PSYCHOMETRIC VALIDATION OF THE

SOUTH AFRICAN ADDICTION TREATMENT ASSESSMENT

MEASUREMENT INSTRUMENT (SAATSA)

PILOT SECOND ITERATION

FINAL REPORT

5 August 2013

Prof. Rajen Govender

KhyAna

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Glossary of Abbreviations

CFA	Confirmatory Factor Analysis
CTT	Classical Test Theory
CFI	Comparative Fit Index
EFA	Exploratory Factor Analysis
IRT	Item Response Theory
MAR	Missing At Random
ML	Maximum Likelihood
MMT	Modern Measurement Theory
MRC	Medical Research Council
MVA	Missing Values Analysis
OBLIMIN	Oblique Minimum
RMSEA	Root Mean Square Error of Approximation
SAATSA	South African Substance Abuse Treatment Assessment
SACENDU	South African Community Epidemiological Network on Drug Use
SEM	Structural Equations Model/Modelling
SQM	Service Quality Measures

CHAPTER1: INTRODUCTION AND OVERVIEW

This report presents the analysis of the second iteration of the pilot testing of the South African Addiction Treatment Service Assessment measurement instrument (SAATSA). The report follows on from the report on the first iteration of the pilot. The first report provided a detailed background to the Services Quality Measures (SQM) initiative as well to various aspects of sampling, measurement theory and methods of analysis. This report omits such detail as it focuses specifically on the psychometric validation of the SAATSA scales as revised following the first pilot iteration. Readers are encouraged to refer to the earlier report for relevant background detail.

First Iteration Pilot Results

The purpose of the second iteration of the SAATSA pilot was to address the methodological and measurement shortcomings identified in the first iteration. These related variously to 1) the fieldwork, data capture and data vetting processes, 2) the collection and capture of relevant background demographic and administrative data, and 3), the measurement of individual SAATSA items and SAATSA scales.

The key findings of the first pilot iteration may be summarised as follows:

1. **Background Information** – background information on patients comprises demographic detail such as race, age, gender, education, etc., and administrative data such as start of treatment, length of treatment, admission type at facility (inpatient/outpatient), prior treatment received, etc. The first iteration obtained this information directly from the respondents, and the analysis noted several problems as follows:
   - There were several concerns about the veracity of this information as supplied by respondents. This was considered problematic as these variables served as essential entry variables into validation and substantive analyses.
   - The analysis indicated problems with high proportions of missing values for many of the essential background variables, a factor which also restricted the application of these variables in the validation process.
   - Despite language being a critical consideration on the development of the SAATSA, there was no measure of primary home language/language spoken most often. Race was used as a proxy for language though it remained an imperfect proxy given the extent of bilingualism/multilingualism in the general population. This had implications for the validation process.
   - The variable measuring educational attainment was in need of revision to ensure better correspondence to standard measurement thereof in official statistics in the country.

2. **SAATSA Items** – the individual SAATSA items presented some problems in terms of the following:
    - Some items were in need of revision of phrasing/expression to ensure better comprehension by respondents as well as to ensure greater correspondence to the SAATSA scales to which they purportedly belonged.
    - All of the SAATSA items contained a "Not Applicable" category as part of their response format. However, for the majority of items this response category was neither conceptually valid nor relevant. As a result, respondent responses using this category had to be coded as missing, resulting in an inflation of the proportion of missing values per item and the exclusion of a number of items and cases from the validation process.

3. **SAATSA Scales** – analysis of the SAATSA scales revealed the following issues for attention:
    - There was an insufficient number of items to estimate and measure some of the scales, in particular the Person Centeredness Scale. Accordingly, this scale was collapsed into the Quality Scale.
    - There was no proper scale to measure Access to Treatment. This was considered a critical oversight in a country where research has indicated considerable inequities in the health system and where access to such treatment is manifestly skewed depending on education, income and spatial location and other variables.

4. **Fieldwork and Data Processing** – the first iteration identified the need for proper protocols for the backchecking of returned surveys and quality management of the data capture and processing activities.

The Revised SAATSA

The results of the first iteration led to the following essential revisions in the second iteration:

1. **Background Information** – Background demographic and administrative data on respondents was collected via the South African Community Epidemiological Network on Drug Use (SACENDU) process rather than from respondents themselves. The SACENDU has been operational since 1996 and collects patient data at treatment centres by clinical staff. It is thus far more accurate and robust than information collected from respondents. In addition, the education and home language variables were revised for the second iteration.
2. **SAATSA Items** – individual SAATSA items identified for improvement were revised in terms of language and expression and subject to a round of cognitive testing before finalisation and inclusion in the SAATSA. All items were also revised by excluding the "Not Applicable" category from their response formats except in those instances where it was conceptually valid and relevant.
3. **SAATSA Scales** – The Access to Treatment Scale was improved by development of new items. These were also subject to cognitive testing for review and refinement purposes.

4. **Fieldwork and Data Processing** – all recommendations in this regard were implemented to ensure improved protocols both in the field and for data capture and processing.

The gains made in the SAATSA as a result of these improvements are discussed in greater detail in the relevant sections of this report.

The SAATSA HIV Scale

During the course of the second iteration of the pilot it was acknowledged that the SAATSA lacked sufficient focus on HIV awareness and prevention issues. These issues were deemed critical for investigation given the nexus between risky sexual behaviour and substance abuse as observed in research in the country, as well as the emergent focus on HIV education and awareness in a number of current treatment programmes. Accordingly, a HIV Scale was constituted using items from existing measures. These items were not subject to cognitive testing due to time constraints but were deemed appropriate for piloting given their previous application in South Africa. The scale was introduced into the SAATSA mid-stream in the fieldwork process, and data was collected from only a portion of the obtained sample. However, the returned data comprised a sufficient number of cases to enable psychometric validation of these items and the scale. The revised SAATSA instrument and the HIV addendum as deployed in the field are attached in Appendix 1.

Organisation of the Report

The report is organised as follows:

Chapter 2: Sampling and Fieldwork – this chapter provides an overview of the sampling and fieldwork issues for the second iteration of the SAATSA pilot.

Chapter 3: Data Validation and Univariate Analysis – this chapter examines the integrity of the SAATSA items by investigating their completeness (using Missing Values Analysis) and univariate normality.

Chapter 4: Sample Characteristics – this chapter summarises the key descriptive characteristics of the second iteration sample. This is provided purely for descriptive and not inferential purposes.

Chapter 5: Psychometric Validation of the SAATSA Scales - this chapter provides the detailed psychometric analysis of the SAATSA scales as tested in the second iteration. Comparisons are also provided where applicable with the analysis from the first pilot iteration.

Chapter 6: Conclusion and Recommendations - this chapter summarises the principal findings of the analysis and presents the various recommendations of the second pilot iteration.

CHAPTER 2: SAMPLING AND FIELDWORK

The sampling and fieldwork for the second pilot iteration followed the logical framework of the first iteration, with implementation of appropriate revisions and improvements where relevant and necessary based on the learning from the first iteration. All fieldwork began in February 2013 and was concluded in June 2013.

Sampling Criteria

Consistent with the first iteration, the sample for the second iteration was stratified according to the following criteria:

1. Facility Administration – Both state owned and operated and non-state owned and operated facilities were selected for the study.
2. Admission Type – To account for the differential nature and characteristics of substance abuse presenting symptoms and treatment programmes in the inpatient and outpatient admission of patients, the sample was structured to capture data for both admission types.
3. Patient Demographics – As far as possible, the sample of facilities accounted for the variable demographic characteristics of the national population and the spatial distribution of these characteristics.

Sampling Design

The second pilot iteration employed purposive sampling, with treatment facilities selected for inclusion based on the access afforded to the research team by relevant organisational/institutional administration and facility staff. Access to the various facilities was negotiated and obtained by the Medical Research Council either by way of the relevant provincial authorities in the case of government owned/administered facilities or the relevant organisational administration in the case of non-governmental facilities and/or facilities owned by the state but operated on an agency basis.

As with the first iteration, the specific sampling strategy was motivated by three primary considerations:

1. Access – systematic and rigorous service quality assessment in the treatment of substance abuse is a nascent development in South Africa. Consequently, there was a requirement for significant effort by the SQM to ensure the co-operation and participation of administrators and practitioners and their respective facilities. To circumvent the protracted communication and media process that would be necessary to secure access at facilities generally, it was decided to engage with organisations and facilities which had provided access in the first iteration.
2. Budget - for reasons of a limited budget, the sampling was restricted to facilities located in two major metropolitan areas in two provinces in the country, these being the Durban and

Pietermaritzburg metropolitan areas in KwaZulu-Natal and the Cape Town metropolitan area in the Western Cape.

3. Logistical Support - these locations were chosen because of the presence of the Medical Research Council in these areas, which ensured proper logistical support for the fieldwork process and research personnel.

The purposive sampling strategy was deemed acceptable for the second iteration as the focus of the research was to conclude the psychometric validation of the SAATSA items and scales and not to generalise any of these findings to treatment centres generally in the country. Accordingly, the second iteration data is in no way representative - nor presented as being representative - of the adult population of substance abuse treatment facilities in either of these provinces or the country as a whole.

Fieldwork And Questionnaire Administration

Fieldwork was once again undertaken by personnel from the MRC, with assistance from staff at the various treatment facilities. Once access had been approved and provided, site visits were undertaken by research personnel to brief the facility staff on the research and the survey and to prepare the facility for the administration of the survey questionnaire, including where necessary the provision of training to staff. All SAATSA surveys were self-administered, while demographic and administrative data was collected directly by clinical staff/clinicians through the established SACENDU process and protocols. The removal of the collection of background information from the SAATSA instrument greatly rationalised the data collection process.

The protocol for selection of patients for the second iteration remained consistent with the previous iteration: patients were identified and directly approached by facility staff and research personnel. The principal requirement was that patients would have been in the facility for a sufficient period of time to make informed judgements about service quality. As the treatment regimens varied in duration at different facilities, it was decided that patients would be targeted only after they had completed the majority portion (70-80% or more) of their designated treatment programme. For those in 16 week programmes, this meant having completed at least 12 weeks, while those in 5 week programmes were required to have completed 4 weeks and those in three week programmes were required to have completed more than two weeks. Further, it was decided that all patients completing the survey would still be enrolled in their treatment programmes (either on an inpatient or outpatient basis) and would not have completed the programme. This was considered important as the assessment of service quality can and does vary considerably based upon whether or not patients are still enrolled or have already left the programme/facility. Treatment facilities furnished all completed questionnaires immediately to the MRC or stored them on their premises for later submission to the MRC. The MRC captured all questionnaires electronically and developed the dataset for analysis of the pilot study in SPSS version 20.

CHAPTER 3: DATA VALIDATION AND UNIVARIATE ANALYSIS

This chapter summarises the base analyses undertaken to ensure that individual SAATSA items met the necessary requirements for inclusion in the psychometric validation analysis. The second iteration began with a base sample of 334 cases.

Overview of Methods

Data validation and univariate analysis for the second iteration followed the methods as applied in the first iteration, these being:

1. **Missing Value Analysis** - This aspect of the analysis focuses on understanding the distributions of individual items in terms of the proportion of missing values for each of these items. The objective is to confirm whether or not such values are missing at random (MAR), in which case the effect on the univariate distribution of variables is normalised and the impact on the variable distribution is minimal. In these instances the missing values are accepted at face value and those cases for which they are applicable are legitimately excluded from any analysis. In contrast, when values are missing systematically, this reflects specific errors or bias in methodology and measurement. If such biases or errors are present, the process considers how to resolve them in advance of the bivariate and multivariate analysis required for psychometric validation purposes.

2. **Univariate Normality** – Univariate normality is a precursor to multivariate normality, an essential requirement for psychometric validation. To ensure applicability in psychometric validation, individual items must therefore demonstrate normality in the distribution of values. Assessments of univariate normality are generally undertaken by examining distribution characteristics such as skewness and kurtosis. Skewness refers to the extent to which the values of the distribution are clustered towards one or the other end of the distribution, while kurtosis refers to the height of the distribution, that is, the extent to which the values are spread out (a flat distribution) or bunched together (a tall distribution). In the real world study variables rarely achieve perfect univariate normality. Fortunately, current analytic techniques are forgiving of variations from true normality, and thus variables may be included even when their distributions reveal non-significant skewness and kurtosis characteristics.

Missing Values Analysis

As with the first iteration, the missing values analysis focused principally on the individual SAATSA items as these are the constituent elements of the SAATSA scales, though it also examined the background demographic and administrative data as these are important to understanding whether

or not the SAATSA scales perform differentially in different sub-populations of the treatment population.

To give effect to the missing values analysis it was decided to employ a generally accepted cut-off criterion: items which have missing values for 5% or more of respondents were deserving of proper consideration, while proportions of missing values less than 5% were regarded as being acceptable and not necessarily deserving of further attention. In all instances, however, the missing values needed to be normally distributed, that is, missing at random, though this requirement is optional for proportions of missing values well below 5% of respondents.

SAATSA/SACENDU Link

The second iteration linked the SAATSA to the SACENDU data collection for the first time, and the MVA consequently began by examining all 334 cases of the base sample to confirm the presence of both data. This analysis revealed 35 cases for which SACENDU data was present but for which SAATSA data was absent. It is likely that these sites had already collected the SACENDU data (at intake) and were awaiting sufficient time as per the research protocols to administer the SAATSA. However, the fieldwork process had to be terminated in order to enable analysis of the data in preparation for a meeting of the SQM National Steering Committee, and hence these cases had SACENDU and not SAATSA data. Consequently, these cases were permanently removed from the sample.

SAATSA Items

A critical measurement shortcoming identified in the first iteration was the presence of the "Not Applicable" response category for all the SAATSA items. When these values were correctly recoded as missing for items where this response category was not conceptually valid or relevant, the result was a significant increase in the proportion of missing values for at least one third of the SAATSA items. In the current iteration this category was only applied for a minority of items (3) where it was deemed to be conceptually valid and relevant. As a result, there was no need to recode any Not Applicable responses as missing except where relevant.

The distribution of the missing values for all of the SAATSA items is presented in Table 1. The table lists the proportions of the sample that had missing values for each of the items in question. The proportions for the HIV items (Item 32 to Item 40) are applicable only to the sub-sample from which this data was collected (130 cases).

Table 1: Proportions of the Sample with Missing Values on SAATSA Items

SAATSA ITEMS	% With Missing Values	
	PILOT 1	PILOT 2
The amount of time I had to wait to get services was acceptable to me.	5.0%	2.0%
I can afford the treatment I want to receive.	14.0%	44.0%
The location of this treatment centre is convenient for me.	-	2.4%
I feel safe travelling to this treatment centre.	-	2.4%
It is easy for me to obtain the treatments offered by this centre.	-	2.4%
My family is able to access services provided by this treatment centre.	-	4.7%
I can afford the transport costs of getting to this treatment centre	-	2.7%
The staff at this treatment centre treat me with respect.	2.3%	2.0%
The people I went to for treatment services at this centre spent enough time with me.	3.1%	1.3%
I have a say in deciding about my substance abuse treatment that I am receiving here.	7.5%	2.0%
The staff told me about services in my area that will help me stay off drugs and alcohol.	3.6%	2.4%
This treatment centre teaches me how to avoid getting HIV.	12.1%	3.0%
The staff at this treatment centre are sensitive to my background.	4.7%	2.4%
My general health is improving.	3.6%	1.0%
I am better able to cope when things go wrong.	4.5%	1.0%
I am better able to accomplish the things I want to do.	5.0%	2.0%
I am less likely to use alcohol or other drugs.	6.1%	1.7%
In the future, I will be more likely to do better at work or at school.	6.1%	15.6%
I am more likely to practice safe sex.	11.5%	3.0%
There is someone who cares about whether I am doing better.	5.0%	0.7%
I have someone who will help me when I have a problem.	4.7%	1.3%
I have people in my life who are a positive influence.	5.3%	1.0%
The people who care about me are supportive of my treatment.	5.6%	1.3%
My friends and family are able to count on me more	5.0%	10.4%
I have friends who are not using alcohol or drugs.	7.5%	2.0%
I have someone who will listen to me when I need to talk.	2.3%	1.3%
I now know that using alcohol and drugs is a problem for me.	2.3%	0.7%
I need to work on my problems with alcohol and/or drugs.	3.4%	1.7%
The treatment centre is helping me to recover from using drugs and alcohol.	2.0%	1.3%
I would recommend this treatment centre to a friend.	3.4%	1.0%
I am more likely to use condoms when I have sex	-	5.4%
I am less likely to use alcohol and drugs before I have sex	-	1.5%
I am less likely to share drug equipment (e.g. pipes and injections)	-	2.3%
I am more likely to only have sex with one partner	-	1.5%
I am more likely to use condoms properly	-	5.4%
I know it is important to know my HIV status	-	1.5%
I know where I can get tested for HIV	-	1.5%
I know where I can get treatment for HIV	-	2.3%

As is evident from Table 1, the number of SAATSA items where the proportion of respondents with missing values exceeds 5% is quite small. The highest recorded proportion of missing values is for Item 2, Item 18 and Item 24. These are the only three SAATSA items which retained the "Not Applicable" response category, and it is expected that they would record relatively higher levels of missing values as these "Not Applicable" responses would have been regarded as missing in the analysis. For the rest, the proportions of missing values are well below the 5% threshold, and this is in sharp contrast to the first iteration of the pilot, where over one third of the SAATSA items produced proportions of cases with missing values in excess of this threshold. The significant improvement from the first to the second iteration vindicates the decision to remove the "Not Applicable" response category from all but the three original SAATSA items. Consequently, the data integrity for the SAATSA items is significantly improved, facilitating improved psychometric validation.

Considering the newly constructed HIV scale, the MVA revealed that two of the eight items had missing values in excess of 5% of cases. However, both these are only marginally over the threshold, and given the small sample size for this scale, any slight movement in cases would have produced much larger movements in the proportions of missing values. Hence even in the instance of these two variables there does appear to be reason for concern. More reassuringly, the other six items all exhibit proportions of missing values well below 5%.

Examining missing values by means of individual items is one way of assessing data integrity. An alternate and related way is to examine the missing data by individual cases. This is useful because often respondents who are missing values on some items may be missing them on others as well. This might be because respondents adopted a response set and simply passed over all or most of the SAATSA items, or because they had difficulty understanding many of them. In both instances there would be some systematic bias in the distribution of missing values and this would render the particular respondent ineligible for inclusion in the final sample. This form of MVA analysis in the seconds iteration revealed 14 cases which had missing values for ten or more of the SAATSA items. Based on the available data it was difficult to determine the reasons why these cases had such high proportions of missing values. However, as the proportions were in excess of one quarter of the total number of SAATSA items, these cases were also deleted from the sample. This resulted in a final sample of 285 cases. The proportions of cases with missing values in the final sample is shown in Table 2, with the respective figures from the first iteration.

Table 2: Proportion of Cases with SAATSA Missing Values

SAATSA ITEMS WITH MISSING VALUES	PILOT 1	PILOT 2
	% Of Cases	% Of Cases
Missing for No Items	64.8%	87.3%
Missing for One Item	21.4%	8.8%
Missing for Two Items	8.5%	2.5%
Missing for Three Items	5.2%	1.4%

As Table 2 indicates, the proportion of cases with no missing values on any of the SAATSA items has improved considerably from the first (64.4%) to the second iteration (87.3%), with fewer than 5% of cases in the second iteration having missing values on two or more items as compared to 13.7% in the first iteration. Overall, then, the MVA examining both individual SAATSA items and individual cases indicates considerable improvement from the first iteration.

Demographic and Administrative Data

Analysis of the demographic and administrative data for the finalised sample of 285 cases revealed only two variables for which the proportion of cases with missing values exceeded 5% (see Table 3). In the first instance this related to respondents indicating the source of payment for their treatment (10.9%) and in the second instance this related to whether or not they had been tested for HIV in the past year (7.7%). In the former case it is likely that respondents simply did not know who was footing the bill for their treatment and in the latter case respondents might have considered this item too sensitive and hence elected to not provide any response. In all other instances the proportions of missing values was below 5%, and hence did not warrant any further investigation for evidence of systematic bias.

Table 3: Missing Values for Demographic Variables and Administrative Data

DEMOGRAPHIC/ADMIN VARIABLES	% OF CASES
Gender	2.1%
Race	0.4%
Home Language	0.4%
Age	0.7%
Marital Status	2.1%
Education	1.4%
Employment Status	2.8%
Referral Source	3.9%
Admission Type	2.8%
HIV Test	7.7%
Prior Treatment	4.9%
Payment Source	10.9%

The low rate of missing values for critical demographic and administrative data validates the decision to utilise the existing SACENDU process for data collection. The SACENDU data forms are completed by clinical staff who are much better placed to provide accurate and reliable information. There is little doubt that this has contributed to the overall completeness of these two types of data.

In summary, the MVA of items and cases indicated the need to remove 35 cases due to no SAATSA data being present and a further 14 cases due to high proportions of missing values on SAATSA items. This realised a sample of 285 cases which were cleared for analysis and which constituted the finalised sample for the analyses of all SAATSA items and scales except those relating to HIV, for which the finalised sample was 130 cases.

Univariate Analysis

Having addressed and resolved the issues relating to missing values, the SAATSA items were then subject to univariate analysis to investigate the properties of the distributions for each item. As discussed earlier, univariate normality is a necessary (though not a sufficient) condition for multivariate normality, and thus essential prior to the psychometric analyses. Of particular importance for this stage of the analysis were the following:

1. The likely presence of ceiling effects, which might have posed a threat in terms of diminished variance for multivariate analysis.
2. The extent of deviation from normality of the univariate distributions, which might have rendered items unsuitable for inclusion in multivariate analysis.

These two issues were addressed in turn in the univariate analysis of the individual SAATSA items.

Ceiling Effects

Ceiling effects refer to the phenomenon whereby the distribution of a variable is skewed towards certain (positive) response categories. In the case of the SAATSA the individual items are worded such that higher scores reflect greater agreement and accordingly more positive evaluations. Hence high concentrations of responses in categories such as Strongly Agree indicate uniformly positive evaluations. Should these responses be in disproportionate proportions relative to the other response categories, it might indicate that respondents were unduly biased towards positive evaluations of the specific aspect of service quality being measured by the item.

Ceiling effects are not uncommon in consumer service quality assessment surveys, especially when they are administered while the service is current, as in the case of the SAATSA. Accordingly, it was to be expected that respondents would provide far more positive than negative evaluations, and the only concern was the extent of such skewness towards these positive evaluations (the ceiling effect). Examination of the distribution of individual SAATSA items (see Appendix B) confirmed the generally positive bias in respondent evaluations of the various aspects of service quality and outcomes. Nevertheless, examination of the variance data revealed that all SAATSA items appeared to possess sufficient variation in the response categories, thereby facilitating their inclusion in the psychometric analysis.

Measures of Normality

Normality was assessed by examining for skewness and kurtosis in SAATSA item distributions. Skewness refers to how biased the scores are towards one or the other end of the distribution (the horizontal property – the SAATSA items can be seen to be skewed towards the right), while kurtosis refers to the height of the distribution (the vertical property – how flat or peaky the distribution is). The analysis of skewness and kurtosis statistics is available through SPSS, which provides values for the estimates for each of these as well as the standard errors for these estimates. In both instances the true (absolute value) is zero, and hence any deviation in the positive or negative direction away from zero indicates deviation from normality either horizontally and vertically. While the skewness and kurtosis values are generally more informative when converted to standardised scores, in large samples such as this one (approximately 300 cases) it is more useful to look at the actual values rather than their standardised counterparts as the increase in sample size typically reduces the standard error of the estimate and thus amplifies the size of the standardised skewness and kurtosis estimates. As a general rule, the raw scores for skewness and kurtosis should be between -2 and +2 for one to conclude that the distributions do not deviate substantially from normality.

Inspection of the estimates for the SAATSA items revealed that while there is a clear noticeable positive bias in all items, none of the items exhibited skewness estimates which exceed the specified range of -2 to +2. We thus conclude that these distributions did not deviate significantly from normality. In contrast to the skewness estimates, the kurtosis estimates for a number of SAATSA items exceed the specified range, specifically for Items 19, 28, 30, 38 and 39. Inspection of these items reveals that the kurtosis was at the positive end of the response scale, resulting in these distributions being taller than expected. However, as the skewness scores for these items do not show deviation outside the acceptable range we may conclude that the items do contain sufficient variance to permit their inclusion in the analysis.

Overall, then, the univariate analysis suggested no glaring issues with the normality of the individual items which would necessitate transformations of the data prior to multivariate analyses, and all SAATSA items were cleared for psychometric analysis.

CHAPTER 4: SAMPLE CHARACTERISTICS

As indicated previously, sampling was purposive and based on negotiated access to treatment facilities. As a result, no conclusions about representivity (inferences) can be drawn from the obtained data, and the outputs are supplied purely for illustrative purposes.

Sample Demographics

The finalised sample of 285 cases ranged in age from 18 to 71 years, with a mean age of 32.7 years. The sample was predominantly male (79.6%), predictably so given the well established higher preponderance of substance abuse amongst males. This male bias was relatively consistent at each of the different treatment facilities sampled. The majority of the respondents were Coloured (56.7%), followed by Black (25%) and then White (12.7% and Asian (5.6%) respondents. The distribution of primary language spoken followed from the distribution of race, with Afrikaans being the most frequent home language (57%), followed by English (18%) and then isiXhosa (12.7%) and isiZulu (12.3%). The distribution of race and home language reflects the demographic characteristics of the two provinces where data was collected, with higher concentrations of Coloured and Afrikaans speakers in the Western Cape and higher proportions of English and Zulu speakers in KwaZulu-Natal. Almost three-fifths (58.4%) of the sample were single, about one quarter were married (27.6%), and the rest were either cohabiting (11.1%) or divorced/widowed (6.5%). Just under half of the sample were unemployed (47.8%), a slightly smaller proportion were employed (43.6%) and the remaining 9% were either students or retired.

The sample was split almost equally between those who had completed education up to Grade 11 only and those who had completed Grade 12 or higher (see Figure 1). Of those who had completed school, almost half (48.5%) had completed the last two years (Grades 11 and 12).

Figure 1: Distribution of Sample by Levels of Educational Attainment
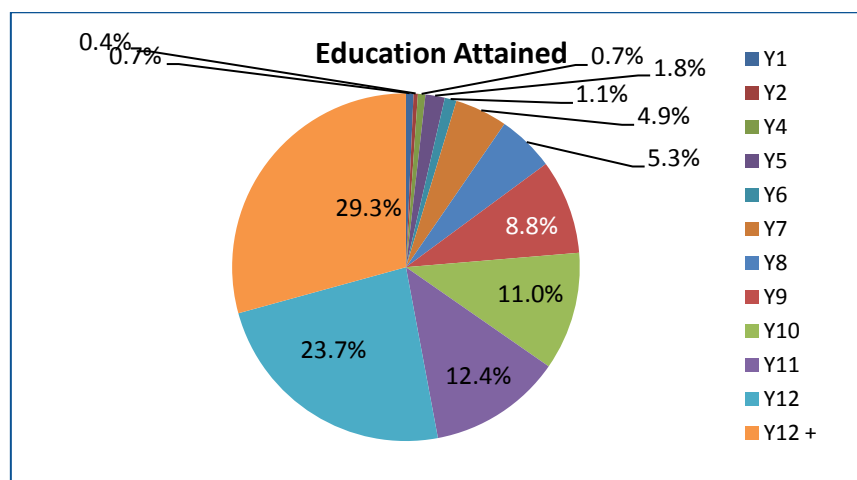
The primary substance abused was crystal meth (tik), followed by alcohol and then marijuana and heroin (see Figure 2). These distributions are consistent with the treatment centres selected for the data collection and do not reflect the true distribution of primary substance abused in the country.

Figure 2: Primary Substance Abused
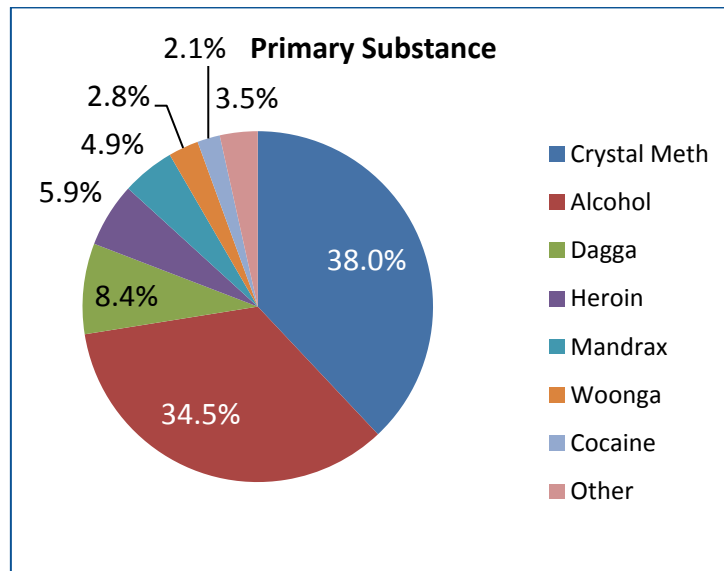
The majority of patients at these centres were not responsible for their own treatment costs, with one third relying on medical aid (33.9%), about one fifth being treated at the state's expense (19.3%), and a similar proportion (17.3%) being dependent on family or employers (see Figure 3).

Figure 3: Source of Payment of Treatment Costs
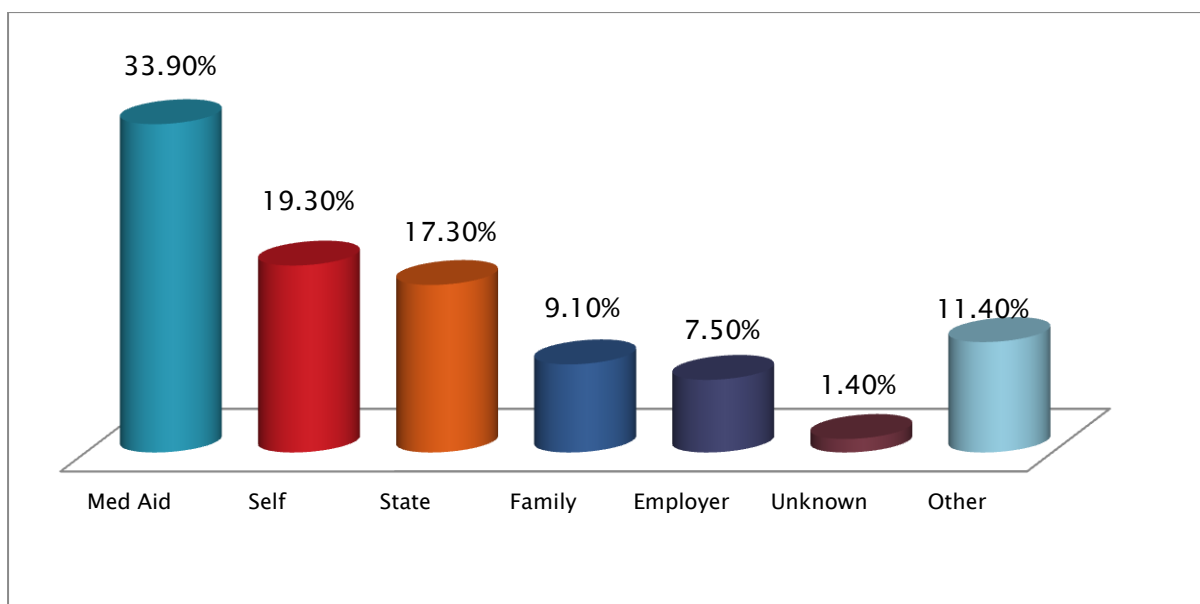
The most frequent referral source for enrolment for treatment was the patient themselves (24.8%), followed by employers (17.5%) and social services (17.2%), and then family (13.9%) and health services (9.5%). Referrals from correctional services accounted for only 4.4% of respondents, unsurprisingly as this referral type is more common in juveniles than in adults.

Finally, the vast majority of respondents (70.1%) were being admitted for substance abuse treatment for the first time, while the remainder has received one or more such treatments in the past. Just over half of the sample (51.7%) has undergone an HIV test in the 12 months, 47.5% did not undergo such tests, and the remaining 1% declined to comment. It is not known whether or not the proportion who did not have a HIV test in the past year may have done so previously.

CHAPTER 5: PSYCHOMETRIC ANALYSIS OF THE SAATSA

This chapter constitutes the substantive portion of the report and discusses the psychometric validation of the various SAATSA scales. The chapter begins with an overview of the psychometric properties of validity and reliability, discusses the methods employed for the psychometric validation process, describes the conceptual structure of the SAATSA, and thereafter details and discusses the psychometric validation of each of the SAATSA scales. To the extent possible, all discussion has been kept non-technical to enable wider comprehension of the report.

Validity and Reliability

All SAATSA items are intended to measure underlying theoretical constructs relevant to process and outcome aspects of substance abuse treatment. These constructs are latent in that they are not directly observed, but they can be measured by use of manifest (observed) variables, all of which are assumed to be representative of the latent variable to varying degrees. The measurement of latent constructs by observed variables is inevitably prone to some degree of error as a result of the design, conceptualisation, and operationalisation of these measured variables. Additionally, such measurement is also vulnerable to the effects of sampling and sampling error. To examine the extent of such error and the overall competence of the observed variables in measuring the latent construct, it is necessary to investigate the psychometric properties of each of these individual measured variables in relation to the latent construct.

Two concepts are important with regard to item/scale psychometric properties: validity and reliability, and these may be summarised as follows:

- Validity, the more important of the two, refers to the extent to which the observed variables are actually measuring the latent construct which they are intended to measure. Validity is typically differentiated in terms of face validity (measured variables on face value appear to measure the latent construct), content validity (measured variables provide coverage of important content domains relevant to the latent construct), convergent validity (measured variables correspond to some independent index in a predictable manner) and construct validity (measured variables capture and measure the conceptual meaning of the latent variable). In psychometric analysis it is construct validity which is of paramount importance, as achieving this typically ensures that the other three types of validity are also achieved.
- Reliability refers to the extent to which measured variables consistently measure whatever it is they are intended to measure. Reliability is a necessary but not a sufficient condition for validity: variables may consistently measure the same thing but this "same thing" may have little correspondence with the actual latent construct of interest. In contrast, construct validity is a fully sufficient condition for reliability: if measured items are highly indicative of some latent variable, they will be so consistently in repeated applications of the variables.

Psychometric validation beings with an examination of validity as any absence thereof would render the scale unusable despite whatever reliability was established.

Methodology For Psychometric Validation

Psychometric validation refers to the process whereby questionnaire items are subject to various tests to investigate their validity and reliability. Such tests may - and often do - derive from different methodological approaches to validation, and while these different methodological approaches are sometimes presented as competing, they are in fact complementary and convergent as they analyse validity and reliability through different lenses. In other words, if applied rigorously, different approaches would and should converge on common solutions.

The first iteration of the pilot used two approaches to psychometric validation, one derived from Classical Test Theory (CTT) and the other derived from Modern Measurement Theory (MMT). The primary difference between the two approaches may be summarised in terms of their emphasis: CTT emphasises the characteristics of the scale over that of the items while MMT emphasises the characteristics of individual items in the scale. Nevertheless, they are complementary and were applied as such in the first iteration. In the analysis of the second iteration data both approaches are employed as previously: MMT (in the form of Item Response Theory) to examine the properties of individual items, and CTT to examine scale characteristics. However, for the purposes of comprehension and brevity, only the CTT tests are reported in detail in this report.

Validity Analysis

Analysis of validity was undertaken first, using exploratory factor analysis (EFA) to investigate the likely underlying dimension (factor) measured by individual SAATSA scales. Unless otherwise stated, all EFA was conducted using Maximum Likelihood Estimation (ML) and Oblique Minimum Rotation (Oblimin). ML was deemed necessary as this provides the most robust test for real world data and Oblimin was employed to account for the fact that a single SAATSA scale might actually comprise more than one latent dimension, hence these latent dimensions should be allowed to correlate with each other. In all EFA the standard criterion for selection of factors was an Eigen value greater than 1. The EFA was conducted using the SPSS software package.

For each scale, the exploratory factor analysis was followed by a confirmatory factor analysis (CFA), to confirm the latent structure observed in the EFA. This analysis was run using the AMOS module in SPSS. Where necessary, data imputation was executed to ensure that missing data was eliminated and thus modification indices could be produced to refine obtained latent models. As is well known in the testing of latent or structural equations models, model outputs are vulnerable to increases in sample size, and this is particularly true for the χ^2 statistic, which becomes significant as the sample size increases past 200 (the final SAATSA sample is 285 cases). Hence models may be rejected

principally because of the sample size and despite being conceptually valid and significant. To overcome this effect, current approaches to latent model testing eschew the traditional over-reliance on the χ^2 statistic and instead employ a series of criteria to provide a much more comprehensive and detailed assessment of the adequacy of the obtained latent model. These criteria are as follows:

1. χ^2/df Ratio – The first criterion is the ratio of the χ^2 to the degrees of freedom for the model (unlike other parametric statistics, the degree of freedom in CFA is derived from the number of parameters in the model and not the number of cases in the analysis). A χ^2/df ratio of between 2 and 4 is deemed to be indicative of a very good fitting model, and anything less than 2 is indicative of an excellent fit.
2. Comparative Fit Index (CFI) – the CFI ranges between 0 and 1, with higher values indicating better fit. A CFI of 0.90-0.94 indicates a good fitting model and a CFI of 0.95 and more indicates a very good fitting model.
3. Root Mean Square Error of Approximation (RMSEA) – The smaller the RMSEA (pronounced "Ramsey") value the better the fit of the model. A RMSEA value of 0.08 or less indicates a good fit while a value of less than 0.05 indicates a very good fitting model.

These criteria were applied simultaneously during the CFA tests to assess the adequacy of the obtained latent models.

Reliability Analysis

Following the validity analysis, reliability analysis was conducted for the SAATSA scales using the Cronbach alpha (α) statistic to examine for internal consistency. The Cronbach alpha ranges between 0 and 1, with higher scores indicating greater levels of reliability and vice versa. The assessment of the returned reliability estimates was conducted using the following scale:

- Above 0.90 – excellent (but with possibility of some redundancy beyond 0.93),
- Between 0.80 and 0.90 – excellent,
- Between 0.70 and 0.80 – good,
- Between 0.65 and 0.70 – satisfactory,
- Less than 0.65 – unsatisfactory.

Additionally, apart from obtaining the Cronbach α for each scale for the entire sample, analyses were conducted to examine the variation in this statistic based on selected respondent demographic characteristics. For the purposes of the analysis, the demographic variables employed in the first iteration are used once again (gender, education and admission type) though the fourth variable (race) has been dropped and replaced by home language (race was used as a proxy for home language in the first iteration). The objective here is to examine whether or not the reliability estimates obtained for different sub-samples differed substantially across these sub-samples. Significant variance in the estimates would highlight the possibility of respondent effects for that particular demographic variable. That is, it would be indicative of the fact that the reliability of the

scale differed significantly depending on whether respondents were male or female, inpatients or outpatients, had a particular level of educational achievement, or as a result of their primary home language. Ideally, all sub-sample reliability statistics should be within the same interval (the intervals presented above), indicating that the scale was (sufficiently and) equally reliable for all respondents regardless of their demographic characteristics. This particular analysis overlaps with the analysis of differential item functioning in IRT.

The SAATSA Scales

For the purposes of consistency, the SAATSA conceptual dimensions (Quality of Treatment, Access to Treatment, Quality of Life, Social Connectedness, Substance Abuse, and HIV) will hereinafter be referred to as scales. The composition of the SAATSA scales by SAATSA items is reflected in Table 4.

Table 4: SAATSA Scales and Items

QUALITY DOMAIN	
SCALES	**ITEMS**
ACCESS	Item 1. The amount of time I had to wait to get services was acceptable to me.
	Item 2. I can afford the treatment I want to receive.
	Item 3. The location of this treatment centre is convenient for me.
	Item 4. I feel safe travelling to this treatment centre.
	Item 5. It is easy for me to obtain the treatments offered by this centre.
	Item 6. My family is able to access services provided by this treatment centre.
	Item 7. I can afford the transport costs of getting to this treatment centre
QUALITY	Item 8. The staff at this treatment centre treat me with respect.
	Item 9. The people I went to for treatment services spent enough time with me.
	Item 10. I have a say in deciding about substance abuse treatment I am receiving here.
	Item 11. The staff told me about services that will help me stay off drugs and alcohol.
	Item 12. This treatment centre teaches me how to avoid getting HIV.
	Item 13. The staff at this treatment centre are sensitive to my background.
	Item 30. I would recommend this treatment centre to a friend.

EFFECTIVENESS DOMAIN	
SCALES	**ITEMS**
QUALITY OF LIFE	Item 14. My general health is improving. Item 15. I am better able to cope when things go wrong. Item 16. I am better able to accomplish the things I want to do. Item 18. In the future, I will be more likely to do better at work or at school. Item 19. I am more likely to practice safe sex.
SOCIAL CONNECTEDNESS	Item 20. There is someone who cares about whether I am doing better. Item 21. I have someone who will help me when I have a problem. Item 22. I have people in my life who are a positive influence. Item 23. The people who care about me are supportive of my treatment. Item 24. My friends and family are able to count on me more Item 25. I have friends who are not using alcohol or drugs. Item 26. I have someone who will listen to me when I need to talk.
SUBSTANCE ABUSE	Item 17. I am less likely to use alcohol or other drugs. Item 27. I now know that using alcohol and drugs is a problem for me. Item 28. I need to work on my problems with alcohol and/or drugs. Item 29. The treatment centre is helping me to recover from using drugs and alcohol.
HIV	Item 33. I am more likely to use condoms when I have sex Item 34. I am less likely to use alcohol and drugs before I have sex Item 35. I am less likely to share drug equipment (e.g. pipes and injections) Item 36. I am more likely to only have sex with one partner Item 37. I am more likely to use condoms properly Item 38. I know it is important to know my HIV status Item 39. I know where I can get tested for HIV Item 40. I know where I can get treatment for HIV

The analysis of the second iteration data follows from the first iteration in terms of adopting some of the recommendations for scale-item composition though in a few instances it also begins with a blank slate to enable a more thorough testing of the SAATSA scales and items.

The Social Connectedness Scale

The Social Connectedness scale (SC) comprises seven individual SAATSA items: Items 20 to 26. The first pilot iteration established that the scale had very good validity, with a single factor solution obtained in the EFA and the CFA analysis confirming this with good model indices. The CFA also suggested that the scale easily could be reduced to four items (Items 20 to 23). Using these four items the internal consistency test returned an excellent reliability estimate of α = 0.88.

For the purposes of the second iteration all seven items were once again subject to validity testing using EFA and CFA. The current EFA returned a single factor with an Eigen value greater than 1 and accounting for 48% of the variance in the items. The loadings of the items on the factor are provided in Table 5, together with the loadings obtained in the first pilot iteration for comparison purposes.

Table 5: Factor Loadings for the Social Connectedness Scale Items

ITEM	PILOT 1	PILOT 2
Q20. There is some-one who cares about whether I am doing better	0.612	0.788
Q21. I have some-one who will help me when I have a problem	0.655	0.841
Q22. I have people in my life who are positive influence	0.635	0.762
Q23. The people who care about me are supportive of my treatment	0.676	0.738
Q24. My friends and family are more able to count on me	0.511	0.596
Q25. I have friends who are not using alcohol or drugs	0.256	0.360
Q26. I have some-one who will listen to me when I need to talk	0.490	0.653

As Table 5 reveals, the loadings for the items in the second iteration are generally higher than those obtained in the first iteration, indicating better measurement of the factor at item level. However, the items with the highest loadings remain the same (items 20-23) in both iterations, confirming that these are the core items by which the latent factor is measured. Importantly, the EFA also revealed that *Item 25: I have friends who are not using alcohol or drugs*, relates very weakly to the factor in both iterations. It was previously argued that this may be due to the fact that while the scale items generally measure the extent of connectedness to others in the immediate familial and interpersonal matrix of the respondent, this particular item relates specifically to the substance abuse behaviours of others and hence would not be expected to be a measure of the strength of interpersonal relationships. The data from both iterations appear to confirm this conclusion.

To investigate and confirm the latent structure of the scale, all seven items were entered into a CFA using structural equations modelling (SEM). The results of the CFA for the seven item scale as well as for a reduced six item scale (excluding Item 25) are provided in Table 6, along with the comparable figures for the first iteration.

Table 6: Confirmatory Model Tests Of The Latent Structure For The Social Connectedness Scale

ITEM	PILOT FIRST ITERATION			PILOT SECOND ITERATION		
	χ^2/df Ratio	RMSEA	CFI	χ^2/df Ratio	RMSEA	CFI
All seven items	**2.86**	**0.072**	**0.979**	**1.84**	**0.055**	**0.984**
Four/Six Items	**2.25**	**0.059**	**0.991**	**1.61**	**0.046**	**0.992**

As the SEM results indicate, the reduced scale excluding Item 25 works best in both iterations. However, while the first iteration indicated an optimal model solution for a scale of only four items, the second iteration supports the retention of six items as the indices for this model are much more robust than those obtained for the four item scale in the first iteration. It is likely that this is due to better data integrity as a result of revisions to items from the first to second iteration. Taken

together, the second iteration confirms a very robust scale with good justification for retention of all six items.

Following from the CFA, the six item Social Connectedness scale was subject to tests of internal consistency, firstly to obtain the reliability estimate for the entire sample and then to examine this reliability estimate by sub-samples. This latter analysis corresponds to the examination for respondent effects, that is, the variance in the reliability of the scale as a result of specific respondent demographic factors. The reliability estimates for the full sample and by sub-sample are presented in Table 7, together with the estimates obtained in the first iteration.

Table 7: Reliability Of The Finalised Social Connectedness Scale - Sample And Sub-Samples

SAMPLE	SUB-SAMPLE	PILOT 1	PILOT 2
FULL SAMPLE		0.88	0.89
GENDER	Male	0.82	0.87
	Female	0.89	0.87
EDUCATION	Grade 11 or less	0.88	0.84
	Grade 12 or more	0.88	0.89
ADMISSION TYPE	Inpatient	0.89	0.85
	Outpatient	0.86	0.89
HOME LANGUAGE	Afrikaans	-	0.86
	English	-	0.91
	isiZulu	-	0.85
	IsiXhosa	-	0.84

The second iteration confirms the overall excellent reliability of the Social Connectedness scale. Moreover, the variability in the reliability estimate for the different subsamples is slight, with none of the estimates indicating problems with the reliability of the scale based on respondent demographics such as gender, admission type, education or home language. The latter two results for education and home language are particularly notable and confirm the value of the cognitive testing of items in the first pilot iteration and the revisions to items following that iteration.

Overall, the results of the second iteration confirm the robustness of the Social Connectedness scale as observed in the first iteration and indicate a strong scale with high validity and excellent reliability. It is thus recommended that this six item scale be retained as currently tested for future measurement and application purposes.

The Quality Of Life Scale

Testing of the Quality of Life scale in the first pilot iteration using IRT analysis revealed some variability in the difficulty (probability) and discriminant function values for the items, with these values being relatively equivalent for Items 10, 11, 12, and 14 and different for Item 15 (Items 14, 15 16, 18 and 19 respectively in the second iteration). A corresponding analysis with the second iteration data returned slightly different results: the Item Characteristic Curves (ICC) and Item Information Curves (IIC) were similar for all items except for Item 18 (*In the future, I will be more likely to do better at work or at school*). The final scale model in the first iteration suggested that this item was problematic as it tended to conflate the assessment of *current* quality of life with expected *future* quality of life. The report had accordingly recommended that the phrasing of the item be revised to more accurately reflect the present time and thus be consistent with the other items. However, this recommendation does not appear to have been implemented, and accordingly, the item still retains this reference to a future rather than a present time. Nevertheless, it was included in the analysis in the second iteration.

An EFA conducted on all five items in the second iteration data returned a single factor with an Eigen value greater than one and accounting for 54% of the variance in the five items. The loadings of the five items on the single factor for both iterations of the pilot are presented in Table 8.

Table 8 Factor Loadings for the Quality of Life Scale Items

ITEM	PILOT 1	PILOT 2
Q14. My general health is improving	0.483	0.664
Q15. I am better able to cope when things go wrong	0.450	0.772
Q16. I am better able to accomplish the things I want to do	0.558	0.724
Q18. In the future, I will be more likely to do better at work or at school.	0.410	0.403
Q19. I am more likely to practice safe sex	0.316	0.590

Comparison of the results across the two iterations reveals that 1), Item 18 remains relatively poorly related to the latent factor, returning low loadings in both iterations, and 2), Item 19, which returned the lowest loading in the first iteration, has improved considerably in the second iteration. This is arguably due to the improvement of this item by removal of the "Not Applicable" response category, which in the first iteration had resulted in a substantial proportion of values which had to be recoded as missing.

Based on the outputs of the EFA, a CFA was conducted specifying two models, the first comprising all five items and the second comprising four items (Item 18 excluded). The results of these analyses (see Table 9) confirm that the scale is far more robust when Item 18 is excluded as the model indices for the four item scale are considerably better than those for the five item scale.

Table 9: Confirmatory Model Tests Of The Latent Structure For The Quality of Life Scale

	PILOT FIRST ITERATION			PILOT SECOND ITERATION		
ITEM	χ^2/df Ratio	RMSEA	CFI	χ^2/df Ratio	RMSEA	CFI
All 5 items	3.56	0.084	0.980	5.18	0.109	0.915
4 Items (exc. Q18)	2.75	0.070	0.973	3.52	0.081	0.993

Using these four items, the scale reliability estimates were obtained for the entire sample and the respective sub-samples (see Table 10). As can be seen in the table, the reliability of the four item scale for the full sample is comparable across both iterations (excellent). Additionally, while some variability is evident for some sub-samples (isiZulu speakers), the scale still has good reliability for this sub-sample while returning good to excellent reliability estimates for all other subsamples.

Table 10: Reliability Of The Finalised Social Connectedness Scale - Sample And Sub-Samples

SAMPLE	SUB-SAMPLE	PILOT 1	PILOT 2
FULL SAMPLE		0.83	0.81
GENDER	Male	0.83	0.80
	Female	0.82	0.84
EDUCATION	Grade 11 or less	0.81	0.79
	Grade 12 or more	0.83	0.81
ADMISSION TYPE	Inpatient	0.83	0.78
	Outpatient	0.83	0.85
HOME LANGUAGE	Afrikaans	-	0.79
	English	-	0.84
	isiZulu	-	0.75
	IsiXhosa	-	0.85

A note on the reliability estimates for the sub-samples of isiZulu or isiXhosa language speakers: the proportion of these respondents in the sample is quite small when compared to the other two linguistic sub-samples. As a result, it is likely that there will be variability in the reliability estimates due mainly to this small sample size, and this is evident here. For instance, if the scale is less reliable for native African language speakers – as is evident by the lower estimate for isiZulu speakers, then it is difficult to understand why it is so reliable for isiXhosa speakers, even more so than it is for native English speakers. A more likely explanation is that the estimates are affected by the sample size, and hence greater variability in the reliability estimates for these sub-samples should be viewed with some caution as it may not be as good a reflection of the reliability estimates for these sub-samples as it is for native English and Afrikaans languages speakers.

Based on the validity and reliability analyses, it is recommended that the Quality of Life scale be constituted by Items 14, 15, 16 and 19 for future application and measurement.

The Substance Abuse Scale

The substance abuse scale is measured by four items (items 17, 27, 28 and 29). As in the first iteration, an additional item was added to the scale (*Item 25: I have friends who are not using alcohol or drugs*) to assess if this item would find a better fit here than in the Social Connectedness scale. The assumption here was that respondents who acknowledged that they had a substance abuse problem and were keen to work on it would make changes in their social milieu that correspond to this acknowledgement and willingness to recover and consequently were more likely to associate with people who did not abuse drugs and alcohol. Should this be true, this item ought to load on the Substance Abuse scale. The IRT analysis in the first iteration revealed a somewhat mixed bag for this scale, with the suggestion of two subscales when Item 25 is added. Significantly this item added very little information to the measurement of the latent factor, a result confirmed in the subsequent exploratory and confirmatory factor analyses for the first iteration.

Analysis of the data for the second iteration generally confirmed the results obtained in the IRT and CTT analyses of the first iteration. Item 25 supplied the lowest information to the factor in the IIC analysis, and this was confirmed by the EFA which returned the lowest factor loading for the item (see Table 11).

Table 11: Factor Loadings for the Substance Abuse Scale Items

ITEM	PILOT 1	PILOT 2
Q17. I am less likely to use alcohol or other drugs	0.392	0.439
Q25. I have friends who don't use alcohol and drugs	0.205	0.304
Q27. I know that using alcohol and drugs is a problem for me	0.501	0.827
Q28. I need to work on my problems with alcohol and/or drugs	0.332	0.620
Q29. The treatment centre is helping me to recover from using drugs and alcohol	0.440	0.744

Additionally, the loading for Item 28 (*I need to work on my problems with alcohol and/or drugs*) has improved considerably from the first to the second iteration, a welcome result considering that this is a scale measuring attitudes towards substance abuse. Indeed, the loadings are generally higher for all of the items in the second iteration when compared to the first. The results of the CFA testing two models – one with five items and one with four items - is presented in Table 12.

Table 12: Confirmatory Model Tests Of The Latent Structure For The Substance Abuse Scale

ITEM	PILOT FIRST ITERATION			PILOT SECOND ITERATION		
	χ^2/df Ratio	RMSEA	CFI	χ^2/df Ratio	RMSEA	CFI
All 5 items	4.231	0.094	0.964	1.82	0.054	0.986
4 Items (exc. Q25)	0.965	0.001	0.999	0.59	0.000	0.999

As in the case of the first iteration, the four item model excluding Item 25 works considerably better in the second iteration, with appreciably better model indices. This result, which is the second such rejection for Item 25, confirms that the item is poor and ill-suited to any of the scales in the SAATSA. Accordingly, it is recommended that it be deleted entirely from the SAATSA measure. The reliability estimates for the four item substance abuse scale are provided in Table 13.

Table 13: Reliability Of The Finalised Substance Abuse Scale - Sample And Sub-Samples

SAMPLE	SUB-SAMPLE	PILOT 1	PILOT 2
FULL SAMPLE		0.75	0.72
GENDER	Male	0.75	0.71
	Female	0.76	0.73
EDUCATION	Grade 11 or less	0.79	0.69
	Grade 12 or more	0.74	0.73
ADMISSION TYPE	Inpatient	0.78	0.72
	Outpatient	0.74	0.74
HOME LANGUAGE	Afrikaans	-	0.73
	English	-	0.73
	isiZulu	-	0.68
	IsiXhosa	-	0.73

As in the first iteration, the reliability estimates for the substance abuse scale are among the lowest for all of the SAATSA scales, and this matter that was flagged for attention and future consideration. Having said that, it must be acknowledged that the current reliability returns are still within the range of "good" estimates, with only marginal dips below the lower cut-off of 0.70 for two subsamples (Respondents with Grade 11 or less educational attainment and those who are native isiXhosa speakers). Overall, the substance abuse scale comprising Items 17, 27, 28 and 29 exhibits good validity and reliability and should be retained as such. However, it is recommended that this scale be revisited in the future to investigate how the reliability may be improved through the reformulation of existing items or, more appropriately, the addition of further items.

The HIV Scale

The SAATSA remains principally a measure of the process and outcomes of substance abuse treatment. However, given 1), the nexus between substance abuse and risky sexual behaviour as observed in research in South Africa, 2), the emphasis on HIV awareness and education already prevalent in many substance abuse treatment facilities, and 3), the emphasis on the need for dual diagnosis (substance abuse and HIV) in clinical treatment, it was decided that a robust measure of issues relating to HIV should be developed and included in the SAATSA. This measure would be an outcome and not a process measure, and thus equivalent to the Substance Abuse, Social Connectedness and Quality of Life scales.

To give effect to this objective, desktop research was undertaken to identify relevant and suitable items for inclusion in a SAATSA HIV scale. This research focussed on a number of countries further afield but emphasized measures which had already been applied in other research in South Africa. Based on this, six items were selected for composition of the scale and the scale was introduced midstream in the second iteration pilot. The pilot was considered appropriate for testing of the scale despite having already begun as it would not have been feasible from a logistical and cost perspective to test the scale on its own. There were two trade-offs to implementing the HIV midway through the second iteration pilot: firstly, the sample size would necessarily be smaller for this scale than for the other scales, though if it was adequate in its own right this would suffice for the psychometric validation, and secondly, there was no time to conduct full cognitive testing of the items as had been done for the other SAATSA items. However, as the items were selected principally because they had been applied previously in South Africa, it was reasonable to assume that they would be suitable in terms of language and lexile level considerations. The six items which constitute the HIV scale are as follows:

- Item 33. I am more likely to use condoms when I have sex
- Item 34. I am less likely to use alcohol and drugs before I have sex
- Item 35. I am less likely to share drug equipment (e.g. pipes and injections)
- Item 36. I am more likely to only have sex with one partner
- Item 37. I am more likely to use condoms properly
- Item 38. I know it is important to know my HIV status
- Item 39. I know where I can get tested for HIV
- Item 40. I know where I can get treatment for HIV

All items employed the same response categories as the other SAATSA items, and were presented in the same format though as an addendum to the existing questionnaire. The sample size for the HIV scale was 132 cases, which is well below the pilot sample of 285 cases. However, as there are only six items in the scale, the sample size is more than adequate in terms of having a variable-to-cases ratio of 1 to 20.

Initial testing of the HIV scale revealed some constraints in the application of robust EFA techniques, mainly due to the small sample size. To circumvent this, an IRT analysis was conducted, and this revealed some notable results. Firstly, it appeared that Item 35 (*I am less likely to share drug*

equipment e.g. pipes and injections), was performing poorly on the scale in terms of the discriminant function as well as in terms of the information supplied to the measurement of the latent construct. This is unsurprising as this item examines sharing of drug equipment such as injections, and the prevalence of substance abuse by injection is low in South Africa relative to the other forms of ingestion such as smoking, which accounts for the majority of primary drugs abused (tik, marijuana and mandrax). Even drugs such as heroin which are traditionally injected elsewhere are in fact smoked in the country. Given this, it is very probable that this item has little relevance for the South African context, which would explain its poor performance on the scale. The item was removed from subsequent analysis.

Further inspection of the IRT outputs suggested two sub-scales; a measure of attitudes relating to behaviours, such as using condoms, having multiple partners, and using drugs (Items 33, 34, 36 and 37); and a measure of HIV status, in terms of knowing one's status and knowing where to get tested and treated (Items 38, 39 and 40). Using this hypothesized scale substructure, it was decided that the latent structure of the scale would be directly tested in confirmatory factory analysis. The results of the confirmatory factor analyses testing one model with a single scale and another with two sub-scales is presented in Table 14.

Table 14: Confirmatory Model Tests Of The Latent Structure For The HIV Scale

SUB-SCALES	PILOT SECOND ITERATION		
	χ^2/df Ratio	RMSEA	CFI
Single scale	**3.27**	**0.134**	**0.923**
2 Sub-scales	**1.68**	**0.073**	**0.979**

As can be seen in the table, the output indices for the two sub-scale model are considerably better than that for the single scale model, particularly in terms of the RMSEA and the comparative fit index. This would appear to confirm the thesis that the HIV scale is constituted of two sub-scales, one measuring attitudes to behaviour (Items 33, 34, 36, and 37) and the other measuring attitudes towards HIV status (Items 38, 39 and 40). The observation of two sub-scales in the HIV measure does not automatically require the application of the measure as two separate scales, as overall the data support the use of the complete measure as a single scale. However, the option does exist for application of the two sub-scales as and when this may be deemed necessary and appropriate. For the purposes of the current analysis, the measure will be regarded as a single scale.

The internal consistency of the HIV scale was assessed for the entire sample as well as for the relevant sub-samples and the respective reliability estimates are presented in Table 15. Please note that due to the small sample size for this analysis and especially for some of the linguistic sub-samples, the applicable reliability estimates should be interpreted with caution.

Table 15: Reliability Of The Finalised HIV Scale - Sample And Sub-Samples

SAMPLE	SUB-SAMPLE	PILOT 2
FULL SAMPLE		0.81
GENDER	Male	0.83
	Female	0.76
EDUCATION	Grade 11 or less	0.82
	Grade 12 or more	0.80
ADMISSION TYPE	Inpatient	0.83
	Outpatient	0.80
HOME LANGUAGE	Afrikaans	0.83
	English	0.74
	isiZulu	0.55
	IsiXhosa	0.87

Notwithstanding the small number of cases for this analysis, the HIV scale has excellent reliability across the entire sample, and this reliability is not significantly variant across the different sub-samples. While the reliability estimates for the home language sub-samples should be regarded cautiously because of the very small number of cases, the rest of the sub-sample analysis can be inspected more closely, and a few results are notable here. Firstly, there is little difference in reliability for either level of education attained or the admission type. Secondly, the lower reliability for females is arguably due to the fact that some of the HIV items relate to the respondent using condoms. In the case of males this relates directly to their own behaviour while for females it does not, and instead refers to having to regulate the behaviour of others. This differential applicability by gender would explain the variance in the two reliability estimates across males and females, the largest for any of the sub-samples barring those with very small number of cases.

Overall, the validity and reliability tests of the HIV scale are quite promising. The small sample size for the present analysis necessitates greater attention to psychometric validation of this scale in the future, though based on current results there is little reason not to provisionally incorporate the HIV scale into the SAATSA.

The Quality Scale

The first pilot iteration tested the quality of service by collapsing a number of different scales into a single Quality scale, most notably the scale measuring access to treatment and the scale purportedly measuring the person-centeredness of substance abuse treatment. The primary reason for this was the many conceptual and measurement problems with the individual items intended to measure these latter two scales, the result of which was the deletion of a number of items from the analysis. This deletion resulted in an insufficient number of items for each of these scales to be tested as

stand-alone scales, and consequently, the remaining items from these scales which were vetted for use in the analysis were collapsed into the quality scale. As argued in the first iteration report, this was not an optimal solution and various recommendations were made regarding the improvement of the individual items as well as the adoption of new items to properly measure the access to treatment scale (this matter with be discussed in more detail in the following section). In terms of the person-centeredness scale, it was recommended that the useable items from this scale be subsumed into the quality scale, and they are treated as such in the second iteration analysis.

The revised Quality scale as analysed in the second iteration comprised seven items: Items 8, 9, 10, 11, 12 , 13 and 30. Most of the items were unchanged from the first iteration, with some exceptions:

- Item 10 (*I have a say in deciding about substance abuse treatment I am receiving here*) underwent significant revision following the first iteration to address conceptual ambiguities relating to the type of choice that may be reasonably expected to be granted to patients entering treatment centres;
- Item 12 *(This treatment centre teaches me how to avoid getting HIV)* was flagged for closer inspection as it may not be applicable to centres which did not offer such a service as a routine part of its treatment programme - hence it would not constitute a good measure of the quality of such treatment at these centres.

Using these seven items, the latent structure of the Quality scale was examined by means of an exploratory factor analysis, and the results indicated a two factor solution accounting for 42% of the variance in the items (see Table 16).

Table 16: Factor Loadings for the Quality Scale Items

ITEMS	FACTOR 1	FACTOR 2
Item 8. The staff at this treatment centre treat me with respect.	0.560	0.385
Item 9. The people I went to for treatment services spent enough time with me.	0.563	
Item 10. I have a say in deciding about substance abuse treatment I am receiving here.	0.557	
Item 11. The staff told me about services that will help me stay off drugs and alcohol.	0.461	
Item 12. This treatment centre teaches me how to avoid getting HIV.	0.604	0.575
Item 13. The staff at this treatment centre are sensitive to my background.	0.566	
Item 30. I would recommend this treatment centre to a friend.	0.620	

The two factor solution obtained here mirrors that obtained in the first pilot iteration. Closer inspection of the current EFA output reveals that there are two items that load onto both factors (only loadings greater than 0.30 are considered): Item 8 and Item 12. In the case of Item 8 the dual loading is not of concern as the loading is considerably higher for the first than the second factor. In contrast, Item 12 loads almost equally on both factors, suggesting that this item may be responsible for returning a two factor rather than a single factor solution. To investigate this, a CFA was undertaken. However, as there is only one item with a substantial loading onto the second factor, it

was not possible to test a two factor solution in SEM (the minimum requirement for a factor is three items). Accordingly, the CFA focussed on testing a single factor solution firstly using all seven items and then six items with Item 12 excluded (see table 17).

Table 17: Confirmatory Model Tests Of The Latent Structure For The Quality Scale

ITEM	PILOT FIRST ITERATION			PILOT SECOND ITERATION		
	χ^2/df Ratio	RMSEA	CFI	χ^2/df Ratio	RMSEA	CFI
All 9/7 items	5.04	0.105	0.858	1.51	0.042	0.979
Eight/Six Items	1.84	0.048	0.977	0.39	0.002	0.999

As the results of the CFA indicate, the Quality scale comprising six items (excluding Item 12) is considerably better than the scale comprising all seven items, with improved model indices all round. Moreover, as the issue of HIV awareness and education is subsumed under the newly developed HIV scale, the removal of this item does not deplete the SAATSA in any way. Using the six items, the reliability estimates for the quality scale were determined for the sample and the sub-samples, and the results are presented in Table 18.

Table 18: Reliability Of The Finalised Quality Scale - Sample And Sub-Samples

SAMPLE	SUB-SAMPLE	PILOT 1	PILOT 2
FULL SAMPLE		0.78	0.74
GENDER	Male	0.79	0.76
	Female	0.74	0.71
EDUCATION	Grade 11 or less	0.79	0.71
	Grade 12 or more	0.75	0.76
ADMISSION TYPE	Inpatient	0.78	0.74
	Outpatient	0.77	0.76
HOME LANGUAGE	Afrikaans	-	0.68
	English	-	0.73
	isiZulu	-	0.78
	IsiXhosa	-	0.75

In reviewing these reliability estimates it is important to bear in mind that the scales as measured in the two iterations are somewhat conceptually different in that the first iteration contained items relating to access to treatment while the second has these removed. Nevertheless, it is instructive to note that the reliability estimates in the second iteration are relatively comparable to that obtained in the first and are good for the entire sample and all sub-samples except one (native Afrikaans speakers), where it is satisfactory. Given the improved validity from the first iteration, these reliability estimates are more than adequate.

All told, the Quality scale as tested in the second iteration is a robust and reliable scale, and it should be retained in its current composition, comprising the Items 8, 9, 10, 11, 13 and 30.

The Access Scale

As argued in the first iteration report, the absence of a fully-fledged scale to measure access is a significant omission in South Africa, where access to proper health services, particularly for the majority of the population, is well documented to be poor and extremely uneven. In considering access to treatment and the development of appropriate items for such a scale, the report recommended coverage of various issues, *inter alia*;

- the spatial location of these facilities, which renders them difficult to access in terms of the time required to travel there as well as the costs associated with such travel;
- the convenience of the locations in terms of the degree of effort required by patients to reach the facility, that is, whether or not transportation is available from their location to the facility, and the complexity of such travel in terms of modes of transport and transport schedules;
- the personal safety of the patient in travelling to and from the treatment centre and their home or workplace. This is an important issue in a country with a high rate of crime, and particularly so given the degree of violence associated with personal crime.

Following the first iteration a new set of items measuring access to treatment was developed based on coverage of the key issues and items which had been employed in studies of access to health care in South Africa. Once the initial pool of items had been selected the items were subject to cognitive testing using the same protocols as applied for all the other SAATSA items. The results of the cognitive testing indicated seven items that could be used to measure access to treatment, as follows:

- Item 1. The amount of time I had to wait to get services was acceptable to me.
- Item 2. I can afford the treatment I want to receive.
- Item 3. The location of this treatment centre is convenient for me.
- Item 4. I feel safe travelling to this treatment centre.
- Item 5. It is easy for me to obtain the treatments offered by this centre.
- Item 6. My family is able to access services provided by this treatment centre.
- Item 7. I can afford the transport costs of getting to this treatment centre

Two of these items were carried over from the first iteration (Items 1 and 2). As this is the first test of this construct as a stand-alone scale there is no comparable scale-level data available from the previous iteration.

Initial analysis of the seven access items using EFA indicated a two factor solution. However, the inclusion of Item 2 *(I can afford the treatment I want to receive)* severely compromised the analysis as more than 40% of the pilot sample indicated that this aspect of access was not applicable to them as they were not directly responsible for payment for their treatment (the data presented in Chapter 4 confirm that just over half of the respondents indicated that treatment costs were being met either by the state or medical aid). This issue had surfaced and was addressed previously, and it was recommended that the item be removed to ensure retention of the full sample. The second iteration confirms the validity of this recommendation.

In addition to the issue of a depleted sample for analysis, there are sound conceptual reasons for removing Item 2 from the Access scale. All of the other scale items relate specifically to issues of access, while this item relates specifically to the issue of affordability. To understand this better, it is useful to adopt the framework commonly applied in assessing service delivery in public administrations around the world. This method – known as the triple A framework - accounts for three distinct yet related constructs of service delivery: availability, accessibility, and affordability. Within this framework, service delivery assessment is conducted by asking:

- is the service available where the citizen is, i.e., are there actually physical structures where the service is provided?
- if available, is the service accessible by the citizen, i.e., can they travel there, can they afford to travel to the service, are the operating hours convenient, etc.
- if accessible, is it affordable for them to use, i.e., , can they afford to pay for the service, etc.

This framework serves all public service delivery, and health is no exception. Based on this, it is clear that the Item 2, despite being important, is inappropriately located in a scale that measures access.

Removing item 2 and testing the remaining items in an EFA retains the entire sample for analysis and, more importantly, demonstrates a stronger conceptual solution, with the analysis indicating a single factor accounting for over 40% of the variance in the items. The loading of the items on this single factor are provided in Table 19.

Table 19: Factor Loadings for the Access Scale Items

ITEM	FACTOR 1
Item 1. The amount of time I had to wait to get services was acceptable to me.	0.671
Item 3. The location of this treatment centre is convenient for me.	0.664
Item 4. I feel safe travelling to this treatment centre.	0.686
Item 5. It is easy for me to obtain the treatments offered by this centre.	0.709
Item 6. My family is able to access services provided by this treatment centre.	0.590
Item 7. I can afford the transport costs of getting to this treatment centre	0.626

As the table reveals, the six items generally load quite well onto the single latent construct. This EFA result confirms the IRT analysis of these items, which showed similar Item Characteristic Curves and Item Information Curves for all of the items except Item 6. At face value, there appears an obvious reason for this: while the other five items measure access specifically in relation to the respondent, Item 6 measures access in terms of the family of the respondent. This may be considered to have relevance only to the extent that 1), the treatment is either inpatient or outpatient as these have differential implications in terms of the family, and 2), the respondent has such family for whom accessing services are deemed important. This is a qualitatively different measurement from that of access for the respondent themselves, hence the weaker loading for this item.

Using the results from the IRT analysis, a CFA was conducted testing two models, one with all six items and a second with only five items (Item 6 excluded). The results of the CFA (see table 20) indicate that while the six item scale is somewhat robust, the five item scale is considerably more so with better model indices all round.

Table 20: Confirmatory Model Tests Of The Latent Structure For The Access Scale

ITEM	χ^2/df Ratio	RMSEA	CFI
	PILOT SECOND ITERATION		
Six items	3.35	0.091	0.949
Five items (excluding Item 6)	1.60	0.046	0.990

Tests of internal consistency for the five item Access scale returned good estimates for the full sample as well as for the various sub-samples (see table 21), with none of the estimates dropping out of the good range. Moreover, the variability of the estimate across different sub-samples is minimal, suggesting that the scale is comparably reliable across all of these groups.

Table 21: Reliability Of The Finalised Access Scale - Sample And Sub-Samples

SAMPLE	SUB-SAMPLE	PILOT 2
FULL SAMPLE		0.76
GENDER	Male	0.76
	Female	0.72
EDUCATION	Grade 11 or less	0.72
	Grade 12 or more	0.78
ADMISSION TYPE	Inpatient	0.75
	Outpatient	0.76
HOME LANGUAGE	Afrikaans	0.74
	English	0.76
	isiZulu	0.77
	IsiXhosa	0.70

Taken together, the IRT and CFA analysis of the access items support the removal of Item 6 along with the removal of Item 2 from the Access scale. The finalised scale was demonstrated to be quite robust in terms of its validity and to have good reliability for the entire sample and within sub-samples. It is thus recommended that scale as constituted by these five items (Items 1, 3, 4, 5 and 7) be considered fit for use (though it should also remain a candidate for further validation as this is the first testing).

Finalised SAATSA Scales

Based on the psychometric analysis, the final SAATSA scales are as reflected in Table 22.

Table 22: Finalised SAATSA Scales

FINAL SAATSA SCALES	
QUALITY DOMAIN	
ACCESS	Item 1. The amount of time I had to wait to get services was acceptable to me.
	Item 3. The location of this treatment centre is convenient for me.
	Item 4. I feel safe travelling to this treatment centre.
	Item 5. It is easy for me to obtain the treatments offered by this centre.
	Item 7. I can afford the transport costs of getting to this treatment centre
QUALITY	Item 8. The staff at this treatment centre treat me with respect.
	Item 9. The people I went to for treatment services spent enough time with me.
	Item 10. I have a say in deciding about substance abuse treatment I am receiving here.
	Item 11. The staff told me about services that will help me stay off drugs and alcohol.
	Item 13. The staff at this treatment centre are sensitive to my background.
	Item 30. I would recommend this treatment centre to a friend.
EFFECTIVENESS DOMAIN	
QUALITY OF LIFE	Item 14. My general health is improving.
	Item 15. I am better able to cope when things go wrong.
	Item 16. I am better able to accomplish the things I want to do.
	Item 19. I am more likely to practice safe sex.
SOCIAL CONNECTEDNESS	Item 20. There is someone who cares about whether I am doing better.
	Item 21. I have someone who will help me when I have a problem.
	Item 22. I have people in my life who are a positive influence.
	Item 23. The people who care about me are supportive of my treatment.
	Item 24. My friends and family are able to count on me more
	Item 26. I have someone who will listen to me when I need to talk.
SUBSTANCE ABUSE	Item 17. I am less likely to use alcohol or other drugs.
	Item 27. I now know that using alcohol and drugs is a problem for me.
	Item 28. I need to work on my problems with alcohol and/or drugs.
	Item 29. The treatment centre is helping me to recover from using drugs and alcohol.

HIV	Item 33. I am more likely to use condoms when I have sex
	Item 34. I am less likely to use alcohol and drugs before I have sex
	Item 36. I am more likely to only have sex with one partner
	Item 37. I am more likely to use condoms properly
	Item 38. I know it is important to know my HIV status
	Item 39. I know where I can get tested for HIV
	Item 40. I know where I can get treatment for HIV

CHAPTER 6: CONCLUSION AND RECOMMENDATIONS

As discussed at the beginning of this report, the purpose of the second iteration of the SAATSA pilot study was to address various issues arising from the first iteration with respect to 1), fieldwork and administration, and 2), methodology and measurement of the SAATSA items and scales. The analysis of the second iteration data indicate that the overwhelming majority of these issues have been successfully addressed and resolved. The specific findings and recommendations emanating from this iteration are summarised in this chapter.

Patient Background/Demographic Information

SACENDU- SAATSA Link

The linking of the SAATSA to the SACENDU has greatly improved the collection and quality of the patient demographic and administrative information, and there are various reasons for this. Firstly, the SACENDU is well established at treatment facilities in the country hence it does not add any further administrative burden to staff, a key factor in ensuring good returns on the data. Secondly, the SACENDU data is collected directly by clinical staff rather than through self-reporting by respondents, and this improves the reliability of the data. And thirdly, the coverage of issues by SACENDU exceeds that of the previous SAATSA background items, and this provides valuable additional information for analysis purposes.

Recommendation: The SACENDU should remain the sole mechanism for collection of patient background information for the SAATSA. Moreover, and as reflected in discussions with MRC staff affiliated to the SACENDU, the imminent review of the SACENDU offers a good opportunity to ensure improved measurement of this data by appropriate and relevant revision of the SACENDU items.

Education and Primary Home Language Variables

The variables measuring home language and educational attainment still require further revision. In the former case the variable must measure *primary* home language by use of a single response item, as this is the principal issue at hand for the study of language effects (item differential effects) for the SAATSA items and scales. In the latter case there needs to be better measurement of the different levels of tertiary education, to enable greater resolution of these for future analysis. These revisions are easily effected using standard variables as applied in the country.

Recommendation: The revision of the two variables measuring primary home language and educational attainment should be done within the context of the review of the SACENDU so that they are incorporated into this revised measure. The home language variable would be a useful addition to the SACENDU as it currently does not ask this question.

SAATSA Items

Item Response Format – Ceiling Effects

The second iteration of the pilot confirmed the findings of the first iteration with respect to the absence of ceiling effects in the measurement of the SAATSA items. As discussed previously, ceiling effects are not uncommon in the assessment of service quality, and particularly so when the assessment is done while the service is current, as in the case of the SAATSA. If they are sufficiently large, such effects can impact validation and scale measurement by constraining the amount of variance available in each item. Reassuringly, the current response format for the SAATSA items does not lend itself to these effects and returns sufficient item variance to permit necessary analyses.

Recommendation: The current response format comprising four categories of agreement with the item statement (Disagree, Somewhat Agree, Agree, Strongly Agree) be retained intact, and that future validation tests continue to examine for such ceiling effects as part of efforts to ensure continuous improvement of the SAATSA items.

Item Response Format – The Not Applicable Category

The removal of the "Not Applicable" category from the response format for all SAATSA items except those items for which such a category is conceptually meaningful and valid has significantly improved the integrity of the data in two important ways. Firstly, it has removed the uncertainty around why some respondents might consider as immaterial (not applicable) the very issues which are incontrovertibly important and central to their treatment for substance abuse. And secondly, the elimination of the need to (correctly) recode such Not Applicable responses as missing has significantly reduced the proportion of missing values for the various SAATSA items, thereby making more of the sample available for validation of the scales and any further analysis.

Recommendation: The current configuration of the response format for the SAATSA items be retained, with the Not Applicable category being available only for the minority of designated items.

SAATSA Scales

The Social Connectedness Scale

The current data confirms the robustness of the Social Connectedness scale as observed in the first iteration and indicates a scale with high validity and excellent reliability. The analysis from both iterations also motivates for the exclusion of one item (*I have friends who are not using alcohol and/or drugs)* from the finalised scale. Taken together, these analyses suggest no pressing need for any further psychometric development of the items or the scale.

Recommendation: The six item Social Connectedness scale as tested in this analysis be retained for future measurement and application purposes, with validation at regular intervals to ensure that the scale psychometric properties remain intact and acceptable.

The Quality of Life Scale

The Quality of Life scale was also found to have good validity and excellent reliability in both iterations of the pilot. The first iteration suggested that one item was problematic (*I expect to do better at work/finding work or at school*) and recommended revision for the second round. Despite these revisions, the second iteration analysis indicated that the item fails to contribute substantively to the measurement of the latent construct, and that removing it greatly improves both validity and reliability.

Recommendation: The four item Quality of Life scale as tested in this analysis be retained for future measurement and application purposes, with validation at regular intervals to ensure that the scale psychometric properties remain intact and acceptable.

The Substance Abuse Scale

The Substance Abuse scale demonstrated good validity and good reliability in both iterations of the pilot. In both instances an additional item was added to the scale to test its applicability (*I have friends who are not using alcohol and/or drugs*). This item failed to add any appreciable value to the scale in either of the analyses and was subsequently dropped from the scale. The removal of the item substantially improved the psychometric properties of the scale.

Recommendation: The four item Substance Abuse scale as tested in this analysis be retained for future measurement and application purposes, with validation at regular intervals to ensure that the scale psychometric properties remain intact and acceptable.

The HIV Scale

The HIV scale is a new addition to the SAATSA outcome measures, and was introduced somewhat late in the fieldwork process, thus returning data for only a small segment of the sample (130 cases). The small sample size posed some constraints for standard exploratory factor analysis, but Item Response Theory analysis was possible and this indicated two sub-scales in the measure: one measuring attitudes towards safer behaviours, and the other relating to HIV status. The IRT analysis also indicated problems with one item (*I am less likely to share drug equipment e.g. pipes and injections*) and it was argued that the item was inappropriate for the South African context where ingestion of substances is principally orally rather than by injection. With this item deleted, the confirmatory factor analysis validated the hypothesis of two subscales with good validity. The overall scale returned excellent reliability estimates for the sample, though the small sub-sample sizes mean that relevant reliability estimates must be regarded with caution.

Recommendation: The seven item HIV scale as tested in this analysis should be integrated into the SAATSA as it provides valuable information that accounts for the nexus between HIV and substance abuse. Notwithstanding any current application, this scale should be regarded as being in the development stage, with further psychometric testing and validation necessary to confirm the observed validity and reliability.

The Quality Scale

The Quality scale was demonstrated to have good validity and good reliability in the first iteration, and this was confirmed in the second iteration, though some changes were effected across the iterations and hence the finalised scales in each are not exactly comparable. The most important revisions were the removal of the items measuring access and their insertion into the new Access scale, as well revision of the wording of some items (*I have a say in deciding about substance abuse treatment I am receiving here*). The current analysis confirmed the improvement as a result of these revisions but also confirmed that one item was inappropriately located in this scale (*This treatment centre teaches me how to avoid getting HIV*). This item was subsequently dropped from the scale, realising a significant improvement in validity and reliability. Additionally, the items measuring person-centeredness were permanently subsumed into the Quality scale.

Recommendation: The six item Quality scale as tested in this analysis be retained for future measurement and application purposes, with validation at regular intervals to ensure that the scale psychometric properties remain intact and acceptable.

The Access Scale

The Access scale is a new addition to the SAATSA and addresses a significant gap in the SAATSA coverage of service quality issues in substance abuse treatment in South Africa. The scale was constituted from some existing items (previously incorporated into the Quality scale) and additional items sourced to reflect critical issues in differential access to treatment by the local population. The current data revealed problems with two items. Firstly, the item *I can afford the treatment I want to receive* correctly provided a "Not Applicable" response category, but when this is accounted for, there is a significant loss of cases (40%) from the analysis. However, this item is important - and hence the data ought to be collected - but it should not be regarded as a part of this particular scale. And secondly, the item *My family is able to access services provided by this treatment centre* does not work well as it refers to access for the family while all other access items refer to access for the respondent, hence the measurement is qualitatively different. Both items were subsequently dropped from the scale, and the finalised scale demonstrated good validity and reliability.

Recommendation: The five item Access scale as tested in this analysis should be adopted as a SAATSA scale as it enhances the measurement of the process aspect of service quality at substance abuse centres in important ways. However, despite the presence of previously tested items the overall scale is relatively new and should be regarded

as being in the development stage, with further psychometric testing and validation necessary to confirm the observed validity and reliability.

South African Addiction Treatment Services Assessment (SAATSA)   **Client number** ☐

Demographic information

1.   What is your home language/s? Please tick all relevant boxes.

Afrikaans	
English	
IsiNdebele	
IsiXhosa	
IsiZulu	
Sepedi	
Sesotho	
Setswana	
Sign language	
SiSwati	
Tshivenda	
Xitsonga	
Other…………………………….	

2.   What is the highest level of school education that you have <u>completed</u>? Please tick the appropriate box.

Grade 1		Grade 8	
Grade 2		Grade 9	
Grade 3		Grade 10	
Grade 4		Grade 11	
Grade 5		Grade 12	
Grade 6			
Grade 7			

2.1 Do you have any post-matric qualifications? Please tick the appropriate box/s if you do.

NTC 1?N1/NIC/(V)	
Degree	
Post graduate degree	
Other, please specify……………………… …………………………………………….	

Please read each statement below and think about the treatment you have received from the staff at this treatment center. Please tick the box that best describes how you feel about these treatments services. Treatment services include in-take; screening; assessment and treatment planning; medications; individual, group and family counseling; information and referral; and information and support groups for the family.

		Disagree	Somewhat Agree	Agree	Strongly Agree	
For the statement below, please indicate the extent to which you agree						
1	The amount of time I had to wait to get services was acceptable to me.					
2	I can afford the treatment I want to receive. (If you are receiving free treatment, then tick 'Does not apply').					Does not apply
3	The location of this treatment centre is convenient for me.					
4	I feel safe travelling to this treatment centre.					
5	It is easy for me to obtain the treatments offered by this centre.					
6	My family is able to access services provided by this treatment centre.					
7	I can afford the transport costs of getting to this treatment centre					
8	The staff at this treatment centre treat me with respect.					
9	The people I went to for treatment services at this centre spent enough time with me.					
10	I have a say in deciding about my substance abuse treatment that I am receiving here.					
11	The staff at this treatment centre told me about services in my area that will help me stay off drugs and alcohol.					
12	This treatment centre teaches me how to avoid getting HIV.					
13	The staff at this treatment centre are sensitive to my background.					
As a result of the treatment I am receiving…						
14	My general health is improving.					
15	I am better able to cope when things go wrong.					
16	I am better able to accomplish the things I want to do.					

		Disagree	Somewhat Agree	Agree	Strongly Agree	
17	I am less likely to use alcohol or other drugs.					
18	In the future, I will be more likely to do better at work or at school. (If you will not be working or attending school in the future, then tick 'Does not apply').					Does not apply
19	I am more likely to practice safe sex.					
20	There is someone who cares about whether I am doing better.					
21	I have someone who will help me when I have a problem.					
22	I have people in my life who are a positive influence.					
23	The people who care about me are supportive of my treatment.					
24	My friends and family are able to count on me more . (If you do not have any friends or family, then tick 'Does not apply').					Does not apply
25	I have friends who are not using alcohol or drugs.					
26	I have someone who will listen to me when I need to talk.					
27	I now know that using alcohol and drugs is a problem for me.					
28	I need to work on my problems with alcohol and/or drugs.					
29	The treatment centre is helping me to recover from using drugs and alcohol.					
30	I would recommend this treatment centre to a friend.					

Thank you for completing the survey!!!

SQM HIV services assessment   Client number

Please read each of the statements and tick the box that best describes you.

1. As a result of the treatment I am receiving, I received information or education about HIV.

Yes	No

2. I am sexually active

Yes	No

Please read the next eight statements and think about the treatment you have received at this centre. Please tick the box that best describes how you feel about these treatment services.

		Disagree	Somewhat Agree	Agree	Strongly Agree
	For the statements below, please indicate the extent to which you agree				
	As a result of the treatment I am receiving…				
1	I am more likely to use condoms when I have sex				
2	I am less likely to use alcohol and drugs before I have sex				
3	I am less likely to share drug equipment (e.g. pipes and injections)				
4	I am more likely to only have sex with one partner				
5	I am more likely to use condoms properly				
6	I know it is important to know my HIV status				
7	I know where I can get tested for HIV				
8	I know where I can get treatment for HIV				

Thank you for answering these questions.

APPENDIX 2

Q1 Time waited for services was acceptable

Mean = 3.26
Std. Dev. = .742
N = 285

Q2 Can afford treatment

Mean = 3.04
Std. Dev. = .785
N = 171

Q3 Centre conveniently located

Mean = 3.03
Std. Dev. = .871
N = 284

Q4 Safe travelling

Mean = 3.17
Std. Dev. = .759
N = 283

Q5 Treatments easily obtained at centre

Mean = 3.18
Std. Dev. = .788
N = 284

Q6 Easy access for family to services at centre

Mean = 3.06
Std. Dev. = .868
N = 277

Q7 Transport affordable to centre

Q8 Staff treat me with respect

Q9 Staff spent enough time with me

Q10 Had say in deciding treatment

Q11 Staff informed about services available in area

Q12 Teaches how to avoid HIV

Q13 Staff sensitive to my background

Q14 General health is improving

Q15 Am better able to cope when things go wrong

Q22 Have people in my life who are positive influence

Mean = 3.4
Std. Dev. = .684
N = 284

Frequency

Q22 Have people in my life who are positive influence

Q23 Those who care support my treatment

Mean = 3.5
Std. Dev. = .621
N = 284

Frequency

Q23 Those who care support my treatment

Q24 Friends/family can count on me

Mean = 3.34
Std. Dev. = .699
N = 267

Frequency

Q24 Friends/family can count on me

Q25 Have friend who dont use

Mean = 2.86
Std. Dev. = 1.009
N = 283

Frequency

Q25 Have friend who dont use

Q26 Have someone who listens

Mean = 3.36
Std. Dev. = .687
N = 285

Frequency

Q26 Have someone who listens

Q27 Realise using is a problem for me

Mean = 3.66
Std. Dev. = .51
N = 285

Frequency

Q27 Realise using is a problem for me

Q28 Need to work on problems with using

Mean = 3.5
Std. Dev. = .774
N = 283

Q29 Centre helping me recover from using

Mean = 3.65
Std. Dev. = .507
N = 284

Q30 Will recommend centre to others

Mean = 3.64
Std. Dev. = .575
N = 284

Q33 More likely to use condoms

Mean = 3.16
Std. Dev. = .967
N = 128

Q34 Less like to use alcohol/drugs before sex

Mean = 2.86
Std. Dev. = 1.175
N = 133

Q35 Less likely to share drug equipment

Mean = 2.83
Std. Dev. = 1.221
N = 130

Q36 Likely to have sex with only one partner

Q37 Likely to use condoms properly

Q38 Realise it is importnt to know my status

Q39 Know where I can get tested for HIV

Q40 Know where I can gert treatment for HIV