

1. Introducción

La Encuesta Mundial de Valores (EMV) también conocida por sus siglas en inglés WVS (World Value Survey) se trata de una encuesta que actualmente incluye a casi 100 países o territorios. Esta encuesta tiene como objetivo principal explorar los valores, opiniones e incluso estereotipos sostenidos por las personas, así como también observar cómo estos se transforman con el paso del tiempo y eventos que tienen una influencia tanto en la sociedad como en sus sistemas de gestión política (World Values Survey [WVS], 2020).

Además de realizarse por rondas en las que cada país o territorio decide si incluirse o no realizando los trabajos concernientes al uso de muestreos probabilísticos que garanticen la representatividad de la población, comprende un arduo trabajo que busca incluir diseños de investigación científicos y rigurosos. Inició en el año 1981, consiste en encuestas representativas a nivel nacional y realiza mediante cuestionarios comunes que pueden ser de carácter presencial o no. Otra de sus características es que se realiza por rondas, lo que quiere decir que cada país se encuesta una vez por ronda y se compromete a realizar su trabajo de campo (World Values Survey [WVS], 2020).

Por su parte, Venezuela participó en la EMV en el período de 2017-2020 a través del Instituto Delphos (quien hizo el trabajo de campo) y para ello realizó un formato de encuesta en forma de tarjetas a ser aplicadas a los participantes. Estas tarjetas incluían preguntas o ítems respecto a un conjunto de áreas de opinión, valores e intereses, que iban desde la evaluación de la familia y la fe hasta los intereses y percepciones respecto a la política y el sistema de gobierno venezolano. Para ese entonces, recogieron un total de 1190 observaciones que en conjunto ayudan a comprender una complejidad de fenómenos insertos en las áreas individual, familiar, social, económica y política.

A continuación, se llevarán a cabo todos los estudios pertinentes a un Análisis Factorial Exploratorio sobre la sección última de la encuesta denominada “Interés por la Política”, con la finalidad de identificar variables latentes o de carácter no observable capaces de explicar el orden y estructura de los datos procesados, lo que permite comprender con un mayor nivel de

profundidad el comportamiento e interrelación de los intereses y opiniones políticas en la población venezolana.

2. ¿Qué variables incluir?

Las variables a incluir en la evaluación de *Interés sobre la Política* en la población venezolana, comprenden un total de 15 indicadores o items que hacen parte de la encuesta enfocados en evaluar cuáles son las percepciones de los ciudadanos respecto al comportamiento de los políticos, el nivel de confianza que depositan en ellos, las decisiones del gobierno y su forma de gestión (p. ej. “*No estoy seguro de creer en la mayoría de los políticos*” o “*En general, el gobierno suele hacer lo correcto*”). En principio, se parte del hecho de que estas variables que comprenden cada uno de los indicadores que hacen parte del constructo de *Interés sobre la Política* han debido ser cuidadosamente revisados por expertos en el área política para corroborar u obtener evidencia de la validez de contenido de la encuesta (Watkins, 2018).

La validez de contenido hace referencia a la calidad en que los indicadores o variables muestreados cubren adecuadamente el área de comportamiento o universo a ser evaluados. Es decir, los indicadores o variables deben ser tanto representativos como relevantes del constructo denominado *Interés sobre la Política*. Por su parte, la representatividad de los indicadores se logra analizando sistemáticamente el constructo para garantizar que los items cubren todos los aspectos importantes que forman parte de él y en la proporción correcta (esto especialmente importante al trabajar con constructos heterogeneos), mientras que la relevancia de los indicadores se logra incluyendo todas las conductas o items que verdaderamente estén cubiertos o se encuentren dentro del dominio evaluado. Si se trata de interés sobre la política, entonces se debe garantizar que los items estén destinados a evaluar dicho interés y no otro como la confianza en las instituciones que accionan en el país, por ejemplo (Martínez, 1996).

El muestreo inadecuado y el incumplimiento de la representatividad y relevancia de los indicadores del constructo puede llevar a conclusiones erróneas sobre el fenómeno estudiado, por lo que es de suma importancia evaluar su validez a partir del juicio de expertos en el área. De acuerdo también a las afirmaciones de Watkins (2018), no es apropiado incluir variables que

sean dependientes de otras en el análisis y se debe tratar con sumo cuidado indicadores formativos que funcionan como determinantes del constructo más que como sus efectos.

3. Participantes

Los participantes han debido ser seleccionados a partir de un muestreo probabilístico que sea representativo de la población venezolana e incluya cada subgrupo de esta población en la proporción adecuada si el estudio lo amerita. Adicionalmente, se debe garantizar que el tamaño de la muestra debe reproducir adecuadamente los valores de la población. Aunque el tamaño de la muestra depende en gran medida por el número de variables evaluadas, el nivel de correlación entre las variables medidas y también su relación con los factores, una aproximación aceptable es tomar una relación de 5:1 o 10:1 entre el número de variables y factores (Watkins, 2018).

Para el caso trabajado sobre la variable de *Interés por la Política*, existen al menos 15 variables que están siendo incluidas en el factor y se recogió un total de 1190 observaciones en la población. De acuerdo a las consideraciones anteriores, este número de observaciones es adecuado también para el procedimiento a emplear que es el Análisis Factorial Exploratorio, debido a que esta técnica estadística requiere de al menos 10 observaciones por cada variable incluida para que su uso resulte recomendable, sustentado en el principio estadístico de que se debe obtener una matriz de tamaño $p < n$. Dado el número de observaciones incluidas ($n=1190$), es posible afirmar que se cuenta con el número necesario de observaciones, aunque se debe tener en cuenta que la idoneidad de los resultados no solo depende de este tamaño muestral.

Por su parte, la muestra de participantes estuvo conformada por adultos (mayores de 18 años) en Venezuela y no necesariamente nacidos en el país pero sí residentes del mismo. Del total de participantes ($n=1190$), aproximadamente el 48% pertenecían al género masculino (571), mientras que aproximadamente el 52% pertenecían al género femenino (619). Adicionalmente, el rango de edad abarcado era desde los 18 hasta los 85 años, con una media de edad de 38 años ($X= 38.31$; $S= 15.25$). Por último, aproximadamente el 99% de los participantes entrevistados había nacido en Venezuela (1176 de ellos) y el otro 1% en otro país (14).

4. Revisión de los datos

Como puede observarse en la Figura 1, el análisis descriptivo de las 15 variables o ítems analizados muestra una media que va desde 2.53 hasta 3.78 para las opciones de respuesta. Cabe destacar que estas opciones de respuestas conformaban un total de cinco alternativas que se expresaban en una escala de acuerdo que iba desde “Muy en Desacuerdo” hasta “Muy de Acuerdo”. La desviación típica de estos valores también se encontraba desde 1.10 hasta 1.25. También puede observarse que en el caso de los dos primeros ítems de la escala la mediana es de 4, mientras que en los dos últimos ítems es de 3.

Debe tenerse en cuenta que las variables se encuentran en un nivel de medición ordinal, por lo que estas medidas deben analizarse con cuidado, ya que una un valor de media de 3.58 no forma parte de la escala, lo que quiere decir que las respuestas se encuentran entre la opción de respuesta 3 y 4, las cuales hacen referencia a “Ni en Desacuerdo ni De Acuerdo” y “De Acuerdo”.

Por otro lado, el valor mínimo de la escala se ubica en 1 que comprende la primera opción de respuesta, mientras que el valor máximo se ubica en 5 que comprende la última opción de respuesta. El rango de todas las respuestas es 4 y también son evidentes valores que indican una asimetría negativa. Esto quiere decir que quizás existen pocas respuestas en el polo negativo que empieza con las alternativas de respuesta “Muy en Desacuerdo” y “Desacuerdo” y muchas respuestas en el polo positivo que terminan con las alternativas de respuesta “De Acuerdo” y “Muy en Desacuerdo”. Para el caso presente, tanto la media podada (trimmed) como la desviación absoluta de la mediana (mad) resultan medidas mucho más adecuadas de la evaluación descriptiva de las variables o ítems al estar representadas en una escala de medida ordinal. El valor de la media podada se encuentra entonces entre 2.46 y 3.91. La desviación absoluta de la mediana indica que el promedio de las desviaciones en valores absolutos de cada rango menos la mediana correspondiente es de 1.48.

Los valores de la kurtosis también sugieren comportamientos de aplanamiento sobre la distribución, por lo que los datos se encuentran relativamente dispersos con relación a la media.

Lo anterior también puede indicar la posibilidad de colas más delgadas (menor probabilidad de valores extremos) y una distribución uniforme en torno a la media.

Figura 1

Estadísticos descriptivos de los ítems de la escala de Interés sobre la Política

	vars	n	mean	sd	median	trimmed	mad	min	max	range	skew	kurtosis	se
pol94	1	1190	3.58	1.19	4	3.68	1.48	1	5	4	-0.58	-0.67	0.03
pol95	2	1190	3.78	1.10	4	3.91	1.48	1	5	4	-0.84	-0.01	0.03
pol96	3	1190	2.58	1.14	2	2.53	1.48	1	5	4	0.38	-0.71	0.03
pol97	4	1190	2.55	1.21	2	2.47	1.48	1	5	4	0.44	-0.81	0.04
pol98	5	1190	3.26	1.23	3	3.31	1.48	1	5	4	-0.26	-1.01	0.04
pol99	6	1190	3.58	1.15	4	3.66	1.48	1	5	4	-0.58	-0.54	0.03
pol100	7	1190	2.53	1.14	2	2.46	1.48	1	5	4	0.43	-0.62	0.03
pol101	8	1190	3.41	1.13	4	3.47	1.48	1	5	4	-0.39	-0.65	0.03
pol102	9	1190	3.56	1.13	4	3.63	1.48	1	5	4	-0.49	-0.57	0.03
pol103	10	1190	3.41	1.11	3	3.44	1.48	1	5	4	-0.23	-0.79	0.03
pol104	11	1190	2.87	1.25	3	2.83	1.48	1	5	4	0.11	-1.06	0.04
pol105	12	1190	3.53	1.10	4	3.59	1.48	1	5	4	-0.48	-0.50	0.03
pol106	13	1190	3.41	1.15	4	3.45	1.48	1	5	4	-0.30	-0.80	0.03
pol107	14	1190	2.69	1.16	3	2.66	1.48	1	5	4	0.22	-0.86	0.03
pol108	15	1190	2.71	1.18	3	2.66	1.48	1	5	4	0.24	-0.84	0.03

Nota. Tomado de la salida del paquete estadístico RStudio.

Posterior al análisis descriptivo, resulta necesario comprobar los supuestos de linealidad, normalidad (univariante y multivariante) y comprobar los estadísticos pertinentes de correlación de acuerdo al nivel de medición de la variable, con la finalidad de evaluar la pertinencia del uso de la técnica estadística de Análisis Factorial Exploratorio.

Para comprobar el supuesto de linealidad y como aspecto inicial de cualquier Análisis Factorial Exploratorio se debe calcular y analizar la matriz de correlaciones de las variables observadas. Como el nivel de medición de las variables o ítems corresponde a un nivel de medida ordinal, en el que no se conservan distancias iguales entre los valores, el estadístico apropiado en este caso resulta el Coeficiente de Correlación de Spearman. El coeficiente debe medir la relación lineal entre las variables, en la que se cumpla el comportamiento estadístico de que no solo las variables aumentan o decrecen cuando otras lo hacen, sino que lo hagan en la misma magnitud. Por su parte, la presencia de correlaciones demasiado altas o bajas será indicador de la falta de idoneidad de un AFE.

Matriz de correlaciones de Spearman entre los indicadores de la escala de Interés sobre la Política



Para los otros recuadros numéricos por encima de la diagonal principal se muestran además algunas correlaciones por encima de cero, pero también debe enfatizarse la presencia de correlaciones negativas menores que cero. Por debajo de la diagonal principal, las líneas graficadas indican el comportamiento de la nube de puntos que intentan explicar un comportamiento lineal. Si bien esto ocurre en algunos casos y es posible observar tal

comportamiento lineal a través de líneas que forman una pendiente (p. ej. en el caso de los ítems nueve y ocho que expresan que *“la gente en el gobierno con frecuencia muestra mal juicio”* y *“los políticos frecuentemente son incompetentes y eficaces”*), no debe obviarse el caso de otras combinaciones o parejas de ítems que parece adoptar una forma no lineal (p. ej. en el caso de los ítems once y doce que expresan *“los políticos a menudo ponen al país por encima de sus intereses personas”* y *“los políticos suelen ignorar a mi comunidad”* y cuyo comportamiento parece ser curvilíneo o anómalo). Por lo tanto, debe tenerse precaución en la interpretación de los datos debido a indicios de que el principio de linealidad puede estar siendo ligeramente no cumplido.

En cuanto al principio de normalidad univariada, el cual toma en cuenta la distribución ajustada a la forma de una distribución normal o no de una sola variable por separado como medida parcial de la normalidad multivariante, en la Figura 1 se observa que no hay indicios de que no se cumple el principio de normalidad debido a que los valores están por debajo de 2 en su mayoría, aunque sí son evidentes valores que indican asimetría negativa, lo que puede explicarse por la magnitud elevada del tamaño de la muestra. Al mismo tiempo, los valores de la curtosis son menores a 7, lo cual sumado a los valores de la asimetría confirman que la distribución de cada variable no parece diferir de una distribución normal.

Para la identificación normalidad multivariada, a través del cálculo de asimetría y curtosis multivariada (Figura 3) se observó que el valor de la curtosis fue de 71.27, lo cual es un valor muy superior al valor límite de referencia igual a 5, por lo que se puede concluir que los datos no cumplen con el supuesto de normalidad multivariante, aunque no haya habido indicios de esto para el caso del cálculo de la normalidad univariada. La simetría multivariante es también bastante elevada lo que confirma el incumplimiento de este supuesto.

Figura 3

Asimetría y curtosis multivariada de los indicadores de la escala de Interés sobre la Política

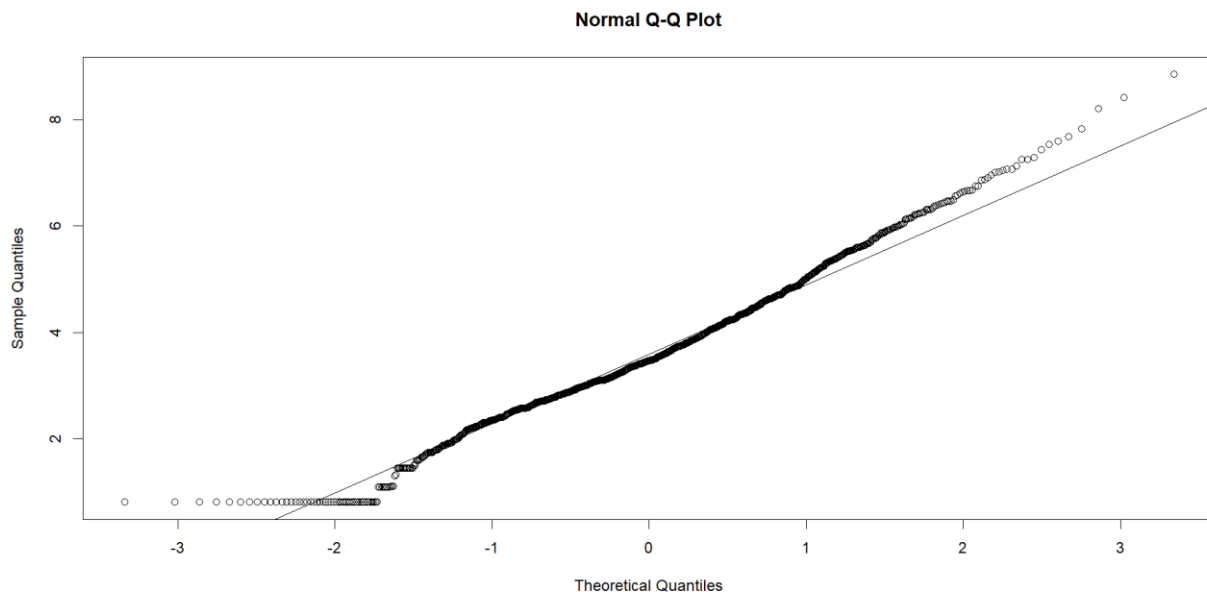
```
Mardia tests of multivariate skew and kurtosis
Use describe(x) the to get univariate tests
n.obs = 1190    num.vars = 15
b1p = 17.1    skew = 3391.7    with probability <= 0
  small sample skew = 3401.32    with probability <= 0
b2p = 348.32    kurtosis = 71.27    with probability <= 0
```

Nota. Tomado de la salida del paquete estadístico RStudio.

De otro modo, el gráfico Q-Q de normalidad (Figura 4) arrojado por la salida del estadístico mardia deja también al descubierto que los cuantiles teóricos y los cuantiles multivariados arrojados por los datos difieren en cierta medida. Aunque el comportamiento en el centro de los datos parece ser el de una distribución normal, las colas muestran una tendencia a la asimetría negativa.

Figura 4

Gráfico Q-Q de normalidad de los indicadores de la escala de Interés sobre la Política

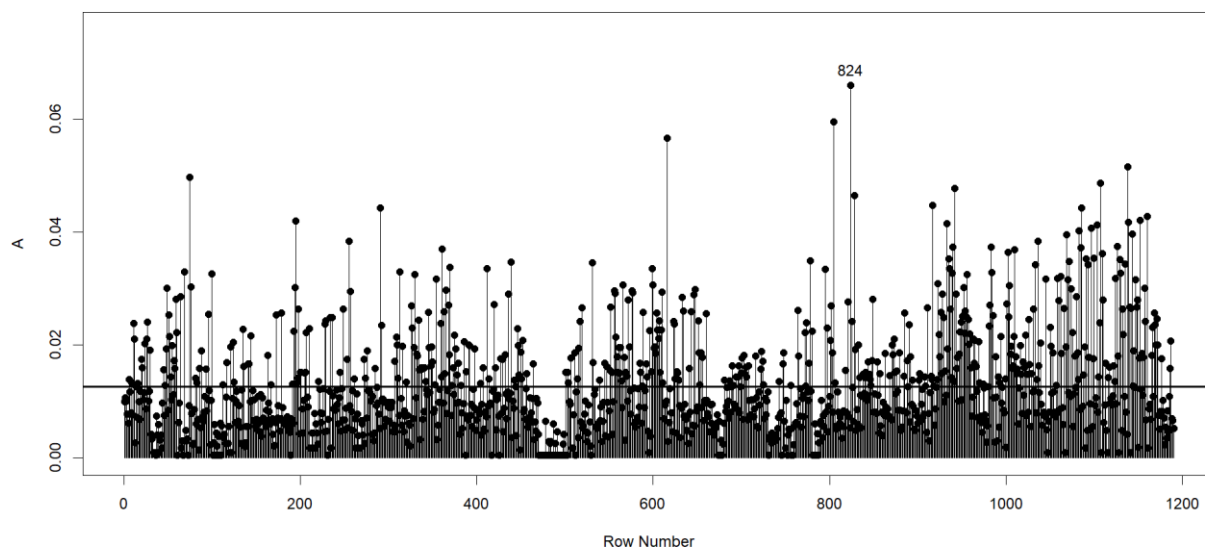


Nota. Tomado de la salida del paquete estadístico RStudio.

Respecto de los valores atípicos, calculados a través de la función bollen se obtiene un gráfico que arroja los datos multivariantes que pudieran influir sobre la calidad del ajuste. En este caso se observa que el dato multivariante 824 es identificado como un valor atípico que no solo se toma en cuenta para un indicador sino para los 15 indicadores incluidos en la escala.

Figura 5

Gráfico bollen de valores atípicos en los datos de la escala de Interés sobre la Política



Nota. Tomado de la salida del paquete estadístico RStudio.

Al analizar individualmente este registro se observa que sus alternativas de respuesta son de carácter extremo, ya que algunos ítems los contesta con la alternativa de “Muy de Acuerdo” y otras con la alternativa “Muy en Desacuerdo”. De hecho, para este caso no existen opciones o alternativas de respuesta intermedias que difieran las ya mencionadas y que tienen unos valores de 5 y 1 respectivamente.

Aunque estadísticamente se haya incumplido con el principio de normalidad multivariante, se asume que la variable de *Interés sobre la Política* sí se distribuye de forma normal y acampanada pero que debido a limitaciones del instrumento, solo se es capaz de aproximarse a la evaluación de este constructo de forma parcial, por lo cual es plausible realizar el cálculo correlaciones policóricas entre las parejas de los ítems para observar su

comportamiento estadística. Al realizar el cálculo, puede observarse como algunos de los ítems muestran comportamientos atípicos que se reflejan en correlaciones iguales a cero (p. ej. como el ítem dos y el ítem once que hablan de ser cuidadosos con los políticos y la creencia de que los políticos ponen por encima los intereses del país antes que los suyos) e incluso correlaciones negativas (p. ej. el ítem cuatro que expresa que el gobierno suele hacer lo correcto y el ítem doce que afirma que el gobierno suele ignorar a la comunidad).

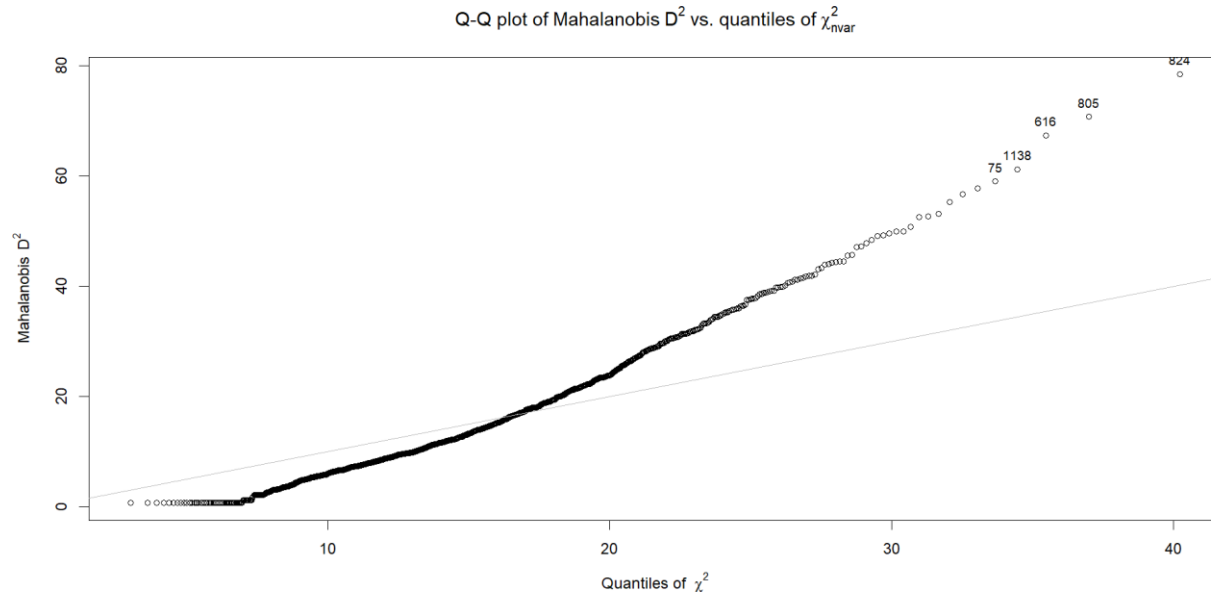
Sin embargo, también se encuentran correlaciones policóricas moderadas entre ítems como el ítem ocho y el ítem nueve, los cuales afirman “la gente en el gobierno con frecuencia muestra mal juicio” y “los políticos frecuentemente son incompetentes e ineficaces” respectivamente. Muchas de las correlaciones se encuentran por encima de 0.30 lo que puede ser indicador de la idoneidad del AFE. También es posible ver ciertos niveles de asociación entre grupos de ítems que hacen referencia a fenómenos similares.

Tomando en cuenta el cálculo de la matriz de correlaciones policóricas con la totalidad de observaciones o registros, es posible también hacer el cálculo de esta misma matriz omitiendo el registro que conforma el valor atípico multivariante. Posteriormente, es posible calcular la diferencia entre ambas matrices, para lo cual si se obtienen valores muy distintos de cero es pertinente concluir que el valor atípico influye en la calidad del ajuste y los estimados de los parámetros del modelo del análisis de factores.

El resultado de la diferencia entre las matrices policóricas tomando en cuenta todos los registros y sin el registro que comprende el valor atípico arroja valores de cero para todas las posibles parejas de ítem, lo que indica que el valor atípico no genera ninguna influencia sobre la calidad del ajuste del modelo. Se hizo lo propio tomando en cuenta el cálculo de la distancia de Mahalanobis con lo cual se pudo observar que aunque hubo otros valores atípicos adicionales (a saber, las observaciones 75, 1138, 616, 805 y 824 como la más alejada del centroide) las diferencias entre la matriz policórica original y la matriz policórica sin estos valores atípicos no superaban la centésima.

Figura 6

Gráfico Q-Q para valores atípicos en los datos de la escala de Interés sobre la Política a través de la distancia de Mahalanobis

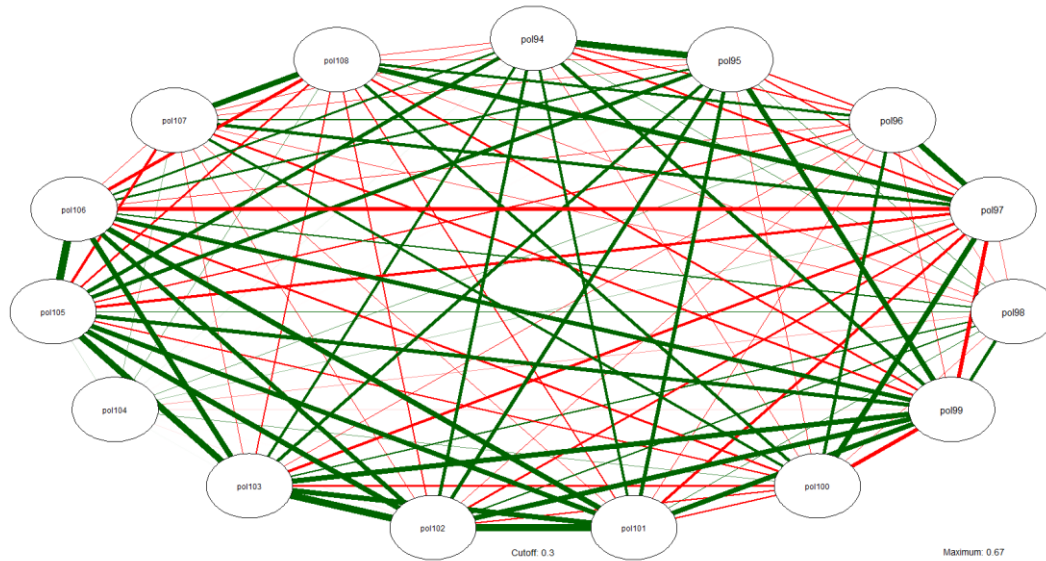


Nota. Tomado de la salida del paquete estadístico RStudio.

Por último, se examinó un gráfico basado en la matriz de correlaciones policóricas (Figura 7) que mostrara los quince ítems y todas las posibles combinaciones entre ellos. En cuanto mayor sea el grosor de las líneas verdes que unen a cada pareja de ítems, estas serán indicativo de una mayor magnitud entre su correlación policórica. En cambio, las líneas rojas muestran todas las correlaciones negativas entre los pares de ítem y puede observarse que mientras mayor grosor tenga la línea roja, a su vez de mayor magnitud es la correlación negativa. Los patrones de comportamiento son los ya reportados en los párrafos anteriores y corresponden a las afirmaciones que expresan los ítems respecto a las percepciones del gobierno y los políticos.

Figura 7

Gráfico basado en la matriz de correlaciones policóricas de datos de la escala de Interés sobre la Política



Nota. Tomado de la salida del paquete estadístico RStudio.

5. ¿Es apropiado el EFA?

Un paso esencial en la ejecución de un análisis de factores es la comprobación de que los datos a utilizar son apropiados para esto, más específicamente, resulta ineludible verificar que las variables a incluir en el análisis están lo suficientemente intercorrelacionadas como para justificar el análisis. Con esta finalidad se computó, en primer lugar, la medida de adecuación muestral *Kaiser-Meyer-Olkin* (KMO), la cual informa de la proporción en la cual las correlaciones totales y parciales son una función de la varianza compartida por todas las variables y no la varianza compartida por un par particular de variables. Y, en segundo lugar, la prueba de esfericidad de *Bartlett*, la cual compara la matriz de datos observados y la matriz de identidad docimando como hipótesis nula que la matriz incluida en el análisis es igual a la matriz de identidad.

Lo encontrado en el análisis sugiere que las variables están lo suficientemente intercorrelacionadas como para llevar a cabo el análisis de factores, más específicamente, se

encontró que la proporción en la cual las correlaciones totales y parciales son una función de la varianza compartida por todas las variables es alta tanto para la escala en general ($KMO = .92$) como para las variables individualmente, siendo en todas mayor a $.88$, lo que, siguiendo lo propuesto por Kaiser (1974), puede ser considerado maravilloso/meritorio. Además de esto se encontró que efectivamente existen diferencias estadísticamente significativas entre la matriz de datos observados y la matriz de identidad ($X^2_{(105)} = 8055.219, p < .001$), lo que permite aseverar que las variables incluidas en el análisis están lo suficientemente correlacionadas.

6. Elección del modelo apropiado (FA o PCA)

Teniendo en cuenta que la finalidad del estudio aquí descrito es identificar constructos latentes que logren explicar la mayor cantidad de información posible sobre las personas a partir de las variables medidas, teniendo como supuesto que el resultado del uso del instrumento de medición utilizado, en este caso la escala presentada en la última sección de la EMVC denominada “Interés por la Política”, es un reflejo de la estructura subyacente del constructo medido, se elige entonces el análisis de factores como el modelo apropiado para esta tarea.

La teoría subyacente a este modelo hace referencia a que el análisis de factores busca separar la varianza total de las variables medidas en dos varianzas: la varianza que es común para todas las variables incluidas en el análisis (la comunalidad o h^2) y la varianza que es propia de cada variable (representada por u^2). La segunda varianza, la única de cada variable, está compuesta por varianza que es fiable o repetible pero que no se comparte con las otras variables (s^2) y por el error de medición (e) (Watkins, 2018). Atendiendo a esto y a que el propósito del estudio es identificar constructos latentes responsables por la variación de las variables medidas, tal como fue mencionado anteriormente, la opción de llevar a cabo el análisis de factores resulta la apropiada.

7. Elección del método de extracción de factores

Son muchos los métodos de extracción que se pueden utilizar en el análisis de factores para estimar el modelo, sin embargo, en este caso el método elegido para esta tarea será

MINRES, el cual usa mínimo cuadrados ordinarios (OLS) para minimizar los elementos no diagonales de la matriz residual de correlación ajustando los autovalores de la matriz de correlación original. MINRES, también llamado eje principal iterado, aprovecha las comunalidades estimadas inicialmente para generar estimaciones intermedias que permitan una estimación mucho más precisa de las comunalidades que son iteradas hasta dar con una solución final que reproduzca de la mejor manera posible la matriz de correlaciones de las variables medidas (Watkins, 2018). Además de estas características que fundamentan la elección de MINRES como el método de extracción de los factores, es importante también tomar en cuenta que este método es especialmente preferido sobre métodos como el de máxima verosimilitud cuando el tamaño de la muestra es pequeño, cuando no se puede asegurar la normalidad multivariante y cuando el número de factores que subyace a las variables medidas no está especificada (Watkins, 2018).

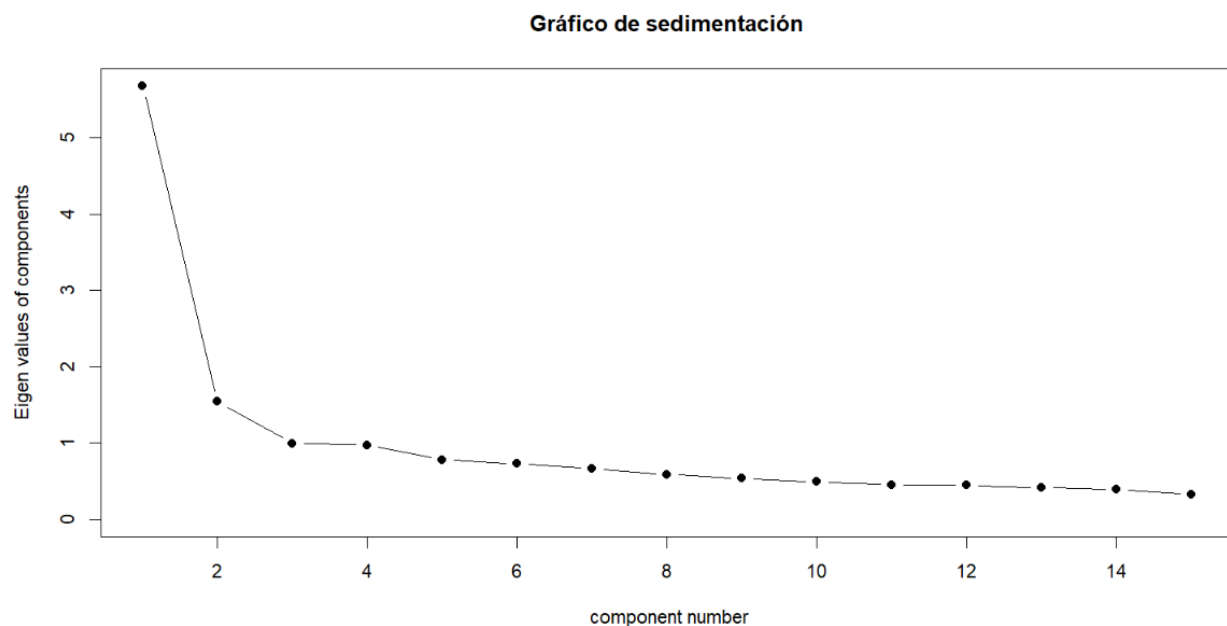
8. ¿Cuántos factores retener?

Tal como se ha mencionado en diferentes ocasiones hasta ahora, la finalidad de la aplicación del análisis de factores con estas variables en esta muestra es identificar factores subyacentes al conjunto de datos que expliquen la mayor cantidad de información posible en torno al comportamiento de interés, en este caso, el interés y la participación política. En este sentido, gran parte del trabajo es determinar el número de factores que permitan cumplir de la mejor manera posible esta tarea. Un aspecto esencial de esto que debe regir la práctica es que un uso exitoso del análisis de factores está caracterizado por la parsimonia y la exhaustividad, es decir, el modelo final debe tener el número suficiente de factores para dar cuenta de la covariación importante entre las variables medidas pero no tantos factores como para que se pierda la esencia de la técnica de explicar un gran número de variables a partir de los factores que las subyacen. A continuación se reportan los resultados encontrados a partir de cuatro estrategias que buscan dar una idea fundamentada del número de factores a retener, a saber: el gráfico de sedimentación, el gráfico de redes, los parciales mínimos promedios (MAP por sus siglas en inglés) y el análisis paralelo.

A través del paquete *psych* (Revelle, 2024) de *R* se computó el gráfico de sedimentación, el cual grafica la magnitud de los autovalores de los factores contra la cantidad total de factores incluidos en el modelo, obteniendo la Figura 6. Catell (1966) propone que para darle sentido a lo que se puede percibir en este gráfico se use la técnica del codo, la cual hace referencia a que el número de factores a retener sea igual al número de factores que están por encima de un quiebre evidente en la pendiente del gráfico, que en este caso serían dos factores.

Figura 6

Gráfico de sedimentación para los ítems de la escala de Interés sobre la Política



Nota. Tomado de la salida del paquete estadístico RStudio.

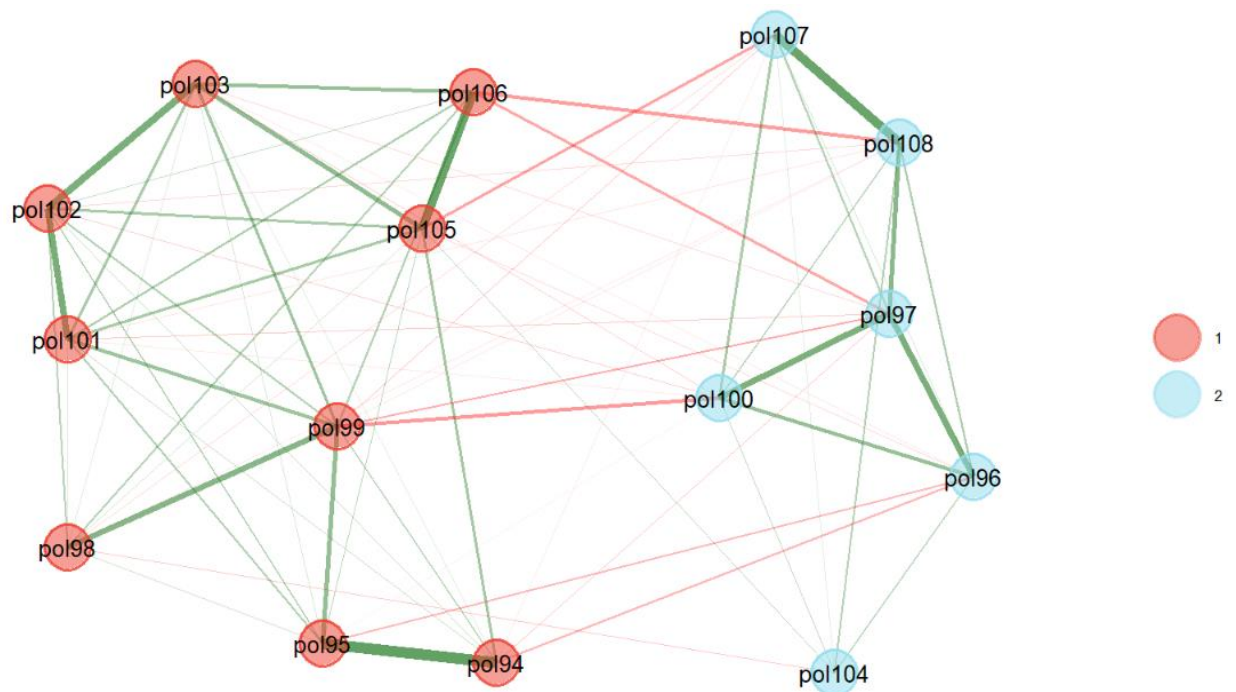
Cómo es razonable pensar, y tal como indica (Gorsuch, 1983), la interpretación del gráfico de sedimentación puede verse influenciada por la subjetividad de quién intente darle sentido para determinar el número de factores a extraer. A pesar de esto, y atendiendo a que este gráfico es de amplio uso dentro del análisis de factores, lo que este gráfico sugiere que puede tomar como un buen indicio del número de factores a extraer.

Otro buen indicio del número de factores a extraer, aunque más indirecto que el resto de estrategias utilizadas, es el gráfico de redes. Este gráfico forma agrupaciones de variables, tomando la matriz de correlaciones policóricas, a partir de la magnitud de las correlaciones entre

todas las variables incluidas en el análisis. Cómo se puede observar en la Figura 7, el gráfico de redes sugiere que a partir de las correlaciones policóricas entre las variable se pueden formar dos agrupaciones de indicadores, un primer conjunto compuesto por 9 ítems e identificado por el color rojo y un segundo conjunto constituido por 6 ítems e identificado con el color azul. Tal cómo se observa en este gráfico, las líneas que conectan las variables dentro de cada agrupación son más gruesas que las líneas entre las variables de diferentes agrupaciones, lo cual indica que la correlación entre las variables de la misma agrupación tienen una mayor magnitud, fundamentando así su reunión en un mismo grupo.

Figura 7

Gráfico de redes de las correlaciones entre los ítems de la escala de Interés sobre la Política



Nota. Tomado de la salida del paquete estadístico RStudio.

Lo obtenido a partir de este gráfico es congruente con lo encontrado en el gráfico de sedimentación a partir de la técnica del codo, esto es, dos podrían ser los factores a extraer para generar el mejor modelo. Sin embargo, teniendo en cuenta las debilidades que implica el uso de gráfico de sedimentación y que el gráfico de redes forma las agrupaciones solo a partir de la

magnitud de las correlaciones policóricas, es necesario continuar en la implementación de estrategias incluso más robustas para determinar el número de factores a retener.

El método empírico de promedios mínimos parciales (MAP por sus siglas en inglés) se basa en el cálculo de una matrix de correlaciones parciales luego de la extracción de cada componente o factor y se calcula el promedio de las correlaciones parciales al cuadrado sin la diagonal para cada una de estas matrices. Cuando se alcance el número correcto de factores extraídos este promedio debería llegar a su valor mínimo, esto debido a que la varianza común máxima se deja fuera de la matriz. A través del paquete *psych* (Revelle, 2024) es posible emplear la función VSS (*very simple structure* o estructura muy simple) con la cual se puede encontrar el valor mínimo del promedio de estas correlaciones parciales para determinar el número de factores a extraer. En este caso, el valor de MAP para las variables incluidas alcanza su mínimo (0.021) en el segundo factor, tal como puede observarse en la Figura 8, lo que indica que el promedio de las correlaciones parciales al cuadrado sin la diagonal es lo más bajo posible cuando se computa un modelo de 2 factores, indicando así que este sería el número ideal de factores a extraer. Convenientemente, esta información brindada por esta estrategia corrobora lo encontrado a partir del gráfico de sedimentación y el gráfico de redes.

Figura 8
Criterios para la selección del número de factores

Statistics by number of factors											
	vss1	vss2	map	dof	chisq	prob	sqresid	fit	RMSEA	BIC	SABIC
1	0.85	0.00	0.025	90	1694.8	2.6e-294	7.2	0.85	0.122	1057	1343
2	0.64	0.77	0.021	76	864.0	6.3e-134	10.9	0.77	0.093	326	567
3	0.53	0.66	0.030	63	521.2	2.5e-73	16.2	0.66	0.078	75	275
4	0.53	0.63	0.036	51	386.3	4.9e-53	16.2	0.66	0.074	25	187
5	0.50	0.57	0.051	40	252.1	1.4e-32	19.1	0.60	0.067	-31	96
6	0.46	0.52	0.063	30	157.3	3.4e-19	22.3	0.54	0.060	-55	40
7	0.35	0.40	0.081	21	38.0	1.3e-02	27.0	0.44	0.026	-111	-44
8	0.33	0.37	0.095	13	9.9	7.0e-01	27.7	0.43	0.000	-82	-41

Nota: VSS: very simple structure; MAP: minimum average partials; BIC: bayesian information criterion; SABIC: sample-adjusted information criterion.

A pesar de que las tres estrategias usadas hasta ahora apuntan a lo mismo, esto es, construir un modelo factorial que esté constituido por dos factores, un análisis adecuado se caracteriza por usar múltiples técnicas para determinar el número de factores a retener y luego juzgar cada modelo a partir de la forma en que se agrupan las variables y la exhaustividad y parsimonia que estos modelos presentan, además de sus indicadores de ajuste por supuesto. Es por esto que para determinar el número de factores a retener también se hace uso del criterio *very simple structure* (estructura muy simple) o VSS por sus siglas en inglés, del criterio de información bayesiana (BIC por sus siglas en inglés) y de este mismo criterios ajustado para el tamaño de la muestra (SABIC por sus siglas en inglés).

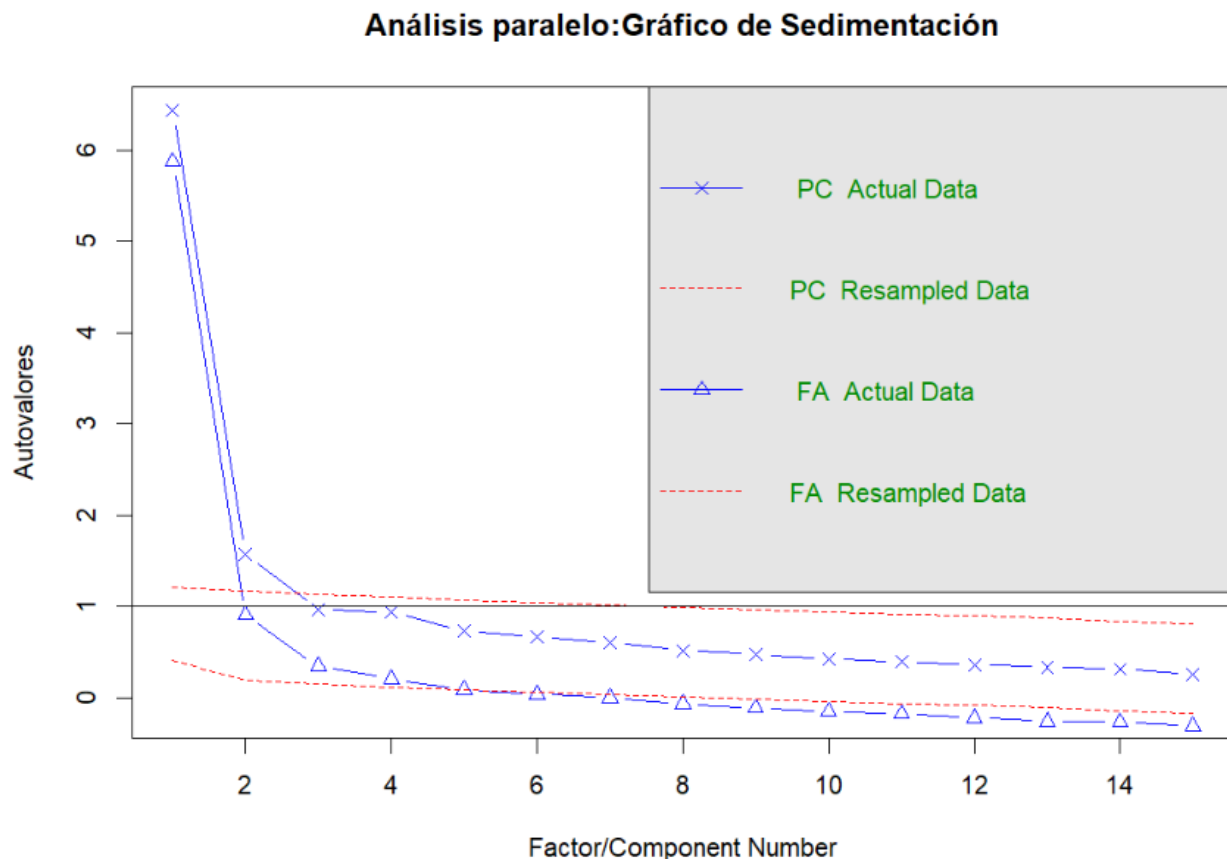
En la Figura 8 se pueden observar los resultados obtenidos para cada criterio con un número de factores desde 1 hasta 8. En cuanto al criterio VSS se encuentra que cuando se maneja una complejidad de 1 el valor máximo de este criterio se da con 1 factor mientras que cuando la complejidad es de 2 el valor máximo se da con 2 factores, esto último siendo congruente con lo encontrado a partir del criterio de los promedios mínimos parciales. Mientras tanto, en el caso de los criterios de información bayesiana, incluyendo el criterio ajustado para el tamaño de la muestra, se ha encontrado que sus valores mínimos se dan cuando son 7 los factores incluidos en el modelo. Hasta ahora la mayoría de criterios utilizados hasta ahora sugieren que el número ideal de factores a retener es 2, exceptuando los 7 factores sugeridos por los criterios de información bayesiana y el único factor sugerido por el criterio VSS con complejidad 1.

La última estrategia empírica a utilizar para determinar el número de factores a retener en este análisis de factores es el análisis paralelo. El análisis paralelo consiste en la simulación estadística de un grupo de datos aleatorios con el mismo número de variables y de registros que el conjunto real de datos, en este caso 15 variables y 1190 registros. Luego de esta simulación, la matriz de correlaciones se somete a un análisis de factores y los autovalores resultante se guardan para repetir este proceso al menos 100 veces y el número final de autovalores se promedian y se comparan con los extraídos a partir de los datos reales. Para conocer el número de factores a extraer se observa la pendiente de datos reales y se cuentan los factores que están por encima de la pendiente que representa los datos simulados. En la Figura 9, ignorando los datos para el análisis de componentes principales, puede observarse que el número de factores

extraídos de los datos reales que está por encima de la pendiente que representa los datos simulados es 4.

Figura 9

Análisis paralelo de los factores a partir de los autovalores y los datos simulados



Nota. Tomado de la salida del paquete estadístico RStudio.

Para resumir lo expuesto hasta aquí, el gráfico de sedimentación, el gráfico de redes, los promedios parciales mínimos y el criterio VSS con complejidad 2 indican que el número ideal de factores a extraer es 2. Mientras tanto, el análisis paralelo para el análisis de factores sugiere que el número de factores a extraer debería ser 4, fundamentándose esto en la comparación con los datos simulados. Por otro lado, el criterio VSS con complejidad 1 indica que de estos datos debería extraerse sólo un factor. Finalmente, ambos criterios de información bayesiana indican que son 7 los factores que deberían retenerse en este análisis de factores. Teniendo todo esto en cuenta y apuntando a la parsimonia y exhaustividad con las que debe contar el modelo resultante,

se construirán modelos de dos, tres y cuatro factores, esto con la finalidad de determinar cual explica la mayor cantidad de información, cual agrupa de manera más sencilla los ítems, y cuál tiene mejor rendimiento en cuanto a los índices de ajuste.

9. Rotación de factores – Oblicua y Promax

Para llevar a cabo un análisis de factores que arroje una solución factorial de mayor facilidad en cuanto a su interpretación es necesario emplear un método de rotación de los ejes que conforman el plano sobre el cual se proyectan las variables y los factores. La rotación en el análisis de factores está diseñada para alcanzar una solución factorial más simple y más significativa en términos teóricos, lo que se logra a partir de la rotación de los ejes para acercarlos más a la ubicación de las variables que se proyectan en el espacio factorial (Watkins, 2018). Existen dos tipos de rotaciones que se pueden emplear y que difieren en términos del ángulo resultante entre los ejes luego de la rotación: la rotación oblicua, en la que no se mantiene el ángulo de 90° entre los ejes, y la rotación ortogonal, en la que los ejes se rotan manteniendo el ángulo recto entre ellos.

Teniendo en cuenta que es razonable pensar que las variables medidas van a generar factores correlacionados entre sí, lo cual se fundamenta en la interrelación que existe entre las variables, y que autores como Fabrigar y Wegener (2012) indican que las rotaciones oblicuas proveen de soluciones que representan de forma más realista los datos, se ha tomado la decisión en este análisis de usar la rotación oblicua de los ejes, más específicamente la rotación Promax, la cual es posiblemente la opción más utilizada y la que tiende a producir soluciones más simples.

10. Interpretación de los resultados del FA

a. Con cuatro factores:

Por su parte, el modelo de cuatro factores, que fue el sugerido por el análisis paralelo, tomando en cuenta un análisis factorial en lugar del análisis de componentes principales, arroja

los siguientes resultados a partir de la rotación oblicua denominada Promax, la cual es favorable para los casos en los que se trabaja con muestras de grandes tamaños y también suele producir soluciones más simples.

Figura 10

Matriz de cargas factoriales para la rotación promax de los datos con cuatro factores

Standardized loadings (pattern matrix) based upon correlation matrix								
	item	MR1	MR2	MR3	MR4	h2	u2	com
pol103	10	0.75				0.59	0.41	1.0
pol102	9	0.75				0.60	0.40	1.0
pol106	13	0.72				0.63	0.37	1.5
pol101	8	0.67				0.54	0.46	1.1
pol105	12	0.66				0.67	0.33	1.6
pol99	6	0.56				0.60	0.40	1.5
pol98	5	0.54				0.32	0.68	1.6
pol97	4		0.74			0.62	0.38	1.0
pol108	15		0.69			0.57	0.43	1.3
pol96	3		0.60			0.44	0.56	1.4
pol107	14		0.56			0.41	0.59	1.2
pol100	7		0.53			0.47	0.53	1.3
pol95	2			0.80		0.72	0.28	1.1
pol94	1			0.63		0.55	0.45	1.2
pol104	11				0.39	0.20	0.80	1.7

Nota. Tomado de la salida del paquete estadístico RStudio.

Tal y como puede observarse, los items 10, 9, 13, 8, 12, 6 y 5 cargan en un primer primer factor, con cargas que van desde 0.54 hasta 0.75. Debido a que el programa fue instruido para que eliminara las cargas menores a 0.30, puede observarse que estos items presentan cargas de magnitud considerable para este primer factor y no para los otros, lo que hace la solución más sencilla de interpretar. Estadísticamente, no se observa la presencia de casos anómalos o ultra-Heywood en los que las cargas factoriales sean iguales o mayores a 1. La tabla de varianza muestra que los cuatro factores logran explicar una varianza total acumulada de 53%, lo cual podría considerarse “aceptable” para el caso de variables en las ciencias sociales.

De forma sustantiva, puede identificarse que los ítems del primer factor hablan sobre la falta de veracidad de la información proporcionada por el gobierno, una actitud de desconfianza hacia este, así como también aspectos que involucran malas decisiones, incompetencia, falta de

respeto hacia los individuos, injusticia y poca ayuda a los comunidades. Como resultado, se observa que el primer factor muestra una estructura coherente respecto de los indicadores que allí se incluyen al apuntar a fenómenos que tienen que ver con actitudes negativas que tienen las personas hacia el gobierno en general. En cambio, los items del segundo factor (items 4, 15, 3, 14 y 7) se relacionan con actitudes positivas hacia los políticos, tales como ser abiertos sobre sus decisiones, hacer lo correcto, ser honestos y veraces y atender las necesidades de la comunidad. El tercer factor (items 2 y 1) se relaciona con la forma de aproximarse al gobierno y creer o confiar o no en él. Por último, el cuarto factor solo contiene un item (item 11) que expresa la posibilidad de que los políticos pongan al país por encima de sus intereses personales.

La matriz de correlaciones también indica que existe una correlación negativa entre el primer y segundo factor, una correlación positiva entre el primero y el tercero y una correlación negativa entre el primero y el cuarto. Adicionalmente, el factor dos presenta una correlación negativa con el tercer factor, así como una correlación positiva entre baja y moderada con el cuarto factor. Finalmente, el tercer factor muestra nuevamente una correlación negativa y baja con el cuarto factor.

Figura 11

Matriz de correlaciones para el modelo de cuatro factores

	with factor correlations of			
	MR1	MR2	MR3	MR4
MR1	1.00	-0.62	0.64	-0.10
MR2	-0.62	1.00	-0.49	0.25
MR3	0.64	-0.49	1.00	-0.12
MR4	-0.10	0.25	-0.12	1.00

Nota. Tomado de la salida del paquete estadístico RStudio.

La complejidad promedio del ítem es de 1.3 y el test de hipótesis para cuatro factores muestra que esta cantidad de factores es suficiente para el modelo analizado. El contraste de hipótesis llevado a cabo a través del estadístico Chi-Cuadrado, el cual compara la matriz de covarianza observada con la matriz de covarianza reproducida por el modelo factorial, arrojó un valor de $X^2_{(105)} = 8055.22$, $p < .001$, lo que al ser un valor tan alto sugiere que el modelo no

presenta calidad de ajuste respecto a los datos. Sin embargo, debe tomarse en cuenta que este estadístico se ve fuertemente distorsionado por el tamaño de la muestra, que al ser de gran tamaño, puede tener como resultado una gran influencia sobre el resultado de la prueba.

Por su parte, la raíz cuadrada media de los residuos (referida como RMSR) arrojó un valor de 0.03 que según el estándar muestra una buena o excelente calidad del ajuste al ser menor que .05. Esto se debe a que los residuos representan las diferencias entre los elementos de la matriz de covarianza observada y los elementos de la matriz de covarianza reproducida por el modelo factorial, por lo que pequeñas diferencias implican una magnitud menor de desvíos respecto del modelo. La raíz cuadrada media de los residuos corregida también muestra un buen o excelente ajuste al arroja un valor de 0.04, lo que sigue siendo menor a .05.

La n observada armónica presenta un valor empírico igual al estadístico $X^2_{(105)} = 166.22$, con una probabilidad menor a $3.9e-14$, al igual que la media total de las observaciones con una probabilidad de Chi-Cuadrado de $X^2_{(105)} = 567.04$ con una probabilidad menor a $4.9e-53$, es decir, valores sumamente pequeños. En consonancia con lo anterior, el índice de Tucker Lewis de confiabilidad factorial proporciona un índice de ajuste incremental que compara el ajuste del modelo propuesto con el ajuste de un modelo nulo o que asume que no hay factores comunes, por lo que se esperan valores mayores a .90 para demostrar un buen ajuste. Para el caso tratado, este índice arrojó un valor de 0.913, lo cual es aceptable o razonable.

El error cuadrático medio de aproximación, también denominado índice RMSEA es ideal nuevamente para evaluar la calidad de ajuste del modelo y para los datos analizados presenta un valor de 0.074 e intervalos con un 90% de confianza que van desde 0.068 hasta 0.081, lo que se traduce en una medida de ajuste razonable considerando que arrojó valores mayores a .05 y menores o iguales a 0.08. Ahora bien, la medida de ajuste bayesiana identificada como BIC (criterio de información Bayesiana) piensa en la complejidad del modelo, por lo que valores reducidos serán ideales e indican un buen ajuste. Para los datos analizados, el BIC arroja un valor igual a 25.12, lo que puede interpretarse como un número considerablemente elevado que indica falta de calidad en el ajuste.

b. Con tres factores:

La iteración con tres factores, realizado con la finalidad de contrastar el modelo que muestre la mejor calidad de ajuste para las variables trabajadas, arroja los siguientes resultados a partir de la rotación oblicua denominada Promax:

Figura 12

Matriz de cargas factoriales para la rotación promax de los datos con tres factores

Standardized loadings (pattern matrix) based upon correlation matrix							
	item	MR1	MR2	MR3	h2	u2	com
pol106	13	0.86			0.61	0.39	1.1
pol103	10	0.81			0.59	0.41	1.0
pol105	12	0.80			0.61	0.39	1.0
pol102	9	0.74			0.59	0.41	1.1
pol101	8	0.69			0.54	0.46	1.1
pol99	6	0.49			0.55	0.45	1.5
pol98	5	0.39			0.22	0.78	1.1
pol108	15		0.67		0.58	0.42	1.5
pol97	4		0.62		0.59	0.41	1.2
pol100	7		0.55		0.47	0.53	1.2
pol96	3		0.55		0.43	0.57	1.4
pol107	14		0.54		0.42	0.58	1.4
pol104	11		0.40		0.11	0.89	1.4
pol95	2			0.90	0.76	0.24	1.0
pol94	1			0.61	0.51	0.49	1.1

Nota. Tomado de la salida del paquete estadístico RStudio.

Tal y como puede observarse, los items 13, 10, 12, 9, 8, 6 y 5 cargan en un primer factor, con cargas que van desde 0.39 hasta 0.86. Debido a que el programa fue instruido para que eliminara las cargas menores a 0.30, puede observarse que estos items presentan cargas de magnitud considerable para este primer factor y no para los otros, lo que podría complejizar la estructura de los datos incluidos en el modelo. Estadísticamente, no se observa la presencia de casos anómalos o ultra-Heywood en los que las cargas factoriales sean iguales o mayores a 1. La tabla de varianza muestra que los cuatro factores logran explicar una varianza total acumulada de 51%, lo cual podría considerarse “aceptable” para el caso de variables en las ciencias sociales.

De forma sustantiva, puede identificarse que los items del primer factor (items 13, 10, 12, 9, 8, 6 y 5 respectivamente) hablan nuevamente sobre la falta de competencia del gobierno

respecto a sus acciones e intenciones respecto de los ciudadanos, así como también la falta de confianza hacia esta institución macro. Hace referencia a características de mal juicio, falta de respeto, injusticia y falta de consideración a las necesidades de la comunidad. Como resultado, se observa que el primer factor muestra una estructura coherente respecto de los indicadores que allí se incluyen al apuntar a fenómenos que tienen que ver con actitudes negativas que tienen las personas hacia el gobierno en general.

De otro modo, los items del segundo factor (items 15, 4, 7, 3, 14 y 11) se relacionan con actitudes positivas hacia los políticos, tales como ser abiertos sobre sus decisiones, hacer lo correcto, ser honestos y veraces y atender las necesidades de la comunidad. El tercer factor (items 2 y 1) se relaciona con la forma de aproximarse al gobierno y creer o confiar o no en él. Respecto al modelo anterior, se observa que el item 11 que habían sido incluido en el cuarto factor y rezaba sobre la posibilidad de que los políticos pusieran al país por encima de sus intereses personales ahora se encuentra en el segundo factor para este modelo de tres factores, lo cual no resulta incoherente si se considera el resto de los items que forman parte de esta factor, representado en su mayoría por actitudes positivas hacia el gobierno.

La matriz de correlaciones también indica que existe una correlación negativa entre el primer y segundo factor y una correlación positiva entre el primero y el tercero. Adicionalmente, el factor dos presenta una correlación negativa con el tercer factor.

Figura 13

Matriz de correlaciones para el modelo de cuatro factores

with factor correlations of			
	MR1	MR2	MR3
MR1	1.00	-0.55	0.69
MR2	-0.55	1.00	-0.45
MR3	0.69	-0.45	1.00

Nota. Tomado de la salida del paquete estadístico RStudio.

La complejidad promedio del ítem es de 1.2 y el test de hipótesis para tres factores muestra que esta cantidad de factores es suficiente para el modelo analizado. El contraste de hipótesis llevado a cabo a través del estadístico Chi-Cuadrado, el cual compara la matriz de covarianza observada con la matriz de covarianza reproducida por el modelo factorial, arrojó un valor de $X^2_{(105)} = 8055.22$, $p < .001$, lo que al ser un valor tan alto sugiere que el modelo no presenta calidad de ajuste respecto a los datos. Sin embargo, debe tomarse en cuenta que este estadístico se ve fuertemente distorsionado por el tamaño de la muestra, que al ser de gran tamaño, puede tener como resultado una gran influencia sobre el resultado de la prueba.

Por su parte, la raíz cuadrada media de los residuos (referida como RMSR) arrojó un valor de 0.03 que según el estándar muestra una buena o excelente calidad del ajuste al ser menor que .05 (pequeñas diferencias implican una magnitud menor de desvíos respecto del modelo). La raíz cuadrada media de los residuos corregida también muestra un buen o excelente ajuste al arroja un valor de 0.04, lo que sigue siendo menor a .05. Ambos valores son iguales a los presentados en el modelo anterior de cuatro factores.

La n observada armónica presenta un valor empírico igual al estadístico $X^2_{(105)} = 268.61$, con una probabilidad menor a $3.3e-27$, al igual que la media total de las observaciones con una probabilidad de Chi-Cuadrado de $X^2_{(105)} = 521.16$ con una probabilidad menor a $2.5e-73$, es decir, valores sumamente pequeños. En consonancia con lo anterior, el índice de Tucker Lewis para el caso tratado arrojó un valor de 0.904, lo cual es aceptable o razonable.

El error cuadrático medio de aproximación, también denominado índice RMSEA es ideal nuevamente para evaluar la calidad de ajuste del modelo y para los datos analizados presenta un valor de 0.078 e intervalos con un 90% de confianza que van desde 0.071 hasta 0.084, lo que se traduce en una medida de ajuste razonable considerando que arrojó valores mayores a .05 y menores o iguales a 0.08. Ahora bien, la medidas de ajuste bayesiana identificadas como BIC proporciona un valor igual a 75.01, lo que puede interpretarse como un número considerablemente elevado (incluso mucho más elevado que en modelo anterior de cuatro factores) que indica falta de calidad en el ajuste.

c. Con dos factores:

Finalmente, la iteración con dos factores, llevada a cabo con la finalidad de contrastar el modelo que muestre la mejor calidad de ajuste para las variables medidas, arroja los siguientes resultados a partir de la rotación oblicua Promax y suprimiendo las cargas menores a .30:

Figura 14

Matriz de cargas factoriales para la rotación promax de los datos con tres factores

Standardized loadings (pattern matrix) based upon correlation matrix						
	item	MR1	MR2	h2	u2	com
pol102	9	0.80		0.59	0.41	1.0
pol103	10	0.78		0.57	0.43	1.0
pol105	12	0.78		0.58	0.42	1.0
pol101	8	0.76		0.55	0.45	1.0
pol95	2	0.68		0.44	0.56	1.0
pol106	13	0.68		0.54	0.46	1.0
pol99	6	0.67		0.55	0.45	1.1
pol94	1	0.61		0.39	0.61	1.0
pol98	5	0.43		0.22	0.78	1.0
pol108	15		0.67	0.55	0.45	1.1
pol97	4		0.62	0.59	0.41	1.3
pol100	7		0.55	0.47	0.53	1.3
pol107	14		0.55	0.41	0.59	1.1
pol96	3		0.54	0.40	0.60	1.2
pol104	11		0.40	0.11	0.89	1.5

Nota. Tomado de la salida del paquete estadístico RStudio.

Como se puede observar en la Figura 14, el primer factor está constituido por 9 ítems con cargas factoriales entre .43 y .80, encontrándose que ninguno de estos ítems tienen cargas mayores a .30 en el segundo factor que constituye este modelo, lo que simplifica la lectura de esta solución factorial. Mientras tanto, el segundo factor está compuesto por 6 ítems, con cargas factoriales que van desde .40 hasta .67, siendo estas menores a las cargas de las variables en el primer factor. Sin embargo, al igual que esas variables, las variables que cargan en el segundo factor no presentan cargas mayores a .30 para el otro factor que forma parte del modelo, lo que indica parsimonia en la solución factorial.

En el primer factor que forma parte de este modelo de dos factores se agrupan los ítems que, en general, tienen como finalidad explorar las actitudes negativas de los encuestados hacia los políticos y el gobierno. Más específicamente, estos ítems expresan desconfianza hacia la información que provee el gobierno y los políticos y también hacia sus formas de actuar, la creencia de que los políticos y el gobierno no cuentan con las capacidades necesarias para cumplir con sus responsabilidades, la creencia de que el gobierno y los políticos carecen de respeto hacia la ciudadanía y además que el gobierno y los políticos actúan injustamente. Mientras tanto, el segundo factor de este modelo reúne los ítems que, en general, exploran actitudes positivas hacia el gobierno y los políticos. Más específicamente, estos ítems expresan la confianza hacia el gobierno y los políticos en cuanto a sus acciones, decisiones e intenciones, además, expresan la creencia de que los políticos y el gobierno ponen los interés y necesidades del país por encima de las de ellos.

Figura 15

Varianza total explicada por el modelo de 2 factores

	MR1	MR2
SS loadings	4.80	2.16
Proportion Var	0.32	0.14
Cumulative Var	0.32	0.46
Proportion Explained	0.69	0.31
Cumulative Proportion	0.69	1.00

Nota. Tomado de la salida del paquete estadístico RStudio.

En la Figura 15 se puede observar la distribución de la varianza de los factores que constituyen esta solución factorial de 2 factores. Como se puede evidenciar, el modelo es capaz de explicar el 46% de la varianza total, lo que es menor tanto al modelo de cuatro factores como al modelo de tres factores, los cuales explican el 53% y 51% de la varianza total respectivamente. Siendo esto así, a partir de la varianza explicada por el modelo ya se podría pensar que este modelo de dos factores probablemente no es la mejor opción.

En la Figura 16 se puede observar que para este modelo de dos factores se obtuvo una matriz de correlaciones entre ellos que da cuenta de la relación negativa existente entre ambos,

más específicamente, se da una relación inversa y moderada de $-.55$ entre el factor 1 y el factor 2, lo que tiene sentido entendiendo que el primer factor agrupa ítems que evalúan actitudes negativas y el segundo factor agrupa ítems que evalúan actitudes positivas, tal como fue descrito anteriormente.

Figura 16

Matriz de correlaciones entre los dos factores del modelo

```
With factor correlations of
      MR1    MR2
MR1  1.00 -0.55
MR2 -0.55  1.00
```

Nota. Tomado de la salida del paquete estadístico RStudio.

Ahora, en cuanto a la idoneidad de este modelo de dos factores a lo primero que se debe hacer referencia es a la complejidad promedio de los ítems, la cual es de 1.1, significando esto que, en promedio, los ítems tienden a cargar considerablemente en aproximadamente 1.1 factores, lo cual indica que la mayoría de ítems están relacionados principalmente con un solo factor, lo que es ideal en estos casos. El contraste de hipótesis llevado a cabo a través del estadístico Chi-Cuadrado, el cual compara la matriz de covarianza observada con la matriz de covarianza reproducida por el modelo factorial, arrojó un valor de $X^2_{(105)} = 8055.22$, $p < .001$, lo que al ser un valor tan alto sugiere que el modelo no presenta calidad de ajuste respecto a los datos. Sin embargo, debe tomarse en cuenta que este estadístico se ve fuertemente distorsionado por el tamaño de la muestra, que al ser de gran tamaño, puede tener como resultado una gran influencia sobre el resultado de la prueba.

Por su parte, la raíz cuadrada media de los residuos (referida como RMSR) arrojó un valor de 0.05 que según el estándar muestra una buena calidad del ajuste al ser igual que .05 (pequeñas diferencias implican una magnitud menor de desvíos respecto del modelo). La raíz cuadrada media de los residuos corregida también muestra un buen ajuste al arrojar un valor de 0.05, lo que también es igual al valor estándar de .05.

La n observada armónica presenta un valor empírico de $X^2_{(105)} = 511.05$, con una probabilidad menor a $1.1e-65$ y la media total de las observaciones presenta un valor Chi-Cuadrado de $X^2_{(105)} = 864$ con una probabilidad menor a $6.3e-134$, es decir, valores sumamente pequeños, lo que podría indicar un pobre ajuste del modelo. De nuevo, es importante tener en cuenta que este estadístico es altamente sensible a la muestra. Mientras tanto, el Índice Tucker Lewis arrojó un valor de .863, lo cual no es ideal entendiendo que son preferibles valores mayores a .90 e incluso .95. Finalmente, el error cuadrático medio de aproximación, también denominado índice RMSEA es ideal nuevamente para evaluar la calidad de ajuste del modelo y para los datos analizados presenta un valor de 0.093 e intervalos con un 90% de confianza que van desde 0.088 hasta 0.099, lo que se traduce en una medida de ajuste mediocre considerando que arrojó valores mayores a .08.

d. Con tres factores modificados (eliminación de ítems):

Como se habrá hecho evidente hasta este punto, el modelo de 3 factores parece ser la opción que teórica y estadísticamente reúne las mejores cualidades para explicar el constructo medido a través de las variables incluidas en el análisis. Aunque el modelo de 4 factores cumple con la mayoría de indicadores de buen ajuste y además explica la mayor cantidad de varianza de los tres modelos computados, en este modelo se presenta un factor de una sola variable, lo cual es una debilidad importantísima en términos de la parsimonia del modelo. Por otro lado, a pesar de que el modelo de 2 factores sí puede ser considerado parsimonioso en tanto está compuesto por dos factores que se diferencia claramente entre sí, este cuenta mayoritariamente con indicadores de ajuste pobres, lo que lo hace el menos idóneo de los tres modelos.

Teniendo esto en cuenta, y en aras de mejorar aún más el modelo de tres factores, los resultados que se presentarán a continuación representan una segunda iteración de este modelo pero con la eliminación de dos variables, a saber, el ítem 104 y el ítem 98, los cuales son los ítems con menores cargas factoriales del modelo y en sus respectivos factores y además son los ítems con menores comunalidades del modelo, representando así los eslabones más débiles de él.

La segunda iteración del modelo de tres factores, esta vez eliminando del conjunto de variables al ítem 104 y 98, arroja los siguientes resultados a partir de la rotación oblicua Promax:

Figura 17

Matriz de cargas factoriales para el modelo de tres factores menos dos ítems

	ítem	MR1	MR2	MR3	h2	u2	com
pol103	9	0.80			0.60	0.40	1.0
pol106	11	0.77			0.60	0.40	1.1
pol102	8	0.75			0.60	0.40	1.1
pol105	10	0.74			0.60	0.40	1.0
pol101	7	0.68			0.55	0.45	1.1
pol99	5	0.42			0.53	0.47	2.0
pol108	13		0.78		0.57	0.43	1.2
pol97	4		0.76		0.62	0.38	1.0
pol107	12		0.64		0.42	0.58	1.1
pol100	6		0.63		0.46	0.54	1.0
pol96	3		0.62		0.42	0.58	1.3
pol95	2			0.90	0.77	0.23	1.0
pol94	1			0.60	0.51	0.49	1.1

Nota. Tomado de la salida del paquete estadístico RStudio.

Como se puede observar en la Figura 17, el primer factor está compuesto por 6 ítems con cargas factoriales que van desde .42 hasta .80. En este factor se agrupan los ítems que tienen como finalidad evaluar las creencias de que los políticos y los gobiernos no son de fiar, de que estos se caracterizan por ser incompetentes, ineficaces y por tener mal juicio, y de que tanto los políticos como el gobierno no se interesan por la ciudadanía. Mientras tanto, el segundo factor reúne 5 ítems con cargas entre .62 y .78, los cuales buscan explorar, en general, actitudes y creencias positivas hacia los políticos y el gobierno, a saber, la creencia de que los políticos y el gobierno saben lo que hacen, de que entienden las necesidades del pueblo, y de que son honestos en cuanto a sus acciones e intenciones. Finalmente, el tercer factor agrupa dos ítems, el 95 y 94, cuyas cargas son .60 y .90 respectivamente, y estos, a diferencia del resto de ítems, evalúan las actitudes propias del encuestado en relación a los políticos y no la creencia general hacia ellos, más específicamente, estos ítems evalúan la seguridad del encuestado para creer en los políticos y además del cuidado del encuestado para confiar en ellos.

Figura 18

Varianza total explicada por el modelo de 3 factores menos dos ítems

	MR1	MR2	MR3
SS loadings	3.25	2.54	1.47
Proportion Var	0.25	0.20	0.11
Cumulative Var	0.25	0.45	0.56
Proportion Explained	0.45	0.35	0.20
Cumulative Proportion	0.45	0.80	1.00

Nota. Tomado de la salida del paquete estadístico RStudio.

En la Figura 18 se puede observar la distribución de la varianza para esta nueva iteración del modelo de 3 factores con dos ítems menos, evidenciándose que hay un aumento importante en la varianza total explicada, la cual es de 56%, mejorando así la varianza total explicada por todos los modelos computados hasta ahora. Específicamente para este modelo, el primer factor explica el 25% de la varianza, el segundo factor explica el 20%, y el tercer factor explica el 11%.

Figura 19

Matriz de correlaciones entre los tres factores del modelo con dos ítems menos

With factor correlations of			
	MR1	MR2	MR3
MR1	1.00	-0.66	0.66
MR2	-0.66	1.00	-0.55
MR3	0.66	-0.55	1.00

Nota. Tomado de la salida del paquete estadístico RStudio.

Como se puede ver en la Figura 19, las correlaciones entre los factores de esta nueva iteración del modelo de 3 factores con dos ítems menos son entre moderadas y altas tanto positivas como negativas. La correlación entre el factor 1 y 2 es negativa e igual a -.66, lo que tiene sentido entendiendo que el factor 1 engloba creencias negativas hacia los políticos y el gobierno y el factor 2 hace lo contrario, evalúa actitudes positivas hacia ellos. Por otro lado, teniendo en cuenta que tanto el factor 1 como el factor 3 evalúan creencias y actitudes, ambas negativas, hacia los políticos y el gobierno, la correlación positiva entre ellos (.66) también tiene sentido.

Ahora, en cuanto a la idoneidad de este modelo de dos factores a lo primero que se debe hacer referencia es a la complejidad promedio de los ítems, la cual es de 1.1, significando esto que, en promedio, los ítems tienden a cargar considerablemente en aproximadamente 1.1 factores, lo cual indica que la mayoría de ítems están relacionados principalmente con un solo factor, lo que es ideal en estos casos. El contraste de hipótesis llevado a cabo a través del estadístico Chi-Cuadrado, el cual compara la matriz de covarianza observada con la matriz de covarianza reproducida por el modelo factorial, sugiere que el modelo de 3 factores es significativamente mejor que el modelo nulo, el cual tiene una función objetiva mucho mayor.

A su vez, la RMSR arrojó un valor de 0.03 que según el estándar (0.05) indica una excelente calidad del ajuste (pequeñas diferencias implican una magnitud menor de desvíos respecto del modelo). La raíz cuadrada media de los residuos corregida también muestra un excelente ajuste al arrojar un valor de 0.04, lo que también es menor al valor estándar de .05.

La n observada armónica presenta un valor empírico de $X^2_{(105)} = 157.61$, con una probabilidad menor a $2.8e-15$ y la media total de las observaciones presenta un valor Chi-Cuadrado de $X^2_{(105)} = 385.56$ con una probabilidad menor a $4.4e-57$, es decir, valores sumamente pequeños, lo que podría indicar un pobre ajuste del modelo. De nuevo, es importante tener en cuenta que este estadístico es altamente sensible a la muestra. Mientras tanto, el Índice Tucker Lewis arrojó un valor de .915, lo cual indica un buen ajuste por estar por encima de .90. Finalmente, el RMSEA presenta un valor de 0.083 e intervalos con un 90% de confianza que van desde 0.075 hasta 0.091, lo que se traduce en una medida de ajuste mediocre considerando que arrojó valores mayores a .08.

11. Conclusión

Queda en evidencia como las posibles iteraciones de los modelos de Análisis Factorial Exploratorio realizadas para explicar el constructo *Interés sobre la Política* a través de la Encuesta Mundial de Valores llevada a cabo en Venezuela presentan ciertas fortalezas y retos desde el punto de vista estadístico y teórico. A grandes rasgos, se puede concluir que ninguno de los modelos muestra una idoneidad lo suficientemente clara a partir de la evidencia recopilada

como para seleccionar uno de los modelos con un número de factores determinados sin que quede lugar a duda de que puede ser utilizado para explicar el fenómeno estudiado. Incluso, esto ya se hace evidente a partir de los datos, a través de los cuales no se puede cumplir con el principio de normalidad multivariante. Sin embargo, si se lleva a cabo un análisis tanto cualitativo como cuantitativo de los resultados obtenidos, se puede determinar que el modelo de tres factores puede ser el más parsimonioso, coherente y provechoso en cuanto a los indicadores que se ven reflejados en su estructura interna del AFE, así como también sus indicadores a nivel estadístico, los cuales reflejan la calidad de su ajuste.

Cabe destacar que deben usarse los datos con precaución, ya que la aproximación al fenómeno es solo de carácter parcial y no pretende ni tiene la evidencia suficiente como para ser exhaustivo de la variable que se estudia. No obstante, el estudio realizado presenta información valiosa a ser tomada en cuenta en futuras investigaciones para seguir ganando comprensión sobre el fenómeno político en Venezuela y cómo es percibido por sus ciudadanos.

Referencias

- Kaiser, H. F. (1974). An index of factorial simplicity. *Psychometrika*, 39, 31-36. DOI: 10.1007/BF02291575
- Fabrigar, L. R., y Wegener, D. T. (2012). *Exploratory factor analysis*. Oxford University Press.
- Watkins, M. (2018). Exploratory Factor Analysis: A Guide to Best Practice. *Journal of Black Psychology*, 44(3), 219-246. DOI: 10.1177/0095798418771807
- Gorsuch, R. L. (1983). *Factor analysis* (2da Ed.). Erlbaum.
- Cattell, R. B. (1966). The scree test for the number of factors. *Multivariate Behavioral Research*, 1, 245-276. DOI:10.1207/s15327906mbr0102_10
- Martínez, R. (1996). *Psicometría: Teoría de los tests psicológicos y educativos*. Editorial Síntesis, S. A.
- World Values Survey Association. (2020). *Who we are*.
<https://www.worldvaluessurvey.org/WVSContents.jsp>