

Informe Estandarización Perú Escala INDI, Parte 2: Análisis de Ítems

Muestra Nivel 3

Martín Vargas Estrada

2025-01-17 22:20:40.001102

Índice

Introducción	2
Profundización Análisis Factorial Confirmatorio (AFC)	2
Resumen de Índices de Bondad de Ajuste para el AFC de la Escala INDI	2
Análisis de Ítems: Outliers y Transformaciones de Escala	5
Análisis de Outliers	5
Resumen de Outliers por ítem	7
Tabla de Outliers según Categoría	7
Tabla de Outliers según Categoría	7
Transformaciones de las Escalas	7
Transformación Logarítmica	8
Transformación Box-Cox	8
Conclusiones	8
Análisis de Ítems	8
Conclusiones Escala C	9
Conclusiones Escala M	9
Conclusiones Escala S	10
Conclusiones Escala D	10

Introducción

Informe de Exploración Psicométrica de los ítems de la prueba INDI obtenidas con muestra de Perú, Nivel 3.

Profundización Análisis Factorial Confirmatorio (AFC)

Como se detalla en secciones previas, el FC del modelo resultó con bajo ajuste en la mayoría de índices; por lo tanto se pasará a profundizar en el análisis a fin de obtener el detalle de los ítems cuya modificación (o incluso supresión) podría generar una mejora en el ajuste del modelo.

Resumen de Índices de Bondad de Ajuste para el AFC de la Escala INDI

Tabla 1: Resumen de Índices de Bondad de Ajuste para el AFC de la Escala INDI

	Índice.Obtenido	Nivel.Obtenido
RMSEA	0.0875521	Deficiente
CFI	0.7438284	Deficiente
TLI/NNFI	0.7315567	Deficiente
SRMR	0.0812253	Aceptable
GFI	0.6399880	Deficiente

Tabla 2: Cargas Factoriales Estandarizadas

Item	Factor1	Factor2	Factor3	Factor4
C1	0.835	0.000	0.000	0.000
C2	0.811	0.000	0.000	0.000
C3	0.650	0.000	0.000	0.000
C4	0.791	0.000	0.000	0.000
C5	0.816	0.000	0.000	0.000
C6	0.814	0.000	0.000	0.000
C7	0.799	0.000	0.000	0.000
C8	0.706	0.000	0.000	0.000
C9	0.757	0.000	0.000	0.000
C10	0.853	0.000	0.000	0.000
C11	0.746	0.000	0.000	0.000
C12	0.591	0.000	0.000	0.000
C13	0.528	0.000	0.000	0.000
C14	0.677	0.000	0.000	0.000
C15	0.749	0.000	0.000	0.000
C16	0.650	0.000	0.000	0.000
C17	0.714	0.000	0.000	0.000
C18	0.713	0.000	0.000	0.000
C19	0.638	0.000	0.000	0.000
C20	0.790	0.000	0.000	0.000
C21	0.816	0.000	0.000	0.000
C22	0.797	0.000	0.000	0.000
C23	0.738	0.000	0.000	0.000
M1	0.000	0.489	0.000	0.000
M2	0.000	0.720	0.000	0.000
M3	0.000	0.488	0.000	0.000

Item	Factor1	Factor2	Factor3	Factor4
M4	0.000	0.653	0.000	0.000
M5	0.000	0.742	0.000	0.000
M6	0.000	0.586	0.000	0.000
M7	0.000	0.670	0.000	0.000
M8	0.000	0.700	0.000	0.000
S1	0.000	0.000	0.782	0.000
S2	0.000	0.000	0.792	0.000
S3	0.000	0.000	-0.247	0.000
S4	0.000	0.000	-0.395	0.000
S5	0.000	0.000	-0.390	0.000
S6	0.000	0.000	-0.355	0.000
S7	0.000	0.000	0.629	0.000
S8	0.000	0.000	0.772	0.000
S9	0.000	0.000	0.598	0.000
S10	0.000	0.000	-0.114	0.000
S11	0.000	0.000	-0.202	0.000
S12	0.000	0.000	-0.384	0.000
S13	0.000	0.000	-0.339	0.000
D1	0.000	0.000	0.000	0.758
D2	0.000	0.000	0.000	0.582
D3	0.000	0.000	0.000	0.823
D4	0.000	0.000	0.000	0.715
D5	0.000	0.000	0.000	0.697
D6	0.000	0.000	0.000	0.795

Para entender lo que las tablas anteriores nos indican es básico tomar en consideración lo siguiente:

1. Aplicando los criterios generalmente aceptados para la interpretación de AFC, consideraremos que todas las cargas factoriales estandarizadas inferiores a 0.4 en valor absoluto indicarían “red flags”, es decir ítems que podrían considerarse para su eliminación o modificación.
2. Es importante tomar en cuenta que los resultados estadísticos son solo parte del proceso de decisión en relación a la gestión de los ítems. La otra parte es el contenido de los ítems. Un ítem puede tener carga factorial negativa pero el signo podría estar totalmente justificado si, al analizar el contenido, comprobamos que en efecto la idea original al redactar el ítem era precisamente evaluar aspectos opuestos al factor o variable latente que se pretende medir. Por ejemplo una escala para medir depresión podría contener un ítem como “A menudo me siento entusiasta al pensar en mi futuro”; sería perfectamente lógico y coherente esperar que tal ítem tuviera una carga factorial negativa, al tratarse de un ítem inverso.
3. La meta al redactar un ítem es describir la variabilidad en la variable a medir. En ese sentido, estadísticamente hablando, la carga factorial no es más que la correlación entre el ítem y la variable o factor latente que se pretende medir. Una consecuencia de plantearse las cargas factoriales de esa forma es que es posible establecer qué proporción de la variabilidad de los datos es explicada por el factor o variable latente que queremos medir. En términos matemáticos, es posible cuantificar esa proporción elevando al cuadrado la carga factorial del caso. Por ejemplo, si un ítem tiene una carga factorial de 0.62, entonces matemáticamente un 38.44 % de su variabilidad (0.62 al cuadrado) estaría siendo explicada por la variable a medir. Es por esto que queremos excluir ítems con carga factorial menores a 0.4; por ejemplo, un ítem con carga factorial de 0.25 solo sería explicado en un 6.25% por la variable a medir. En otras palabras, el 93.75% de los datos de tal ítem se deberían a alguna otra variable (o combinación de variables), distinta a la variable que queremos medir.

Ahora bien, sabedores de lo anterior, pasemos a señalar las conclusiones más saltantes:

1. Los ítems que componen la subescala “C” muestran cargas factoriales altas, lo cual indica que son

- estadísticamente válidos.
2. Los ítems de la subescala “M” también muestran cargas factoriales altas, lo cual indica que son estadísticamente válidos.
 3. Los ítems de la subescala “S” tienen resultados variopintos. Para empezar, los ítems S1, S2, S7, S8 y S9 se comportan como ítems directos (a mayor puntaje en el ítem, mayor valor o presencia de la variable a medir), mientras que los ítems S3, S4, S5, S6, S10, S11, S12, y S13 se comportan como ítems inversos. Todos los ítems directos muestran cargas factoriales aceptables, pero en especial los ítems inversos S3 y S11 parecen buenos candidatos para la supresión, mientras que S4, S5, S6, S10, S12, y S13 requerirían al menos la modificación con vistas a su mejoramiento.
 4. Entre los ítems de la subescala “D”, casi todos ostentan buenas cargas factoriales (siendo además todos ítems directos).

Ya desde el punto de vista conceptual, podemos pasar a analizar el *contenido* de los ítems en cuestión. En especial, nos centraremos en los ítems “problemáticos”.

Tabla 3: Cargas Factoriales Estandarizadas con Contenido

Item	Factor1	Factor2	Factor3	Factor4
C1 Comprende un cuento breve	0.835	0.000	0.000	0.000
C2 Cuenta una historia apoyándose en imágenes o dibujos	0.811	0.000	0.000	0.000
C3 Discrimina palabra corta - palabra larga (oralmente)	0.650	0.000	0.000	0.000
C4 Articula las palabras de manera clara	0.791	0.000	0.000	0.000
C5 Identifica colores primarios y secundarios	0.816	0.000	0.000	0.000
C6 Memoriza canciones y/o rimas	0.814	0.000	0.000	0.000
C7 Se comunica adecuadamente a través del lenguaje	0.799	0.000	0.000	0.000
C8 Permanece sentado, prestando atención a un cuento	0.706	0.000	0.000	0.000
C9 Inicia y detiene sus movimientos en respuesta a una instrucción (ej. Simón dice, Juego de las estatuas)	0.757	0.000	0.000	0.000
C10 Compara dimensiones o magnitudes (ej. grande-pequeño, lleno- vacío)	0.853	0.000	0.000	0.000
C11 Clasifica objetos según su forma y color	0.746	0.000	0.000	0.000
C12 Reconoce números entre el 1 y el 10	0.591	0.000	0.000	0.000
C13 Recita la serie numérica hasta 30	0.528	0.000	0.000	0.000
C14 Espera por algo que desea (ej. un juguete, lonchera) cuando se le solicita	0.677	0.000	0.000	0.000
C15 Se concentra cuando dibuja o pinta	0.749	0.000	0.000	0.000
C16 Es capaz de esperar turnos	0.650	0.000	0.000	0.000
C17 Cuenta objetos de una colección mayor a tres elementos	0.714	0.000	0.000	0.000
C18 Identifica y señala un cuadrado, un triángulo y un círculo	0.713	0.000	0.000	0.000
C19 En el juego utiliza los objetos de manera simbólica	0.638	0.000	0.000	0.000
C20 Comenta con detalles lo que le sucedió en un momento pasado (ej. el fin de semana, el día anterior)	0.790	0.000	0.000	0.000
C21 Infiere emociones básicas e intenciones de los personajes de historias o cuentos	0.816	0.000	0.000	0.000
C22 Comprende intenciones de los demás	0.797	0.000	0.000	0.000
C23 Se anticipa a lo que puede suceder	0.738	0.000	0.000	0.000
M1 Agarra el lápiz adecuadamente	0.000	0.489	0.000	0.000
M2 Rasga y troza papel	0.000	0.720	0.000	0.000
M3 Agarra la tijera adecuadamente	0.000	0.488	0.000	0.000
M4 Modela formas simples (ej. pelota, serpiente) y complejas (figura humana) con plastilina o arcilla	0.000	0.653	0.000	0.000
M5 Camina sobre una línea recta manteniendo el equilibrio	0.000	0.742	0.000	0.000
M6 Se desplaza saltando en un pie	0.000	0.586	0.000	0.000
M7 Lanza y recibe una pelota con sus manos	0.000	0.670	0.000	0.000
M8 Se viste y desviste por si solo	0.000	0.700	0.000	0.000

Item	Factor1	Factor2	Factor3	Factor4
S1 Tiene un buen relacionamiento con los compañeros de clase	0.000	0.000	0.782	0.000
S2 Ayuda a sus compañeros cuando lo necesitan	0.000	0.000	0.792	0.000
S3 Evita relacionarse con otros en diferentes situaciones	0.000	0.000	-	0.000
			0.247	
S4 Pasa el tiempo solo, aislado del grupo	0.000	0.000	-	0.000
			0.395	
S5 Agrede físicamente a sus compañeros (ej. escupe, pateo, empuja)	0.000	0.000	-	0.000
			0.390	
S6 Agrede verbalmente a sus compañeros (ej. insulta, grita, se burla)	0.000	0.000	-	0.000
			0.355	
S7 Reconoce y expresa asertivamente sus emociones	0.000	0.000	0.629	0.000
S8 Comparte juguetes y materiales	0.000	0.000	0.772	0.000
S9 Está atento a las emociones de los otros	0.000	0.000	0.598	0.000
S10 Se muestra tímido y/o no comparte sus experiencias con el grupo	0.000	0.000	-	0.000
			0.114	
S11 Tiene un aspecto triste, cansado y/o preocupado	0.000	0.000	-	0.000
			0.202	
S12 Tiene una postura desafiante hacia las maestras	0.000	0.000	-	0.000
			0.384	
S13 Se frustra y/o enoja con facilidad	0.000	0.000	-	0.000
			0.339	
D1 Se adapta a las rutinas de clase	0.000	0.000	0.000	0.758
D2 Busca ayuda cuando lo necesita	0.000	0.000	0.000	0.582
D3 Muestra curiosidad e interés durante las actividades	0.000	0.000	0.000	0.823
D4 Practica hábitos de aseo/higiene durante la jornada	0.000	0.000	0.000	0.715
D5 Se mueve con autonomía en el aula	0.000	0.000	0.000	0.697
D6 Persiste hasta completar una actividad	0.000	0.000	0.000	0.795

Podemos apreciar ahora que:

- Los ítems de la subescala “C” tienen validez aparente: su contenido parece relacionado con el factor a medir.
- Los ítems de la subescala “M” tienen validez aparente: su contenido parece relacionado con el factor a medir.
- Los ítems inversos de la subescala “S” parecen efectivamente medir la variable latente en sentido opuesto, por lo que es plenamente justificado que tengan cargas factoriales negativas. En ese sentido, evidencian validez aparente. Dicho esto, la mayoría de ítems inversos muestran también cargas factoriales relativamente bajas. En especial, S3 (“Evita relacionarse con otros en diferentes situaciones”); S10 (“Se muestra tímido y/o no comparte sus experiencias con el grupo”); y S11 (“Tiene un aspecto triste, cansado y/o preocupado”). El resto de ítems inversos, si bien se ubican en valores iguales o ligeramente por encima de 0.4 en valor absoluto, ciertamente evidencian una carga factorial menor a las sus contrapartes directas. Las razones de esta relativa ineficiencia de los ítems inversos de la subescala “S” puede deberse o bien a aspectos de contenido (los ítems no son adecuadamente interpretados o quizás haya un efecto cultural/lingüístico), o a aspectos metodológicos (puede que haya existido un efecto debido a la secuencia de administración).
- Los ítems de la subescala “D” parecen reflejar apropiadamente la dimensión a evaluar.

Análisis de Ítems: Outliers y Transformaciones de Escala

Análisis de Outliers

NULL Tabla consolidada de casos con valores atípicos

Columna	Fila	Valor_Outlier	Contexto
C13	42	6	Costa
C13	64	6	Costa
C13	82	6	Costa
C13	100	6	Costa
C13	158	6	Costa
C13	161	6	Sierra
C13	172	6	Costa
C13	199	6	Selva
C13	200	6	Selva
C13	201	6	Costa
C13	207	6	Costa
C13	210	6	Costa
C13	215	6	Costa
C13	306	6	Sierra
C13	401	6	Sierra
C13	507	6	Sierra
C13	599	6	Sierra
C13	785	6	Sierra
C13	980	6	Costa
S6	13	6	Costa
S6	200	6	Selva
S6	205	6	Costa
S6	265	6	Sierra
S6	345	6	Sierra
S6	346	6	Sierra
S6	367	6	Sierra
S6	406	6	Sierra
S6	428	6	Sierra
S6	435	6	Sierra
S6	543	6	Sierra
S6	785	6	Sierra
S6	828	6	Selva
S6	895	6	Selva
S6	911	6	Selva
S6	964	6	Sierra
S12	145	6	Sierra
S12	146	6	Sierra
S12	199	6	Selva
S12	200	6	Selva
S12	205	6	Costa
S12	346	6	Sierra
S12	367	6	Sierra
S12	406	6	Sierra
S12	435	6	Sierra
S12	554	6	Sierra
S12	620	6	Sierra
S12	628	6	Costa
S12	763	6	Sierra
S12	785	6	Sierra
S12	867	6	Sierra
S12	964	6	Sierra

Para resumir, podemos ofrecer las tablas siguientes, que resumen qué ítems muestran valores atípicos y cuántos casos hay, así como la frecuencia de valores atípicos según las categorías de Región (Costa, Sierra o Selva), y Departamento:

Resumen de Outliers por ítem

Columna	Valor_Outlier	Frecuencia
C13	6	19
S12	6	16
S6	6	16

Tabla de Outliers según Categoría

Contexto	Frecuencia
Sierra	28
Costa	15
Selva	8

Tabla de Outliers según Categoría

Contexto	Frecuencia
Cusco	24
Lima Metropolitana	14
Loreto	7
Piura	6

Las conclusiones apuntan a que la única escala que contiene valores extremos en la escala S, en especial los ítems S6 y S12. Hay que señalar que los parámetros elegidos para definir “valor extremo” (tres veces el rango intercuartílico, en lugar del nivel usual de rango y medio) son extremos en verdad; la distribución de los puntajes es bastante dispersa.

Una vez más, esto apunta en una dirección bastante clara: los subgrupos son tan distintos entre sí que incluirlos todos dentro de la misma “muestra general” hace que la variabilidad se eleve bastante y la aparición de outliers se haga más probable.

Finalmente, el total de los casos con valores extremos (51 casos) no llega a sobrepasar el 5% del total de la muestra, por lo que se considera que no valdría la pena tomar otras medidas de gestión de valores extremos, lo cual llevaría necesariamente al recálculo de todos los índices.

Transformaciones de las Escalas

Pasaremos a ejecutar dos transformaciones de la data. Las transformaciones son algoritmos que aplican modificaciones sistemáticas a cada uno de los datos a fin de obtener data transformada cuya distribución, potencialmente, podría acercarse a una distribución normal.

Ejecutaremos dos transformaciones: logarítmica y Box-Cox. Se eligió esas dos porque son las que usualmente generan los mejores resultados.

Transformación Logarítmica

Tabla 8: Transformación Logarítmica de las Escalas

Escala	Índice	p.valor	Interpretación
C_sum	S-W	0.00, ***	No Normalidad
C_sum	K-S, L	0.00, **	No Normalidad
M_sum	S-W	0.00, ***	No Normalidad
M_sum	K-S, L	0.00, ***	No Normalidad
D_sum	S-W	0.00, ***	No Normalidad
D_sum	K-S, L	0.00, ***	No Normalidad

Los resultados indican que, aun después de la transformación logarítmica, la distribución de los datos sigue sin satisfacer los criterios de normalidad.

Pasemos ahora a verificar los resultados de la transformación Box-Cox.

Transformación Box-Cox

A continuación, mostramos los resultados de la transformación Box-Cox.

Tabla 9: Transformación Box-Cox de las Escalas

Escala	Índice	p.valor	Interpretación
C_sum	S-W	0.00, ***	No Normalidad
C_sum	K-S, L	0.00, **	No Normalidad
M_sum	S-W	0.01, *	No Normalidad
M_sum	K-S, L	0.00, **	No Normalidad
D_sum	S-W	0.00, ***	No Normalidad
D_sum	K-S, L	0.00, ***	No Normalidad

Los resultados de la transformación Box-Cox tampoco logran que los datos transformados se acerquen significativamente a una distribución normal.

Conclusiones

- En general, se verifica que estas dos transformaciones sistemáticas de la data no logran acercarla significativamente a la normalidad.
- Si bien es posible seguir intentando con otros algoritmos (raíz cuadrada, inversa, etc.), es nuestra recomendación pasar al análisis ulterior, ya que en estos casos el tamaño muestral justifica matemáticamente el uso de herramientas paramétricas.
- Más allá de lo metodológico, y viendo la distribución gráfica los datos (Ver Parte 1), se recomienda verificar las razones por las que una muestra tan grande evidencia una distribución no normal.
- El presente análisis se planteará en lo subsiguiente enfocarse en esta cuestión, la cual se convierte en relevante desde el punto de vista conceptual.

Análisis de Ítems

En esta sección pasaremos a analizar la calidad de los ítems de la escala INDI, en su aplicación a la muestra peruana que nos ocupa.

Cabe mencionar que se recalculó los puntajes sumatorios en vista de que existen ítems inversos.

Luego se calculó el Índice de Discriminación de los ítems.

Conclusiones Escala C

Table 1: Índices de Discriminación Escala C

Ítem	Índice.D	Interpretación
C1	0.86	Excelente discriminación
C2	0.79	Excelente discriminación
C3	0.57	Excelente discriminación
C4	0.85	Excelente discriminación
C5	0.85	Excelente discriminación
C6	0.87	Excelente discriminación
C7	0.86	Excelente discriminación
C8	0.74	Excelente discriminación
C9	0.85	Excelente discriminación
C10	0.92	Excelente discriminación
C11	0.74	Excelente discriminación
C12	0.42	Excelente discriminación
C13	0.23	Discriminación aceptable
C14	0.73	Excelente discriminación
C15	0.82	Excelente discriminación
C16	0.75	Excelente discriminación
C17	0.79	Excelente discriminación
C18	0.80	Excelente discriminación
C19	0.70	Excelente discriminación
C20	0.82	Excelente discriminación
C21	0.81	Excelente discriminación
C22	0.82	Excelente discriminación
C23	0.70	Excelente discriminación

- Pese a estar constituida por la mayor cantidad de ítems en el INDI, la Escala C es la que ostenta los ítems con mayores niveles de discriminación.
- Los resultados, al igual que los demás análisis, apuntan a que hay muy poca distinción posible entre los ítems de esta escala.

Conclusiones Escala M

Table 1: Índices de Discriminación Escala M

Ítem	Índice.D	Interpretación
M1	0.43	Excelente discriminación
M2	0.76	Excelente discriminación
M3	0.47	Excelente discriminación
M4	0.55	Excelente discriminación
M5	0.76	Excelente discriminación
M6	0.52	Excelente discriminación
M7	0.66	Excelente discriminación
M8	0.66	Excelente discriminación

- En general, la Escala M contiene ítems con buenos niveles de discriminación.
- Tan solo los ítems M1 y M3 tienen un nivel menor de discriminación, de todos modos está dentro de un rango más que aceptable.

Conclusiones Escala S

Table 1: Índices de Discriminación Escala S

Ítem	Índice.D	Interpretación
S1	0.57	Excelente discriminación
S2	0.64	Excelente discriminación
S3_i	0.04	Mala discriminación
S4_i	0.07	Mala discriminación
S5_i	0.07	Mala discriminación
S6_i	0.06	Mala discriminación
S7	0.58	Excelente discriminación
S8	0.59	Excelente discriminación
S9	0.64	Excelente discriminación
S10_i	0.09	Mala discriminación
S11_i	0.05	Mala discriminación
S12_i	0.03	Mala discriminación
S13_i	0.02	Mala discriminación

- En general, tan solo los ítems S1, S2, S7, S8 y S9 tienen un buen nivel discriminación. El resto de ítems (todos ellos inversos) ofrecen un nivel discriminante bastante bajo.

Conclusiones Escala D

Table 1: Índices de Discriminación Escala D

Ítem	Índice.D	Interpretación
D1	0.54	Excelente discriminación
D2	0.34	Buena discriminación
D3	0.63	Excelente discriminación
D4	0.46	Excelente discriminación
D5	0.37	Buena discriminación
D6	0.64	Excelente discriminación

- En general, los ítems que componen la Escala D encuentra dentro de niveles aceptables. El resto de ítems ofrecen un nivel discriminante entre bueno y excelente.