Asian Journal of Population Sciences p-ISSN 2822-1613
(A Peer-Reviewed, Open Access Journal; Indexed in NepJOL) e-ISSN 2822-1621
Email: population@pncampus.edu.np
eJournal Site: http://ejournals.pncampus.edu.np/ejournals/ajps/

[ORIGINAL RESEARCH ARTICLE]

Evaluation of Age-Sex Data Collected in the 1991, 2001 and 2011 Population and Housing Censuses of Namibia

Immanuel Shipanga* and Tobias Shinyemba*
*School of Science, Department of Computing, Mathematics and Statistical Sciences, University of Namibia, Windhoek, Namibia.

Corresponding Author & Email
Immanuel Shipanga; ishipanga@unam.na

Submitted 19 August 2022; Reviewed 7 December 2022; Accepted 16 December 2022
DOI: https://doi.org/10.3126/ajps.v2i1.51087

Published by
Department of Population Studies
Prithvi Narayan Campus
Tribhuvan University
Pokhara, Nepal

Scan to Access eJournal

Abstract

This paper examined the accuracy of age-sex data reported in the Namibia population censuses with the aim to establish patterns and trends in the quality of age-sex reporting. To date, Namibia has conducted three population and housing censuses. However, it is not clear if there has been any advancement in the quality of data collected on age and sex over the years. The study followed a cross-sectional study design using the 1991, 2001 and 2011 population censuses data from Namibia Statistics Agency. We computed Whipple's and Myers's Blended indexes to measure age heaping. The age-sex ratio and UN joint score were used to ascertain the possibility of population undercounts and displacement between age groups. The results show that age heaping was fairly accurate in all three censuses. Analysis of single-age distribution showed a minimal digit preference and avoidance across all three censuses. Likewise, Myers blended index portrayed that terminal digits 0, 1, 2, 5, 8 and 9 were preferred the most by both sexes in all three censuses. The Whipple's Index showed a two-point decline from 106 in 1991 to 104 in 2001 and 2011 suggesting an improvement in age reporting. The UN joint score for Namibia found that the data collected in all three censuses were inaccurate in terms of age-sex displacement.

Keywords: Age heaping, Sex Ratio, Namibia, Digit preference, Age misreporting, Census, Data Quality

INTRODUCTION

A population census is the most direct way of establishing the number of people within the designated boundaries of an area. Population censuses are the primary sources of data on population composition, structure, and characteristics (Weeks, 2020). Good and accurate population data are vital for evidence-based planning and good governance. When data from the population census are used correctly, they serve as a good tool for decision-making and development planning (Mba, 2014; Pardeshi, 2010). Similarly, data aggregated by age and sex form an important basis in demographic analysis of fertility, and mortality, and serve as a basis for economic and social planning. Without this information planning for economic and social programs such as educational programs; health programs; job creation; and many others cannot be successfully realised.

The importance of population census data cannot be overemphasised enough. Demographic analyses based on inaccurate data often lead to misleading (and sometimes disastrous) results. Evidence from the literature confirms that demographic data, particularly those collected in most Sub-Saharan African countries are subjected to errors that may either be small or large (Fayehun et al., 2020; Siegel & Swanson, 2004; Yusuf et al., 2014). Although precision on census data is a challenge inherent in most data collected globally, censuses result for most African countries are afflicted by gross errors and deficiencies in age data classified by sex for various reasons. This has been well documented in numerous studies such as that of Myers (1940); Onsembe (2003), (Mba, 2004); Srinivasan (2011); Singh et al. (2022); and Fayehun et al. (2020). According to Yusuf et al. (2014), challenges in conducting a complete count in a particular population often lead to inaccuracies in the data collected. Inaccuracies or errors in census data mainly occur due to coverage errors, and response and non-response errors (Srinivasan, 2011). Equally, content errors in addition to enumeration errors ruin the quality and validity of the data intended to be used for planning purposes. The key concern of every data user is the type and enormity of errors acquired within the data set, especially among the variables of interest (Mba, 2014).

Globally, developed countries are at an advanced level of demographic data collection compared to their developing counterparts, and this is evident in the quality of the data they produce. The available literature on data quality in developing countries reveals a gloomy picture. In countries such as India census data quality on age and sex is still said to have a rough or very rough quality on the Whipple's index (WI) scale (Yadav et al., 2020). The phenomenon was likewise observed in the demographic health surveys (DHS) data collected in five South Asian countries Afghanistan, India, Nepal, Bangladesh, and Pakistan using the WI (Singh et al., 2022). Lyons-Amos and Stones (2017) evaluated the data quality of the DHS collected in 34 Sub-Saharan Africa from 1987 to 2015. They found that data quality reporting to have improved in Ghana (4.28%) and Nigeria (4.28%), and a notable increase in age misreporting was observed in Sierra Leone (4.46%), Chad (7.38%) and Ethiopia (7.58%) (Lyons-Amos & Stones, 2017). For Namibia, their findings showed the presence of age preference on terminal digits ending with 0s or 5s in all four DHS (Lyons-Amos & Stones, 2017). In Southern

Africa, countries like Lesotho (Mba, 2003) and Botswana (Bainame & Letamo, 2014) are reported to have good quality census data in terms of age reporting.

Thus far, Namibia has carried out three successful censuses in the years 1991, 2001 and 2011. However, it is not clear whether there has been any improvement in the quality of data collected in the past censuses. To the best of our knowledge, there are no studies done to examine the quality of age-sex data collected in the last three censuses. Thus, this paper aims to examine the quality and accuracy of reporting age and sex data in the last three censuses with the view of establishing patterns and trends. This is vital for policy and research intervention, and most importantly, for consideration as the country prepares for yet another decennial census undertaking.

MATERIALS AND METHODS

Data source

Following a cross-sectional design, the study used 1991, 2001 and 2011 Namibia Population and Housing Censuses (NPHC) data obtained from the Namibia Statistics Agency (NSA). NPHCs are conducted in every ten years interval following the country's independence in 1990. All three censuses were conducted, implemented, and published by the National Planning Commission (NPC). However, the 2011 census results were published by NSA after its establishment by the Statistics Act. 9 of 2011 and established in 2012. A *de facto* population count was employed in all three Censuses. The reference nights were 21st October 1991, 28th August 2001, and 28th August 2011 respectively. The 1991 NPHC excluded the town of Walvis Bay since it was still under the South African administration at the time. In 1991 a total of 1,409,920 individuals were enumerated followed by 1,830,330 and 2,113,077 in 2001 and 2011 respectively.

Techniques

The principal technique for identifying errors in the age data is to examine single-year distributions. The study used the rate of age not stated (per 1,000 population) to assess the quality of the age-sex data, while Whipple's and Myers's Blended Indexes were applied to assess age heaping in the data. Age heaping is defined as the propensity of innumerate respondents to report their age round to a close convenient even number in the multiple of 5 or 10 (Yadav et al., 2020). Data in five years age groups were analysed using Age Sex Ratio and the United Nations Age Sex Accuracy Index to establish possible undercounting of the population and displacement of individuals between age groups.

Whipple's Index (WI)

This technique assesses age preference/heaping on terminal digits (age) ending in multiples of 0s and 5s over a 5-year age range or on terminal digits ending in 0s over a 10-year age range (Myers, 1940). The outcome of this index varies between a minimum of 100, indicating no concentration at digits 0 and 5, and a maximum of 500, if only 0 and 5 were preferred (Pardeshi, 2010). The United Nations recommended the scale for measuring age heaping using WI to vary from <105 as highly accurate, 105 to 110 as fairly accurate, 110 to 125 as

approximately accurate, 125 to 175 as roughly accurate, and >175 as very roughly inaccurate data (Srinivasan, 2011).

Myers's Blended Index (MBI)

The MBI measures age preference or avoidance on all ten terminal digits 0 to 9 years instead of focusing on terminal digits ending in 0s and 5s in the WI. This is done through a 'blended' method to avoid possible bias due to the mortality effect (Srinivasan, 2011). The blended totals for the ten digits are expected to be nearly ten per cent of the total. The extent of preference and avoidance of a particular digit is reported as the deviations from ten per cent of the proportion of the population recorded on a given terminal digit (Mba, 2004). The deviations of each sum from ten per cent of the total ignoring the sign are summed up to the last terminal digit and divided by two to obtain a summary index of age preference. MBI yields both the reference index and summary index of preference for all ending digits. The theoretical range of MBI ranges from 0 to 90. An index outcome of 0 represents no evidence of age heaping, while an index of 90 signifies heaping on all reported ages at a single digit (Srinivasan, 2011).

Age-sex ratio (ASR) and United Nations Joint Score (UNJS)

The ASR measures are slightly different from the WI and MBI measures. The distinction is drawn from the fact that the ASR and the UNJS focus on data compiled in 5-year age groups, whereas the WI and MBI both concentrates on data provide in single years.

Sex Ratio is computed by taking the total male populations of the given 5-year age groups over that of females in the same age group, given by:

$$Sex\ Ratio = \frac{{}_5P_x^m}{{}_5P_x^f} \times 100$$

Similarly, the ASR per age group is derived as "the ratio of twice the population in that age group to the sum of the population in each of the adjacent age groups" (Moultrie et al., 2013, p. 9). ASR is computed for both sexes, and its general formula for age groups is given:

$$Age\ Sex\ Ratio = \frac{{}_5P_x}{\frac{1}{2}\left({}_5P_{x-5} + {}_5P_{x+5}\right)} \times 100$$

Moultrie et al. (2013) suggested that the ASR value should be closer to 100, with a deviation from 100 (in the lack of reasonable exogenous factors such as migration and distressing historical events that affected age groups) indicating undercounting or bizarre errors within the dataset.

Sex ratios provide an opportunity to estimate the extent of age misreporting in 5-year age groups in a certain population. We expect the index to have a slight excess number of males in the younger ages and an excess number of females in the older ages as a result of sex-age mortality differentials. As a

norm female deaths are typically lower compared to those of males in many populations, as a result, the values of the sex ratio will reflect a declining ratio in successive age groups (Mba, 2004). A deviation from this pattern is an indication of poor enumeration or age misreporting.

Equally, the UNJS is a pooled or complex statistic that shows the overall age displacement in the given five-year age groups and it indicates disparities in age misreporting among males and females (Srinivasan, 2011, p. 66). Depending on the overall UNJS obtained, it is recommended to interpret the overall quality of the data in terms of age-sex as follows, a value below 20 indicates an accurate data quality, a value between 20 and 40 indicates an inaccurate data quality, while a value above 40 indicates a highly inaccurate data quality (Srinivasan, 2011).

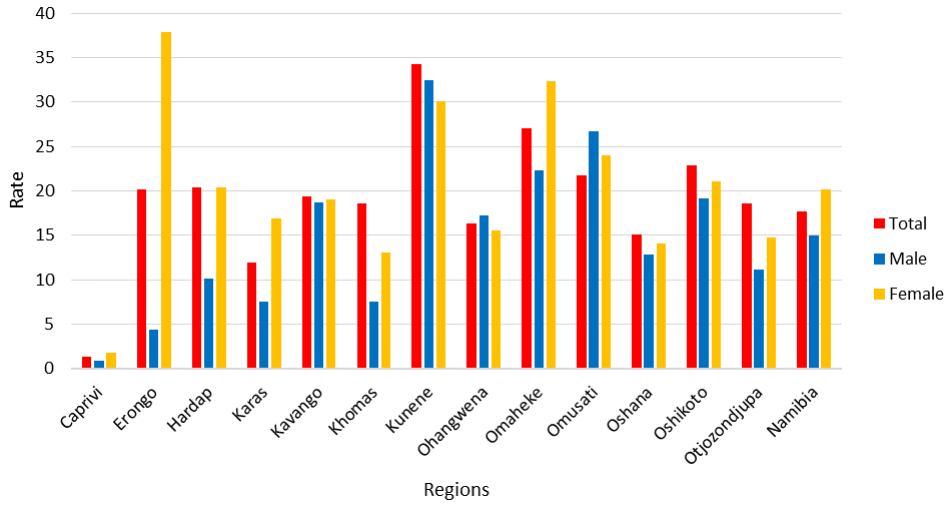
A statistical package R version 3.6.3 ("Holding the Windsock"), as well as the SINAGE and AGESEX spread sheets developed by the United States Census Bureau, were used to analyse the data presented in this paper.

RESULTS

To begin with our investigation on assessing the quality of the NPHCs data, we considered first analysing the variable "*age not stated"*. For some apparent reasons, the 1991 and 2011 NPHCs metadata did not contain *age not stated* category. Thus, making it impossible to estimate the trends in the *age not stated* over the study period. In the population census data set of 2001, a population of 1,830,330 was enumerated and the study found that approximately 18 per 1,000 persons in the total population could not state their age to enumerators (Figure 1). The figure translated loosely to about 32,325 persons of which 20 per 1,000 persons were females. Observing the regional profiles (Namibia had 13 administrative regions when the three censuses were conducted, and more than half of the population lived in rural areas during the same time) in Figure 1, *age not stated* was more prevalent in rural-based regions compared to their urban counterpart, with Kunene region ranking the highest (34 per 1,000 persons), followed by Omaheke (27 per 1,000 persons), Oshikoto (23 per 1,000 persons) and Omusati (22 per 1,000 persons) region. The region with the least *age not stated* was Caprivi (1.6 per 1,000 persons).

Figure 1

Age not stated per 1,000 inhabitants by sex and regions for 2001 NPHCs



Patterns of digit preference in single-age distributions

Figure 2 shows the pattern of single-age terminal digits preference and avoidance over the three Namibia decennial population censuses. The overall results revealed that there was a minimal digit preference or avoidance in Namibia over the three censuses period. Terminal digits 0, 1, 2, 5, 8 and 9 were most preferred by respondents of both sexes in all three censuses. While the remaining terminal digits 3, 4, 6 and 7 exhibited a negative sign confirming a general pattern of digit avoidance over the census counts, with terminal digits ending with 4 and 7 have the highest value of digit avoidance.

When controlled for sex, the results showed little to no variation in digit preference between males and females (Figure 2). Where a difference existed, the data showed a point to two-point digit preference and/or avoidance by females compared to males across the three census years. For 1991 NPHC, females preferred ages with terminal digits ending with 0, 1, 2, 5, 8 and 9, and avoided ages ending with terminal digits 3, 4, and 7, compared to their male counterparts. However, during the 2011 NPHC males preferred ages ending with terminal digits with 0, 1, 2, 5, 8 and 9 in contrast to females.

Additionally, findings in Table 1 further showed fluctuating results of the MBI of age preference from 41 in 1991 to 43 in 2011. Conversely, WI presents 2 points decline from 106 in 1991 indicating that the data was fairly accurate to 104 in 2001 and 2011. Thus, the data reveals that there was little variation in age misreporting by sex in the 1991, 2001 and 2011 NPHCs. Age misreporting was observed in both males and females, although it was more prevalent in females' reported ages than those of males.

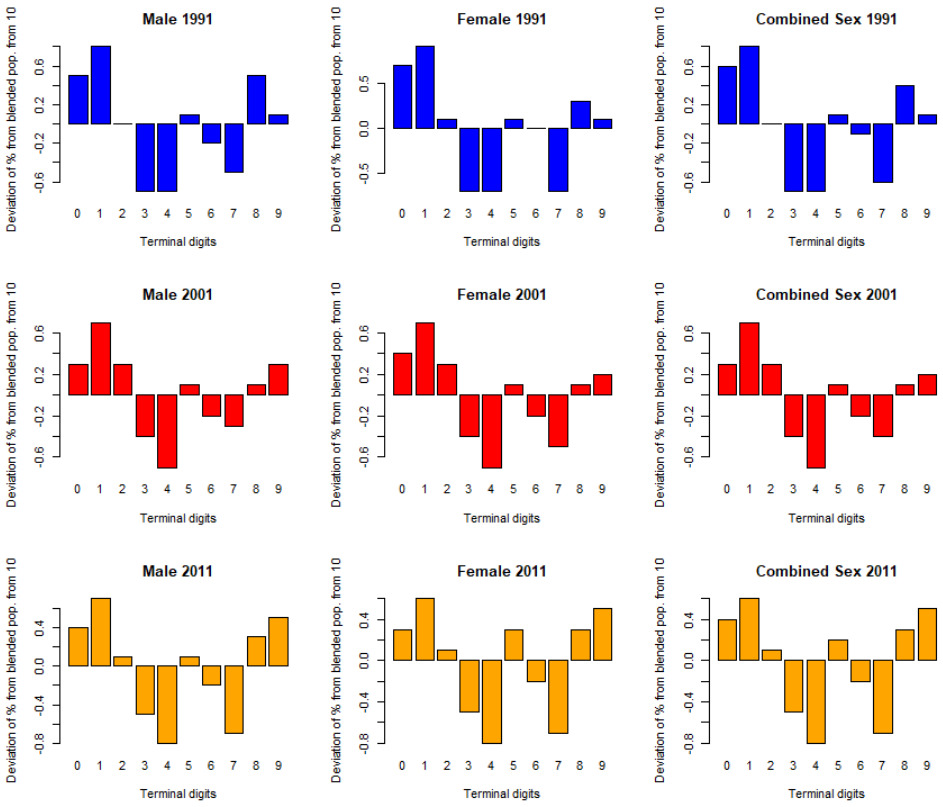
Table 1

Myer's Blended and Whipple's Indexes by sex for 1991, 2001 and 2011 NPHCs

Indexes	1991 Census			2001 Census			2011 Census		
	Male	Female	Combined sexes	Male	Female	Combined sexes	Male	Female	Combined sexes

MBI	39	42	41	33	37	35	42	44	43
WI	105	107	106	103	105	104	104	104	104

Figure 2
Age preference or avoidance by sex for 1991, 2001 and 2011 NPHCs



Regional patterns in Myer's Blended Index

Table 2 shows MBI ranking scores for males across all political regions. Overall MBI value for males displayed an improved score in 1991 (MBI=39) and 2001 (MBI=33) population censuses. However, the score worsened in the 2011 population census with an MBI score of 42. For regions, the Oshana region showed an improvement, moving from the 11th position in 1991 to the 6th and 5th position in 2001 and 2011 respectively. Similarly, Otjozondupa and Caprivi regions displayed progress in their rankings. Otjozondjupa moved from 5th position in 1991 to 2nd position in 2011, although it dropped momentarily to 7th position in 2001. Whereas Caprivi region was ranked 13th in 2001 with an MBI value of 71, and later ranked 8th in 2011 with an MBI value of 52.

On the other hand, the Khomas region performed poorly for both sexes as observed in Table 2 and Table 3. In 1991, the male population was ranked at 4th position and later dropped to 9th and 11th positions in 2001 and 2011 in that order. A similar trend can be noticed among the female population of the Khomas

region, with the region ranked at 4th position in 1991 and later ranked at 6th and 13th position in 2001 and 2011 respectively. Kavango region was the least-ranked region across all three census periods (ranked 12th in 1991 and 2001 and 13th position in 2011) for mal**e** population.

Table 2
MBI scores for male population by regions for 1991, 2001 and 2011 NPHCs

	1991			2001			2011	
Rank	**Region**	**MBI**	**Rank**	**Region**	**MBI**	**Rank**	**Region**	**MBI**
1	Hardap	30	1	Karas	19	1	Karas	28
2	Karas	30	2	Erongo	22	2	Otjozondupa	32
3	Erongo	31	3	Hardap	27	3	Hardap	33
4	Khomas	33	4	Ohangwena	30	4	Erongo	39
5	Otjozondupa	37	5	Omaheke	31	5	Oshana	44
6	Ohangwena	43	6	Oshana	32	6	Oshikoto	45
7	Omaheke	44	7	Otjozondupa	34	7	Omaheke	47
8	Omusati	55	8	Oshikoto	37	8	Caprivi	52
9	Oshikoto	55	9	Khomas	39	9	Ohangwena	53
10	Caprivi	61	10	Kunene	39	10	Omusati	53
11	Oshana	62	11	Omusati	43	11	Khomas	54
12	Kavango	81	12	Kavango	60	12	Kunene	66
13	Kunene	81	13	Caprivi	71	13	Kavango	71
	Namibia	**39**		**Namibia**	**33**		**Namibia**	**42**

The regional ranking of females' MBI scores is presented in Table 3. The study found that the overall MBI value for Namibia population slightly increased from 42 in 1991 to 37 in 2001, and it later decreased to 44 in 2011. Regarding regions, Otjozondupa region remained in the top three ranked regions for the entire study period. Whereas Omaheke region made a surprise improvement to be ranked 1st in the 2001 census from the 6th position in the 1991 census and it further dropped to the 5th position in the 2011 census. Conversely, the Oshikoto region performed well in terms of age accuracy reported by the female population in both WPI and MBI rankings. In the 1991 census, Oshikoto region was ranked 8th with an MBI value of 51 which improved to the 6th position in the 2011 census with an MBI value of 37. Yet again, the Kavango region remained at the bottom of the table, even though it has shown some improvements in MBI score from the highest value of 104 in 1991, to 81 in 2001 and 64 in 2011.

Table 3
MBI scores for female population by regions for 1991, 2001 and 2011 NPHCs

1991			2001			2011		
Rank	**Region**	**MBI**	**Rank**	**Region**	**MBI**	**Rank**	**Region**	**MBI**

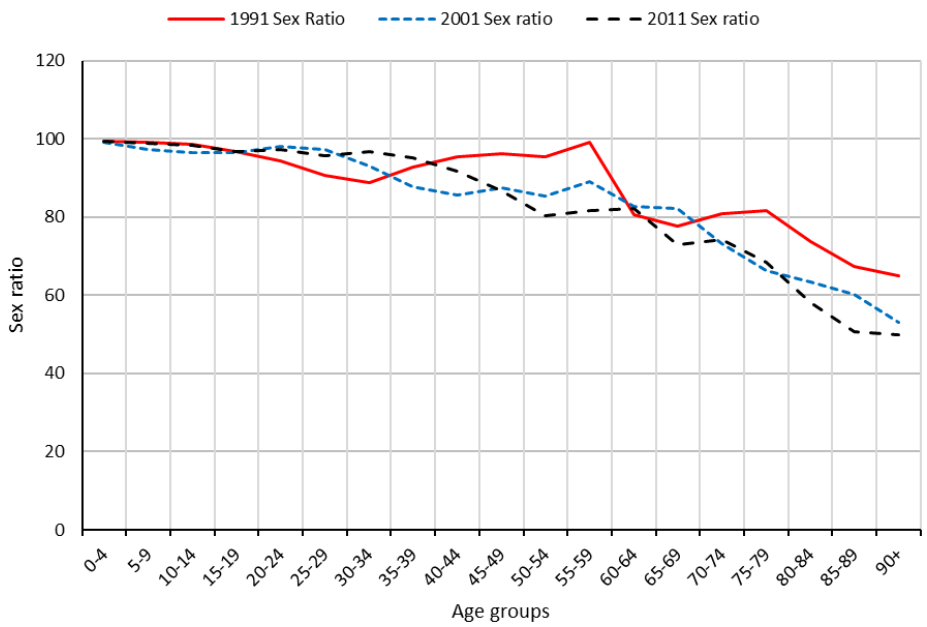
1	Hardap	15	1	Omaheke	18	1	Otjozondupa	32
2	Karas	20	2	Erongo	21	2	Hardap	35
3	Otjozondupa	23	3	Otjozondupa	23	3	Karas	36
4	Khomas	27	4	Oshana	29	4	Erongo	37
5	Erongo	29	5	Hardap	33	5	Omaheke	37
6	Omaheke	32	6	Khomas	33	6	Oshikoto	37
7	Oshana	51	7	Karas	36	7	Oshana	41
8	Oshikoto	51	8	Ohangwena	44	8	Ohangwena	50
9	Ohangwena	60	9	Kunene	45	9	Omusati	50
10	Omusati	61	10	Oshikoto	48	10	Caprivi	51
11	Kunene	62	11	Omusati	52	11	Kunene	53
12	Caprivi	70	12	Kavango	81	12	Kavango	64
13	Kavango	104	13	Caprivi	88	13	Khomas	68
	Namibia	**42**		**Namibia**	**37**		**Namibia**	**44**

Age Sex Accuracy Index

Age and sex data from NPHCs of 1991, 2001 and 2011 provided an opportunity for assessing the accuracy of age and sex data categorised in five-year age bands. Namibia being a developing country, we expected the ratio of males per 100 females to be higher in younger age groups and we anticipated it to decline with age due to high mortality among males (Lenart et al., 2019).

Figure 3

Sex ratio for 1991, 2001 and 2011 NPHCs



Overall, Figure 3 shows that Namibia had an overall sex ratio of 94.8, 94.2, and 93.7 in 1991, 2001, and 2011 correspondingly that were both lower than the African sex ratio model range of 100 to 107 (Mba, 2003; UN, 1983)

implying that there were fewer males per 100 females in Namibia across the study period. Over the two decades, the sex ratio at birth remained at 99 males per 100 females, as the predicted sex ratio decreased with successive age groups as mortality increased in the population. The sex ratio at age 35-39, 40-44, 55-59 and 70-74 in 1991 NPHCs shows that the preceding age groups had a lower sex ratio, this is an indication of errors present in the self-reported age-sex data. Similar findings were also observed in 2001 and 2011 NPHCs in the age groups 20 – 24, 35 – 39, 55 – 59 and 20-24, 30-34, 55-59, 60-64, and 70-74 respectively.

For the age ratios presented in Table 4, the study findings show a substantial fluctuation from the point of balance (100) indicating inaccuracy in all three NPHCs age distributions. The result further disclosed that in 1991 NPHC, there was under-enumeration in the following age groups: 5-9, 20-24, 30-34, 35-39, 55-59, 65-69, 75-79, 80-84 and 85-89. Similarly, the study also found that there was over-enumeration of the population in the age groups: 15-19, 50-54, 60-64 and 70-74. Across the three censuses, the results found a minimal to no significant difference between females' and males' age ratios; however, females' data tend to perform slightly poorer than that of males.

Table 4
Age sex ratios for 1991, 2001 and 2011 NPHCs

Age groups	1991		2001		2011	
	Male	**Female**	**Male**	**Female**	**Male**	**Female**
0-4	NA	NA	NA	NA	NA	NA
5-9	97.4	97.3	104.5	105	90.2	90.3
10-14	99.1	98.5	102.3	102.7	103.2	102.5
15-19	107.6	107.7	99.6	100.3	102.9	104.1
20-24	94.9	94.7	99.5	98.2	104.2	103.1
25-29	100.7	102.4	103.6	102.3	96.5	97.9
30-34	96	98.8	95.7	96.1	99.4	98
35-39	95.2	93.6	98.8	101.4	101.3	100.9
40-44	100	98.8	94.9	97.1	95.5	95.4
45-49	94.6	93.7	95.9	93.8	100.7	101.2
50-54	105.6	107.7	101	104.3	93	98.1
55-59	88.2	79	88.4	83.5	93.9	93.3
60-64	106.6	117.5	111.2	115.8	106.4	101
65-69	90	93.6	91.6	87.9	91.5	99.3
70-74	129.2	126	104.8	108.5	101.9	98
75-79	78.2	75.7	87	90.8	90.3	89.4
80-84	91.2	97.1	127.4	130.1	104	110.4
85-89	59.5	62.3	56.6	56.7	66.4	71.3
90+	NA	NA	NA	NA	NA	NA

NA = Not Applicable

The UNJS for Namibia given in Table 5 shows a declining (improving) trend. The scores were 31.3 in 1991, 27.7 in 2001 and 22.0 in 2011, suggesting that the age and sex data were inaccurately reported in all three NPHCs. The result shows an improvement in the quality of data collected on age and sex in Namibia with a decline in the joint score of 9.3 points between 1991 and 2011. Although we observed modest progress in reporting age and sex data, the data quality index for 2011 NPHC yielded a UNJS of 22 points, implying that the data was inaccurate.

Table 5
Summary Evaluative Indexes for 1991, 2001 and 2011 NPHCs

Census Year	Age ratio score		Sex Ratio Score	UNJS
	Male	Female		
1991	9.7	10.5	3.7	31.3
2001	8.6	9.8	3.1	27.7
2011	6.4	5.4	3.4	22.0

DISCUSSION

The study examined the quality of data collected on age and sex in the past three NPHCs by assessing variable age not reported, age reported in single years and five years age bands by sex. WI and MBI were used to assess whether there is digit preference or avoidance in the three NPHCs data. It is premised to appraise the quality of data collected on age and sex because often most demographic measures derived from the population's age and sex composition are proven to be exposed to numerous errors that eventually influence the resulting demographic estimates. Most common errors occur as a result of individuals not being able to recall their date of birth/actual age, leading to either age not being reported or to incorrect reporting of ages by respondents or by enumerators (Mba, 2014; Srinivasan, 2011).

The paper shed light on the continuum misreporting of age data in NPHCs by describing age not reported and age heaping using different demographic methods. Age not stated was only found to be present in the 2001 NPHC data. For some reason, the data sets for NPHCs for 1991 and 2011 did not contain age not stated category. The distribution of age not stated was predominant more in the rural area-based regions, signifying the presence of higher percentage of less educated population in these areas. However, there were also a considerable number of people in urban areas who were not unable to state or recall their ages, especially in the Erongo and Khomas regions. This can be explained by the fact that these two regions have the highest number of migrants per capita in the entire country, most migrants flocking to these regions have lower to no level of educational attainment (Greiner, 2011; Indongo et al., 2013). Among those with age not stated, 89 per cent were women in the Erongo region and 71 per cent of women in the Khomas region. The current study results were similar to those of Fayehun et al. (2020) in Nigeria who pointed out that the improvement in age reporting was linked to increased educational attainment. Singh et al. (2022) also

found age heaping to be higher among respondents with no education in South Asia countries. However, these findings and conclusions differed from those of A'Hearn et al. (2022) who argued that age heaping does not occur as a result of people's intellectual abilities but rather as a result of social, economic and functional transformation.

Generally, all demographic measures used in our study suggested that the distribution of age and sex data collected in the three NPHCs was fraught with distortions. However, we wish to point out that the intensity of the revealed biases was not momentous to nullify demographic estimates derived from these censuses. Despite the decrease in 2011 NPHC data quality, overall the data was still of acceptable quality when compared with other African and developing countries, such as The Gambia (Mba, 2004), Nigeria and Malawi (Mba, 2014), and India Basannar et al. (2022). Moreover, the fact that the findings from both censuses are not far away from each other, gives accounts of consistency and confidence in the reported age-sex distributions. Despite this, caution should be exercised when producing estimates from the data as these inherent biases can result in poor planning and incorrect estimation of demographic rates (Pardeshi, 2010).

Other than the inherent bias in the data, the reported distortions and the good quality of NPHCs data quality could be associated with the education and literacy level of the population. On average, the NPHCs recorded the literacy level at 78 per cent, 65 per cent in 1991 (UNESCO, 2017), 81 per cent in 2001 and 89 per cent in 2011 (NSA, 2013). Although less expected, a high proportion of age-sex inaccuracies were reported in urban area-based regions of Erongo and Khomas Region (a shared 97 per cent in 2011) where literacy level is likely to be higher compared to that of rural-based regions. In India Basannar et al. (2022) ascribed the quality of age and sex reporting to the high level of education among the study participants.

The MBI results showed that terminal digits ending with a 1 were most preferred compared to all other digits across all three censuses. In absence of proper records of civil registration for the black population before Namibia's Independence in 1990 meant that people from the older generations found it difficult to recall their exact date, month, and year of birth. Similarly, the data revealed that most respondents selected the first date of the month as their date of birth, leading to the digit terminal 1 being a preferred digit by respondents. The preference of 1 in 1991, 2001 and 2011 is comparable to the result of the Malawi 1998 and 2008 census that found the digit preference 8 similar to the terminal digit of the census year (Fajardo et al., 2014; Palamuleni, 1995). While in Nigeria respondents preferred to report their ages in terminal digits ending with 0 (Fayehun et al., 2020) whereas in India respondents favoured terminal digits ending with 0 and 5 (Basannar et al., 2022).

Expectedly, the regional performance of MBI for both males and females in 1991 showed that the top five ranked regions had a high number of people living in urban areas, compared to the bottom five regions with a rural-based population. This implies that those who lived in urban areas were more likely to report their age correctly as opposed to those in rural areas. These findings were

similar to those of Agrawal and Khanduja (2015). Despite this, the index changed in the subsequent censuses, in 2001 and 2011 Khomas region which is highly urbanised took a heavy toll by dropping in ranking for both male and female indexes. Similarly, the Erongo region followed this trend. Migration might have played a role in this trend as these two regions offer more economic opportunities compared to other regions in the country. Hence, internal migration influenced the results; however, its impact is not addressed in this paper. Equally, it is worth noting that before 1990, Namibia was under the racial apartheid policy of South Africa Affairs Act No. 25 of 1969 that limited black people's movements to areas in the police zone (Geldenhuys, 1981). In this legislation, black Namibians (referred to therein as natives) were restricted to tribal reserves or homelands while urban locations (police zone) were reserved for European settlements. Only a few black Namibians were allowed to be in urban areas primarily for providing cheap labour (UN, 1988). This explains why Namibia's urban population was only 28 per cent in 1991 (NPC, 1994).

The lifting of the apartheid laws allowed the free movement of people to urban-based regions namely Khomas and Erongo regions in search of socio-economic advancements. An examination of the urban-rural population structures of 2001 and 2011 showed the expansion of the urban population in the age group of 20 to 24 years and a reduction in the rural population in the same age group. In 2011 the urban population grew to 43 per cent, resulting in the staggering urban growth of 4.96 per cent as people flood the informal settlement on the periphery of urban centres in particular Windhoek in the Khomas region and Walvis Bay in Erongo region (NPC, 2015). Although, these regions are predominantly urban by locality, the proportion of the semi-urban population requires them to be given special attention when it comes to census education like rural areas to improve data quality.

CONCLUSION

The study aimed to examine the quality and accuracy of reporting age and sex data during the last three censuses with the sole purpose of establishing patterns and trends. Within the limits of the appraising tools employed in this study, we deduced that age-sex data reported by females were slightly more inaccurate in both censuses compared to that of males. The finding of sex differences was similar to the result of Mba (2004) obtained in The Gambia. There was not much difference in adult literacy rates among males and females in the country. In 2001, the literacy rates were 86 per cent for males and 83 per cent for females, while the literacy level was 89 and 88 for males and females respectively in the same year (UNESCO, 2017). Surely the literacy levels alone cannot explain the difference in the data accuracy amongst sexes, therefore there is a need to explore more factors that might have led to age misreporting in both censuses. Thus, we speculated that women might have had their information reported by somebody else other than themselves (the head of household) compared to men. Therefore, the study recommends an investigation be done to examine leading factors that instigate females to report their ages inaccurately compared to their male counterparts.

ACKNOWLEDGEMENTS

The authors would like to express their gratitude to the Namibia Statistics Agency for availing the data sets on their website.

REFERENCES

A'Hearn, B., Delfino, A., & Nuvolari, A. (2022). Rethinking age heaping: a cautionary tale from nineteenth-century Italy. *The Economic History Review*, *75*(1), 111-137. https://doi.org/https://doi.org/10.1111/ehr.13087

Agrawal, G., & Khanduja, P. (2015). Influence of literacy on India's tendency for age misreporting: evidence from Census 2011. *Journal of Population and Social Studies [JPSS]*, *23*(1), 47-56.

Bainame, K., & Letamo, G. (2014). Evaluation of Data Quality of the Botswana 2011 Population and Housing Census. *Botswana Notes and Records*, *46*, 26-36. https://doi.org/https://www.jstor.org/stable/90024359

Basannar, D. R., Singh, S., Yadav, J., & Yadav, A. K. (2022). Quantifying age heaping and age misreporting in a multicentric survey. *Indian J Community Med*, *47*(1), 104-106. https://doi.org/10.4103/ijcm.ijcm_1179_21

Fajardo, J., Attanasio, L., & Ha , J. (2014). *An Assessment of the Age Reporting in the IPUMS-I Microdata. 2014* Annual Meeting of the Population Association of America,

Fayehun, O., Ajayi, A. I., Onuegbu, C., & Egerson, D. (2020). Age heaping among adults in Nigeria: evidence from the Nigeria Demographic and Health Surveys 2003–2013. *Journal of Biosocial Science*, *52*(1), 132-139. https://doi.org/10.1017/S0021932019000348

Geldenhuys, D. (1981). South Africa's black homelands: Some alternative political scenarios. *Journal of Contemporary African Studies*, *1*(1), 105-139. https://doi.org/10.1080/02589008108729377

Greiner, C. (2011). Translocal networks and socio-economic stratification in Namibia. *Africa*, *81*(4), 606-627. https://doi.org/10.1017/S0001972011000477

Indongo, N., Angombe, S., & Nickanor, N. M. (2013). *Urbanisation in Namibia: Views from semi-formal and informal settlements*.

Lenart, P., Kuruczova, D., Joshi, P. K., & Bienertová-Vašků, J. (2019). Male mortality rates mirror mortality rates of older females. *Scientific Reports*, *9*(1), 10589. https://doi.org/10.1038/s41598-019-47111-w

Lyons-Amos, M., & Stones, T. (2017). Trends in Demographic and Health survey data quality: an analysis of age heaping over time in 34 countries in Sub Saharan Africa between 1987 and 2015. *BMC Research Notes*, *10*(1), 760. https://doi.org/10.1186/s13104-017-3091-x

Mba, C. J. (2003). Assessing the reliability of the 1986 and 1996 Lesotho census data. *Journal of Social Development in Africa*, *18*(1). https://doi.org/10.4314/jsda.v18i1.23821

Mba, C. J. (2004). Challenges of population census enumeration in Africa: an illustration with the age-sex data of the Gambia. *Institute of African Studies Research Review*, *20*(1), 9-19.

Mba, C. J. (2014). Examining the accuracy of age-sex data: An evaluation of recent Sub-Saharan African Population Censuses. In C. O. Odimegwu & J. Kekovole (Eds.), *Continuity and Change in Sub-Saharan African Demography* (1 ed., pp. 356). Routledge. https://doi.org/https://doi.org/10.4324/9781315879444

Moultrie, T. A., Dorrington, R. E., Hill, A. G., Hill, K., Timæus, I. M., & Zaba, B. (2013). *Tools for Demographic Estimation*. International Union for the Scientific Study of Population. demographicestimation.iussp.org

Myers, R. J. (1940). Errors and bias in the reporting of ages in census data. *Transactions of the Actuarial Society of America*, *41*(2), 395-415.

NPC. (1994). *1991 Population and Housing Census Report*. https://d3rp5jatom3eyn.cloudfront.net/cms/assets/documents/1991_Population_and_Housing_Census_Basic_Analysis_with_Highlights.pdf

NPC. (2015). *Policy Brief: Population Dynamics*. https://www.npc.gov.na/download/pbriefs/NPC%20Population%20Dynamics.pdf

NSA. (2013). *Namibia 2011 Population & Housing Census Main Report*. https://cms.my.na/assets/documents/p19dmn58guram30ttun89rdrp1.pdf

Onsembe, J. (2003). *Improving Data Quality in the 2000 Round of Population and Housing Censuses*. UNFPA Country Technical Services Team.

Palamuleni, M. E. (1995). Age misreporting in Malawian censuses and sample surveys: An application of the United Nations joint age and sex score. *Southern African Journal of Demography*, *5*(1), 11-17. https://doi.org/https://journals.co.za/doi/pdf/10.10520/AJA16824482_68

Pardeshi, G. S. (2010). Age heaping and accuracy of age data collected during a community survey in the yavatmal district, maharashtra. *Indian J Community Med*, *35*(3), 391-395. https://doi.org/10.4103/0970-0218.69256

Siegel, J., & Swanson, D. (2004). *The Methods and Materials of Demography* (2 ed.). Elsevier Academic Press.

Singh, M., Kashyap, G. C., & Bango, M. (2022). Age heaping among individuals in selected South Asian countries: evidence from Demographic and Health Surveys. *Journal of Biosocial Science*, *54*(4), 725-734. https://doi.org/10.1017/s0021932021000249

Srinivasan, K. (2011). *Training Manual on Demographic Techniques*. https://india.unfpa.org/sites/default/files/pub-pdf/TrainingManualonDemographicsTechniques%28forwebsite%29.pdf

UN. (1983). *Manual X. Indirect techniques for Demographic estimations* (Vol. 81). United Nations. https://www.un.org/en/development/desa/population/publications/pdf/mortality/Manual_X.pdf

UN. (1988). *Decolonization*. https://www.un.org/dppa/decolonization/sites/www.un.org.dppa.decolonization/files/decon_num_35.pdf

UNESCO. (2017). *National Literacy Programme, Namibia*. https://uil.unesco.org/case-study/effective-practices-database-litbase-0/national-literacy-programme-namibia

Weeks, J. R. (2020). *Population: An introduction to Concepts and Issues* (13 ed.). Cengage Learning.

Yadav, A., Vishwakarma, M., & Chauhan, S. (2020). The quality of age data: Comparison between two recent Indian censuses 2001–2011. *Clinical Epidemiology and Global Health, 8*(2), 371-376. https://doi.org/10.1016/j.cegh.2019.09.005

Yusuf, F., Martins, J. M., & Swanson, D. A. (2014). *Methods of Demographic Analysis*. Springer. https://doi.org/10.1007/978-94-007-6784-3