

# Understanding Component Structures to Support Quality Primary Health Care Using Factor Analysis in O. R. Tambo District Municipality, South Africa

Nombulelo Chitha<sup>1</sup>, Wezile Chitha<sup>1</sup>, John Sungwacha Nasila<sup>1,2</sup>, Zukiswa Jafta<sup>1</sup>, Buyiswa Swartbooi<sup>1</sup>, Siyabonga Sibulawa<sup>1</sup>, Onke Mnyaka<sup>1</sup>, Natasha Williams<sup>1</sup>, Benjamin Ben-I-Sasa Longo-Mbenza<sup>2,3,4,\*</sup>

<sup>1</sup>Health Systems Enablement & Innovation Unit, University of the Witwatersrand, Johannesburg, South of Africa

<sup>2</sup>Department of Mathematical Sciences & Computing, Walter Sisulu University, Mthatha, South of Africa

<sup>3</sup>Cardiology, University of Kinshasa, Kinshasa, Democratic Republic of Congo

<sup>4</sup>Public Health, Lomo University of Research, Kinshasa, Democratic Republic of Congo

## Email address:

nchitha@wits.co.za (N. Chitha), wezile.chitha@wits.co.za (W. Chitha), nasila.john@gmail.com (J. S. Nasila), zjafta@gmail.com (Z. Jafta), bswartbooi@witshealth.co.za (B. Swartbooi), sibulawasiyabonga@yahoo.com (S. Sibulawa), omnyaka@hsei.co.za (O. Mnyaka), nwilliams@witshealth.co.za (N. Williams), longombenza@gmail.com (B. Ben-I-Sasa Longo-Mbenza)

\*Corresponding author

## To cite this article:

Nombulelo Chitha, Wezile Chitha, John Sungwacha Nasila, Zukiswa Jafta, Buyiswa Swartbooi, Siyabonga Sibulawa, Onke Mnyaka, Natasha Williams, Benjamin Ben-I-Sasa Longo-Mbenza. Understanding Component Structures to Support Quality Primary Health Care Using Factor Analysis in O. R. Tambo District Municipality, South Africa. *World Journal of Public Health*. Vol. 6, No. 3, 2021, pp. 67-80. doi: 10.11648/j.wjph.20210603.11

Received: May 23, 2021; Accepted: June 4, 2021; Published: August 4, 2021

**Abstract:** This publication intends to provide a step-by-step description of the application of factor analysis performed at the two levels and interpretation of the results based on information seeking behaviour of medical professionals of five selected hospitals in O. R. Tambo District municipality in the Eastern Cape Province of South Africa. The data for analysis was collected on different variables using a closed-ended likert scale questionnaire. This study was a cross-sectional, comparative, and correlational survey conducted between January and April 2017, in the Mthatha Hospital Complex, O. R. Tambo District Municipality. The clustering of indicators for extraction of factors was well-defined owing to high loadings across all questions. The analysis was executed on a split data approach. The data were split by gender. The analysis was performed on the separate derived data sets. Descriptive analyses, correlations and component factor analyses were performed. The data consisted of 96.3% South African participants and 3.7% Non-South African. In addition, the sample was composed of 17.5% Males and 82.5% females; 13.8% medical doctors and 86.2% professional nurses. The percentage age distribution was: ≤30.00 (21.9%), 31.00 - 37.00 (20.5%), 38.00 - 44.00 (18.5%), 45.00 - 54.00 (19.2%) and 55.00+ (19.9%). The percentage hospital participation distribution was: Holly Cross Hospital (13.9%), Dr Malizo Mpehle Hospital (28.5%), St Barnabas Hospital (21.8%), Zithulele Hospital (18.8%) and St Elizabeth Hospital (17.0%). Use of materials sources available within the hospital as sources of information; Improvement of patient care through collaborative consultations. Use of information acquired through workshops, seminars and journals to improve the participants' knowledge; Use of internet and hospital facilities as sources of information; Use of reference materials and the medical need dictates the source of information required; Causes of limited availability sources of information; Use of printed material and colleagues to access information. Under the females' data, the following factors were extracted: Using both external as well as internal sources to solicit information; Lack of: a physical library, limited online access, slow internet and poor online searching skills contribute to non-availability of important medical information; Lack of awareness of sources of information, time taken to access information, non-existent of sources of information; The factor analysis has shown that whereas there were more females than males according to the split data, more factors were established for males than were for females.

**Keywords:** Health, Information, Systems

## 1. Introduction

In the simplest terms, a health information system (HIS) is a system that captures, stores, transmits, or otherwise manages health data or activities [1–3]. These systems are used to collect, process, use, and report health information. In turn, information from a health information system can be used to drive policy- and decision-making, research, and ultimately health outcomes. Here's what you need to know about the key components of an HIS, the various types, and benefits of HIS. There are many different types of healthcare information systems, including:

- 1) operational and tactical systems for easy classification of information;
- 2) clinical and administrative systems for managing patient details on an administrative level;
- 3) subject and task based systems such as Electronic Medical Records (EMRs) or Electronic Health Records (EHRs);
- 4) financial systems for tracking revenue and managing billing submissions.

These systems are designed to assist healthcare providers with managing daily tasks and patient information. Often, continues “these types of systems are broken up into different software solutions, but what if you could have all of these systems packaged into one convenient software solution [2].

The researchers of this medical publication are of the view that energising primary care approach in South Africa is of critical importance to realizing high quality, easily accessible, and effective health care for all South Africans. The effective use of health information is of great importance as it leads to improved healthcare in terms of collaborative work by both nurses and doctors. Furthermore, its amicability is informed by the way and manner the information is received and utilised. One of the obvious end use is the improvement objective by primary care practitioners to facilitate quality improvement (QI) that can help improve their ability to deliver high quality care and improve patient health outcomes. This has been supported by two specialist doctors by a research conducted in research in Saudi Arabia. According to consumers periodically seek information about their health and medicines to increase their knowledge and understanding, and to support self-management. Gender has an impact on people's attitudes and behaviour; gaining a better understanding of how males and females differ in the way in which they find and use health and medicine information may help to enhance patient-health care provider communication and the information that they receive. Based on a review of the literature, women overall are more active. This has been of practical observation in view of the data collected for this study, where the greater majority was constituted by females information seekers than men. A number of differences do exist, notably in: sources of

information consulted, information desired and degree of engagement in information seeking. Furthermore states that Differences exist in how males and females source information about their health care and medicines [4, 5]. Women are more active information seekers than men, which may be reflected in how they utilize information sources such as the Internet. Gender differences are also apparent in desired information with respect to managing medical conditions. Health care professionals must explore and be privy to these gender differences to ensure that consumer health and medicine information needs are acknowledged and addressed as part of a patient-centred approach to health care [5].

## 2. Methods

This was a cross-sectional research planned and executed by medical professionals in the Mthatha Hospital Complex. Five hospital were targeted for data collection. Professionals involved included specialist doctors, medical officers, senior nurses and registrars.

The data collected was divided into six categories namely: Bio-data, Task (Area of service, time), Source of information, what informs the choice of information source, hindrance to obtaining information and purpose for which information was required.

The team used a questionnaire to get the required information. The participants were drawn either doctors or professional nurses. The researchers who designed the questionnaire were experienced doctors, professional nurses who were involved in teaching at medical schools and doctors who were participating in research work where questionnaires were used and were also practicing doctors. issues indicated in the questionnaire were well considered and the end result well discussed. Members of the research team had read publications and likened their views with those in publications read.

The survey was specifically and extensively designed using a statistical stratified random sampling with proportional allocation model at each hospital to recruit a study sample with similar and representative characteristics of OR Tambo District medical professionals comprising of nurses, doctors, registrars and senior nurses, whose results were comparable with global data on medical information data collection. Each hospital contributed with a number of participants calculated by population representation. Out of a total of 166 participants, Holy Cross Hospital contributed 23 (13.9%) participants, Dr Malizo Mpehle Hospital 47 (28.5%), St Barnabas Hospital 36 (21.8%), Zithulele Hospital 31 (18.8%) and St. Elizabeth Hospital 28 (17.0%). The margin of sampling error of estimation was put to 2%.

In addition, the research team consulted many other professionals including a Statistician, other researchers and modellers. These other professionals came up with

suggestions which were meant to improve the data collection and possible analysis and interpretations. Healthcare IT people were conducted for advice. To a great extent, the idea of QI was seriously considered as the most welcome beneficiary of scrutiny of the study.

The final output of the data for this publication was divided into two parts: one comprising of the males' data and the other consisting of the females' data specifically for this journal article. The analysis was meant to be a comparative one where, the objective was to compare the outcome based on gender. This was well-meditated in view of the output which has shown have drawn interesting lines of significant nature in the theory of medical services provision in view of the importance of information reception and utilisation.

The analysis adopted a Factor Analysis for obvious reasons as this was the only possible procedure of creating the required components of the collected data in view of the constructed objectives. As predicted, different factors have been established in two ways: the number of factors determined according to gender and the type of such factors.

Finally, the analysis, particularly for this publication has taken into account the splitting of the data on the basis of gender. This has been the big comparative parameter of this particular assay.

### 3. Results

#### 3.1. Analysis Among the Males' Data

Table 1 below presents the means and standard deviations of males data used in the analysis for publication of this journal article. The reader is informed that the means should be interpreted bearing in mind that the larger the mean the more the weight the variable carries and vice-versa. Furthermore, the higher the mean, the more important the variable is, since its rating was higher than other variables. It is observed from the table that the following variables and their corresponding means were rated to be more important than others:

I use protocols/guidelines kept in the ward or my pocket

**Table 1.** Table showing the means of indicators in the descriptive analysis.

Descriptive Statistics	Mean	Std. Deviation	Analysis N
I talk to colleagues as a source of information	2.7333	.45774	15
I consult doctors as a source of information	2.6667	.48795	15
Consulting information sources assists me in treating my patients	4.4667	1.12546	15
Consulting information sources assists me in the prognosis of my patients	4.3333	1.11270	15

**Table 2.** Table showing the means of indicators in the descriptive analysis.

Descriptive Statistics	Mean	Std. Deviation	Analysis N
I talk to colleagues as a source of information	2.5909	.73414	22
I consult doctors as a source of information	2.6818	.47673	22
Access to information improves quality of patient care	4.8636	.35125	22
Consulting information sources assists me in the prognosis of my patients	4.5909	.59033	22

The following correlation table shows some of the correlations of some pairs of variables. The complete table

book as a source of information (2.8000), Trustworthiness informs my choice of information source (2.8000), I am aware of all information sources that are available in my hospital (4.4000), I use all information sources that are available in my hospital (4.1333) I always consult information sources for clinical decisions (4.0000), I use my mobile phone to access online information (4.4667), I use my book collection to access information (4.4667), I understand the importance of information in my work (4.6667), I consult information sources for professional development (4.7333), I consult information sources for teaching (4.2667), I only consult information sources for personal purposes (3.3333), Access to information improves quality of patient care (4.5333), Consulting information sources assists me to diagnose my patients (4.5333), Consulting information sources assists me in treating my patients (4.4667), Consulting information sources assists me in the prognosis of my patients (4.3333), etc.

On the other hand, the analysis shows the following means for some highly rated indicators under the females' data:

I consult print material (books, journals) from the library (3.0000), I use my mobile phone to access online information (4.3636), I use my book collection to access information (4.4091), I understand the importance of information in my work (4.5909), I consult information sources for professional development (4.6364), I consult information sources for teaching (4.3636), Access to information improves quality of patient care (4.8636), Consulting information sources assists me to diagnose my patients (4.6364), Consulting information sources assists me in treating my patients (4.6364), Consulting information sources assists me in the prognosis of my patients (4.5909), etc. From a close scrutiny of the means, we note that the means show a trend where both genders performed very highly among the last set of variables where concerns were about obtaining information from medical professionals as compared to other sets of variables. Under those indicators, the means ranged from 2.80 to 4.6 for males and 3.0 to 4.86 for females. The variabilities were equally varied.

I use protocols/guidelines kept in the ward or my pocket

**Table 1.** Table showing the means of indicators in the descriptive analysis.

Descriptive Statistics	Mean	Std. Deviation	Analysis N
I talk to colleagues as a source of information	2.7333	.45774	15
I consult doctors as a source of information	2.6667	.48795	15
Consulting information sources assists me in treating my patients	4.4667	1.12546	15
Consulting information sources assists me in the prognosis of my patients	4.3333	1.11270	15

**Table 2.** Table showing the means of indicators in the descriptive analysis.

Descriptive Statistics	Mean	Std. Deviation	Analysis N
I talk to colleagues as a source of information	2.5909	.73414	22
I consult doctors as a source of information	2.6818	.47673	22
Access to information improves quality of patient care	4.8636	.35125	22
Consulting information sources assists me in the prognosis of my patients	4.5909	.59033	22

could not be accommodated due to the large size of the correlations. Correlation analyses were performed on all pairs

of variables across the data. Some of the correlations were positive and significant, others showed no relationship between the relevant pairs, others were positive and strongly correlated, others pairs were negatively but insignificantly correlated, etc.

The following are some of the bivariate correlations:

*I read newspapers as a source of information and I talk to people outside of your work as a source of information* (-.059), *I use computer at work to access internet as a source of information and I talk to people outside of your work as a source of information* (.411), *I read newspapers as a source of information and I use my mobile phone to access internet as a source of information* (.288\*\*). Other correlations for other pairs of variables can equally be read from the following table.

This is the proportion of each variable's variance that can be explained by the factors (e.g., the underlying latent continua). It is also noted and symbolised as  $h^2$  and can be

defined as the sum of squared factor loadings for the variables.

The communalities for the  $i^{\text{th}}$  variable are computed by taking the sum of the squared loadings for that variable. This was expressed as follows:

$$\hat{h}_i = \sum_{j=1}^m \hat{f}_{ij}^2$$

This means that the 98.6% of the variance in the variable "I talk to colleagues as a source of information" has been explained by the extraction of the factors, and further that, the variable "I consult senior nurses as a source of information with an extracted communality of .983, means that the factor extraction explained 98.3% of the variance in the indicator. In addition, factor extraction explained 86.6% of the variable "I use my mobile phone to access internet as a source of information" with an extracted communality of .866.

**Table 3.** Table showing correlations of between pairs of indicators in the males' data.

Correlations					
Respondent's gender			I talk to people outside of your work as a source of information	I read newspapers as a source of information	I use my mobile phone to access internet as a source of information
Male	I talk to people outside of your work as a source of information	Pearson Correlation	1	-.059	.160
		Sig. (2-tailed)		.784	.455
		N	25	24	24
	I read newspapers as a source of information	Pearson Correlation	-.059	1	.246
		Sig. (2-tailed)	.784		.235
		N	24	26	25
	I use my mobile phone to access internet as a source of information	Pearson Correlation	.160	.246	1
		Sig. (2-tailed)	.455	.235	
		N	24	25	27
	I use computer at work to access internet as a source of information	Pearson Correlation	.411	-.115	.020
		Sig. (2-tailed)	.064	.621	.933
		N	21	21	21
Female	I use reference books kept in the hospital as a source of information	Pearson Correlation	-.318	.109	.021
		Sig. (2-tailed)	.131	.613	.920
		N	24	24	25
	I use protocols/guidelines kept in the ward or my pocket book as a source of information	Pearson Correlation	.267	.378	.171
		Sig. (2-tailed)	.198	.057	.402
		N	25	26	26
	I talk to people outside of your work as a source of information	Pearson Correlation	1	.472**	.226*
		Sig. (2-tailed)		.000	.030
		N	102	91	92
	I read newspapers as a source of information	Pearson Correlation	.472**	1	.288**
		Sig. (2-tailed)	.000		.005
		N	91	107	94
	I use my mobile phone to access internet as a source of information	Pearson Correlation	.226*	.288**	1
		Sig. (2-tailed)	.030	.005	
		N	92	94	109
	I use computer at work to access internet as a source of information	Pearson Correlation	.492**	.319**	.109
		Sig. (2-tailed)	.000	.005	.338
		N	77	77	80
	I use reference books kept in the hospital as a source of information	Pearson Correlation	.046	.307**	.231*
		Sig. (2-tailed)	.674	.003	.028
		N	85	91	91
	I use protocols/guidelines kept in the ward or my pocket book as a source of information	Pearson Correlation	.077	.016	.392**
		Sig. (2-tailed)	.458	.877	.000
		N	96	101	104

Table 3. Continued.

Correlations			I use computer at work to access internet as a source of information	I use reference books kept in the hospital as a source of information	I use protocols/guidelines kept in the ward or my pocket book as a source of information
Respondent's gender					
Male	I talk to people outside of your work as a source of information	Pearson Correlation	.411	-.318	.267
		Sig. (2-tailed)	.064	.131	.198
		N	21	24	25
	I read newspapers as a source of information	Pearson Correlation	-.115	.109	.378
		Sig. (2-tailed)	.621	.613	.057
		N	21	24	26
	I use my mobile phone to access internet as a source of information	Pearson Correlation	.020	.021	.171
		Sig. (2-tailed)	.933	.920	.402
		N	21	25	26
	I use computer at work to access internet as a source of information	Pearson Correlation	1	.046	.166
		Sig. (2-tailed)		.840	.461
		N	22	22	22
Female	I use reference books kept in the hospital as a source of information	Pearson Correlation	.046	1	.375
		Sig. (2-tailed)	.840		.065
		N	22	26	25
	I use protocols/guidelines kept in the ward or my pocket book as a source of information	Pearson Correlation	.166	.375	1
		Sig. (2-tailed)	.461	.065	
		N	22	25	27
	I talk to people outside of your work as a source of information	Pearson Correlation	.492**	.046	.077
		Sig. (2-tailed)	.000	.674	.458
		N	77	85	96
	I read newspapers as a source of information	Pearson Correlation	.319**	.307**	.016
		Sig. (2-tailed)	.005	.003	.877
		N	77	91	101
Female	I use my mobile phone to access internet as a source of information	Pearson Correlation	.109	.231*	.392**
		Sig. (2-tailed)	.338	.028	.000
		N	80	91	104
	I use computer at work to access internet as a source of information	Pearson Correlation	1	.398**	-.140
		Sig. (2-tailed)		.000	.209
		N	88	81	82
	I use reference books kept in the hospital as a source of information	Pearson Correlation	.398**	1	.208*
		Sig. (2-tailed)	.000		.036
		N	81	107	102
	I use protocols/guidelines kept in the ward or my pocket book as a source of information	Pearson Correlation	-.140	.208*	1
		Sig. (2-tailed)	.209	.036	
		N	82	102	121

\*\* . Correlation is significant at the 0.01 level (2-tailed).

\* . Correlation is significant at the 0.05 level (2-tailed).

Form the women section, in the same breath, this study observes that the following two indicators with their respective communalities are:

"Not knowing how to formulate search questions" is a factor hindering me seeking information (.910) and "Lack of awareness about information sources" is a factor hindering me seeking information (.954). The indication here is that the extraction of factors for the females' data explained 91.0%

and 95.4% of the variances in the two indicators respectively.

This study confirms the hypothesis that the percentages of variances explained by the two factor extractions are greater among the males' data than they are among the females' data. This demonstrates the strength component extraction among the men as compared to the extraction among the women.

Table 4. Table showing the extracted communalities for males' data.

Communalities		
	Initial	Extraction
I talk to colleagues as a source of information	1.000	.986
I consult doctors as a source of information	1.000	.987
Consulting information sources assists me in treating my patients	1.000	.994
Consulting information sources assists me in the prognosis of my patients	1.000	.972

Extraction Method: Principal Component Analysis.

a. Respondent's gender=Male.

### 3.2. Females

**Table 5.** Table showing the extracted communalities for females' data.

Communalities		
	Initial	Extraction
I talk to colleagues as a source of information	1.000	.784
I consult doctors as a source of information	1.000	.777
Consulting information sources assists me in treating my patients	1.000	.857
Consulting information sources assists me in the prognosis of my patients	1.000	.870

Extraction Method: Principal Component Analysis.

a. Respondent's gender=Female.

*Observation:* Thus the total variance for  $X$  can be expressed as trace ( $\Sigma$ ), but by the same Property, this is also the total variance for  $Y$ .

The "Total Variance Explained" demonstrates the number of components that the analysis was able to determine using eigen values amounting to the percentage of variance extracted. Under this comparative bi-analysis setup, 9 components were established for the men's setup. According to Exploratory Factor Analysis among participating males, a

component is only considered to be one of the factors extracted if the initial Eigen value is greater than or equal to 2.00. Accordingly, factors from 10 up to 47 are not significant and thus, not important and not included.

On the other side of the females' data, the analysis determined 8 components since those that had an eigen-value at least equal to 2 were only 8. It is evident that from the ninth component, the eigen-values start to reduce below 2.

**Table 6.** Table presenting the Total Variance Explained for males' data.

Total Variance Explained									
Component	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative%	Total	% of Variance	Cumulative%	Total	% of Variance	Cumulative%
1	8.995	19.138	19.138	8.995	19.138	19.138	6.898	14.677	14.677
2	7.770	16.533	35.671	7.770	16.533	35.671	6.835	14.544	29.220
3	5.455	11.607	47.278	5.455	11.607	47.278	5.484	11.668	40.888
4	5.256	11.183	58.461	5.256	11.183	58.461	5.453	11.603	52.491
5	3.666	7.800	66.260	3.666	7.800	66.260	4.475	9.521	62.013
6	3.175	6.756	73.016	3.175	6.756	73.016	3.800	8.085	70.097
7	2.834	6.031	79.047	2.834	6.031	79.047	3.175	6.756	76.853
8	2.671	5.683	84.730	2.671	5.683	84.730	3.125	6.649	83.503
9	2.013	4.283	89.013	2.013	4.283	89.013	2.590	5.510	89.013
10	1.559	3.317	92.330						
45	-1.170E-15	-2.490E-15	100.000						
46	-1.365E-15	-2.904E-15	100.000						
47	-3.698E-15	-7.867E-15	100.000						

Extraction Method: Principal Component Analysis.

Respondent's gender=Female.

**Table 7.** Table presenting the Total Variance Explained for females' data.

Total Variance Explained									
Component	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative%	Total	% of Variance	Cumulative%	Total	% of Variance	Cumulative%
1	9.190	19.554	19.554	9.190	19.554	19.554	5.425	11.542	11.542
2	5.929	12.614	32.168	5.929	12.614	32.168	4.839	10.296	21.838
3	4.674	9.945	42.113	4.674	9.945	42.113	4.605	9.797	31.635
4	4.308	9.166	51.279	4.308	9.166	51.279	4.590	9.765	41.400
5	3.332	7.090	58.369	3.332	7.090	58.369	4.544	9.667	51.067
6	3.063	6.518	64.886	3.063	6.518	64.886	4.201	8.937	60.005
7	2.697	5.738	70.624	2.697	5.738	70.624	4.199	8.934	68.939
8	2.137	4.548	75.172	2.137	4.548	75.172	2.930	6.233	75.172
9	1.894	4.029	79.201						
45	-1.008E-15	-2.145E-15	100.000						
46	-1.286E-15	-2.737E-15	100.000						
47	-1.534E-15	-3.263E-15	100.000						

Extraction Method: Principal Component Analysis.

a. Respondent's gender=Female.

The cumulative total percentages of 89.013% and 75.172% are high by any standard and show that significant fractions

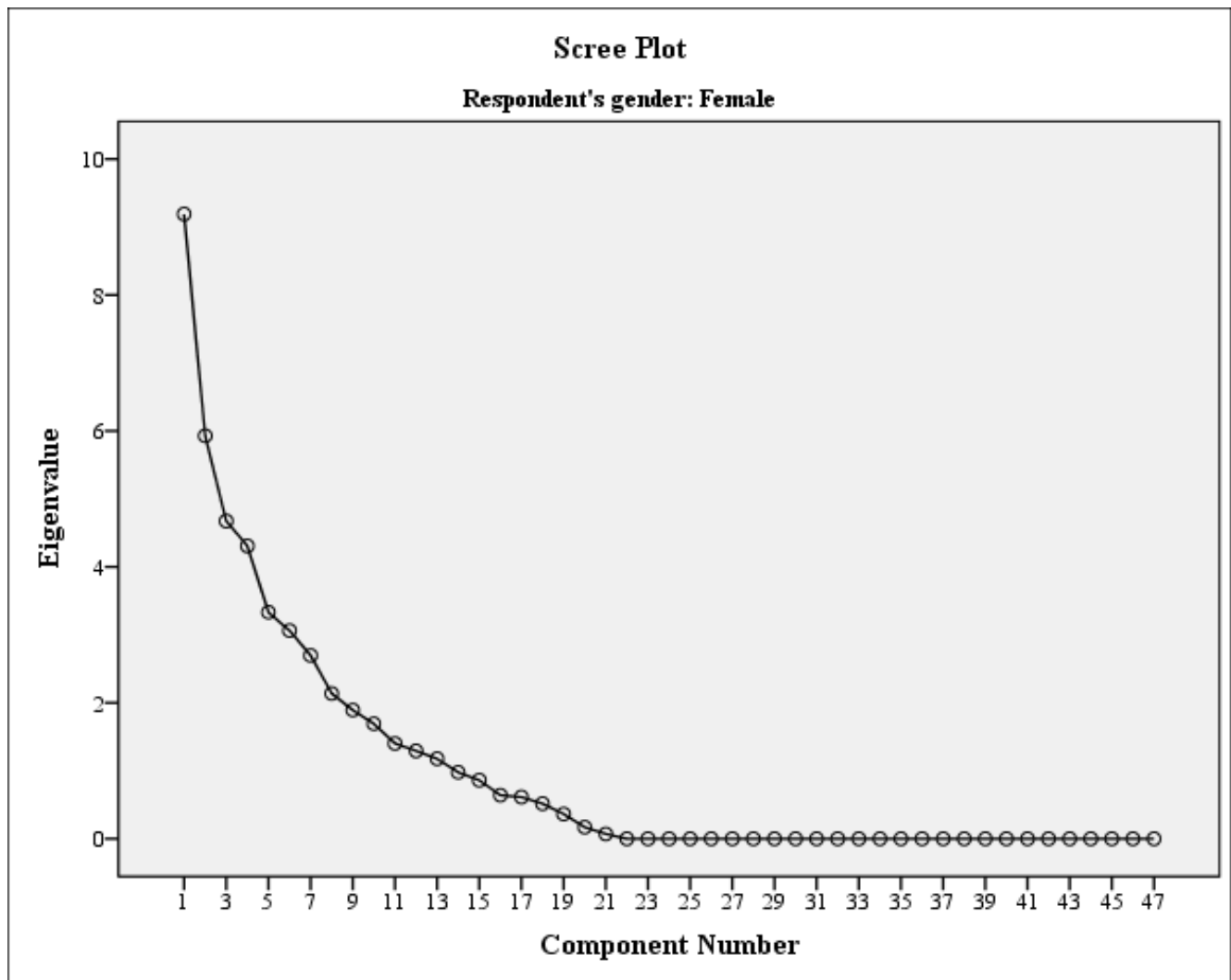
of the variances were accounted for. Therefore, according to this analysis, nine and eight factors were respectively extracted for males and females using principal component factor analysis.

The scree plot is a graph of the eigen-values against all the factors. In addition to identification and subsequent determination of the number of factors extracted, the graphs are useful for determining the number of factors to retain. The point of interest is where the curve starts to flatten (usually referred to as the Elbow). It can be seen that the curve begins to flatten between factors 9 and 10 for males while it begins to flatten between factors 8 and 9 for females. Note also that for both plots, factor 10 for males and factor 9

for females onwards have an eigenvalue of less than 2.00 and so only nine factors for males and eight factors for females were retained under the data extraction analyses for males and females.

Start with the following. If we multiply an  $n \times n$  matrix by an  $n \times 1$  vector we will get a new  $n \times 1$  vector back.

Notice that before we factored out the  $\eta \tilde{\eta} \tilde{\eta}$  we added in the appropriately sized identity matrix. This is equivalent to multiplying things by one and so doesn't change. It is usually of significant interest to know that the (total variance explained and Scree Plot) reveal the same number of components extracted under any given analysis. The two work complementarily with each other.



*Figure 1. Figure showing the scree plot for the males' data.*

*Table 8. Table presenting the Rotated Component Matrix for males' data.*

	Component								
	1	2	3	4	5	6	7	8	9
"Lack of awareness about information sources" is a factor hindering me seeking information	.936								
"Print material not being relevant" is a factor hindering me seeking information	.788								
"Lack of Physical library" is a factor hindering me seeking information	.528					.404		.537	
I use my book collection to access information									.864
I talk to colleagues as a source of information		.537							-.614

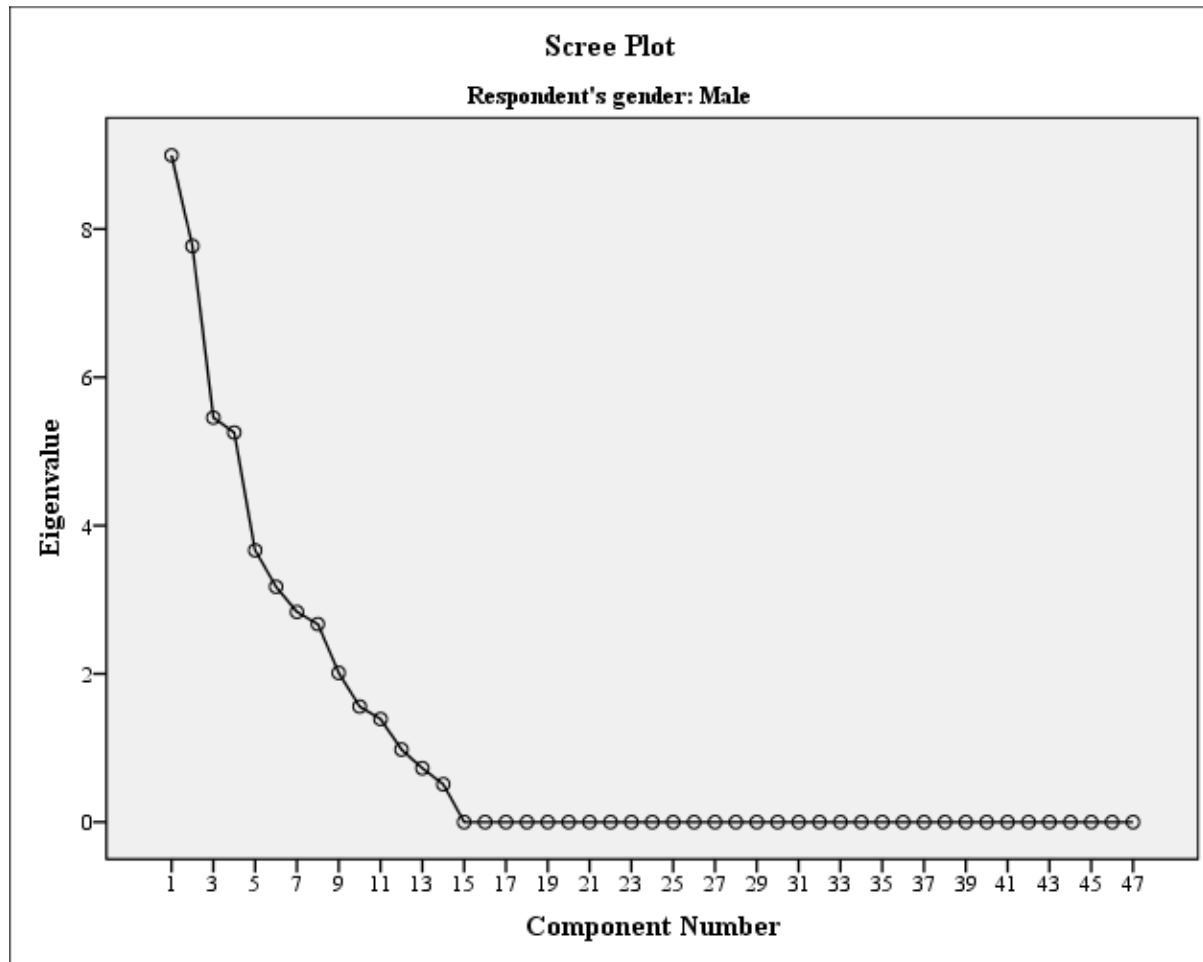


Figure 2. Figure showing the scree plot for the females' data.

The rotated component matrix below shows that the first component is constituted of the following indicators against its loading in brackets:

"Lack of awareness about information sources" is a factor hindering me seeking information (.936);

"Not knowing how to formulate search questions" is a factor hindering me seeking information (.860);

"Print material not being relevant" is a factor hindering me seeking information (.788);

I attend training workshops organized by district office as a source of information (.781);

"Lack of skill to search the online resources" is a factor hindering me seeking information (.772);

I use computer at work to access online information (.716);

"No Print material" is a factor hindering me seeking information (.639);

I consult doctors as a source of information (-.632).

#### Component 1

*Lack of technological knowhow in searching for information through the internet.*

*The second component constitutes of:*

Consulting information sources assists me in treating my patients (.946);

Consulting information sources assists me in the prognosis of my patients (.930);

Consulting information sources assists me to diagnose my patients (.897);

I use my mobile phone to access online information (.858);

I always consult information sources for clinical decisions (.696);

Cost of the information informs my choice of information source (.604);

I use my mobile phone to access internet as a source of information (-.445, .492).

#### Component 2

*External information support to diagnose and treat patients*

*The third component constitutes of:*

I consult the hospital policy manual as a source of information (.824);

I use protocols/guidelines kept in the ward or my pocket book as a source of information (.724);

I talk to people outside of my work as a source of information (.696);

"Lack of online access" is a factor hindering me seeking information (.451, -.685);

I only consult information sources for personal purposes (.553, .592);

"Slow internet" is a factor hindering me seeking information (.530, -.559);



I use all information sources that are available in my hospital (.515);

"Time taken to access information" is a factor hindering me seeking information;

Trustworthiness informs my choice of information source (.449);

#### *Component 3*

*Use of materials sources available within the hospital as sources of information*

*The fourth component constitutes of:*

I consult doctors as a source of information;

"Slow internet" is a factor hindering me seeking information (.530, -.559, -.475);

I use all information sources that are available in my hospital (.515, .468);

Access to information improves quality of patient care (.923);

I consult information sources for professional development (.897);

I understand the importance of information in my work (.808);

I am aware of all information sources that are available in my hospital (.616);

I attend training workshops organized by non-governmental organisations as a source of information (-.581);

"Cost of accessing information" is a factor hindering me seeking information (.470, .486);

I consult information sources for teaching (.561);

Familiarity informs me of the information source/awareness of the information source (.452).

#### *Component 4*

*Improvement of patient care through collaborative consultations*

#### *Component 5*

*Use of information acquired through workshops, seminars and journals to improve the participants' knowledge*

#### *Component 6*

*Use of internet and hospital facilities as sources of information*

#### *Component 7*

*Use of reference materials and the medical need dictates the source of information required*

#### *Component 8*

*Causes of limited availability sources of information*

#### *Component 9*

*Use of printed material and colleagues to access information*

The research team constructed the identification of other components and stated them here below.

#### *Component 1*

I talk to people outside of my work as a source of information (.779);

I attend training workshops organized by district office as a source of information (.778);

I attend training workshops organized by provincial office as a source of information (.759);

I use library books as a source of information (.751);

I attend training workshops organized by non-governmental organisations as a source of information (.673);

I consult doctors as a source of information (.652);

I consult the hospital policy manual as a source of information (.538);

I use computer at work to access internet as a source of information (.498);

I use Mind-set facilities available in the hospital as a source of information (.499).

#### *Component 1*

Using both external as well as internal sources to solicit information.

#### *Component 2*

A practice of consultative/collaborative medical practice when attending to patients.

I always consult information sources for clinical decisions (.924);

Consulting information sources assists me to diagnose my patients (.711);

I use all information sources that are available in my hospital (.710);

Consulting information sources assists me in the prognosis of my patients (.656);

I consult print material (books, journals) from the library (.655);

I use reference books kept in the hospital as a source of information (.651);

I read newspapers as a source of information (.615);

Consulting information sources assists me in treating my patients (.612).

#### *Component 2*

A practice of consultative/collaborative medical practice when attending to patients.

#### *Component 3*

The procedure of receiving information dictates the source of information to be used.

#### *Component 4*

Whereas availability of information improves patient care and treatment, there is serious limitation of information due to lack of skill for searching information, poor internet services and non-availability of library services.

#### *Component 5*

Lack of: a physical library, online access, slow internet and poor online searching skills contribute to non-availability of important medical information.

#### *Component 6*

Use of locally available sources of information including: colleagues, hospital internet and protocol guidelines kept in the ward.

#### *Component 7*

Lack of awareness of sources of information, time taken to access information, non-existent of sources of information.

#### *Component 8*

Varied sources of searching for useful medical information.

Table 9. Table presenting the Rotated Component Matrix for females' data.

	Component							
	1	2	3	4	5	6	7	8
I talk to people outside of your work as a source of information	.779							
I attend training workshops organized by district office as a source of information	.778							
"Not knowing how to formulate search questions" is a factor hindering me seeking information								.540
I attend seminars run in the hospital as a source of information								-.477

Extraction Method: Principal Component Analysis. Rotation Method: Varimax with Kaiser Normalization.

a. Respondent's gender=Female

## Component Plot in Rotated Space

Respondent's gender: Male

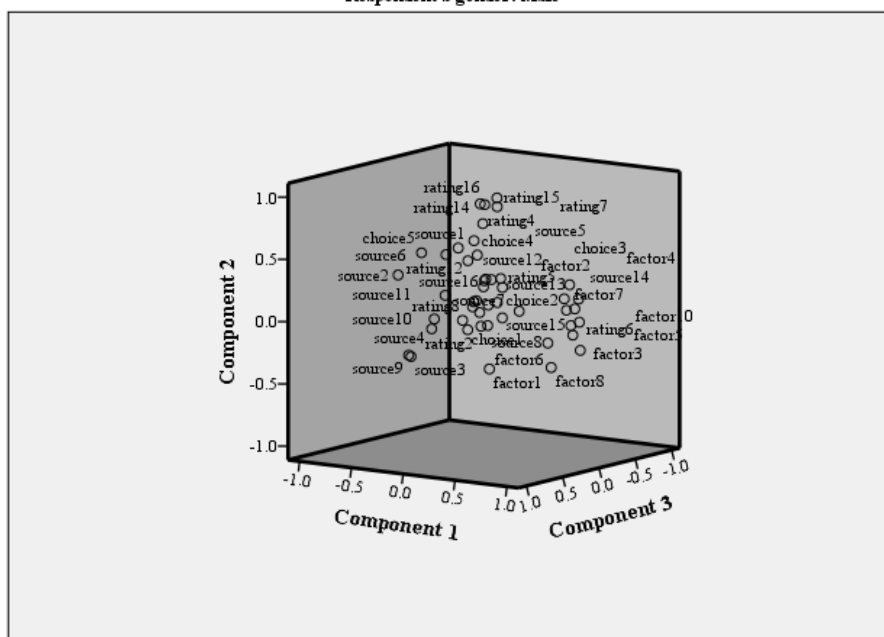


Figure 3. Figure showing the component plot in rotated for the males' data.

## Component Plot in Rotated Space

Respondent's gender: Female

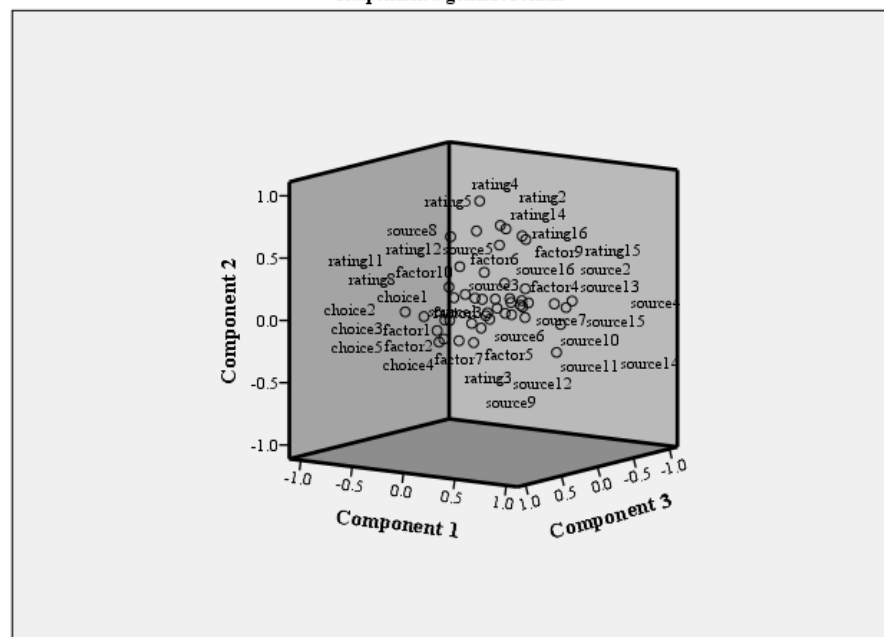


Figure 4. Figure showing the component plot in rotated for the females' data.

### 3.3. Confirmatory Factor Analysis

The researchers performed confirmatory factor analysis on the two types of data namely: the males' data and the females' data. Two separate types of analyses were performed: confirmatory factor analyses under path models and boots trap analyses models test statistics.

### 3.4. Confirmatory Path Model for Males' Data

Figure 5 below shows the confirmatory path model diagram for the males' data. Smart PLS, courtesy of enabled this analysis [4, 6–12]. The figure presents the factor to indicator loadings, cause effect relationship between factors and R-squared values on the factors. In more research statistical terminologies, the analysis is composed of: Standardised Regression Weights (measuring the strengths between factor variables), R-Squared values (which express the variance explained by explanatory factor variables) and Factor Loadings (which show the loading strength between each indicator and the corresponding variable).

The significant R-Squared values show that the amount of

variance explained by the explanatory factors. The higher the R-squared, the stronger the relationship between the pairs of any selected pairs of factor variables. This logic equally applies to loadings between factor variables and indicators. Consider the following pairs of factors and their corresponding R-squared values, indicators and factor variables, and regression weights derived between pairs of factor variable as some examples:

Factor-indicator loadings for factor 2: all the indicators are significantly loaded on factor 2. All indicators are significantly loaded on factor 4 and all indicators are significantly loaded on factor 5. Other indicator loadings can easily be read from the figure below.

R-squareds for factor 2, factor 3, factor 4, factor 5, factor 6, factor 8 and factor 1 explain large percentages of the variances.

The observed standardised regression weights show significance between pairs of most factor. The regression weight between factor 2, factor 3, factor 6, factor 5 show significance predictions. This means, that as factor 2 improves, so does factor 3, and as factor 3 improves in value, factor 4 equally improves in value.

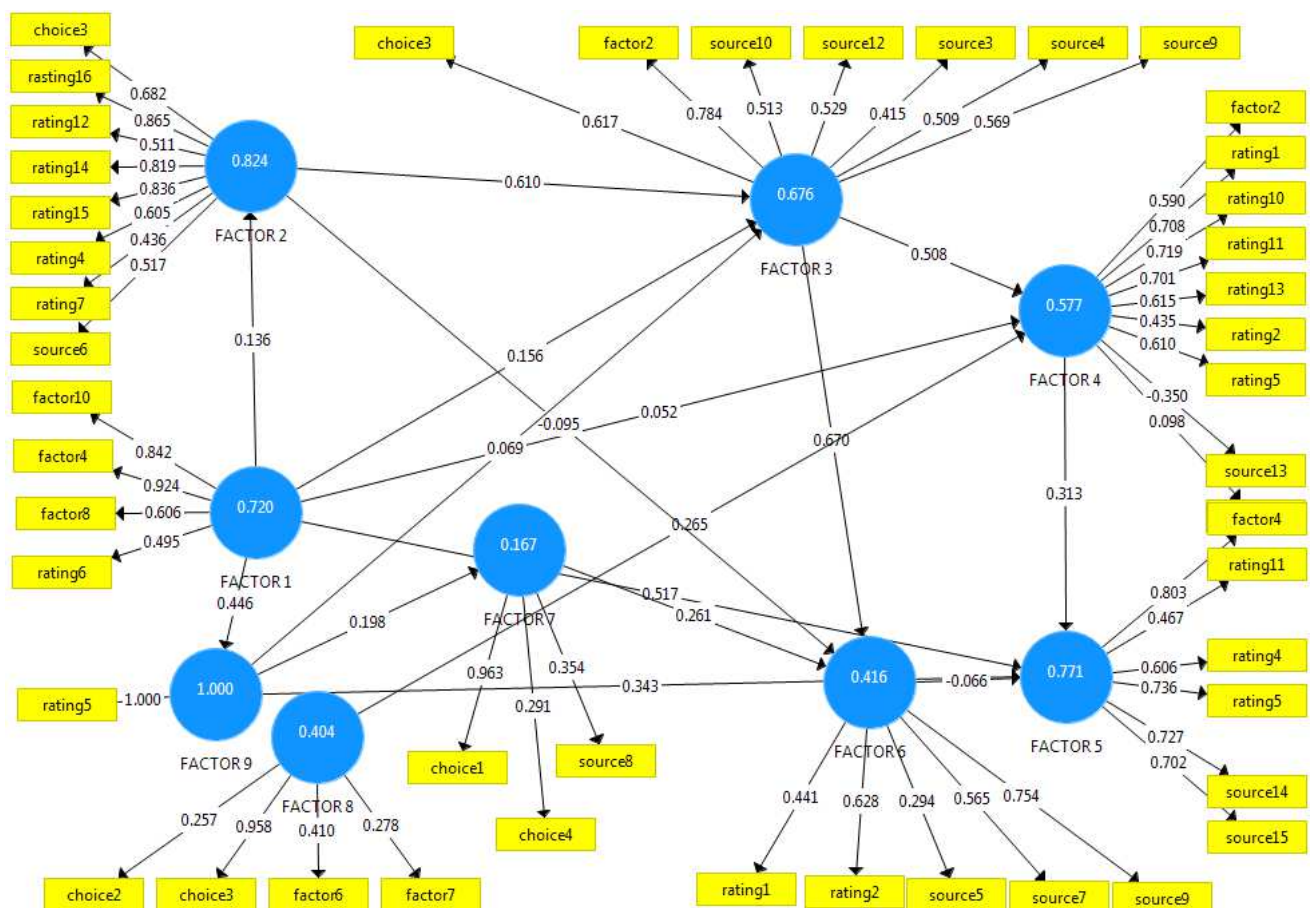


Figure 5. Figure showing the Confirmatory Factor Path Model for Males' Data.

### 3.5. Confirmatory Boots Trap Model for Males' Data

The following figure (figure 6) shows boots traps for the above males' factor analysis in view of the testing for

significance of the observed loadings. As a guiding rule, the testing is done at the 5% level of significance and of course at the 95% confidence level. The observed t-test statistics are classified as significant if either the p-value is smaller than

0.05 or the t-statistic is larger than 1.96. Observing this traditional rule, one notes the following indicators and factors have been interpreted as under this rule.

The researchers noted that choice 1, rating 16, rating 12, rating 14 up to source 6 have been significantly influenced by

factor 2. Similarly, factor 4, factor 10 and factor 8 have strongly been influenced by factor 1. On the other hand, rating 4, rating 5, source 14 and source 15 have significantly been influenced by factor 5. Other observations can easily be read from the figure below.

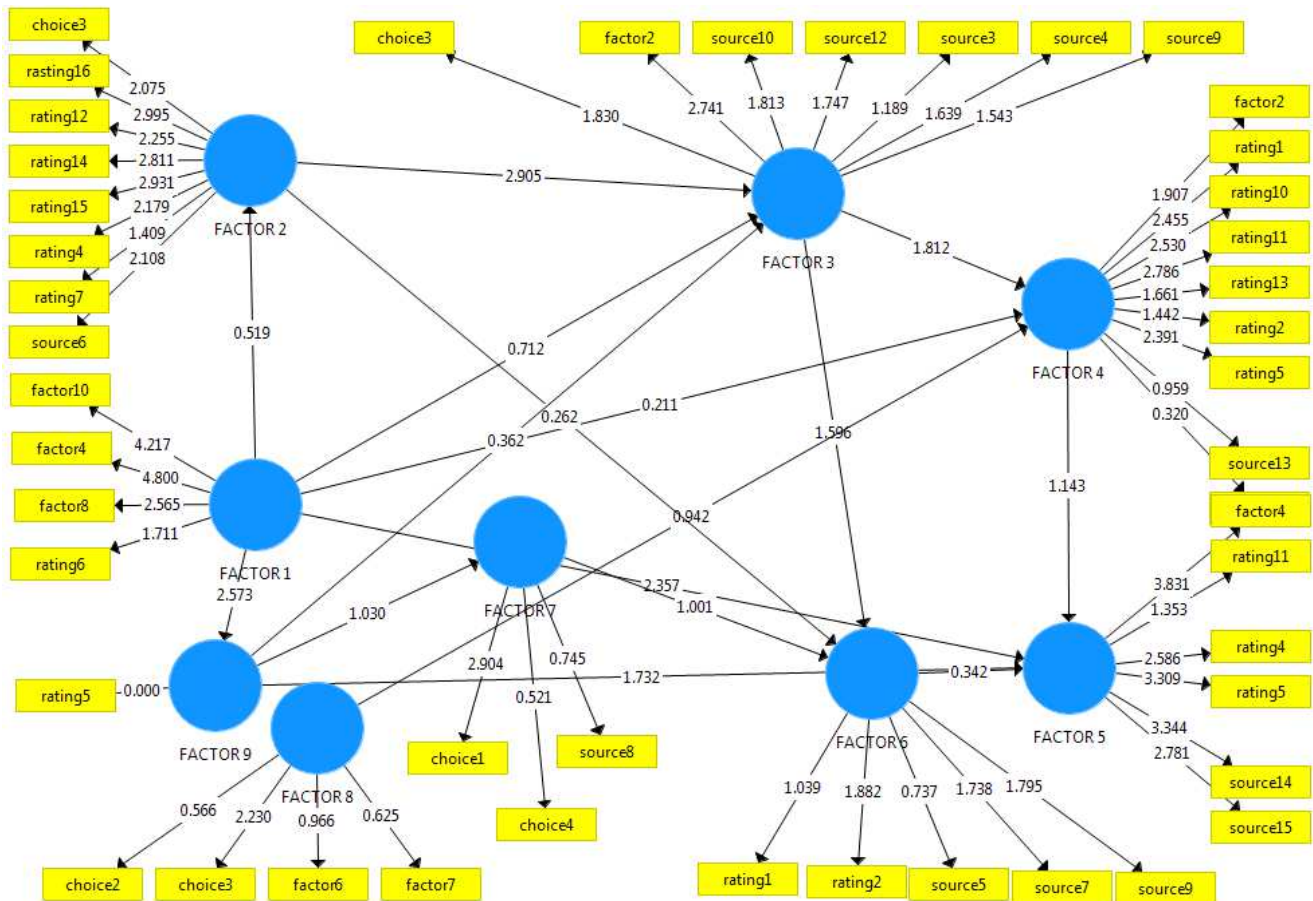


Figure 6. Figure showing the Confirmatory Boots Trap Model for Males' Dales' Data.

### 3.6. Confirmatory Factor Path Model for Females' Data

Figure 7 below shows the confirmatory path model diagram for the females' data. The figure presents the factor to indicator loadings, cause effect relationship between factors and R-squared values on the factors. In more research statistical terminologies, the analysis is composed of: Standardised Regression Weights (measuring the strengths between factor variables), R-Squared values (which express the variance explained by explanatory factor variables) and Factor Loadings (which show the loading strength between each indicator and the corresponding variable).

The researchers noted that most of the retained indicators demonstrate significance due to high observed loadings. It is observed that all the retained indicators had each extracted factor-loadings of values greater than 0.500. These high loadings show the high significance of the influence the indicators have on the factors. On the other hand, the high loading form the hypotheses point of view, they demonstrate the strong predictive relationship between a factor and the indicator.

### 3.7. Confirmatory Boots Trap Model for Females' Data

The figure below presents only one type of statistical output – t-statistics. These t-statistics are based on Standardised Regression Weights (measuring the strengths between factor variables), R-Squared values (which express the variance explained by explanatory factor variables) and Factor Loadings (which show the loading strength between each indicator and the corresponding variable as explained above).

The interpretations will follow the same rule for the three types of statistics calculated. The statistical rule of thumb is: either the test-statistic demonstrates significance or not. The level of significance is the default 0.05. The critical value at 0.05 level of significance is 1.96. Any test-statistic value observed from the figure below that is greater than 1.96 shows significance at the chosen 0.05 level of significance. It is understood therefore that: all factor loadings from factor 2 are statistically significant, source 3, source 4, source 6 down to source 2 are significantly loaded by factor 1. Choices 3, 4



and choice 5 are significantly loaded by factor 4. Other loadings can be determined similarly. With regard to R-squared, factor 2 significantly explains factor 5, factor 7

significantly explains factor 5 and factor 6 significantly explains factor 2, and others can be read from the figure below.

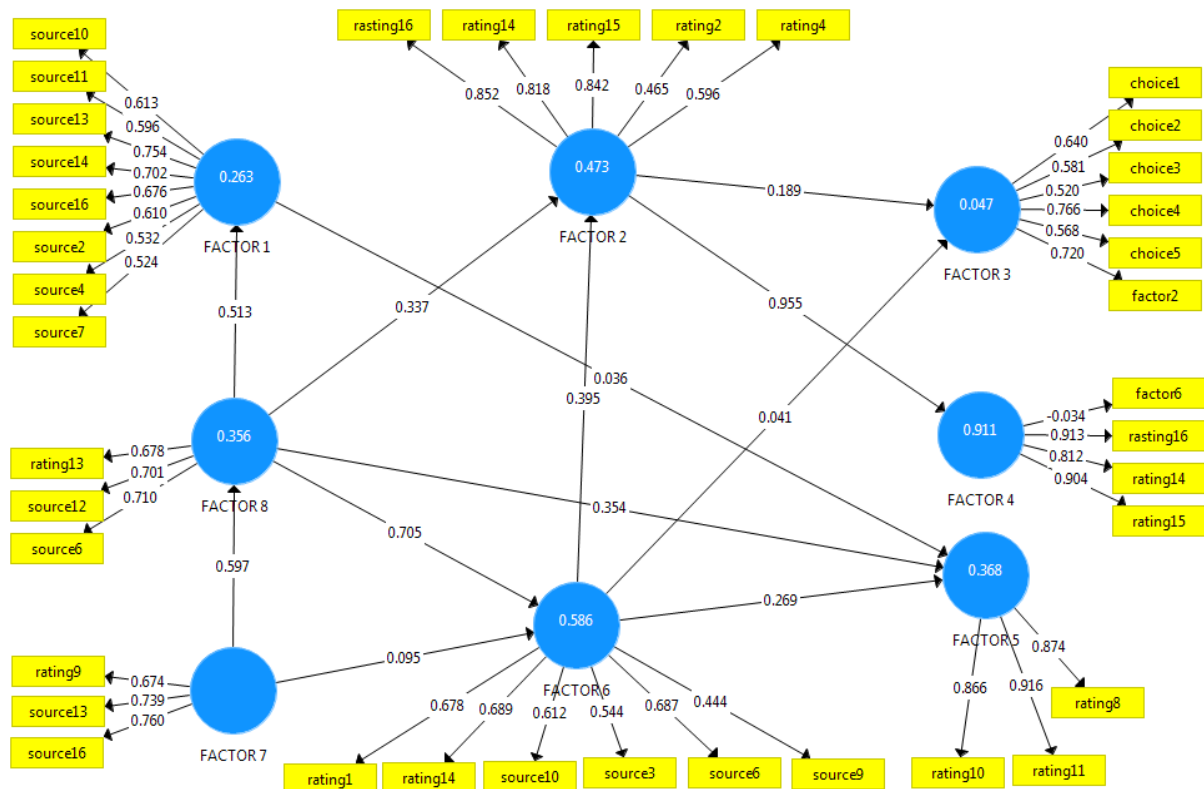


Figure 7. Figure showing the Confirmatory Factor Path Model for Females' Data.

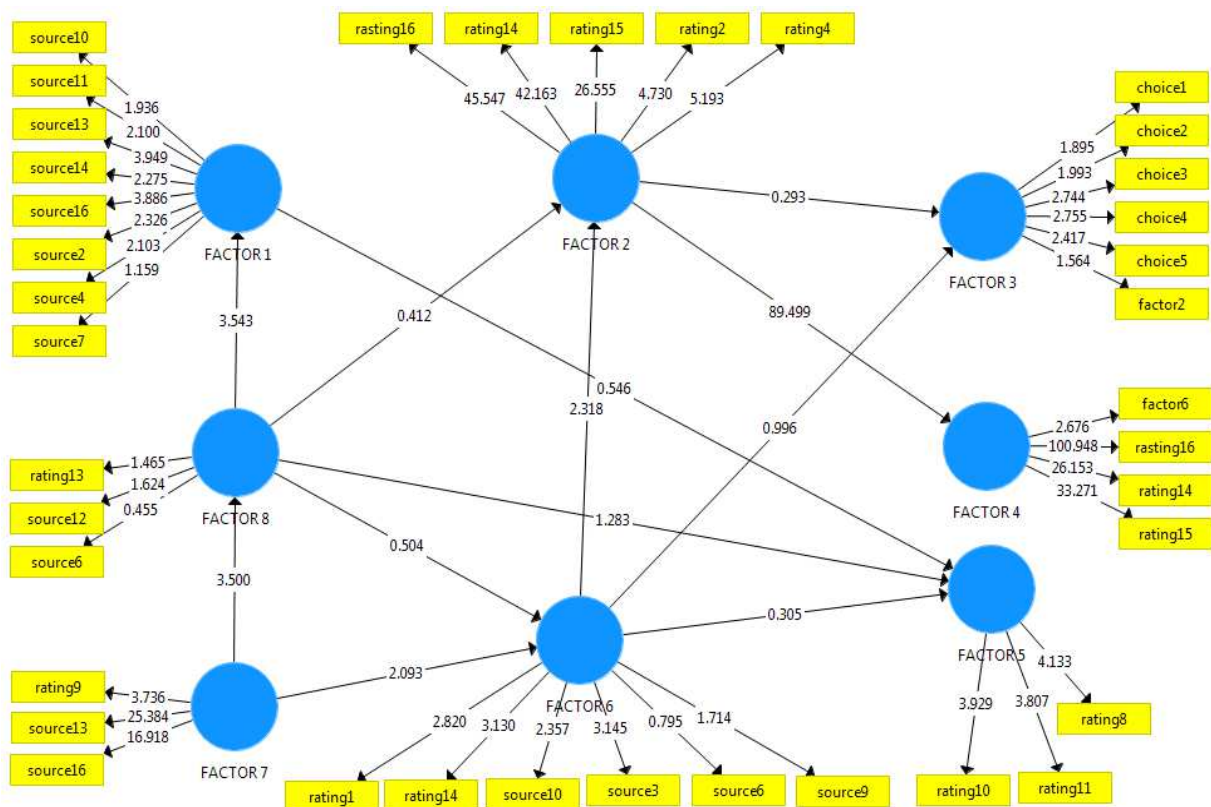


Figure 8. Figure showing the Confirmatory Boots Trap Path Model for Females' Data.

## 4. Discussion and Conclusion

The researchers noted that most of the retained indicators demonstrate significance due to high observed loadings. It is observed that all the retained indicators had each extracted factor-loadings of values greater than 0.500. These high loadings show the high significance of the influence the indicators have on the factors. On the other hand, the high loading from the hypotheses point of view, they demonstrate the strong predictive relationship between a factor and the indicator.

The main objective of this study was to understand component structures of data collected on medical professionals and split by gender. The data were collected from five hospitals in O. R. Tambo District Municipality in the Eastern Cape Province of the republic of South Africa.

The significant R-Squared values show that the amount of variance explained by the explanatory factors. The higher the R-squared, the stronger the relationship between the pairs of any selected pairs of factor variables. This logic equally applies to loadings between factor variables and indicators. Consider the following pairs of factors and their corresponding R-squared values, indicators and factor variables, and regression weights derived between pairs of factor variable as some examples:

F1 → F5 with an R-squared of 36.8%, F2 → F4 with an R-squared of 91.1%, F8 → F1 with an R-squared of 26.3%, F8 → F2 with an R-squared of 47.3%, F8 → F6 with an R-squared of 58.6%. these high percentages show strong relationships between the extracted factors. Though the stated percentages seem large by eye measurement, this judgement is not enough to draw a decision. The best is to perform a boots trap using Smart PLS, which will produce t-statistics and corresponding p-values. The following path diagram is a boots trap consisting t-statistics for testing the strengths of the loadings, the regression weights and the factors loadings.

The reader can witness that the descriptive analysis for the data over the different hospitals has been done, descriptive analysis over different variables have equally been accomplished. It can furthermore be observed that factor analysis revealed that when data were divided according to gender, nine factors were extracted from the males' data, while eight factors emerged from the females' data. This cross-sectional study revealed that there were differences arising from the two data sets depending on gender. It was observed, with a scrutiny, that the factors addressed different issues from the research point of view.

The total variance explained further shows that the 89.01% was explained by the extraction of the factors under the males data whereas, 75.17% of the variance was attributed to the females' data analysis. The scree plots further demonstrated

this noticeable differences. Finally, the component plot in rotated space shows this marked differences.

## References

- [1] Almunawar MN, Anshari M (2012) Health Information Systems (HIS): Concept and Technology.
- [2] Higgins C, Jesse C, Deborah P, Rober M, Janice G, Meyers D (2015) Using Health Information Technology to Support Quality Improvement in Primary Care | PCMH Resource Center. <https://pcmh.ahrq.gov/page/using-health-information-technology-support-quality-improvement-primary-care>. Accessed 3 Jun 2021.
- [3] Khalifa M, Alsweilem O (2015) Hospital information systems (HIS) acceptance and satisfaction: a case study of a Tertiary Care Hospital. *Procedia Computer Science* 63: 198–204.
- [4] Ek S (2015) Gender differences in health information behaviour: a Finnish population-based survey. *Health Promot Int* 30: 736–745.
- [5] Tong V, Raynor DK, Aslani P (2014) Gender differences in health and medicine information seeking behaviour: a review.
- [6] Jiménez-Pernett J, Olry de Labry-Lima A, García-Gutiérrez JF, Salcedo-Sánchez M del C, Bermúdez-Tamayo C (2010) Sex differences in the use of the Internet as a source of health information among adolescents. *Telemed J E Health* 16: 145–153.
- [7] Finney Rutten LJ, Blake KD, Greenberg-Worisek AJ, Allen SV, Moser RP, Hesse BW (2019) Online Health Information Seeking Among US Adults: Measuring Progress Toward a Healthy People 2020 Objective. *Public Health Rep* 134: 617–625.
- [8] Garrido T, Raymond B, Jamieson L, Liang L, Wiesenthal A (2004) Making the business case for hospital information systems--a Kaiser Permanente investment decision. *J Health Care Finance* 31: 16–25.
- [9] Wienert J (2019) Understanding Health Information Technologies as Complex Interventions with the Need for Thorough Implementation and Monitoring to Sustain Patient Safety. *Front ICT*. <https://doi.org/10.3389/fict.2019.00009>.
- [10] Chen C-J, Kendall J, Shyu Y-IL (2010) Grabbing the rice straw: health information seeking in Chinese immigrants in the United States. *Clin Nurs Res* 19: 335–353.
- [11] Feldman SS, Buchalter S, Hayes LW (2018) Health Information Technology in Healthcare Quality and Patient Safety: Literature Review. *JMIR Med Inform*. <https://doi.org/10.2196/10264>.
- [12] Haux R, Ammenwerth E, Herzog W, Knaup P (2002) Health care in the information society. A prognosis for the year 2013. *Int J Med Inform* 66: 3–21.