Reporte del Modelo de medición para las escalas del Informe de responsabilidades profesionales

Servicio Profesional Docente 2017

Elaborado el 1 de septiembre de 2017

Unidad del Sistema Educativo Nacional

Dirección General de Medición y Tratamiento de Datos

Contenido

[Resultados 3](#_Toc492302116)

[Reactivos que conforman las Escalas 3](#_Toc492302117)

[Resumen de análisis psicométrico 3](#_Toc492302118)

[Modelo Factorial Confirmatorio 5](#_Toc492302119)

[El modelo de medición 8](#_Toc492302120)

[Parámetros del Modelo de Crédito Parcial (MCP) 8](#_Toc492302121)

[Ajuste del Modelo de Crédito Parcial (MCP) 11](#_Toc492302122)

[Mapa de Wright 13](#_Toc492302123)

[Nivel de cumplimiento 15](#_Toc492302124)

[Bibliografía 17](#_Toc492302125)

# Resultados

Las escalas reportadas presentan respuestas en las puntuaciones más altas, lo cual indica que se trata de cuestionarios con mucha deseabilidad. Esto tiene repercusión en la construcción de la medida, pues los reactivos se ubican en los puntajes mayores de la escala, así la población de ATP o AI ubicada en puntuaciones menores está caracterizada con mayor error dentro de la medida respecto a las puntuaciones mayores.

Es conveniente contar con preguntas en los cuestionarios que permitan discriminar a la población evaluada brindando descripciones en todo el continuo que conforma la medida.

# Reactivos que conforman las Escalas

Se construyeron las escalas del Informe de responsabilidades profesionales del Asesor técnico pedagógico (ATP) respondidas por el Asesor Técnico Pedagógico (ATP) y por la autoridad inmediata del ATP. En el documento se hace referencia a estas escalas con los nombres cortos “ATP” y “AI” respectivamente.

Cada una de las escalas se compone de 30 reactivos organizados en cinco bloques con 6 categorías de respuesta tipo Likert, denominadas del 0 al 5 en donde el cero es el menor valor y el 5 el mayor valor en sentido del constructo teórico.

Se considera como caso válido en los datos aquellos que presentan al menos una respuesta. Es asi que de los 2849 casos o folios de sustentantes reportados, la escala ATP se construyó con 2740 casos y la AI con 2748.

# Resumen de análisis psicométrico

En el archivo Excel se presentan los resultados del alfa de Cronbach, índice de discriminación, la proporción de respuesta y el valor pplus de cada reactivo. A continuación se presentan tablas resumen con los principales valores a destacar.

* La correlación de Pearson permite valorar la asociación entre pares de reactivos, estos valores se consideran aceptables cuando son superiores a 0.20.

|  |  |
| --- | --- |
| Escala | Matriz de correlación de Pearson |
| ATP | Tiene valores mayores a 0.20, excepto el reactivo ATP26 con ATP3, ATP6 y ATP9 |
| AI | Tiene valores mayores a 0.20. |

* La consistencia interna de los cuestionarios se valora con el estimador de alfa de Cronbach, el cual debe ser cercano a uno.

|  |  |
| --- | --- |
| Escala | Valor de alfa de Cronbach |
| ATP | Valor de 0.95, todos los reactivos contribuyen |
| AI | Valor de 0.97, todos los reactivos contribuyen |

* El índice de discriminación permite distinguir los reactivos a ATP o AI con alta puntuación de los reactivos ATP o AI con baja puntuación, sus valores deben estar entre 0.3 y 0.7 para indicar una buena discriminación del reactivo.

|  |  |
| --- | --- |
| Escala | Índice de discriminación |
| ATP | Valores en el rango de 0.39 a 0.68. |
| AI | Valores en el rango de 0.58 a 0.74. Valores arriba de 0.7 coloreados en rojo (17 reactivos). |

**Nota:** Todos los valores de p-plus están por arriba de 0.7. Lo cual indica que las respuestas están en las puntuaciones mayores.

# Modelo Factorial Confirmatorio

El análisis factorial confirmatorio es una forma de verificar el supuesto de unidimensionalidad de la escala. Se hizo este análisis para dos escalas, una para los reactivos del Informe de responsabilidades profesionales del Asesor técnico pedagógico (ATP) y la otra para el Informe de responsabilidades profesionales del ATP a responder por su autoridad inmediata (AI). A continuación se presentan los resultados del análisis. Cabe mencionar que los resultados de la estimación del modelo mostraron advertencias, los reactivos tienen categorías con menos del 4% de casos y la matriz de correlación en la iteración inicial del algoritmo de estimación es no definida positiva.

El modelo correlacionado del que se habla en este apartado corresponde al análisis factorial confirmatorio de ambas escalas (ATP y AI) en donde se estima además la correlación entre ellas, el gráfico de este modelo se encuentra en este documento y en el archivo de Excel que forma parte de este reporte, dado que contiene un gran número de variables observadas es importatnte incorporar una imagen que permita visualizar en un tamaño adecuado la información que contiene, por tal motivo se incorpora la imagen en el archivo Excel. Este modelo se estima para indagar la asociación entre los dos cuestionarios evaluados mediante la estimación de la correlación entre lo que responden el ATP y su AI respecto al cuestionario que ambos responden cada quién desde su perspectiva.

El índice RSMEA es 0.053 lo que indica un ajuste cercano (Brown, 2015).

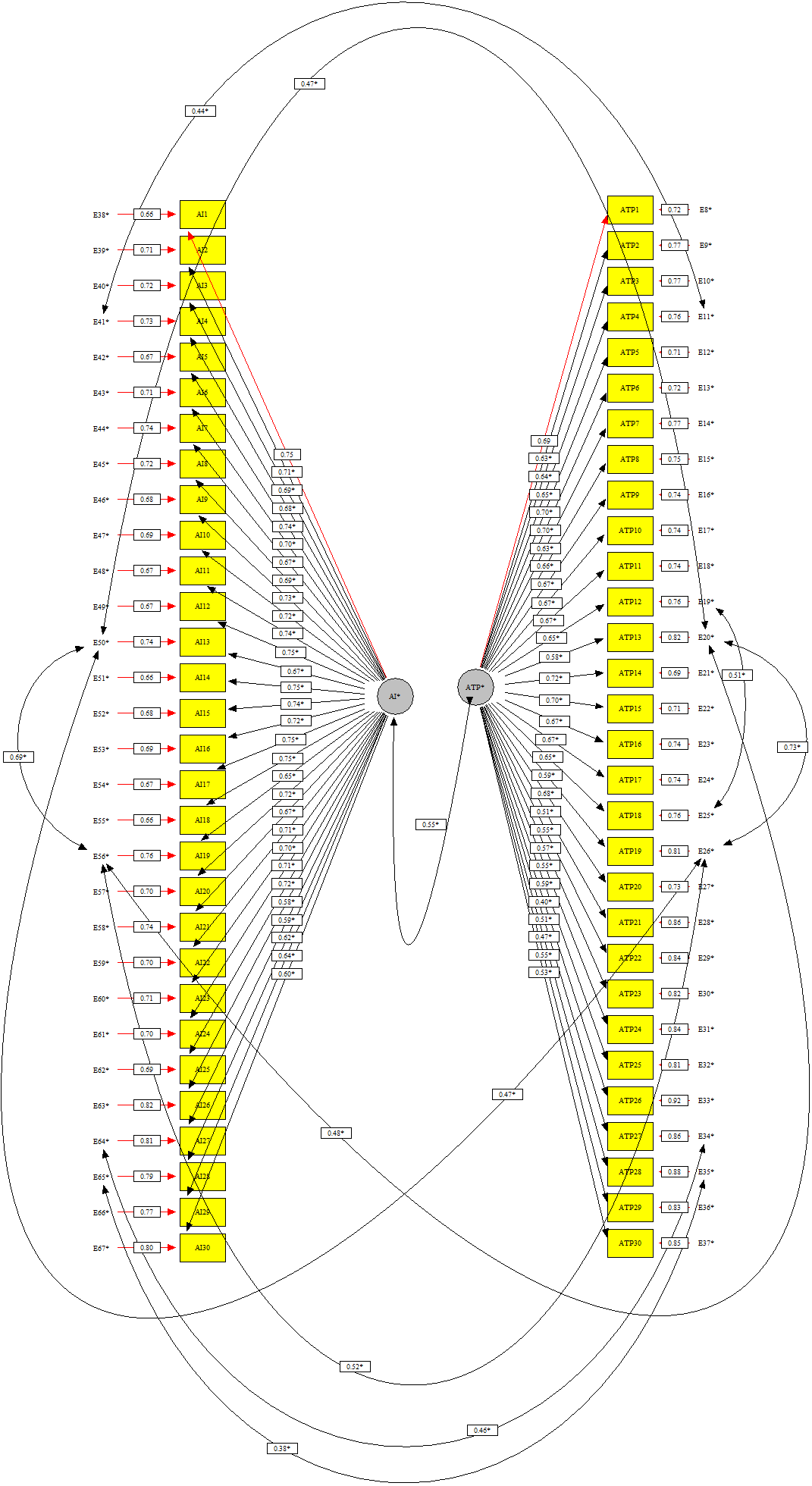
La correlación estimada entre las dos escalas (ATP y AI) es de 0.55, lo que en análisis de ciencias sociales es una correlación alta. En la siguiente tabla se muestran las correlaciones entre los errores estimados de los reativos que fueron particularmente altas según el modelo. Se observa que los reactivos 4, 13, 19, 27 y 28 de ambos cuestionarios se correlacionan, es decir, estas preguntas tienen respuestas con tendencia similar tanto para los ATP como para las AI. También es de observarse los reactivos 19 y 13, se correlacionan etre sí para ambas escalas, a diferencia de los reactivos ATP18 y ATP12 que en el modelo sólo hay correlación en la escala ATP. Es de notar la alta correlación entre los reactivos 13 y 19, que se correlacionan entre los reactivos de la misma escala y con los de la otra (ATP19-ATP13,AI19-AI13,AI19-ATP13, ATP19-AI13).

| **Reactivos** | **Correlación** |
| --- | --- |
| AI4-ATP4 | 0.438 |
| ATP18-ATP12 | 0.512 |
| ATP19-ATP13 | 0.727 |
| AI13-ATP13 | 0.472 |
| AI19-ATP13 | 0.478 |
| AI13-ATP19 | 0.474 |
| AI19-ATP19 | 0.523 |
| AI27-ATP27 | 0.463 |
| AI28-ATP28 | 0.375 |
| AI19-AI13 | 0.688 |

Las cargas factoriales del modelo en su mayoria son superiores a 0.5, lo que sugiere que los reactivos se asocian a un factor, es decir, a una escala. Hair, Black, Babin y Anderson (2009) recomiendan cargas factoriales mayor o igual a 0.5 y cargas entre 0.3 y 0.4 son mínimamente aceptables. A continuación se hace una breve descripción de las cargas factoriales por escala.

**Asesor técnico pedagógico:** La estimación de las cargas factoriales del modelo reflejan que la escala asociada a los reactivos de los ATP explica entre el 26 % y 51 % de la variabilidad de cada reactivo, esto al tener cargas entre 0.51 y 0.7 [[1]](#footnote-2), salvo los reactivos ATP26 y ATP28 que sus cargas factoriales son de 0.4 y 0.47, respectivamente. En la mayoría de los reactivos (ATP1-ATP4, ATP6-ATP13, ATP16-ATP30) la variabilidad explicada por el modelo es menor a la variabilidad no explicada por el modelo, es decir, de estos reactivos se explica menos del 50% de su variabilidad y la estimación de su error es más grande que el valor de la carga factorial.

**Autoridad inmediata del asesor téncnico pedagógico:** La estimación de las cargas factoriales del modelo reflejan que la escala explica entre el 34 % y 56 % de la variabilidad de cada reactivo, esto al tener cargas entre 0.58 y 0.75. En los reactivos AI4, AI7, AI8, AI13, AI19, AI21, AI26, AI27, AI28, AI29, AI30 la variabilidad explicada por el modelo es menor a la variabilidad no explicada por el modelo, es decir, de estos reactivos se explica menos del 50% de su variabilidad y la estimación de su error es más grande que el valor de la carga factorial.



El índice RSMEA es 0.053.

# El modelo de medición

Para formar la escala se empleó el Modelo de Rasch de Crédito Parcial (MCP) basado en la teoría de respuesta al ítem descrito por Wu y otros (2007), utilizando el Software ConQuest 4 (Adams, Wu, & Wilson, 2015). De manera similar a lo que se describe en el Reporte Técnico de PISA, Capítulo 16 (OECD, 2009).

# Parámetros del Modelo de Crédito Parcial (MCP)

Las tablas de la siguiente página presentan los resultados obtenidos del MCP que son: el número de casos, los valores infit y el puntaje promedio para cada reactivo por cada categoría. El reactivo ATP26 correspondiente a la pregunta 26 no presentó respuesta en la categoría 1. En cuanto a todos los demás reactivos, las primeras dos y en algunos hasta la tercer categoría, presentan menos de 10 respuestas lo que da lugar a que los umbrales sean probremente estimados y potencialmente inestables (Dimitrov, 2012). Esto último se puede verificar en el archivo Excel.

El infit es un estadístico que proporciona información sobre que tanto se ajustan los datos al modelo teórico. Se espera tener valores cercanos a uno, Wilson (2005) plantea que los valores aceptables para el infit deben estar en el rango entre 0.75 y 1.33; en este análisis se usarán como referencia estas cotas.

En cuanto al puntaje promedio por categorías, se espera que vayan creciendo conforme avanzan las categorías de respuesta. Esta condición es necesaria para evitar desorden entre las mismas. (Dimitrov, 2012).

**Cuestionario ATP**

Inicialmente el número de casos era de 2740 y el número mayor de valores pérdidos por reactivo fue 2. En el caso del infit todos los valores se encuentran dentro del rango excepto el reactivo ATP27 que tuvo un valor de 1.43 (diez décimas mayor al límite superior indicado 1.33). Respecto al puntaje promedio por categoría nueve reactivos presentan un comportamiento no adecuado en la categoría 2 (ATP5, ATP7, ATP9, ATP11, ATP17-ATP20 y ATP22), y uno en la categoría 1 (ATP8).



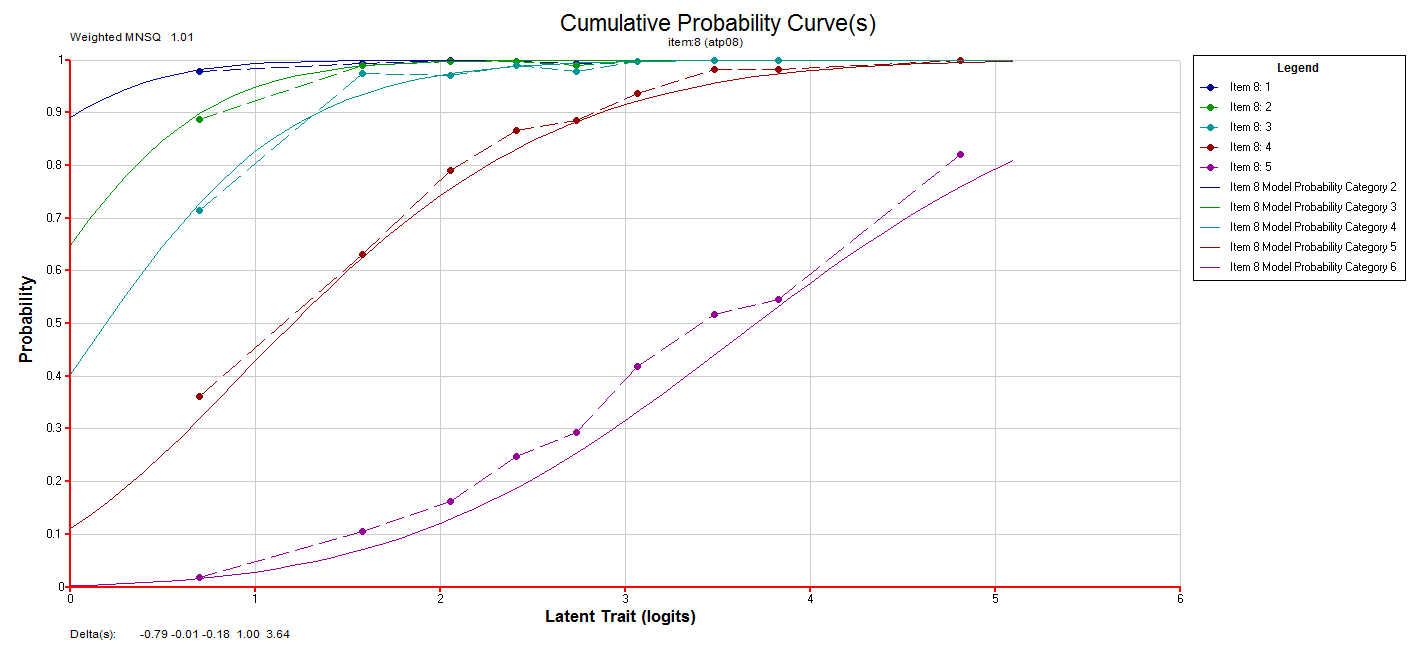
**Cuestionario AI**

Inicialmente el número de casos era de 2748 y el número mayor de valores pérdidos por reactivo fue 4. En el caso del infit todos los valores se encuentran dentro del rango excepto el reactivo AI27 que tuvo un valor de 1.46 (trece décimas mayor al límite superior indicado 1.33). ). Respecto al puntaje promedio por categoría cinco reactivos presentan un comportamiento no adecuado en la categoría 2 (AI2, AI9, AI14, AI21 y AI24).

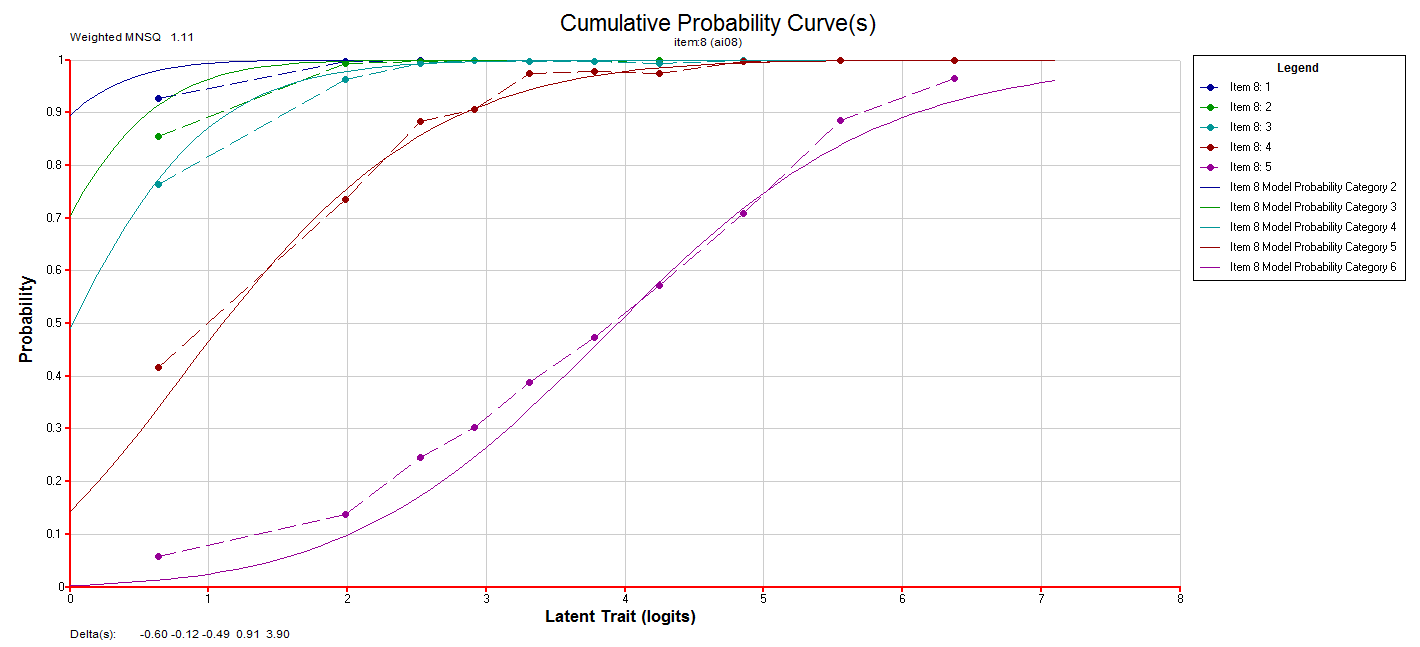


# Ajuste del Modelo de Crédito Parcial (MCP)

En las siguientes gráficas se presentan las curvas de probabilidad acumulada (CPC) de un reactivo de cada cuestionario como ejemplo; estas gráficas muestran la distribución de las categorías, donde las curvas separan las regiones de probabilidad, tal que se elige una categoría superior adyacente cuando se tiene mayor probabilidad de cambiar de categoría. En las gráficas se presentan las puntuaciones más bajas muy cercanas, lo cual indica que las puntuaciones 0, 1, 2 y 3 podrían ser parte de una sola categoría de respuesta. Esta situación está presente en todos los reactivos que conforman ambos cuestionarios.

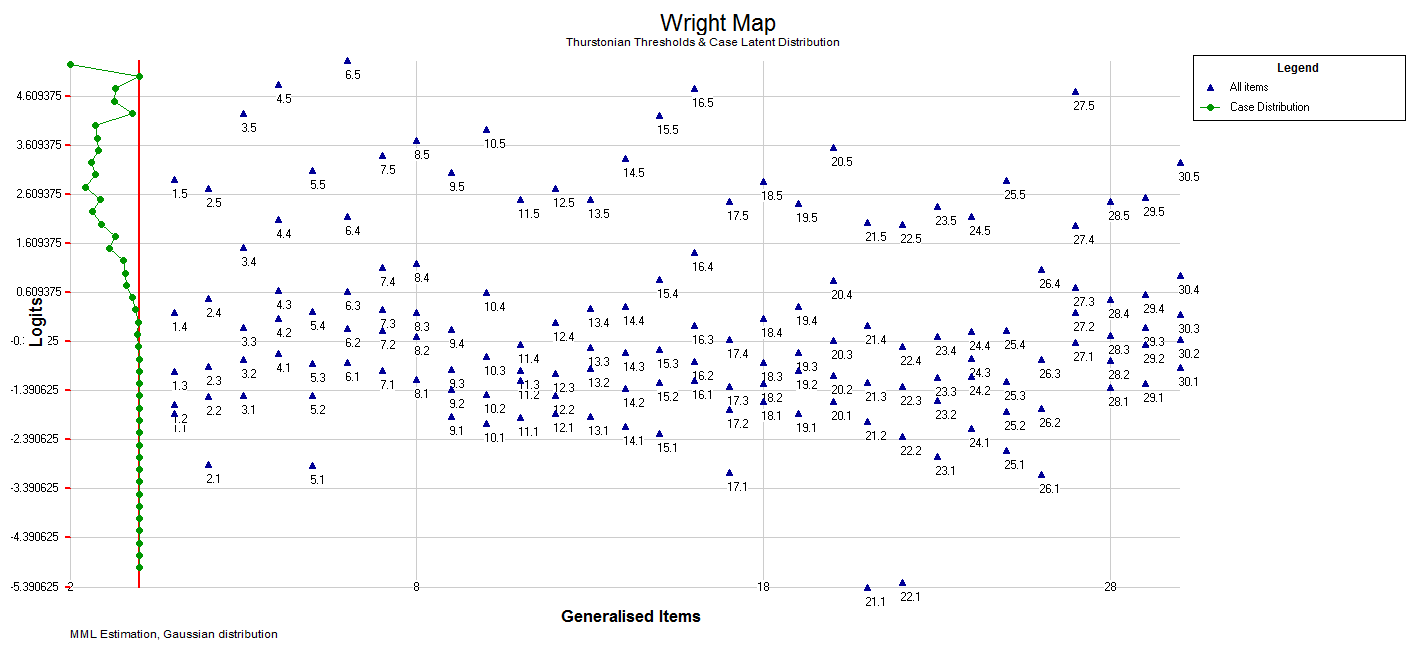
CPC del cuestionario ATP reactivo 8.

CPC del cuestionario AI reactivo 8.

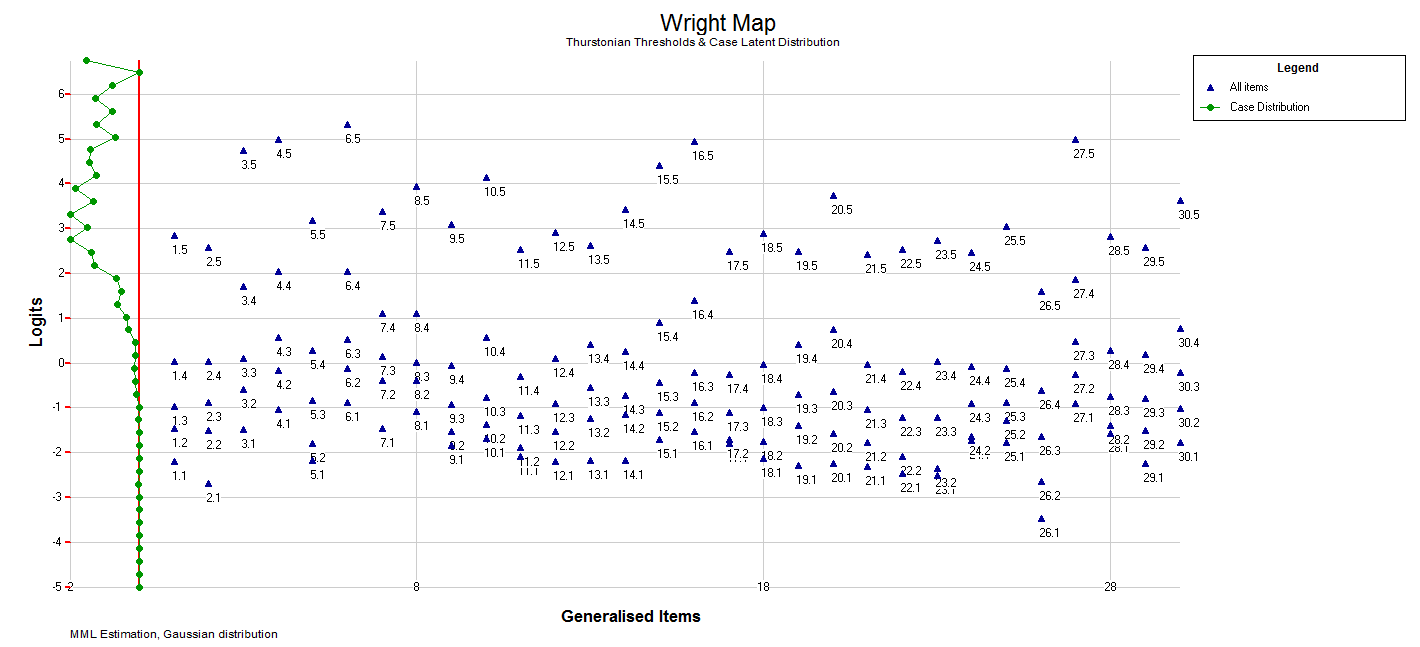


# Mapa de Wright

El siguiente mapa de Wright corresponde al cuestionario ATP del modelo de crédito parcial considerando una probabilidad de respuesta de 0.5 y con la escala en puntajes logit. La población de ATP se ubica en las puntuaciones mayores, en dirección a la mayor cantidad de constructo.



El siguiente mapa de Wright corresponde al cuestionario de AI del modelo de crédito parcial en puntuaciones logit.



# Nivel de cumplimiento

Cada una de las escalas construidas con puntuaciones logit obtenida por el modelo de crédito parcial se traslada a una puntuación tal que la población tiene media 1 y desviación estándar de 0.5 y está ubicada en el intervalo [0, 2]. Cuando puntuación trasladada sale del intervalo, se ubica dentro del intervalo [0,2] simplemento moviendo la puntuación al extremo correspondiente. La asignación del nivel para cada ATP se construye por la suma de las puntuaciones trasladadas de ambos instrumentos, la cual se ubica en el intervalo [0, 4] considerando que cada uno de ellos contribuye en igual magnitud. El Nivel asignado está constituido por cuatro categorías denotadas como: NI, NII, NIII y NIV

La asignación del nivel de cumplimiento deberá asignarse de acuerdo a la siguiente tabla:

|  |  |
| --- | --- |
| **Suma de puntuaciones de ambos instrumentos** | **Nivel de cumplimiento** |
| De cero a uno [0, 1] | NI |
| Mayor a uno hasta dos (1, 2] | NII |
| Mayor a dos hasta tres (2, 3] | NIII |
| Mayor a tres hasta cuatro (3, 4] | NIV |

Si un instrumento no fue presentado, se le asignó una puntuación de cero.

En el archivo de datos con las puntuaciones trasladadas y la clasificación de éstas en niveles se incluyen las variables empleadas durante el proceso de cada cuestionario distinguiéndolas por la terminación “9” y “10” en la clasificación que presenta el archivo la información recibida.

La siguiente tabla enlista y describe las variables agregadas al archivo de datos. Las que se distinguen por terminar en llevan la terminación “9” y “10” son: “delta”, “ee”, “media”, “de” y “deltay”, las otras son comunes.



Finalmente, en la siguiente tabla se muestra el número de sustentantes por nivel de cumplimiento y su respectivo porcentaje.



Siendo notable que el 56.32% de la población se encuentra en los dos primeros niveles.

# Bibliografía

Adams, R., Wu, M., & Wilson, M. (2015). *Acer ConQuest: Generalised Item Response Modelling Software.* [Computer software]. Version 4. Camberwell, Victoria: Australian Council for Educational Research.

Brown, T. (2015). *Confirmatory Factor Analysis for Applied Research* (Segunda edición ed.). New York : The Guilford Press.

Dimitrov, D. M. (2012). *Statistical Methods for Validation of Assessment Scale Data in Counseling and Related FIelds.* Alexandría : Wiley.

ETS. Educational Testing Service. (s.f.). *An adjustment for Sample Size in DIF Analysis*. Recuperado el 19 de Abril de 2016, de http://www.rasch.org/rmt/rmt203e.htm

Hair, J., Black, W., Babin, B., & Anderson, R. (2009). *Mutivariate Data Analysis* (Séptima edición ed.). Pearson.

OECD. (2009). *PISA 2006. Technical Report. Programme for International Student Assessment.* Paris: OECD.

Wilson, M. (2005). *Constructing Measures. An Item Response Approach.* Mahwah, New Jersey: Lawrence Erlbaum Associates, Inc., Publishers.

Wilson, M. (2005). *Constructing Measures: An Item Response Modeling Approach.* London: Lawrence Erlbaum Associates.

Wu, M. L., Adams, R. J., Wilson , M. R., & Haldane, S. A. (2007). *ACER ConQuest Version 2: Generalised item response modelling software.* Camberwell: Australian Council for Educational Research.

INSTITUTO NACIONAL PARA LA EVALUACIÓN DE LA EDUCACIÓN

Av. Barranca del Muerto 341, Col. San José Insurgentes, Deleg. Benito Juárez, C.P. 03900, México, D.F.

Tel.: (55) 5482 -0900

UNIDAD DE EVALUACIÓN DEL SISTEMA EDUCATIVO NACIONAL

DIRECCIÓN GENERAL DE MEDICIÓN Y TRATAMIENTO DE DATOS

1. La variabilidad del reactivo explicada por el modelo se obtiene al elevar al cuadrado la carga factorial asociada al reactive (Brown, 2015). [↑](#footnote-ref-2)