

Análisis factorial exploratorio mediante el uso de las medidas de adecuación muestral kmo y esfericidad de bartlett para determinar factores principales

Exploratory factor analysis using the kmo and bartlett sphericity sample adequacy measures to determine main factors

AUTORES: Kelvin Pizarro Romero^{1*}
Omar Martínez Mora²

DIRECCIÓN PARA CORRESPONDENCIA: khpizarro2016@gmail.com

Fecha de recepción: 04 / 10 / 2020

Fecha de aceptación: 28 / 12 / 2020

RESUMEN

El objetivo del presente trabajo es demostrar la importancia de la técnica estadística del Análisis Factorial Exploratorio (AFE). La metodología utilizada fue a través del test KMO (Kaiser, Meyer y Olkin), aplicabilidad del análisis factorial de las variables estudiadas, por medio de la prueba de esfericidad de Bartlett, así como también, el porcentaje de varianza explicada, matrices policórica y componentes de factores principales. La base de datos presenta 27 variables, para lograr un mejor modelo solo se conservaron 18. Los resultados alcanzados, donde la escala de medición es en escala de likert determinaron que el mejor modelo, amparado en la mayor varianza acumulada, fue la de cuatro factores con un porcentaje del 51 %.

Palabras clave: *análisis factorial exploratorio, Bartlett, licker, KMO, varimax.*

ABSTRACT

The objective of this research is to demonstrate the importance of the statistical technique of Exploratory Factor Analysis (EFA). The methodology used was through the KMO test

¹Doctorando de Estadística y Matemática Aplicada, Universidad Nacional de Tumbes, khpizarro2016@gmail.com

²Doctorando de Estadística y Matemática Aplicada, Universidad Nacional de Tumbes, emartinez@utmachala.edu.ec

(Kaiser, Meyer and Olkin), applicability of the factor analysis of the variables studied, by means of the Bartlett test of sphericity, as well as the percentage of explained variance, polychoric matrices and components of main factors. The database shows 27 variables, in order to achieve a better model, only 18. The results achieved, where the measurement scale is on the Likert scale, determined that the best model, based on the highest accumulated variance, was that of four factors with a percentage of 51%.

Key words: *exploratory factor analysis, Bartlett's, licker, KMO, varimax.*

INTRODUCCIÓN

La Técnica del análisis factorial exploratorio, es una de las técnicas multivariante, que influye en realizar la búsqueda para poder reducir la dimensión de los datos, se utiliza para varios propósitos, todos ellos encaminados a explicar un fenómeno o constructo desde un menor número de variables llamados factores. Este método se puede utilizar en variables continuas que son la mayoría generalmente, así mismo lo puede realizar en variables dicotómica (Khan, 2006).

El AFE tiene como objetivo definir la estructura subyacente de un conjunto de datos, analizando las interrelaciones o correlaciones entre variables, ya sean estos ítems o preguntas de una prueba, respuesta o encuesta entre muchas otras. Las variables analizadas son agrupadas en conjunto de componentes o factores, los cuales vendrían siendo la estructura subyacente. Este método AFE busca componentes o factores que expliquen la mayor parte de varianza total.

El estadístico Francis Galton (1889) comienza la realización de este método AFE, bajo un concepto estadístico de correlación y regresión hacia la media, propuso el concepto de rasgo latente para revelar por qué unos conjuntos de variables se correlacionaban unas con otras. Cada factor se encarga de agrupar a los ítems que se encuentran intercorrelacionados, y a su vez al mismo tiempo, comparativamente independientes de los conjuntos “factores” de ítems restantes (Khan, 2006).

En términos generales, el análisis factorial exploratorio (AFE) prácticamente realiza un enfoque en destinar a un conjunto de métodos estadísticos multivariados de interdependencia con el propósito principal de poder desarrollar una estructura de factores subyacentes a un conjunto amplio de datos. Por tanto, para Arias y Justicia (2003) el análisis factorial exploratorio indica que este método AFE se usa para tratar de descubrir la estructura interna de una numerosa cantidad de variables para el estudio.

Las siglas AFE puede hacer referencia a un determinado número de conjunto de técnicas estadísticas como también a un método único de interdependencia (Khan, 2006), que se emplea con el propósito de reducir la dimensión de un gran número de indicadores operativos en un número inferior de variables conceptuales (Blalock, 1966). Esto compone una técnica frecuentemente utilizada en las ciencias sociales.

El desarrollo básico de este método AFE se debe atribuir a otro estadístico como lo fue Carl Pearson (1901), ya que desarrolló el coeficiente de correlación y bosquejó, principios en que se fundamenta el análisis factorial de ejes principales.

El objetivo de este trabajo es demostrar un AFE observando las diferencias en cuanto a las medidas de adecuación muestral KMO, test de esfericidad de Bartlett, porcentaje de varianza explicada, Matrices policórica y componentes de factores principales. Para esto se trabajará un base de datos en escala de lickert mediante el programa estadístico r-estudio.

METODOLOGÍA.

Para poder desarrollar un AFE se recomienda que la matriz sea rectangular, o sea tener más registros que variables esto con el fin de asegurar un menor error de muestreo. Si la trabaja con muestras pequeñas probabilidad aumenta de que las correlaciones varíen de una muestra respecto a otra, y así mismo se obtiene factores inestables y los resultados sean engañosos, por este motivo siempre se recomienda muestras mayores a 300 registros (Blalock, 1966).

También debe existir multicolinealidad que es la que indica la relación de dependencia lineal fuerte entre más de dos variables y esta debe ser métrica. En el AFE utilizaremos método de rotación ortogonal varimax, ya que este método es el más utilizado en las ciencias sociales, tiene como finalidad maximizar la suma de las varianzas de las cargas factoriales de cada una de las variables.

Este trabajo desarrolla el análisis factorial exploratorio con el software estadístico r-estudio, en una data que se propone observar el Constructo “Perfil Sicológico_del joven adulto” aplicado a 610 jóvenes de la ciudad de Guayaquil, desde diferentes teorías sicológicas para describir al hombre.

Este test tiene 4 dimensiones: human: humanismo (tiene 9 ítems), materia: materialismo (tiene 9 ítems), cognitivo: cognitivo (tiene 6 ítems) y cogemoc: cognitivo emocional (tiene 3 ítems).

Tabla 1. Estructura de la Base de datos

test<-read.delim ("clipboard")	→ Importar la data excel
> str(test)	→ Estructura de la data llamada test
'data.frame': 610 obs. of 27 variables:	
\$ HUMAN1 : int 3 3 3 3 3 4 5 3 3 1 ...	
\$ COGNITIVO1: int 4 2 3 4 3 3 5 5 3 3 ...	
\$ MATERIA1 : int 2 3 3 3 3 2 1 3 3 2 ...	
\$ MATERIA2 : int 2 2 4 3 3 2 1 1 2 1 ...	
\$ HUMAN2 : int 2 3 2 3 3 2 1 3 2 2 ...	
\$ HUMAN3 : int 4 3 4 4 4 3 5 3 5 5 ...	
\$ COGNITIVO2: int 2 3 2 4 3 2 1 3 2 3 ...	
\$ MATERIA3 : int 2 2 1 3 3 1 5 5 3 1 ...	
\$ HUMAN4 : int 4 3 4 4 3 2 5 3 4 5 ...	
\$ COGNITIVO3: int 3 2 3 4 3 2 1 3 4 1 ...	
\$ MATERIA4 : int 2 3 3 3 3 2 5 3 4 4 ...	
\$ HUMAN5 : int 2 3 2 4 2 3 1 3 2 4 ...	
\$ COGNITIVO4: int 3 2 3 4 3 4 5 4 3 3 ...	
\$ MATERIA5 : int 2 3 3 3 3 1 1 1 3 3 ...	
\$ HUMAN6 : int 2 3 1 4 1 2 1 3 2 1 ...	
\$ MATERIA6 : int 2 3 2 3 3 1 1 2 3 2 ...	
\$ HUMAN7 : int 3 3 3 4 2 4 1 3 2 2 ...	
\$ COGNITIVO5: int 3 3 4 4 3 1 5 3 2 2 ...	
\$ HUMAN8 : int 3 3 3 3 2 4 1 3 3 1 ...	
\$ MATERIA8 : int 1 5 1 3 3 1 5 4 3 1 ...	
\$ HUMAN9 : int 3 3 3 3 4 2 1 3 4 5 ...	
\$ COGNITIVO6: int 3 4 4 4 3 5 5 2 4 5 ...	
\$ MATERIA9 : int 3 2 2 3 3 2 1 5 4 5 ...	
\$ COGEMOC1 : int 3 3 3 4 3 2 5 4 2 2 ...	
\$ MATERIA7 : int 3 1 1 3 3 1 5 1 2 1 ...	
\$ COGEMOC3 : int 2 3 4 1 2 4 5 2 2 3 ...	
\$ COGEMOC2 : int 3 3 4 4 3 3 5 2 3 3 ...	

Fuente: Elaborado por los autores.

Observamos la tabla 1 que indica la estructura de la data llamada en mi caso “test” contiene 610 observaciones y 27 variables, además se aprecia que se encuentra en escala de Likert con una valoración de 1 hasta un máximo de 5.

Esta base de datos cumple con lo indicado para un análisis factorial exploratoria, porque podemos observar que posee más de 300 participantes lo resultados útiles y relativamente estables tal como lo indica (Tabachnick y Fidell, 2001).

Tabla 2. Nombre de las variables y dimensión de la Base de datos

> names(test)	→ Los nombres de los 27 ítems “variables”
[1] "HUMAN1" "COGNITIVO1" "MATERIA1" "MATERIA2" "HUMAN2" "HUMAN3" "COGNITIVO2" "MATERIA3" "HUMAN4" "COGNITIVO3" "MATERIA4"	
[12] "HUMAN5" "COGNITIVO4" "MATERIA5" "HUMAN6" "MATERIA6" "HUMAN7" "COGNITIVO5" "HUMAN8" "MATERIA8" "HUMAN9" "COGNITIVO6"	
[23] "MATERIA9" "COGEMOC1" "MATERIA7" "COGEMOC3" "COGEMOC2"	
.	
> dim(test)	→ La dimensioe de la base de datos
[1] 610 27	

Fuente: Elaborado por los autores

En la tabla 2 con esas funciones también se puede observamos los 27 nombres de las variables y la dimensión de los datos en estudio.

Vamos a determinar bajo tres criterios estadísticos (correlaciones policóricas, la prueba de esfericidad de Bartlett y la prueba KMO), si es posible proceder a un análisis factorial y explicaremos el porqué del mismo.

Comenzaremos por elaborar la matriz de correlación policórica para determinar si existen altas o bajas correlaciones, para realizar este paso necesitamos instalar en el r-studio el paquete psych y ejecutar los siguientes códigos.

Tabla 3. Funciones para visualizar ventana emergente de matriz policórica

> library(psych)	
> r.poly=polychoric(test)	
> r.poly	
> R=r.poly\$rho	→ Número de filas.
> View(R)	

Fuente: Elaborado por los autores

A continuación se imprime a través de las funciones de la tabla 3, la matriz de correlaciones policórica con su código “con su función polychoric” que es una de las técnicas para poder estimar la correlación entre dos variables latentes continuas, a esta matriz de correlación policórica la hemos denotado en este caso (R) y se mostrara una con la función (view) una ventana emergente para su mayor visualización, el cual por ser una matriz amplia se la colocó en este trabajo en 2 partes.

Tabla 4. Matriz de Correlación Policórica

	HUMAN1	COGNITIV O1	MATERIA1	MATERIA 2	HUMAN2	HUMAN3	COGNITIV O2	MATERIA 3	HUMAN4	COGNITIV O3	MATERIA 4	HUMAN5
HUMAN1	1.00000	0.17259	0.31657	0.19459	0.43859	0.26280	0.105823	0.306919	0.348961	0.057080	0.328954	0.300425
COGNITIVO1	0.17259	1.00000	-0.01577	-0.07268	0.02776	0.18151	0.426529	0.001243	0.096859	0.416573	0.028674	0.082005
MATERIA1	0.31657	-0.01577	1.00000	0.39883	0.64690	0.16838	0.241684	0.632092	0.148865	0.136213	0.587110	0.302594
MATERIA2	0.19459	-0.07268	0.39883	1.00000	0.30652	-0.00960	0.070409	0.351455	-0.019232	0.043031	0.327270	0.096697
HUMAN2	0.43859	0.02776	0.64690	0.30652	1.00000	0.24465	0.202935	0.523063	0.265504	0.109462	0.457316	0.480651
HUMAN3	0.26280	0.18151	0.16838	-0.00960	0.24465	1.00000	0.151295	0.195969	0.387663	0.176647	0.218556	0.317018
COGNITIVO2	0.10582	0.42653	0.24168	0.07041	0.20294	0.15129	1.000000	0.222391	0.085179	0.564790	0.278821	0.175207
MATERIA3	0.306919	0.001243	0.632092	0.351455	0.523063	0.195969	0.222391	1.000000	0.177943	0.132145	0.499320	0.361831
HUMAN4	0.348961	0.096859	0.148865	-0.019232	0.265504	0.387663	0.085179	0.177943	1.000000	0.070725	0.094488	0.414554
COGNITIVO3	0.057080	0.416573	0.136213	0.043031	0.109462	0.176647	0.564790	0.132145	0.070725	1.000000	0.259093	0.123644
MATERIA4	0.328954	0.028674	0.587110	0.327270	0.457316	0.218556	0.278821	0.499320	0.094488	0.259093	1.000000	0.281006
HUMAN5	0.300425	0.082005	0.302594	0.096697	0.480651	0.317018	0.175207	0.361831	0.414554	0.123644	0.281006	1.000000
COGNITIVO4	0.102235	0.520707	0.152863	0.011906	0.124249	0.226366	0.415090	0.178332	0.095379	0.434901	0.147560	0.200490
MATERIA5	0.223896	0.038299	0.593269	0.398426	0.381933	0.072147	0.160797	0.493819	0.086576	0.093548	0.450104	0.234336
HUMAN6	0.247688	0.058190	0.369510	0.164337	0.532203	0.250175	0.125357	0.507317	0.270905	0.066939	0.199741	0.611520
MATERIA6	0.265233	0.074740	0.594905	0.411677	0.441540	0.180924	0.263439	0.610942	0.150461	0.200193	0.538897	0.336997
HUMAN7	0.297302	0.061510	0.292369	0.124010	0.502518	0.281590	0.131016	0.233872	0.319701	0.094210	0.176718	0.494606
COGNITIVO5	0.077822	0.462137	0.150290	0.064678	0.119158	0.087026	0.399450	0.177010	0.039395	0.429355	0.230650	0.080899
HUMAN8	0.266947	0.046892	0.234917	0.052422	0.407236	0.265328	0.052849	0.202675	0.302860	0.063049	0.123878	0.364269
MATERIA8	0.220542	0.007289	0.530334	0.400328	0.451633	0.097044	0.150541	0.492710	0.122409	0.097824	0.461091	0.230005
HUMAN9	0.251480	0.080035	0.266440	0.192828	0.359712	0.275957	0.129355	0.207565	0.327958	0.104109	0.135744	0.450706
COGNITIVO6	0.024351	0.509380	-0.055816	-0.133356	-0.08748	0.115858	0.273393	-0.031701	0.025095	0.350936	0.042592	0.030266
MATERIA9	0.26993	0.11002	0.45614	0.21874	0.37932	0.21021	0.19945	0.45368	0.18529	0.19008	0.36912	0.31300
COGEMOC1	0.03082	0.36396	0.03643	0.18344	-0.00154	-0.04251	0.31692	0.04302	-0.07540	0.25011	0.05685	-0.06601
MATERIA7	0.15085	-0.13048	0.44928	0.44943	0.34234	-0.01318	0.16214	0.40523	0.01117	0.08469	0.42104	0.11609
COGEMOC3	-0.13048	0.08960	-0.12436	0.03956	-0.17335	-0.11635	0.07995	-0.03103	-0.10847	0.05131	-0.04839	-0.08363
COGEMOC2	-0.01468	0.23363	-0.06398	0.19888	-0.00105	-0.06429	0.17953	0.00736	-0.01661	0.23376	0.00789	0.00600

Fuente: Elaborado por los autores.

Tabla 4. Matriz de Correlación Policórica

	COGNITIV O4	MATERIA 5	HUMAN6	MATERIA 6	HUMAN7	COGNITIV O5	HUMAN8	MATERIA 8	HUMAN9	COGNITIV O6	MATERIA 9	COGEMO C1	MATERIA 7	COGEMO C3	COGEMO C2
HUMAN1	0.10223	0.22390	0.24769	0.26523	0.29730	0.07782	0.26695	0.22054	0.25148	0.02435	0.26993	0.03082	0.15085	-0.13048	-0.01468
COGNITIVO1	0.52071	0.03830	0.05819	0.07474	0.06151	0.46214	0.04689	0.00729	0.08003	0.50938	0.11002	0.36396	-0.13048	0.08960	0.23363
MATERIA1	0.15286	0.59327	0.36951	0.59490	0.29237	0.15029	0.23492	0.53033	0.26644	-0.05582	0.45614	0.03643	0.44928	-0.12436	-0.06398
MATERIA2	0.01191	0.39843	0.16434	0.41168	0.12401	0.06468	0.05242	0.40033	0.19283	-0.13336	0.21874	0.18344	0.44943	0.03956	0.19888
HUMAN2	0.12425	0.38193	0.53220	0.44154	0.50252	0.11916	0.40724	0.45163	0.35971	-0.08748	0.37932	-0.00154	0.34234	-0.17335	-0.00105
HUMAN3	0.22637	0.07215	0.25017	0.18092	0.28159	0.08703	0.26533	0.09704	0.27596	0.11586	0.21021	-0.04251	-0.01318	-0.11635	-0.06429
COGNITIVO2	0.41509	0.16080	0.12536	0.26344	0.13102	0.39945	0.05285	0.15054	0.12935	0.27339	0.19945	0.31692	0.16214	0.07995	0.17953
MATERIA3	0.17833	0.49382	0.50732	0.61094	0.23387	0.17701	0.20268	0.49271	0.20756	-0.03170	0.45368	0.04302	0.40523	-0.03103	0.00736
HUMAN4	0.09538	0.08658	0.27091	0.15046	0.31970	0.03939	0.30286	0.12241	0.32796	0.02510	0.18529	-0.07540	0.01117	-0.10847	-0.01661
COGNITIVO3	0.43490	0.09355	0.06694	0.20019	0.09421	0.42935	0.06305	0.09782	0.10411	0.35094	0.19008	0.25011	0.08469	0.05131	0.23376
MATERIA4	0.14756	0.45010	0.19974	0.53890	0.17672	0.23065	0.12388	0.46109	0.13574	0.04259	0.36912	0.05685	0.42104	-0.04839	0.00789
HUMAN5	0.20049	0.23434	0.61152	0.33700	0.49461	0.08090	0.36427	0.23001	0.45071	0.03027	0.31300	-0.06601	0.11609	-0.08363	0.00600
COGNITIVO4	1.00000	0.19250	0.17474	0.18926	0.18269	0.44050	0.12943	0.11391	0.15731	0.50533	0.23085	0.21520	-0.01334	0.07979	0.22095
MATERIA5	0.19250	1.00000	0.43341	0.63140	0.28451	0.13488	0.11954	0.43238	0.28908	-0.00487	0.44216	-0.02348	0.34017	0.03610	0.02765
HUMAN6	0.17474	0.43341	1.00000	0.42909	0.55712	0.10116	0.36763	0.28781	0.38607	-0.01819	0.26800	-0.08656	0.09590	-0.03561	-0.02399
MATERIA6	0.18926	0.63140	0.42909	1.00000	0.31964	0.24050	0.17616	0.50999	0.25909	0.03037	0.54700	0.01688	0.42649	0.01381	0.02092
HUMAN7	0.18269	0.28451	0.55712	0.31964	1.00000	0.17989	0.41388	0.21197	0.35010	0.02542	0.23240	-0.09625	0.04963	-0.13557	0.04279
COGNITIVO5	0.44050	0.13488	0.10116	0.24050	0.17989	1.00000	0.08175	0.20368	0.08541	0.26259	0.12747	0.35809	0.13643	0.08014	0.23916
HUMAN8	0.12943	0.11954	0.36763	0.17616	0.41388	0.08175	1.00000	0.14419	0.28170	0.01050	0.20237	-0.05547	0.05239	-0.23263	0.03583
MATERIA8	0.11391	0.43238	0.28781	0.50999	0.21197	0.20368	0.14419	1.00000	0.30431	-0.10186	0.37520	0.17083	0.50691	-0.06621	0.10740
HUMAN9	0.15731	0.28908	0.38607	0.25909	0.35010	0.08541	0.28170	0.30431	1.00000	-0.03325	0.20262	0.02503	0.12424	-0.06335	0.07463
COGNITIVO6	0.50533	-0.00487	-0.01819	0.03037	0.02542	0.26259	0.01050	-0.10186	-0.03325	1.00000	0.19637	0.19552	-0.26215	0.05466	0.18505
MATERIA9	0.23085	0.44216	0.26800	0.54700	0.23240	0.12747	0.20237	0.37520	0.20262	0.19637	1.00000	0.05652	0.20101	-0.05012	-0.10737
COGEMOC1	0.21520	-0.02348	-0.08656	0.01688	-0.09625	0.35809	-0.05547	0.17083	0.02503	0.19552	0.05652	1.00000	0.16362	0.18189	0.31460
MATERIA7	-0.01334	0.34017	0.09590	0.42649	0.04963	0.13643	0.05239	0.50691	0.12424	-0.26215	0.20101	0.16362	1.00000	0.02467	0.12340
COGEMOC3	0.07979	0.03610	-0.03561	0.01381	-0.13557	0.08014	-0.23263	-0.06621	-0.06335	0.05466	-0.05012	0.18189	0.02467	1.00000	0.06943
COGEMOC2	0.22095	0.02765	-0.02399	0.02092	0.04279	0.23916	0.03583	0.10740	0.07463	0.18505	-0.10737	0.31460	0.12340	0.06943	1.00000

Fuente: Elaborado por los autores.

La tabla 4 representa matriz de correlaciones policórica, donde indique correlaciones bajas entre las variables, la mayoría de las correlaciones entre variables son inferiores de 64%. Debido a que existen correlaciones se puede realizar un análisis factorial.

Ahora realizaremos con la función cortest bartlett, la prueba Test de esfericidad de Bartlett.

Tabla 5. Valores del Test de Bartlett

> n = nrow(test)	→	Tamaño de la muestra
> cortest.bartlett(R,n)		
\$chisq	\$p.value	\$df
[1] 6779.59	[1] 0	[1] 351

Fuente: Elaborado por los autores

El test de esfericidad de Bartlett permite nos permite evaluar la hipótesis nula que asegura que las variables no están correlacionadas, para lo cual debemos comparar la matriz de

intercorrelación de los datos conseguidos con una matriz de identidad en la que todos los términos de la diagonal son unidades y los demás términos con valores numéricos ceros. Si los resultados conseguidos de la comparación resultan significativos a un nivel $p < .05$, se rechaza la hipótesis nula y se considera que las variables están intercorrelacionadas para realizar el AFE (Everitt y Wykes, 2001).

Con los principios estadísticos mencionados y observando los resultados de la tabla 5, donde la prueba de esfericidad con el Test de Bartlett que se aplicó, nos informa que el *p.value* es muy pequeño, el software imprime un valor “casi cero” señalando prácticamente que la probabilidad es muy pequeña. Entonces por ser inferior a “0.05”, debemos rechazar la “Ho” la hipótesis nula, debido a que el determinante no es 1 por que su p-valor que es muy bajo, debemos aceptar la “H1” la hipótesis alternativa.

La matriz de correlación es distinta de la matriz identidad, señalando que por encima y por debajo de los “1” “diagonal principal” las correlaciones son distintas de cero, por lo tanto indica que se pueden agrupar en factores.

Si podemos aplicar el siguiente paso de análisis factorial, porque existen variables que se correlacionan de manera moderada se pueden agrupar en factores, existe poca multicolinealidad, o sea una relación de dependencia lineal entre más de dos variables.

A continuación realizamos la valoración de KMO.

Tabla 6. Valor grupal de los ítems del indicar KMO.

```

> KMO(Kaiser-Meyer-Olkin)
> KMO(R)
Kaiser-Meyer-Olkin factor adequacy
Call: KMO(r = R)
Overall MSA = 0.87
    
```

Fuente: Elaborado por los autores

Tabla 7. Valores individuales por ítem del KMO

```

> #KMO(kaiser-meyer-olkin)
> # -----
> KMO(R)
Kaiser-Meyer-Olkin factor adequacy
Call: KMO(r = R)
Overall MSA = 0.87
MSA for each item =
  HUMAN1 COGNITIVO1  MATERIA1  MATERIA2  HUMAN2  HUMAN3 COGNITIVO2  MATERIA3  HUMAN4 COGNITIVO3  MATERIA4  HUMAN5 CO
GNITIVO4
  0.88    0.81    0.90    0.90    0.90    0.86    0.87    0.90    0.85    0.86    0.88    0.86
  0.87
  MATERIA5  HUMAN6  MATERIA6  HUMAN7 COGNITIVO5  HUMAN8  MATERIA8  HUMAN9 COGNITIVO6  MATERIA9  COGEMOC1  MATERIA7
  0.86    0.82    0.92    0.88    0.85    0.92    0.92    0.89    0.77    0.87    0.76    0.88
  0.73
  COGEMOC2
  0.69
    
```

Fuente: Elaborado por los autores.

Mediante la prueba del estadístico Káiser-Meyer-Olkin KMO, índice en el cual indica un promedio de los términos de la diagonal de la matriz de correlaciones, la cual contiene los valores negativos de los coeficientes de correlación parcial de las variables.

La lógica que determina según (Sierra Bravo, 1981) del índice KMO es que las variables logran formar parte de otros factores comunes, es porque entonces los coeficientes de correlación parcial deberían ser pequeños, y por lo tanto los valores de la diagonal de la matriz deben ser elevados, indicando que si es elevados la proporción de coeficientes grandes en la matriz existe una mayor interrelación entre las variables de estudio.

El estadístico Káiser-Meyer-Olkin "KMO" nos ayuda para evaluar la adecuación del conjunto de datos de la muestra al análisis factorial, para este caso la tabla 6 indica un valor de 0.87, valor que es superior a 0.5.

Según Káiser el valor 0.87 es considerado bueno, ya que lo ha definido dentro del rango de (0.8 a 0.9), porque se encuentra cercano a 1 o sea es meritorio. Podemos observar en la tabla 7 los valores de medidas de adecuación de la muestra por ítem de todas las variables de la data test, donde se evidencia valores que son superiores a 0.73 %, excepto la variable "COGEMOC2" que tiene un valor de 0.69 pero se acerca a 0.7 donde se considera intermedio.

El valor de KMO 0.87 en este ejercicio, es superior a 0.5 y además se considera bueno, entonces se procede a la aplicación del análisis factorial al conjunto de datos.

Una vez realizado los 3 criterios (*Correlación policórica, la prueba de esfericidad de Bartlett y la prueba KMO*) se determinó que si es posible realizar un AFE.

Método de extracción y determinación del número de factores

Vamos a proceder con el gráfico de sedimentación con el software r-studio, para determinar en número que dimensiones o factores en el lugar de estudio.

Para ello necesitamos instalar el paquete parallel; una vez instalado el paquete y ejecutado con la función library, esta función ya quede habilitada y pueda operar en el software.

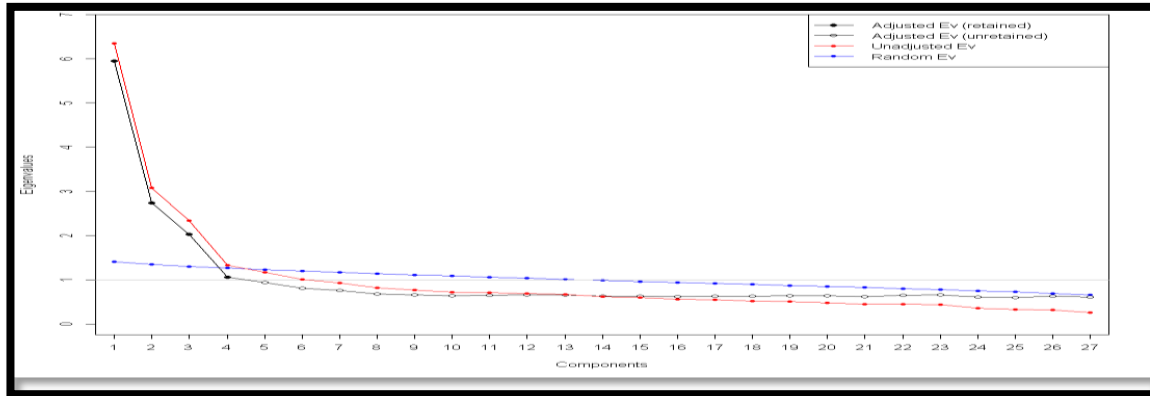
El análisis paralelo, se lo conoce también como análisis paralelo de Horn, este nombre hace referencia al psicólogo (John L. Horn, 1965), quien fue creador de este método. Es un método estadístico utilizado para determinar el número de componentes a mantener en un análisis de factores principales o llamado también factores retenidos. En un análisis factorial exploratorio, el cual se lo aplicará con la data llamada test, con 5000 interacciones, mostrando gráfica, el ancho y alto de 100, tal como indica la tabla 8.

Tabla 8. Componentes para el Gráfico de Sedimentación.

```
install.packages("paran")
library(paran)
>paran(test, iterations=5000,graph=TRUE,color=TRUE, width=100, height=100)
```

Fuente: Elaborado por los autores.

Tabla 9. Gráfico de Sedimentación.



Fuente: Elaborado por los autores.

En el gráfico de sedimentación de la tabla 9, en el eje vertical u ordenada se representan los autovalores, y en el eje horizontal o abscisa, este representa al número de factores.

En la gráfica resultante se traza horizontalmente una línea recta base a la altura de los últimos autovalores, aquellos que se encuentran por encima de esa línea base, indicarán el número de factores que se tiene que retener.

En este caso en demostración se observa la gráfica de sedimentación, en la línea negra que representa “Adjusted Eigenvalue retaine” que son los valores propios ajustados retenidos, se observa los puntos de la línea negra que están superior al 1 en el eje de la ordenadas “Eigenvalues” indicándonos que son 4 factores para este caso en estudio.

Tabla 10. Resultados de los valores de los Eigenvalue Adjusted, Unadjusted y Estimated bias.

Using eigendecomposition of correlation matrix.			
Computing: 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%			
Results of Horn's Parallel Analysis for component retention			
5000 iterations, using the mean estimate			

Component	Adjusted	Unadjusted	Estimated
	Eigenvalue	Eigenvalue	Bias

1	5.947454	6.352988	0.405533
2	2.736089	3.083066	0.346977
3	2.033837	2.335860	0.302023
4	1.062517	1.326768	0.264251
Adjusted eigenvalues > 1 indicate dimensions to retain.			
(4 components retained)			

Fuente: Elaborado por los autores.

En la tabla 10 observamos los valores correspondientes al gráfico de sedimentación de la tabla 9 que indica que son 4 los componentes de retenidos. Los valores impresos son pertenecientes a la gráfica de sedimentación, indicando a que altura se ubica cada punto de la línea negra que representa a los Adjusted Eigenvalue retaine, la línea de color rojo en cambio identifica a los Adjusted Eigenvalue unretaine y por último la línea azul que representa a los Random Eigenvalue.

A continuación procedemos a crear un factorial2 con la función (fa.poly), de 4 factores, con (correlación policórica) esto debido a que estamos trabajando con una data que tiene escala de Likert y con el método de rotación VARIMAX, método que nos va a permitir maximizar la suma de varianzas de las cargas factoriales dentro de cada factor que es lo que busca este método AFC, obsérvese más adelante la tabla 11 de funciones.

Además el método varimax o llamado también Método de rotación ortogonal que logra minimizar el número de variables que tienen cargas elevadas en cada factor, ayuda a Simplificar la interpretación de los factores en el trabajo de estudio.

Tabla 11. Funciones para obtener valores de las cargas factoriales por ítem.

```
factorial2=fa.poly(test, nfactors=4, cor="poly",rotate="varimax")
factorial2
Factor Analysis using method = minres
Call: fa.poly(x = test, nfactors = 4, rotate = "varimax", cor = "poly")
Standardized loadings (pattern matrix) based upon correlation matrix
```

Fuente: Elaborado por los autores.

Tabla 12. Valores de las pesos de las cargas factoriales por ítem.

	MR1	MR3	MR2	MR4	h2	u2	com
HUMAN1	0.26	0.41	0.07	0.01	0.247	0.75	1.8
COGNITIVO1	-0.09	0.10	0.75	0.08	0.579	0.42	1.1
MATERIA1	0.78	0.26	0.03	-0.03	0.674	0.33	1.2
MATERIA2	0.53	0.02	-0.07	0.33	0.402	0.60	1.7
HUMAN2	0.53	0.57	-0.01	0.09	0.610	0.39	2.0
HUMAN3	0.07	0.45	0.19	-0.13	0.265	0.74	1.6
COGNITIVO2	0.22	0.08	0.59	0.14	0.422	0.58	1.4
MATERIA3	0.70	0.27	0.08	-0.03	0.571	0.43	1.3
HUMAN4	0.02	0.55	0.06	-0.05	0.311	0.69	1.0
COGNITIVO3	0.13	0.05	0.63	0.10	0.427	0.57	1.2
MATERIA4	0.66	0.13	0.15	-0.01	0.480	0.52	1.2
HUMAN5	0.21	0.69	0.08	-0.03	0.535	0.47	1.2
COGNITIVO4	0.09	0.17	0.69	-0.01	0.518	0.48	1.2

MATERIA5	0.68	0.17	0.07	-0.06	0.495	0.51	1.2
HUMAN6	0.32	0.63	0.03	-0.05	0.503	0.50	1.5
MATERIA6	0.76	0.21	0.16	-0.07	0.657	0.34	1.3
HUMAN7	0.17	0.66	0.07	0.01	0.472	0.53	1.2
COGNITIVO5	0.16	0.05	0.57	0.23	0.412	0.59	1.5
HUMAN8	0.08	0.57	0.02	0.02	0.328	0.67	1.0
MATERIA8	0.64	0.18	0.02	0.25	0.502	0.50	1.5
HUMAN9	0.18	0.52	0.05	0.14	0.330	0.67	1.4
COGNITIVO6	-0.11	-0.01	0.67	-0.20	0.493	0.51	1.2
MATERIA9	0.53	0.22	0.22	-0.24	0.437	0.56	2.2
COGEMOC1	0.07	-0.15	0.41	0.45	0.398	0.60	2.3
MATERIA7	0.62	-0.03	-0.11	0.39	0.544	0.46	1.8
COGEMOC3	0.00	-0.22	0.14	0.09	0.077	0.92	2.0
COGEMOC2	-0.02	-0.01	0.29	0.46	0.292	0.71	1.7

Fuente: Elaborado por los autores

Los factores principales se denotan la tabla 12 con los nombres MR1-MR2-MR3-MR4 y los pesos de las cargas factoriales de cada ítem.

Para determinar a qué factor pertenece cada ítem, hay que observar la tabla 12. Por tomar el caso del ítem HUMAN1, donde indica que el mayor valor es 0.41 que está ubicado justo en la comuna del factor MR3, indicando así que su mayor carga la tiene en ese factor. También comunalidades señaladas con las siglas h², la estimaciones de los factores únicos sin mucha injerencia denotada con U² y además aparece todas las cargar factoriales de los ítems.

Tabla 13. Valores de varianzas por factor.

	MR1	MR3	MR2	MR4
SS loadings	4.62	3.36	3.01	0.99
Proportion Var	0.17	0.12	0.11	0.04
Cumulative Var	0.17	0.30	0.41	0.44
Proportion Explained	0.39	0.28	0.25	0.08
Cumulative Proportion	0.39	0.67	0.92	1.00

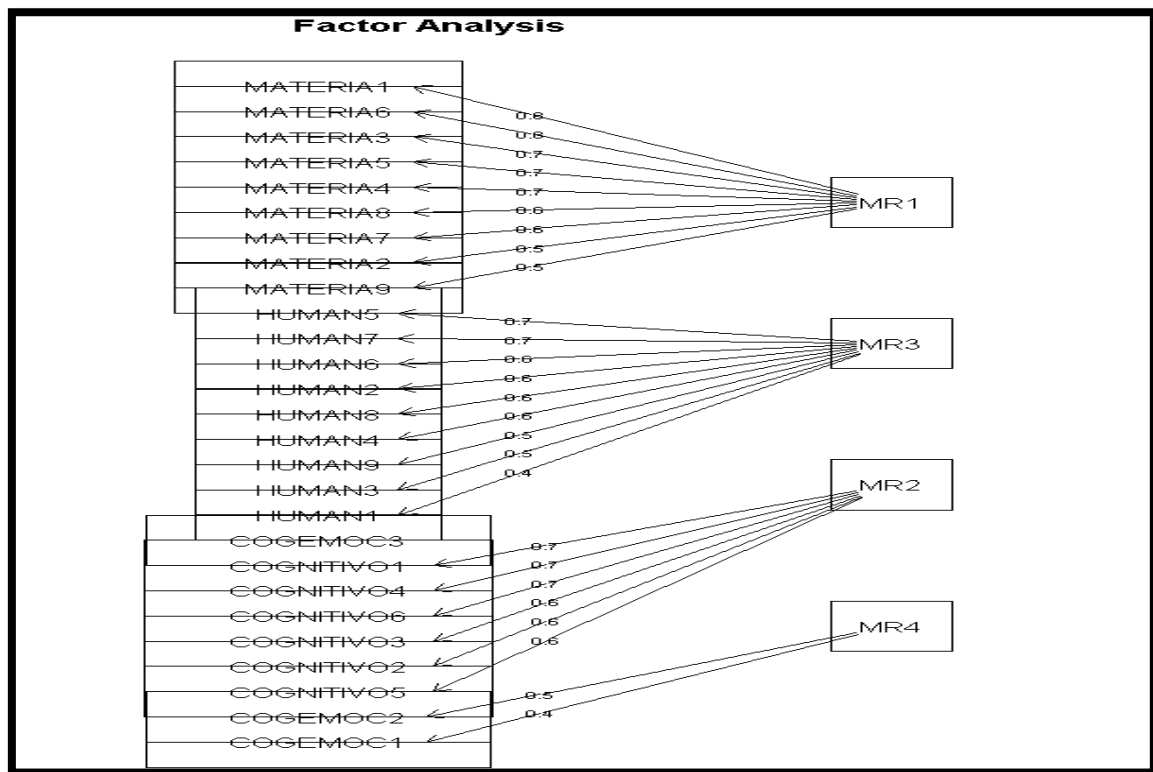
Fuente: Elaborado por los autores

El porcentaje de la varianza acumulada que representan a los 4 factores corresponde a 44%, como se observa en la tabla 13, el peso que aporta la varianza individual de cada factor es de 17 % que corresponde al factor MR1, el factor MR3 12%, el factor MR2 11 % y por último el factor MR4 aporta a la varianza acumulada con un valor de 4 %.

La SS loadings “sumatoria de las cargas factoriales” es la suma de todos los valores elevados al cuadrado del factor MR1 es de 4,62, del factor MR3 es de 3,36, del factor MR2 es de 3,01 y el último factor MR4 es de 0,99.

También se muestra la proporción acumulada del factor MR1 es de 0,39, del factor MR3 es de 0,67, del factor MR3 es de 0,92 y el último factor MR4 es de 1.

Tabla 14. Gráfico de distribución de ítems por cada factor.



Fuente: Elaborado por los autores

Aquí podemos observar la gráfica de análisis de factores de la tabla 14 como cada ítem se agrupan a cada Factor, con su respectivo valor del peso de carga factorial, correspondiente que se indicó en el cuadro anterior.

- En el factor rm1 se agrupan (9) los ítems materia 1, materia 2, materia 3, materia 4, materia 5, materia 6, materia 7, materia 8, materia 9.
- En el factor rm3 (9) human1, human2, human3, human4, human5, human6, human7, human8, human9.
- En el factor rm2 (6) se agrupan cognitivo1, cognitivo2, cognitivo3, cognitivo4, cognitivo5, cognitivo6.
- En el factor rm4 (2) se agrupan cogemoc1 y cogemoc2.

Ahora precedemos a crear el modelo de los 4 factores con sus respectivos ítems que nos indicó el gráfico de sedimentación, con la ayuda del paquete lavan.

Tabla 15. Creación del modelo con 4 factores con cada ítem.

```

> library(lavaan)
> # Creando el modelo con 4 factores
> modelo <- 'MR1 = ~ MATERIA1+MATERIA6+MATERIA3+MATERIA5+MATERIA4+MATERIA8+MATERIA7+MATERIA9+MATERIA2
+ MR3 = ~ HUMAN5+HUMAN7+HUMAN8+HUMAN6+HUMAN4+HUMAN9+HUMAN2+HUMAN3+HUMAN1
+ RM2= ~ COGNITIVO1+COGNITIVO4+COGNITIVO6+COGNITIVO3+COGNITIVO2+COGNITIVO5
+ MR4= ~ COGEMOC2+COGEMOC1
+ '
> modelo
[1] "RC1 = ~ MATERIA1+MATERIA6+MATERIA3+MATERIA5+MATERIA4+MATERIA8+MATERIA7+MATERIA9+MATERIA2\nRC3 = ~ HUMAN5+HUMAN7+HUMAN8+HUMAN6+HUMAN4+HUMAN9+HUMAN2+HUMAN3+HUMAN1\nRC2= ~ COGNITIVO1+COGNITIVO4+COGNITIVO6+COGNITIVO3+COGNITIVO2+COGNITIVO5\nRC4= ~ COGEMOC2+COGEMOC1\n"
    
```

Fuente: Elaborado por los autores

Una vez creado el modelo en la tabla 15, procedemos a ver la validez del mismo con las métricas Chi Cuadrado, SRMR, RMSEA, TLI, CFI, GFI, AGFI de este trabajo en desarrollo, para esto aplicamos la función que se denota `fit <- cfa(modelom, data = test, ordered = TRUE)` `> summary(fit, fit.measures = TRUE)`

Tabla 16. Creación de validez y summary de las métricas del modelo con 4 factores.

npar	fmin	chisq	df
136.000	1.220	1488.773	293.000
pvalue	chisq.scaled	df.scaled	pvalue.scaled
0.000	1421.488	293.000	0.000
chisq.scaling.factor	baseline.chisq	baseline.df	baseline.pvalue
1.151	27210.493	325.000	0.000
baseline.chisq.scaled	baseline.df.scaled	baseline.pvalue.scaled	baseline.chisq.scaling.factor
9983.093	325.000	0.000	2.784
cfi	tli	nnfi	rfi
0.956	0.951	0.951	0.939
nfi	pnfi	ifi	rni
0.945	0.852	0.956	0.956
cfi.scaled	tli.scaled	cfi.robust	tli.robust
0.883	0.870	NA	NA
nnfi.scaled	nnfi.robust	rfi.scaled	nfi.scaled
0.870	NA	0.842	0.858
ifi.scaled	rni.scaled	rni.robust	rmsea
0.884	0.883	NA	0.082
rmsea.ci.lower	rmsea.ci.upper	rmsea.pvalue	rmsea.scaled
0.078	0.086	0.000	0.080
rmsea.ci.lower.scaled	rmsea.ci.upper.scaled	rmsea.pvalue.scaled	rmsea.robust
0.075	0.084	0.000	NA
rmsea.ci.lower.robust	rmsea.ci.upper.robust	rmsea.pvalue.robust	rmr
NA	NA	NA	0.072
rmr_nomean	srmr	srmr_bentler	srmr_bentler_nomean
0.074	0.074	0.072	0.074
crmr	crmr_nomean	srmr_plus	srmr_plus_nomean
0.074	0.077	NA	NA
cn_05	cn_01	gfi	agfi
137.595	145.087	0.970	0.956
pgfi	mfi		
0.662	0.375		

Fuente: Elaborado por los autores

A continuación analizaremos los valores de métricas que se obtuvo en la tabla 16.

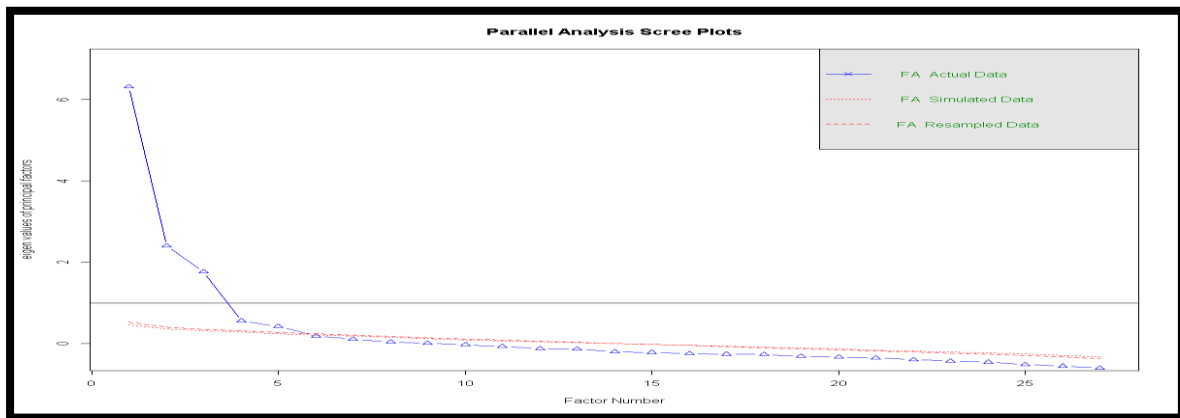
➤ Chi Cuadrado nos indica un valor de 1488.773.

- RMSEA, este es un indicador de ajuste absoluto, el valor a considerar para un buen ajuste es que tiene que ser inferior a "0.05" y aceptable $<0.05, 0.08 >$, en este caso ejercicio nos tiene un valor superior de "0.08" señalando que puede ser aceptable modelo.
- SRMR, otro indicador de los denominados corregidos parsimonia, el valor siempre tiene que ser inferior a "0.08" para que el modelo tenga un buen ajuste, como en este caso que nos indica un valor de 0.074 que es inferior a "0.08" dando señales que es un buen modelo.
- CFI, mantiene el principio de los indicadores incrementales con el NFI, aquí se considera un ajuste bueno aceptable cuando el valor $<0.90, 0.95 >$ y bueno cuando sea mayor a 0.95, en este modelo de ejercicio nos dio un valor de 0.956 , indicando que es bueno.
- TLI, este indicador mide el ajuste comparativo o incremental, aquí se considera un buen ajuste cuando el valor es mayor a 0.90, en este ejercicio desarrollado da un valor de 0.951 , o sea es un buen ajuste.
- GFI, este índice de bondad de ajuste dio en este ejercicio un valor de 0.970 , se considera un ajuste casi perfecto ya que este valor está cercano a (1) lo cual lo hace ver como un ajuste perfecto.
- AGFI, es el índice ajustado de bondad de ajuste, "extensión del GFI" aquí tenemos en este caso un valor de 0.956 , también es buen ajuste, porque lo recomendado en este indicador es que el valor sea mayor o igual a 0.90.

Análisis Factorial exploratorio

Todas las métricas de la tabla 16, nos han indicado una validez buena respecto al modelo, pero la varianza acumulada fue 44%. Procedemos al Análisis Factorial Exploratorio, ya que obtuvimos un valor bajo en la varianza acumulada con 4 factores, utilizando función `fap=fa.parallel(test,fa="fa",cor="poly")`, para determinar el número de factores.

Tabla 17. Gráfico Parallel Analysis Scree Plots



Fuente: Elaborado por los autores

Observando y valorando el gráfico de sedimentación de la tabla 17, nos indica que factores podemos conservar cuyo número este situado antes del punto de inflexión. A partir de este criterio se concluye que se tendría para *este caso 3 factores*.

En el gráfico de sedimentación Adjusted EV “Eigenvalue” que se dibuja con la línea azul que representa a los factores, nos indica que los autovalores ajustados son 3.

Los triángulos que se encuentran en la línea azul, hasta la altura del 1, son los factores principales que el r-studio a determinado mediante esta gráfica.

Tabla 18. Valores para la Gráfica de los Eigen values.

```
> #Identificando el N° de factores
> # -----
> fap=fa.parallel(test,fa="fa",cor="poly")
Parallel analysis suggests that the number of factors = 5 and the number of components = NA
> fap
Call: fa.parallel(x = test, fa = "fa", cor = "poly")
Parallel analysis suggests that the number of factors = 5 and the number of components = NA

Eigen Values of

eigen values of factors
[1] 6.31 2.40 1.77 0.56 0.42 0.19 0.11 0.04 0.01 -0.03 -0.07 -0.12 -0.13 -0.20 -0.22 -0.24 -0.26 -0.27 -0.31 -0.33 -0.35 -0.40 -0.43
[25] -0.52 -0.55 -0.61

eigen values of simulated factors
[1] 0.45 0.36 0.31 0.28 0.25 0.22 0.19 0.15 0.13 0.10 0.08 0.05 0.03 0.00 -0.02 -0.04 -0.06 -0.09 -0.11 -0.14 -0.16 -0.19 -0.21
[25] -0.26 -0.29 -0.33

eigen values of components
[1] 6.95 3.31 2.53 1.37 1.17 0.99 0.93 0.81 0.74 0.69 0.68 0.64 0.63 0.57 0.56 0.50 0.49 0.47 0.44 0.42 0.38 0.38 0.37 0.29 0.25 0.25 0.19

eigen values of simulated components
[1] NA
```

Fuente: Elaborado por los autores

También se puede leer en la descripción que el análisis paralelo sugiere 5 el número de factor, obsérvese la tabla 18. Analizaremos rápidamente solo el porcentaje de varianza

explicado con 3 y 5 factores. Con el fin de determinar cuál de los constructos tiene un valor más elevado en la varianza acumulada.

Tabla 19. *Análisis de valores de varianzas con 3 factores*

```
> factorial2=fa.poly(test, nfactors=3, cor="poly",rotate="varimax")
> factorial2
          MR1 MR3 MR2
SS loadings      4.55 3.53 3.03
Proportion Var   0.17 0.13 0.11
Cumulative Var   0.17 0.30 0.41
Proportion Explained 0.41 0.32 0.27
Cumulative Proportion 0.41 0.73 1.00
```

Fuente: Elaborado por los autores

El porcentaje de la varianza acumulada que representan a los 3 factores es de 41 %, como se indica observando la tabla 19 el peso que aporta la varianza individual de cada factor es de 17 % que corresponde al factor MR1, el factor MR3 13 %, y el factor MR2 11 %.

Tabla 20. *Análisis de valores de varianzas con 5 factores*

```
> factorial2=fa.poly(test, nfactors=5, cor="poly",rotate="varimax")
> factorial2
          MR1 MR3 MR2 MR4 MR5
SS loadings      4.68 3.32 3.01 1.00 0.69
Proportion Var   0.17 0.12 0.11 0.04 0.03
Cumulative Var   0.17 0.30 0.41 0.45 0.47
Proportion Explained 0.37 0.26 0.24 0.08 0.05
Cumulative Proportion 0.37 0.63 0.87 0.95 1.00
```

Fuente: Elaborado por los autores

El porcentaje de la varianza acumulada que representan a los 5 factores es de 47 %, el peso que aporta la varianza individual de cada factor es de 17 % que corresponde al factor MR1, el factor MR3, 12 %, el factor MR2, 11 %, el factor MR4, 4 % y el factor MR5 un 3 %, valores porcentuales que se observó en la tabla 20.

Las varianzas acumuladas de los 2 modelos realizados con 3 y 5 factores indicaron valores inferiores a 60 %, se procede a continuación a eliminar los ítems que tengan una correlación ítem subtest (ítem factor) menor a 65 %, para poder encontrar el mejor modelo factorial mediante el AFE.

Para realizar esta búsqueda de las correlaciones ítem subtest, procederemos a agrupar los factores con sus respectivos ítems, para luego identificar los valores que estén inferiores a 65 %.

Tabla 21. Creación de factores y validación de Subtest-Test, o Factor-Test

```

#Factor 1
M1=test[,3];M6=test[,16];M3=test[,8];M5=test[,14];M4=test[,11];M8=test[,20];M7=test[,25];M
2=test[,4];M9=test[,23]
fact1=data.frame(M1,M2,M3,M4,M5,M6,M7,M8,M9)
fact1

#Factor 3
H5=test[,12];H7=test[,17];H6=test[,15];H2=test[,5];H8=test[,19];H4=test[,9];H9=test[,21];H3=t
est[,6];H1=test[,1]
fact3=data.frame(H1,H2,H3,H4,H5,H6,H7,H8,H9)
fact3

#Factor 2
C1=test[,2];C4=test[,13];C6=test[,22];C3=test[,10];C2=test[,7];C5=test[,18]
fact2=data.frame(C1,C2,C3,C4,C5,C6)
fact2

#Factor 4
CG2=test[,27];CG1=test[,24]
fact4=data.frame(CG2,CG1)
fact4

Validación Subtest-Test, o Factor-Test

> subtest1=rowSums(fact1)
> subtest2=rowSums(fact2)
> subtest3=rowSums(fact3)
> subtest4=rowSums(fact4)
> Sum.total=rowSums(test)
> subtest=data.frame(subtest1,subtest2,subtest3,subtest4)
> Subtest.test=cor(subtest,Sum.total)
> Subtest.test

[,1]
subtest1 0.8172652
subtest2 0.5556505
subtest3 0.7632489

```

subtest4 0.2938496

Fuente: Elaborado por los autores

En la validación de la tabla 21 donde el subteste. test o llamada también Factor-test, se evidencia que existe una baja correlación en el “factor y el test” del subtest2 y también del subtest3 que son inferiores a 0.60. En este indicador tiene que ser alto porque se mide el subtest respecto al “test general” pero en este caso no lo es.

Tabla 22. Validación item-subtest o item-factor

```

> item.subtest1=cor(subtest1,fact1)
> item.subtest1
      M1      M2      M3      M4      M5      M6      M7      M8      M9
[1,] 0.7784453 0.5929772 0.7316904 0.6898506 0.7081716 0.7782545 0.6202129 0.698842 0.5985382
> item.subtest2=cor(subtest2,fact2)
> item.subtest2
      C1      C2      C3      C4      C5      C6
[1,] 0.72795 0.6897825 0.7084406 0.7388601 0.6793456 0.6506587
> item.subtest3=cor(subtest3,fact3)
> item.subtest3
      H1      H2      H3      H4      H5      H6      H7      H8      H9
[1,] 0.5473282 0.6899352 0.5429947 0.5877526 0.7258917 0.6983558 0.6851765 0.595622 0.5930344
> item.subtest4=cor(subtest4,fact4)
> item.subtest4
      CG2      CG1
[1,] 0.804215 0.7996168
    
```

Fuente: Elaborado por los autores

En el análisis de la validación de la tabla 22 de item-subtest o item-factor, varios ítems muestran una baja correlación respecto a sus factores, de lo cual se identificó 8 que se encuentran con valores inferiores a 0,65.

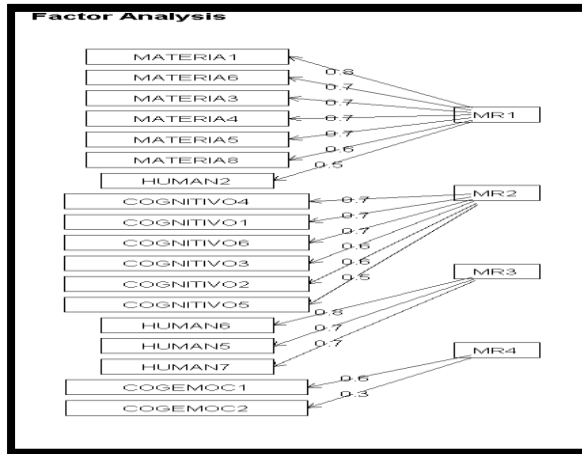
Las variables a eliminar para la nueva base de datos serán las variables que tiene un valor inferior a 0,65 o sea presentan correlación baja respecto a su factor como son en este caso en desarrollo “M2-M7-M9” baja correlación con su factor MR1 y “H1-H3-H4-H8-H9” que también tiene correlación con el factor MR3. Además observando la tabla 7 se procedió a eliminar también la variable COGEMOC2, que presentó un valor de 0.69 cuando se valoró mediante la prueba del KMO.

Una vez eliminados los ítems indicados, se procederá a la búsqueda AFE para encontrar un mejor modelo factorial.

Nuevo Modelo Factorial

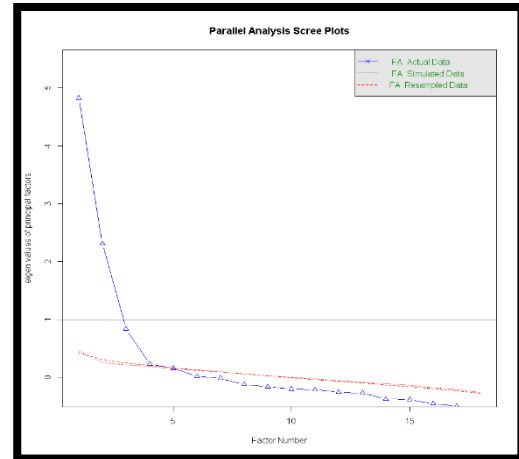
Observando la tabla 22, se procedió con la eliminación de 8 ítems con valores inferiores 0.65, y más la eliminación del ítem COGEMOC2, se conservó 18 ítems para la base de datos. Se inicia un AFE con función `fap=fa.parallel(test,fa="fa",cor="poly")`.

Tabla 23. Gráfico de distribución de ítems por cada factor.



Fuente: Elaborado por los autores

Tabla 24. Gráfico Parallel Analysis Scree Plots



Fuente: Elaborado por los autores

El gráfico de sedimentación de la tabla 24, recomendó bajo el AFE 3 factores obteniendo un valor en la varianza acumulada de 48 %, además también se realizó el trabajo con 4 factores, obteniendo una varianza acumulada de 51 %, valores que se indica en la tabla 24.

Tabla 24. Valores de varianzas con 3 factores y 4 factores.

3 FACTORES	MR1	MR2	MR3	4 FACTORES	MR1	MR2	MR3	MR4
SS loadings	3.70	2.92	2.04	SS loadings	3.63	2.67	2.07	0.85
Proportion Var	0.21	0.16	0.11	Proportion Var	0.20	0.15	0.12	0.05
Cumulative Var	0.21	0.37	0.48	Cumulative Var	0.20	0.35	0.47	0.51
Proportion Explained	0.43	0.34	0.24	Proportion Explained	0.39	0.29	0.22	0.09
Cumulative Proportion	0.43	0.76	1.00	Cumulative Proportion	0.39	0.68	0.91	1.00

Fuente: Elaborado por los autores

Tabla 25. Valores de los coeficiente con 4 factores principales

$F1 = 0.8197931M1 + 0.7631348M3 + 0.7144546M4 + 0.7158375M5 + 0.7725256M6 + 0.6980171M8 + 0.06966073H2$
$F2 = 0.72795C1 + 0.6897825C2 + 0.7084406C3 + 0.7388601C4 + 0.6793456C5 + 0.6506587C6$
$F3 = 0.8174046H5 + 0.7831908H7 + 0.8492573H6$
$F4 = 0.804215CG2 + 0.7996168CG1$

Fuente: Elaborado por los autores

Una vez analizado otros modelos, se determinó observando la tabla 24 que el mejor modelo fue de (4) *factores*, debido a que se obtuvo una varianza acumulada del 51% con los coeficientes que se muestran en la siguiente tabla 25.

RESULTADOS

En el desarrollo del trabajo que fue aplicar un Análisis Factorial exploratorio con los instrumentos citados para cada uno de los análisis propuestos mediante las medidas de adecuación muestral KMO, test de esfericidad de Bartlett, porcentaje de varianza acumulada, validación de la correlación de matrices policórica, peso de cargas factoriales por ítem, método de rotación varimax, gráfico de sedimentación para la determinación de componentes de factores principales, también con el análisis de los indicadores, lo mismo con las métricas Chi Cuadrado, SRMR, RMSEA, TLI, CFI, GFI, AGFI, validación item-subtest, factor-tes. Se pudo determinar mediante el AFE, que el mejor encontrado fue el que presentaba una varianza acumulada de 51%, el cual se compone de 4 factores principales, donde el mayor peso en cuanto a porcentaje de la varianza se encuentra ubicado en el primer factor MR1 con un 20 %, el segundo MR2, tercero MR3 y el cuarto factor MR4, que obtuvieron valores porcentuales de 15 %,12 % y 5 % respectivamente. Para obtener ese mayor valor de la varianza acumulada de los 4 factores principales, se consideró eliminar a 9 ítems que tenían una baja correlación respecto a su factor de la base de datos que en su principio presentaba 27 ítems, teniendo así una nueva base de datos con 18 ítems, esto se determinó mediante el del AFE.

Se indica que en su mayoría de variables excepto 1, se agrupaban con nombres similares al mismo factor, esto debido por la carga del peso factorial que representaba individualmente. Análisis factorial, se puede aplicar en cualquier campo, cuando el estudio lo que requiere es reducir factores o identificar los factores principales de cada variable.

CONCLUSIONES

En el AFE estos indicadores han demostrado mediante el trabajo desarrollado que se puede aplicar en cualquier campo, cuando el estudio lo que requiere es reducir variables y poder identificar los factores principales, esto sirve de gran ayuda para facilitar el trabajo del investigador cuando tiene un número cantidad de variables. El AFE se puede aplicar a datos que presenten escale de Likert, dicotómica o de tasa y razón.

El investigador dentro de su área de conocimiento es quien decide si desea eliminar o conservar ítems que representen una baja correlación con el factor, esto debido a que el especialista puede considerar que dentro de su investigación ese ítem representa suma importancia, respetando el criterio del especialista el deberá determinar bajo criterio si elimina el ítem o reformula dando otro enfoque a la pregunta.

REFERENCIAS BIBLIOGRÁFICAS.

- Abridged ACRA scale of learning strategies for Escala de estrategias de aprendizaje ACRA-Abreviada para alumnos universitarios Jesús de la Fuente Arias *, Fernando Justicia Justicia . (2003), (January)
- Alonso, A. (2008). Técnicas de análisis multivariante de datos. (I. Capella & M. Diego, Eds.) (Primera). Madrid-España: PEARSON PRENTICE HALL.
- Cuadras, C. M. (2014). Nuevos Métodos de Análisis Multivariante. (CMC Editions, Ed.) CMC Editions (Primera). Barcelona - España: CMC Editions. <http://doi.org/10.1017/CBO9781107415324.004>
- Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*, 4(3), 272–299. <http://doi.org/10.1037/1082-UNEMI2017>
- Hair, J. F., Black, W. C., & Babin, B. J. (2010). *Multivariate Data Analysis: A Global Perspective*. Pearson Education.
- Hair, Anderson, R., Tatham, R., & Black, W. (2008). *Análisis Multivariante*. (A. Otero, Ed.) (Quinta). Madrid-España: Prentice Hall Iberia.
- Kaiser, H.F. (1958). The varimax criterion for analytic rotation in factor analysis. *Psychometrika*, 23, 187-200.
- Lloret-Segura, S., Ferreres-Traver, A., Hernández-Baeza, A., & Tomas Marco, I. (2014). El análisis factorial exploratorio de los ítems