at the 0.01 level. There were significant correlations between the LSP total score and most of the 12 HoNOS scale scores. These ranged from -0.28 (*self-injury*) to -0.61 (*social skills*). The strongest associations were with those scores of the HoNOS that measure 'living skills'. There were also relatively strong correlations with the *problems with hallucinations/delusions* scores ($r=-0.47$). Significant correlations were observed between the HoNOS *social skills* sub-scale and *total* scores and the RFS Scores (Scales 1-4, and GRFI).

Convergence of Selected Service Provider and Consumer Measures: Table B3.9 displays the correlations between the composite scores of the selected consumer and service provider measures. In general the correlations were low. Of the three selected service provider measures, the HoNOS total scores consistently demonstrated significant, though low, correlations with the summary scores of the selected consumer measures.

The results suggest that, in so far as being global measures of mental health, the scores obtained from the three selected consumer measures are closely associated. Different patterns of associations are observed across the sub-scale scores for the three measures. This is due to the fact that the measures assess different, albeit related, dimensions of consumers' mental health status, and general health in the case of the SF-36.

Similarly the composite scores of the three selected service provider measures demonstrated strong correlations, as did those sub-scale scores most closely related to the dimensions of social disability. Again, each measure purports to assess different aspects of mental health with the HoNOS including a greater range of assessment dimensions.

There were weak associations between the consumer measure scores and the service provider measure scores. In light of previously reported findings, this result is not unexpected. The service provider and consumer measures differ in content, and are influenced by different perspectives: that of consumers and that of service providers. This point will be discussed further, below.

4.3.1.2 Analyses of Differences in Mental Health Status across Diagnostic Groups and Practice Settings

The analyses described in this section set out to examine how groups of consumers defined by *diagnostic groups* and *practice settings* differed in relation to the selected measures scores; and whether differences between groups were consistent with those observed using alternative measures of mental health status.

Analyses of Mental Health Status Using Alternative Methods

Simple measures (visual analogue scales and modified clinical global impressions scales) were used to describe the mental health status of consumers independently of the six selected measures.

Visual Analogue Scales (VAS): Four visual analogue scales (VAS) provided a subjective report of the mental *well-being*, *distress*, *depression*, and *anxiety* experienced by consumers in the seven days preceding administration. Table B3.10 displays summary statistics for the four scales and the total scores.

Significant correlations were obtained between both the *scale* and *total* scores (Pearson's *r*, range: 0.53 to 0.86). Test-retest correlation coefficients were also strong (Pearson's *r*, range: 0.71 to 0.83; see Table B3.11). Internal consistency coefficients (α) for the total scale were 0.86 at Time 1, and 0.91 at Time 1(Retest). Given this high level of internal consistency a *total score* was calculated by reversing coding response to the well-being scale, and summing the four scales. In this manner we hoped to obtain a simple summary statistic for self-rated mental health status.

There were significant correlations between each of the four VAS *scale* and *total* scores, and the composite scores of the selected consumer measures in the expected directions (Range:-0.39 to -0.72) (see Table B3.12). The strongest association was between the VAS *total* scores and the MCS scores of the SF-36 ($r = -0.72$).

One-way ANOVAs indicated significant differences in VAS *total* scores across *diagnostic groups*, $F(2,180)=3.0$, $p=0.05$, and *practice settings*, $F(2,180)=3.9$, $p<.05$. (see Figures 6 and 7). The *affective disorders group* ($M=199.2$, $SD=93.7$) reported significantly higher total scores than the *schizophrenia group* ($M=162.5$, $SD=83.2$) ($p=.056$); and, the *private psychiatry group* ($M=208.5$, $SD=87.3$) reported significantly higher scores than the *public psychiatry group* ($M=162.4$, $SD=111.7$).

There were no significant multivariate effects obtained for either independent variable on the linear

combination of the four individual VAS scales. However, univariate effects were observed on a number of variables. There were significant differences between the mean *anxiety* scores of the *affective disorders group* ($M=56.0$, $SD=26.7$) and *schizophrenia group* ($M=43.5$, $SD=24.8$): $F(2,174)=3.8$, $p<.05$. There were also significant differences across *practice settings* for *well-being*, $F(2,174)=3.1$, $p<.05$, and *depression*, $F(2,174)=4.6$, $p<.05$. The *private psychiatry group* recorded significantly lower *well-being* scores ($M=46.7$, $SD=24.4$) than the *public psychiatry group* ($M=58.5$, $SD=27.0$); and the *public psychiatry group* recorded significantly lower *depression* scores ($M=35.1$, $SD=28.5$) than the *private psychiatry group* ($M=50.6$, $SD=26.0$).

Modified Clinical Global Impressions Scale (CGI): Table B3.13 displays frequencies and summary statistics for the *modified CGI scale*. Surprisingly, twenty-two percent of consumers were classified by their service provider as ‘Normal, not at all ill’ (7%) or ‘Borderline mentally ill’ (15%). That is, almost a quarter of the people involved in this study were not (at least in the opinion of their service providers) currently ill. This might reflect different perspectives of ‘mental illness’; these consumers may have been receiving a mental health service for reasons relating to ‘disability’ rather than illness, or required ongoing support to prevent relapse. This result, however, raises questions about the sample of consumers, a point which will be discussed later.

There were significant correlations between composite scores of the service provider measures and the scores for the *CGI*: 0.44, *HoNOS total score*; -0.31 *LSP total score*; and, -0.53, *Global Role Functioning Index* (RFS). Higher *CGI* ratings representing more severe levels of mental illness, were associated with higher *total* scores on the *HoNOS*, and lower *total* scores on the *LSP* and *RFS*. There was a weak correlation between the *CGI* and the *VAS total* scores ($r=0.13$) suggesting a poor association between service providers’ and consumers’ global assessments of severity of illness.

CGI scores differed significantly across *diagnostic groups*, $F(2,171)=3.4$, $p<.05$, and *practice settings*, $F(2,171)=8.0$, $p<.001$ (see Table 3.14 and Figures 8 and 9). On average, people with *schizophrenia* ($M=3.7$, $SD=1.1$) were considered to have greater severity of illness than people with *affective disorders* ($M=3.2$, $SD=1.3$). People receiving services in *private psychiatry settings* ($M=4.0$, $SD=1.0$) were considered as having greater severity of illness than people receiving services in either the *public* ($M=3.4$, $SD=1.2$) or *general practice* ($M=3.3$, $SD=1.2$) settings. This is consistent with the data reported in section 4.1.

Given the results of the *VAS* and *CGI* data we might expect to observe similar patterns of findings across *diagnostic groups* and *practice settings* for data obtained from the six selected measures. That is, consumers in the *affective disorders group* should report higher levels of mental health problems than people in the *schizophrenia group*; and those in the *private psychiatry group* should report higher levels of mental health problems and depression, and lower levels of well-being, than people in the *public psychiatry group*.

For the service provider measures, we might expect to observe scores suggesting more mental health problems, and greater difficulties with general social functioning, for people in the *schizophrenia group* as opposed to the *affective disorders group*; and people in the *private psychiatry setting* having more mental health problems than people in the *public psychiatry setting*.

Differences in Scores for the Selected Measures of Mental Health Status across Diagnostic Groups and Practice Settings

Univariate and multivariate ANOVAs, and Kruskal-Wallis tests, were used to investigate whether the scores of the six selected measures differed significantly across the three *diagnostic groups* and the three *practice settings*, and whether these results were consistent with the findings reported above. Where a significant multivariate effect was observed, univariate analyses were undertaken. Post-hoc analyses with Scheffe tests were used to investigate differences across the levels of the independent variables. Summaries of the results of analyses are included in Tables B3.15 (a)-(d).

Selected Consumer Measures

Differences across *diagnostic groups* and *practice settings* were observed on a number of scale scores for the consumer measures. The general pattern of results suggested significant differences

between the *affective disorders group* and the *schizophrenia group* (and to a lesser extent the *anxiety disorders group*) on the scores of the BASIS-32 (*average, relation to self/others, depression/anxiety, and daily living/role functioning* scores), MHI (all scores), and SF-36 (*mental component summary, bodily pain, social functioning, and mental health* scores), with the former group reporting scores indicative of significantly higher levels of mental health problems.

Differences were also observed across *practice settings* for the scores of the BASIS-32 (*average, relation to self/others, depression/anxiety, and daily living/role functioning*), MHI (all except *emotional ties*) and SF-36 (*mental component summary, general health, vitality, role emotional, social functioning, and mental health* scores). Here the results were generally indicative of significant differences between the *private psychiatry* and *public psychiatry groups* with the former group reporting scores suggestive of higher levels of mental health problems.

Figures 10-15 (Appendix 4) illustrates the observed differences between *diagnostic groups* and *practice settings* on the relevant dimensions of the consumer measures.

Selected Service Provider Measures

The data from the service provider measures provide different pictures of the sample of consumers. Significant differences were observed between the ratings for people in the *schizophrenia group* and people in the *affective* and *anxiety disorders group* on the scores of the LSP and RFS (*working/productivity* and *independent living skills*), with scores for the former group indicative of significantly higher levels of problems of general functioning.

There were no significant differences across *practice settings* on the RFS scores. However, differences were observed for all LSP scores: in each case significant differences were observed between the ratings for people in the *public psychiatry group* and people in the *private psychiatry group* (and to a lesser degree the *general practice group*), with scores for the former *group* indicative of significantly higher levels of difficulties with general social functioning.

Data from the HoNOS scores suggested that the people in the *schizophrenia group* were experiencing greater *problems with hallucination/delusions, general functioning, living conditions, and activities of daily living*; people with *affective disorders* were experiencing greater problems with *depressed mood*; and people with *anxiety disorders* were experiencing greater problems with *other behavioural and emotional problems*.

The *private psychiatry group* received significantly higher ratings in relation to *problems with self-injury* and *depressed mood*, while the *public psychiatry group* were rated as being more likely to experience *problems with hallucination/delusions*. Again, these results are consistent with the findings reported earlier in relation to the characteristics of the sample. Of the three service provider measures the HoNOS data invites a number of comparisons of the group which is indicative of the breadth of dimensions it covers.

Figures 16-20 (Appendix 4) illustrate the observed differences between *diagnostic groups* and *practice settings* on the relevant dimensions of the service provider measures.

The two sets of data (consumer and service provider) provide different perspectives on the mental health status of this group of consumers. Such results reflect the different contents of the measures: in general, the consumer measures concentrate on assessing how the consumer is 'feeling', whereas the service provider measures rely on observational data and what the service provider feels the consumer is experiencing and 'can do'. As well, the LSP and RFS focus on dimensions of mental health relevant to general social and role functioning.

Other explanations for the differences in rating patterns include the possibility that service providers generally rated people with schizophrenia as having more severe levels of mental illness because of the long-term prognosis for people with schizophrenia in comparison to other forms of mental illness. Yet people with schizophrenia assessed themselves quite differently to people with other disorders, in a manner which suggested comparatively lower levels of mental health problems. This may be related to their condition, with one symptom of schizophrenia being lack of insight. An alternative

explanation is that the people with schizophrenia were currently well and able to complete the form. In making their judgments about how well they were feeling they may have based their assessments on comparisons to how bad things have been in the past. Likewise, service providers may be likely to base their assessments on their past experiences with a particular consumer.

Finally, it could be argued that not only were the measures designed to assess different aspects of mental health, but that consumers and service providers adopt different processes when making assessments of mental health status. After all the data provided by the service provider is based on their experience of the consumer and on what the consumer tells them. The consumer's data on the other hand is influenced by their illness history, their feelings at the time of completing the questionnaire, and their expectations as to how the information will be interpreted.

The findings reported here are specific to the present sample. The aim of the study was not to investigate differences in outcomes across different practice settings, per se; and the sample is clearly not representative of the population of people who would be expected to receive such services. The findings do, however, raise interesting questions for future investigations with regard to the validity of the selected measures.

4.3.2 Reliability

The analyses described in this section set out to:

- Examine the test-retest reliabilities (using Pearson's correlation coefficients) for the consumer measures;
- Examine the internal consistency (using Cronbach's coefficient alpha) of the measures; and
- Examine the inter-rater reliabilities of the service provider measures by measuring the association between the scores of the service provider measures (i.e., the HoNOS, LSP and RFS) provided by service providers, and scores from the same measures completed by *consumers*.

4.3.2.1 Test-Retest Reliabilities

Table 3.16 (a)-(c) display the test-retest reliabilities calculated for the present data set for the three consumer selected measures, as well as previously published figures. In general the results for the present data set are consistent with those reported from previous data. For instance, all coefficients for the MHI are above 0.90; coefficients for the BASIS-32 range from 0.84 to 0.90; and those of the SF-36 range from 0.59 to 0.92.

4.3.2.2 Internal Consistency

Results of analyses of internal consistencies are displayed for the composite and scale scores of the selected measures in Tables B3.16 (a)-(c) and B3.17, along with previously published values. In general the results are consistent with those reported in earlier studies (see Section 1.3.). For instance, coefficients for the MHI range from 0.80 to 0.97; coefficients for the BASIS-32 range from 0.65 to 0.90; and those of the SF-36 range from 0.77 to 0.91. Coefficients obtained for the LSP and RFS were in the expected range. Although the alpha for the *total* score of the HoNOS was 0.73, the internal consistency coefficients for the sub-scales (except the *social skills* sub-scale) were quite low.

4.3.2.3 Inter-Rater Reliabilities of the Service Provider Measures

Table B3.18 displays the correlations between the composite scores of the selected service provider measures: service provider and consumer versions. In general the correlations between the service provider assessments and the consumer assessments on the same measures were low.

Strictly speaking this was not really a test of 'inter-rater' reliability because these particular measures (i.e., the HoNOS, LSP and RFS) were designed to be completed by service providers and not consumers. However, the results still provide a useful examination of the convergence between

service providers' and consumers' assessments of consumers' mental health status. As well the results reinforce the observation of the differences in 'how' consumers and service providers respond to these measures.

Interesting, too, is the data presented in Table B3.19 which displays the correlations between the composite scores of the selected consumer measures with the composite scores of the consumer-versions of the selected service provider measures. Correlations were significant (except those for the physical component score of the SF-36), ranging from -0.56 to 0.79 (all p 's<.01). Even though the measures assess different dimensions of mental health, most of the associations were strong and in the predicted directions.

It would appear from this data, and that reported above, that not only do the measures assess different aspects of mental health, but that consumers and service providers differ in how they make their assessments. The implications of these observations for the implementation of measures in routine clinical practice will be discussed below.

The results presented here in relation to the measures as measures of health status provide some indication of the validity and reliability of the measures. It has been shown that the scores of the measures converge to the extent that the sub-scales measure similar dimensions. Furthermore the composite score of the three consumer measures exhibited high levels of association.

As tools for comparing sub-groups of consumers according to *diagnostic groups* and *practice settings*, the measures displayed consistency and produced results which were compatible with expectations as well as being in agreement with data recorded on alternative measures of mental health status. A clear difference emerged between the data obtained from the three consumer measures and that obtained from the LSP and RFS. This is to be expected given the different orientations of these measures. The data from the HoNOS, on the other hand, provided detail for a wider range of mental health dimensions.

The consumer measures displayed test-retest reliabilities in the expected range, and all scores displayed relatively good internal consistency coefficients, except some of the HoNOS sub-scales. The low internal consistencies for the HoNOS sub-scales are probably due to the small number of items included in each scale. Observations of low internal consistencies for the sub-scales are not a great problem for this measure as it is likely that the 12 scale items are probably more useful for practical purposes.

Low levels of convergence were observed between the scores of the consumer outcomes measures and those of the service providers, which appeared not to be due to simple differences in scale formats. This raises questions about the need to obtain multiple perspectives of mental health status.

4.4 Selected Measures as Measures of Change in Mental Health Status

The analyses described in this section set out to:

- Examine the consumers' and service providers' reports of global change in the consumers' mental health status;
- Examine change scores on the selected measures from Time 1 to Time 2;
- Examine the association between change scores for the selected measures; and
- Examine the association between global ratings of change and change scores from the selected measures.

4.4.1. Descriptive Statistics for the Global Change Ratings

Tables B4.1 and B4.2 display frequencies and summary statistics for the *Global Change Ratings* (consumer and service provider ratings, respectively). Data from the ratings were used to define *change groups: better, no change or worse*. Of consumers, 60% reported some change for the better in their mental health over the three-month period between ratings, 19% reported no change, and 21% reported feeling worse. For service providers, 67% reported a change for the better in the

mental health of their consumers, 16.5% reported no change, and 16.5% reported a change for the worse. While there was a statistically significant convergence between ratings (*Spearman's rho*=0.25, $p<.05$), a direct correspondence between service providers' and consumers' assessment of change was observed in only half of the cases (see Table 4.3).

4.4.2 Differences in Changes in Scores for the Selected Measures from Time 1 to Time 2 across Change Groups

The analyses described in this section had two objectives. The first was to determine whether differences between Time 1 and Time 2 scores for the six selected measures were statistically significant. The second objective was to examine differences between Time 1 and Time 2 scores across the three *change groups* (as defined by consumers and service providers, respectively) to investigate whether there were different patterns of changes in scores amongst the groups. For instance, we would expect that people who indicated an improvement in general mental health would obtain scores on the selected measures that reflected this. Similarly, we would expect that people in the *no change* and *worse groups* would obtain scores consistent with their feelings in relation to changes in general mental health.

Mixed design repeated measures ANOVAs (with *change group* as the between subjects independent variable and *time of administration* as the within-subjects independent variable), were conducted to examine the statistical significance of observed differences between Time 1 and Time 2 scores of the selected measures. Wilcoxon matched-pairs signed rank tests were used to analyse change on the 12 HoNOS item scales. The results of these analyses are summarised in Tables 4.4 and 4.5.

Selected Consumer Measures

On average the scores of the BASIS-32 (except the *impulsive/addictive behaviour* and *psychosis* scores) improved from Time 1 to Time 2 (see Table 4.5a). Similarly there were significant differences between the mean Time 1 and 2 scores for the *depression*, *well-being* and *distress* scores of the MHI.

A significant *time of administration x change group* interaction was observed for the *depression/anxiety* scale of the BASIS-32, $F(2,74)=3.7$, $p<.05$. Post-hoc analyses indicated a significant difference between the Time 1 means for the *worse* and *no change* groups. At Time 2 there was also a significant difference between the means for the *better* and *worse* groups. In this case, the mean scores of the *better group* were significantly lower at Time 2 than at Time 1.

For the SF-36, significant interactions between *time of administration* and *change group* were observed for the *mental health*, *vitality* and *mental component summary* scores (see Figure 21). In each case the *better group* achieved statistically significant gains in scores. Changes from Time 1 to Time 2 for the other *groups*, however, were not statistically significant.

Significant *time x change group* interactions were observed on all of the scores of the MHI (except the *emotional ties* variable). In each case the mean scores for the *better group* improved significantly from Time 1 to Time 2. The scores for the *no change group* did not differ significantly over time. In the case of *anxiety* and *general positive affect* scores the means of the *worse groups* indicated a significant decline from Time 1 to Time 2. Similar trends are observed on the other scores however they did not reach statistical significance.

There is evidence, then, on each of the consumer measures, of people in the *better group* reporting 'better' scores at Time 2, and people in the *no change group* reporting scores that did not differ significantly over time. Only on the scores of the MHI was there any evidence of people in the *worse group* reporting scores indicative of a decline in mental health status.

Selected Service Provider Measures

There were no main effects for time of administration on the scores of the LSP. There were however significant *time of administration x change group* interactions on the *responsibility* sub-scale and the *LSP total score*. In each case the scores of the people in the *worse group* significantly 'declined'

from Time 1 and Time 2, whereas differences for the other groups were not statistically significant (see Figure 23).

Significant differences were observed between the Time 1 and 2 scores for *independent living skills*, *immediate* and *extended social network relationships*, and the *global role functioning index* of the RFS. In each case scores improved from over time. There was also a significant *time of administration x change group* interaction on the *working/ productivity* scores of the RFS, $F(2,58)=3.7$, $p<.05$. Post-hoc analyses indicated a significant difference between the means of the *better* groups and the *worse* and *no change* groups at Time 2. Differences between the groups at Time 1 were not statistically significant, and the effects at Time 2 are due to decreases in the scores for the *no change* and *worse* groups.

Significant 'gains' in scores were observed for the *impairment*, *social skills* and *symptoms* sub-scale and *total* scores for the HoNOS representing a general improvement on these dimensions for the whole sample. As well, significant *time of administration x change groups* interactions were observed on the *symptoms* sub-scale and *total* scores. In each case ratings for the *better* group indicated a significant gain. Post-hoc analyses indicated that the means of the *better* group were significantly lower at Time 2 than at Time 1 (indicative of an improvement on these dimensions).

There were also significant differences between Time 1 and Time 2 ratings for the whole sample on the *cognitive problems*, *physical illness/disability problems*, *problems with depressed mood*, *other mental/behavioural problems*, and *problems with relationships* scales. Further analyses across each *change group* suggested that these 'gains' from Time 1 to Time 2 were isolated to the *better* group for which there were also significant gains for the *non-accidental self-injury* and *problems with activities of daily living* scores. People in the *no change* group were rated as improving on the *problems with relationships* scores and declining on the *problems with occupation and activities* scores. People in the *worse* group were rated as declining on the *overactive, aggressive, disruptive or agitated behavior* scale.

As with the data reported in relation to the selected consumer measures the data from each of the service provider measures shows at least some degree of convergence with the 'global impressions' ratings provided by service providers. Both sets of data (the selected measures and the *global change ratings*) indicated an increase in mental well-being across the sample.

Interpretation of the results is difficult given the relatively small sample size at Time 2; the large discrepancies in sample sizes between change groups; and the lack of further information concerning the *clinical*, as opposed to the *statistical*, significance of observed differences between the scores at the two points in time. As well, direct comparisons of the measures are not possible as the scores for each measure are derived from different samples of the whole group: all participants completed only *two* of the *three* possible measures. Differences observed in the manner in which the measures describe the sample in terms of mental health status, and changes therein, could be attributed to this potentially confounding factor.

4.4.3 Change in Scores across Diagnostic Groups and Practice Settings

Following on from the results reported in section 4.3, analyses were undertaken to determine whether there were differences across *diagnostic groups* or *practice settings* in scores on the selected measures collected at Time 1 and Time 2, and the *global change ratings*.

4.4.3.1 Differences in Global Change Ratings across Diagnostic Groups and Practice Settings

Univariate ANOVAs indicated that differences in Global Change Ratings (consumer and service provider, respectively) across *diagnostic groups* and *practice settings* were not statistically significant.

4.4.3.2 Differences in Scores for the Selected Measures across Diagnostic Groups and Practice Settings

Mixed design repeated measures ANOVAs, with *diagnostic groups* or *practice settings* as the between subjects independent variable and *time of administration* as the within-subjects independent variable, were conducted to examine the statistical significance of observed differences between Time 1 and Time 2 scores of the selected measures. Wilcoxon matched-pairs signed rank tests were used to analyse change on the 12 HoNOS item scales.

Consumer Measures

For the BASIS-32 scores there were no significant main effects observed for *diagnostic groups* or *practice settings*, nor were there any significant interactions between these factors and the *time of administration* variable. Inspection of the means also indicated that the differences between *diagnostic groups* and *practice settings* observed for the original sample at Time 1 (i.e., all of the people who participated initially) and reported in section 4.3, were not evident for the sub-sample of people who continued with the study. This is a general finding that is evident across the measures as will be shown below and demonstrates the effect that attrition of participants can have when using such data to compare groups of people over time.

In the case of the MHI and SF-36 significant interactions between the factors of *time of administration* and *diagnostic groups* were observed for the scores for *depression, general positive affect, loss of behavioural/emotional control, psychological well-being, psychological distress, and mental health index* (MHI) (see Figure 24); and *social functioning, mental health, and mental component summary scores* (SF-36) (see Figure 25). Post-hoc analyses indicated significant improvements on these variables for the people in the *affective disorders group*. Significant differences between the means of the *affective disorders* and *schizophrenia groups* at Time 1 were observed on the scores for *depression, general positive affect, psychological well-being, and the MHI Index* of the MHI, as well as the *social functioning scores* of the SF-36. Again the results for differences at Time 1 reported here for the sub-sample of people who continued with the study, differ to those reported for the original sample of participants.

Significant differences between *practice settings* were observed on most scores for the MHI. In each case there were significant differences in the grand means for the *private psychiatry group* and those of the *public psychiatry* and (except for *psychological distress*) *general practice settings*. Post-hoc analyses suggested that these differences were significant at Time 1 only, consistent with earlier results.

For the SF-36 the means of the *public psychiatry group* were significantly higher than those of the *private psychiatry* and *general practice groups* for the *role-physical* and *general health* scores; the mean scores of the *private psychiatry group* were significantly lower than those of the *public psychiatry* and *general practice groups* for the *role-emotional, mental health, and mental component summary scores*.

There was also a significant *time of administration x practice setting* interaction for the *social functioning* scores of the SF-36, $F(2,74)=3.3, p<.05$. In this case a statistically significant improvement was observed between the Time 1 and Time 2 scores in the *general practice setting*, but not in the other *settings*.

Service Provider Measures

Significant effects for *diagnostic groups* were observed on the *communication* and *self-care* scores of the LSP: in each case the grand means of the *schizophrenia group* were lower than those for the *affective disorders groups*. As well a significant *practice setting* effect was observed for the *self-care* score: the grand means of the *public psychiatry group* were significantly different to those of the *private psychiatry group* (only at Time 1). Once again these results are in contrast to those reported in Section 4.3 in relation to group differences observed for the original sample.

Significant *time x diagnostic groups* interactions were also observed for the *immediate social network*

relationships scores of the RFS, $F(2,63)=4.7$, $p<.05$. This effect was due to a non-significant increase in scores for people in the anxiety group.

A significant *time x practice setting* interaction was observed for the *GRFI* scores, $F(2,64)=4.2$, $p<.05$, of the RFS. This effect was due to an improvement in scores for people in the *private setting*.

Significant *time x diagnostic groups* interaction effects were also observed for the *symptoms* sub-scale scores, $F(2,67)=4.5$, $p<.05$, of the HoNOS. In this case post-hoc analyses indicated significant improvements on these variables for the *affective disorders* group.

For the 12 item scores of the HoNOS, analyses across *diagnostic groups* and *practice settings* indicated that people with *affective disorders* received improved ratings on the dimensions of *cognitive problems*, *physical illness/disability problems*, *problems with depressed mood*, *other mental/behavioural problems*, and *problems with relationships* (see Figure 26); that people in the *schizophrenia* and *public psychiatry* groups experienced a gain on the *physical health* dimension; and that people in the *general practice* group were rated as improving on the *other mental/behavioural problems* dimension.

As with the data from the consumer measures, many of the differences observed between the *diagnostic groups* and *practice settings* at Time 1 for the original sample were not evident in the data for the group of people who continued through the study. This raises a number of practical issues in relation to attrition rates and the comparability of group data over time; a point that will be discussed below.

4.4.4 Correlational Analyses

An alternative method of examining change on these measures is to investigate the correlations between change scores (i.e., the simple difference between Time 2 and Time 1 scores) both across measures and within the same set of scale items. As well, the relationships between change scores and the *Global Change Ratings* were investigated.

4.4.4.1 Association between Change Scores for the Selected Measures

There were significant correlations across change scores for the selected consumer measures, except on the *physical component summary* scores of the SF-36 (see Table B4.7). There were also significant correlations between total scale change scores for the HoNOS and LSP, and LSP and RFS, but not for the HoNOS and RFS (see Table B4.8).

There were weak correlations between change scores for the consumer measures and those for service provider measures, consistent with the results reported above in relation to the measures as measures of health status. The strongest association was between the LSP *total* scores and the *mental component summary* scores for the SF-36 ($r=0.44$ $p<.01$) (see Table 4.9).

4.4.4.2 Association between Change Scores on the Selected Measures and Global Change Ratings

Table B4.10 displays the Pearson's correlation coefficients, r , and r^2 values (a measure of the proportion of variance in the dependent variable scores accounted for by the independent variables) to describe the association between the *Global Change Ratings* and change scores for the composite scores of the selected measures. There were significant correlations between *Global Change Ratings* and change scores for all measures, except the RFS. The change scores for the MHI *Index* and *mental component summary* score (SF-36) had stronger associations with *Global Change Ratings* (*consumers*) and accounted for a greater percentage of variance than the BASIS-32 *Average Scores*. Both the LSP and HoNOS *total* scores demonstrated significant associations with the *Global Change Ratings* (*service providers*).

Standard multiple regressions were performed with *Global Change Ratings* as the dependent variables, and the sub-scale change scores of the consumer and service provider measures as predictor

variables. Separate analyses were undertaken for each set of change scores. For the analyses involving the MHI and SF-36, two multivariate models were tested (see Table B4.11). In the case of the MHI the first model tested included the five sub-scale scores; the second model included the scores for *psychological well-being* and *distress*. In the case of the SF-36 the first model tested included the eight sub-scale scores; the second model included the scale scores for the *mental health*, *social functioning*, and *role-emotional* scales, which are the scales most related to mental health. The analyses are presented as a further means of measuring the degree of association between the change scores of the selected measures and global ratings of change.

The results indicated significant multivariate relationships for the MHI (two- and five-factor model), SF-36 (three-factor model), HoNOS and LSP Sub-Scale Change Scores, respectively, on the *Global Change Ratings*.

4.4.4.3 Internal Consistency of Change Scores

The alpha coefficients computed on the change scores of the consumer measures indicated moderate to high levels of internal consistency (see Table B4.12). Coefficients for the service provider measures, especially the HoNOS and RFS, were not as strong as those for the consumer measures (see Table B4.13). The results generally support the hypothesis that change in scores was consistent across items that were supposed to measure the same construct.

4.5 Effect of Computer Administration

Wherever feasible, data was collected using computerised versions of the selected measures. Eighty-six consumer (47%) and 26 service provider (14.9%) assessments were conducted using the computerised questionnaires at Time 1. The remaining assessments were completed with pen-and-paper versions of the measures. Most of the computer assessments were conducted in the *public* ($N=54$) and *general practice* ($N=50$) settings.

Univariate and multivariate ANOVA were undertaken to investigate the influence of the two methods of administration on the composite scores of the selected measures. Results indicated that there were no effects for *method of administration* for the scores of the selected measures of outcome.

Following administration of the computerised questionnaires consumers and service providers were asked, *How do you feel about using a computer to answer these questions?* This was followed by two specific questions:

- Is it a good idea to use a computer to get information about the way people are feeling?
- Did it feel comfortable to use a computer to answer these questions?

To both questions consumers simply responded *yes* or *no*. Results in Table B5.1 indicate that consumers and service providers did not feel uncomfortable in using the computers. Participants also indicated a high level of acceptance for using computers to collect health status information.

The use of computers for outcomes assessment was an issue that arose within the focus groups as well as in the normal course of data collection (including the submissions received from the service provider organizations). Feedback from participants highlighted the questions of the user-friendliness of computer software, as well as the availability of such equipment for consumers as well as service providers. Although computers were considered a useful resource the costs involved were considered to be a major impediment to their introduction.

4.6 Discussion of Results of Analyses of the Selected Measures as Measures of Mental Health Status and Change in Mental Health Status

In order for a measure of consumer outcomes to be useful in decision-making processes the information derived from the measure must be valid and reliable. The results obtained in this study support the hypotheses, stated in Section 2.0, with regard to the selected measures as measures of mental health status and of change in mental health status.

The results of correlational analyses were generally consistent with predictions. The validity data indicated strong associations between the scores of the selected consumer measures (BASIS-32, MHI and SF-36) in the expected directions. While the associations between the LSP and RFS scores were strong there was greater variability in correlations between the HoNOS scores and those of the other selected service provider measures. This reflects the more diverse range of dimensions of mental health covered by the HoNOS. The test-retest reliability and internal consistency data for each measure are similar to data reported elsewhere.

There was a low convergence between consumers' and service providers' assessments. This is not unexpected given the different contents of the two sets of measures. It does, however, reinforce the point that the assessment of mental health status is not straightforward. Any assessment process requires the contribution of all of the relevant stakeholders, wherever possible. No doubt these results also reflect the heterogeneity of the group; the consequences of sampling from a range of sites, *diagnostic groups* and *practice settings*.

The lack of strong convergence between consumers' and service providers' assessments was also reflected in the data obtained for the three *diagnostic groups* and *practice settings*. The data provided by the selected consumer measures suggested that the people with *affective disorders* and receiving services in the *private psychiatry setting* had greater levels of illness in comparison to the *schizophrenia* and *public psychiatry groups*. The service provider data provided a different picture of the sample: people with *schizophrenia* tended to obtain scores indicative of greater levels of illness and disability.

The results were consistent with expectations given the distribution of *diagnostic groups* across *practice settings*. A large proportion of people in the *affective disorders group* were receiving services in a private acute inpatient setting. On the other hand most people with *schizophrenia* were receiving services in a public outpatient service.

Differences between consumers' and service providers' assessments (as well as differences across all of the measures) reflect the different contents and orientations of the measures (the consumer measures generally assessed how the consumer was feeling, whereas the service provider measures generally assessed functional ability). They may also reflect treatment approaches to alternative diagnostic groups. There is also the possibility that consumers and service providers undertake different processes in completing such measures.

Taking into account the information reported in Section 1 as well as the current data, we would suggest that the measures are valid and reliable as measures of mental health status (given the limits of the study; see below). However, consideration needs to be given to the context in which they are to be applied. As well, attention needs to be given to the validity and reliability of different aspects of the information derived from these measures. In particular we would caution the application of the sub-scale scores for these measures without further work being undertaken to confirm their construct validity in relation to mental health populations.

A simple examination of *sensitivity to change* (i.e., the validity and reliability of these measures as measures of change in mental health) was undertaken in this study. It may be argued that we have examined *perceived* as opposed to *actual* change, and essentially compared two methods of doing so, as no attempts were made to examine specific treatment effects. Change was also examined analysing the consistency between change scores for each measure.

Almost 81% of consumers indicated some change in their mental health according to data collected using a general global impressions scale (60% *better*, 21% *worse*). Approximately 84% of service providers indicated that their consumers had changed (67% *better*, 17% *worse*). Ratings did not

differ significantly across *diagnostic groups* or *practice settings*. On the basis of this data three *change groups* were identified and the associations between these three groups and change recorded on the selected measure scores, were investigated.

There was evidence of convergence between the scores of each of the selected measures and the three *change groups*, to varying degrees. The strongest results were obtained for the MHI and SF-36 data. The HoNOS, RFS and LSP also demonstrated useful results. On each of these scales we observed significant interaction effects between *time of administration* and *change groups* which appeared to indicate that people who stated that they had experienced, or observed, a global improvement in their own, or their consumers', mental health over the three month period, recorded scores on the selected measures which were consistent with this. On the other hand, people who indicated no global change recorded similar scores at Time 1 as at Time 2. In the case of the data collected for the MHI, people who had reported a global deterioration in their mental health recorded significant 'declines' on a number of scores for this measure (see also the data for the LSP).

Changes in scores were also observed across *diagnostic groups* and *practice settings* from Time 1 to Time 2 on some of the scores of the MHI, SF-36, HoNOS, and RFS, with the trend suggesting significant gains in scores for people in the *affective disorders* group.

Most differences in scores, both across *time*, *change groups*, *diagnostic groups*, and *practice settings*, were small, and it is uncertain how *clinically* significant these changes are without recourse to other information sources. As well, the study suffered from a large attrition rate complicating the interpretation of the findings.

Using the simple difference scores, significant associations were observed for the *Global Change Ratings* (consumer and service provider, respectively) and the composite scores of the selected measures, except the RFS. Change for the better was associated with 'positive' changes on the MHI Index, MCS (SF-36), HoNOS and LSP Total scores; while changes for the worse were associated with 'negative' changes on those scores. Significant multivariate effects were observed for the sub-scale change scores of the MHI (*psychological well-being* and *distress*; and the five sub-scale scores), the three mental health scores of the SF-36 (*social functioning*, *mental health* and *role-emotional* scores), as well as the HoNOS and LSP sub-scale scores, respectively. These results are consistent with the data reported in relation to differences in selected measure scores across *change groups*.

Correlational analyses indicated reasonable amounts of internal consistency for the change scores (except the HoNOS); significant associations among change scores for consumer and service provider measures, respectively; and relatively low convergence between consumer and service provider change scores.

Interpretation of these results is not straightforward. For instance, as explained above, the comparability of the results across measures is compromised by the fact that the data for each measure represents a different sub-sample of the total group of participants. As well, the large attrition rate complicates the findings.

Moreover, it could be argued that the data provides some support for the use of simple global measures, such as visual analogue scales and clinical global impression scales, as measures of outcomes. After all a close relationship has been observed between the mental health status scores on the consumer measures and the scores of the visual analogue scales. As well, changes on a number of measure scores over time were consistent with simple ratings of global change. Pertinent to this issue were comments raised by both consumers and service providers within the focus groups suggesting that simple measures may be just as effective in measuring outcomes. As one service provider stated, "...I think that the other question that you could ask people is 'are you better now?', and frankly we start treatment because they come in and they say they're not well, but we never ask them at the end of it are you better, or do you think that you're better yet..." (Group 3)

We are not suggesting that measures such the six selected questionnaires be abandoned in favour of visual analogue scales or clinical global impressions scales. But the measures chosen for the purpose of routine outcomes assessment should be the simplest and most straightforward to meet

the decision-making requirements of service providers. The degree of specificity required in a measure will be determined by the decision-making needs of the service provider and consumer.

While not specifically a trial of the usefulness or feasibility of using computer technology to collect outcomes information the study presented a good opportunity to investigate the suitability of using computers, and the results indicated that the introduction of this technology into the data collection process did not unduly influence the quality of the information collected. Issues that did arise were in relation to the ability of service providers and consumers to access and to use such equipment and software in an appropriate and efficient manner.

This study did not set out to provide a definitive examination of the validity and reliability of the six selected measures. Yet the data does provide useful insights into the measures, and, perhaps more importantly, it raises a number of fruitful questions about the actual process of implementing routine outcomes assessment.

4.7 General Discussion

Does the information gathered here assist one to choose amongst the selected measures for the routine assessment of consumer outcomes? To answer this question we must consider all of the data collected (*qualitative* as well as *quantitative*), as well as taking into account current trends in the field. We must also be aware of the potential limitations of this study and factors that could restrict the generalizability of the results obtained (discussed below).

With respect to the specific measures it is worthwhile to review the relevant findings.

Behaviour and Symptom Identification Scale: As we have seen the BASIS-32 was designed as a measure of outcomes for use in inpatient settings. The original data reported on it was collected with a sample of people who were predominantly suffering from affective disorders. The data reported here indicates that the scores of the BASIS-32 correlate well with those of the MHI and SF-36 in the predicted directions; is consistent with findings from alternative measures; and showed some indication of convergence with change ratings. While representatives of three service provider organisations considered this a useful measure, it was not the first choice of consumers in this study. A number of adverse comments were received in relation to the language and content of the measure (which were also raised in relation to the other measures).

Issues that require further consideration include the language of the measure, the ability of people who are very ill to complete this form, the construct validity of the sub-scales, and the utility of the measure for assessing outcomes for specific treatments.

Of the three consumer measures, the BASIS-32 did not appear to converge with the alternative ratings of change to the extent that the others did. However, this could be due to a range of factors addressed below.

Medical Outcomes Study Short Form (SF-36): An issue that continually arose in administering this measure was the definition of 'health'. Consumers at times interpreted the phrase 'health' or 'general health' as relating to their *physical* and not mental health. Many consumers, therefore, considered this measure as irrelevant to measuring mental health. The language and content of items, range of available responses, and rating period were all raised as issues for further consideration in light of the needs of people with mental health problems.

The statistical data suggested that the measure was strongly associated with the other consumer measures (in so far as it measured mental health) and there was evidence of convergence with alternative measures of mental health status and change in mental health.

Life Skills Profile: Service providers in the public psychiatry settings considered The Life Skills Profile a very useful and suitable measure. This is consistent with the original purpose of the measure. The general applicability of this measure to other contexts such as the assessment of outcomes for

people with anxiety or affective disorders, or in the general practice setting, requires further investigation.

The statistical data were consistent with expectations and there was some indication that the scores obtained from the measure converged with global ratings of change in mental health.

Role Functioning Scale: Of the service provider measure, the RFS is the less comprehensive in the range of dimensions that it assesses. Again the original purpose of this measure was for evaluating outcomes in outpatient, community-based services. There was evidence of convergent validity and internal consistency, as well as convergence with alternative measures of change. This is a measure that would benefit from further investigation.

Of the six selected measures the two which appear to offer potential for widespread use are the *Mental Health Inventory* (MHI) and the *Health of the Nation Outcome Scales* (HoNOS). The MHI was well received by consumers and is specifically concerned with *mental* as opposed to *physical* health. Its relatively strong performance in predicting global change ratings also justifies this measure. Moreover the MHI has recently been examined in relation to, and recommended as a suitable measure of outcomes for, adolescent mental health services (Ostroff et al., 1996). And it has also been shown to be a valid measure in non-Anglo-Celtic cultures (Liang et al., 1992).

Difficulties with this measure included the 1-month rating period, as well as the language and response format of the questions. Further work is needed to examine the construct validity of the sub-scale scores for an Australian mental health population, as the measure was not originally intended as an outcomes measure for mental health services, and most of the normative data have been collected from a US sample.

The HoNOS is potentially a very useful measure of outcomes given the data reported above; the large amount of research and work that is currently being undertaken on it; and, the growth in its acceptability amongst service providers both here in Australia as well as the U.K. The HoNOS also appeared to provide information consistent with global change ratings; information that was covered by the lengthier LSP; as well as addressing a wider range of mental health dimensions.

Also, the fact that the HoNOS comes with a specific training program, and is amenable to computer technology, makes this an attractive option as a general outcomes measure, in comparison to the LSP and RFS.

Difficulties with this measure that need to be addressed relate to the possibility of floor effects in the data; the restriction in range of responses; and, the utility of the four sub-scale scores.

In suggesting that these two measures display the best potential as general measures of consumer outcomes, we are recognising the principle that any assessment process benefits from the perspectives of both the consumer and the service provider. Wherever possible, outcomes information should be collected from both consumers and service providers.

The preceding statements should not be interpreted as being critical of the other selected measures. The BASIS-32, SF-36, LSP and RFS have their place when used in the appropriate situations. For instance, the SF-36 is potentially very useful in those situations in which physical health issues are relevant, such as in general medical practice. Moreover it is currently being used as an outcomes measure in a number of different assessment settings (e.g., Smith et al., 1996a & 1996b). The LSP on the other hand has proven utility in community based services where concerns of disability are of high importance (Hambridge & Rosen, 1994). The BASIS-32, too, has been widely used in the United States and offers great potential as a measure of outcomes (Eisen & Dickey, 1996).

Clearly there is a need for services to consider the exact nature of their decision-making needs, and consequently, which measures of outcomes may be most appropriate. Furthermore, the recommendations for using any of these measures is subject to a number of caveats discussed below.

Perhaps the question of which of the six selected measures is most suitable for implementation as a national measure of outcomes misses the point. The challenge lies in encouraging policy makers,

service managers and providers and consumers, to think seriously about the value of implementing systems of routine outcomes assessment, and to adopt the most appropriate methods and processes for doing so. At this stage, it is doubtful whether choosing any single measure or group of measures would materially advance this process. Indeed, we simply do not know enough about the utility of these measures, nor enough about the proposed uses of a 'national' measure, to conclusively recommend any of the six selected measures for such a purpose. More work needs to be undertaken to investigate the utility of these measures, as well as alternative measures and systems for outcomes assessment. Of primary importance is the need to consider the *process* of outcomes assessment. Essentially this relates to the *why, what* and *how* of outcomes assessment. Perhaps the high attrition rate observed in this study indirectly reinforces this fact: you can't even pay people to do these questionnaires if they cannot see any worth in doing so.

Central to all of this is the question of the purpose of outcomes assessment. One of the principal concerns of consumers and service providers alike in relation to outcomes assessment is what will be done with the data. Consumers and service providers generally had little difficulty suggesting uses for information collected from routine outcomes assessment for clinical and service purposes. There was support for the utilisation of such data to inform broad planning requirements. However, there was a reluctance to relinquish control of the use of this data to others, and, in some quarters, suspicion as to how the data would be used.

4.8 Some Practical Issues

This study encountered a number of challenges in its implementation and in the analyses of the data obtained. It is worthwhile to examine these closely in the context of their potential implications for the routine assessment of consumer outcomes.

Sample Size and Attrition: Amongst the difficulties experienced in this study were the small sample size, and the representativeness of the sample. Given the small size of the sample the design was limited in the range of analytical techniques that could be applied to the data. For instance, factor analyses that have been conducted on the various measures have used US or UK sample data (e.g., SF-36, BASIS-32, MHI, HoNOS). Further investigation of the factor structures of the measures and the construct validity of the sub-scale scores with data obtained from a large sample of people with mental illness in an Australian context needs to be undertaken.

In relation to the representativeness of the sample it may be argued that the people who participated in the study represent specific subsets of the population of people with mental health problems and that other groups were not included (see also the report by the Mr Tony Hennessey, Consumer Consultant, in relation to the constitution of the focus groups, Appendix 2). For instance, there was a selection bias in the fact that consumers were only approached to participate if it was felt, by their service provider, that they could participate. If people are unwell, especially those who have been recently admitted to hospital or have just commenced treatment, they may either be unable or unwilling to complete a questionnaire even if they are assisted to do so. This raises questions of how to measure outcomes in such situations, and of when it is appropriate to use self-report measures for outcomes assessment.

Our experience suggests that consumers, although generally very co-operative in the initial stages, need to be able to identify some relevance and benefits to completing such measures if they are to continue being co-operative.

This has ramifications for how such data can be used. For instance, if one was to attempt to evaluate the *treatment outcomes* of a service, low participation and high attrition rates would place limits on the interpretability of the data. The conclusions of this study in relation to the sensitivity to change properties of the measures were impeded by the relatively high attrition rate. It has been suggested that an 80 to 85 percent follow-up rate is the minimally acceptable level in practice (Eisen et al., 1991). Obviously the 60 percent follow-up rate in the present study would, in the real world, be unacceptably low. The main reason for consumers discontinuing with the study appeared to be simply a lack of interest. After all the situation they were presented with was odd and seemingly unconnected with the service that they were receiving. Another likely factor for some consumers

was disapproval with being used in a process which was apparently beyond their control and which seemed to suggest an even greater intrusion into their lives.

Perhaps a bigger hurdle to introducing outcomes assessment will be retaining the co-operation of service providers. For instance, those who were involved in this study all had considerable constraints on their time. As well, they were concerned about the implications of this process for their consumers and did not wish to involve their consumers in an exercise that did not have clear and demonstrable benefits.

Standardised implementation and training: Of all the measures, this study has probably been most unkind to the HoNOS, and to those people who used it. The authors of the HoNOS expressly recommend training for this instrument, which we were unable to provide in this study. However, the issue of training is pertinent to all of the measures. The validity and reliability of data obtained from these measures will be greatly influenced by the level of standardization in the administration, scoring, analyses and reporting of these measures. Moreover, training is required to address and deal with the range of peripheral issues concerning the implementation of any system of assessment: dealing with confidentiality, dealing with consumers needs, introducing the process, etc.

Attribution of change: A concern for consumers and service providers was the extrapolation of data collected on these measures to judgments concerning the effectiveness of service provision. That is, attributing change in a person's mental health condition to the efficacy of treatment. While this may be a legitimate purpose of collecting outcomes data, we would be concerned if any of the selected measures were to be adopted for this purpose without further, more detailed, research being undertaken.

The first concern raised was that to attribute change to treatment effects would require valid and reliable measures, as well as a focus on the actual treatment interventions. A number of service providers expressed distrust in the use of the selected measures for such a purpose. A further objection was that if attributing change was one of the purposes for implementing such measures, there may be an incentive for consumers or service providers to unconsciously or consciously bias the results to suit their purposes.

Attribution of change is an integral part of the definition of outcome contained in the AHMAC 'Sunshine Statement'. In light of our experience in this study and those issues discussed in Section 1, we are of the opinion that re-examining and altering the definition of 'outcomes' to reduce the emphasis on attribution may assist the promotion of the use of routine consumer outcomes measurement.

We simply do not know enough about the measures to determine their usefulness in facilitating decisions concerning the attribution of change to specific factors, especially given the potential impact that such information could have. Certainly, the statistical data recorded here does not allow us to make any conclusions as to the suitability of these measures for *attributing* change to any specific factor. Also, such issues carry with them political as well as clinical implications. In order to make appropriate judgments concerning clinically meaningful change for an individual, and to attribute such change to the influence of various factors such as treatment efficacy, requires more information than the six selected measures appear able to provide.

The appropriateness of assessment: The timing and frequency of assessments, as well as the necessity of using questionnaires for outcomes assessments, were raised as important issues. A number of service providers and consumers who participated in the focus groups could see no benefit in completing any of the selected questionnaires, although they may have been supportive of the concept of routine outcomes assessment. While the general response to the idea of routine assessment of outcomes was positive, in most cases this was conditional on the provision that the measures be shown to provide meaningful and useful information. Moreover the determination of the frequency of administration etc., clearly needs to be addressed. The obvious strategy would be to allow service providers and consumers to determine the frequency of administration of the measures. Such a strategy, however, does raise difficulties with the aggregation and analyses of such data from individuals over different sites and time periods between ratings.

Convergence of service providers' and consumers' perspectives on mental health assessments: As the results of this study indicate, there was low correspondence between the assessments of mental

health made by consumers and service providers. Further research is necessary to investigate this lack of convergence. Furthermore, different perspectives on what constitutes an outcome in mental health deserve more detailed inquiry. For instance, many consumers involved in the focus group discussions articulated a link between their outcomes, their satisfaction with service provision, and their attitudes to service effectiveness (*“What I want from a service provider versus what I didn’t get.”*).

Special Groups of Consumers: The sample of people involved in this study clearly represented specific subsets of the total population of people with mental illness. This may place limitations on the generalizability of the findings, especially the statistical results. In addition to diagnostic groups, and differences in characteristics across practice settings, there are other groups of consumers whose special needs present challenges to outcomes assessment. These groups include children and adolescents, Aboriginal and Torres Strait Islanders, consumers from non-English speaking backgrounds, people in rural or remote areas, the elderly and the incarcerated. And the circumstances of each group raise different questions in relation to the suitability of specific measures of mental health outcomes.

Carers: It is clear that carers are an important and reliable source of information regarding the mental health outcomes of the individual consumers with whom they are associated (Massey and Wu, 1994). Apart from the LSP, none of the selected measures are specifically designed to facilitate the carer’s contribution to the outcomes assessment process. The involvement of carers in routine outcomes assessment obviously deserves further consideration.

Technological advances to assist in the collection, analyses, interpretation and reporting of outcomes data: Computers offer a number of potential benefits in facilitating the collection and analyses of outcomes data. Software is available for the standardized analyses and reporting of statistical findings. Such processes have the potential to minimize bias on the part of any one party by providing an objective assessment. However, difficulties may arise in relation to the costs of such systems, ensuring the confidentiality of information, etc. (Butcher, 1994).

The cost of routine outcomes assessment: A significant concern for service providers and consumers, is the cost of implementing routine outcomes assessment both in terms of time and money. Obviously, the real cost will depend on the needs of the service. Further research is needed with respect to the costing of routine outcomes assessment.

Bias and Artifact: As was discussed above, a range of factors can unduly influence scores recorded on a measure (see Table A4). A better understanding is required of those social psychological factors which may be of particular relevance to the outcomes assessment process, as well as of the strategies that may be employed to deal with them. For instance, in the focus group discussions a number of consumers suggested that they may feel reluctant to indicate any form of dissatisfaction with a service that they depend upon. Furthermore, service providers also indicated that at times they may feel compelled to answer in a particular manner depending on the circumstances.

The design of this study did not adequately address the possibility of bias in the assessment process. This is partly due to the dynamics of the research situation as opposed to assessment in the real world. For instance, consumers and service providers were told that their responses to the questions would not be shown to the other party. However, in those circumstances where information provided by consumers could be accessed by service providers it is likely that many consumers may feel reluctant to respond in any manner that might negatively influence their relationship with their service provider.

A further potential bias was the unfamiliarity of the process on the part of consumers and service providers many of whom had not heard of outcomes assessment, and were concerned about its potential consequences. That the participants did not necessarily view this as something that they had a stake in, has implications for how the results can be generalised to practice situations.

The results of the study may have been particularly influenced by a selection bias, viz., the ‘volunteer factor’. Potential participants may not have been approached by their service providers, or chosen to participate, if it was felt that they were not well enough to complete the questionnaires.

Other biases that the present design did not account for were regression to the mean, spontaneous remission, and inert placebo. This is partly due to the inclusion of only two waves of data collection. Including more than two collection points will go some way to controlling for such factors. The issue of the systematic effects of bias in the assessment of outcomes in the real world is one that deserves closer research attention.

Measurement of Change: It should be remembered that the study did not specifically set out to measure 'real' change. Basically, the analyses compared two methods for rating change, and the level of agreement between those methods. Moreover, the relatively short time period for assessment of consumers posed an impediment to measuring change. For instance, there was no real expectation that a number of consumers recruited in the public psychiatry setting who had long-term chronic disorders, were going to change on the dimensions assessed by the measures in such a short period of time. Future studies need to examine the progress of consumers over a much longer time frame.

An aspect of all of the measures requiring further consideration included the format for responding to, and scoring, the items. The measures could be seen to suffer from restriction in the range of available responses. This has the potential of limiting the statistical techniques available for analyzing the data (e.g., in the case of the SF-36 Stevenson, 1995, points out that the categorical nature of the data of certain sub-scales imposes restrictions upon the use of linear statistical techniques), and of decreasing the chances of observing meaningful change over time on the dimensions of interest. This point was also raised by consumers in the focus groups. Many suggested that the measures placed limitations on the manner in which they could respond, and, consequently, impacted upon the quality of the information that could be derived from the measures. Alternative formats for responding to such questions, such as sliding scales, visual analogue scales etc., were considered helpful.

In a way the assessment of change in this study was made even more artificial in that consumers and service providers did not have access to their Time 1 ratings when they came to completing the assessment at the three month follow-up. This would be unlikely in real life practice. Instead of completing the forms without reference to earlier ratings, in practice service providers may compare their impressions of a consumer with what they have recorded in the past. This may influence completion of the forms with the intention of showing either change for the better, worse or not at all. The process then becomes different from that which one would wish to present.

A more effective method for assessing the feasibility and utility of these measures, and for addressing issues of change would be to study the implementation of a system of outcomes assessment over a long period of time (two to five years), and over a large sample of people. Also adopting a strategy of examining specific sets of conditions (e.g., depression, anxiety and schizophrenia) and designing separate outcomes approaches that are meaningful in light of our knowledge of these conditions and which meet the needs of consumers (e.g., see Smith et al., 1997; Smith et al., 1996a & 1996b), may be a more productive process.

Finally, one might question the ability of these measures to assess change on the basis that the questions on all measures require a person to provide an assessment of their mental health status, or that of their consumers', at a specific point in time, or as an average for a specified period. They do not, in general (c.f., question 2 of the SF-36), directly ask for assessments of *change*. Instead *change* is to be extrapolated from consecutive ratings of mental health status. Perhaps it would be useful for more direct measures of change to be trialed as alternative approaches for outcomes assessments.

The limitations experienced by this research project are instructive for the assessment of outcomes in the real world. They exemplify some of the hurdles that services are likely to face when attempting to implement measures of outcomes. The description of these issues is not meant to be exhaustive, but it is illustrative of the complex factors that surround the assessment of consumer outcomes and which must be addressed should any system of measurement be successfully implemented. Inspection of the results of the focus groups quickly indicates that there are other aspects to consumer outcomes that require closer examination. Clearly there would be great benefit in studying these issues in detail.

5.0 Conclusions and Recommendations

In opening her presentation to the 1996 AHOCH Conference on Mental Health Outcomes, Professor Beverley Raphael argued that 'outcomes assessment' has been promoted as a panacea by which the worth of mental health services may be measured. Yet, in many instances discussions of outcomes assessment have revolved around a simplistic framework of 'measures'. Raphael then went on to explain that the practicalities of outcomes assessment are more complicated than the mechanical completion and recording of numbers on a form. In summing up Raphael states: "There are no simplistic solutions. Key requirements must, however, include a conceptual framework recognising the theoretical and real issues affecting those with mental health problems and disorders..." (Raphael et al., 1997). The data collected in this study provide further corroboration of these views.

A great deal of effort is currently being expended to improve the quality of mental health services in Australia. Specifically this entails assisting the consumers of such services to achieve better health outcomes. An important part of this process is the improvement of information systems by which consumer outcomes can be evaluated. This study has investigated a selection of six measures to determine which may be most suitable to implement on a national basis as a method for obtaining routine data about consumer outcomes.

The aims of the Consumer Outcomes Project were quite diverse. Consequently, the findings of this project address a broad range of issues relating to the psychometric properties of the six selected measures, as well as to more general issues relating to the processes of outcomes measurement and the measurement of change.

A number of issues have been explored from a variety of perspectives. The data reported here will inform aspects of Phase 2 of Stage 2 of the AHMAC National Mental Health Working Group's research program on consumer outcomes in mental health, as well as the long term processes of implementing outcomes assessment in clinical practice in Australia. Many of the issues go to the core of the outcomes assessment process and warrant closer examination.

A series of general recommendations concerning the routine assessment of consumer outcomes in mental health have been developed. In addition a number of recommendations with regard to the six selected measures have also been formulated. These recommendations are offered as suggestions for the work of Phase 2 and any other work in this regard. Essentially we are advising that the focus of future work move from *measures* to *methods* and address the practical issues relating to the implementation of systems for the routine assessment of consumer outcomes.

General Recommendations

- 1. The Commonwealth and State governments should aim to develop an environment that encourages and supports routine outcomes assessment in accordance with National Standards for Mental Health Services and guidelines for clinical practice. The choice of a national consumer outcomes measure should not be the primary goal at this stage.**

A measure is not a method (Andrews et al., 1994). There is more to consumer outcomes assessment than the regular administration of the HoNOS, LSP, MHI or any other questionnaire. Using standardised outcomes measures may provide stakeholders with a 'common language' to report on consumer outcomes (Johnson, 1995). However, they will be of limited utility if the methodologies for implementing and analysing the data are not considered in the same instance.

The implementation of a process for routine outcomes assessment needs to address the following questions:

- What are the information needs of service providers, consumers, carers, service managers, funding agencies? What are the outcomes of interest?
- How will the information be collected, analyzed, reported?
- Who will own the information?

- Will routine outcomes assessment add value to:
 - * Treatment planning and coordination of services for consumers and carers?
 - * The task of day-to-day service provision?
 - * Service management?
- How will such information assist in determining the contribution of service provision to consumer outcomes?

There is a need for closer consultation with consumers, service providers and carers in this matter. If a measure is meant to be used by service providers and consumers to inform treatment progress, then it is only reasonable to expect that consumers and service providers will have a great deal of input into the design and implementation of such a system. Similarly where the needs of service managers are the issue, they should be consulted in relation to the sort of outcomes data, and manner of presentation of such, that they require in order to facilitate decision-making processes.

The introduction of outcomes measurement will be enhanced by the involvement of the people who will be most directly impacted upon by its implementation. But in order to win the support of stakeholders we must first be able to show that the benefits of outcomes assessment outweigh the potential costs, and that the concerns of consumers and service providers in relation to such things as confidentiality, attribution of change etc., can be appropriately addressed and managed.

2. Measures of outcomes should be considered primarily as tools for improving the quality of services. Outcomes assessment is not the only measurement exercise that a modern service requires. Alternative uses of outcomes data such as resource decision-making will, however, require more specific and detailed examinations of the suitability of these measures and the data obtained from them.

This statement recognizes the separate yet related outcomes assessment activities described in Section 1: micro, meso and macro levels of assessment (Sutherland and Till, 1993). A danger lies in confounding these processes. For instance, blanket application of one or two outcomes measures that assist macro-level decision making may be viewed as the beginning and end of outcomes assessment, to the detriment of effective micro-level information gathering. Furthermore, there is the risk of broad dimensions applicable to macro-level assessment being applied to individuals. For instance, a general concept such as psychological well-being might well be a dimension of interest at a population or group level, but for a specific individual in a particular situation a higher score on a questionnaire may be meaningless without reference to treatment goals.

By linking outcomes assessment to the specific goals of treatment and of service provision, such problems may be avoided and greater accuracy achieved for assessments of change for individuals and groups of individuals.

The question of how useful information derived from these measures can be, for informing decisions concerning resource allocation, and specifically funding decisions, requires further consideration.

3. Outcomes assessment should be considered as a goal in itself and needs to be differentiated from other information gathering tasks and developments.

Much of the literature pertaining to outcomes assessment is either prefaced by, or set in the context of, discussions about cost management, managed care, diagnostically related groups (DRG's), casemix, etc. Many of these concepts are associated with particular ideologies concerning the management of mental health systems. However, outcomes assessment as a task and as a goal should be considered independently of such issues.

The justification for outcomes assessment can be simply stated in the following manner:

- Services that do not possess internal processes of evaluation will be impeded in their progress; and,
- Services with efficient outcomes assessment systems will promote professional development of service providers, and empower consumers and carers by providing them with a means of observing the progress of treatment and evaluating the effectiveness of the services they receive.

Moreover, effective and transparent systems of outcomes assessment will, over time:

- Benefit our understanding of mental health by providing large and valuable databases with which to investigate mental illness in the practice context, and to track those systems of intervention that are most effective in dealing with different types of mental illness; and
- Provide information to facilitate service planning with the equitable targeting of resources.

4. There is a need for more applied research into service effectiveness and the dimensions of consumer outcomes.

In order to understand and effectively utilise outcomes data there is a need to further investigate such factors as those described in section 4.8. Any future research should be conducted in the context of the practical implementation of systems of outcomes assessment over a reasonable length of time.

The focus for Phase 2 of the current research program should shift from 'measures' to 'methods' and concentrate on the issue of *how* routine outcomes assessment will actually work. This will entail addressing many of the issues raised through the focus groups, as well as:

- Consultation: consulting with service providers and consumers to determine their exact needs and how routine outcomes assessment can fit into existing service frameworks. Determining the appropriate measures for the situation. Dealing with issues of what outcomes to measure, ethical issues (e.g., confidentiality), the attribution of change, etc.
- Education: clearly defining the purpose of the process and providing a conceptual map of how outcomes data will be of direct benefit to all stakeholders (consumers, service providers, government, etc.) in the process.
- Training: providing service providers with training, clear guidelines and support for collecting outcomes data. As well as assisting consumers in any way they require.
- Analysis and Reporting of Outcomes Data: provide support in the analyses of the data obtained and the design of methods for reporting back to service providers and consumers the outcomes of the assessments.
- Ongoing Support, Feedback, and Review: assisting service providers and consumers over time; following-up and reviewing the progress of the system; testing the validity and reliability of the data obtained over time and in relation to other indices of outcomes.

Table A17 provides a cursory description of how a service might implement a system for the routine assessment of consumer outcomes. The implementation of routine outcomes assessment need not be conceptually complicated. However, it is highly probable that any simplistic attempts to introduce measures in lieu of the careful examination of the practicalities of the whole assessment process, will frustrate the implementation of any worthwhile system of outcomes data collection.

The reactions of service providers and consumers to this study highlight the importance of closer attention to process issues. It would be fair to say that many service providers in this study had little awareness of the outcomes movement prior to the study. While a substantial fund of good will towards the concepts of outcomes assessment was evident, some service providers indicated suspicion and cynicism in regard to the use of the selected measures.

Some consumers expressed bemusement with the process and its potential implications. The routine use of questionnaires looks, at least initially, like an unnatural and artificial intrusion into the accepted dynamics of the consumer-service provider relationship.

There is a need for greater clarification as to the global purposes of this whole exercise:

- Is the aim of implementing routine assessment of consumer outcomes on a national basis to improve the information collected by the Commonwealth and State governments for their own planning purposes?
- Is the aim of the Commonwealth and State governments to encourage changes in manner in which mental health services are delivered and evaluated?

A clear statement in relation to the purpose of routine outcomes assessment within the national context would be helpful.

Table A17: The implementation of routine use of measures of outcomes - A Short Example of how an individual service might approach introducing routine assessment of consumer outcomes.

The staff of Service X decide to implement routine outcomes measurement as a way to improve the monitoring of consumers' treatment progress, and to evaluate of the effectiveness of their service. A committee of staff members is convened to oversee the process. They invite members of the consumer and carer advisory group/s to participate.

The first task for the committee is to ascertain what the implementation of routine outcomes assessment will mean, and to identify what can and cannot be achieved. Members of the committee agrees that the first use of the information is to focus the staff, consumers and carers on obtaining the best possible results for the consumers of the service. They discuss whether the outcomes measures can be used between consumers and clinicians to help in setting treatment goals and plans, or whether they can be used to identify individuals or groups who warrant further attention. They consider how the results might be used to look for impacts of changes in the service, and to demonstrate to management, funders or community, the effectiveness of what they do and/or what could be achieved with additional resources.

Many pragmatic issues are discussed: who will collect information; how frequently information will be gathered; how confidentiality will be maintained; how consumers will be involved in the information gathering; how measures can be integrated with other data collection processes to ensure that the paperwork is not onerous; how technology can help; and finally, which outcomes measure(s) will be used.

The group decides how to ensure that the maximum amount of benefit is obtained for the effort expended: what feedback to consumers, service providers, teams, managers etc., is going to be helpful; how can these be provided; and, what should happen to ensure that the information can be accessed when the service has a special need.

Having outlined the needs of their consumers and the general range of service and treatment goals, the committee investigates a range of measures of outcomes that may be useful to them. They also discuss what additional information needs to be collected and how this system can be integrated with their existing processes for collecting information. The staff of Service X decide they need assistance with respect to the analyses and reporting of the data obtained from the measures and decide to consult with an external organisation.

An implementation strategy is developed with the staff and consumers, which ensures appropriate training and consultation. Following the implementation phase, the group arranges to monitor problems such as the completeness and quality of the information collected. They explore any problems discovered and make appropriate adjustments until the process works well.

Once established, as has happened elsewhere, the initial results suggest that more improvement is taking place than anyone expected. The people who are responding least are identified and arrangements made to ensure that additional attention is provided. Research is undertaken to relate data collected from the outcomes measures with service utilisation data, and consumer, carer, service provider impressions; and, to determine any relationships between outcomes data and the resources expended on various programs and treatment approaches.

The data from the measures of outcomes are used as one tool, albeit an important tool, in treatment and service planning. The system is regularly reviewed and assessed. Recommendations from staff, consumers and carers are taken into account and implemented where appropriate. This might include refining or replacing the original outcomes measures in light of more up-to-date developments in the field.

Recommendations in Relation to the Six Selected Measures

Having suggested that the focus for Phase 2 be shifted from 'measures' to 'methods', the consultancy acknowledges the need for introducing guidelines for choosing standardised measures for the routine assessment of consumer outcomes in mental health. In order to facilitate the development of effective processes, it is necessary to commence with the measures that appear at this time to offer the best potential as routine measures of outcomes.

We simply do not know enough about the utility of these measures, nor enough about the proposed uses of a 'national' measure, to conclusively recommend any of the six selected measures for such a purpose. More work needs to be undertaken to investigate the utility of these measures, as well as alternative measures and systems for outcomes assessment. However, a better understanding of the properties of the measures (validity, reliability, sensitivity to change) will only emerge if they are utilised in practice situations over a reasonable period of time and across large numbers of people.

Consequently the following recommendations are made, with regard to the six selected measures:

1. The six selected measures trialed in this study should be considered as a pool of measures which have the potential to facilitate the assessment of consumer outcomes in various situations depending on the needs of consumers and the focus of service provision.
2. Services should be encouraged to implement routine assessment of consumer outcomes. Where services already have a preferred measure or approach as part of their system, unless change is clearly warranted, they should continue with this and review it in due course. Where the needs of a service do not suggest a particular approach or instrument, it is recommended that the utility and applicability of the six selected measures should be considered, and in particular that,
 - a) The MHI should be considered for trialing for an interim period as a suitable self-report measure for the routine assessment of consumer outcomes in appropriate situations.
 - b) The HoNOS should be considered for trialing for an interim period as a service provider-rated measure of outcomes across all settings.
 - c) In situations where high levels of disability are expected, the LSP or RFS may be considered acceptable alternatives.
3. The measures should be chosen on the basis of the identified information needs of services and consumers in accordance with the principles stated above.
4. The implementation of any measure should be subject to review following a specified period of time to assess the adequacy of the measure. The implementation of any measure should involve consultation with consumers or their representatives, and the suitable training of staff, as described above.

The recommendations concerning the MHI and HoNOS should not be taken to suggest that the other measures are not useful. Clearly the data collected in this study indicates that the other measures are valid and reliable and may be useful in different circumstances. For instance, the SF-36 may be very useful when issues of physical health are clearly relevant (e.g., in general practice); and many consumers and service providers considered the BASIS-32, LSP and RFS to be equally reliable and valid. As well, the strong associations between the scores of the measures suggest a certain amount of equivalence in data collected by the measures.

What Does it All Mean?

Surveying the findings and conclusions of this study may appear daunting given the breadth of the findings and the limitations discussed in relation to the generalizability of the statistical results. It may also seem confusing that we recommend greater attention to the pragmatics of implementing outcomes and less attention on specific measures, while in turn, suggesting that two of the measures appear to be useful. Yet two of the aims of the study were to indicate the potential pitfalls which may confront services involved in implementing routine outcomes assessment, and to provide advice as to which of the six measures would be most useful to facilitate such processes.

The introduction of routine outcomes assessment is part of a process of change within mental health services that has been occurring over the last two decades. It is part of a wider push in all aspects

of human services for improved evaluation of service value and effectiveness. While many may be wary of such developments, in general they offer the promise of improvements in the standards of mental health care delivery.

The primary purpose of this study was to collect data that would inform Phase 2 of the consumer outcomes research program. But the implications of the findings are of value to a much wider audience. The findings of this study have demonstrated that there is more to the implementation of measures of outcomes than simply choosing and administering questionnaires. Further research is necessary to examine how useful such measures are to clinical and resource-allocation decision-making processes. As well there should be a shift towards discussion of the information needs of local and regional services; and, of how data collected at the grass-roots level can be meaningfully translated into information that is relevant to higher levels of management in mental health.

Routine outcomes assessment can assist consumers, service providers and policy makers to make rational care-related choices based on a better insight into the likely effect of such choices on the lives of consumers (Ellwood, 1988). The present research program being undertaken on behalf of the AHMAC National Mental Health Working Group is a positive step towards the implementation of such systems in Australian mental health services, and, as a consequence, towards improving the quality of service provision. However, the research to date suggests that the 'simplistic' application of outcomes measures will not benefit service providers and consumers in the long term. The results of this study encourage a reconsideration of the 'simplistic' approaches to outcomes assessment and a reassessment of how routine outcomes assessment can be usefully implemented by individual services.

6.0 References

- Aitken, R.C.B. (1969). Measurement of feelings using visual analogue scales. *Proceedings of the Royal Society of Medicine*, 62, 989-993.
- Albrecht, T.L., Johnson, G.M., and Walther, J.B. (1993). Understanding communication processes in focus groups. In D.L. Morgan, *Successful Focus Groups: Advancing the State of the Art*. Newbury Park: Sage Publications.
- American Psychological Association. (1985). *Standards for educational and psychological testing*. Washington, D.C.: American Psychological Association.
- Andrews, G., Peters, L., and Teeson, M. (1994). *Measurement of consumer outcome in mental health: A report to the National Mental Health Information Strategy Committee*. Sydney: Clinical Research Unit for Anxiety Disorders.
- Anthony, W.A., and Nemeck, P.B. (1984). Psychiatric rehabilitation. In A.S. Bellack (Ed) *Schizophrenia: Treatment, management, and rehabilitation*. New York: Grune and Stratton.
- Bagley, C., and Genuis, M. (1991). Psychology of computer use: Evaluation of a computerized questionnaire in a population of young adult males. *Perceptual and Motor Skills*, 72(1), 287-288.
- Benjamin, K.L., Perfetto, E.M., and Greene, R.J. (1995). Public policy and the application of outcomes assessments: paradigms versus politics. *Medical Care*, 33(Suppl4), AS299-306.
- Berwick, D.M., Murphy, J.M., Goldman, P.A., Ware, J.E., Barsky, A.J., and Weinstein, M.C. (1991). Performance of a five-item mental health screening test. *Medical Care*, 29(2), 169-176.
- Blatt, S.J., Sanislow III, C.A., and Zuroff, D.C. (1996). Characteristics of effective therapists: Further analyses of data from the National Institute of Mental Health treatment of depression collaborative research program. *Journal of Consulting and Clinical Psychology*, 64(6), 1276-1284.
- Bond, A., Shine, P., and Bruce, M. (1995). Validation of visual analogue scales in anxiety. *International Journal of Methods in Psychiatric Research*, 5, 1-9.
- Bowen, D.J., Kestin, M., McTiernan, A., Carrell, D., and Green, P. (1995). Effects of dietary fat intervention on mental health in women. *Cancer Epidemiol Biomarkers Prev*, 4(5), 555-559.
- Brekke, J.S. (1992). An examination of the relationships among three outcome scales in schizophrenia. *The Journal of Nervous and Mental Disease*, 180(3), 162-168.
- Brown, R.B., McCartney, S., and Bell, L. (1995). Why the NHS should abandon the search for the universal outcome measure. *Health Care Analysis*, 3, 191-195.
- Butcher, J.N. (1994). Psychological assessment by computer: Potential gains and problems to avoid. *Psychiatric Annals*, 24(1), 20-24.
- Campbell, D.T. (1996). Regression artifacts in time-series and longitudinal data. *Evaluation and Program Planning*, 19(4), 377-389.
- Cella, D.F., and Perry, S.W. (1986). Reliability and concurrent validity of three visual-analogue mood scales. *Psychological Reports*, 59, 827-833.
- Chandler, D., Meisel, J., Hu, T.W., McGowen, M., and Madison, K. (1996). Client outcomes in a three-year controlled study of an integrated service agency model. *Psychiatric Services*, 47(12), 1337-1343.
- Ciarlo, J.A., Edwards, D.W., Kiresuk, T.J., Newman, F.L., and Brown, T.R. (1986). *The assessment of client/patient outcome techniques for use in mental health programs* (Contract No. 278-80-005). Washington D.C.: National Institute of Mental Health.
- Clark, A., and Friedman, M.J. (1983). Nine standardized sales for evaluating treatment outcome in a mental health clinic. *Journal of Clinical Psychology*, 39(6), 939-950.
- Conte, H.R., Plutchik, R., Picard, S., Karasu, T.B., and Vaccaro, D. (1988). Self-report measures as predictors of psychotherapy outcome. *Comprehensive Psychiatry*, 29(4), 355-360.
- Corrigan, P.W., Buican, B., and McCracken, S. (1996). Can severely mentally ill adults reliably report their needs? *Journal of Nervous and Mental Disease*, 184(9), 523-529.
- Cronbach, L. (1984). *Essentials of psychological testing*. Cambridge: Harper & Row.
- Cronbach, L.J., and Furby, L. (1970). How should we measure "change" - or should we? *Psychological Bulletin*, 74, 68-80.
- Cummins, R.A., and Baxter, C. (1994). Choice of outcomes measures in service evaluations for people with an intellectual disability. *Evaluation Journal of Australasia*, 6(1), 22-30.
- Davies, A.R., Sherbourne, C.D., Peterson, J.R., and Ware, J.E. (1988). *Scoring manual: Adult health status and patient satisfaction measures used in RAND's Health Insurance Experiment*. Santa Monica: RAND Corporation.

- Dawson, S., Manderson, L., and Tallo, V.L. (1993). *A manual for the use of focus groups*. Boston: INFDC.
- Deyo, R.A., and Patrick, D.L. (1995). The significance of treatment effects: The clinical perspective. *Medical Care*, 33(4), AS286-AS291.
- Dickerson, F.B., Boronow, J.J., Ringel, N., and Parente, F. (1997). Lack of insight among outpatients with schizophrenia. *Psychiatric Services*, 48(2), 195-199.
- Dorfman, R.A., Lubben, J.E., Mayer-Oakes, A., Atchison, K., (1995). Screening for depression among a well elderly population. *Social Work*, 40(3), 295-304.
- Eisen, S.V. (1995). Assessment of subjective distress by patients' self-report versus structured interview. *Psychological Reports*, 76, 35-29.
- Eisen, S.V., and Dickey, B. (1996). Mental health outcome assessment: the new agenda. *Psychotherapy*, 33(2), 181-189.
- Eisen, S.V., Dill, D.L., and Grob, M.C. (1994). Reliability and validity of a brief patient-report instrument for psychiatric outcome evaluation. *Hospital and Community Psychiatry*, 45, 242-247.
- Eisen, S.V., Grob, M.C., and Dill, D.L. (1991). Tapping the patient's perspective. In S.M. Mirin, J.T. Gossett, and M.C. Grob *Psychiatric Treatment: advances in outcomes research*. Washington, D.C., American Psychiatric Press.
- Ellwood, P.M. (1988). Outcomes management: a technology of patient experience. *New England Journal of Medicine*, 318(23), 1549-1556.
- Epstein, M., and Wadsworth, Y. (1994). *Understanding and involvement: Consumer evaluation of acute psychiatric hospital practice*. Brunswick: Victorian Mental Illness Awareness Council.
- Fernandez de Larrinoa Palacios, P., Bulbena Vilarrasa, A., and Dominguez Panchon, A.I. (1992). Spanish adaptation of the Life Skills Profile. Profile of activities of daily living. Structure and factorial composition. *Actas Luso Esp Neurol Psiquiatr Cienc Afines*, 20(2), 71-75.
- Fuller, T.D., Edwards, J.N., Vorakitphokatorn, S., and Sermsri, S. (1993) Using focus groups to adapt survey instruments to new populations: Experience from a developing country. In D.L. Morgan, *Successful focus groups: Advancing the state of the art*. Newbury Park: Sage Publications.
- Goldney, R.D., Fisher, L.J., Walmsley, S., Kent, P., and Cooper, A.W. (1996). A pilot study of the Health of the Nation Outcome Scale as a measurement of outcome in a private psychiatric hospital. *Australasian Psychiatry*, 4(6), 319-321.
- Goodman, S.H., Sewell, D.R., Cooley, E.L., and Leavitt, N. (1993). Assessing levels of adaptive functioning: the Role Functioning Scale. *Community Mental Health Journal*, 29, 119-131.
- Green, R.S., and Gracely, E.J. (1987). Selecting a rating scale for evaluating services to the chronically mentally ill. *Community Mental Health Journal*, 23(2), 91-102.
- Gross, D., Fogg, L., and Conrad, B. (1993). Designing interventions in psychosocial research. *Annals of Psychiatric Nursing*, 7(5), 259-264.
- Grummon, K., Rigby, E.D., Orr, D., Procidano, M., and Reznikioff, M. (1994). Psychosocial variables that affect the psychological adjustment of IVDU patients with AIDS. *Journal of Clinical Psychology*, 50(4), 488-502.
- Guy, W. (1976). ECDEU assessment manual for psychopharmacology. US Dept. Health, Education and Welfare Publication (ADM) 76-338. Rockville, Md: National Institute of Mental Health.
- Guyatt, G., Walter, S., and Norman, G. (1987). Measuring change over time: Assessing the usefulness of evaluative instruments. *Journal of Chronic Diseases*, 40, 171-178.
- Guyatt, G.H., Deyo, R.A., Charlson, M., Levine, M.N., and Mitchell, A. (1989). Responsiveness and validity in health status measurement: a clarification. *Journal of Clinical Epidemiology*, 42(5), 403-408.
- Hambridge, J.A., and Rosen, A. (1994). Assertive community treatment for the seriously mentally ill in suburban Sydney: a programme description and evaluation. *Australian and New Zealand Journal of Psychiatry*, 28, 438-445.
- Hays, R.D., Wells, K.B., Sherbourne, C.D., Rogers, W., and Spritzer, K. (1995). Functioning and well-being outcomes of patients with depression compared with chronic general medical illnesses. *Archives of General Psychiatry*, 52, 11-19.
- Jacobson, N.S., and Traux, P. (1991). Clinical significance: a statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology*, 59(1), 12-19.
- Jenkins, R. (1990). Towards a system of outcome indicators for mental health care. *British Journal of Psychiatry*, 157, 500-514.
- Jerrell, J.M., and Ridgeley, M.S. (1995). Evaluating changes in symptoms and functioning of dually diagnosed clients in specialized treatment. *Psychiatric Services*, 46(3), 233-238.

- Johnson, L.D. (1995). *Psychotherapy in the age of accountability*. New York: W.W. Norton.
- Kemp, D.R. (1993). *International handbook on mental health policy*. Westport, Conn.: Greenwood Press.
- Keppel, G. (1982). *Design and analysis: a researcher's handbook*. Prentice-Hall: New Jersey.
- Kirkby, K.C., and Lambert, T.J. (1996). Computer aids to treatment in psychiatry. *Australian and New Zealand Journal of Psychiatry*, 30, 142-145.
- Knodel, J. (1993) The design and analysis of focus group studies: a practical approach. In D.L. Morgan, *Successful focus groups: Advancing the state of the art*. Newbury Park: Sage Publications.
- Kobak, K.A. (1996). Computer-administered symptom rating scale. *Psychiatric Services*, 47(4), 367-369.
- Kornblith, A.B., Hollis, D.R., Zuckerman, E., Lyss, A.P., Canellos, G.P., Cooper, M.R., Herndon, J.E. II, Phillips, C.A., Abrams, J., Aisner, J., et al. (1993). Effect of megestrol acetate on quality of life in a dose-response trial in women with advanced breast cancer. *Journal of Clinical Oncology*, 11(1), 2081-2089.
- Kornblith, A.B., Thaler, H.T., Wong, G., Vlamis, V., Lepore, J.M., Loseth, D.B., Hakes, T., Hoskins, W.J., and Portenoy, R.K. (1995). Quality of life of women with ovarian cancer. *Gynecological Oncology*, 59(2), 231-42.
- Kraemer, H.C., and Telch, C.F. (1992). Selection and utilization of outcome measures in psychiatric clinical trials. *Neuropsychopharmacology*, 7(2), 85-94.
- Liang, J., Wu, S.C., Krause, N.M., Chiang, T.L., and Wu, H.Y. (1992). The structure of the Mental Health Inventory among Chinese in Taiwan. *Medical Care*, 30(8), 659-676.
- Locke, S.D., and Gilbert, B.O. (1995). Method of psychological assessment, self-disclosure, and experiential differences: A study of computer, questionnaire, and interview assessment formats. *Journal of Social Behaviour and Personality*, 10(1), 255-263.
- Lord, F.M., and Novick, M.N. (1968). *Statistical theories for mental test scores*. Reading, Mass.: Addison-Wesley Publishing Company.
- Luria, R.E. (1975). The validity and reliability of the visual analogue mood scale. *Journal of Psychiatric Research*, 12, 51-57.
- Massey, O.T., & Wu, L. (1994). Three critical views of functioning: Comparisons of assessments made by individuals with mental illness, their case managers, and family members. *Evaluation and Program Planning*, 17(1), 1-7.
- Marks, I. (1996). A computerized system to measure treatments outcomes and cost. *Psychiatric Services*, 47(8), 811-812.
- Maruff, P., Wood, S., Currie, J., McArthur-Jackson, C., Malone, V., and Benson, E. (1994). Computer-administered visual analogue mood scales: rapid and valid assessment of mood in HIV positive individuals'. *Psychological Reports*, 74, 39-42.
- McCormack, H.M., Horne, D.J., and Sheather, S. (1988). Clinical applications of visual analogue scales: a critical review. *Psychological Medicine*, 18, 1007-1019.
- McHorney, C.A., Ware, J.E.J., and Raczek, A.E. (1993). The MOS 36-item Short-Form Health Survey (SF-36): II. Psychometric and clinical tests of validity in measuring physical and mental health constructs. *Medical Care*, 31, 247-263.
- McPheeters, H.L. (1984). Statewide mental health outcome evaluation: a perspective of two Southern States. *Community Mental Health Journal*, 20(1), 44-55.
- Meltzer, H.Y. (1995). Multiple-outcome criteria in schizophrenia: an overview of outcome with clozapine. *European Psychiatry*, 10(Suppl.1), 19s-25s.
- Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist*, 50(9), 741-749.
- Mirin, S.M., and Namerow, M.J. (1991). Why study treatment outcome? *Hospital and Community Psychiatry*, 42(10), 1007-1013.
- Mirin, S.M., Gossett, J.T., and Grob, M.L. (1991). *Psychiatric treatment: Advances in outcome research*. American Psychiatric Press: Washington.
- Monti, P.M., Corriveau, D.P., and Currna, J.P. (1982). Assessment of social skills in the day hospital: Does the clinician see something than other than the researcher sees? *International Journal of Partial Hospitalization*, 1(3), 245-250.
- Morgan, D.L. (1993). *Successful focus groups: Advancing the state of the art*. Sage Publications: Newbury Park.

- Murrelle, L., Ainsworth, B.E., Bulger, J.D., Holliman, S.C., et al. (1992). Computerized mental health risk appraisal for college students: User acceptability and correlation with standard pencil-and-paper questionnaires. *American Journal of Health Promotion*, 7(2), 90-92.
- National Institute of Mental Health. (1985). Clinical Global Impressions. *Psychopharmacology Bulletin*, 21(4), 839-843.
- National mental health policy*. (1992). Canberra: Australian Government Publishing Service.
- National standards for mental health services* (1996). Canberra: Australian Government Publishing Service.
- Newman, F.L. (1980). Strengths, uses, and problems of global scales as an evaluation instrument. *Evaluation and Program Planning*, 3, 257-268.
- Nunnally, J.C. (1978). *Psychometric theory*. McGraw-Hill: New York.
- Ostroff, J.S., Woolverton, K.S., Berry, C., and Lesko, L.M. (1996). Use of the Mental Health Inventory with adolescents: a secondary analysis of the RAND Health Insurance study. *Psychological Assessment*, 8(1), 105-107.
- Parker, G., and Hadzi-Pavlovic, D. (1995). The capacity of a measure of disability (the LSP) to predict hospital readmission in those with schizophrenia. *Psychological Medicine*, 25(1), 157-63.
- Parker, G., Rosen, A., Emdur, N., and Hadzi-Pavlovic, D. (1991). The Life Skills Profile: psychometric properties of a measure assessing function and disability in schizophrenia. *Acta Psychiatrica Scandinavica*, 83, 145-152.
- Pedhazur, E.J., and Pedhazur-Schmelkin, L. (1991). *Measurement, design and analysis: an integrated approach*. Lawrence Erlbaum Associates: Hillsdale, N.J.
- Perreault, M., Rogers, W.L., Leichner, P., and Sabourin, S. (1996). Patients' requests and satisfaction with services in an outpatient psychiatric setting. *Psychiatric Services*, 47(3), 287-292.
- Pfennings, L., Cohen, L., and van der Ploeg, H. (1995). Preconditions for sensitivity in measuring change: visual analogue scales compared to rating scales in a Likert format. *Psychological Reports*, 77, 475-480.
- Piersma, H.L., and Boes, J.L. (1995). Agreement between patient self-report and clinician-ratings: concurrence between the BSI and the GAF amongst psychiatric inpatients. *Journal of Clinical Psychology*, 51(2), 153-157.
- Portenoy, R.K., Kornblith, A.B., Wong, G., Vlamis, V., Lepore, J.M., Loseth, D.B., Hakes, T., Foley, K.M., and Hoskins, W.J. (1994). Pain in ovarian cancer patients. Prevalence, characteristics, and associated symptoms. *Cancer*, 74(3), 907-915.
- Priebe, S., and Gruyters, T. (1994). Patients' and caregivers' initial assessments of day-hospital treatment and course of symptoms. *Comprehensive Psychiatry*, 35, 234-238.
- Pulice, R.T., McCormick, L.L., and Dewees, M. (1995). A qualitative approach to assessing the effects of system change on consumers, families, and providers. *Psychiatric Services*, 46(6), 575-579.
- Raphael, B., Hugh, G., and Stewart, G. (1997). Mental health and mental health outcomes. In *Integrating health outcomes measurement in routine health care conference 1996 proceedings*. AIHW Cat. No. HOC 1. Canberra: Australian Institute of Health and Welfare.
- Relman, A. (1988). Assessment and accountability: the third revolution in health care. *New England Journal of Medicine*, 319, 1220-1222.
- Rogosa, D., and Willett, J. (1983). Demonstrating the reliability of the difference score in the measurement of change. *Journal of Educational Measurement*, 20(4), 335-343.
- Rogosa, D., Brandt, D., and Zimowski, M. (1982). A growth approach to the measurement of change. *Psychological Bulletin*, 92(3), 726-748.
- Rosen, A., Hadzi-Pavlovic, D., and Parker, G. (1989). The Life Skills Profile: A measure assessing function and disability. *Schizophrenia Bulletin*, 15, 325-337.
- Ross, H.E., Swinson, R., Larkin, E.J., & Doumani, S. (1994). Diagnosing comorbidity in substance abusers. *Journal of Nervous and Mental Disease*, 182(10), 556-563.
- Russell, G.K.G., Peace, K.A., and Mellsop, G.W. (1986). The reliability of a micro-computer administration of the MMPI. *Journal of Clinical Psychology*, 42(1), 120-122.
- Sansoni, J. (1995). Measurement issues and the SF-36 instrument. *Health Outcomes*, 9-14.
- Sawyer, M., Sarris, A., Quigley, R., Baghurst, P., and Kalucy, R. (1990). The attitude of parents to the use of computer-assisted interviewing in a child psychiatry service. *British Journal of Psychiatry*, 157, 675-678.
- Sawyer, M., Sarris, A., and Quigley, R. (1993). Attitude of clinicians toward use of computers in hospital settings. *Professional Psychology Research and Practice*, 24(2), 220-223.

- Sederer, L.I., and Dickey, B. (Eds.) *Outcomes assessment in clinical practice*. Baltimore: Walter and Williams, 1996.
- Sederer, L.I., Dickey, B., and Hermann, R.C. (1996). The imperative of outcomes assessment in psychiatry. In L.I. Sederer and B. Dickey (Eds.) *Outcomes assessment in clinical practice*. Baltimore: Walter and Williams, 1996, p.p. 1-8.
- Seligman, M.E.P. (1995). The effectiveness of psychotherapy: the Consumer Reports study. *American Psychologist*, 50(12), 965-974.
- Siedman, A.D., Portenoy, R., Yao, T.J., Lepore, J., Mont, E.K., Kortmansky, J., Onetto, N., Ren, L., Grechko, J., Beltangady, M., et al. (1995). Quality of life in phase II trials: a study of the methodology and predictive value in patients with advanced breast cancer treated with paclitaxel plus granulocyte colony-stimulating factor. *Journal of the National Cancer Institute*, 87(17), 1316-1322.
- Skinner, H.A., and Allen, B.A. (1983). Does the computer make a difference? Computerized versus face-to-face versus self-report assessment of alcohol, drug and tobacco use. *Journal of Consulting and Clinical Psychology*, 51(2), 267-275.
- Smith, G.R., Fischer, E.P., Nordquist, C.R., Mosley, C.L., and Ledbetter, N.S. (1997). Implementing outcomes management systems in mental health settings. *Psychiatric Services*, 48(3), 364-368.
- Smith, G.R., Ross, R.L., and Rost, K.M. (1996a). Depression outcomes module (DOM). In L.I. Sederer and B. Dickey (Eds.) *Outcomes assessment in clinical practice*. Baltimore: Walter and Williams, 1996, p.p.82-84.
- Smith, G.R., Ross, R.L., and Rost, K.M. (1996b). Substance abuse outcomes module (SAOM). In L.I. Sederer and B. Dickey (Eds.) *Outcomes assessment in clinical practice*. Baltimore: Walter and Williams, 1996, p.p.85-88.
- Sonnaburg, K. (1996). Meaningful measurement in psychotherapy. *Psychotherapy*, 33(2), 160-170.
- Spinhoven, P., Labbe, M.R., and Rombouts, R. (1993). Feasibility of computerized psychological testing with psychiatric outpatients. *Journal of Clinical Psychology*, 49(3), 440-7.
- Stedman, T. (1996). Approaches to measuring quality of life and their relevance to mental health. *Australian and New Zealand Journal of Psychiatry*, 30, 731-740.
- Stevenson, C. (1995). MOS SF-36 questionnaire- interim Australian norms. *Health Outcomes*, 14-17.
- Streiner, D.L., and Norman, G.R. (1989). *Health measurement scales*. Oxford University Press: Oxford.
- Sullivan, C.W., & Grubea, J.M. (1991). Who does well in a day treatment program? Following patients through 6 months of treatment. *International Journal of Partial Hospitalization*, 7, 101-107.
- Sutherland, H.J., and Till, J.E. (1993). Quality of life assessments and levels of decision making: Differentiating objectives. *Quality of Life Research*, 2, 297-303.
- Thompson, C. (1992). Research questionnaires in routine clinical practice: Are they useful? In K. Hawton and P. Cowen (Eds) *Practical problems in clinical psychiatry*. New York: OUP, p.p. 231-241.
- Trauer, T., Duckmanton, R.A., and Chiu, E. (1995). The Life Skills Profile: a study of its psychometric properties. *Australian and New Zealand Journal of Psychiatry*, 29, 492-499.
- Tsevat, J., Solzan, J.G., Kuntz, K.M., Ragland, J., Currier, J.S., Sell, R.L., and Weinstein, M.C. (1996). Health values of patients infected with human immunodeficiency virus. *Medical Care*, 34(1), 44-57.
- Veit, C.T., and Ware, J.E.J. (1983). The structure of psychological distress and well-being in general populations. *Journal of Clinical and Consulting Psychology*, 51, 730-745.
- Walters, G., Kirkby, K., Marks, I., Whiteford, H., Andrews, G., and Swinson, R. (1996). Outcome measurement: sharing experiences in Australia. *Australasian Psychiatry*, 4(6), 316-318.
- Ware, J.E. (1996). The MOS 36-Item Short Form Health Survey (SF-36). In L.I. Sederer and B. Dickey (Eds.) *Outcomes assessment in clinical practice*. Baltimore: Walter and Williams, 1996, p.p.61-64.
- Ware, J.E., Snow, K.K., Kosinski, M., and Gandek, B. (1993). *SF-36 Health Survey manual and interpretation guide*. Boston, Massachusetts: The Health Institute, New England Medical Center, 1993.
- Ware, J.E., and Sherbourne, C.D. (1992). The MOS 36-Item short-form health survey (SF-36): 1. Conceptual framework and item selection. *Medical Care*, 30, 473-481.

- Weinstein, M.C., Berwick, D.M., Goldman, P.A., Murphy, J.M., and Barsky, A.J. (1989). A comparison of three psychiatric screening tests using receiver operating characteristic (ROC) analysis. *Medical Care*, 27(6), 593-607.
- Weisband, S., and Keisler, S. (1996). Self-disclosure on computer forms: Meta-analysis and implications. (*Internet transmission*)
- Weisberg, H.I. (1979). Statistical adjustments and uncontrolled studies. *Psychological Bulletin*, 86(3), 1149-1164.
- Wells, K.B. (1995). Cost-containment and mental health outcomes: Experience from U.S. studies. *British Journal of Psychiatry*, 166(Suppl. 27), 43-51.
- White, J., White, K., and Razani, J. (1984). Effects of endogeneity and severity on consistency of standard depression rating scales. *Journal of Clinical Psychiatry*, 45, 260-261.
- Williams, R.H., and Zimmerman, D.W. (1996). Are simple gain scores obsolete? *Applied Psychological Measurement*, 20(1), 59-69.
- Wing, J., Curtis, R., and Beevor, A. (1996). Health of the Nation Outcome Scales: Report on research. London: Royal College of Psychiatrists.
- World Health Organization (1992). *The ICD-10 classification of mental and behavioural disorders. Clinical descriptions and diagnostic guidelines*. Geneva : World Health Organization.
- Yokley, J., Coleman, D.J., and Yates, B.T. (1990). Cost-effectiveness of three child mental health assessment methods: Computer-assisted is effective and inexpensive. *Journal of Mental Health Administration*, 17(1), 99-107.
- Zealley, A.K., and Aitken, R.C.B. (1976). Measurement of mood. *Proceedings of the Royal Society of Medicine*, 62, 993-996.

Appendix 1: Executive Summary-Measurement of Consumer Outcome (Andrews et al., 1994)

This report provides a review of the issues, literature, instruments available, and implementation strategies to facilitate the development of consumer outcome measurement for Commonwealth and State funded mental health services.

Preamble

1. Consumer outcome is defined as the effect on a patient's health status attributable to an intervention by a health professional or health service.
2. The objectives of the National Mental Health Policy that pertain to consumer outcome are
 - Objective 30 "To institute regular reviews of outcomes of services provided to persons with serious mental health problems and mental disorders as a central component of mental health service delivery" and;
 - Objective 32 "To encourage the development of national outcome standards for mental health services, and systems for assessing whether services are meeting these standards."
3. The central role of a health system is defined as the treatment of persons who are ill. The central role of good clinical care in this process is emphasized, and the routine measurement of outcome is seen as an integral part of good clinical care.
4. The potential consumers are the 25% to 30% of the population who will meet criteria for a mental disorder in any year. The actual consumers, or patients of the health services, are the one third of the potential consumers who actually seek treatment from the medical sector in any year. Three quarters of those who do seek treatment are seen by general practitioners with the remainder being shared between private psychiatrists and the public sector mental health and addictive services. This report is focused on outcome measurement in consumers with serious mental disorders. Three percent of the population meet criteria for a serious mental disorder in any year. Their principal diagnoses are schizophrenia (10%), mood disorder (44%), anxiety disorder (25%), or substance use disorder (21%). Currently, only half get treatment.
5. Outcome measurements are of interest to five groups: consumers, carers, clinicians, local health service administrators and State and Federal health service administrators. Privacy considerations are paramount. Administrators should only have access to deidentified or grouped data, carers should only have access to their consumer's data with the consumer's permission, clinicians should have access to data about patients in their care, and consumers to their own records.
6. The repeated measurement of diagnoses, symptoms, disability and risk factors using standardised measures is regarded as likely to improve treatment as well as providing the information necessary for identifying the outcome of treatment. Measurement, provided it does not detract from treatment time, does not interfere with therapy. Information on outcome will be of considerable value to health planners and administrators at state and national levels.
7. The presumption that changes in outcome can be automatically attributed to a treatment can be considered to be true for routine clinical practice, but this presumption would require considerable supporting data if changes in outcome were to be considered evidence of the efficacy of a treatment or the efficacy of a clinic or service.
8. Outcome data will be the first clinical data routinely available to administrators. First, because aggregated disability data will establish the load on the service and second, because the changes in these disability measures will provide an outcome performance indicator for individual clinics and clinicians. If such a system is going to comply with privacy requirements then only deidentified or aggregated data should be sent from a patient's clinical record to administration or head office.
9. Population outcome indicators, while measures of health, are not favoured because they would not allow the cause of any change to be attributed. A multidimensional measure of individual outcome could serve to guide clinical practice and, when aggregated with measures from other consumers, serve to guide the delivery of services.
10. Consumer outcome measures are reviewed in regard to their applicability, acceptability, practicality, reliability, validity and sensitivity to change. The advantages and disadvantages of questionnaires and interviews, completed by clinicians or consumers are delineated.

Measures which are suitable for use in routine clinical practice are likely to be brief, low cost, multidimensional measures which require minimal training for their administration, scoring and interpretation. Other more complex measures are reviewed only if they are likely to be of value for collateral studies on outcome.

Methodology

1. Fifty three per cent of 377 providers, non-government organisations and carer and consumer groups responded to a survey, either by completing the questionnaire or by offering views. All persons known to be expert in the measurement of disability and outcome, both in Australia and overseas, were contacted for advice. Peak bodies were consulted and their opinions given special weight.
2. A keyword search on Medline returned 350 references which were reviewed for possible outcome measures. Outcome measures were also identified from specialty texts, citation lists and other sources. Ninety five measures were reviewed by two independent raters in relation to their function and psychometric properties.

Results of the Consultation Process

1. The replies to the survey show that disability and quality of life is viewed as more important areas of outcome than either symptoms or satisfaction with services. Consulting the consumer, either by questionnaire or interview is deemed better than either clinician opinion or carer rating. Respondents thought that the principal use of such data is to improve service efficacy.
2. The professional bodies representing psychiatrists, mental health nurses, clinical psychologists and occupational therapists recognise the importance of the measurement of consumer outcome, but express reservations as to how it might be implemented.
3. The views of the participants at the National Workshop on Consumer Outcome in Mental Health were diverse, but there is consensus about the need to measure outcome, especially with a reliable multidimensional measure.

Results: Review of Existing Generic Measures

1. The reviews of symptom measures are restricted to those applicable to broad groups of patients. Eighteen such measures are reviewed.
2. Level of functioning is an important target to measure and 20 measures of functioning are reviewed, with one, the Role Functioning Scale being recommended.
3. Seventeen quality of life measures are reviewed but none meet the criteria for a consumer outcome measure suitable for routine use.
4. Five measures of family burden are reviewed. None are recommended for routine use but some would be valuable in studies focused on family burden.
5. Eight measures of satisfaction with services are reviewed but all are too narrow in focus to serve as the sole consumer outcome measure.
6. Twenty-seven multidimensional measures met criteria for review. Five met minimum criteria but one, the Life Skills Profile, is not recommended on the grounds of cost. Four are recommended for use as routine measures of consumer outcome. These four are the Behaviour and Symptom Identification Scale (BASIS), Health of the Nation Outcome Scales (HoNOS), Medical Outcomes Study Short Form (SF36), and the Mental Health Inventory (MHI).

Recommendations

1. Field test all five measures and explore the level to which they are applicable, acceptable, practical, reliable, valid and sensitive to change in the appropriate consumer groups and service settings.
2. Establish pilot sites in each state to implement the measurement of consumer outcomes.
3. Develop training packages for consumers, clinicians and administrators as to the meaning of the information being produced by the measures of consumer outcome.
4. Establish a mechanism to coordinate the above activities and facilitate the nationwide implementation of routine consumer outcome measurement.

Appendix 2: Consumer Consultant's Report on Focus Groups Process, Mr Tony Hennessey

I am pleased to be given the opportunity to present a report on this research project. I found that the high level of participation by consumers in the ninety (90) minute discussion groups to be a major indicator of the importance this project is to them. It is indeed a credit to the researchers to have gained the confidence of the consumers in this study. It has allowed the consumers to be frank and honest with their responses. Past research has shown that the consumers of mental health services have expertise in assessing not only their service providers, but most importantly service delivery. As highlighted in the National Mental Health Strategy, consumer input is seen as a valuable asset.

Part of the problems faced by consumers is that their voice is often not heard, or worst still, not taken on its full merit. Too often consumers have been involved in projects, only to find that nothing has changed!

This has had the profound effect of consumers being critical of research projects, and this was highlighted by many consumers involved in this particular project. The "not another questionnaire" comment which was made by some of the consumers seemed to be reduced somewhat once they were informed about the nature and confidentiality of this project, and helped to lower the myth of "will this make a difference?"

The broad feeling of these sessions was that this is an opportunity for their voices to be heard, and although some of the points which follow may sound critical of the project, this is not so! They are only points that I feel should be taken into consideration when reading the outcomes of this project. I must also mention that the level of honesty, and inclusion of the consumers, is a major factor leading to the success of this project.

As we are heading towards deinstitutionalization it is critical that not just service providers, but also consumers, can see evidence that the community can provide the services that are required. Accountability is very important not only to funding bodies and service providers, but also to consumers and the general community.

As I was not present at the time that the questionnaires were completed, I am only able to comment on the consumer focus group workshops. Of particular interest was the response in relation to "how people felt about filling in the questionnaires?" Almost all were voicing similar concerns. I was concerned that most of the participants were affected by the process of filling out the questionnaires.

I have to highlight the honesty that was shared during this project. I felt very comfortable with the approach of my fellow researchers. This also showed at the workshops through the participation level of the consumers and the respect shown to my fellow researchers. It also must be noted that since this part of the project has been completed, several consumers have approached me for feedback. This once again highlights their keenness to be part of the project, and indeed the importance and impact they feel this project will have on future service delivery.

All in all I was very impressed by the style and format of the workshops. It is encouraging to see that not only are consumers involved in this project, but that they were financially rewarded as well.

General issues raised during the consumer workshops:

- The language used in the Questionnaires;
- The relevance of the questions to individual illness;
- The actual outcome of filling out these Questionnaires;
- How can you rate your wellness by a standard form?; and
- The impact on the consumer of filling out Questionnaires.

The language used in the Questionnaires

Consumers voiced that the language used in the questionnaires was out of date. More modern questionnaires with appropriate language are required.

The relevance of the questions to individual illness

A consensus regarding the questionnaires was the inability of a general questionnaire to address not just different diagnoses, but that they do not provide open-ended questions, as well as not taking into account wellness and people's individuality.

The actual outcome of filling out these Questionnaires

It is taken for granted that the information gathered will benefit service providers, but how will it benefit consumers unless a tailor made questionnaire is developed?

How can you rate your wellness by a standard form

Several factors influence this statement, not just the broad range of mental illness, but the actual mood of the consumer at the time of filling out the questionnaire. The questions themselves do not address day to day influences that we are all prone to. The scoring of each question is difficult on a number scale, which could be altered by having a sliding scale instead.

The impact on the consumer of filling out Questionnaires.

It was raised that actually filling out the questionnaire can put people back into their illness. This must be taken into consideration as many consumers have mood swings that must have an impact on their answers, but more importantly sets them back in their recovery. The project did have an informal debrief, and this point must be taken into account when future projects are undertaken.

Other issues to be taken into account when considering the outcomes of this project.

- Was the sample a fair representation of mental health consumers?
- All participants at the consumer workshops presented in my opinion to be well.
- Environment.
- Group dynamics.

• Was the sample a fair representation of mental health consumers?

It is a very difficult thing to get a true representation of any sample group and this is compounded by the fact that mental illness has such a broad diagnostic range, and that is influenced by the degree of their illness.

The sample of consumer groups needs to be expanded on. The age and gender, racial, cultural, forensic, geographical and multi-diagnosis populations were not taken into account.

• All participants at the consumer workshops presented in my opinion to be well

The sample of consumers presented in quite definite advanced recovery. Some participants spoke of becoming unwell due to the process of filling out a questionnaire, but none were unwell while filling them out.

Environment

The university was a very difficult place to access, and even I felt threatened by the locality of the workshops. Several consumers had difficulty locating the correct building.

The workshop at PARU had a conflict of interest for the consumers as they access the services in the building, and confidentiality is always an issue.

Consumers attending the workshop at the Qld Association for Mental Health were more relaxed as the environment was not only known to them, but the building has a relaxed feel to it.

Group dynamics

The group dynamics played a large part in the easy flow of information. The negative effect is that the group gets into a “my story is better than yours” syndrome. People are often led in a different direction due to peer/group persuasion.

Care was taken to include all participants in the discussions held, and my role was to encourage discussion. This was, I believe, achieved. I am aware that groups sometimes are dominated by one or two speakers, but this was not the case in this study. At the end of each workshop I endeavoured to engage with consumers, who, for various reasons, did not participate in the Focus Groups as fully as other participants. I did this in order to provide them with the opportunity to have their say about this process.

Appendix 3: Tables of Results

Table B1.1 (a): Australian General Population (Persons) Norms for the SF-36 (Eight Sub-scales) (Source: Stevenson, 1995, Table 4.1.3.)

	Physical Functioning	Role Physical	Bodily Pain	General Health	Vitality	Social Functioning	Role-Emotional	Mental Health
<i>Unstandardised values</i>								
Mean	83.6	79.7	76.9	71.5	63.9	84.6	83.7	75.7
Standard deviation	23.4	35.1	25.2	20.8	19.4	22.2	30.9	16.4
Standard error of mean	0.4	0.6	0.4	0.4	0.3	0.4	0.5	0.3
Observations contributing to mean	6823	6807	6803	6848	6805	6825	6804	6803
25 th percentile	75.0	75.0	62.0	57.0	50.0	75.0	66.7	68.0
50 th percentile	95.0	100.0	84.0	75.0	65.0	100.0	100.0	80.0
75 th percentile	100.0	100.0	100.0	87.0	80.0	100.0	100.0	88.0
Range	0-100	0-100	0-100	0-100	0-100	0-100	0-100	4-100
% Ceiling	37.7	66.8	38.7	6.9	1.9	53.3	71.9	4.0
% Floor	0.7	12.1	1.0	0.1	0.4	0.7	8.1	0.1
<i>Age standardised values</i>								
Mean	83.6	79.6	76.8	71.5	63.9	84.6	83.6	75.7
Standard error of mean	0.4	0.6	0.4	0.3	0.3	0.4	0.5	0.3
Upper 95% confidence interval for mean	84.3	80.7	77.7	72.2	64.5	85.3	84.7	76.3
Lower 95% confidence interval for mean	82.9	78.5	76.0	70.9	63.2	83.8	82.6	75.1

Table 1.1(b): U.S. General Population Norms (N=2474) for the Physical and Mental Component Summary Scores (SF-36) (Source: Ware et al., 1994, Table 8.1.)

	<i>Physical Component Score</i>	<i>Mental Component Score</i>
Mean	50.0	50.0
25 th Percentile	42.8	45.0
50 th Percentile	52.6	52.5
75 th Percentile	56.0	57.4
Standard Deviation	10.0	10.0
Range	8-7	10-7

**Table B1.2: U.S. General Population Norms (N=5089) for the Mental Health Inventory
(Source: Davies & Ware (1981, Table C.3))**

	Anxiety	Depression	Loss Behav/ Emotional Control	General Positive Affect	Emotional Ties	Life Satisfaction	Psych Distress	Psych. Well-Being	MHI Index (38-Item)
Mean	19.1	8.0	15.9	41.3	9.1	4.3	47.5	59.2	77.5
Standard Deviation	6.8	3.0	5.6	8.8	2.6	1.0	15.4	12.2	25.7
Number	5089	5089	5089	5089	5089	5089	5089	5089	5089

Table 1.3: Distribution of Genders Across Diagnostic Groups and Practice Settings

	<i>Gender</i>		Total
	Female	Male	
Diagnostic Groups			
Anxiety Disorders	25	14	39
Affective Disorders	50	28	78
Schizophrenia	23	43	66
Practice Settings			
Private Psychiatry	38	14	52
Public Psychiatry	24	44	68
General Practice	36	27	63
Total	98	85	183

Table B1.4: Summary of differences between those people who remained in the study and those who dropped out between Time 1 and Time 2

	<i>M1</i>	<i>M2</i>	<i>Effect</i>
Duration of Contact with Service Provider	4.7	3.3	$F(1,171)=27.4^{**}$
Hospital Admissions	1.3	1.7	$F(1,172)=11.1^{**}$
Days in Hospital	2.1	3.6	$F(1,172)=12.8^{**}$
Modified Clinical Global Impressions Scale	3.3	3.7	$F(1,173)=6.9^*$
Global Role Functioning Index (RFS)	18.9	16.2	$F(1,110)=6.3^*$

* $p < .05$, ** $p < .01$

Table B2.1: Consumers' Responses to the Utility Questionnaire Items

<i>Were the words used in the questions difficult to understand?</i>						
		None	Some	Most	All	Total
Measures	BASIS-32	95	23	2	1	121
	%	78.5	19.0	1.7	0.8	
	SF-36	103	15		1	119
	%	86.6	12.6		0.8	
MHI		105	13		3	121
	%	86.8	10.7		2.5	
Total		303	51	2	5	361
	%	83.9	14.1	0.6	1.4	
<i>Did the questions ask you things that are important to you?</i>						
		None	Some	Most	All	Total
Measures	BASIS-32	5	46	48	22	121
	%	4.1	38.0	39.7	18.2	
	SF-36	4	34	54	27	119
	%	3.4	28.6	45.4	22.7	
MHI		4	30	45	43	122
	%	3.3	24.6	36.9	35.2	
Total		13	110	147	92	362
	%	3.6	30.4	40.6	25.4	
<i>Did the questions ask you things that relate to how you feel?</i>						
		None	Some	Most	All	Total
Measures	BASIS-32	4	39	45	32	120
	%	3.3	32.5	37.5	26.7	
	SF-36	3	33	52	30	118
	%	2.5	28.0	44.1	25.4	
MHI			23	52	47	122
	%		18.9	42.6	38.5	
Total		7	95	149	109	360
	%	1.9	26.4	41.4	30.3	
<i>Is this questionnaire a useful way for your doctor or specialist to find out whether you are getting better?</i>						
		Not Useful at all	Slightly Useful	Reasonably Useful	Very Useful	Total
Measures	BASIS-32	7	16	45	52	120
	%	5.8	13.3	37.5	43.3	
	SF-36	6	18	52	42	118
	%	5.1	15.3	44.1	35.6	
MHI		1	15	47	58	121
	%	0.8	12.4	38.8	47.9	
Total		14	49	144	152	359
	%	3.9	13.6	40.1	42.3	

Table B2.2: Consumers' Preferences for Selected Measures

<i>Comparison of BASIS-32 and MHI</i>			
Better Measure	BASIS-32 44.4%	MHI 55.6%	Total (N) 45
<i>Comparison of BASIS-32 and SF-36</i>			
Better Measure	BASIS-32 58%	SF-36 42%	Total (N) 50
<i>Comparison of MHI and SF-36</i>			
Better Measure	MHI 71.4%	SF-36 28.6%	Total (N) 49

Table B2.3: Summary of Issues Raised in Comments to Selected Measures by Consumers

<i>Measures</i>	
<i>Positive</i>	<i>Negative</i>
<p>BASIS-32</p> <ul style="list-style-type: none"> • Format helpful. • One week is an appropriate time frame. • Comprehensive. • User friendly, i.e., short, simple and easily understood. • Format and content of this questionnaire allows for independent completion. • Relevant questionnaire on a personal level. • Useful in providing the service provider with information that may not be covered comprehensively during a clinical interview. • Can provide feedback on a personal level as to changes in your condition. 	<ul style="list-style-type: none"> • Some responses were too restrictive. • Some of the language was difficult to comprehend and/or open to individual alternative interpretations. • an interviewer would be necessary in order to assist the completion of the questionnaire. • Questions predictable; may effect validity of information obtained. • Not relevant in an inpatient environment. • Not perceived as relevant by some participants. • Will be responded to depending on wellness. • May have a negative impact on affect during the completion process.
<p>MHI</p> <ul style="list-style-type: none"> • The consistent format was considered helpful. • The questionnaire was considered to be comprehensive. • User friendly, i.e., simple to complete and easily understood. • Format and content allows for independent completion. • Relevant questionnaire on a personal level. • Useful in providing the service provider with necessary information that may not be covered comprehensively during a clinical interview. • Provide feedback on a personal level as to changes in condition. 	<ul style="list-style-type: none"> • Some of the responses were too restrictive. • Some of the language was considered inappropriate. • Content of the questions was often considered to be repetitive. • One month was perceived as a difficult time period to respond accurately to. • Content was not considered comprehensive enough. • Not perceived as relevant by some participants. • Some questions may have a negative impact on affect during the completion process.
<p>SF-36</p> <ul style="list-style-type: none"> • Format was considered easy to follow. • Questionnaire was considered to be comprehensive. • Questionnaire was considered to be user friendly, i.e. simple to complete and easily understood. • The inclusion of questions pertinent to both physical health and mental health was perceived positively by some participants. • Relevant questionnaire on a personal level. • Useful in providing the service providers with information that may not be covered comprehensively during a clinical interview. • Can provide feedback on a personal level as to changes in condition. 	<ul style="list-style-type: none"> • Format was considered confusing. • Some of the language was considered to be difficult to comprehend and/or open to individual interpretation. • An interviewer would be necessary to assist the completion of the questionnaire. • Questionnaire did not differentiate physical health from mental health comprehensively. • Questionnaire was considered to have a very negative focus. • Questionnaire was perceived to be too general to be used as a measure of outcome. • Not perceived as relevant by some participants. • Some open-ended questions would be useful.

Table B2.5: Service Providers' Responses to the Utility Questionnaire Items

		<i>Were the words used in the questions difficult to understand?</i>				
		None	Some	Most	All	Total (N)
Measures	HoNOS	71	34	12	2	119
	%	59.7	28.6	10.1	1.7	
	LSP	95	8	5	9	117
	%	81.2	6.8	4.3	7.7	
	RFS	78	21	5	6	110
	%	70.9	19.1	4.5	5.5	
Total		244	63	22	17	346
	%	70.5	18.2	6.4	4.9	

		<i>Were the questions relevant to measuring the treatment outcomes for your client?</i>				
		None	Some	Most	All	Total
Measures	HoNOS	6	61	38	14	119
	%	5.0	51.3	31.9	11.8	
	LSP	10	49	46	13	118
	%	8.5	41.5	39.0	11.0	
	RFS	3	58	29	18	108
	%	2.8	53.7	26.9	16.7	
Total		19	168	113	45	345
	%	5.5	48.7	32.8	13.0	

		<i>Were the questions relevant to measuring the treatment outcomes for people with this particular diagnosis (i.e., that of your client), in general?</i>				
		None	Some	Most	All	Total
Measures	HoNOS	4	61	39	15	119
	%	3.4	51.3	32.8	12.6	
	LSP	7	52	43	16	118
	%	5.9	44.1	36.4	13.6	
	RFS	2	57	34	16	109
	%	1.8	52.3	31.2	14.7	
Total		13	170	116	47	346
	%	3.8	49.1	33.5	13.6	

		<i>Were the questions relevant to measuring the treatment outcomes for all people with mental illness?</i>				
		None	Some	Most	All	Total
Measures	HoNOS	2	61	42	13	118
	%	1.7	51.7	35.6	11.0	
	LSP	4	60	42	12	118
	%	3.4	50.8	35.6	10.2	
	RFS	3	64	30	13	110
	%	2.7	58.2	27.3	11.8	
Total		9	185	114	38	346
	%	2.6	53.5	32.9	11.0	

Table 2.5 (cont.): Service Providers' Responses to the Utility Questionnaire Items

<i>Is this questionnaire an effective way of measuring treatment outcomes for your client?</i>						
		Not Effective at all	Slightly Effective	Reasonably Effective	Very Effective	Total
Measures	HoNOS	15	52	46	6	119
	%	12.6	43.7	38.7	5.0	
	LSP	11	47	54	6	118
	%	9.3	39.8	45.8	5.1	
	RFS	12	53	41	4	110
	%	10.9	48.2	37.3	3.6	
Total		38	152	141	16	347
	%	11.0	43.8	40.6	4.6	
<i>Is this questionnaire an effective way of measuring treatment outcomes for people with this particular diagnosis (i.e., that of your client), in general?</i>						
		Not Effective at all	Slightly Effective	Reasonably Effective	Very Effective	Total
Measures	HoNOS	13	50	54	2	119
	%	10.9	42.0	45.4	1.7	
	LSP	12	46	53	7	118
	%	10.2	39.0	44.9	5.9	
	RFS	2	59	46	2	109
	%	1.8	54.1	42.2	1.8	
Total		27	155	153	11	346
	%	7.8	44.8	44.2	3.2	
<i>Is this questionnaire an effective way of measuring treatment outcomes for all people with mental illness?</i>						
		Not Effective at all	Slightly Effective	Reasonably Effective	Very Effective	Total
Measures	HoNOS	11	54	53		118
	%	9.3	45.8	44.9		
	LSP	5	54	51	8	118
	%	4.2	45.8	43.2	6.8	
	RFS	5	60	44	1	110
	%	4.5	54.5	40.0	.9	
Total		21	168	148	9	346
	%	6.1	48.6	42.8	2.6	
<i>Is this questionnaire a useful way of measuring treatment outcomes for people all people with mental illness?</i>						
		Not Effective at all	Slightly Effective	Reasonably Effective	Very Effective	Total
Measures	HoNOS	9	54	52	4	119
	%	7.6	45.4	43.7	3.4	
	LSP	4	49	58	7	118
	%	3.4	41.5	49.2	5.9	
	RFS	6	55	44	5	110
	%	5.5	50.0	40.0	4.5	
Total		19	158	154	16	347
	%	5.5	45.5	44.4	4.6	

Table B2.6: Service Providers' Preferences for Selected Measures

<i>Comparison of HoNOS and LSP.</i>			
Better Measure	HoNOS 50.0%	LSP 50.0%	Total (N) 54
<i>Comparison of HoNOS and RFS.</i>			
Better Measure	HoNOS 52.1%	RFS 47.9%	Total (N) 48
<i>Comparison of LSP and RFS.</i>			
Better Measure	LSP 69.4%	RFS 30.6%	Total (N) 49

Table B2.7: Summary of Issues Raised in Comments to Selected Service Provider Measures by Service Providers

		<i>Measures</i>	
		<i>Positive</i>	<i>Negative</i>
HoNOS	<ul style="list-style-type: none"> • Considered to be a comprehensive measure. • Considered to be user friendly by some participants. • Questions were considered to be framed positively. 	<ul style="list-style-type: none"> • Responses were considered too restrictive. • Format hinders completion. • Time-frame for this questionnaire was considered too restrictive. • Language considered difficult and confusing at times. • Too global and therefore insensitive. • Failed to address the issue of comorbidity. • Questions were perceived to be framed negatively. • Not considered relevant by some service providers. 	
LSP	<ul style="list-style-type: none"> • The consistent format was considered helpful. • Three months is a more appropriate time frame to consider when completing a measure. • The questionnaire was considered to be very comprehensive. • The questionnaire was considered to be very user friendly, i.e., simple and easily understood. • The questions were perceived to be framed in a positive manner. • Provides an accurate picture of the consumer's level of functioning. • If additions made to accommodate for country or interstate needs, this would increase the usefulness of this measure. 	<ul style="list-style-type: none"> • Some of the available responses were considered restrictive. • Difficult to respond to some of the questions accurately, may effect the validity of the information obtained. • Some of the language is ambiguous. • The questionnaire was considered too global and insensitive. • Only perceived as an effective measure for lower functioning population. • Does not address the issue of comorbidity. • The questions were viewed to be framed in a negative manner. • Not perceived as relevant by some service providers. • Not useful in a community setting. 	
RFS	<ul style="list-style-type: none"> • Considered to be a comprehensive measure. • User friendly measure, i.e., is short, simple and easily understood. 	<ul style="list-style-type: none"> • Available responses were considered restrictive. • Difficult to respond to some of the questions accurately, which would therefore decrease the validity of the information obtained. • Time-frame was considered inappropriate by some Service Providers. • Wording of some of the questions was perceived to be difficult and/or ambiguous. • Too global a measure, and therefore not particularly sensitive. • Not a particularly relevant measure. • Not considered to be useful in a community setting. • Ignored relevant issues such as culture and ethnicity. 	

Table B3.1: Correlations between the Composite Scores for the Selected Consumer Measures

	BASIS-32 Average Score	MHI Index Score	SF-36 Scores	
			Mental Component Summary Score	Physical Component Summary Score
BASIS-32 Average Score	---	-0.81**	-0.79**	-0.22
MHI Index Score			0.86**	0.53**
SF-36 Scores				
Mental Component Score			---	0.20*
Physical Component Score				---

** $p < .01$

Table B3.2: Correlations between the Scores of the BASIS-32 and MHI

MHI Scores	BASIS-32 Scores					BASIS-32 Average
	Relation to Self/ Others	Depression /Anxiety	Daily Living/ Role Functioning	Impulsive/ Addictive Behaviour	Psychosis	
Anxiety	0.55**	0.67**	0.68**	0.43**	0.39**	0.66**
Depression	0.64**	0.73**	0.73**	0.44**	0.44**	0.72**
Loss Behavioural/ Emotional Control	0.76**	0.76**	0.77**	0.52**	0.51**	0.80**
General Positive Affect	-0.74**	-0.72**	-0.76**	-0.38**	-0.41**	-0.75**
Emotional Ties	-0.54**	-0.37**	-0.33**	-0.31*	-0.32*	-0.44**
Psychological Well-Being	-0.77**	-0.72**	-0.73**	-0.39**	-0.42**	-0.75**
Psychological Distress	0.70**	0.77**	0.78**	0.49**	0.48**	0.78**
MHI Index	-0.76**	-0.80**	-0.81**	-0.48**	-0.48**	-0.81**

* $p < .05$, ** $p < .01$

Table B3.3: Correlations between the Scores of the BASIS-32 and SF-36

SF-36 Scores	BASIS-32 Scores					BASIS-32 Average
	Relation to Self/ Others	Depression /Anxiety	Daily Living/ Role Functioning	Impulsive/ Addictive Behaviour	Psychosis	
Physical Functioning	-0.07	-0.23	-0.11	-0.04	-0.05	-0.14
Role Physical	-0.24	-0.35**	-0.31*	-0.22	-0.16	-0.31*
Bodily Pain	-0.14	-0.33**	-0.12	-0.05	-0.01	-0.17
General Health	-0.51**	-0.59**	-0.52**	-0.34**	-0.37**	-0.57**
Role Emotional	-0.53**	-0.53**	-0.59**	-0.30*	-0.41**	-0.58**
Social Functioning	-0.67**	-0.73**	-0.74**	-0.33*	-0.46**	-0.74**
Vitality	-0.56**	-0.76**	-0.67**	-0.28*	-0.36**	-0.66**
Mental Health	-0.61**	-0.77**	-0.70**	-0.41**	-0.54**	-0.73**
Mental Component Score	-0.70**	-0.77**	-0.79**	-0.40**	-0.55**	-0.79**
Physical Component Score	-0.16	-0.35**	-0.18	-0.07	-0.03	-0.22

* $p < .05$, ** $p < .01$

Table B3.4: Correlations between the Scores of the MHI and SF-36

MHI Scores	SF-36 Scores									
	Physical Functioning	Role Physical	Bodily Pain	General Health	Role Emotional	Social Functioning	Vitality	Mental Health	Mental Component Score	Physical Component Score
Anxiety	-0.41**	-0.59**	-0.46**	-0.58**	-0.63**	-0.66**	-0.67**	-0.84**	-0.80**	-0.47**
Depression	-0.46**	-0.54**	-0.49**	-0.45**	-0.59**	-0.70**	-0.71**	-0.88**	-0.81**	-0.47**
Loss Behavioural/ Emotional Control	-0.45**	-0.61**	-0.46**	-0.62**	-0.58**	-0.73**	-0.70**	-0.82**	-0.79**	-0.53**
General Positive Affect	0.35**	0.45**	0.53**	0.50**	0.50	0.69**	0.79**	0.85**	0.80**	0.46**
Emotional Ties	0.27*	0.34**	0.27*	0.09	0.17	0.16	0.42**	0.37**	0.27*	0.31*
Psychological Well-Being	0.37**	0.50**	0.53**	0.49**	0.51**	0.68**	0.80**	0.86**	0.80**	0.48**
Psychological Distress	-0.46**	-0.62**	-0.48**	-0.59**	-0.64**	-0.73**	-0.72**	-0.88**	-0.84**	-0.52**
MHI Index	0.45**	0.60**	0.52**	0.58**	0.61**	0.74**	0.78**	0.91**	0.86**	0.53**

* $p < .05$, ** $p < .01$

Table B3.5: Correlations between the Composite Scores for the Selected Service Provider Measures

	HoNOS	LSP	Global RFI
HoNOS	-	-0.65**	-0.58**
LSP		-	0.78**
RFS			-

** $p < .01$

Table B3.6: Correlations between the Scores of the HoNOS and LSP

HoNOS Scores	LSP Scores					
	Communication	Non-turbulence	Responsibility	Self-care	Social Contact	Total
Aggression	-0.15*	-0.63**	-0.32*	-0.07	0.01	-0.34**
Self-harm	-0.18	-0.38**	-0.15	-0.20	-0.14	-0.28*
Substance Use	-0.23*	-0.52**	-0.37*	-0.46**	-0.26	-0.50**
Cognition	-0.45**	-0.22	-0.30*	-0.17**	-0.34**	-0.40**
Physical Health	-0.04	-0.1	-0.06	-0.29*	-0.23	-0.17
Hallucinations/ Delusions	-0.42**	-0.49**	-0.42**	-0.29*	-0.27*	-0.47**
Depression	0.15	-0.05	0.21	-0.22	-0.03	-0.12
Other Symptoms	-0.16	-0.14	-0.23	-0.17	-0.27*	-0.23
Social Relations	-0.42**	-0.26*	-0.15	-0.58**	-0.37**	-0.52*
General Functioning	-0.33**	-0.40**	-0.34**	-0.29*	-0.45**	-0.55**
Living Conditions	-0.40**	-0.24	-0.34**	-0.38*	-0.27*	-0.36**
Activities	-0.58**	-0.50**	-0.37**	-0.35**	-0.48*	-0.58**
Behaviour	-0.24*	-0.75**	-0.42**	-0.34**	-0.18	-0.54**
Impairment	-0.27*	-0.20	-0.21	-0.45**	-0.37**	-0.35**
Social Skills	-0.58**	-0.48**	-0.39*	-0.45**	-0.55**	-0.61**
Symptoms	-0.21	-0.34**	-0.21	-0.07	-0.31**	-0.30*
Total	-0.51**	-0.61**	-0.44**	-0.44**	-0.54**	-0.65**

* $p < .05$, ** $p < .01$

Table B3.7: Correlations between the Scores of the HoNOS and RFS

HoNOS Scores	RFS Scores				
	Working/ Productivity	Independent Living/Self Care	Immediate Social Network Relationships	Extended Social Network Relationships	Global Role Functioning Index
Aggression	-0.25	-0.10	-0.16	-0.36**	-0.25
Self-harm	0.07	0.05	-0.06	-0.06	-0.00
Substance Use	-0.39**	-0.23	-0.32*	-0.27*	-0.36**
Cognition	-0.38**	-0.53**	-0.30*	-0.37**	-0.48**
Physical Health	-0.09	-0.27*	-0.13	-0.02	-0.16
Hallucinations/Delusions	-0.40**	-0.32**	-0.22	-0.25	-0.36**
Depression	-0.04	-0.05	-0.09	-0.02	-0.06
Other Symptoms	-0.18	-0.01	-0.12	-0.14	-0.16
Social Relations	-0.37**	-0.22	-0.38**	-0.46**	-0.42**
General Functioning	-0.36**	-0.66**	-0.42**	-0.37**	-0.61**
Living Conditions	-0.30*	-0.36**	-0.24	-0.26*	-0.36**
Activities	-0.40**	-0.39*	-0.26*	-0.41	-0.43**
Behaviour	-0.31*	-0.15	-0.28*	-0.35**	-0.32*
Impairment	-0.28*	-0.49**	-0.26*	-0.23	-0.39**
Social Skills	-0.57**	-0.57**	-0.46**	-0.54**	-0.64**
Symptoms	-0.35*	-0.2*	-0.25	-0.23	-0.34*
Total	-0.52**	-0.51**	-0.4**	-0.47**	-0.58**

* $p < .05$, ** $p < .01$

Table B3.8: Correlations between the Scores of the LSP and RFS

RFS Scores	LSP Scores					Total
	Communication	Non-turbulence	Responsibility	Self-care	Social Contact	
Working/Productivity	0.51**	0.66**	0.59**	0.65**	0.55**	0.74**
Independent Living/ Self Care	0.44**	0.60**	0.59**	0.65**	0.50**	0.69**
Immediate Social						
Network Relationships	0.47**	0.59**	0.42**	0.38**	0.45**	0.58**
Extend Social						
Network Relationships	0.51**	0.57**	0.50**	0.47**	0.60**	0.65**
GRFI	0.56**	0.70**	0.62**	0.64**	0.61**	0.78**

* $p < .05$, ** $p < .01$

Table B3.9: Correlations between the Composite Scores of the Selected Consumer and Service Provider Measures

Selected Service	Selected Consumer Measures					
	BASIS-32		MHI		SF-36	
Provider Measures	Average Score	Index	Psychological Distress	Psychological Well-Being	Mental Component Score	Physical Component Score
HoNOS Total	0.31**	0.27*	0.28*	-0.21	-0.27*	-0.25*
LSP Total	0.13	0.07	-0.11	0.01	-0.07	0.14
RFS Total (GRFI)	-0.01	0.15	-0.17	0.10	0.03	0.13

* $p < .05$, ** $p < .01$

Table B3.10: Summary statistics for the Visual Analogue Scales

	<i>M</i>	<i>SD</i>	<i>Min</i>	<i>Max</i>	<i>N</i>
VAS Total	183.04	91.36	0.00	361.00	183
Well-Being	53.27	26.02	0.00	100.00	181
Depression	41.82	27.81	0.00	96.00	181
Anxiety	50.30	28.03	0.00	100.00	182
Distress	44.50	27.57	0.00	100.00	181

Table B3.11: Intercorrelations and test-retest reliabilities (in parentheses) for the Visual Analogue Scales

	VAS Total	Well-Being	Depression	Anxiety	Distress
VAS Total	(0.83)	-0.83**	0.83**	0.82**	0.86**
Well-Being		(0.82)	-0.62**	-0.56**	-0.67**
Depression			(0.74)	0.53**	0.59**
Anxiety				(0.71)	0.67**
Distress					(0.80)

** $p < .01$

Table B3.12: Intercorrelations between the Visual Analogue Scales (VAS) and the Composite Scores of the Selected Consumer Measures

VAS Scores	Selected Consumer Measures					
	BASIS-32		MHI		SF-36	
	Average Score	Index	Psychological Distress	Psychological Well-Being	Mental Component Score	Physical Component Score
VAS Total	0.66**	-0.68**	0.68**	-0.69**	-0.72**	-0.25
Well-Being	-0.58**	0.63**	-0.62**	0.56**	0.65**	0.23
Depression	0.64**	-0.64**	0.61**	-0.58**	-0.61**	0.19
Anxiety	0.48**	-0.49**	0.51**	-0.39**	-0.54**	-0.26
Distress	0.56**	-0.51**	0.53**	-0.42**	-0.66**	-0.17

** $p < .01$

Table B3.13: Frequencies and Summary Statistics for the Modified Clinical Global Impressions Scale

	Frequency	Valid %	Cumulative %		
Normal	13	7.4	7.1		
Borderline	26	14.9	22.3		
Mild	37	21.1	43.4		
Moderate	72	41.1	84.6		
Marked	21	12.0	96.6		
Severe	6	3.4	100.0		
Extreme	-	-			
Total	175	100			
	<i>M</i>	<i>SD</i>	<i>N</i>	<i>Min.</i>	<i>Max.</i>
Modified Clinical Global Impressions Scores	3.5	1.2	175	1	6

Table B3.14: Modified CGI Scores across Diagnostic Groups and Practice Settings

	<i>M</i>	<i>SD</i>	<i>N</i>	<i>F</i>
Total	3.5	1.2	174	
<i>Diagnostic Groups</i>				
Anxiety	3.6	1.1	37	3.4*
Affective	3.2	1.3	75	
Schizophrenia	3.7	1.1	62	
<i>Practice Settings</i>				
Private	4.0	1.0	47	8.0**
Public	3.4	1.2	65	
GP	3.1	1.2	62	

** $p < .01$

Table 3.15(a): Significant Differences across Diagnostic Groups for the Scores of the Selected Consumer Measures

Measure	Group Means			Effect
	Anxiety Disorders	Affective Disorders	Schizophrenia	
<i>BASIS-32</i>				
Average	1.1	1.5 ^a	1.0	$F(2,118)=5.8^{**}$
Relation to Self/Others	1.3	2.0 ^b	1.2	$F(2,118)=7.4^{**}$
Depression/Anxiety	1.5	2.0 ^a	1.1	$F(2,118)=9.1^{**}$
Daily Living/Role Functioning	1.4	2.0 ^b	1.4	$F(2,118)=6.0^{**}$
<i>MHI</i>				
Mental Health Index	141.6	128.2 ^a	159.1	$F(2,119)=8.4^{**}$
Psychological Distress	73.0	77.5 ^a	60.0	$F(2,119)=6.1^{**}$
Psychological Well-Being	48.6	39.5 ^b	52.8	$F(2,119)=11.3^{**}$
Anxiety	31.0	29.4	23.2 ^c	$F(2,119)=6.7^{**}$
Depression	11.8	13.4 ^a	9.7	$F(2,119)=8.1^{**}$
General Positive Affect	33.9	27.9 ^a	37.9	$F(2,119)=11.3^{**}$
Emotional Ties	8.0	6.0 ^d	6.8	$F(2,119)=3.6^*$
Loss of Emotional/ Behavioural Control	23.5	27.5 ^a	20.9	$F(2,119)=5.5^{**}$
<i>SF-36</i>				
Mental Component Score	30.4	31.7	39.5 ^c	$F(2,117)=4.5^*$
Bodily Pain ⁺	57.1	52.2 ^a	69.3	$F(2,116)=3.6^*$
Social Functioning ⁺	52.3	46.5 ^a	66.7	$F(2,116)=4.7^{**}$
Mental Health ⁺	48.4	50.2 ^e	61.7	$F(2,116)=3.5^*$

⁺Omnibus *F*-test n.s.

* $p < .05$, ** $p < .01$

- ^a Significant difference between the mean scores for the Affective Disorders Group and the Schizophrenia Group on this variable.
- ^b Significant difference between the mean score for the Affective Disorders Group and the mean scores of the Schizophrenia and Anxiety Disorders Groups, respectively, on this variable.
- ^c Significant difference between the mean score for the Schizophrenia Group and the mean scores for the Affective and Anxiety Disorders Group, respectively, on this variable.
- ^d Significant difference between the mean scores for the Affective Disorders Group and the Anxiety Disorders Group on this variable.
- ^e Post-hoc analyses revealed no significant differences between group means.

Table 3.15(b): Significant Differences across Practice Settings for the Scores of the Selected Consumer Measures

Measure	Group Means			Effect
	Private Psychiatry	Public Psychiatry	General Practice	
<i>BASIS-32</i>				
Average	1.6 ^a	1.0	1.2	$F(2,118)=5.3^{**}$
Relation to Self/Others	2.0 ^a	1.2	1.5	$F(2,118)=7.1^{**}$
Depression/Anxiety	2.1 ^a	1.1	1.6	$F(2,118)=9.2^{**}$
Daily Living/Role Functioning	2.0 ^a	1.4	1.6	$F(2,118)=4.4^{**}$
<i>MHI</i>				
Mental Health Index	122.0 ^b	157.9	143.4	$F(2,119)=9.8^{**}$
Psychological Distress	82.8 ^b	60.9	68.8	$F(2,119)=8.3^{**}$
Psychological Well-Being	38.7 ^a	52.2	46.2	$F(2,119)=8.9^{**}$
Anxiety	32.2 ^a	23.6	27.3	$F(2,119)=8.1^{**}$
Depression	13.8 ^a	9.9	11.8	$F(2,119)=7.7^{**}$
General Positive Affect	27.1 ^a	37.7	32.3	$F(2,119)=10.3^{**}$
Loss of Emotional/ Behavioural Control	29.2 ^b	21.3	23.2	$F(2,119)=7.2^{**}$
<i>SF-36</i>				
Mental Component Score	25.6 ^a	40.4	33.8	$F(2,117)=10.3^{**}$
General Health	43.0	65.7 ^c	52.8	$F(2,116)=8.7^{**}$
Vitality	39.2 ^a	54.2	43.3	$F(2,116)=4.7^*$
Role Emotional	25.6 ^a	55.6	45.5	$F(2,116)=5.4^{**}$
Social Functioning	38.3	68.9 ^c	52.0	$F(2,116)=9.3^{**}$
Mental Health	41.6 ^a	62.5	53.4	$F(2,116)=7.3^{**}$

$p < .05$, $** p < .01$

- ^a Significant difference between the mean scores for the Private Psychiatry Group and the Public Psychiatry Group on this variable.
- ^b Significant difference between the mean score for the Private Psychiatry Group and the mean scores for the Public Psychiatry and General Practice Groups, respectively, on this variable.
- ^c Significant difference between the mean score for the Public Psychiatry Group and the mean scores for the Private Psychiatry and General Practice Groups, respectively, on this variable.

Table 3.15(c): Significant Differences across Diagnostic Groups for the Scores of the Selected Service Provider Measures

Measure	Group Means			Effect
	Anxiety Disorders	Affective Disorders	Schizophrenia	
<i>HoNOS</i>				
Hallucinations/ Delusions ⁺	44.5	46.0	85.6 ^b	$\chi^2(2)=53.5^{**}$
Depression ⁺	49.9	75.7 ^c	46.1	$\chi^2(2)=21.4^{**}$
Other Symptoms ⁺	75.2 ^d	58.6	53.6	$\chi^2(2)=6.6^*$
General Functioning ⁺	54.0	53.7	71.0 ^a	$\chi^2(2)=7.6^*$
Living Conditions ⁺	50.6	52.1	74.8 ^b	$\chi^2(2)=18.6^{**}$
Activities ⁺	52.8	54.0	71.3 ^b	$\chi^2(2)=9.7^{**}$
<i>LSP</i>				
Total	136.9	134.8	125.1 ^b	$F(2,115)=6.0^{**}$
Communication	21.8	22.1	20.1 ^b	$F(2,115)=7.7^{**}$
Self-Care	36.5	35.9	32.8 ^b	$F(2,115)=6.5^{**}$
Social Contact	17.2	16.9	14.8 ^b	$F(2,115)=4.9^*$
<i>RFS</i>				
GRFI	19.6	19.4	15.9 ^b	$F(2,107)=5.8^{**}$
Working/Productivity	4.7	4.4	3.2 ^b	$F(2,107)=8.7^{**}$
Independent Living Skills	5.8	5.4	4.3 ^b	$F(2,107)=7.2^{**}$

$p < .05$, $** p < .01$

- ^a Significant difference between the mean scores for the Affective Disorders Group and the Schizophrenia Group on this variable.
- ^b Significant difference between the mean score for the Schizophrenia Group and the mean scores of the Affective and Anxiety Disorders Group, respectively, on this variable.
- ^c Significant difference between the mean score for the Affective Disorders Group and the mean scores for the Schizophrenia and Anxiety Disorders Groups, respectively, on this variable.
- ^d Significant difference between the mean scores for the Schizophrenia Group and the Anxiety Disorders Group on this variable.
- ⁺ Means displayed are group mean rank data.
- ⁺⁺ Omnibus F-test n.s.

Table 3.15 (d): Significant Differences across Practice Settings for the Scores of the Selected Service Provider Measures

Measure	Group Mean			Effect
	Private Psychiatry	Public Psychiatry	General Practice	
<i>HoNOS</i>				
Self-harm +	68.1 ^a	55.6	59.7	$\chi^2(2)=6.4^*$
Hallucinations/ Delusions +	51.8	72.4 ^c	55.0	$\chi^2(2)=12.1^{**}$
Depression +	77.6 ^d	47.3	60.9	$\chi^2(2)=15.4^{**}$
<i>LSP</i>				
Total	139.3	125.4 ^c	134.7	$F(2,115)= 8.7^{**}$
Communication	22.3 ^a	20.6	21.7	$F(2,115)=4.9^{**}$
Self-Care	37.2 ^c	32.8	36.0	$F(2,115)=10.5^{**}$
Responsibility	18.7 ^a	16.8	17.8	$F(2,115)=5.8^{**}$
Social Contact	17.4 ^c	14.9	17.0	$F(2,115)=6.0^{**}$

* $p < .05$, ** $p < .01$

- ^a Significant difference between the mean scores for the Private Psychiatry Group and the Public Psychiatry Group on this variable.
- ^b Significant difference between the mean score for the Private Psychiatry Group and the mean scores for the Public Psychiatry and General Practice Groups, respectively, on this variable.
- ^c Significant difference between the mean score for the Public Psychiatry Group and the mean scores for the Private Psychiatry and General Practice Groups, respectively, on this variable.
- ^d Significant difference between the mean scores for each group.
- + Means displayed are group mean rank data.

Table B3.16a: Test-Retest Reliabilities and Internal Consistency Coefficient (Observed and Reported) for the BASIS-32

	Observed T-R Reliability	Reported T-R Reliability*	Observed Alpha	Reported Alpha*
Relation to Self and Others	0.90	0.80	0.90	0.76
Daily Living/ Role Functioning	0.85	0.81	0.87	0.79
Depression/ Anxiety	0.87	0.78	0.82	0.74
Impulsive and Addictive Behaviours	0.84	0.65	0.67	0.68
Psychosis	0.84	0.76	0.65	0.43
Total	0.90	0.85	0.94	0.89

*Source: Eisen et al. (1994); Sederer and Dickey (1996)

Table B3.16b: Test-Retest Reliabilities and Internal Consistency Coefficient (Observed and Reported) for the MHI

	Observed T-R Reliability	Reported T-R Reliability*	Observed Alpha	Reported Alpha*
Anxiety	0.94	0.63	0.92	0.90
Depression	0.92	0.56	0.89	0.86
Loss of Behavioural/ Emotional Control	0.94	0.58	0.91	0.83
General Positive Affect	0.93	0.62	0.93	0.92
Emotional Ties	0.90	0.59	0.80	0.81
Psychological Well-Being	0.94	0.63	0.93	0.92
Psychological Distress	0.96	0.62	0.96	0.94
Mental Health Index	0.97	0.64	0.97	0.96

*Source: Veit & Ware (1983, Table 6)

Table B3.16c: Test-Retest Reliabilities and Internal Consistency Coefficient (Observed and Reported) for the SF-36

	Observed T-R Reliabilities	Reported T-R Reliabilities*	Observed alpha	Reported alpha*
Physical Functioning	0.92	0.81-0.90	0.91	0.88-0.94
Role Physical	0.73	0.60-0.69	0.86	0.76-0.90
Bodily Pain	0.72	0.43-0.78	0.77	0.79-0.88
General Health	0.89	0.80-0.83	0.83	0.78-0.90
Role Emotional	0.59	0.60-0.63	0.77	0.80-0.96
Social Functioning	0.81	0.60	0.84	0.63-0.85
Vitality	0.89	0.68-0.80	0.84	0.62-0.96
Mental Health	0.90	0.75-0.80	0.89	0.67-0.95
Physical Component Score	0.90	0.89	0.77	0.92-0.94
Mental Component Score	0.88	0.80	0.84	0.87-0.89

*Source: Ware et al. (1993)

Table B3.17: Internal Consistency Coefficients (Observed and Reported) for the Selected Service Provider Measures

	Observed alpha	Reported alpha*
<i>HoNOS*</i>		
Behaviour	0.40	
Impairment	0.30	
Symptoms	0.06	
Social Skills	0.69	
Total	0.73	
<i>LSP**</i>		
Self-Care	0.85	0.88
Non-Turbulence	0.88	0.85
Social Contact	0.78	0.79
Communication	0.64	0.67
Responsibility	0.75	0.77
Total	0.94	0.93
<i>RFS***</i>		
Global Role Functioning Index	0.86	0.92

*No internal consistency coefficients reported at this stage.

** Source for reported alpha coefficients: Parker et al. (1991); Trauer et al. (1995).

*** Source for reported alpha coefficients: Goodman et al. (1993).

Table B3.18: Correlations between the Composite Scores of the Service Provider Measures and Consumer-Versions of Service Provider Measures

Consumer Version of Service Provider Measures	Service Provider Measures		
	HoNOS Total	LSP Total	Global RFI
<i>HoNOS</i>			
Behaviour	0.29*		
Impairment	0.25*		
Symptoms	0.20		
Social Skills	0.08		
Total	0.15		
<i>LSP</i>			
Communication		0.13	
Non-turbulence		0.16	
Responsibility		0.26*	
Self-Care		0.36**	
Social Contact		0.31*	
Total		0.25	
<i>RFS Total</i>			
Working/Productivity			0.22
Independent Living Skills			0.32*
Immediate Social Network			0.33**
Extended Social Network			0.24
GRFI			0.30*

* $p < .05$, ** $p < .01$

Table B3.19: Correlations between the Composite Scores for the Selected Consumer Measures and Consumer Versions of the Service Provider Measures

Consumer Version of Service Provider Measures	Consumer Summary Score					
	BASIS-32		MHI		SF-36	
	Average Score	Index	Psychological Distress	Psychological Well-Being	Mental Component Score	Physical Component Score
HoNOS Total	0.79**	-0.68**	0.61**	-0.69**	-0.56**	-0.50**
LSP Total	-0.71**	0.63**	-0.60**	0.65**	0.51**	0.23
RFS Total (GRFI)	-0.61**	0.65**	-0.65**	0.61**	0.60**	0.28

** $p < .01$

Table B4.1: Frequencies and Summary Statistics for Consumers' Global Change Ratings

	Frequency	Valid %	Cumulative %		
Change Ratings					
Extremely Better	12	10.9	10.9		
Much Better	36	32.7	43.6		
Slightly Better	18	16.4	60.0		
No Change	21	19.1	79.1		
Slightly Worse	13	11.8	90.9		
Much Worse	10	9.1	100		
Extremely Worse	-	-	-		
Total	110	100.0			
Change Groups					
Better Group	66	60.0	60		
No Change Group	21	19.1	79.1		
Worse Group	23	20.9	100		
Total	110	100.0			
	<i>M</i>	<i>SD</i>	<i>N</i>	<i>Min.</i>	<i>Max.</i>
Global Change Ratings	3.2	1.5	110	1	6

Table B4.2: Frequencies and Summary Statistics for Service Providers' Global Change Ratings

	Frequency	Valid %	Cumulative %		
Change Ratings					
Very Much Improved	10	10.3	10.3		
Much Improved	22	22.7	33.0		
Minimally Improved	33	34.0	67.0		
No Change	16	16.5	83.5		
Minimally Worse	10	10.3	93.8		
Much Worse	5	5.2	99.0		
Very Much Worse	1	1.0	100.0		
Total	97	100.0			
Change Groups					
Better Group	65	67.0	67.0		
No Change Group	16	16.5	83.5		
Worse Group	16	16.5	100.0		
Total	97	100.0			
	M	SD	N	Min.	Max.
Global Change Ratings	3.1	1.4	97	1	7

Table B4.3: Crosstabulation of Change Categories based Consumer and Service Provider responses to the Global Change Rating Scale

Change Categories (Service Provider)	Change Categories (Consumer)			
	Better	No Change	Worse	Total
Better	37	9	8	54
No Change	4	4	5	13
Worse	5	3	3	11
Total	46	16	16	78

Table B4.4: Summary of Significant Differences in Scores Across Time and Change Groups

	<i>Effect for Time</i>	<i>Interaction: Time x Change Category</i>
<i>BASIS-32</i>		
Average	$F(1,77)=8.9^{**}$	
Relation to Self/Others	$F(1,77)=9.5^{**}$	
Daily Living/Role Functioning	$F(1,77)=13.2^{**}$	
Depression/Anxiety	$F(1,77)=6.1^{**}$	^{ac} $F(2,74)=3.7^*$
Impulsive/Addictive Behaviour		
Psychosis		
<i>MHI</i>		
Anxiety		^{be} $F(2,62)=7.8^{**}$
Depression	$F(1,65)=5.0^*$	^{bf} $F(2,62)=6.9^{**}$
General Positive Affect		^b $F(2,62)=10.3^{**}$
Loss of Behavioural/Emotional Control		^{bf} $F(2,62)=6.4^{**}$
Psychological Distress	$F(1,64)=4.4^*$	^{bf} $F(2,61)=6.0^*$
Psychological Well-Being	$F(1,62)=5.4^*$	^a $F(2,59)=6.2^*$
MHI Index		^{bf} $F(2,61)=8.6^{**}$
<i>SF-36</i>		
Mental Health		^a $F(2,72)=4.8^*$
Vitality		^a $F(2,72)=4.8^*$
Mental Component Score		^a $F(2,72)=4.4^*$

* $p < .05$, ** $p < .01$

- ^a post-hoc analyses indicated a significant positive difference between scores at Time 1 and 2 for the *better* group.
- ^b post-hoc analyses indicated a significant positive difference between scores at Time 1 and 2 for the *better* group, and a significant negative difference between scores for the *worse* group.
- ^c post-hoc analyses indicated significant differences between the means of the *worse* group and those of the *better* and *no change* groups.
- ^d post-hoc analyses indicated significant differences between the means of the *worse* group and those of the *better* groups.
- ^e post-hoc analyses indicated significant differences between the means of the *worse* and *better* groups at Time 2.
- ^f post-hoc analyses indicated significant differences between the means of the *worse* group and the *better* and *no change* groups at Time 2.

Table B4.5a: Summary of Significant Differences in Scores Across Time and Change Groups

	Effect for Time	Interaction: Time x Change Category
<i>Life Skills Profile</i>		
Responsibility		^{bc} $F(2,59)=3.4^*$
Total		^{be} $F(2,59)=5.0^*$
<i>Role Functioning Scale</i>		
Working Productivity		^d $F(2,58)=3.7^*$
Independent Living Skills	$F(1,65)=9.1^{**}$	
Immediate Social Network Relationships	$F(1,65)=5.3^*$	
Global Role Functioning Index	$F(1,65)=7.1^{**}$	
<i>HoNOS</i>		
Impairment	$F(1,69)=15.4^{**}$	
Symptoms	$F(1,69)=8.2^{**}$	^a $F(2,61)=7.6^{**}$
Total	$F(1,69)=9.8^{**}$	^a $F(2,61)=6.9^{**}$

* $p<.05$, ** $p<.01$

- ^a post-hoc analyses indicated a significant positive difference between scores at Time 1 and 2 for the *better* group.
- ^b post-hoc analyses indicated significant negative difference between scores at Time 1 and 2 for the *worse* group.
- ^c post-hoc analyses indicated significant difference between the means of the *no change* and *worse* groups at Time 2.
- ^d post-hoc analyses indicated a significant difference between the means of the *worse* groups and the *better* and *no change* groups at Time 2; and a significant difference between the means of the *better* and *no change* groups.
- ^e post-hoc analyses indicated a significant difference between the means of the *better* and *no change* groups at Time 2.

Table B4.5b: Wilcoxon Signed Ranks Tests for Between Time 1 and 2 score for the 12 HoNOS Scale Items across the Whole Sample

	Sum of Ranks				Z
	Negative Ranks	Positive Ranks	Ties	Total	
Cognition	273.0	105.0	43	70	-2.2*
Physical Health	471.5	123.5	36	70	-3.1**
Depressed Mood	706.5	328.5	25	70	-2.3*
Other Symptoms	757.5	232.5	26	70	-3.2**
Relationships	879.0	346.0	21	70	-2.8**

* $p<.05$, ** $p<.01$

Table B4.6: Mean Scores for the Selected Consumer Measures for Time 1 and 2 Across Diagnostic Groups

	Diagnostic Groups						<i>Interaction Effect: Time x Diagnostic Groups</i>
	Anxiety Disorders		Affective Disorders		Schizophrenia		
	M1	M2	M1	M2	M1	M2	
<i>MHI</i>							
Depression	10.9	10.0	13.0	10.6	9.5	10.3	^{ab} $F(2,63)=4.5^*$
General Positive Affect	33.8	33.9	28.1	32.9	38.1	34.2	^{ab} $F(2,63)=7.4^{**}$
Loss of Behavioural Control							$F(2,63)=4.1^*$
Psychological Well-Being	49.6	49.6	39.8	45.9	51.9	50.1	^{ab} $F(2,63)=4.7^*$
Psychological Distress							$F(2,62)=3.5^*$
Mental Health Inventory Index Score	145.8	150.1	130.8	145.0	160.6	152.8	^{ab} $F(2,63)=4.6^*$
<i>SF-36</i>							
Social Functioning	52.6	59.2	47.0	63.6	69.5	55.0	^{ab} $F(2,74)=8.0^{**}$
Mental Health	53.1	52.8	49.1	63.0	60.5	56.6	^a $F(2,63)=6.3^{**}$
Mental Component Summary Score	31.8	34.0	31.0	39.6	37.8	34.8	^a $F(2,74)=4.9^*$

* $p<.05$, ** $p<.01$

^a post-hoc analyses indicated a significant difference between Time 1 and Time 2 mean scores for the Affective Disorders Group.

^b post-hoc analyses indicated significant differences between the means of the Schizophrenia and Affective disorders Groups at Time 1.

Table B4.7: Correlations between Summary Change Scores for Selected Consumer Measures

	BASIS-32	MHI	SF-36	
	Average Scores	Index	Mental Component Score (MCS)	Physical Component Score (PCS)
BASIS-32	-	-0.66**	-0.69**	-0.27
MHI Index		-	0.70**	0.17
<i>SF-36</i>				
Mental Component Score (MCS)			-	0.05
Physical Component Score (PCS)				-

* $p<.05$, ** $p<.01$

Table B4.8: Correlations between Composite Change Scores for Selected Service Provider Measures

	HoNOS	LSP	Global RFI
HoNOS	-	-0.53**	-0.12
LSP		-	0.58**
RFS			-

** $p < .01$

Table B4.9: Correlations between Composite Change Scores for Consumer and Service Provider Measures

Service Provider Measure Change Scores	Consumer Measure Change Scores		
	BASIS-32 Average Scores	MHI Index	SF-36 Mental Component Score
HoNOS	0.29	-0.15	-0.29
LSP	-0.29	0.04	0.44**
RFS	-0.12	-0.18	0.18

* $p < .05$, ** $p < .01$

Table B4.10: R and R² values for Change Scores of the Selected Consumer Measures and Consumers' Global Change Ratings

	<i>r</i>	<i>r</i> ²
Consumer Measures		
BASIS-32 Average Scores	0.26**	0.07
MHI Index Score	-0.46**	0.21
Mental Component Score	-0.34**	0.12
Physical Component Score	-0.04	0.002
Service Provider Measures		
HoNOS Total	0.44**	0.19
LSP Total	-0.45**	0.20
RFS Total	-0.17	0.03

* $p < .05$, ** $p < .01$

Table B4.11: Summary of Multiple Regression analyses of Change Scores for the Selected Measures on Global Change Ratings

Measure	Model Prediction	
	R^2	F
Consumer		
BASIS	0.12	$F(5,71)=1.9$
MHI		
5-Factors	0.28	$F(5,59)=4.5^{**}$
2-Factors	0.20	$F(2,61)=7.3^{**}$
SF-36		
8-Factors	0.16	$F(8,66)=1.6$
3-Factors	0.12	$F(3,71)=3.2^*$
Service Provider		
HoNOS	0.24	$F(4,59)=4.5^{**}$
LSP	0.25	$F(5,56)=3.7^{**}$
RFS	0.04	$F(4,55)<1$

* $p<.05$, ** $p<.01$

R^2 = proportion of variance in DV accounted for by some linear combination of the IVs.

F = test of significance of the R^2

Table B4.12: Internal Consistency Scores for Consumer Measures Using Change Scores

Measures	Cronbach's alpha (α)
<i>BASIS-32</i>	
Relation to Self/Others	0.80
Depression/Anxiety	0.77
Daily Living/role Functioning	0.81
Impulsive/Addictive Behaviour	0.62
Psychosis	0.44
Average	0.93
<i>MHI</i>	
Anxiety	0.78
Depression	0.76
Loss Behavioural/Emotional Control	0.80
General Positive Affect	0.84
Emotional Ties	0.43
Psychological Well-Being	0.87
Psychological Distress	0.91
MHI Index	0.93
<i>SF-36</i>	
Physical Functioning	0.74
Role-Physical	0.82
Bodily Pain	0.61
General Health	0.61
Vitality	0.67
Social Functioning	0.72
Role-Emotional	0.70
Mental Health	0.76
Mental Component Score	0.77
Physical Component Score	0.50

Table B4.13: Internal Consistency Scores for service provider Measures Using Change Scores

	Measures	Cronbach's alpha (α)
<i>HoNOS</i>	Behaviour	0.37
	Impairment	0.00
	Symptoms	0.47
	Social Skills	0.69
	Total	0.71
<i>LSP</i>	Self-Care	0.57
	Non-Turbulence	0.59
	Social Contact	0.66
	Communication	0.36
	Responsibility	0.63
	Total	0.84
<i>RFS</i>	GRFI	0.46

Table B5.1: Participants' Responses to Use of a Computer.

		Yes	No	N
1. Is it a good idea to use a computer to get information about the way people are feeling?	<i>Consumers</i>	83.5%	16.5%	79
	<i>Service Providers</i>	90.5%	9.5%	21
2. Did it feel comfortable to use a computer to answer these questions?	<i>Consumers</i>	93.8%	6.3%	80
	<i>Service Providers</i>	81.8%	18.2%	22

Appendix 4: Figures of Results

Figure 1: Distribution of Service Providers by Professional Groups

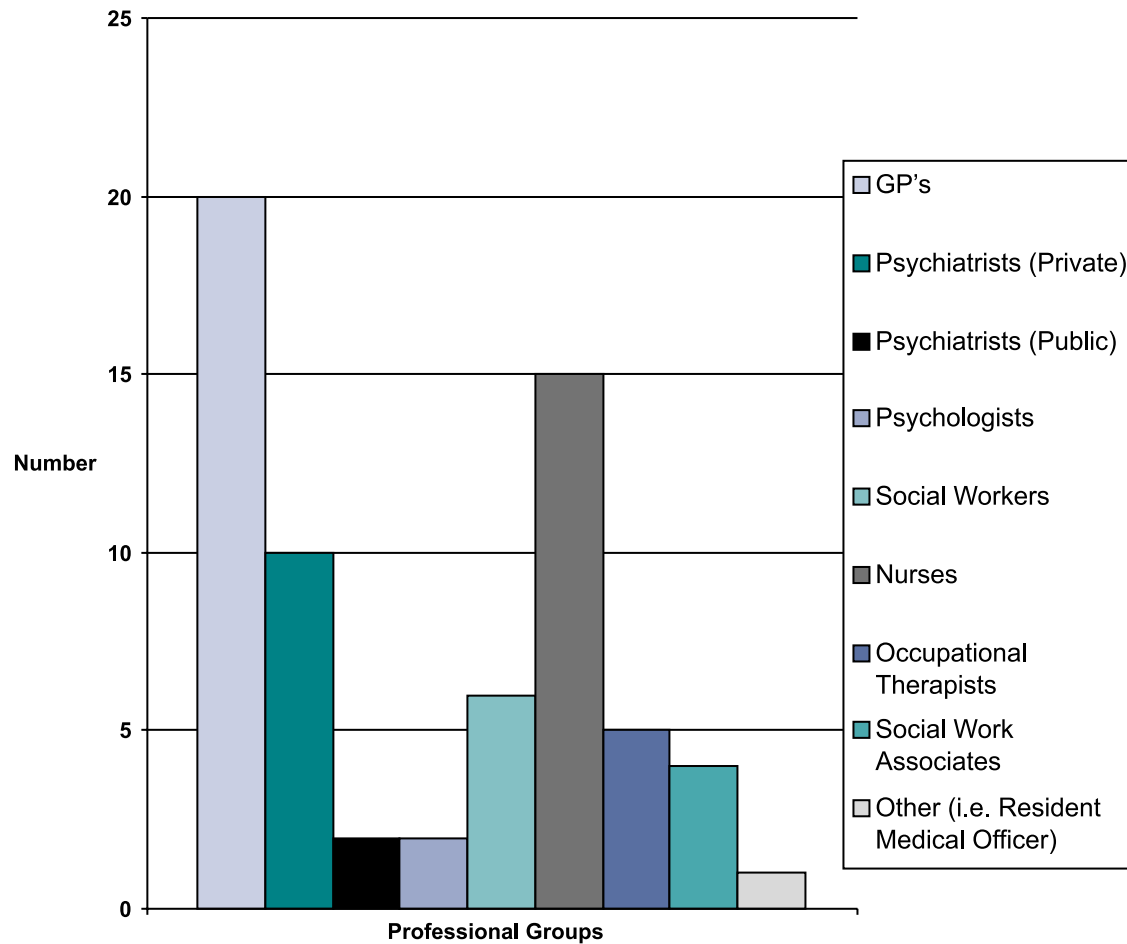


Figure 2: Distribution of Consumers By Diagnostic Groups Across Practice Settings

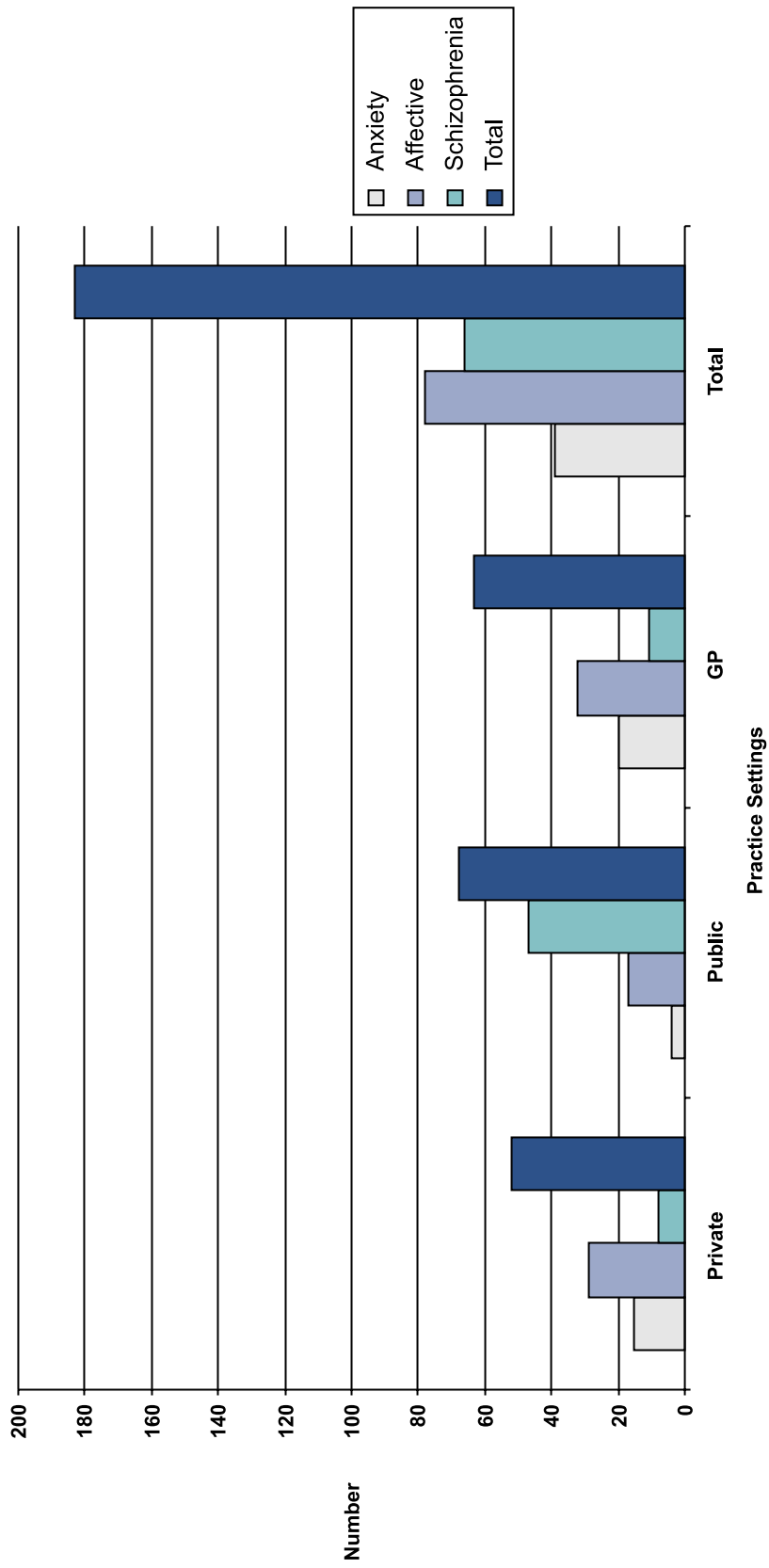


Figure 3: Frequency of Comorbidity Ratings Across Practices Setting

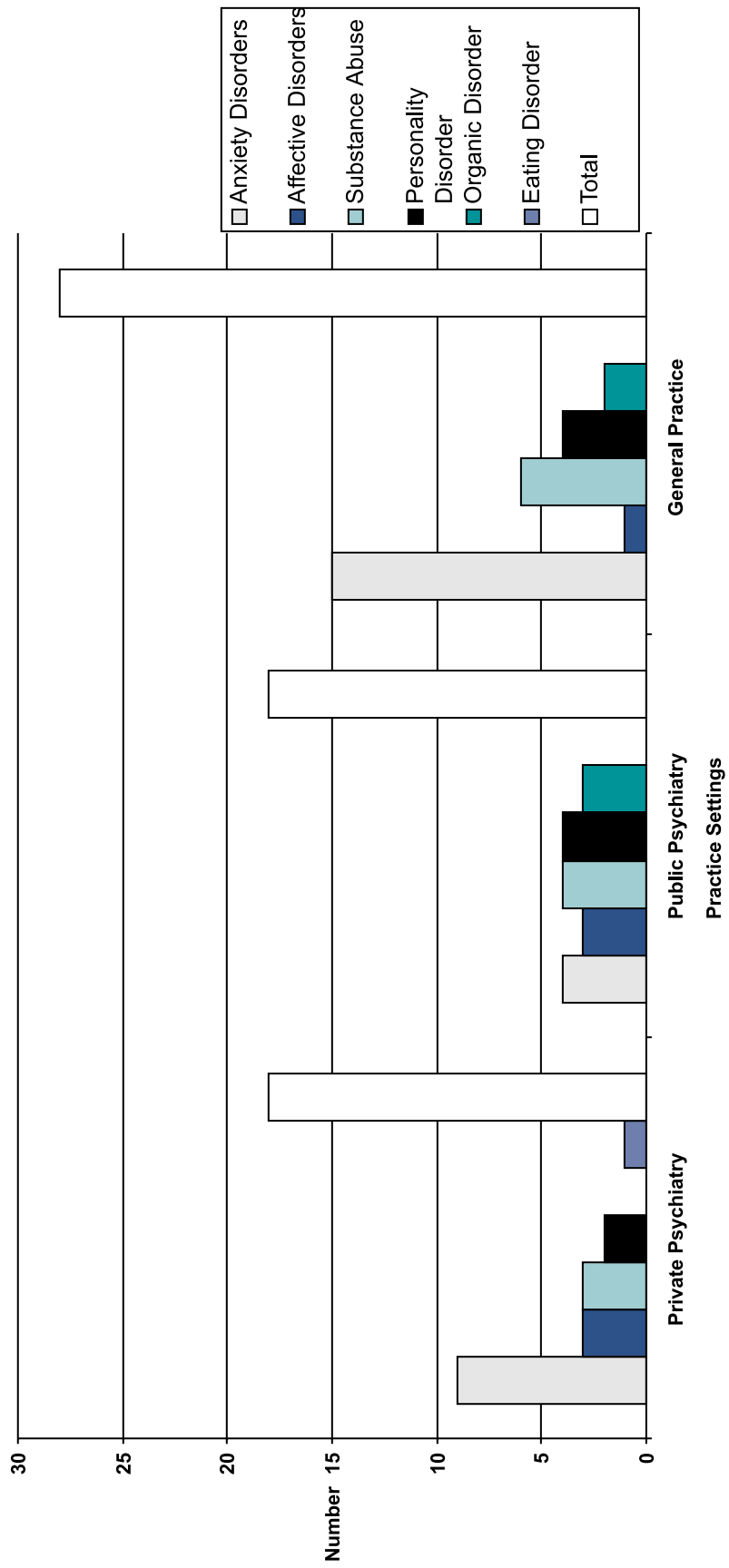


Figure 4: Mean General Utility Scores (Consumers) for the Selected Consumer Measures

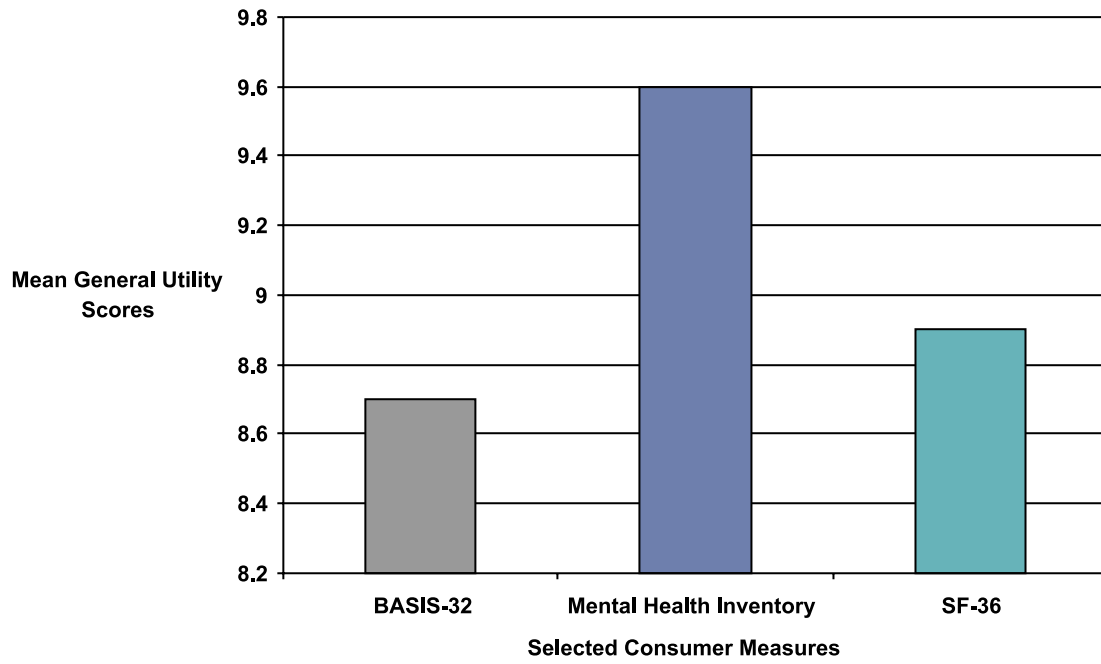


Figure 5: Mean General Utility Scores (Service Providers) for the Selected Service Provider Measures Across Practice Settings

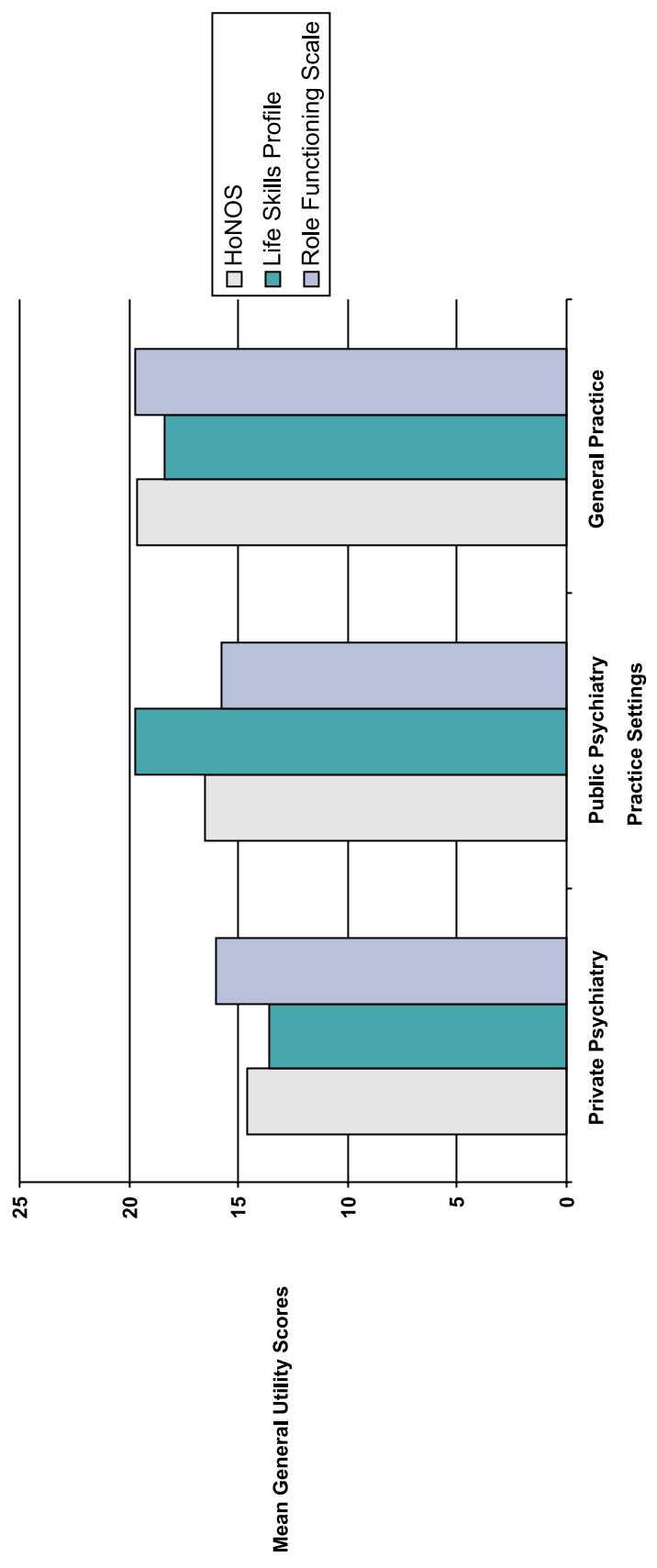


Figure 6: Mean Visual Analogue Scale Total Scores Across Diagnostic Groups

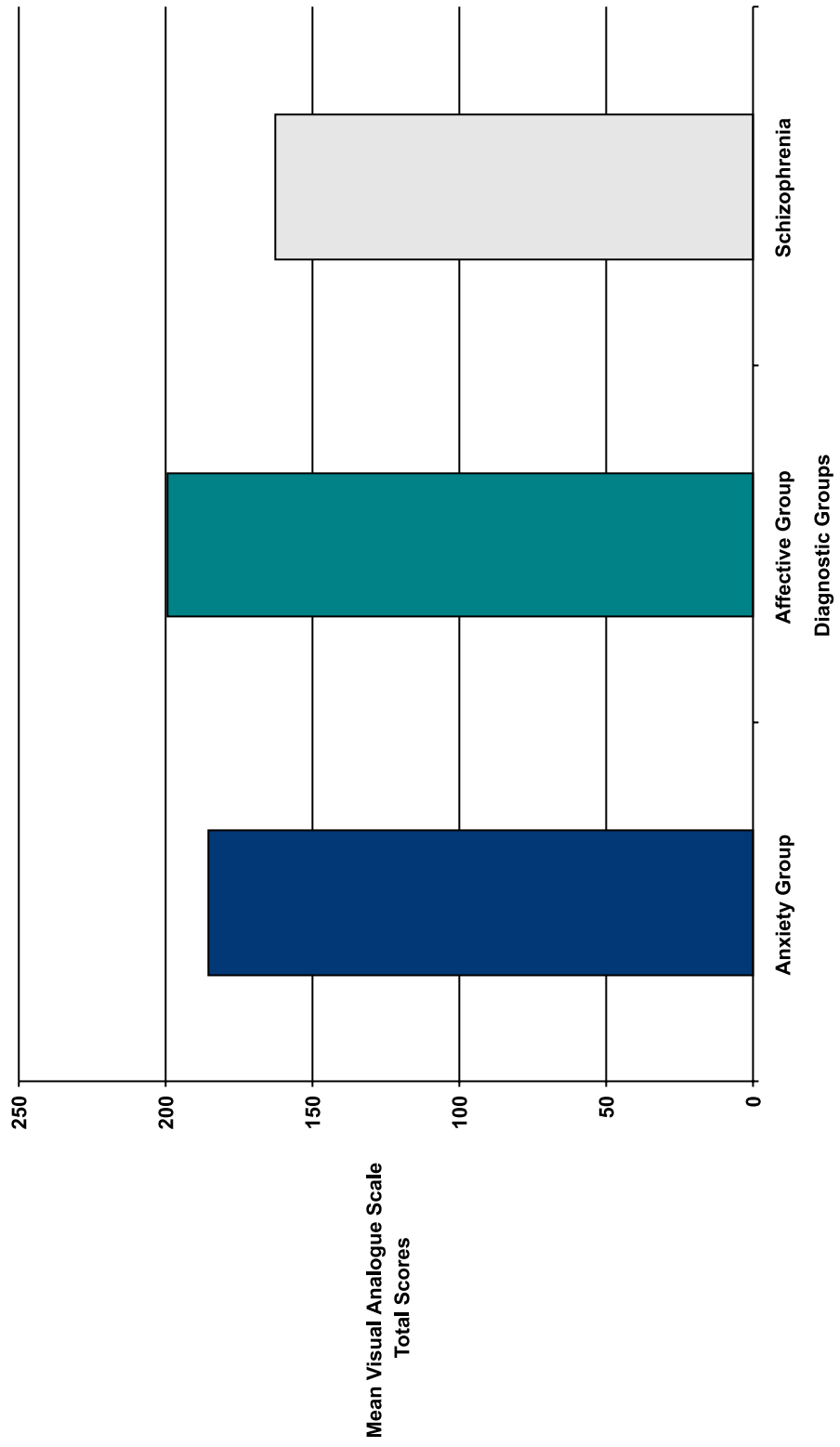


Figure 7: Mean Visual Analogue Scale Total Scores Across Practice Settings

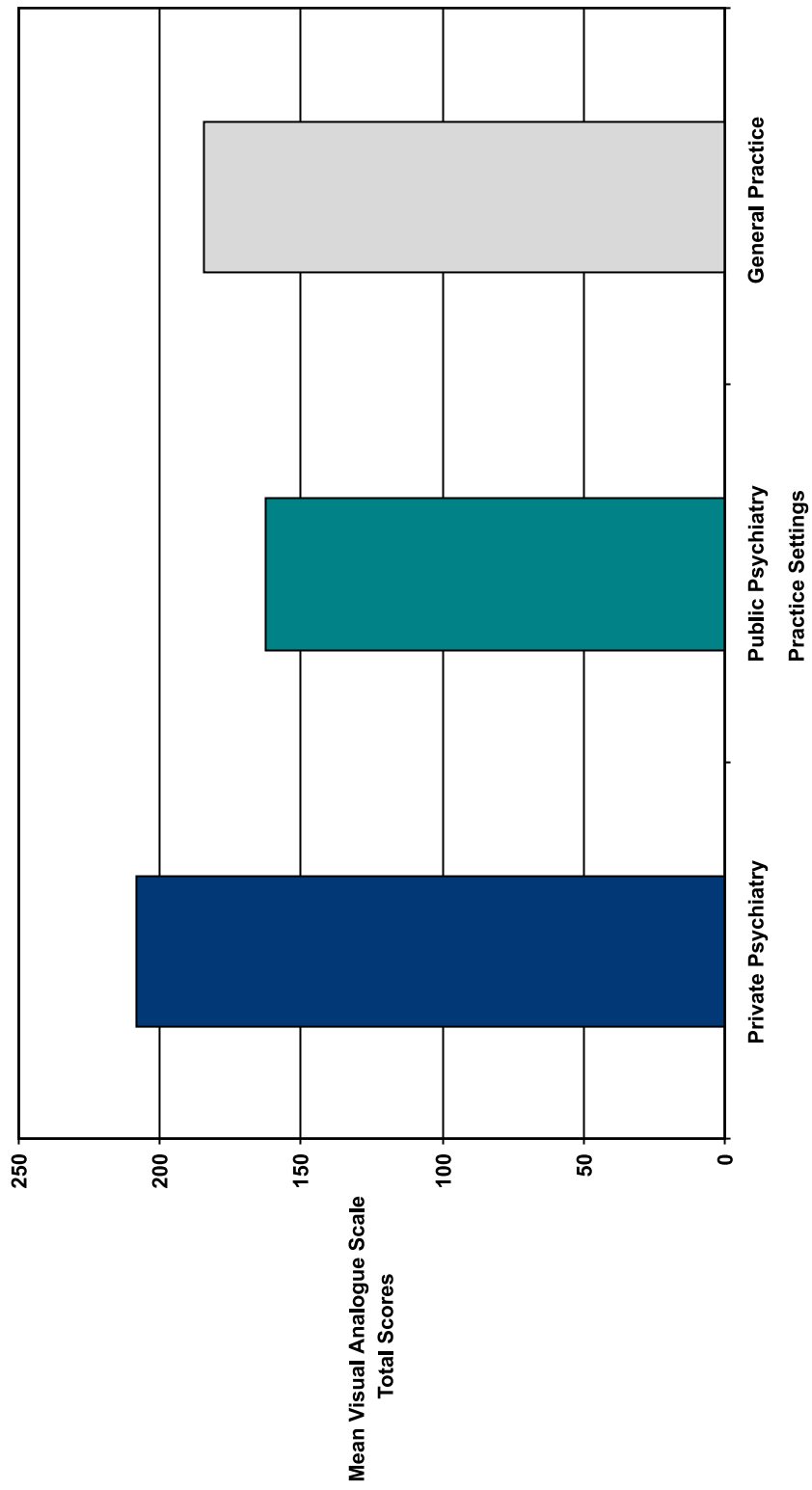


Figure 8: Mean Scores for the Modified Clinical Global Impressions Scale Across Diagnostic Groups

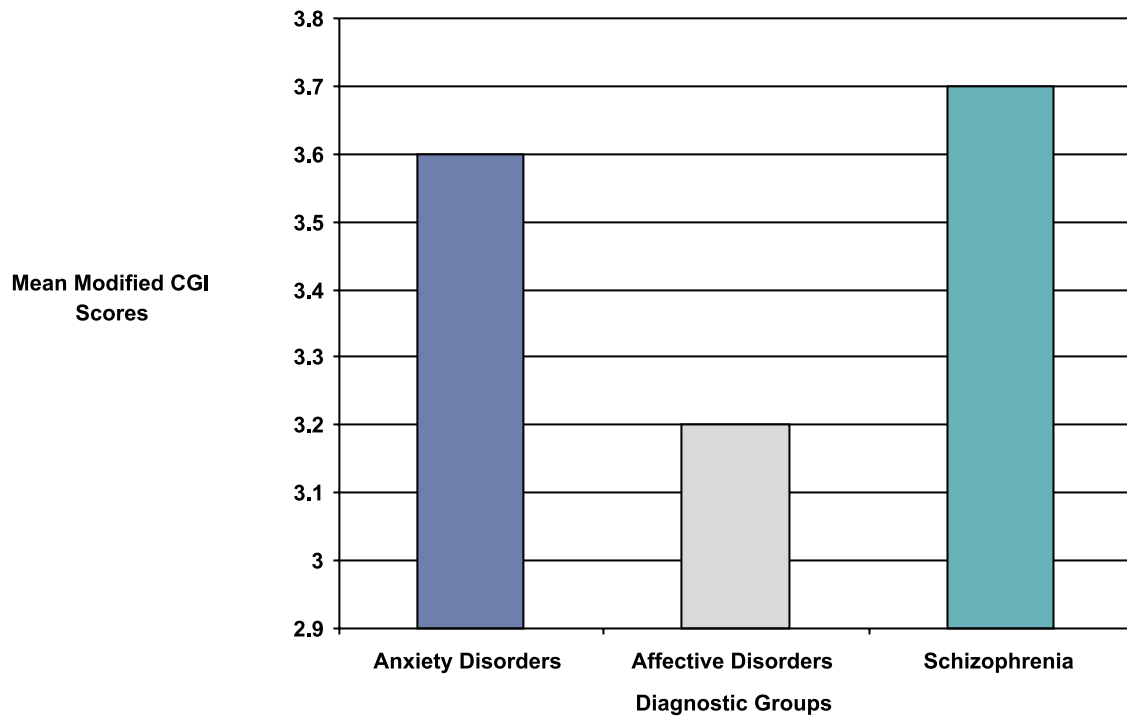


Figure 9: Mean Scores for the Modified Clinical Global Impressions Scale Across Practice Settings

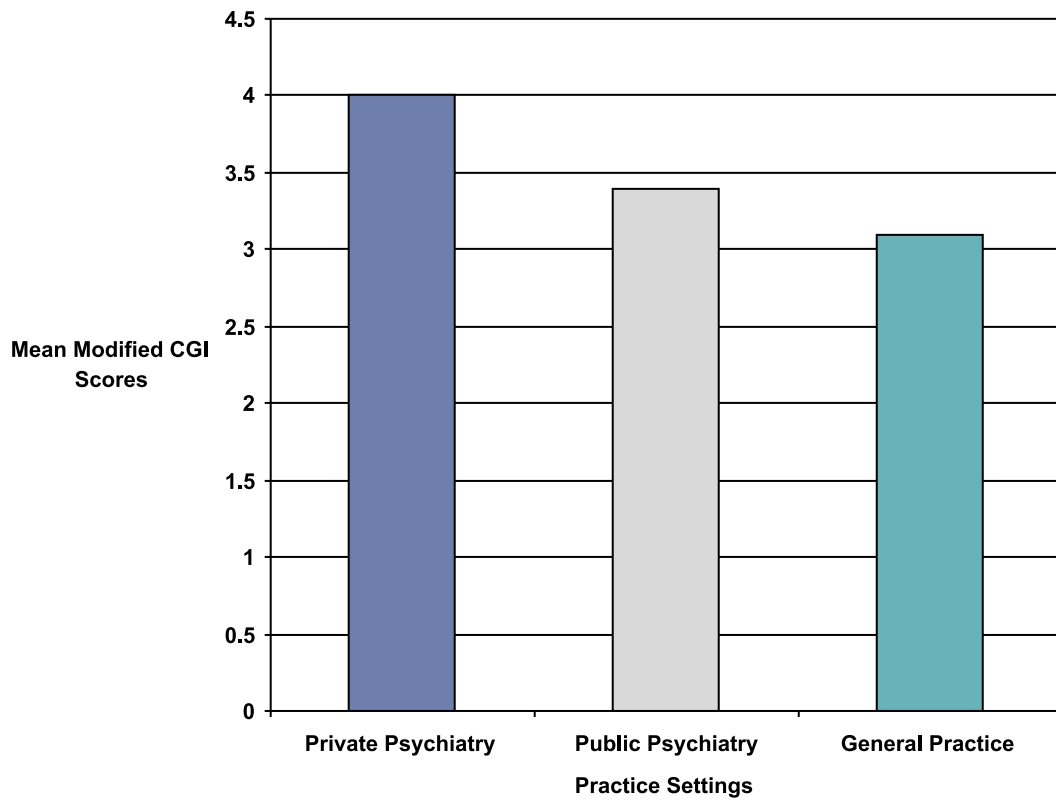


Figure 10: Mean Scores for BASIS-32 Across Diagnostic Groups

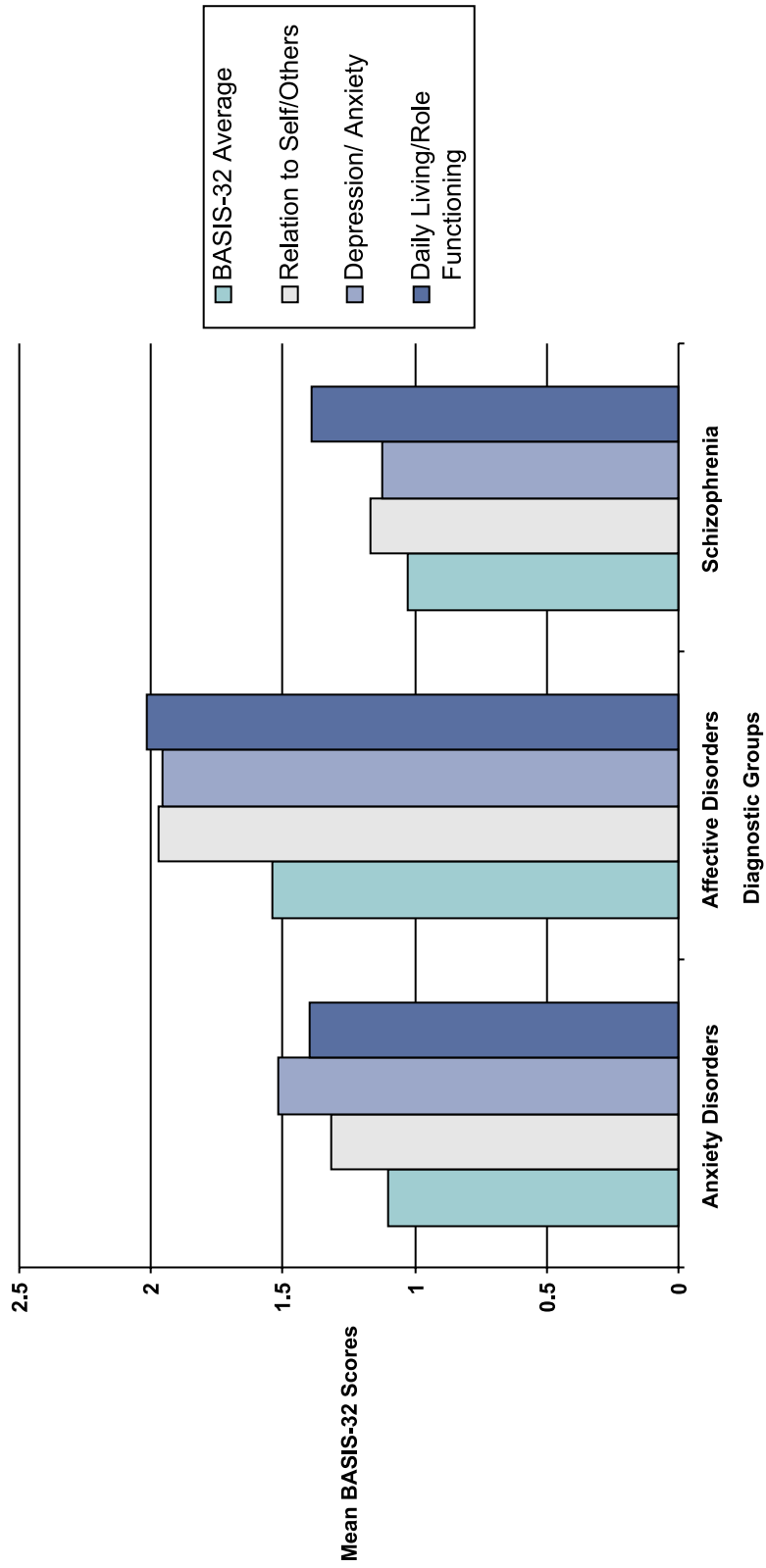


Figure 11a: Mean Scores for the Mental Health Inventory Across Diagnostic Groups

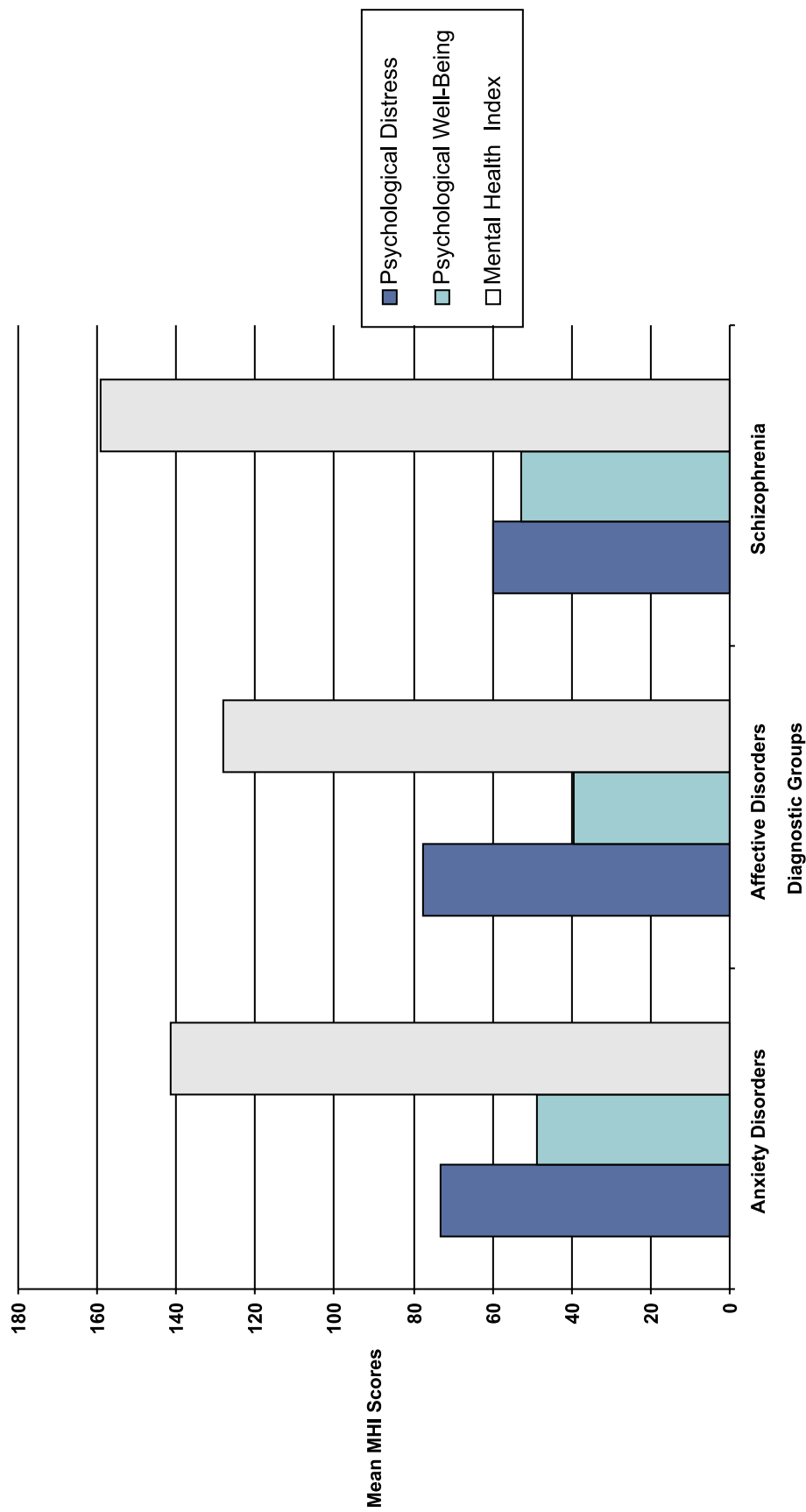


Figure 11b: Mean Scores for the Mental Health Inventory Across Diagnostic Groups

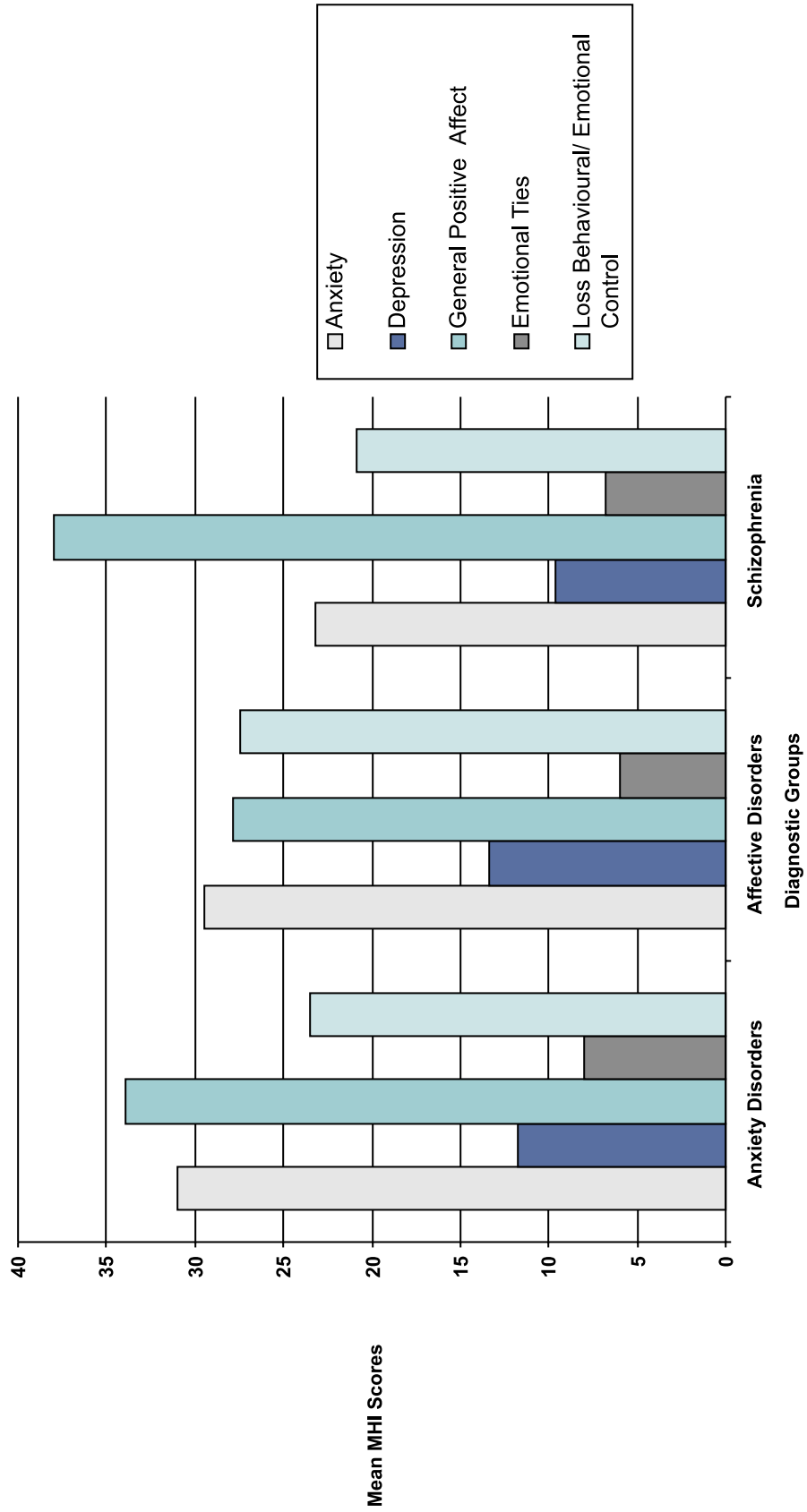


Figure 12: Mean Scores for the SF-36 Across Diagnostic Groups

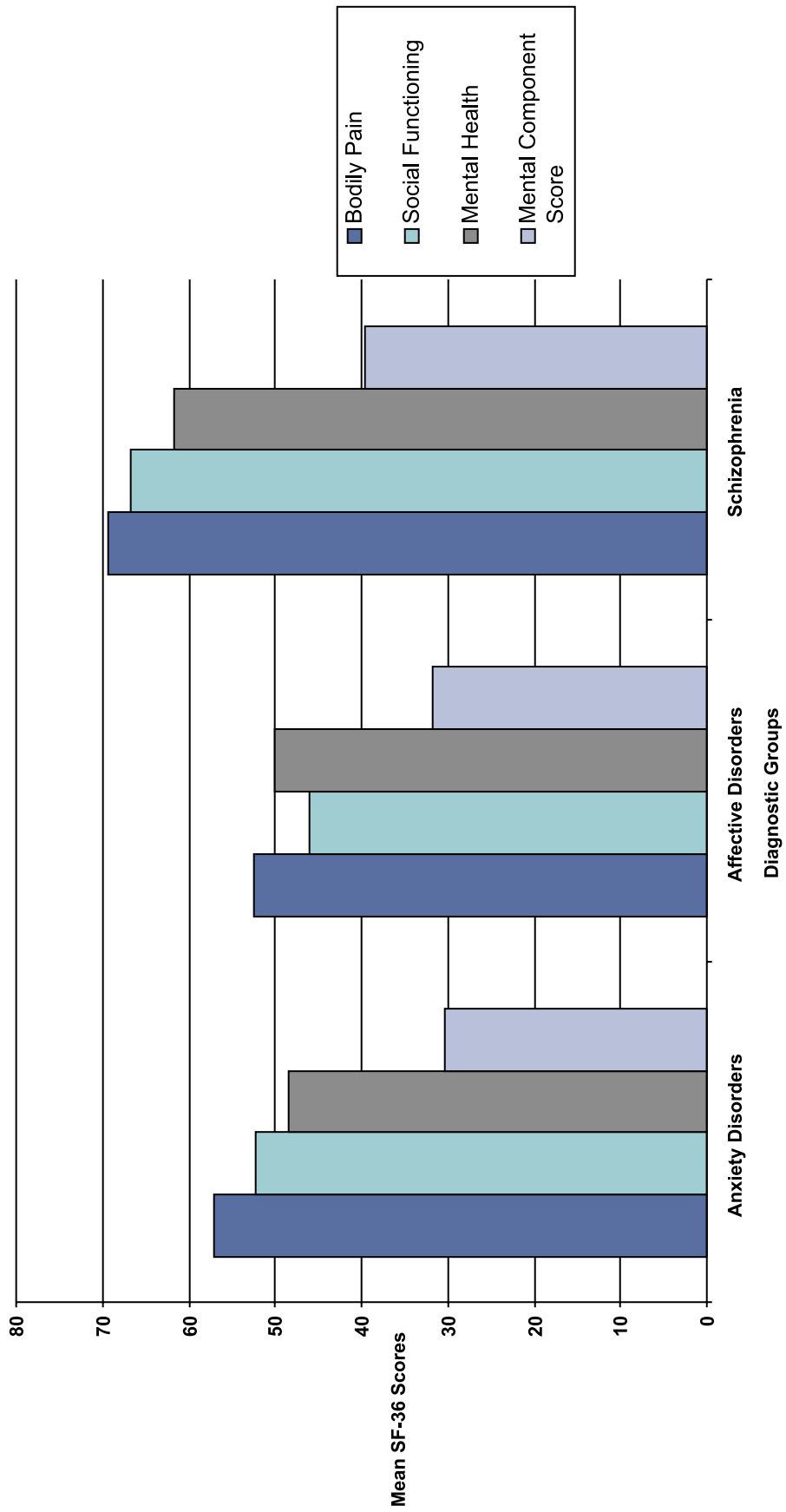


Figure 13: Mean Scores for the BASIS-32 Across Practice Settings

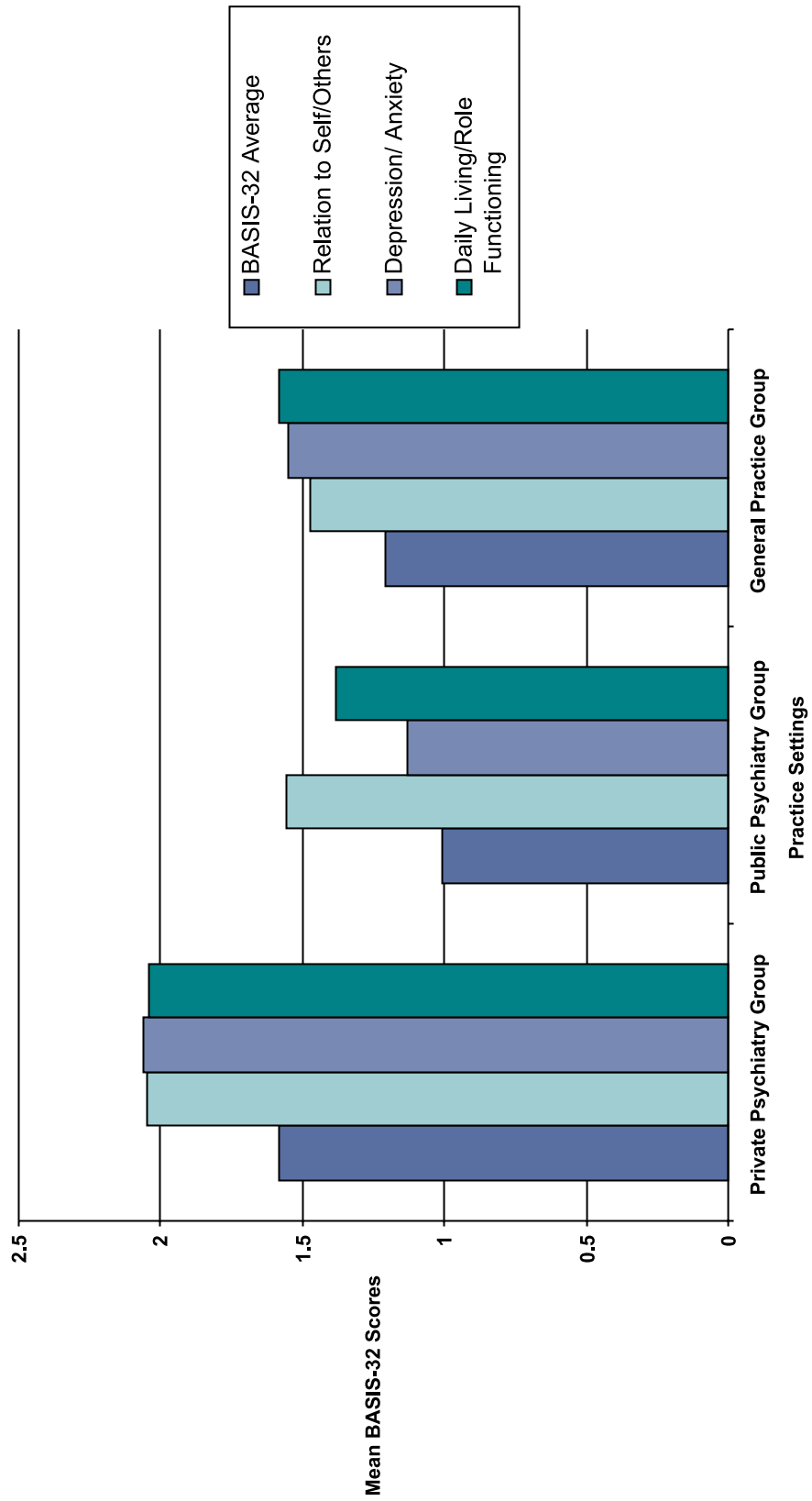


Figure 14a: Mean Scores for the Mental Health Inventory Across Practice Settings

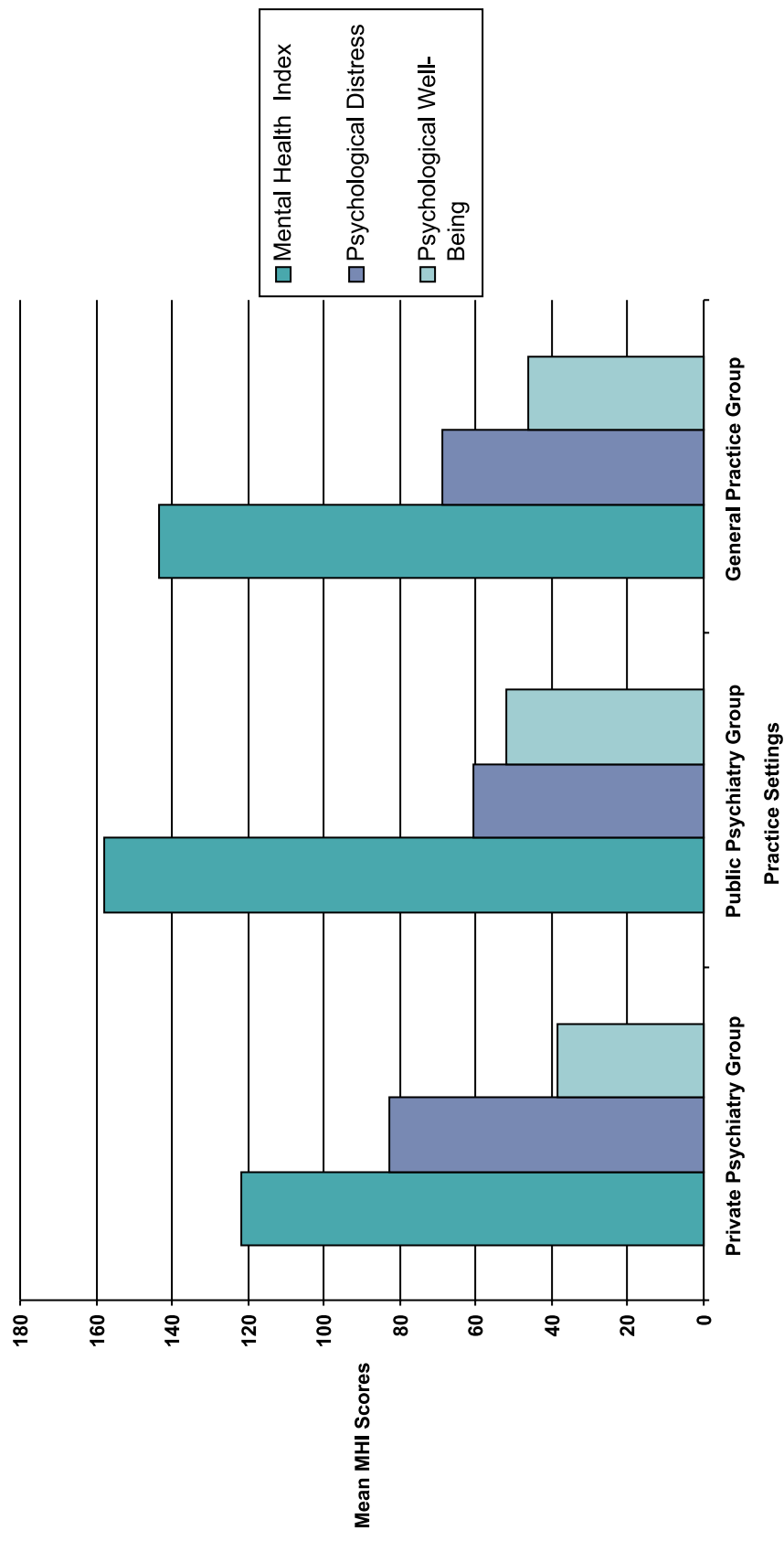


Figure 14b: Mean Scores for the Mental Health Inventory Across Practice Settings

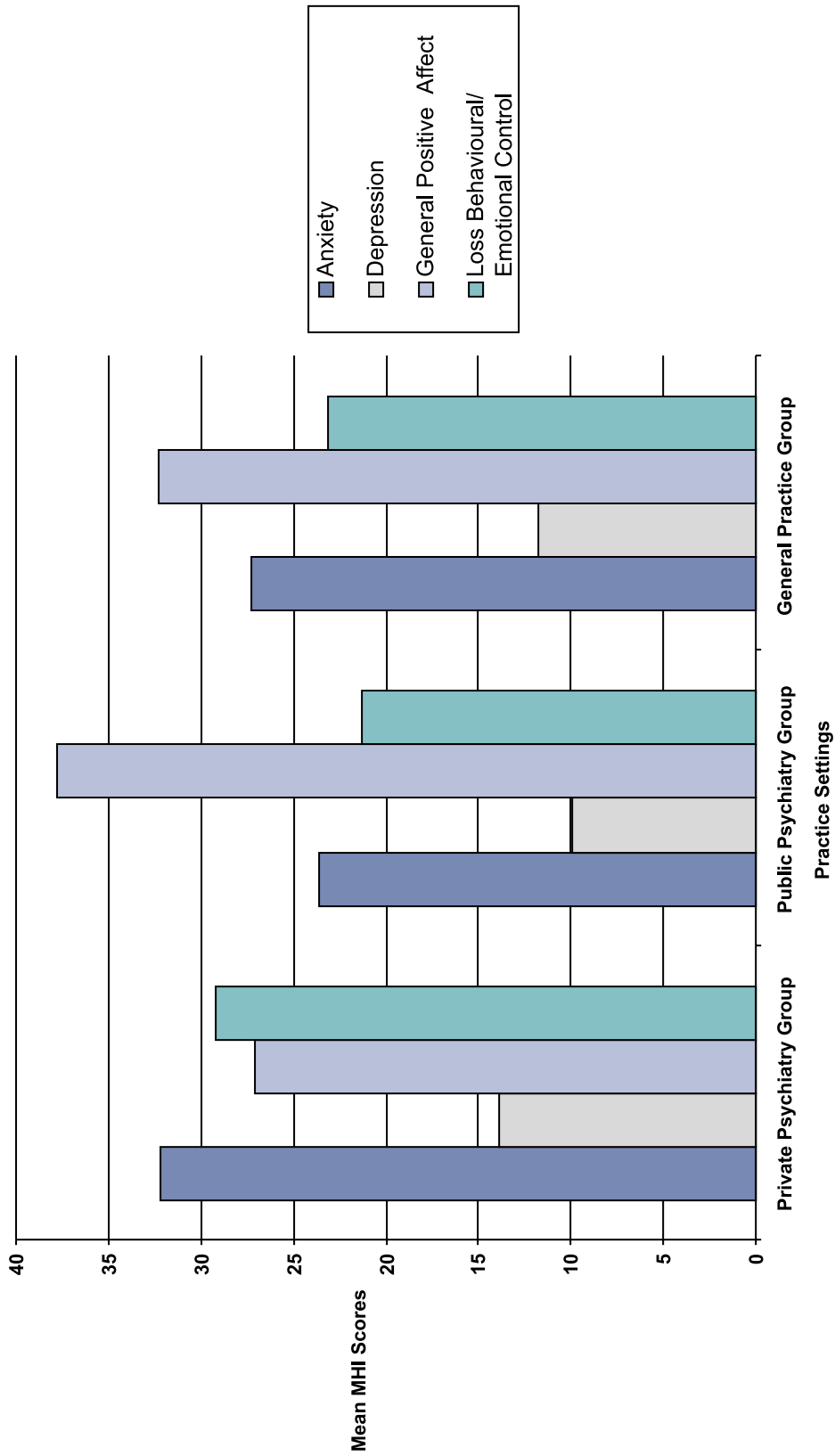


Figure 15: Mean Scores for the SF-36 Across Practice Settings

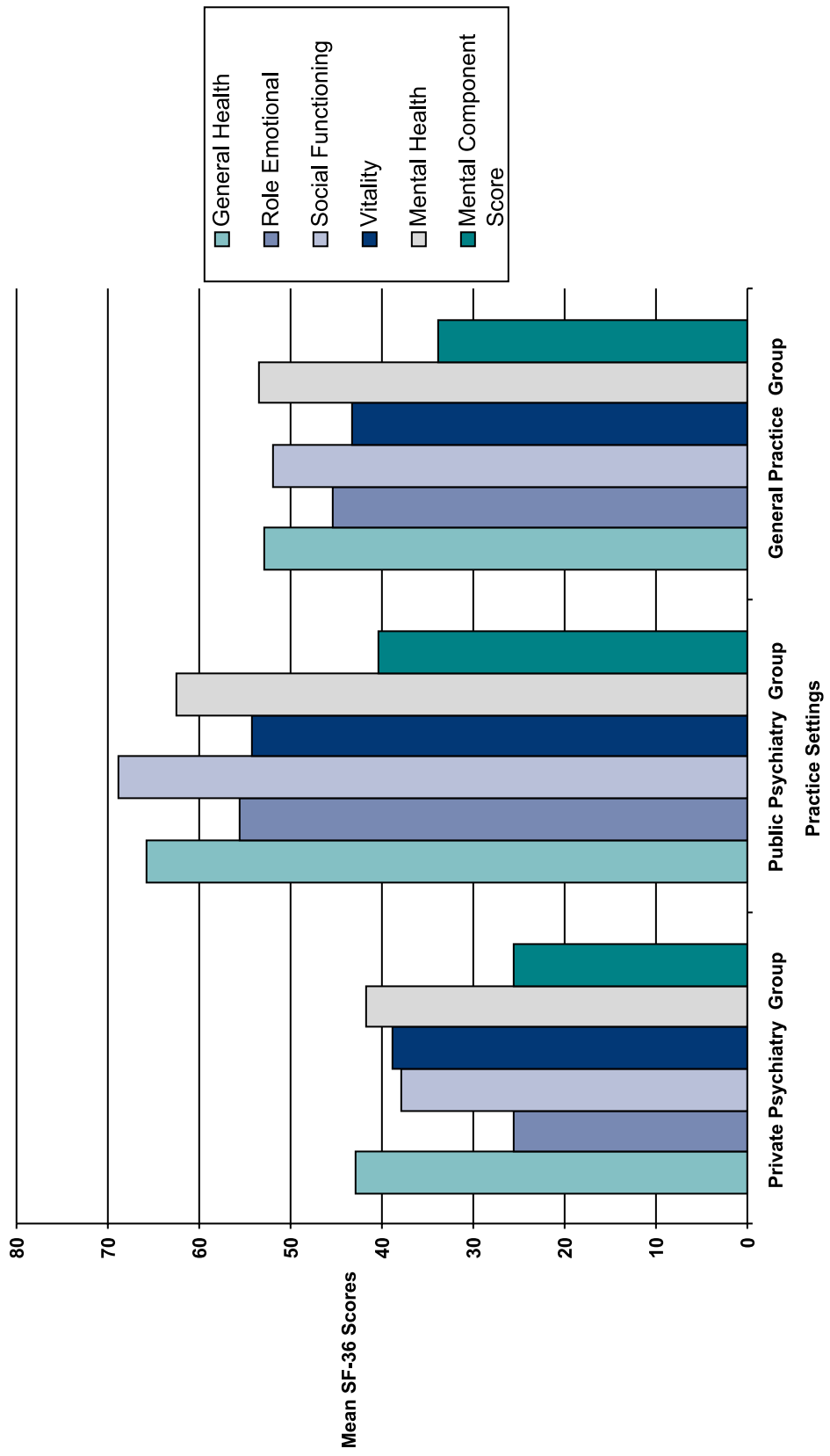


Figure 16: Mean Scores for the HoNOS Across Diagnostic Groups

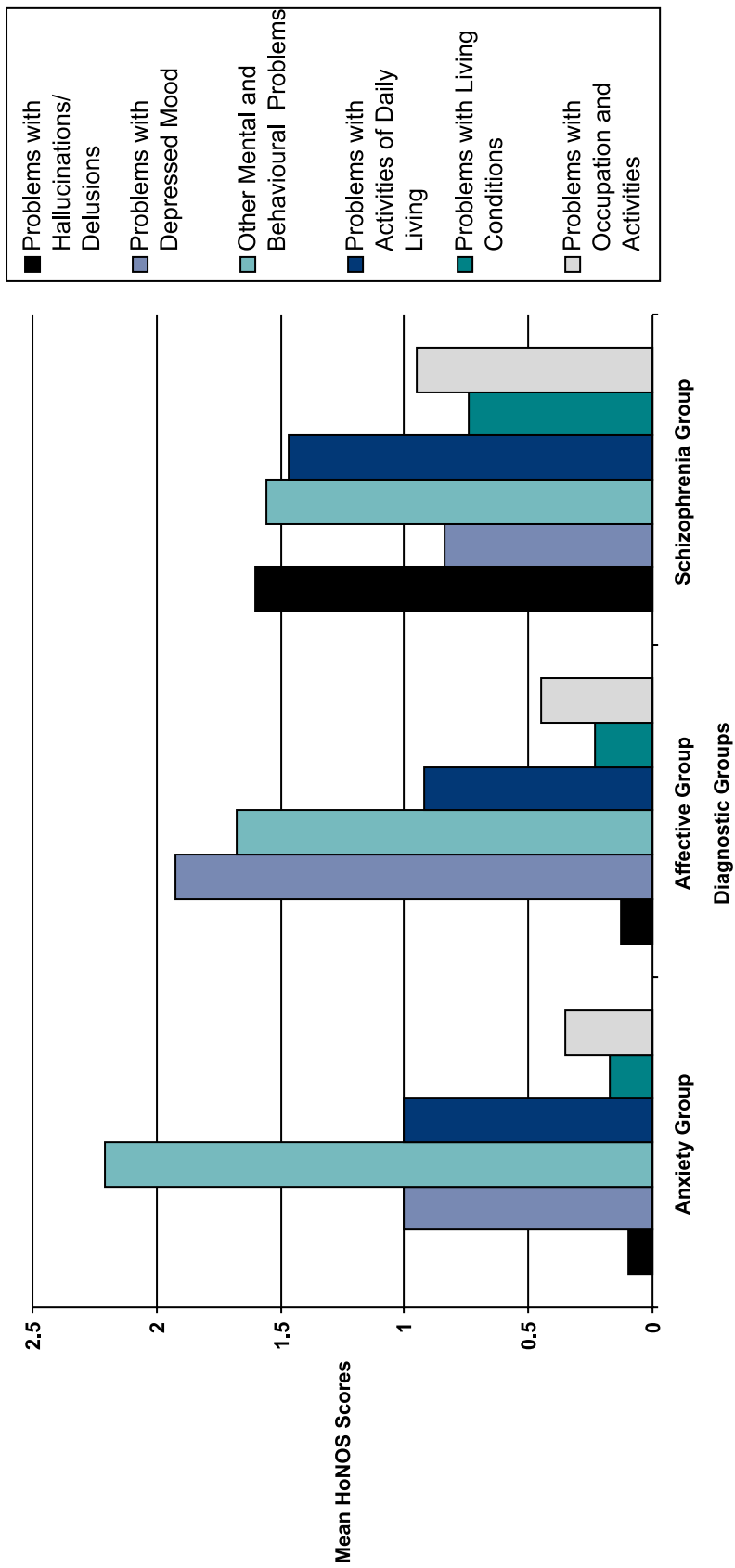


Figure 17: Mean Life Skills Profile Scores Across Diagnostic Groups

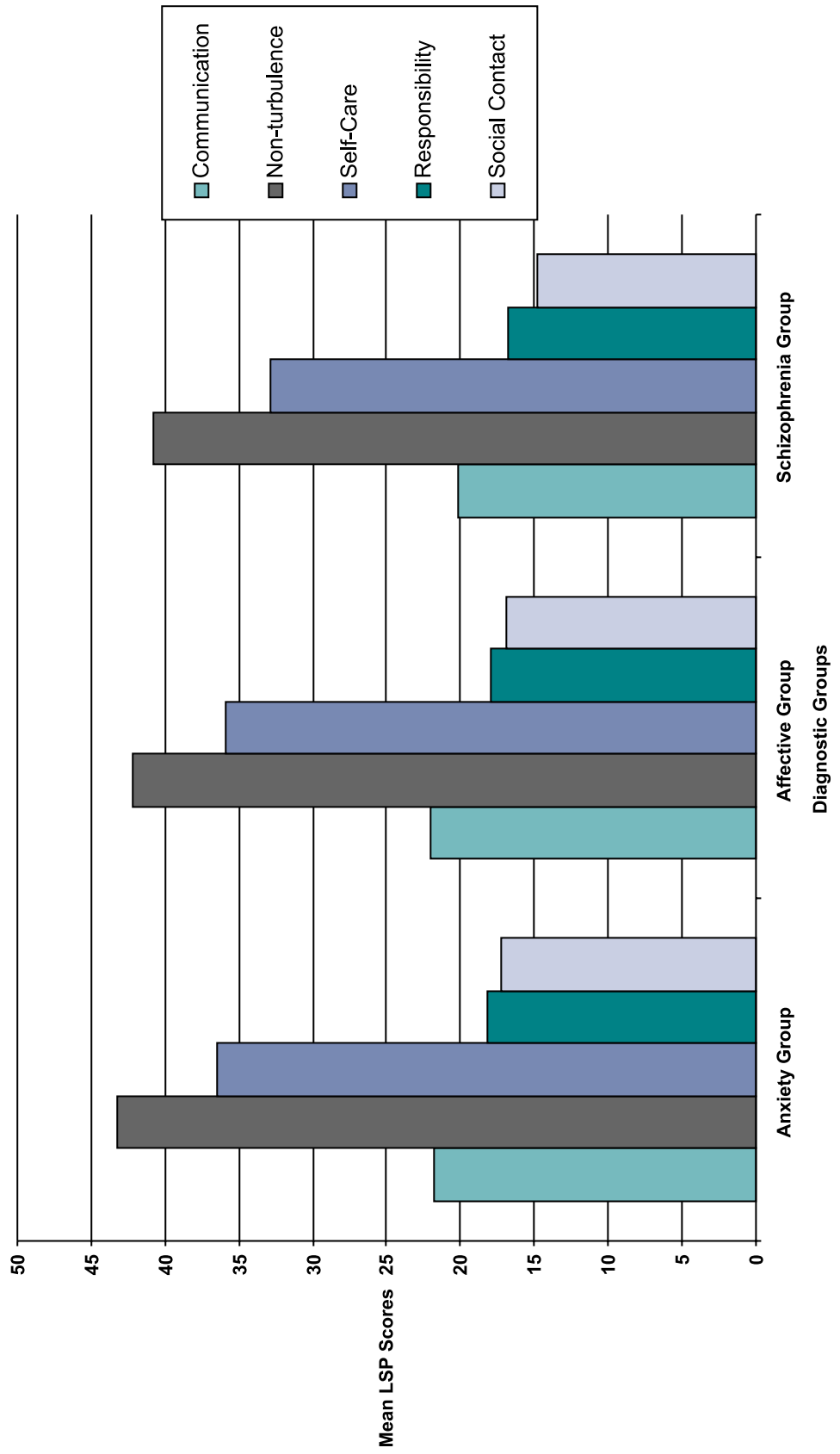


Figure 18: Mean Scores for the Role Functioning Scale Across Diagnostic Groups

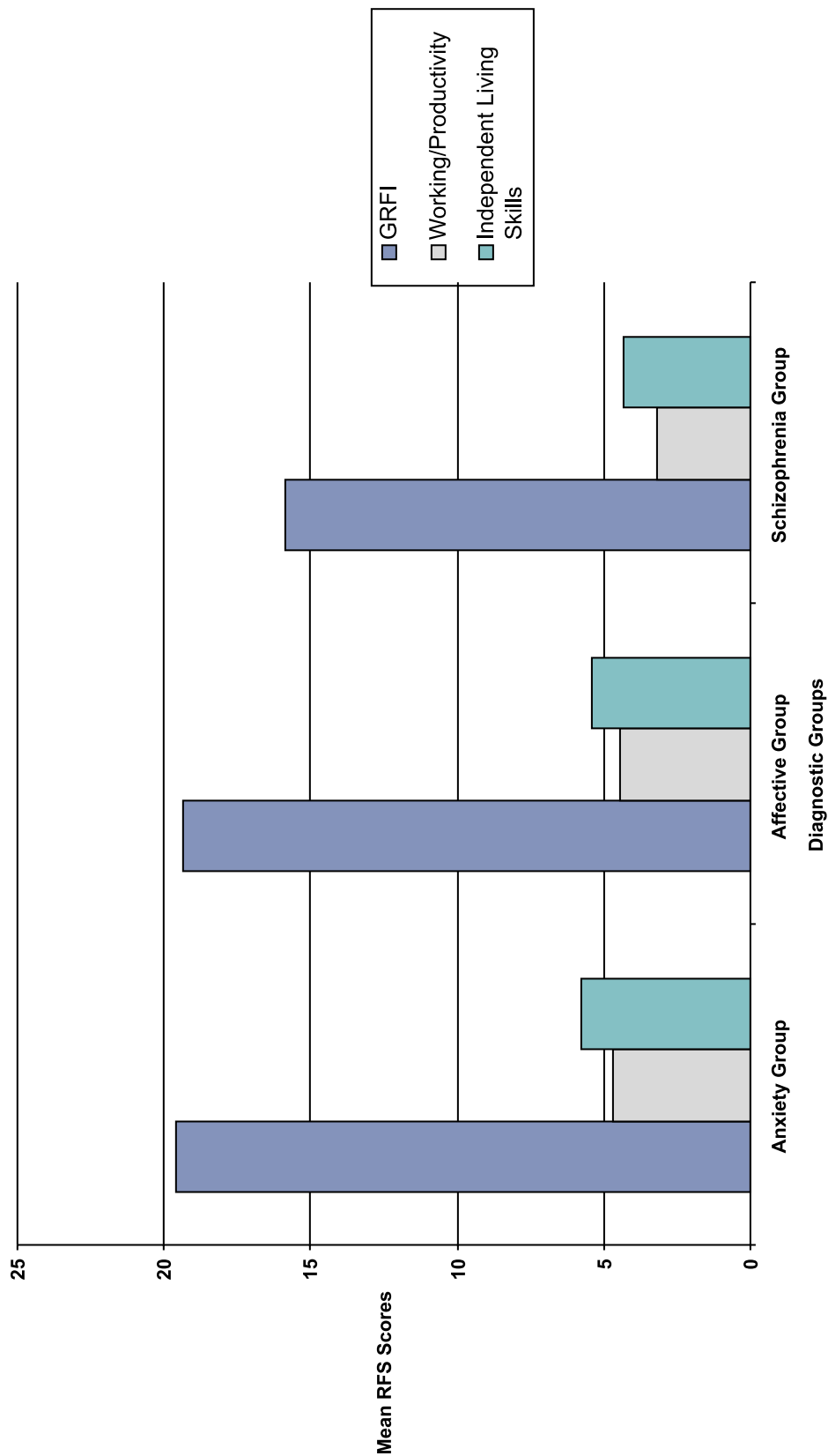


Figure 19: Mean Scores for the HoNOS Across Practice Settings

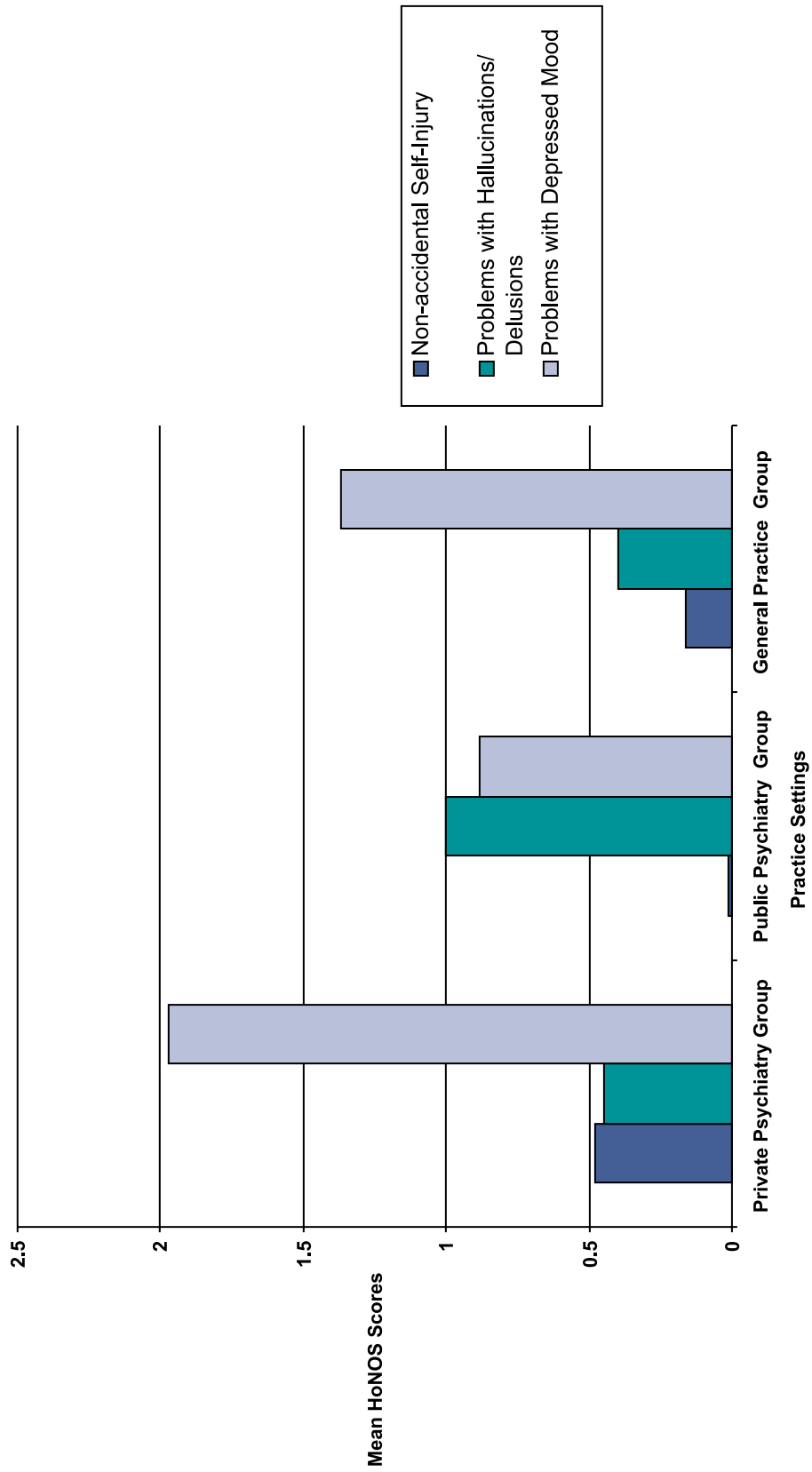


Figure 20: Mean Scores for the Life Skills Profile Across Practice Settings

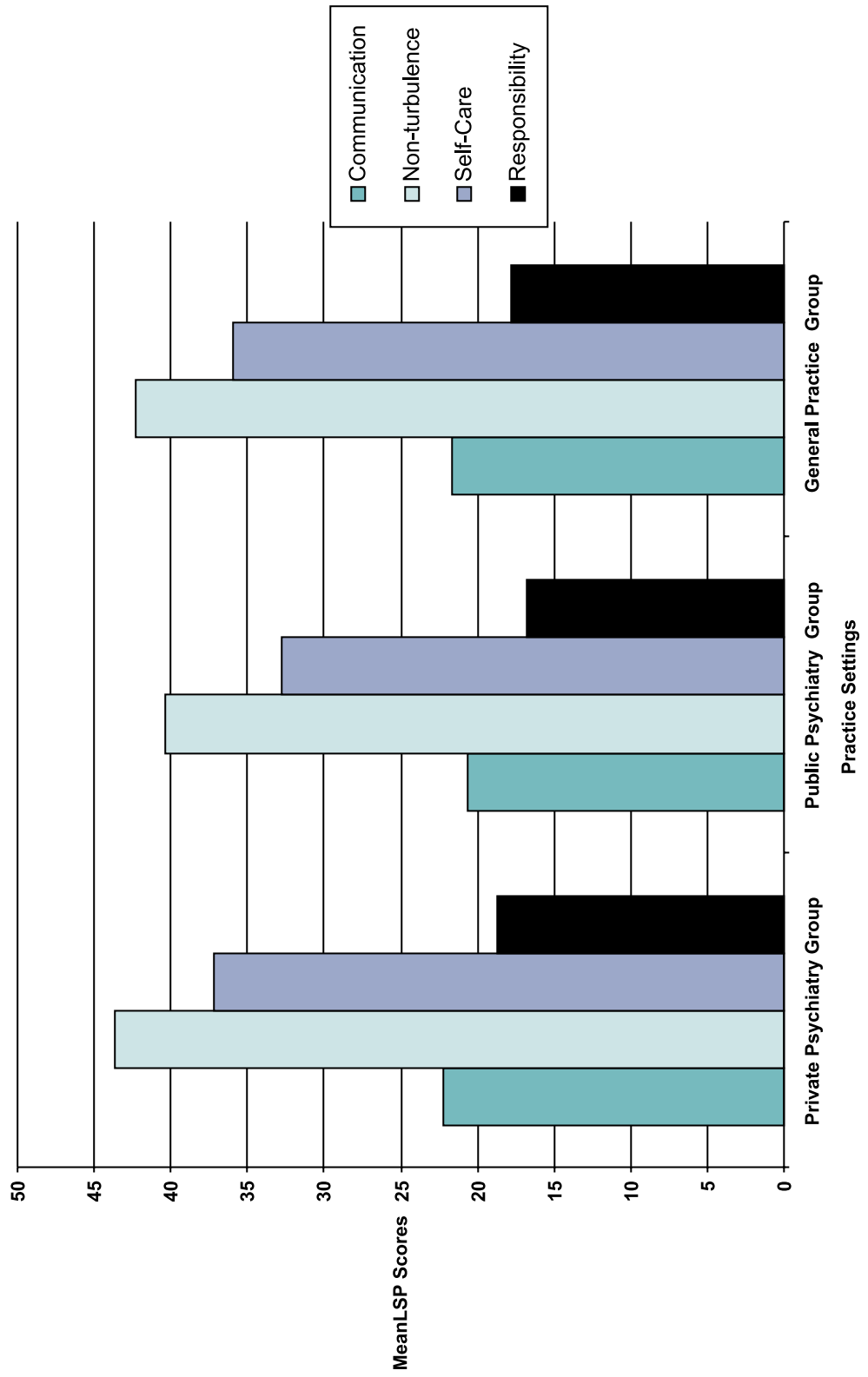


Figure 21: Mean Mental Component Summary Scores (SF-36) Across Change Groups

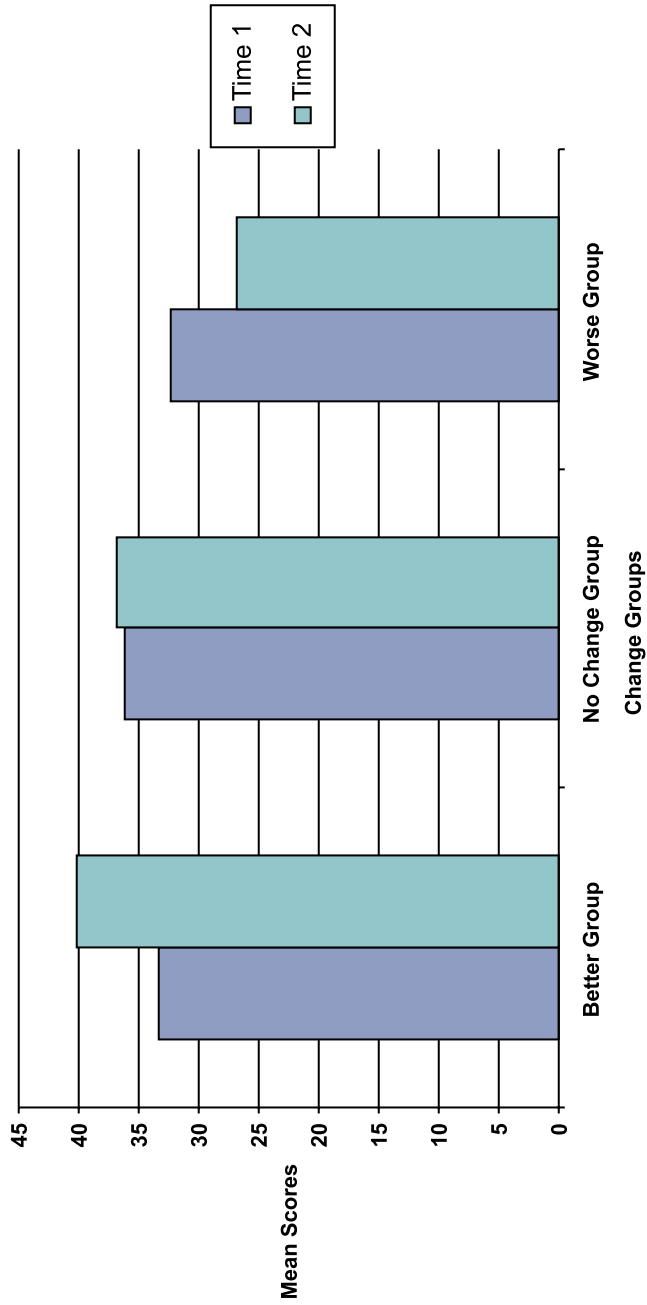


Figure 22: Mean Mental Health Inventory Index Scores for Time 1 and 2 Across Change Groups

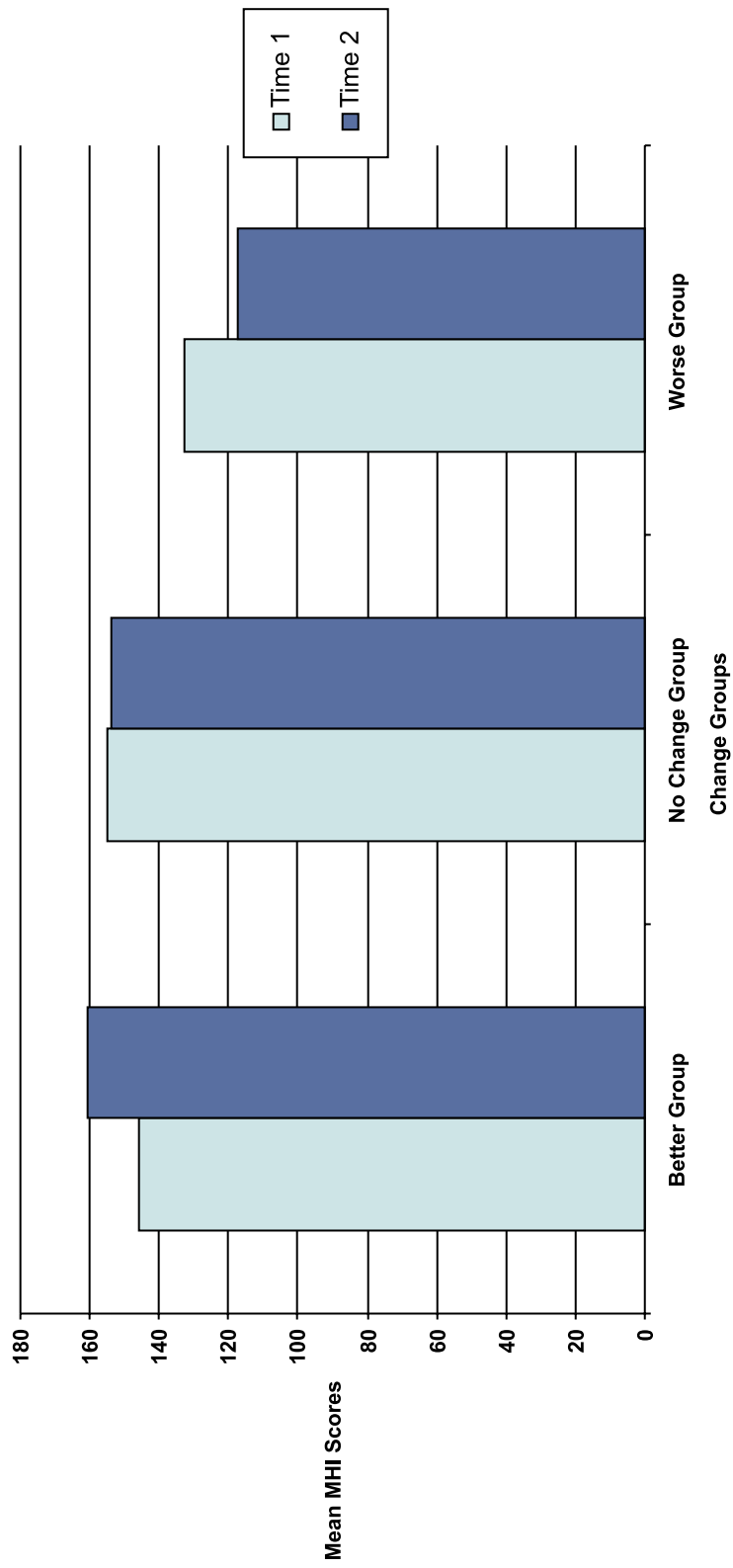


Figure 23: Time 1 and 2 Life Skills Profile Total Scores Across Change Groups

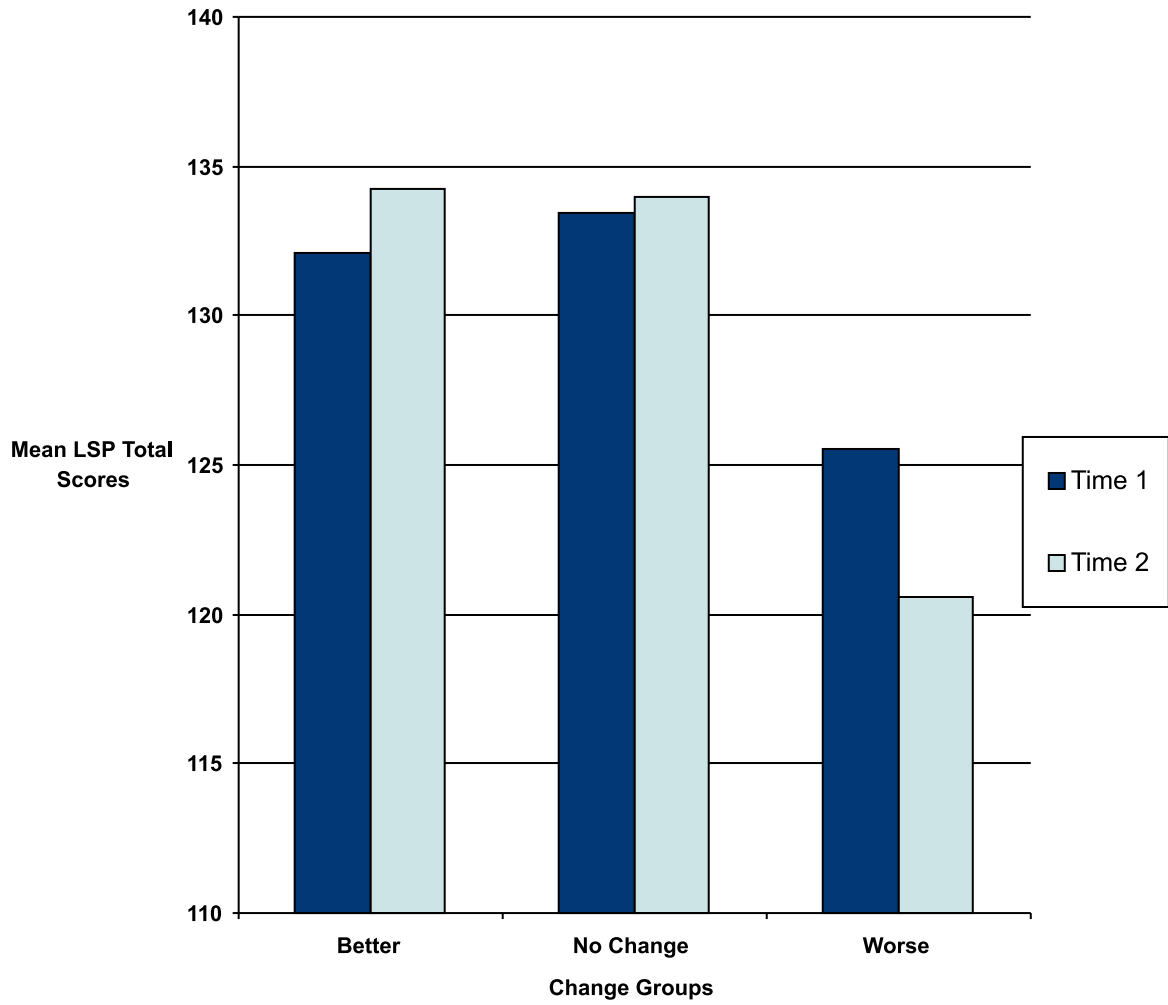


Figure 24: Time 1 and 2 Mental Health Inventory Index Scores Across Diagnostic Groups

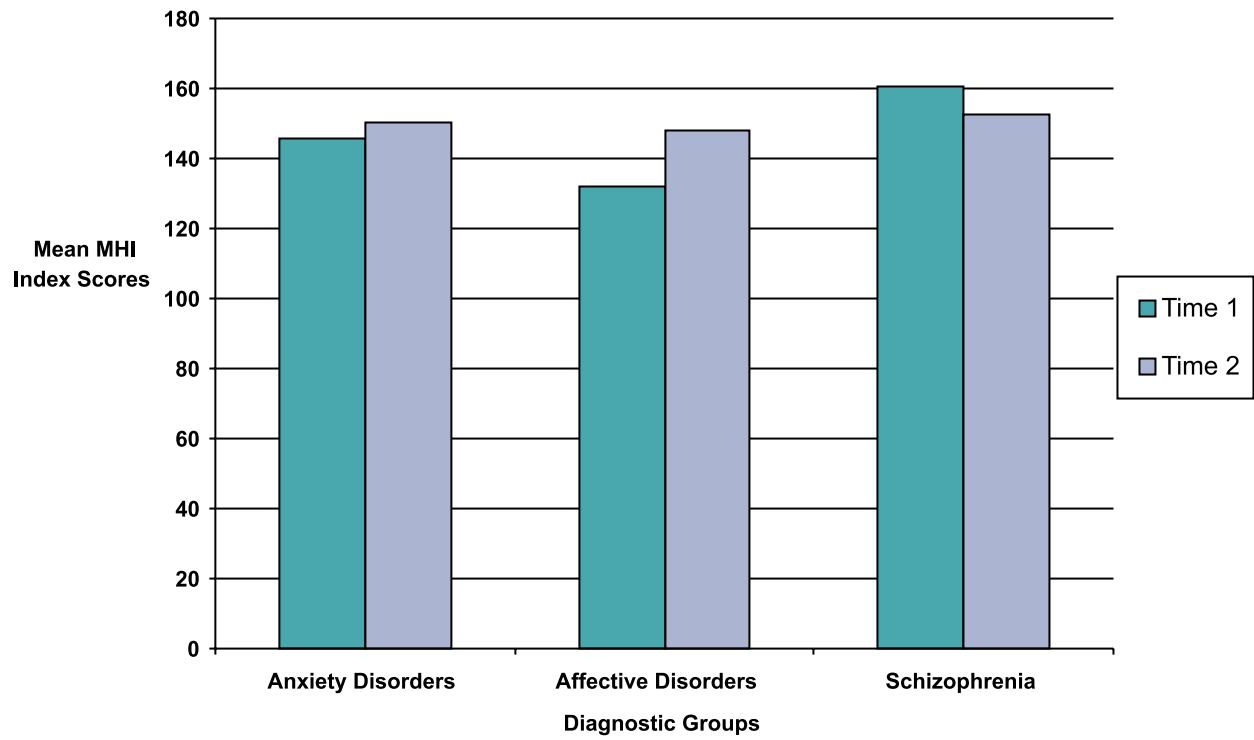


Figure 25: Time 1 and 2 Mental Component Summary Scores (SF-36) Across Diagnostic Groups

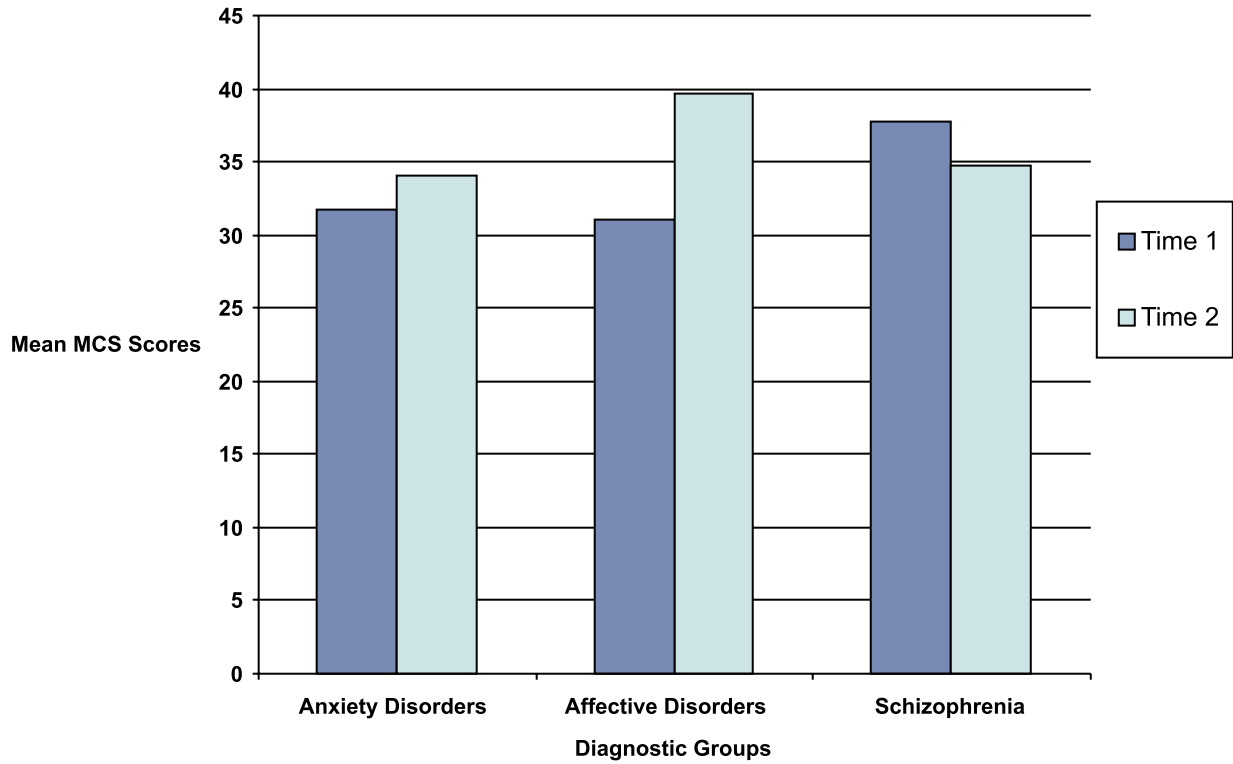
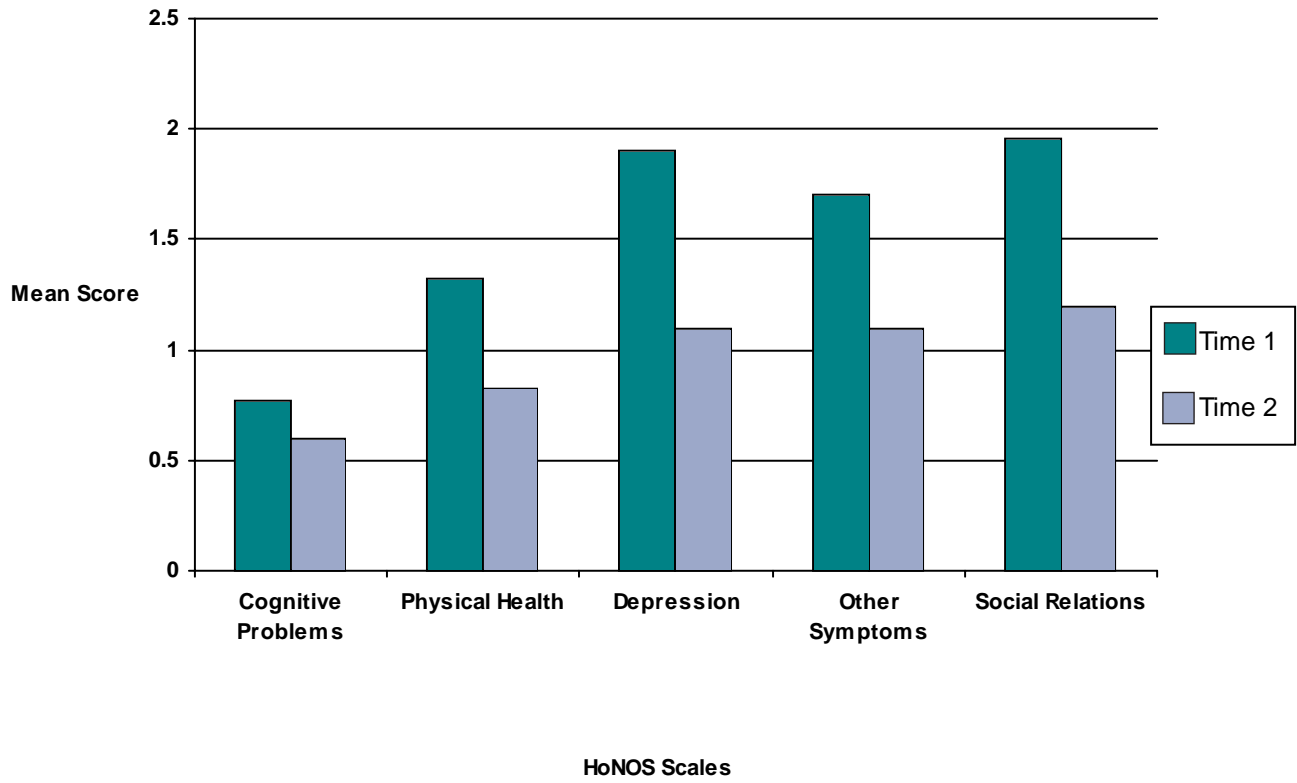


Figure 26: Mean Time 1 and 2 HoNOS Scores for the Affective Group



Appendix 5: Selected Measures

The Behaviour and Symptom Identification Scale (BASIS-32) (Example Only)

Instructions:

On the following pages are lists of problems, and areas of life functioning, in which some people experience difficulties.

Put a tick in the box that best describes the degree of difficulty you have been experiencing in each area during the PAST WEEK.

For instance, Question 1 asks you: “To what extent have you been experiencing difficulties in the area of: Managing day to day life.”

You should respond by placing a tick in the box which corresponds to the choices of: No Difficulty; A little; Moderate; Quite a bit; or, Extreme.

To what extent are you experiencing difficulty in the area of:

	No difficulty	A Little	Moderate	Quite a bit	Extreme
1. Managing day-to-day life. (e.g. getting to places on time, handling money, making everyday decisions)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2. Household responsibilities. (e.g. shopping, cooking, laundry, keeping room clean, other chores).	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3. Work. (e.g. completing tasks, performance level, finding/keeping a job).	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4. School (e.g. academic performance, completing assignments, attendance).	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

To what extent are you experiencing difficulty in the area of:

	No difficulty	A Little	Moderate	Quite a bit	Extreme a bit
5. Leisure time or recreational activities.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
6. Adjusting to major life stresses. (e.g. separation, divorce, moving, new job, new school, a death).	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
7. Relationships with family members.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
8. Getting along with people outside of the family.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
9. Isolation or feelings of loneliness.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

To what extent are you experiencing difficulty in the area of:

	No difficulty	A Little	Moderate	Quite a bit	Extreme
10. Being able to feel close to others.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
11. Being realistic about yourself or others.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
12. Recognising and expressing emotions appropriately.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
13. Developing independence, autonomy.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
14. Goals or direction in life.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
15. Lack of self-confidence, feeling bad about yourself.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

To what extent are you experiencing difficulty in the area of:

	No difficulty	A Little	Moderate	Quite a bit	Extreme
16. Apathy, lack of interest in things.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
17. Depression, hopelessness.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
18. Suicidal feelings or behaviour.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
19. Physical symptoms (e.g. headaches, aches & pains, sleep disturbance, stomach aches, dizziness).	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
20. Fear, anxiety or panic.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

To what extent are you experiencing difficulty in the area of:

	No difficulty	A Little	Moderate	Quite a bit	Extreme
21. Confusion, concentration, memory.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
22. Disturbing or unreal thoughts or beliefs.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
23. Hearing voices, seeing things.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
24. Manic, bizarre behaviour.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
25. Mood swings, unstable moods.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
26. Uncontrollable, compulsive behaviour. (e.g. eating disorder, hand-washing, hurting yourself)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Specify: _____

To what extent are you experiencing difficulty in the area of:

	No difficulty	A Little	Moderate	Quite a bit	Extreme
27. Sexual activity or preoccupation.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
28. Drinking alcoholic beverages.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
29. Taking illegal drugs, misusing drugs.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
30. Controlling temper, outbursts of anger, violence.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
31. Impulsive, illegal or reckless behaviour.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
32. Feeling satisfaction with your life.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

**The Mental Health Inventory
(MHI)
(Example Only)**

Instructions:

Please read each question and tick the box by the ONE statement that best describes how things have been FOR YOU during the past month.

There are no right or wrong answers.

1. How happy, satisfied, or pleased have you been with your personal life during the past month? (*Tick one*)

- Extremely happy, could not have been more satisfied or pleased
- Very happy most of the time.
- Generally satisfied, pleased
- Sometimes fairly satisfied, sometimes fairly unhappy
- Generally dissatisfied, unhappy
- Very dissatisfied, unhappy most of the time

2. How much of the time have you felt lonely during the past month? (*Tick one*)

- All of the time
- Most of the time
- A good bit of the time
- Some of the time
- A little of the time
- None of the time

3. How often did you become nervous or jumpy when faced with excitement or unexpected situations during the past month? (*Tick one*)

- Always
- Very often
- Fairly often
- Sometimes
- Almost never
- Never

4. During the past month, how much of the time have you felt that the future looks hopeful and promising? (*Tick one*)

- All of the time
- Most of the time
- A good bit of the time
- Some of the time
- A little of the time
- None of the time

5. How much of the time, during the past month, has your daily life been full of things that were interesting to you? (*Tick one*)

- All of the time
- Most of the time
- A good bit of the time
- Some of the time
- A little of the time
- None of the time

6. **How much of the time, during the past month, did you feel relaxed and free of tension? (Tick one)**
- All of the time Some of the time
- Most of the time A little of the time
- A good bit of the time None of the time
7. **During the past month, how much of the time have you generally enjoyed the things you do? (Tick one)**
- All of the time Some of the time
- Most of the time A little of the time
- A good bit of the time None of the time
8. **During the past month, have you had any reason to wonder if you were losing your mind, or losing control over the way you act, talk, think, feel, or of your memory? (Tick one)**
- No, not at all
- Maybe a little
- Yes, but not enough to be concerned or worried about it
- Yes, and I have been a little concerned
- Yes, and I am quite concerned
- Yes, and I am very much concerned about it
9. **Did you feel depressed during the past month? (Tick one)**
- No, not at all
- Maybe a little
- Yes, but not enough to be concerned or worried about it
- Yes, and I have been a little concerned
- Yes, and I am quite concerned
- Yes, and I am very much concerned about it

- 10. During the past month, how much of the time have you felt loved and wanted? (*Tick one*)**
- | | |
|---|---|
| <input type="checkbox"/> All of the time | <input type="checkbox"/> Some of the time |
| <input type="checkbox"/> Most of the time | <input type="checkbox"/> A little of the time |
| <input type="checkbox"/> A good bit of the time | <input type="checkbox"/> None of the time |
- 11. How much of the time, during the past month, have you been a very nervous person? (*Tick one*)**
- | | |
|---|---|
| <input type="checkbox"/> All of the time | <input type="checkbox"/> Some of the time |
| <input type="checkbox"/> Most of the time | <input type="checkbox"/> A little of the time |
| <input type="checkbox"/> A good bit of the time | <input type="checkbox"/> None of the time |
- 12. When you got up in the morning, this past month, about how often did you expect to have an interesting day? (*Tick one*)**
- | | |
|---------------------------------------|---------------------------------------|
| <input type="checkbox"/> Always | <input type="checkbox"/> Sometimes |
| <input type="checkbox"/> Very often | <input type="checkbox"/> Almost never |
| <input type="checkbox"/> Fairly often | <input type="checkbox"/> Never |
- 13. During the past month, how much of the time have you felt tense or “high-strung”? (*Tick one*)**
- | | |
|---|---|
| <input type="checkbox"/> All of the time | <input type="checkbox"/> Some of the time |
| <input type="checkbox"/> Most of the time | <input type="checkbox"/> A little of the time |
| <input type="checkbox"/> A good bit of the time | <input type="checkbox"/> None of the time |
- 14. During the past month, have you been in firm control of your behaviour, thoughts, emotions or feelings? (*Tick one*)**
- | | |
|---|--|
| <input type="checkbox"/> Yes, very definitely | <input type="checkbox"/> No, not too well |
| <input type="checkbox"/> Yes, for the most part | <input type="checkbox"/> No, and I am somewhat disturbed |
| <input type="checkbox"/> Yes, I guess so | <input type="checkbox"/> No, and I am very disturbed |
- 15. During the past month, how often did your hands shake when you tried to do something? (*Tick one*)**
- | | |
|---------------------------------------|---------------------------------------|
| <input type="checkbox"/> Always | <input type="checkbox"/> Sometimes |
| <input type="checkbox"/> Very often | <input type="checkbox"/> Almost never |
| <input type="checkbox"/> Fairly often | <input type="checkbox"/> Never |

16. **During the past month, how often did you feel that you had nothing to look forward to? (*Tick one*)**
- | | |
|---------------------------------------|---------------------------------------|
| <input type="checkbox"/> Always | <input type="checkbox"/> Sometimes |
| <input type="checkbox"/> Very often | <input type="checkbox"/> Almost never |
| <input type="checkbox"/> Fairly often | <input type="checkbox"/> Never |
17. **How much of the time, during the past month, have you felt calm and peaceful? (*Tick one*)**
- | | |
|---|---|
| <input type="checkbox"/> All of the time | <input type="checkbox"/> Some of the time |
| <input type="checkbox"/> Most of the time | <input type="checkbox"/> A little of the time |
| <input type="checkbox"/> A good bit of the time | <input type="checkbox"/> None of the time |
18. **How much of the time, during the past month, have you felt emotionally stable? (*Tick one*)**
- | | |
|---|---|
| <input type="checkbox"/> All of the time | <input type="checkbox"/> Some of the time |
| <input type="checkbox"/> Most of the time | <input type="checkbox"/> A little of the time |
| <input type="checkbox"/> A good bit of the time | <input type="checkbox"/> None of the time |
19. **How much of the time, during the past month, have you felt downhearted and blue? (*Tick one*)**
- | | |
|---|---|
| <input type="checkbox"/> All of the time | <input type="checkbox"/> Some of the time |
| <input type="checkbox"/> Most of the time | <input type="checkbox"/> A little of the time |
| <input type="checkbox"/> A good bit of the time | <input type="checkbox"/> None of the time |
20. **How often have you felt like crying, during the past month? (*Tick one*)**
- | | |
|---------------------------------------|---------------------------------------|
| <input type="checkbox"/> Always | <input type="checkbox"/> Sometimes |
| <input type="checkbox"/> Very often | <input type="checkbox"/> Almost never |
| <input type="checkbox"/> Fairly often | <input type="checkbox"/> Never |
21. **During the past month, how often did you feel that others would be better off if you were dead? (*Tick one*)**
- | | |
|---------------------------------------|---------------------------------------|
| <input type="checkbox"/> Always | <input type="checkbox"/> Sometimes |
| <input type="checkbox"/> Very often | <input type="checkbox"/> Almost never |
| <input type="checkbox"/> Fairly often | <input type="checkbox"/> Never |

22. **How much of the time, during the past month, were you able to relax without difficulty? (Tick one)**

- | | |
|---|---|
| <input type="checkbox"/> All of the time | <input type="checkbox"/> Some of the time |
| <input type="checkbox"/> Most of the time | <input type="checkbox"/> A little of the time |
| <input type="checkbox"/> A good bit of the time | <input type="checkbox"/> None of the time |

23. **How much of the time, during the past month, did you feel that your love relationships, loving and being loved, were full and complete? (Tick one)**

- | | |
|---|---|
| <input type="checkbox"/> All of the time | <input type="checkbox"/> Some of the time |
| <input type="checkbox"/> Most of the time | <input type="checkbox"/> A little of the time |
| <input type="checkbox"/> A good bit of the time | <input type="checkbox"/> None of the time |

24. **How often, during the past month did you feel that nothing turned out for you the way you wanted it to? (Tick one)**

- | | |
|---------------------------------------|---------------------------------------|
| <input type="checkbox"/> Always | <input type="checkbox"/> Sometimes |
| <input type="checkbox"/> Very often | <input type="checkbox"/> Almost never |
| <input type="checkbox"/> Fairly often | <input type="checkbox"/> Never |

25. **How much have you been bothered by nervousness, or your “nerves”, during the past month? (Tick one)**

- | | |
|---|---|
| <input type="checkbox"/> Extremely so, to the point where I could not take care of things | <input type="checkbox"/> Bothered some, enough to notice |
| <input type="checkbox"/> Very much bothered | <input type="checkbox"/> Bothered just a little by nerves |
| <input type="checkbox"/> Fairly often | <input type="checkbox"/> Not bothered at all by this |

26. **During the past month, how much of the time has living been a wonderful adventure for you? (Tick one)**

- | | |
|---|---|
| <input type="checkbox"/> All of the time | <input type="checkbox"/> Some of the time |
| <input type="checkbox"/> Most of the time | <input type="checkbox"/> A little of the time |
| <input type="checkbox"/> A good bit of the time | <input type="checkbox"/> None of the time |

27. **How often, during the past month, have you felt so down in the dumps that nothing could cheer you up? (Tick one)**

- | | |
|---------------------------------------|---------------------------------------|
| <input type="checkbox"/> Always | <input type="checkbox"/> Sometimes |
| <input type="checkbox"/> Very often | <input type="checkbox"/> Almost never |
| <input type="checkbox"/> Fairly often | <input type="checkbox"/> Never |

28. During the past month, did you think about taking your own life? (*Tick one*)

- Yes, very often
- Yes, fairly often
- Yes, a couple of times
- Yes, at one time
- No, never

29. During the past month, how much of the time have you felt restless, fidgety, or impatient? (*Tick one*)

- All of the time
- Most of the time
- A good bit of the time
- Some of the time
- A little of the time
- None of the time

30. During the past month, how much of the time have you been moody or brooded about things? (*Tick one*)

- All of the time
- Most of the time
- A good bit of the time
- Some of the time
- A little of the time
- None of the time

31. How much of the time, during the past month, have you felt cheerful, light-hearted? (*Tick one*)

- All of the time
- Most of the time
- A good bit of the time
- Some of the time
- A little of the time
- None of the time

32. During the past month, how often did you get rattled, upset, or flustered?

(Tick one)

- | | |
|---------------------------------------|---------------------------------------|
| <input type="checkbox"/> Always | <input type="checkbox"/> Sometimes |
| <input type="checkbox"/> Very often | <input type="checkbox"/> Almost never |
| <input type="checkbox"/> Fairly often | <input type="checkbox"/> Never |

33. During the past month, have you been anxious or worried? *(Tick one)*

- Yes, extremely so to the point of being sick or almost sick
- Yes, very much so
- Yes, quite a bit
- Yes, some, enough to bother me
- Yes a little bit
- No, not at all

34. During the past month, how much of the time were you a happy person?

(Tick one)

- | | |
|---|---|
| <input type="checkbox"/> All of the time | <input type="checkbox"/> Some of the time |
| <input type="checkbox"/> Most of the time | <input type="checkbox"/> A little of the time |
| <input type="checkbox"/> A good bit of the time | <input type="checkbox"/> None of the time |

35. How often during the past month did you find yourself having difficulty trying to calm down? *(Tick one)*

- | | |
|---------------------------------------|---------------------------------------|
| <input type="checkbox"/> Always | <input type="checkbox"/> Sometimes |
| <input type="checkbox"/> Very often | <input type="checkbox"/> Almost never |
| <input type="checkbox"/> Fairly often | <input type="checkbox"/> Never |

36. During the past month, how much of the time have you been in low or very low spirits? *(Tick one)*

- | | |
|---|---|
| <input type="checkbox"/> All of the time | <input type="checkbox"/> Some of the time |
| <input type="checkbox"/> Most of the time | <input type="checkbox"/> A little of the time |
| <input type="checkbox"/> A good bit of the time | <input type="checkbox"/> None of the time |

36. How often, during the past month, have you been waking up feeling fresh and rested? (*Tick one*)

- | | |
|--|---|
| <input type="checkbox"/> Always, every day | <input type="checkbox"/> Some days, but usually not |
| <input type="checkbox"/> Almost every day | <input type="checkbox"/> Hardly ever |
| <input type="checkbox"/> Most days | <input type="checkbox"/> Never wake up feeling rested |

38. During the past month, have you been under or felt you were under any strain, stress or pressure? (*Tick one*)

- Yes, almost more than I could stand or bear
- Yes, quite a bit of pressure
- Yes, some more than usual
- Yes, some-but about normal
- Yes, a little bit
- No, not at all

**The Medical Outcomes Study Short Form
(SF-36)
(Example Only)**

Instructions:

Please read each question and tick the box by the ONE statement that best describes how things have been FOR YOU during the past month.

There are no right or wrong answers.

1. In general, would you say your health is,

(circle one)

- Excellent 1
- Very good 2
- Good..... 3
- Fair 4
- Poor 5

2. Compared to one year ago, how would you rate your health in general now?

(circle one)

- Much better now than one year ago 1
- Somewhat better now than one year ago 2
- About the same as one year ago 3
- Somewhat worse now than one year ago 4
- Much worse now than one year ago 5

The following questions are about activities you might do during a typical day. Does your health now limit you in these activities? If so, how much?

(circle one number on each line)

ACTIVITIES	Yes, Limited A Lot	Yes, Limited A Little All	No, Not Limited At All
a. Vigorous activities, such as running, lifting heavy objects, participating in strenuous sports	1	2	3
b. Moderate activities, such as moving a table, pushing a vacuum cleaner, bowling, or playing golf	1	2	3
c. Lifting or carrying groceries	1	2	3
d. Climbing several flights of stairs	1	2	3
e. Climbing one flight of stairs	1	2	3
f. Bending, kneeling or stooping	1	2	3
g. Walking more than one kilometre	1	2	3
h. Walking half a kilometre	1	2	3
i. Walking 100 metres	1	2	3
j. Bathing or dressing yourself	1	2	3

4. During the past 4 weeks, have you had any of the following problems with your work or other regular daily activities as a result of your physical health?

(circle one number on each line)

	YES	NO
a. Cut down on the amount of time you spent on work or other Activities	1	2
b. Accomplished less than you would like	1	2
c. Were limited in the kind of work or other activities	1	2
d. Had difficulty performing the work or other activities (for example, it took extra effort)	1	2

5. During the past 4 weeks, have you had any of the following problems with your work or other regular daily activities as a result of any emotional problems (such as feeling depressed or anxious)?

(circle one number on each line)

	YES	NO
a. Cut down on the amount of time you spent on work or other activities	1	2
b. Accomplished less than you would like	1	2
c. Didn't do work or other activities as carefully as usual	1	2

6. During the past 4 weeks, to what extent has Your physical health or emotional problems interfered with your normal social activities with family, friends, neighbours or groups?

(circle one)

- Not at all 1
- Slightly 2
- Moderately 3
- Quite a bit 4
- Extremely 5

7. How much bodily pain have you had during the past 4 weeks?

(circle one)

- No bodily pain 1
- Very mild 2
- Mild 3
- Moderate 4
- Severe 5
- Very severe 6

8. During the past 4 weeks, how much did pain interfere with your normal work (including both work outside the home and housework)? (circle one)

- Not at all 1
- A little bit 2
- Moderately 3
- Quite a bit 4
- Extremely 5

9. These questions are about how you feel and how things have been with you during the past 4 weeks. For each question, please give the one answer that comes closest to the way you have been feeling.

How much of the time during the past 4 weeks, (circle one number on each line)

	All of the Time	Most of the Time	A Good Bit Of the Time	Some of the Time	A Little of the time	None of the Time
a. Did you feel full of life?	1	2	3	4	5	6
b. Have you been a very nervous person?	1	2	3	4	5	6
c. Have you felt so down in the dumps that nothing could cheer you up?	1	2	3	4	5	6
d. Have you felt calm and peaceful?	1	2	3	4	5	6
e. Did you have a lot of energy?	1	2	3	4	5	6
f. Have you felt down?	1	2	3	4	5	6
g. Did you feel worn out?	1	2	3	4	5	6
h. Have you been a happy person?	1	2	3	4	5	6
i. Did you feel tired?	1	2	3	4	5	6

10 During the past 4 weeks, how much of the time has your physical health or emotional problems interfered with your social activities (like visiting with friends, relatives. etc.)?

(circle one)

- All of the time 1
- Most of the time 2
- Some of the time 3
- A little of the time 4
- None of the time 5

11. How TRUE or FALSE is each of the following statements for you?

(Circle one number on each line)

	Definitely True	Mostly True	Don't Know	Mostly False	Definitely False
a. I seem to get sick a little easier than other people	1	2	3	4	5
b. I am as healthy as anybody I know	1	2	3	4	5
c. I expect my health to get worse	1	2	3	4	5
d. My health is excellent	1	2	3	4	5

LIFE SKILLS PROFILE

(Example Only)

Instructions:

Please respond to the following questions to assess your client's general functioning (i.e. not during crises when he or she is ill, or becoming ill, but his or her general state over the past three months).

Answer all items by marking the box that corresponds to the description that you feel best describes the behaviour of your client.

1. Does this person generally have any difficulty with initiating and responding to conversation?
- | | | | |
|---------------------------------|-------------------------------------|---------------------------------------|--------------------------------------|
| <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| No difficulty with conversation | Slight difficulty with conversation | Moderate difficulty with conversation | Extreme difficulty with conversation |
2. Does this person generally intrude or burst in on others' conversation (e.g. interrupts you when you are talking)?
- | | | | |
|--------------------------|--------------------------|--------------------------|--------------------------|
| <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Not intrusive at all | Slightly intrusive | Moderately intrusive | Extremely intrusive |
3. Does this person generally withdraw from social contact?
- | | | | |
|--------------------------|--------------------------|--------------------------|-----------------------------------|
| <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Does not withdraw at all | Withdraws slightly | Withdraws moderately | Withdraws totally or near totally |
4. Does this person generally show warmth to others?
- | | | | |
|--------------------------|--------------------------|--------------------------|--------------------------|
| <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Considerable warmth | Moderate warmth | Slight warmth | No warmth at all |
5. Is this person generally angry or prickly towards others?
- | | | | |
|--------------------------|--------------------------|--------------------------|--------------------------|
| <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Not angry at all | Slightly angry | Moderately angry | Extremely angry |

6. Does this person generally take offence readily?

Doesn't take
offence

Somewhat ready
to take offence

Quite ready to
take offence

Extremely ready
to take offence

7. Does this person generally make eye contact with others when in conversation?

Appropriate eye
contact

Slightly reduced
eye contact

Moderately reduced
eye contact

Extremely reduced
or no eye contact

8. Is it generally difficult to understand this person because of the way he or she speaks (e.g. jumbled, garbled or disordered)?

Not at all
difficult

Slightly difficult

Moderately
difficult

Extremely
difficult

9. Does this person generally talk about odd or strange ideas?

No odd ideas

Slightly odd ideas

Moderately odd
ideas

Extremely odd
ideas

10. Is this person generally well groomed (e.g. neatly dressed, hair combed)?

Well groomed

Moderately well
groomed

Poorly groomed

Extremely
poorly groomed

11. Is this person's appearance (facial appearance, gestures) generally appropriate to his or her surroundings?

Unremarkable or
appropriate

Slightly bizarre
or inappropriate

Moderately bizarre
or inappropriate

Extremely bizarre
or inappropriate

12. Does this person wash himself or herself without reminding?

- | | | | |
|--------------------------|--------------------------|--------------------------|--------------------------|
| <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Generally | Occasionally | Rarely | Never |

13. Does this person generally have an offensive smell (e.g. due to body, breath or clothes)?

- | | | | |
|--------------------------|--------------------------|--------------------------|--------------------------|
| <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Not at all | Smells slightly | Smells moderately | Smells a lot |

14. Does this person wear clean clothes generally, or ensure that they are cleaned if dirty?

- | | | | |
|--|---------------------------------------|-----------------------------------|--|
| <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Maintains
cleanliness of
clothes | Moderate
cleanliness of
clothes | Poor
cleanliness of
clothes | Very poor
cleanliness of
clothes |

15. Does this person generally neglect her or his physical health?

- | | | | |
|--------------------------|--|---|--|
| <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| No neglect | Slight neglect of
physical problems | Moderate neglect
of physical
problems | Extreme neglect
of physical
problems |

16. Does this person generally maintain an adequate diet?

- | | | | |
|--------------------------|--------------------------|--------------------------|--------------------------|
| <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| No problem | Slight problem | Moderate problem | Extreme problem |

17. Does this person generally look after and take her or his own prescribed medication (or attend for prescribed injections on time) without reminding?

- | | | | |
|-----------------------------|--------------------------|--------------------------|--------------------------|
| <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Reliable with
medication | Slightly unreliable | Moderately
unreliable | Extremely
unreliable |

18. Is this person willing to take psychiatric medication when prescribed by a doctor?

- | | | | |
|--------------------------|--------------------------|--------------------------|--------------------------|
| <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Always | Usually | Rarely | Never |

19. Does this person co-operate with health services (e.g. doctors and/or other health workers)?

- | | | | |
|--------------------------|--------------------------|--------------------------|--------------------------|
| <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Always | Usually | Rarely | Never |

20. Is this person generally inactive (e.g. spends most of the time sitting or standing around doing nothing)?

- | | | | |
|--------------------------|--------------------------|--------------------------|--------------------------|
| <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Appropriately active | Slightly inactive | Moderately inactive | Extremely inactive |

21. Does this person generally have definite interests (e.g. hobbies, sports, activities) in which he or she is involved regularly?

- | | | | |
|--------------------------|--------------------------|--------------------------|--------------------------|
| <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Considerable involvement | Moderate Involvement | Some Involvement | Not involved at all |

22. Does this person attend any social organisation (e.g. church, club or interest group but excluding psychiatric therapy groups)?

- | | | | |
|--------------------------|--------------------------|--------------------------|--------------------------|
| <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Frequently | Occasionally | Rarely | Never |

23. Can this person generally prepare (if needed) her or his own food/meals?

- | | | | |
|---------------------------------------|--------------------------|--------------------------|---|
| <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Quite capable of preparing food/meals | Slight limitations | Moderate limitations | Totally incapable of preparing food/meals |

24. Can this person generally budget (if needed) to live within his or her means?

- | | | | |
|-------------------------------|--------------------------|--------------------------|-----------------------------------|
| <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Quite capable
of budgeting | Slight limitations | Moderate
limitations | Totally incapable
of budgeting |

25. Does this person generally have problems (e.g. friction, avoidance) living with others in the household?

- | | | | |
|--------------------------|--------------------------|--------------------------|------------------------------|
| <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| No obvious
problems | Slight problems | Moderate | Extreme problems
problems |

26. What sort of work is this person generally capable of (even if unemployed, retired or doing unpaid domestic duties)?

- | | | | |
|------------------------------|------------------------------|-----------------------------------|------------------------------|
| <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Capable of
full-time work | Capable of
part-time work | Capable only
of sheltered work | Totally incapable
of work |

27. Does this person behave recklessly (e.g. ignoring traffic when crossing the road)?

- | | | | |
|--------------------------|--------------------------|--------------------------|--------------------------|
| <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Not at all | Rarely | Occasionally | Often |

28. Does this person destroy property?

- | | | | |
|--------------------------|--------------------------|--------------------------|--------------------------|
| <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Not at all | Rarely | Occasionally | Often |

29. Does this person behave offensively (includes sexual behaviour)?

- | | | | |
|--------------------------|--------------------------|--------------------------|--------------------------|
| <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Not at all | Rarely | Occasionally | Often |

30. Does this person have habits or behaviours that most people find unsociable (e.g. spitting, leaving lighted cigarette butts around, messing up the toilet, messy eating)?

Not at all Rarely Occasionally Often

31. Does this person lose personal property?

Not at all Rarely Occasionally Often

32. Does this person invade others' space (rooms, personal belongings)?

Not at all Rarely Occasionally Often

33. Does this person take things which are not his or hers?

Not at all Rarely Occasionally Often

34. Is this person violent to others?

Not at all Rarely Occasionally Often

35. Is this person violent to him or her self?

Not at all Rarely Occasionally Often

36. Does this person get into trouble with the police?

Not at all Rarely Occasionally Often

37. Does this person abuse alcohol or other drugs?

Not at all Rarely Occasionally Often

38. Does this person behave irresponsibly?

- | | | | |
|--------------------------|--------------------------|--------------------------|--------------------------|
| <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Not at all | Rarely | Occasionally | Often |

39. Does this person generally make and/or keep up friendships?

- | | | | |
|----------------------------------|--|--|--------------------------------------|
| <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Friendships made or kept up well | Friendships made or kept up with slight difficulty | Friendships made or kept up with considerable difficulty | No friendships made or none kept up. |

HEALTH OF THE NATION OUTCOMES SCALES

(Example Only)

Instructions:

Please rate your client on each of the **twelve** subscales, according to the following guidelines:

1. **Rate** each scale in order from 1 to 12 scales.
2. **Do not** include information rated in an earlier item.
3. Rate the **MOST SEVERE** problem that has occurred over the previous **TWO WEEKS**.
4. All scales follow the format
 - 0= no problem
 - 1=minor problem requiring no action
 - 2=mild problem but definitely present
 - 3=moderately severe problem
 - 4=severe to very severe problem
5. Write the number corresponding to your response in the “Rating” box for each scale.

1. Overactive, aggressive, disruptive behaviour

Include such behaviour due to any cause e.g. drugs, alcohol, dementia, psychosis, depression, etc.

Do not include bizarre behaviours, rated at Scale 6.

0. No problems of this kind during the period rated.
1. Some irritability, quarrels, restlessness, etc.
2. Includes occasional aggressive gestures, pushing or pestering others; threats or verbal aggression; lesser damage to property (e.g. broken cup, window); marked overactivity or agitation.
3. Physically aggressive to others or animals (short of rating 4); persistently threatening manner; more serious overactivity or destruction of property.
4. At least one serious physical attack on others or on animals; destructive of property (e.g. fire setting); persistent serious intimidation or obscene behaviour.
9. Not known or not applicable.

Rating :

2. Non-accidental self-injury

Do not include accidental self-injury (due e.g. to dementia or severe learning disability); the cognitive problem is rated at Scale 4 and the injury at Scale 5.

Do not include illness or injury as a direct consequence of drug/alcohol use rated at Scale 3; (e.g. cirrhosis of the liver or injury resulting from drunk driving are rated at Scale 5).

0. No problem of this kind during the period tested.
1. Fleeting thoughts about ending it all but little risk; no self-harm.
2. Minor risk during period; includes non-hazardous self-harm, e.g. wrist-scratching.
3. Moderate to serious risk of deliberate self-harm; includes preparatory acts e.g. collecting tablets.
4. Serious suicidal attempt and/or serious deliberate self-injury during period.
9. Not known or not applicable.

Rating :

3. Problem-drinking or drug-taking

Do not include aggressive/destructive behaviour due to alcohol or drug use, rated at Scale 1.

Do not include physical illness or disability due to alcohol or drug use, rated at Scale 5.

0. No problem of this kind during the period rated.
1. Over-indulgence but within social norm.
2. Loss of control of drinking or drug-taking, but not seriously addicted.
3. Marked dependence on alcohol or drugs with frequent loss of control, drunk driving, etc.
4. Incapacitated by alcohol/drug problems.
9. Not known or not applicable.

Rating :

4. Cognitive problems

Include problems of memory, orientation and understanding associated with any disorder: learning disability, dementia, schizophrenia, etc.

Do not include temporary problems (e.g. hangovers) resulting from drug/alcohol abuse, rated at Scale 3.

0. No problem of this kind during the period rated.
1. Minor problems with memory or understanding, e.g. forgets names occasionally.
2. Mild but definite problems, e.g. has lost the way in a familiar place or failed to recognise a familiar person, sometimes mixed up about simple decisions.
3. Marked disorientation in time, place or person, bewildered by everyday events; speech is sometimes incoherent, mental slowing.
4. Severe disorientation, e.g. unable to recognise relatives, at risk of accidents, speech incomprehensible; clouding or stupor.
9. Not known or not applicable.

Rating :

5. Physical illness or disability problems

Include illness or disability from any cause that limits or prevents movement, or impairs sight or hearing, or otherwise interferes with personal functioning.

Include side-effects from medication; effects of drug/alcohol use; physical disabilities resulting from accidents or self-harm associated with cognitive problems, drink-driving, etc.

Do not include mental or behavioural problems rated at Scale 4.

0. No physical health problem during the period rated.
1. Minor health problem during the period (e.g. cold, non-serious fall, etc.).
2. Physical health problem imposes mild restriction on mobility and activity.
3. Moderate degree of restriction on activity due to physical health problem.
4. Severe or complete incapacity due to physical health problem.
9. Not known or not applicable.

Rating :

6. Problems associated with hallucinations and delusions

Include hallucinations and delusions irrespective of diagnosis.

Include odd and bizarre behaviour associated with hallucinations or delusions. Do not include aggressive, destructive or overactive behaviours attributed to hallucinations or delusions, rated at Scale 1.

0. No evidence of hallucinations or delusions during the period rated.
1. Somewhat odd or eccentric beliefs not in keeping with cultural norms.
2. Delusions or hallucinations (e.g. voices, visions) are present but there is little distress to patient or manifestation in bizarre behaviour, i.e. clinically present but mild.
3. Marked preoccupation with delusions or hallucinations, causing much distress and/or manifested in obviously bizarre behaviour, i.e. moderately severe clinical problem.
4. Mental state and behaviour is seriously and adversely affected by delusions or hallucinations, with severe impact on patient and/or others.
9. Not known or not applicable.

Rating :

7. Problems with depressed mood

Do not include overactivity or agitation, rated as Scale 1.

Do not include suicidal ideation or attempts rated at Scale 2. Do not include delusions or hallucinations, rated at Scale 6.

0. No problems associated with depressed mood during the period rated.
1. Gloomy; or transient mood changes.
2. Mild but definite depression and distress: e.g. feelings of guilt; sleep disturbance; loss of appetite; loss of self-esteem.
3. Depression with inappropriate self-blame, preoccupied with feelings of guilt.
4. Severe or very severe depression, with guilt or self-accusation.
9. Not known or not applicable.

Rating :

8. Other mental and behavioural problems

Rate only the most severe clinical problem not considered at items 6 and 7 as follows. Specify the type of problem by entering the appropriate letter: A phobic; B anxiety; C obsessive-compulsive; D stress; E dissociative; F somatoform; G eating; H sleep; I sexual; J other, Specify_____.

0. No evidence of these problems during period rated.
1. Minor non-clinical problems.
2. A problem is clinically present, but there are relatively symptom-free intervals and patient/client has a degree of control, i.e. mild level.
3. Constant preoccupation with problem. Occasional severe attack or distress, with loss of control (e.g. has to avoid anxiety provoking situations altogether, call in a neighbour to help, etc.) i.e. moderately severe level of problem.
4. Severe persistent problem dominates most activities.
9. Not known or not applicable.

Rating :

9. Problems with relationships

Rate the patient's most severe problem associated with active or passive withdrawal from social relationships, and/or non-supportive, destructive or self-damaging relationships.

0. No significant problems during the period.
1. Minor non-clinical problem.
2. Definite problems in making or sustaining supportive relationships; patient complains and/or problems are evident to others.
3. Persisting major problems due to active or passive withdrawal from social relationships, and/or relationships that provide little or no comfort or support.
4. Severe and distressing social isolation due to inability to communicate socially and/or withdrawal from social relationships.
9. Not known or not applicable.

Rating :

10. Problems with activities of daily living

Rate the overall level of functioning in activities of daily living (ADL): e.g. problems with basic activities of self-care such as eating, washing, dressing, toilet also complex skills such as budgeting, organising where to live, occupation and recreation, mobility and use of transport, shopping, self-development, etc.

Include any lack of motivation for using self-help opportunities, since this contributes to a lower overall level of functioning.

Do not include lack of opportunities for exercising intact abilities and skills, rated at Scales 11-12.

0. No problems during period rated; good ability to function in all areas.
1. Minor problems only; e.g. untidy, disorganised.
2. Self-care adequate, but major lack of performance of one or more complex skills (see above).
3. Major problems in one or more area of self-care (eating, washing, dressing, toilet) as well as major inability to perform several complex skills.
4. Severe disability or incapacity in all or nearly all areas of self-care and complex skills.
9. Not known or not applicable.

Rating :

11. Problems with living conditions.

Rate the most severe problem with the quality of living conditions and daily domestic routine.

Are the basic necessities met (heat, light, hygiene)? If so, is there help to cope with disabilities and a choice of opportunities to use skills and develop new ones?

Do not rate the level of functional disability itself, rated at Scale 10.

N.B. Rate patient's usual accommodation. If in acute ward, rate the home accommodation. If accommodation not available, rate 9.

0. Accommodation and living conditions are acceptable; helpful in keeping any disability rated at Scale 10 to the lowest level possible, and supportive of self-help.
1. Accommodation is reasonably acceptable although there are minor or transient problems (e.g. not ideal location, not preferred option, doesn't like the food, etc.).
2. Significant problems with one or more aspects of the accommodation and/or regime: e.g. restricted choice; staff or household have little understanding of how to limit disability, or how to use or develop new or intact skills.
3. Distressing multiple problems with accommodation: e.g. some basic necessities absent; housing environment has minimal or no facilities to improve patient's independence.
4. Accommodation is unacceptable: e.g. lack of basic necessities, patient is at risk of eviction, or 'roofless', or living conditions are otherwise intolerable and making patient's problems worse.
9. Not known or not applicable.

Rating :

12. Problems with occupation and activities.

Rate the most severe problem with quality of day-time environment. Is there help to cope with disabilities, and opportunities for maintaining or improving occupational and recreational skills and activities? Consider factors such as stigma, lack of qualified staff, access to supportive facilities, e.g. staffing and equipment of day care centres, workshops, social clubs, etc.

Do not rate the level of functional disability itself, rated at Scale 10.

N.B. Rate the patient's usual situation. If in acute ward, rate activities during period before admission. If information not available, rate 9.

0. Patient's day-time environment is acceptable: helpful in keeping any disability rated at Scale 10 to the lowest level possible, and supportive of self-help.
1. Minor or temporary problems e.g. late social security or health benefit payments; reasonable facilities available but not always at desired times, etc.
2. Limited choice of activities: e.g. there is a lack of reasonable tolerance (e.g. unfairly refused entry to public library or baths, etc.); or handicapped by lack of a permanent address; or insufficient carer or professional support; or helpful day setting available but for very limited hours.
3. Marked deficiency in skilled services available to help minimise level of existing disability; no opportunities to use intact skills or add new ones; unskilled care difficult to access.
4. Lack of any opportunity for day time activities makes patient's problems worse.
9. Not known or not applicable.

Rating :

The Role Functioning Scale

(Example Only)

Please read the following descriptions for each area of role functioning and give a rating to indicate the most appropriate description for your client by placing a tick in the corresponding box. Rate the client for their role functioning over the PAST MONTH.

1. Working Productivity

(Rate the client primarily in the most expected role: i.e., homemaker, student, wage earner.)

- Productivity severely limited; often unable to work or adapt to school or homemaking; virtually no skills or attempts to be productive.
- Occasional attempts at productivity unsuccessfully; productive only with constant supervision in sheltered work, home or special classes.
- Limited productivity; often with restricted skills/abilities for homemaking, school, independent employment (e.g. requires highly structured routine).
- Marginal productivity (e.g. productive in sheltered work or minimally productive in independent work; fluctuates at home, in school; frequent job changes).
- Moderately functional in independent employment, at home, or in school. (Consider very spotty work history or fluctuations in home, in school with extended periods of success).
- Adequate functioning in independent employment, home or school; often not applying all available skills/abilities.
- Optimally performs homemaking, school tasks, or employment related functions with ease and efficiency.

2. Independent Living, Self Care

(Management of household, eating, sleeping, hygiene care)

- Lacking self-care skills approaching life endangering threat; often involves multiple and lengthy hospital services; not physically able to participate in running a household.
- Marked limitations in self-care/independent living; often involving constant supervision in or out of protective environment (e.g. frequent utilisation of crisis services).
- Limited self-care/independent living skills; often relying on mental/physical health care; limited participation in running household.
- Marginally self-sufficient; often uses REGULAR assistance to maintain self-care/independent functioning; minimally participates in running household.
- Moderately self-sufficient; i.e., living independently with ROUTINE assistance (e.g. home visits by nurses, other helping persons, in private or self-help residences).
- Adequate independent living and self-care with MINIMAL support (e.g. some transportation, shopping assistance from neighbours, friends, other helping persons).
- Optimal care of health/hygiene; independently manages to meet personal needs and household tasks.

3. Immediate Social Network Relationships

(Close friend, spouse, family.)

- Severely deviant behaviours within immediate social networks (i.e. often with imminent physical aggression or abuse to others or severely withdrawn from close friends, spouse, family; often rejected by immediate social network).
- Marked limitations in immediate interpersonal relationships (e.g., excessive dependency or destructive communication or behaviours).
- Limited interpersonally; often no significant participation/communication with immediate social network.
- Marginal functioning with immediate social network (i.e. relationships are often minimal and fluctuate in quality).
- Moderately effective continuing and close relationship with at least one other person.
- Adequate personal relationship with one or more immediate members of social network (e.g., friend or family).
- Positive relationships with spouse or family and friends; assertively contributes to these relationships.

4. Extended Social Network Relationships

(Neighbourhood, community church, clubs, agencies, recreational activities.)

- Severely deviant behaviours within extended social networks (i.e., overtly disruptive, often leading to rejection by extended social networks).
- Often totally isolated from extended social networks, refusing community involvement or belligerent to helpers, neighbours, etc.
- Limited range of successful and appropriate interactions in extended social networks (i.e., often restricts community involvement to minimal survival level interactions).
- Marginally effective interactions; often in a structured environment; may receive multiple public system supports in accord with multiple needs.
- Moderately effective and independent in community interactions; may receive some public support in accord with need.
- Adequately interacts in neighbourhood or with at least one community or other organisation or recreational activity.
- Positively interacts in community; church or clubs, recreational activities, hobbies or personal interests, often with other participants.

Appendix 6: Additional Measures Used in the Study

General Participant Information (Consumer) (Example Only)

1. Age:
2. Gender: Male Female
3. Were you born in Australia? Yes No
4. If yes, do you belong to the Aboriginal Torres Strait Island cultures? Yes No
5. If No, where were you born:
 New Zealand United Kingdom Asia
 Europe Pacific Islands Americas
 Africa
6. What is your first language: English Other
Specify: _____
7. What is your present marital status:
 Married In a de facto relationship Divorced
 Widowed Never Married.
8. Employment:
 Currently employed Unemployed
 Full-time student Homemaker
9. If a student, please specify:
 High school Tertiary (e.g. TAFE, University).
10. Nature of employment (*if employed*):
11. If you are unemployed, how long has it been since you were last employed?
 less than one month 1-3 months 3-6 months
 6-12 months 1-2 years More than 2 years

Well-Being Rating
(Revised)
(Example Only)

Please answer the following questions by placing a mark on the horizontal line in the position that you believe best represents how you feel.

1. How would you rate your **mental well-being** in the past 7 days?

This is the _____ This is the **best** I have
worst I have ever felt in my
ever felt in my life. ever felt in my life.

2. In the last seven days how **depressed** have you felt?

This is the _____ This is the **most**
least depressed I have ever felt
have ever felt in my life. **depressed** I have ever
in my life. felt in my life.

3. In the last seven days how **anxious** have you felt?

This is the _____ This is the **most**
least anxious I have ever felt
have ever felt in my life. **anxious** I have
in my life. ever felt in my life.

4. In the last seven days how **distressed** have you felt?

This is the _____ This is the **most**
least distressed I have ever felt
have ever felt in my life. **distressed** I have
in my life. ever felt in my life.

Global Change Scale
(Consumer)
(Example Only)

Rate the total change that you have experienced in your *mental health* in the past 3 months (that is, since we last surveyed you).

Compared to your condition three (3) months ago how much have changed? (*tick one of the boxes*)

- Extremely Better:** I feel *extremely* better than I did 3 months ago.
- Much Better:** I feel *much* better than I did 3 months ago.
- Slightly Better:** I feel only *slightly* better than I did 3 months ago.
- No Change:** I do not feel either better or worse than I did 3 months ago.
- Slightly Worse:** I feel only *slightly* worse than I did 3 months ago.
- Much Worse:** I feel *much* worse than I did 3 months ago.
- Extremely Worse:** I feel *extremely* worse than I did 3 months ago.

Utility Questionnaire

(Consumer Version) (Example Only)

1. Did you have any of the following problems answering the questions in this questionnaire? (*tick the boxes that are applicable*)

- Some of the words used in the questions were confusing.
- There were too many questions to answer.
- The questions did not relate to how I feel.
- I did not understand the purpose of the some of the questions.
- Other (give one or two):

2. Were the words used in the questions difficult to understand?

- None** **Some** **Most** **All**

3. Did the questions ask you things that are important to you?

- None** **Some** **Most** **All**

4. Did the questions ask you things that relate to how you feel?

- None** **Some** **Most** **All**

5. Is this questionnaire a useful way for your doctor or specialist to find out whether you are getting better?

- None** **Some** **Most** **All**

6. Other Comments:

General Information Section
(Service Provider)
(Example Only)

Service Provider Information:

1. I.D.(*to be completed by researcher*):
2. Professional description:
 - General Practitioner Psychiatrist (Private Practice)
 - Psychiatrist (Public Hospital/Community Clinic)
 - Psychologist Social Worker
 - Nurse Other. Specify:

Client Information:

1. I.D. (*to be completed by researcher*):
2. To which of the following diagnostic categories does your client belong?
 Schizophrenia Affective Disorders Anxiety Disorders
3. What is this person's primary diagnostic condition?
Specify:

4. Is there a secondary mental health diagnosis (i.e. comorbidity)?
Specify:

Modified Clinical Global Impressions
(Service Provider)
(Example Only)

1. Severity of Illness Scale

Considering your total clinical experience with this particular population, *how mentally ill* is the patient at this time? (*tick one of the boxes*)

- Normal, not at all ill.
- Borderline mentally ill.
- Mildly ill.
- Moderately ill.
- Markedly ill.
- Severely ill
- Among the most extremely ill patients.

2. Global Change Scale

Please rate the total change in your client's condition (that is, the *primary mental health condition* for which you are seeing this person) in the past three months (that is, since you were last surveyed for this study with regard to this person).

Compared to your client's condition three months ago how much has he/she changed?
(*tick one of the boxes*)

- Very much improved:** there has been an *very substantial reduction* (if not a complete eradication) in the range and severity of the client's symptoms, and no presentation of new symptoms.
- Much improved:** there has been a *substantial reduction* in the range and severity of the client's symptoms.
- Minimally improved:** there has been *some progress* in reducing either the range or severity of the client's symptoms.
- No change:** there has been no alteration in either the range or the severity of the client's symptoms: the client's condition has remained stable.
- Minimally worse:** there has been *some deterioration* with regard to either the range or severity of the client's symptoms.
- Much worse:** there has been a *substantial deterioration* with regard to either the number or severity of the client's symptoms.
- Very much worse:** there has been a *very substantial deterioration* with regard to either the severity of the client's symptoms which may or may not be complicated by new symptoms.

3. How many times have you had contact (e.g. face-to-face, telephone) with this person, in relation to their mental illness over the last three months?

Response:

**Utility Questionnaire
(Service Provider Version)
(Example Only)**

Please read and answer the following questions about the questionnaire that you have just completed.

1. Were the words used in the questions difficult to understand?

<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
None	Some	Most	All

2. Were the questions relevant to measuring the treatment outcomes for your client?

<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
None	Some	Most	All

3. Were the questions relevant to measuring the treatment outcomes for people with this particular diagnosis (i.e. that of your client) in general?

<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
None	Some	Most	All

4. Were the questions relevant to measuring treatment outcomes for all people with mental illness?

<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
None	Some	Most	All

5. Is this questionnaire an effective way of measuring treatment outcomes for your client?

<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Not Effective at all	Slightly Effective	Reasonably Effective	Very Effective

6. Is this questionnaire an effective way of measuring treatment outcomes for people with this particular diagnosis (i.e. that of your client) in general?

<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Not Effective at all	Slightly Effective	Reasonably Effective	Very Effective

7. Is this questionnaire an effective way of measuring treatment outcomes for all people with mental illness?

<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Not Effective at all	Slightly Effective	Reasonably Effective	Very Effective

8. Is this questionnaire a useful way of measuring treatment outcomes for all people with mental illness?

<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Not Useful at all	Slightly Useful	Reasonably Useful	Very Useful



**Diagnostic Screen- Primary Health Care Criteria for ICD-10
(Example Only)**

This section to be completed by the researcher.

Date:

Consumer I..D.:

Name:

The client that you have nominated for inclusion in this study has been diagnosed with **Depression**. Please take a few moments to read the following checklist.

Depression

Clinical features

Low or sad mood.

Loss of interest or pleasure.

Associated symptoms are frequently present:

- disturbed sleep
- guilt or low self-worth
- fatigue or loss of energy
- poor concentration
- disturbed appetite
- suicidal thoughts or acts

Movement and speech may be slowed, but agitation may also be present.

Symptoms of anxiety or nervousness are also frequently present.

In severe cases, mood congruent hallucinations or delusions may be present.

Is your patient's clinical picture consistent with the above?

Yes No

