# Section eight: Appendices

- Development of scales and indices
- Laboratory HIV testing procedures
- Evaluation of the age-sex distributions
- Reliability and validity of HIV prevalence rate, socio-demographic profiles, coefficient of variation and the design effects
- List of fieldwork supervisors, interviewers and coders

# 8. Appendices

## Appendix 1: Development of scales and indices

### Religious involvement index

Based on the principle of subjective rating of importance and religious practice, a measure with two items, how important is religion to you (rated from 1=extremely important to 5=not important at all) and how often do you practise your religion (rated from 1=regularly [once or more a week] to 4=never) was included (Koenig, McCullough & Larson, 2001).

Cronbach's alpha for the religious involvement index was .74 for this sample.

### Job satisfaction scale

A 16-item job satisfaction scale was developed using items from existing scales (Brown, Kitchell, O'Neill, Lockliear, Vosler, Kubek & Dale, 2001; Lester 1987; Van Saane, Sluiter, Verbeek & Frings-Dresen 2003), focus groups with educators and expert interviews, for example, 'Teaching provides possibilities for promotion' (rated from 1=disagree to 3=agree).

Following a pilot survey, item analysis and principal component analysis with varimax rotation was used and a 6-factor solution was found to be the most appropriate.

In the main study, principal component analysis with varimax rotation yielded the same six components accounting for 61% of the total variance. The first factor (eigenvalue: 3.51) accounted for 22% in the variance in the responses and contained items concerning career advancement and recognition. The second factor (eigenvalue: 1.70) accounted for 10.6% in the variance in the responses and included items related to peer support. The third factor (eigenvalue: 1.31) accounted for 8.2% of the variance in the responses and included items related to teaching structure (working hours/load/policies). The fourth factor (eigenvalue: 1.20), explaining 7.5% of the variance in responses, contained items about discipline and respect. The fifth factor (eigenvalue: 1.07), accounting for 6.7% of the variance in responses, reflected items on community enhancement, and the sixth factor (eigenvalue: 1.01), explaining 6.3% of the variance, including items on job security.

Cronbach's alpha for the overall job satisfaction scale was .71. For the factors 1 to 6 coefficient alphas ranged from .47 to .73.

*Table A1: Items and factor loadings for the job satisfaction scale*

| Factor | Items | Loading |
|---|---|---|
| **Factor 1: Career advancement and recognition** | (Alpha: .65) | 4 |
| Teaching provides possibilities for promotion | | .79 |
| Teaching provides ample career development opportunities | | .80 |
| I have the opportunity to participate in decision-making on my institution's policies | | .49 |
| I receive recognition for my work as an educator | | .49 |
| **Factor 2: Peer support** | (Alpha: .73) | 3 |
| I get along well with my colleagues | | .82 |
| My colleagues and I support each other | | .87 |
| My colleagues and I are united in our dedication towards teaching | | .72 |
| **Factor 3: Teaching structure (working hours/load/policies)** | (Alpha: .47) | 3 |
| My workload is not too high | | .78 |
| I am satisfied with the content of the policies that affect my job | | .50 |
| Teaching offers reasonable working hours (despite extra-curricular activities) | | .65 |
| **Factor 4: Discipline and respect** | (Alpha: .58) | 2 |
| My learners respect me as a educator | | .82 |
| Most of my learners are well disciplined | | .82 |
| **Factor 5: Community enhancement** | (Alpha: .64) | 2 |
| Teaching provides me with opportunities to assist in shaping the future of young people | | .83 |
| Teaching provides me with opportunities to empower people with meaningful knowledge and information | | .85 |
| **Factor 6: Job security** | (Alpha: -.42) | 2 |
| Teaching provides me with job security | | .56 |
| I am afraid that I will be forced to take up a teaching position in an area/school/college where I do not want to teach | | -.89 |

### Job stress index

A 6-item job stress index was developed using items from existing scales (for example, Boyle, Borg, Falzon & Baglioni 1995; Fimian & Fastenau 1990; Spielberger 1994;

Spielberger & Reheiser 1994), focus groups with educators, expert interviews, and pilot study analysis of the Job Stress Scale (Spielberger 1994), for example, 'I earn an inadequate salary' (rated from 1=disagree to 3=agree).

In the main study, principal component analysis with varimax rotation yielded three components accounting for 67% of the total variance. The first factor (eigenvalue: 1.94) accounted for 32.4% in the variance in the responses and contained items concerning problems with teaching methods and administration. The second factor (eigenvalue: 1.09) accounted for 18.2% in the variance in the responses and included items related to problems with the educational systems. The third factor (eigenvalue: 1.002) accounted for 16.7% of the variance in the responses and included items related to low socio-economic status.

Cronbach's alpha for the overall job satisfaction scale was .52. For the factors 1 to 6 coefficient alphas ranged from .09 to .82.

*Table A2: Items and factor loadings for the job stress scale*

| Factor | Items | Loading |
|---|---|---|
| **Factor 1: Problems with teaching methods and administration** | (Alpha: .82) | 2 |
| I experience stress arising from the implementation of new curricula, pass requirements and reporting systems | | .91 |
| I experience stress with the preparation/assessment involved in applying the OBE approach | | .91 |
| **Factor 2: Problems with educational system** | (Alpha: .32) | 2 |
| Performing tasks not in my job description | | .83 |
| I experience negative attitudes towards the education department | | .61 |
| **Factor 3: Low socio-economic status** | (Alpha: .09) | 2 |
| I earn an inadequate salary | | .63 |
| The teaching profession needs more status and respect from the community | | .81 |

### Educator support index

A 9-item educator support index measured support in relation to the educator's role as educator (4 items) and for AIDS work or education (5 items), for example, 'I have the support of the school governing body for AIDS work/education'. Items were rated from 1=disagree to 3=agree.

Cronbach's alpha for this index was .82 for this sample.

**Violence at school index (Cronbach's alpha .68)**

A 9-item violence at school index was developed on the basis of literature review, focus groups of educators, expert interviews and pilot study analysis in order to measure the occurrence of the most common forms of violence in and on the way to school in the past 12 months, for example, 'Learners and/or educators have been found carrying weapons onto your institution's premises' (response options were 'yes' or 'no').

Cronbach's alpha for this index was .68 for this sample.

**Alcohol use scale, AUDIT**

Alcohol use was assessed with the ten item Alcohol Use Disorder Identification Test (AUDIT) (Babor, Higgins-Biddle, Saunders & Monteiro 2001). Standard drinking units were adjusted to the South African context (one unit 12 g alcohol), and sex differences were included for binge drinking, namely for men five or more and for women four or drinks on one occasion.

Cronbach's alpha for the AUDIT in this sample was .78.

Other substance use was assessed using two sections of the The Alcohol, Smoking and Substance Involvement Screening Test (ASSIST) (Henry-Edwards, Humeniuk, Ali, Poznyak, & Monteiro 2003). The two sections included past three months substance use (question 2), failure to do what expected to do in the past three months due to substance use (question 5), and the use of drugs by injection (question 8). In the use of tobacco products question 'snuff', as a commonly used tobacco product in South Africa (see for example, Peltzer 1999) was added and in the cannabis question the term 'dagga', a common term used in South Africa for 'cannabis', was added.

**Self-efficacy scale for HIV risk behaviour**

A 4-item self-efficacy scale for HIV risk behaviour was developed on the basis of existing scales (Smith, McGraw, Costa, & McKinlay 1996), literature review, focus groups of educators, expert interviews and pilot study analysis in order to measure self-efficacy of condom use in social situations, for example, 'If you decide to have sex, how sure are you that you would have a condom with you when you need it?' (response options were rated from 1=not sure at all to 3=totally sure).

Cronbach's alpha for this scale was .78 for this sample.

**Health-related quality of life**

The CDC Health Related Quality of Life (HRQOL-4) includes a global question that measures self-reported health (that is, 'Would you say that in general your health is excellent, very good, fair or poor?). Also included in the HRQOL core are two questions reporting the number of days during the previous 30 days in which the respondent's physical or mental health was not good. The sum of these two measures results in the total number of 'unhealthy days' (ranging from 0 to 30 days). The fourth question of the HRQOL core measures the number of days in the previous 30 days in the respondent

experienced activity limitation because of poor physical or mental health (CDC 2000). The CDC HRQOL-4 measure had an acceptable test-retest reliability and strong internal validity in a representative sample in the US (Andresen, Catlin, Wyrwich & Jackson-Thompson 2003).

### HIV/AIDS knowledge index

A 14-item HIV/AIDS knowledge index was developed on the basis of existing scales (Carey & Schroder 2002), literature review (Nelson Mandela/HSRC Study of HIV/AIDS 2002), focus groups of educators, expert interviews and pilot study analysis in order to measure HIV transmission and prevention knowledge (6 items, for example, 'People can protect themselves from HIV by using a condom correctly every time they have sex'), HIV/AIDS myths (4 items, for example, 'Having sex with a virgin can cure HIV/AIDS'), knowledge on HIV and TB (2 items, for example, 'Patients with TB also have HIV'), and knowledge about antiretroviral treatment (2 items, for example, 'Once one has started taking antiretroviral treatment for HIV/AIDS one has to take it forever'). Response options were 'true', 'false' or 'do not know'.

Cronbach's alpha for this index was .52 for this sample.

### HIV and sexuality communication comfort index

Five items on how comfortable one feels about HIV and sexuality education was developed from existing scales (Boscarino & DiClemente 1996) and focus groups with educators, for example, 'I am comfortable talking to learners about sexual matters', rated from 1=agree to 3=disagree.

Cronbach's alpha for this scale was .61 for this sample.

### HIV risk perception scale

The Health Belief model and Social-Cognitive learning theory both stress the importance of perceptions about the seriousness of a health threat, perceptions about one's personal vulnerability to a health threat, and one's perceived ability to reduce one's risk, as key determinants of health behaviour. A 3-item HIV risk perception scale was developed on the basis of literature review (Bengel, Belz-merk & Farin 1996) and focus groups with educators, including one item on personal HIV risk, one item on community HIV risk and one item on seriousness of HIV/AIDS in the community, for example, 'How likely is it that you will become infected with HIV?' Response options included a 5 or 6-point Likert scale from 1=no risk at all to 6=extremely at risk.

Cronbach's alpha for this scale was .58 for this sample.

### Tuberculosis social distance scale

The degree of prejudice against people with TB was measured with a 5-item social distance scale (intention to engage in physical and social contact with people with tuberculosis). Response options were 'yes', 'no' or 'do not know'. Responses to items

were summed so that the higher the score, the more socially acceptable were the people with TB considered by the respondents surveyed, that is, the more willing the respondents were to engage in social contact by sharing meals, working/studying or by physical contact such as hugging, kissing or having sexual relationships with people with TB (Jaramillo 1999).

The scale had an internal consistency of .69 (Cronbach's alpha).

**HIV/AIDS stigma index**

A 5-item stigma index was developed on the basis of literature review (FHI 2000; Nelson Mandela/HSRC Study of HIV/AIDS 2002) and focus groups with educators. Only items with a direct possible experience with a person living with HIV/AIDS were included in relation to the family (for example, 'Did you care for a family member who was ill due to AIDS?') and school situation (for example, 'Have you avoided a colleague who disclosed his/her positive HIV status to you?') Response options were 'yes' or 'no'. These questions were to be analysed with three corresponding questions, namely knowing a family member, learner or educator living with HIV or AIDS.

Cronbach's alpha for this scale was .68 for this sample.

## Appendix 2: Laboratory HIV testing procedures

### HIV Testing

All blood specimens were tested for HIV on the Abbott AXSYM third generation HIV 1 / 2 g0 testing system. There are many advantages of the AXSYM instrument over the old testing system, including the fact that testing of the primary tube is done directly, there is a barcode reading facility and it is fully automated. This has improved the turn-around time, eliminated the potential for pipetting and other 'human' errors and the instrument 'down-time' has been negligible. The instrument maintenance plan is practical and the instrument has proven to be very user-friendly.

The following points explain how quality control of the blood testing was ensured:

- Kit controls: HIV 1 and HIV 2 positive control, HIV negative control, which are run daily;
- Additional commercial internal control: low positive HIV positive Accurun (BBI) control (also run daily);
- A subset of AXSYM positive and negative samples are retested on a different second line 3rd generation Elisa (BioRad Genetic System HIV 1 Elisa);
- All weakly reactive AXSYM results are tested on the BioRad second-line Elisa before a result is provided. If the BioRad is positive, the sample is considered positive and if negative on the BioRad, the sample is considered negative;
- Regular servicing of the AXSYM by Abbott;
- Performance of daily, weekly and monthly maintenance of the AXSYM instrument as recommended by the supplier.

In addition, CLS is also enrolled in a national (NHLS) and international (NRL and CAP) external quality control programs for regular testing of HIV proficiency panels. CLS has achieved 100% scores since enrolment in these programmes.

Participants who did not wish to provide a blood specimen instead provided a specimen of oral fluid, which was obtained by using the 'Orasure' oral fluid collection device. These specimens were tested using the Vironostika HIV Uni-Form II Oral Fluid testing system. This is an Elisa for the qualitative determination of antibodies to human immunodeficiency virus type 1 and/ or 2 (anti-HIV-1, anti-HIV-2 and anti-HIV-1 group O) only with oral fluid specimens.

Quality control for this testing system consists of the following:

- Five negative controls are run on each plate: two in the first row, one each in the fourth, seventh and tenth rows;
- Two positive controls (one high and one low positive): these are run in the last 2 wells of the 12th row;
- No international EQA programs for oral fluids are currently available.

## Appendix 3: Evaluation of the age-sex distributions

Analysts often take the quality of survey data for granted and do not provide sufficient appraisal of the quality of the data. The quality of the demographic variables in a survey often provides an insight into the quality of the survey data in general.

Two important variables in this regard are age and sex. Because of its cross-cutting nature with regard to planning, age and sex information are collected in virtually all surveys. Apart from being directly or indirectly linked to all socio-economic phenomena and planning, the quality of the age-sex distribution from a survey often provides an indication of the quality of various aspects of the data. Although the quality of HIV testing may not necessarily be related to age, invariably such test results are presented by age, and if there errors in age-sex reports, the distribution of HIV prevalence by age and other age related variables in the data might be distorted. Age-related estimates from the survey data might therefore be in error. Confidence intervals though give some indications of the precision of the sample related estimates; they do not provide clues about content errors which are sometimes more substantial in survey data than sampling errors (for example, age-misreporting, tendency to report one's marital status as widowed when actually divorced). In view of these reasons, a critical appraisal of the age-sex distribution is essential.

This section provides an appraisal of the age-sex distribution of educator sample.

### Methods

There are several approaches that may be used in appraising the quality of an age-sex distribution:

*Internal consistency analysis:* This entails examining the quality of the age-sex distribution within the survey data and, if necessary, comparison with other demographic phenomena within the data (for example, fertility and mortality). The question usually posed by the analyst in this approach is: 'Are the data internally consistent?' For example is the age-sex distribution consistent with fertility and mortality? (An age-sex distribution is determined entirely by fertility, mortality and migration). A thorough consistency check cannot be performed on the present data because the demographic information is limited as it was not the objective of the survey to collect detailed demographic information.

*External consistency check.* This entails comparing the age-sex distribution obtained from the survey with other external sources. The external sources chosen for comparison (often what are available) do not necessarily mean these are 'gold standards' as similar patterns of error might operate in the external sources. Even when this is the case, the approach is still useful in gauging the magnitude of probable errors in the survey data relative to those in the external source.

*Use of reference standards.* Several standard age distributions are available (see Udjo forthcoming) for evaluating observed age-sex distributions. Such reference standards were constructed for the general populations and not for specific occupation groups within populations. The use of reference standards is therefore not useful in the present study since the sample was based on educators.

The appraisal of the age-sex distributions of the sample therefore combines a limited internal and external consistency check as described above. The external source is Statistics South Africa's General Household Survey (GHS) (Stats SA 2003). From the occupation question in the 2003 GHS, educators can be extracted and tabulated by age and sex. However, because of the nature of the coding system employed in both data sets, and the universe of study in the ELRC data, this part of the analysis is confined to educators in primary, secondary and special institutions. Statistics South Africa utilised the International Standard Classification of Occupations (ISCO), which is very detailed. The ELRC study did not include university educators and the coding classifications employed is not as detailed as the ISCO coding; hence to avoid overlap and misclassification, the analysis of the GHS and ELRC data were restricted to educators in the institutions indicated above. The coverage of educators in the GHS and ELRC surveys is also different in another important aspect. While the GHS covers educators in both private and public institutions, the ELRC only covers educators in public schools. According to figures from the DoE, however, educators in private institutions constituted only 4% of educators in South Africa in 2001 (see DoE 2003). It may be argued that this 4% would only make a negligible impact in the overall distribution of educators in the country and therefore, one could use the distribution of educators in the GHS survey as an external check on the distribution of educators in the ELRC survey in primary, secondary and special institutions.

The appraisal of the quality of the age-sex distributions was based on examination of the median age, overall sex ratio (number of males per 100 females), single- and five-year age-sex distributions.

**Median age**

The median age is a summary measure of age distributions and is one of several indicators of the 'youngness' of a population distribution. The observed value of the median age of a population can be used to evaluate the accuracy of reported age distributions on the basis of a priori knowledge of fertility levels in that population. Although the present study was confined to persons who would normally be in a restricted age range (that is, defined categories) rather than all ages in the population, the accuracy of the observed median age of defined categories (such as workers and educators) could be inferred from a comparison of the observed mean age of the defined categories from other external sources of data of the same population. Such comparison is shown in Table A3 below.

As seen in the table, the median age of South Africa's population – as observed in the 2003 GHS (Stats SA 2003) – was 23 years. This value is plausible given the trend and prevailing levels of fertility in the country, and is somewhat consistent with the median age of the population in 2004 (24 years) estimated by Udjo (2004). On this basis, we assume that the median age of defined categories in the population is also plausible and hence this external source of information may be used to evaluate the observed values from the educators study.

As seen in Table A3, the observed median age of educators in the 2004 ELRC study is 39 years. This is consistent with the corresponding value observed in the 2003 GHS (41 years), though the former is two years lower than the former. A similar consistency is shown when the median age of educators from the general household survey is compared with the corresponding values from the ELRC study by population group

and province. In particular, it is a happy coincidence that the observed median age of educators from the 2003 GHS and the 2004 ELRC study are the same for Western Cape, Eastern Cape, North West, and Limpopo provinces (see Table A3).

*Table A3: Median age (years) of educators in the 2003 GHS and 2004 ELRC surveys*

| | Median age | |
|---|---|---|
| | GHS 2003** | ELRC 2004*** |
| Total (all persons) | 23 | — |
| Total (educators)* | 41 | 39 |
| **Population group**: | | |
| Africans | 40 | 39 |
| Coloured | 43 | 41 |
| Indian | 43 | 41 |
| White | 45 | 44 |
| **Province**: | | |
| Western Cape | 41 | 41 |
| Eastern Cape | 39 | 39 |
| Northern Cape | 41 | 39 |
| Free State | 44 | 42 |
| KwaZulu-Natal | 40 | 38 |
| North West | 39 | 39 |
| Gauteng | 43 | 40 |
| Mpumulanga | 43 | 40 |
| Limpopo | 40 | 40 |

** Educators in primary, secondary and special institutions*
***Source: Stats SA General household survey (2003)*
**** ELRC Educators survey (2004)*

### Single-year age distribution

The median estimates discussed above provide an overview of the 'youngness' and quality of the observed distributions but do not enable detailed examination of the quality of the observed age distributions. More detailed evaluation of the quality of the observed age distributions can be gained through examination of single-year age distributions. The observed single-year age distributions in the 2003 general household survey and the 2004 ELRC study are shown in Figures A1 and A2 for males and females. One common error in age reporting is age heaping or digit preference (that is, the tendency to report ages ending in certain digits). This kind of age-misreporting if present, can be detected in graphical presentation of single-year age distributions. As seen in Figure A1, heaping

in ages ending in zeros and even numbers (with the exception of an unusual heaping on age 43) was quite marked among males in the 2003 GHS survey as evident from the erratic pattern or large 'peaks' in the single-year age distribution. In contrast, with the exception of a few small peaks, heaping is hardly discernible among males in the 2004 ELRC study. In contrast to the 2003 GHS, the observed single-year age distribution for males in the 2004 ELRC study is fairly smooth thus indicating that the quality of the single-year age distribution for males is a lot better than in the 2003 GHS. The comparison of the observed single-year age distribution for females in the 2003 GHS and the 2004 ELRC study (Figure A2) shows a similar pattern as the comparison for males. The distributions indicate that the quality of the single-year age distribution for females is a lot better than in the 2003 GHS.

*Figure A1: Reported single-year age distribution of educators in the general population and the ELRC study, males*

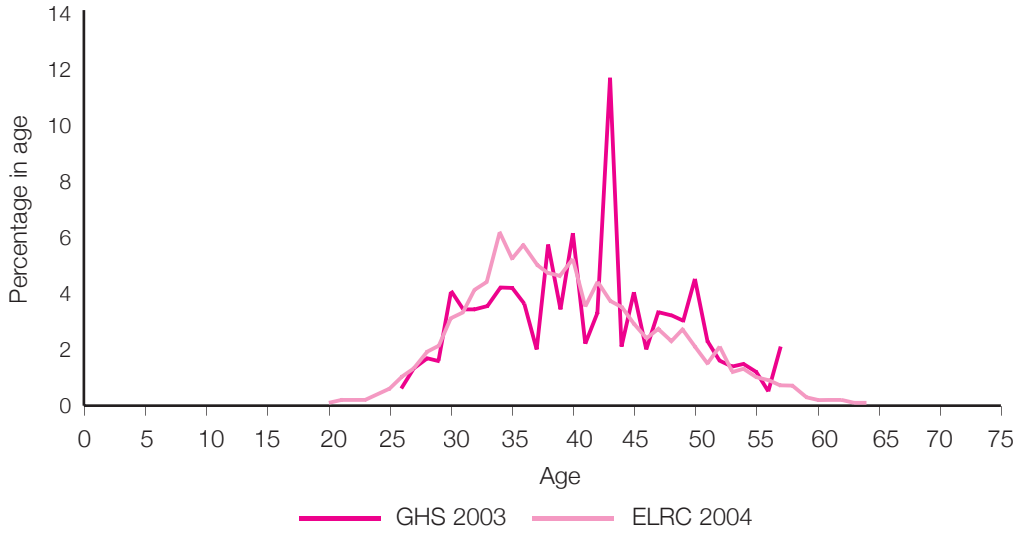
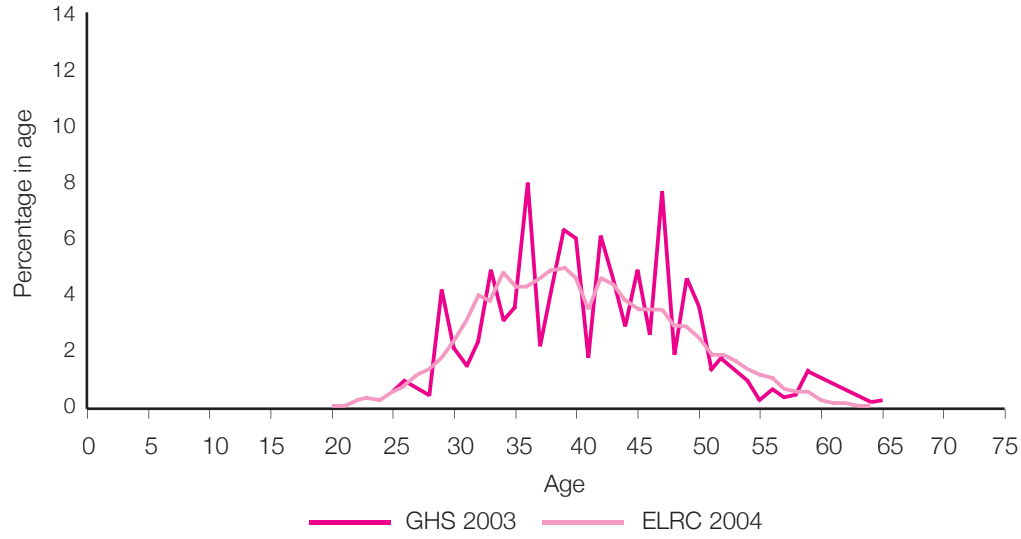



*Figure A2: Reported single-year age distribution of educators in the general population and the ELRC study, females*



**Five-year age distributions**

For practical purposes, analysis of age-related variables of survey data are often done in broad age groups as variables cross-classified by single years of age are too cumbersome for analytical purposes though they provide useful insight about the quality of the data. Commonly, variables are cross-classified by five-year age groups and sometimes by ten-year age intervals or other intervals. Cross-classification of variables in five-year age or other intervals also has the advantage of self-smoothening as some of the erratic fluctuations in single-year age cross-classifications may be minimised though age-shifting (tendency to report one's true age group in another age group) may be present in five-year age distributions. For these reasons, Figures A3 and A4 compare the observed five-year age distributions of educators in the 2003 GHS and in the 2004 ELRC study by sex. The five-year age distributions for male educators in the 2003 GHS and the 2004 ELRC show some slight differences. Since the time period between the two surveys is very short (approximately one year apart) the proportions in each age group should be almost similar in the absence of errors in the two surveys because there is very little overlap in the age cohorts. However, one observes from Figure A3 that the proportion of male educators aged 35–39 in the 2004 ELRC study is about 6% higher that the corresponding proportion in the 2003 GHS. Also, the proportion of male educators aged 40-44 in the 2004 ELRC study is about 5% lower than the corresponding proportion in the 2003 GHS. A close examination of the bulge in the five-year age distribution of educators at age 40–44 in conjunction with the single-year age distribution in the 2003 GHS suggest that the bulge is largely due to age shifting from adjacent or other age groups. (Note from Figure A1, the relative excess males at age 43). On this basis, it would appear that the 2004 ELRC study provides a better indication of the proportion of male educators aged 40–44 than the 2003 GHS. On the other hand, the bulge in the five-year age distribution of male educators at age 35–39 in the ELRC study is suggestive of age shifting on that age group, thus probably an exaggeration of the 'true' proportion of male educators in that age group. Other than these two age groups (40–44 in the 2003 GHS and 35–39 in the ELRC study) and excluding the youngest and oldest age groups, the proportion of male educators in five-year age group in the ELRC study is fairly consistent with the corresponding proportion in the 2003 GHS. In general, the distribution of male educators in five-year age groups shows a relatively more regular and hence better quality in the ELRC study than in the 2003 GHS.

*Figure A3: Five-year age distribution of male educators, GHS 2003 and ELRC 2004*

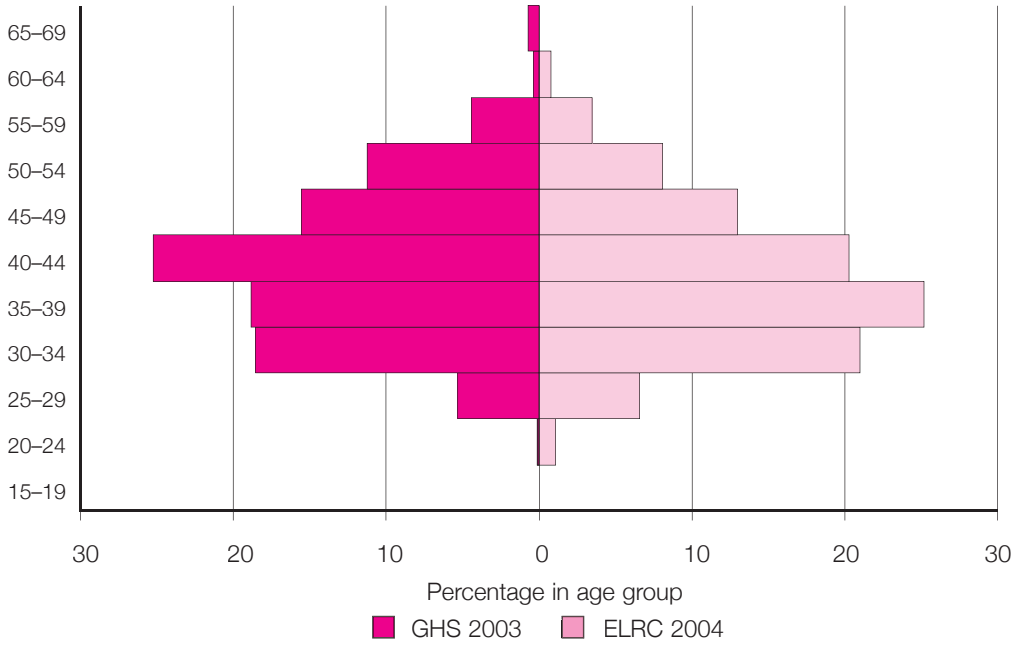
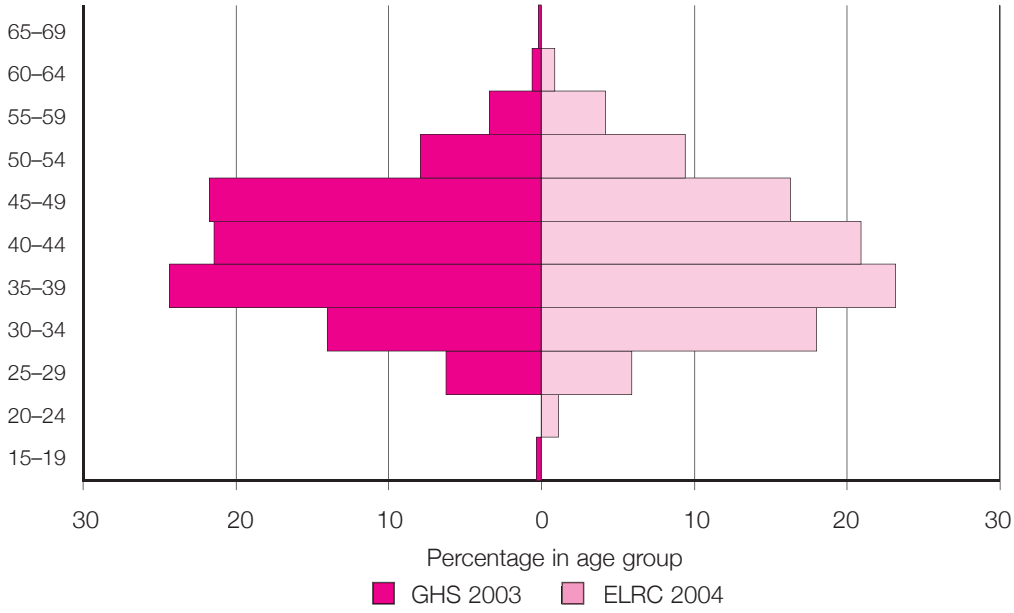



*Figure A4: Five-year age distribution of female educators, GHS 2003 and ELRC 2004*




With regard to females, the five-year age distributions from the 2003 GHS and the 2004 ELRC study (Figure A4) are not consistent with each other at the 30–34 and 45–49 age groups. The proportion of female educators aged 30–34 in the 2004 ELRC study is about 4% higher than the corresponding proportion in the 2003 GHS, while at age 45–49, the proportion of female educators in the 2004 ELRC study is about 5% lower

than the corresponding proportion in the 2003 GHS. There is a bulge in the five-year age distribution at 35–39 in both the 2003 GHS and the 2004 ELRC study, largely arising from age shifting and consequently resulting in a proportionate deficit of females at age 30–34 (relative to males in that age group). This is, however, less marked in the 2004 ELRC study compared to the 2003 GHS. Aside these seemingly aberrations and excluding the youngest and oldest age groups, the proportion of female educators in five-year age group in the ELRC study is fairly consistent with the corresponding proportion in the 2003 GHS. In general and similar to males, the distribution of female educators in five-year age groups shows a relatively more regular and hence better quality in the ELRC study than in the 2003 GHS.

**Overall sex ratios**

Another aspect of evaluating age distributions is examination of observed sex composition of the age distributions. A summary measure in this regard is overall sex ratios and is a powerful indicator of the quality of observed sex composition of a population given knowledge about fertility, mortality and migration in that population. One needs to be cautious however when comparing overall sex ratios in the 2003 GHS with those in the 2004 ELRC survey in view of controversies around observed sex ratios from surveys and censuses in South Africa (see, for example, Dorrington 1999; Phillips, Anderson & Tsebe 1999; Sadie 1999; Shell 1999; Udjo 1999). In a recent work for example, Udjo (forthcoming) has argued that observed overall sex ratio from the October household surveys (now replaced by the general household surveys) and censuses of 91 in South Africa (derived from the listing of members of household by sex) are not reliable as it exaggerates the number of females in the South African population. Furthermore, Udjo argues that since the census results (including overall sex ratios) have been used for deriving weights for sampling and subsequently weighting of survey data, the error in sex ratios are carried over to survey data (and indicators derived from them) since sex is either used as an explicit or implicit stratification variable in sampling and weighting.

The observed overall sex ratio of educators in the specified institutions from the 2003 GHS and the 2004 ELRC study are shown in Table A4. As seen in the table, the observed overall sex ratio is low (91) suggesting substantial excess females over males in the population and is inconsistent with fertility, mortality and migration levels and trend in the population (see Udjo forthcoming). Table A4 also indicates that the overall sex ratio among educators on the basis of the 2003 GHS is much lower (64) than in the general population (91) suggesting large excess female over male educators. Is the excess females in the 2003 GHS entirely due to the nature of the occupation (that is, female-dominated occupation) or a combination of errors in the data and the nature of the occupation? The 2004 ELRC study shows even lower overall sex ratio (49) than in the 2003 GHS (64).

*Table A4: Overall sex ratios of educators*

| | GHS 2003 | ELRC 2004 |
|---|---|---|
| Total (all persons) | 91.4 | — |
| Total (educators) | 64.1 | 48.5 |
| **Population group:** | | |
| Africans | 66.6 | 51.8 |
| Coloured | 75.4 | 51.2 |
| Indian | 102.7 | 51.3 |
| White | 46.7 | 26.2 |
| **Province:** | | |
| WC | 113.2 | 47.5 |
| EC | 67.1 | 38.9 |
| NC | 24.3 | 47.0 |
| FS | 62.4 | 55.2 |
| KZN | 82.0 | 43.0 |
| NW | 91.3 | 46.6 |
| GT | 35.8 | 38.4 |
| MP | 52.7 | 43.1 |
| LP | 68.1 | 77.9 |

Large differences in overall sex ratio of educators can also be seen when the 2003 GHS is compared with the 2004 ELRC study by population group and province. For example, the observed overall sex ratio in the 2003 GHS is at least twice the overall sex ratio in the 2004 ELRC study among Indian educators and educators in the Western Cape. This pattern appears to suggest that the 2003 GHS had a wider coverage of male educators (relative to female educators) than in the 2004 ELRC study. Part of the explanation for the difference in coverage may be due to the fact that whereas the educators in the 2003 GHS were from private and public education institutions, those from the ELRC were from only from public schools and there are probably more female educators.

## Conclusion

Although the information pertaining to educators in the 2003 GHS is by no means a 'gold standard', it provides a useful external gauge of the quality of the observed age-sex distribution from the 2004 ELRC study. This external consistency check suggests that the age distribution from the 2004 ELRC study is of reasonably good accuracy: The median age of the age distribution from the 2004 ELRC study is plausible and consistent with those from the 2003 GHS. The observed single-year age distribution from the 2004

ELRC study is fairly smooth and of better quality than the 2003 GHS. The five-year age distribution from the 2004 ELRC study is also more regular and of better quality than the 2003 GHS. However, the observed sex composition of educators suggest that the 2003 GHS apparently had a wider coverage of male educators than in the 2004 ELRC study. Since the quality of age-sex distribution from a survey provides an indication of the quality of various aspects of the data, it might be inferred from the results of the age analysis that the 2004 ELRC data is of reasonable good quality.

## Appendix 4: Reliability and validity of HIV prevalence rate, socio-demographic profiles, coefficient of variation and the design effects

| Variables | n | Prevalence Rate (r) | SE_r | CV_r | Deff | Deft |
|---|---|---|---|---|---|---|
| **Sex** | | | | | | |
| Male | 5 405 | 12.75 | 0.57 | 0.045 | 1.72 | 1.27 |
| Female | 11 467 | 12.76 | 0.42 | 0.033 | 1.98 | 1.36 |
| **Age in years** | | | | | | |
| 25–35 years | 5 032 | 20.95 | 0.75 | 0.036 | 1.84 | 1.31 |
| 36–49 years | 9 167 | 10.64 | 0.41 | 0.039 | 1.77 | 1.29 |
| 50 years and above | 2 441 | 4.02 | 0.47 | 0.117 | 1.44 | 1.16 |
| **Race group** | | | | | | |
| African | 11 854 | 16.30 | 0.41 | 0.025 | 1.73 | 1.27 |
| Whites | 2 151 | 0.42 | 0.13 | 0.310 | 0.73 | 0.83 |
| Coloureds | 2 298 | 0.71 | 0.21 | 0.296 | 0.93 | 0.93 |
| Indian/Asian | 523 | 0.98 | 0.39 | 0.398 | 1.25 | 1.08 |
| **Type of school** | | | | | | |
| Primary | 9 402 | 12.34 | 0.48 | 0.039 | 2.01 | 1.37 |
| Secondary/High | 5 939 | 12.48 | 0.69 | 0.055 | 3.18 | 1.72 |
| Combined /Intermediate | 1 439 | 16.54 | 1.53 | 0.093 | 2.40 | 1.50 |
| **Location of institution** | | | | | | |
| Urban formal | 6 928 | 6.29 | 0.52 | 0.083 | 2.97 | 1.66 |
| Urban informal | 1 100 | 13.91 | 1.32 | 0.095 | 1.56 | 1.21 |
| Non-urban or rural | 8 779 | 16.77 | 0.51 | 0.030 | 1.89 | 1.33 |
| **Race and sex** | | | | | | |
| African – males | 3 847 | 15.91 | 0.68 | 0.043 | 1.57 | 1.21 |
| Other – males | 1 558 | 0.89 | 0.27 | 0.303 | 0.98 | 0.96 |
| **Race and sex** | | | | | | |
| African – females | 7 998 | 16.46 | 0.50 | 0.0304 | 1.69 | 1.26 |
| Other – females | 3 469 | 0.75 | 0.16 | 0.21 | 1.02 | 0.97 |

| Variables | n | Prevalence Rate (r) | SE_r | CV_r | Deff | Deft |
|---|---|---|---|---|---|---|
| **Marital status** | | | | | | |
| Single | 4 525 | 22.85 | 0.75 | 0.033 | 1.63 | 1.23 |
| Married | 12 302 | 8.82 | 0.35 | 0.040 | 1.97 | 1.36 |
| **Position at school** | | | | | | |
| Junior | 12 523 | 14.05 | 0.43 | 0.031 | 2.07 | 1.39 |
| Senior | 4 034 | 8.73 | 0.54 | 0.062 | 1.59 | 1.22 |
| **Province** | | | | | | |
| WC | 2 128 | 0.99 | 0.29 | 0.293 | 1.46 | 1.17 |
| EC | 1 845 | 13.81 | 0.96 | 0.070 | 2.46 | 1.52 |
| NC | 888 | 4.35 | 0.90 | 0.207 | 0.56 | 0.72 |
| FS | 1 123 | 12.49 | 1.27 | 0.102 | 1.68 | 1.25 |
| KZN | 3 566 | 21.90 | 1.05 | 0.048 | 2.77 | 1.61 |
| NW | 1 411 | 10.51 | 0.97 | 0.092 | 1.32 | 1.11 |
| GT | 2 724 | 6.47 | 0.61 | 0.094 | 1.63 | 1.13 |
| MP | 1 303 | 19.03 | 1.54 | 0.081 | 1.96 | 1.35 |
| LP | 1 896 | 8.54 | 0.72 | 0.084 | 1.84 | 1.31 |

*Note: Abbreviations used: N = number of cases in the sample, r = response rate, SE_r = standard error of the response rate, CV_r = coefficient of relative variation, Deft = design factors (square root of Deff), and Deff = design effect*

## Appendix 5: List of fieldwork supervisors, interviewers and coders

*Provincial co-ordinators*

| | |
|---|---|
| Ramlagan Shandir | KZN |
| Van Wyk Brian | NC |
| Henda Nomvo | WC |
| Petros George | MP |
| Nqeketo Ayanda | EC |
| Louw Julia | GT |
| Zungu-Dirwayi Nompumelelo | WC |
| Letlape Lebogang | FS |
| Francis Shantinie | KZN |
| Tamasane Tsiliso | NW |
| Thaba Fhumulani | LP |
| Sithole Nhlanhla | GT |

*Fieldwork supervisors*

| | |
|---|---|
| Arendse Dorothy Sophia | Mia Shahnaaz |
| Anthony P | Mkansi Gaza Louwy |
| Arrison Marlene | Mojapelo Julia Lifutso |
| Botha Elsie | Moleme BJ |
| Da Gama Harmony | Moodley Vulimal Kanama |
| Fynn R | Ndaba L |
| Galo Julia Tabina | Nel Elmarie |
| Gilday Colleen | Nkoana-Makhetha Perpetua |
| Gum Mothiba Ethel | Nqqovu Weziwe |
| Irwin D | Phooko Matilda Mashaw |
| Kgwedi Dimakatso Octavia | Ramashala Modikane Jane |
| Khosa Nikiwe Sellina | Schnoor JJ |
| Lite Eric Clifford | Scholing MJ |
| Lokwe Theresa | Sekgobela Njale C |
| Maake Rosina M | Thwane Namokonyane Martha |
| Magubane F | Tshipa Gertrude |
| Masisi SS | Tuku-Stuurman Letty Thandiswa |
| Matheson Mauveen | Tyalimpi Norah Nandipha |

*Interviewers*

Adams Ena Nancy
Amadlel L
Amod Eugene Bernice
Arendse Dorothy Sophia
Baloyi A
Barnaschone LE
Beelders Luddy Evelyn
Benefeld Gameda
Bless Sheila
Bloko Nosisa Albertina
Boikanyo Sylvia Bosiswe
Bopape E
Bopape Edith
Boshoff Stephanie C
Briekwa Mina Barbara
Buckton Martha Magdalena
Burger Anna Mara
Buthelezi H
Calvert Faith Virginia
Cebekulu H
Cebekulu Nonhlanhla D
Crouch Dorah
Damonse G
Dasa Lillian Singiswa
De Bruin Elna
Diago Makwadi E
Diaz Gideon
Dike Pauline
Dipela C
Dipela Catherine
Diseko Malebato Margaret
Dladla D
Du toit Hilda Maria
Duba ND
Dube Eunice Matshiliso
Duma N
Dziba Nonkonzo Nancy
Edwana Wandisa Edna
Ellis Tandi Olga Margaret
Els SW
Francis Joy Frances
Gabohele V
Galo Julia Tabina
Galo Nozipho Benedicte
Ganie F
Gelant Jennipher Jean
Geldenhuys Mamona Ester
Ghela M
Gidagaj G
Gidagaj Getrude
Gounden B
Govender M
Govender S
Govinden A
Gqabi N A
Gumede E
Gumede Ntombile Adelaide
Gwayi Ayanda
Gxoyiya Nomatolo Millicent
Havenga Aletta M
Hayes Martha Cornelia
Hermanus Sarah
Hlabeli KJ
Hlengwa F J
Hlubi Rose
Israel SR
Jacobs Melvina Nicolette
Jali J
James Sophia Elizabeth
Johnson Blossom

Joubert Rosina Martalena
Joyi Sannah
Jwaai Mary Magaret N
Kanyile BJ
Kgaphole Joan Malefula
Kgwetiane M
Kgwetiane Mmadifedi
Khanyile E
Khume Miriam Linkeng
Khutoane Makojwana Sophy
Kitsa MS
Koma Tsiane Johanna
Krog J
Kubeka C
Kubeka Sybil M
Kwabe J
Lamprecht MSI
Laubscher Anna C E
Le John L
Ledwaba Freda Ntombizodwa
Lehaiwa J
Lehaiwa Jane
Lekubu Mmatsebe G
Lekwakwe Evatonia D
Leoane TA
Letsoalo G
Letsoalo Gladys
Lewaba Rebone N
Lewaba RN
Liebenberg G U B
Lobelo Mpho Gloria
Lombard Anneli
Lutula N
Mabaso K
Mabaso Khatazile
Mabelane Morufe Eva Ingrid
Mabizela GB
Mabote Boniswa M
Mabunda Nellie Dorothy
Mabunda Sizeni
Mabunda Tsakani Julia
Madihlaba E
Madihlaba Eva
Madiope Christene Madipofa
Mafa Hlamalahi Grace
Mafekiso Nada Nakanye
Magoro Johanna Mmemeng
Mahlaba Gloria R
Mahlaba GR
Mahlangu Matshediso Caroline
Mahonga Gladys Nomgcobo
Maidi Johannes Mfanyana
Maite Kathleen Maud Makgogo
Makathini B
Makhafola Nontsikelelo E
Makhubele Josephine Juku
Malgas Zenzile E
Malinga S
Mangena M
Mangena Miriam
Mango Nobahle Virginia
Mangoale Refilwe Miriam
Manne Florence N
Maomela R
Maomela Rachel
Mapeyi Priscilla Nkosazana
Maphatchwane E
Marole D
Marole Dorah
Marole Elizabeth Nkepile
Mashaba Thembeni Audrey
Mashini EN
Mashini Rachel
Masingi Sheillah H

Masipa Motlatjo Josephine
Maswanganyi Mlunghisi S
Maswanganyi MS
Mathabeng C
Mathibe Mamolefe Rebecca
Matome M Daphne
Matsimella Joycelyn Cornelia
Maunye Theira J
Maunye TJ
Mawila Gabaza Dinah
May Jean Marion
Mbatha E
Mbatha M
Mbebe Gcebile Monica
Mbebe T Veronica Maria
Mbekewi J
Mbele J
Mbhewe T
Mbolekwa Nontsikelelo Miriam
Mboweni Mashoto C
Mchunu N
McKonie Vuyiswa Princess Marina
Mehl Jennifer Elizabeth
Meiring Mary Elizabeth
Mfazwe Patricia
Mgijima Nosipho Dorothy
Mgobhozi C
Mgole Pearl Thokozile
Mhlantla Nomalungelo Alicia
Mhlongo Rita Ruth
Mjacu Nombulelo Dorothy
Mkandawire Winnie Clara
Mkanzi Noluthando
Mkize P
Mngomezulu Mabocha Mary
Mnguni Anna Qothi
Mnotoza Monica N
Moagi D
Moagi Dorothy
Mochoari NA
Modiba Mamtla M
Modiba MM
Mogale Letta
Moholo N
Mokeyane CI
Mokoena Exelda
Mokoena TJ
Molege N
Molepo Nqwayana Damaria
Moloi Ntombikayise Martha
Monnye WM
Moodley C
Motale Francis S
Motau Mabele Bertha
Motemekoane Elias Mahlatsi
Mothabeng C
Mothiba Mokgaetji D
Motihodi Pule Aubrey
Motsei Thabitha Manakedi
Motumi Rakgadi L
Mpahlele Nondumiso Mavis
Mphuti Evelyn B
Mqadi M
Mqadi T
Mqhlangu Matshediso Caroline
Mseti Nolwandle E N
Msimang L
Msimang Theresa
Mthembu Betty Lizzie
Mtise Helena
Mtshakaza Nomana
Mushi Poppy Leah
Musi MC
Musi Mojele Caroline

Naicker P
Naidoo L
Nair Farieda
Nasando Virginia
Ndaba C
Ndabeni Frances Ntombazana
Ndema Thantaswa Elsa
Ndhlovu Eva
Ndlovu C
Ndlovu R
Ndzotyana Irene Tozama
Nene K
Ngalwa Lucy Lindy
Ngcelwane Patricia Ntombomzi
Ngcobo Aurelia ES
Ngcobo G
Ngidi B
Ngobeni Mantwana Magaret
Ngobeni Peggy Mokgaetji
Ngobo S
Nkele Mpheteng Shirley
Nkgoeng Ketiwe Gloria
Nkoane Julia
Nkosi Nomantombazana Doris
Nkuzo Eunice Nomvuyo
Nongogo NON
Nongxa Natalia N
Nqodi Nomalungi Louisa
Ntlatleng Nomalitho Ester
Ntoi F
Ntshinga Doris Balekwa
Ntuli Musa Christian
Nxokwana Funeka Asenath
Nxumalo D
Nyembenya Nomangwanya Florida
Nzimande N
Nzumande T
October Louisa
Ogle I
Olebogeng Rebeccah Maureene
Padi LM
Papane MS
Papo Nomhle SL
Papo NSL
Phala Thokozile Philda
Pillay H
Polinyane S
Porogo Kanate Joseph
Pule Josephine
Qelo Maria Nomtunzi
Qupe Nompumelelo Florence
Qwabe Octavia
Qwele Irene Funeka
Radebe Margaret Busisiwe
Radebe Rebecca Smangela
Ramagaga MA
Ramaila Sharon Lerato
Rampeng Johanna Gontse
Ranisi Sannah Thenjiwe
Resenga Jeanette M
Resenga JM
Rhode Ragma
Rikhotso KE
Rikhotso Khahlela E
Russel Keromang Jacobeth
Sambo M
Sambo Matheline
Schoonraad Jacobus Frederick Smith
Sebe Ntebaleng Michael
Sechoaro E
Seekoei Gaitswe Irene
Seforo Phoki Jeconnett
Sekgobela Njale C
Sekhu Malebo Hilda

Setshedi Nonceba Ethel
Setsubi Sebokane Josephine
Shange T
Shipalana Helen
Sibande Nonlanhla Cathrine
Sikiti NL
Simons SH
Sisusa Nonceba Elfreida
Sithole AN
Sithole Thabiso Esther
Sitsila ND
Skosana Eugenia
Soke Mapula F
Soke MF
Sopazi Beauty Nomthandazo
Stemela Nozipho Angelina
Suttie Alice Winsome
Tarwa U
Tarwa Unopa
Tau MM
Taye Vangile
Thabethe Ntombizodwa Jane
Thango L
Thinta Mbulelo Thomas
Thipe Mercy Semakaleng
Thoka SG
Tikolo Nobantu Olive
Timotheus Wilhelmina Johanna
Tlakula Nofumana Vivienne
Tlale ST
Tsaoane Nanaki Maria
Tshabalala S
Tshoana Matlale Caroline
Tsie G
Tsotetsi Madisebo Paulina
Tsubane E
van Wyk Anna Maria
Vilikazi Maria Rasi
Vollenhoven C
Wallis Isella ingrid
Wareham Elain Even
Xulu E
Yabo Florence Fihliwe
Yoyo Nozipho
Zim Thandiwe Kathleen
Zimu Monica Thokozile
Zonke-Bungane Lulama Henrietta

***Questionnaire coders***

Choane Charity P
Molefe Johannes R
Ntuli Esram M
Phalatse Neo C
Rapeu Thomas PS