



*Universidad de Buenos Aires*  
*Facultad de Psicología*

**TESIS DE DOCTORADO**

**CONSTRUCCIÓN DE UN TEST  
DE COMPORTAMIENTO PROSOCIAL  
Y SU MODELIZACIÓN CON  
LA TEORÍA DE LA RESPUESTA AL ITEM**

**DOCTORANDA: Lic. SOFÍA ESMERALDA AUNÉ**

**DIRECTOR DE TESIS: Prof. HORACIO FÉLIX ATTORRESI**

**CO-DIRECTOR DE TESIS: Dr. FACUNDO JUAN PABLO ABAL**

**CONSEJERA DE ESTUDIOS: Prof. Dra. MARÍA SILVIA GALIBERT**

**Julio de 2017**

# *Dedicatoria*

*A mis padres Elba y Alberto  
A mi hermana del alma Julieta*

## *Agradecimientos*

Al Profesor Horacio Félix Attorresi, por haber confiado en mí, por ser un guía en el mundo de la Investigación y en la Vida. Por haber tenido un papel imprescindible para que esta Tesis sea una realidad, acompañando paso a paso la realización de la misma.

Al Dr. Facundo Juan Pablo Abal por transmitir de manera generosa y aportar para la realización de esta Tesis su amplio saber en Investigación Psicológica en general y en Psicometría en particular, siendo un especialista en Teoría de la Respuesta al Ítem.

Al Equipo de Investigación dirigido por el Profesor Horacio Félix Attorresi, por haberme acompañado en este proceso.

A la Universidad de Buenos Aires, por ser el marco en el que enseñé y aprendí, y en el cual se realizó este Doctorado.

A la Agencia Nacional de Promoción Científica y Tecnológica, por otorgarme una Beca Inicial de Doctorado.

Al Consejo Nacional de Investigaciones Científicas y Técnicas, por otorgarme una Beca de Finalización de Doctorado.

## Resumen

La presente Tesis tuvo como objetivo general la construcción de una escala para medir conducta prosocial en estudiantes terciarios y universitarios y su modelización con la Teoría de la Respuesta al Ítem (TRI). Este instrumento fue construido a partir de una exhaustiva revisión bibliográfica y depurado por juicio experto e indicadores de la Teoría Clásica de Test. Está conformado por 15 ítems que evalúan la conducta prosocial a través de dos dimensiones correlacionadas: Ayuda y Confortar. Para su validación se utilizó una muestra de 1540 participantes, de los cuales se depuraron los sujetos con alta deseabilidad social y los que fueron detectados como casos anómalos. Los datos se analizaron desde el marco de la TRI con los Modelos Logísticos de Uno, Dos y Tres Parámetros, el Modelo de Respuesta Graduada y el Modelo de Crédito Parcial. Los ítems ajustaron de forma aceptable a dichos modelos. Mientras que la Dimensión Ayuda resultó más informativa para niveles medio-altos o altos del rasgo, la Dimensión Confortar fue más informativa para niveles bajos. Se aportaron además evidencias adicionales de confiabilidad y de validez. Se realizó una discusión teórica y metodológica de los resultados obtenidos. Por último, se expusieron las limitaciones del estudio y posibles líneas de investigación futuras.

## Abstract

The general objective of this Thesis was the construction of a scale to measure prosocial behavior in tertiary and university students and its modeling with the Item Response Theory (IRT). This instrument was constructed from an exhaustive bibliographical review and debugged by expert judgment and Classical Test Theory indicators. It is made up of 15 items that evaluate prosocial behavior through two correlated dimensions: Helping and Comfort. For its validation a sample of 1540 participants was used, of which the subjects with high social desirability and those that were detected as outliers were deleted. Data were analyzed from the IRT framework with the One, Two and Three Parameters Logistic Models, the Graded Response Model and the Partial Credit Model. The items adjusted in an acceptable manner to these models. While Helping dimension proved more informative for medium-high or high trait levels, Comfort dimension was more informative for low trait levels. Further evidence of reliability and validity was provided. A theoretical and methodological discussion of the results obtained was made. Finally, the limitations of the study and possible future lines of research were explained.

## ÍNDICE

Dedicatoria	2
Agradecimientos	3
Resumen	4
<i>Abstract</i>	5
Introducción	10

### PRIMERA PARTE

#### ANTECEDENTES TEÓRICOS Y METODOLÓGICOS PARA LA MEDICIÓN DE LA CONDUCTA PROSOCIAL

1. Estado del Arte de la Conducta Prosocial	14
1.1. Qué es la Conducta Prosocial	14
1.2. Tipos de Conducta Prosocial	18
1.3. Variables Relacionadas con la Conducta Prosocial: Consideraciones Generales	21
1.3.1. Variables Sociodemográficas y su Relación con la Conducta Prosocial	23
1.3.1.1. La Conducta Prosocial y los Ciclos Vitales	23
1.3.1.2. La Conducta Prosocial y el Género	26
1.3.1.3. Conducta Prosocial: Su relación con el Ámbito Académico y la Profesión	27
1.3.2. Algunas Variables Disposicionales Relacionadas con la Conducta Prosocial: Empatía, Agresividad, Felicidad y Gratitud	28
1.3.2.1. La Empatía	28
1.3.2.2. La Agresividad	32
1.3.2.3. La Felicidad	33
1.3.2.4. La Gratitud	37
1.4. Motivaciones de la Conducta Prosocial	39
1.5. Conducta Prosocial y Cultura	41
1.6. Instrumentos para la Evaluación de la Conducta Prosocial	43

1.6.1. Instrumentos de Autoevaluación	43
1.6.2. Instrumentos de Heteroevaluación	46
1.6.3. Instrumento Mixto (Auto y Heteroevaluación)	48
2. Psicometría	49
2.1. Evaluación Psicológica	49
2.2. Teoría Clásica de los Test	51
2.2.1. Evidencias de Confiabilidad en el marco de la TCT	53
2.3. Teoría de la Respuesta al Ítem	54
2.3.1. Evidencias de Confiabilidad en el Marco de la TRI	57
2.3.2. Aplicaciones de la TRI a Tests de Comportamiento Típico	57
2.3.3. Modelos de la TRI para Ítems Dicotómicos	59
2.3.3.1. Modelo Logístico de Un Parámetro o Modelo de Rasch	63
2.3.3.2. Modelo Logístico de Dos Parámetros	65
2.3.3.3. Modelo Logístico de Tres Parámetros	66
2.3.3.4. Comparación entre los Modelos Logísticos de Uno, Dos y Tres Parámetros	67
2.3.4. Modelos de la TRI para Ítems Politómicos	68
2.3.4.1. Modelos de Respuesta Graduada	70
2.3.4.2. Modelo de Crédito Parcial	72
2.3.4.3. Comparación entre el Modelo de Respuesta Graduada y Modelo de Crédito Parcial	75
2.3.5. Estimación de Parámetros	76
2.3.6. Evaluación del Ajuste de los Datos a los Modelos	78
2.4. Funcionamiento Diferencial del Ítem	79
2.4.1. Métodos que Aplican la TRI	81
2.4.2. Métodos de Tablas de Contingencia	83
2.4.2.1. Prueba de Mantel-Haenszel	83
2.4.2.2. Pruebas de Breslow-Day	86
2.4.2.3. Diferencia de Proporciones Estandarizada (DPE)	88
2.4.3. TCT, TRI y Ausencia de DIF	89

SEGUNDA PARTE  
ESTUDIO PSICOMÉTRICO

3. Objetivos e Hipótesis	90
3.1. Objetivos	90
3.1.1. Objetivo General	91
3.1.2. Objetivos Específicos	91
3.2. Hipótesis	93
4. El Proceso de Construcción de la Escala de Conducta Prosocial	95
5. Metodología	96
5.1. Diseño General	96
5.1.1. Tipo de Estudio	96
5.1.2. Instrumentos	96
5.1.3. Pautas Éticas	97
6. Procedimientos y Resultados para la Concreción de los Objetivos Propuestos	99
6.1. Fase I: Construcción de Ítems, Estudio Piloto y Crítica de Jueces Expertos	99
6.2. Fase II: Análisis Factorial Exploratorio y Confirmatorio. Depuración Secundaria de Ítems.	102
6.3. Fase III: Estudios de Funcionamiento Diferencial del Ítem Según el Género	111
6.4. Fase IV: Modelización de la Escala de Conducta Prosocial Con la Teoría de la Respuesta al Ítem	117
6.5. Fase V: Fuentes Externas de Validez	200
6.6. Fase VI: Caracterización de los Participantes en su Conducta Prosocial	213

TERCERA PARTE

DISCUSIÓN

7.1. La Conducta Prosocial, Su Evaluación y Su Dimensionalidad	220
7.1.1. La Dimensión Confortar	222
7.1.2. La Dimensión Ayuda	224
7.2. Características Generales de los Participantes	225
7.3. Limitaciones y Futuras Líneas de Investigación	227

REFERENCIAS BIBLIOGRÁFICAS	229
APÉNDICES	245
Apéndice A	246
Apéndice B	249
Apéndice C	255
Apéndice D	259
Apéndice E	262
Apéndice F	263

*Las citas y referencias siguen los criterios planteados por las Normas A.P.A*

## INTRODUCCIÓN

La investigación en Psicología depende de cuestiones históricas y epistemológicas (Macbeth, Cortada de Kohan, Razumiejczyk, & López Alonso, 2006). En este sentido, la Psicología a lo largo de su historia se ha concentrado mucho más frecuentemente en las emociones negativas y las patologías mentales. Específicamente, después de la Segunda Guerra Mundial, las investigaciones se concentraron en el estudio de los desórdenes mentales. Es que muchas personas habían sido afectadas por la guerra y la incidencia de patologías se había convertido en un problema acuciante (Seligman, Parks, & Steen, 2004). Sin embargo, en los últimos años los investigadores comenzaron a interesarse en el bienestar subjetivo, siendo la llamada Psicología Positiva un área en crecimiento (Donaldson, Dollwet, & Rao, 2015) ya que se empezó a concebir que la ausencia de malestar no es igual a felicidad (Seligman, et al., 2004). Se podría estar en presencia de un cambio de paradigma, un cambio de gran alcance que implica una forma diferente de concebir la realidad (Macbeth, Cortada de Kohan, & Razumiejczyk, 2007). La Psicología Positiva es el marco necesario para realizar investigaciones que promuevan calidad de vida y bienestar (Mikulic et al., 2006). Además, se destaca el creciente énfasis en lo social de los nuevos desarrollos psicológicos (Macbeth et al., 2006).

La conducta prosocial es uno de los constructos centrales estudiados por la Psicología Positiva. La realización de conductas prosociales aumenta la posibilidad de generar una reciprocidad positiva y solidaria en las relaciones interpersonales o sociales. Es en la interacción con los otros donde los comportamientos se construyen, vinculándose desde el principio con las redes de relaciones de la persona (Crespi & Mikulic, 2009). Se incluye en el grupo de las capacidades para establecer vínculos interpersonales tales como el sentido del humor y la empatía, y se encuentra en relación con la madurez moral (Casullo & Fernández Liporace, 2007). Se ha comprobado que los efectos benéficos se extienden al actor de la misma y hacia la sociedad entera (Aknin, Broesch, Hamlin, & van de Vondervoort, 2015; Caprara, Alessandri, & Eisenberg, 2012).

Su delimitación conceptual es un tema arduo para la investigación psicológica (Caprara, Steca, Zelli, & Capanna, 2005). Esta dificultad deriva en parte de la complejidad del constructo, que, más allá de las definiciones consensuadas, genera debates en torno a cómo medirlo y cuáles son sus componentes (Batson, 1991; Eisenberg

& Fabes, 1998; Schroeder, Penner, Dovidio, & Piliavin, 1995). Hirschberger, Ein-Dor y Almakias (2008) destacan la ambivalencia implicada en la conducta prosocial, en tanto su realización podría satisfacer una gran variedad de necesidades humanas, reflejando algunas veces egoísmo y, en otras, genuina empatía. En este sentido, resulta especialmente difícil la distinción de la conducta prosocial con el altruismo (Batson & Powell, 2003).

Por otra parte, resulta contrastante la gran cantidad de investigaciones acerca de la conducta prosocial en la infancia y en la adolescencia con los escasos desarrollos enmarcados en etapas posteriores. Este hecho se refleja también en que sólo recientemente se han diseñado instrumentos para medir conducta prosocial en la etapa adulta. Podríamos afirmar, así, que el estudio de la prosocialidad en la adultez es un campo incipiente, con una gran variedad de aristas por explorar.

La propuesta que se presenta en esta Tesis Doctoral es, justamente, la construcción de un instrumento que mida conducta prosocial en la adultez. Dicho instrumento se modelizó con la Teoría de la Respuesta al Ítem (TRI). Gran parte de la evolución de la psicometría mundial se nutrió de las innovaciones teóricas y tecnológicas introducidas en los últimos cincuenta años por la TRI. Los modelos generados desde este enfoque psicométrico posibilitaron una manera alternativa a la perspectiva clásica de formalizar la relación existente entre la puntuación observada de un individuo en un test y la cantidad de rasgo que éste posee. En las últimas décadas, se han generado nuevos modelos para el análisis de ítems puntuados politómicamente y se ha buscado ampliar su aplicación a constructos evaluados mediante tests de comportamiento típico (Abal, Lozzia, Aguerri, Galibert, & Attorresi, 2010).

A pesar de que la TRI se encuentra en auge desde hace varios años en Europa y EE.UU., Argentina se ha mantenido al margen de muchos de estos nuevos desarrollos de la Psicometría Moderna. Dentro de los desarrollos nacionales, la utilización de distintos modelos de la TRI para la construcción de tests de comportamiento típico es aún un área muy poco explorada. En el aspecto metodológico, la aplicación de modelos de la TRI para ítems politómicos a una escala de conducta prosocial es inédita en el país.

La implementación de los modelos de la TRI en Argentina y en Latinoamérica en general es reciente y poco generalizada (Tornimbeni, Pérez, & Olaz, 2008) y más aún cuando se trata de tests de comportamiento típico, es decir de aquellos que evalúan formas usuales de comportarse (Abal, et al., 2010). En Colombia, Alfaro Viracachá y su equipo

utilizaron TRI en la construcción de pruebas educativas y de evaluación de méritos en general. En el mismo país Gómez Rada (2004) diseñó, construyó y validó un instrumento que evalúa clima organizacional en empresas colombianas desde la TRI. En Perú se realizó un análisis psicométrico del Test de Matrices Progresivas Avanzadas de Raven (Escurrea Mayaute & Delgado Vásquez, 2010) y del Cuestionario de Honey y Alonso de Estilos de Aprendizaje (Escurrea Mayaute, 2011) mediante TRI. Los chilenos Rojas, Manríquez, Gatica y Salcedo Lagos (2004) diseñaron una multiplataforma adaptativa basada en la TRI para la educación a distancia.

En la Argentina se destacan las contribuciones de Richaud de Minzi (2005, 2008, 2013), Cortada de Kohan (Cortada de Kohan & Cupani, 2008; Cortada de Kohan & Tornimbeni, 2004) y Attorresi y su equipo (Attorresi, Lozzia, Abal, Galibert, & Aguerri, 2009). Se ha aplicado la TRI para el estudio de la calidad de vida relacionada con la salud (Berra, Tebé, Esandi, & Carignano, 2013), para la implementación de la educación adaptativa a distancia (Rossel, 2006), para la creación de sistemas tutores inteligentes (Badaracco & Martínez, 2010), para la valoración psicométrica de la *Psychological Entitlement Scale* (Mola, Saavedra, Reyna, & Belaus, 2013) y para la modelización de un test de altruismo (Abal, Lozzia, Blum, Galibert, Aguerri, & Attorresi, 2010; Abal, Lozzia, & Galibert, 2008), entre otros usos.

En el aspecto metodológico Cervantes, Cepeda-Cuervo y Camargo (2008), investigadores colombianos, proponen un procedimiento para la selección de los niveles de desempeño en ítems de instrumentos contruidos con los modelos de la TRI. Cobra importancia en Uruguay la estimación no paramétrica de Curvas Características de Ítems mediante regresión no paramétrica (Fortaleza, Luzardo & Padula, 2010; Fortaleza, Padula & Luzardo, 2010). En Chile Asún y Zuñiga (2008) estudian las ventajas de los modelos politómicos de la TRI en la medición de actitudes sociales. En Argentina se ha realizado la modelización de un test de actitud hacia la matemática mediante los Modelos Politómicos de Respuesta Graduada (MRG) y Crédito Parcial (MCP) (Abal, 2013; Abal, Auné, & Attorresi, 2014; Abal, Auné, Lozzia, & Attorresi, 2015) y el estudio de la relación entre los parámetros de la TRI y la deseabilidad social (Abal, Auné, & Attorresi, 2015). Además, se ha aplicado la TRI al Funcionamiento Diferencial de los Ítems en estudios de simulación (Aguerri, 2012; Aguerri, Prieto-Marañón, Galibert, & Attorresi, 2013) con el fin de aportar evidencias acerca de qué métodos son más precisos para diferentes condiciones de aplicación.

El desarrollo de esta Tesis se ha organizado en tres partes. La primera recorre los antecedentes teóricos y metodológicos que sirven de fundamento para la elaboración de la Tesis. En la segunda parte se caracterizan y se aplican los métodos psicométricos propuestos para alcanzar los objetivos formulados. En la tercera parte se discuten los resultados obtenidos.

En la primera parte de la Tesis se desarrolla el marco teórico que contextualiza los aportes realizados. En el primer apartado se aborda el estado del arte del constructo conducta prosocial en forma detallada, en el segundo apartado se exponen los conceptos psicométricos fundamentales para la realización de esta Tesis con especial énfasis en la Teoría de Respuesta al Ítem, mientras que el último capítulo se refiere al Funcionamiento Diferencial de los Ítems. En la segunda parte se describe el proceso de la investigación que se llevó adelante, se especifican las pautas generales del diseño y se obtienen los resultados. En la tercera y última parte se realiza una discusión acerca de los resultados, relacionándolos con el marco teórico realizado en la primera parte.

**PRIMERA PARTE**

**ANTECEDENTES TEÓRICOS Y METODOLÓGICOS**

**PARA LA MEDICIÓN DE LA CONDUCTA PROSOCIAL**

*1. Estado del Arte de la Conducta Prosocial*

*1.1. Qué es la Conducta Prosocial*

El valor heurístico general de la conducta prosocial (Caprara et al., 2005) permite la relación con una diversidad de temáticas y conceptos. Es así que se ha generado una importante producción científica en los últimos años, vinculándola con variables novedosas y construyendo instrumentos para su evaluación (Auné, Abal & Attorresi, 2016b).

Se aplica el término prosocial a aquellos comportamientos realizados con la intención de beneficiar a otros (Dunfield, 2014; Eisenberg, Eggum, & Di Giunta, 2010). Son conductas voluntarias (Eisenberg & Fabes, 1998) que se adoptan para cuidar, asistir, confortar y ayudar (Caprara, et al., 2005). Puede diferenciarse de otros comportamientos porque es una respuesta basada en la observación e interpretación de las necesidades de otro individuo. Las personas se comprometen en actos prosociales si creen que podrán manejar las emociones asociadas al reconocimiento de las necesidades de otros y establecer relaciones y acciones conducentes a aliviar esas necesidades (Caprara et al., 2012).

Los actos orientados a los otros parecen ser una parte esencial y única de la vida social humana. Para Marean (2015) la tendencia a cooperar incluso con individuos no emparentados, es decir el logro de conexiones sociales flexibles, fue una de las claves para la supervivencia y la hegemonía del Homo Sapiens sobre otros Homos. Para este autor existiría incluso una predisposición genética que facilita este tipo de conductas.

La literatura coincide en señalar el origen de la investigación moderna en conducta prosocial en un hecho aparentemente insólito ocurrido en la ciudad de Nueva York el 13 de marzo de 1964: una mujer fue asaltada y asesinada en presencia de treinta y ocho personas que no intervinieron en la situación. Esto provocó un interés por conocer las causas esta falta de iniciativa (Gómez Jiménez & Gaviria Stewart, 2010). Al poco tiempo surgen los primeros trabajos sobre el altruismo. Darley y Latané (1968), describen el efecto espectador *-bystander effect-* según el cual existe una relación inversa entre el número de observadores de una situación de emergencia y la ayuda prestada. Es decir, cuantas más personas presencien una emergencia es menos probable que una de ellas efectivamente reaccione e intente resolver dicha situación. El desconcierto que se produce en la persona dificulta la interpretación de lo que acontece lo cual se combina con la dilución de responsabilidad entre los observadores y así nadie lleva a cabo la acción de ayudar.

Se encuentran en la bibliografía definiciones más estrictas y otras más amplias de conducta prosocial. En este trabajo se considera como definición estricta aquella que supone una cantidad extensa de requisitos para que una conducta sea de ese tipo. Un ejemplo es la definición de Olivar (1998), que entiende como conducta prosocial solamente a aquel comportamiento que favorece a otra persona, grupos o metas sociales y aumenta la probabilidad de generar una reciprocidad positiva, de calidad y solidaria en las relaciones interpersonales o sociales, salvaguardando la identidad, creatividad e iniciativa de las personas o grupos implicados. En cambio, las definiciones amplias son más simples y los requisitos para que una conducta sea clasificada como prosocial son menores. Es así que González Portal (2000) demarca el término en su acepción más abarcativa, entendiendo por tal a toda conducta social positiva con o sin motivación altruista. Esta definición pone el acento en la acción en sí misma, y deja en un segundo plano el análisis acerca de cuál es su motivación. La autora justifica dicha elección argumentando que los intentos de dilucidar si ha existido o no un móvil de tipo altruista por parte del benefactor se han visto irremediabilmente conducidos a un callejón sin salida (González Portal, 2000).

Para González Portal (2000) el término de conducta prosocial surgió, justamente, como alternativa históricamente justificada con el fin de romper con lo restrictivo del concepto altruismo, tan ligado desde sus orígenes a un criterio motivacional específico y del que aún hoy no ha podido desprenderse. De hecho, un programa de investigación

entero se basa en la hipótesis de la empatía-altruismo. La preocupación empática se concibe como una respuesta específica ante una situación en la cual un observador es testigo de una situación difícil por la cual está atravesando otra persona. Es así que se genera el comportamiento altruista, realizado principalmente como un intento de reducir el sufrimiento ajeno (Batson, 1991). El altruismo se considera como un resultado de anteponer la necesidad de los otros a la propia, persiguiendo el beneficio exclusivo del receptor y no del actor (Attorresi, Abal, Lozzia, Picón Janeiro, Aguerri, & Galibert, 2008; Howard & Piliavin, 2000). López (1994) diferencia los conceptos de altruismo y conducta altruista, definiendo el altruismo como una disposición, una orientación hacia el bien de los otros. La conducta altruista es toda acción voluntaria realizada con la intención de ayudar a los demás, provocando o manteniendo efectos positivos. No se realiza con la intención de recibir recompensas y debe implicar más costes que beneficios externos. Chacón (1986) y Hinde y Groebel (1991) también destacaron el sacrificio del benefactor como requisito para definir una conducta como altruista.

Según Batson y Powell (2003), la naturaleza misma del altruismo involucra una discusión teórica entre quienes afirman que beneficiar al otro no es más que un objetivo a mediano plazo en un camino en última instancia egoísta y quienes defienden la existencia del altruismo. Los primeros señalan los beneficios derivados de ayudar: el material, social, y las auto-recompensas recibidas; los castigos materiales, sociales y de conciencia evitados y la reducción de la aversión-excitación. Por ejemplo, Cialdini y sus colegas (Cialdini, Shaller, Houlihan, Arps, Fultz, & Beaman, 1987) han argumentado que la mayoría de los casos de comportamiento prosocial pueden ser explicados por motivos egoístas, tales como tratar de mejorar la propia imagen pública o aliviar un estado de ánimo negativo. Más recientemente, Solomon, Greenberg y Pyszczynski (1991) destacan en este sentido que proporcionar ayuda a los necesitados, especialmente a los que se considera dignos de ella dentro de la cultura, es un ejemplo de cómo cumplir con los estándares de valor provee a los individuos de un sentimiento de autoestima personal. Concordantemente, Salovey, Mayer y Rosenham (1991) puntualizaron que el comportamiento de ayuda puede beneficiar a los individuos que proveen asistencia, no sólo mejorando el estado de ánimo actual, sino también en el largo plazo.

Otros investigadores han propuesto que los motivos para el comportamiento prosocial están ligados a una necesidad de pertenencia. Desde esta perspectiva, las acciones prosociales podrían reforzar el sentimiento de comunidad, satisfaciendo los

deseos de ser valorado, necesitado y apreciado por un grupo social (Baumeister & Leary, 1995). Cuadrado, Tabernero y Steinel (2016) investigaron los efectos de la inclusión y exclusión social en los niveles de realización de conducta prosocial. En el caso de un individuo rechazado por un grupo social, tiene similares niveles en la variable cuando el beneficio se orienta a sujetos diferentes de los que han ejercido el rechazo, y niveles más bajos hacia el grupo que lo excluye en base a características personales. Por el contrario, si la exclusión se funda en la falta de competencia en una determinada tarea, percibe que la situación puede revertirse y realiza más conductas prosociales para, así, poder ser incluido. Cuando gozan de inclusión en el grupo, los individuos populares por personalidad realizan más conductas prosociales que aquellos populares porque se los reconoce capaces.

En cambio, los defensores de la existencia del altruismo sostienen que simplemente porque se deriven beneficios de ayudar a otro, esto no prueba que dichos beneficios son el objetivo final. Estos beneficios pueden ser consecuencias no deseadas de alcanzar el objetivo final de beneficiar a la otra persona. Si es así, la motivación podría ser altruista, no egoísta. Entonces, desde esta concepción, que toman para sí Batson y Powell (2003), lo que determina que una acción sea considerada altruista no es si se derivan o no beneficios de ella, sino cuál era su motivación inicial. Por lo tanto, para dichos autores, el altruismo es un concepto motivacional, siendo su meta aumentar el bienestar de otra persona. Contrariamente el egoísmo es la motivación cuyo fin es acrecentar el propio bienestar. Desde este punto de vista, no existiría una relación estricta entre el comportamiento prosocial y el altruismo, en el sentido de que la conducta prosocial no necesitaría estar motivada por el altruismo. Es así que toda conducta altruista puede ser considerada prosocial, pero no toda conducta prosocial puede ser considerada altruista (González Portal, 2000; López, 1994). También se ha distinguido el comportamiento prosocial altruista del basado en normas y del comportamiento prosocial estratégico (Böckler, Tusche, & Singer, 2016). En el ámbito local se han construido escalas para la medición de la motivación altruista coherentes con la conceptualización batsoniana (Abal, et al., 2010; Abal, Lozzia, & Galibert, 2008; Attorresi et al., 2008).

En el marco de esta Tesis, se entiende a la conducta prosocial como un fenómeno complejo que involucra acciones de los individuos basadas en creencias y sentimientos y que describe la forma en que éstos se orientan hacia los otros al realizar conductas solidarias.

En cuanto a la prevalencia de la conducta prosocial, se encuentran resultados diversos. Inglés et al. (2008) consideraron a un sujeto con conducta prosocial cuando excede un desvío y medio por encima de la media de puntaje y hallaron que la prevalencia fue del 17.35% en una muestra de estudiantes adolescentes españoles, mientras que Redondo Pacheco y Guevara Melo (2012) consideraron a un sujeto con conducta prosocial cuando excede el puntaje directo 70 y la estimaron en 34.9% en una muestra de adolescentes colombianos. Ambos estudios utilizaron el Inventario de Habilidades Sociales para Adolescentes (*Teenage Inventory of Social Skills* [TISS] de Inderbitzen & Foster, 1992).

### *1.2. Tipos de Conducta Prosocial*

Se han desarrollado diferentes tipologías de la conducta prosocial (Auné & Attorresi, 2014). Las primeras categorizaciones (González Portal, 2000; Olivar, 1998) se caracterizan por ser exhaustivas e intentar distinguir en detalle entre las diferentes conductas prosociales. De esta forma, Olivar (1998) distingue entre: Ayuda Física, Servicio Físico, Dar, Ayuda Verbal, Consuelo Verbal, Confirmación y Valorización Positiva del Otro, Escucha Profunda, Empatía, Solidaridad y Presencia Positiva y Unidad. Se puede observar la gran cantidad de categorías y la minuciosidad en la distinción. Asimismo, González Portal (2000) también propone una taxonomía pormenorizada, que realiza mediante la contraposición de diferentes tipos de conductas prosociales. Se presenta en detalle esta categorización, dado que es también una revisión de los tipos de conducta prosocial delimitados en la literatura hasta ese momento:

- Conducta Prosocial de Ayuda Directa *versus* Conducta Prosocial de Ayuda Indirecta: En la ayuda directa el observador interviene personalmente en la situación. En cambio, ayuda indirecta implica que busca la colaboración de otra persona que es quien interviene directamente.
- Conducta Prosocial Solicitada *versus* Conducta Prosocial No Solicitada: Si la conducta se realiza en respuesta específica a un pedido será una conducta prosocial solicitada; en cambio, si se realiza de *motu* propio, se clasifica como no solicitada.

- Conducta Prosocial de Ayuda Identificable *versus* Conducta Prosocial de Ayuda No Identificable: Este criterio trata acerca de si es posible identificar al benefactor, o si el acto de ayuda permanece como anónimo. Criterios personales, situacionales y temporales demarcan el grado en que la conducta prosocial realizada es identificable.
- Conducta Prosocial de Ayuda en Situación de Emergencia *versus* Conducta Prosocial de Ayuda en Situación de No Emergencia: Se considera ayuda en situación de emergencia a aquella que ocurre en una situación que implica una amenaza o daño real, y en la cual, el peligro puede incrementarse con el transcurrir del tiempo. La ayuda en situación de no emergencia, en cambio, se presenta relacionada con sucesos ordinarios, previsibles y no ambiguos.
- Conducta Prosocial en Situación de Emergencia *versus* Conducta Prosocial Institucionalizada: La conducta prosocial institucionalizada o conducta rol es decidida en un contexto apacible, a partir de una motivación intrínseca. En cambio, las situaciones de emergencia ocurren repentinamente y el factor temporal es clave en la elección de ayudar.
- Conducta Prosocial Espontánea *versus* Conducta Prosocial No Espontánea: En la conducta prosocial espontánea o no planificada la ayuda que se presta es simple, constituyendo un hecho aislado. Supone un contacto breve con un desconocido, con el cual no existe una interacción futura. Además, la decisión sobre la ayuda debe hacerse muy rápidamente. Por el contrario, en la conducta prosocial no espontánea o planificada hay interacciones repetidas, el benefactor busca ayudar y supone un mayor costo de tiempo. Son ejemplos de conducta prosocial planificada el trabajo voluntario sistemático, la contribución regular de dinero a instituciones de beneficencia y la asistencia en el marco de roles profesionales.

Carlo y Randall (2002) en base a teoría e investigación previa, dividen al comportamiento prosocial en seis categorías: Comportamiento Altruista, entendido como ayuda voluntaria a otros motivada en la preocupación por su necesidad y bienestar;

Comportamiento Complaciente es ayudar a otros en respuesta a un pedido; Comportamiento Emocional es ayudar a otros bajo circunstancias fuertemente emocionales; Comportamiento Público es aquél motivado, al menos en parte por un deseo de ganar la aprobación y respeto de otros; Comportamiento Anónimo es aquél realizado sin conocimiento del beneficiario; Comportamiento en Situaciones de Urgencia es ayudar en crisis o circunstancias de emergencia.

Por el contrario, las teorizaciones más recientes tienden a la concentración en menor cantidad de categorías, pero más generales y abarcativas. Dunfield (2014) distingue entre Ayudar, Compartir y Confortar. La autora plantea que Ayudar alivia la Necesidad Instrumental del beneficiario, Compartir intenta satisfacer el Deseo Material de otra persona, mientras que Confortar se orienta a aplacar el Dolor Emocional. Warneken y Tomasello (2009) proponen como factores Confortar, Compartir, Informar y Ayuda Instrumental. López de Cózar, Esteban Martínez, Calatayud Miñana y Alamar Rocati (2008) plantean como subcategorías en su test de Conducta Prosocial en el ámbito del tráfico vial a la Prosocialidad Vital -conductas prosociales generales propias de la vida cotidiana-, la Prosocialidad Vial -comportamientos prosociales específicos del contexto de tráfico-, el Altruismo, la Asertividad y la Empatía. Esta última conceptualización se enmarca dentro de la psicología del tránsito, definida como el estudio de los procesos psicológicos subyacentes al comportamiento humano en el contexto del tránsito, el transporte y la movilidad (Ledesma, Poó, & Montes, 2011). Caprara et al. (2005), también al momento de construir un test, diferencian entre Compartir, Cuidar, Ayudar y Ser Empático. Hay y Cook (2007) clasifican la conducta prosocial en tres categorías: Sentimientos por el Otro, Trabajar con Otro y Atender a Otro.

Específicamente, el comportamiento de ayuda se ha descrito como multideterminado por una amplia gama de factores causales, circunstancias moderadoras y mecanismos mediadores (Penner, Dovidio, Piliavin, & Schroeder, 2005). En el dominio de la ayuda se incluyen comportamientos de rescate, donación, asistencia, voluntariado y apoyo social (Pearce & Amato, 1980) y cada una de estas categorías incluye una amplia gama de comportamientos específicos. Por ejemplo, se han considerado como prosociales conductas tan diversas como donar órganos (Besser, Amir, & Barkan, 2004), ser voluntario en organizaciones sociales (Penner, 2002) o corregir una información incorrecta (Amato 1983).

### *1.3. Variables Relacionadas con la Conducta Prosocial: Consideraciones Generales*

Los comportamientos prosociales son formas generales de beneficiar a la sociedad que se asocian con variables salugénicas (Caprara et al., 2012; Martela & Ryan, 2016). Las personas con mayor razonamiento moral (Caprara et al., 2005; Lemos & Richaud de Minzi, 2010), autocontrol (Bandura, 1991, 1999; Caprara & Pastorelli, 1993; Eisenberg, Fabes, Guthries, & Reiser, 2000), estabilidad emocional (Bekkers, 2006), autoconcepto (Inglés, Martínez-González, García-Fernández, Torregrosa y Ruiz-Esteban, 2012), confianza (Uslaner, 2002), creencia de autoeficacia social (Caprara & Steca, 2005), simpatía (Batson, 1991; Batson & Powell, 2003), agradabilidad (Ashton, Paunonen, Helmes, & Douglas, 1998) y extraversión (Bekkers, 2006; Inglés, Hidalgo, Méndez, & Inderbitzen, 2003) tienden a mostrar mayor grado de comportamientos prosociales.

Los efectos benéficos de la conducta prosocial no son sólo para el receptor, sino que se extienden al actor de la misma y, aún más, hacia la sociedad entera (Caprara et al. 2012). En este sentido, la gratitud y el reconocimiento derivados de realizar conductas solidarias y de cuidado a otros, contribuyen a desarrollar sentimientos positivos acerca del sí mismo, a la aceptación por parte de los otros y el apoyo por parte de esos otros cuando se lo necesita (Caprara & Steca, 2005; Van Willigen, 2000).

Por el contrario, que una persona exhiba una baja frecuencia de conductas prosociales, entendidas en sentido amplio, constituye un factor de riesgo de problemas de conducta y desórdenes afectivos (Scourfield, John, Martin, & McGuffin, 2004). Además, una alta frecuencia de conducta prosocial aplaca la conducta agresiva (Eisenberg et al., 2000; Loudin, Loukas, & Robinson, 2003; Mestre Escrivá, Frías Navarro, Samper García, & Náchter, 2003; Mestre Escrivá, Samper García, & Frías Navarro, 2002).

Existen evidencias de diferencias individuales estables en la realización de conductas prosociales (Eisenberg & Fabes, 1998). Los rasgos, valores y creencias de autoeficacia representan predictores fuertes de conducta prosocial, aunque a diferentes niveles (Caprara, et al., 2012). Como ya se ha mencionado, la agradabilidad parece ser un rasgo importante en la determinación de la conducta prosocial (Graziano, Bruce, Sheese, & Tobin, 2007). Los sujetos con alta agradabilidad presentan una mayor disposición a sacrificar sus intereses a favor de los demás, responder constructivamente a conflictos interpersonales, cooperar en tareas grupales, mostrar autocontrol y percibir positivamente

a los otros (Caprara, Allesandri, Di Giunta, Panerai, & Eisenberg, 2010; Graziano, Habashi, Sheese, & Tobin, 2007). Asimismo, tienden a ser más altruistas, sencillos, confiados, bondadosos, modestos y conformes (McCrae & Costa, 1999). La agradabilidad, sin embargo, podría no ser un predictor consistente en las conductas prosociales con motivaciones egoístas (Carlo, Okun, Knight, & de Guzman, 2005).

La capacidad empática y las creencias de autoeficacia empática también se han planteado en numerosos estudios como variables predictoras y moduladoras de la conducta prosocial (Caprara, et al.; Mestre Escrivá et al., 2002). La capacidad empática suele facilitar la conducta prosocial, pero también se reconoce la necesidad de que el que empatiza disponga de habilidades para responder al sufrimiento ajeno, entendido desde una perspectiva amplia como malestar o niveles de bienestar inferiores a los deseables (Batson, Van Lange, Ahmad, & Lishner, 2007).

Además de la empatía, otras variables disposicionales relacionadas con la conducta prosocial incluyen la anomia, el autoritarismo, la autonomía, el respeto, la inteligencia, el maquiavelismo, el cuidado, la religiosidad, la autoestima, la conveniencia social, la responsabilidad social, la sumisión y otras (Batson & Powell, 2003). Ninguna de ellas, por sí misma, es un predictor claro. Sin embargo, Staub (1974) encontró que una medida integrada de disposición, es decir un índice de orientación prosocial (combinando medidas de sentimientos de responsabilidad personal, responsabilidad social, el razonamiento moral, los valores prosociales, y un bajo nivel de maquiavelismo) fue un buen predictor de la conducta prosocial. Otro desarrollo que relaciona conducta prosocial con personalidad fue realizado por Penner, Fritzsche, Craiger y Freifeld (1995). Estos autores definen personalidad prosocial como una tendencia continua a preocuparse por el bienestar y los derechos de los demás y realizar acciones al respecto. Más recientemente, Caprara et al. (2012) delimitan la existencia de una tendencia a actuar prosocialmente a través del tiempo y las situaciones, denominada prosocialidad.

Las variables disposicionales podrían ser más o menos importantes dependiendo de variables situacionales (Aydinli, Bender, & Chasiotis, 2013). Por caso, los hallazgos de Graziano et al. (2007) demostraron que la agradabilidad fue un predictor consistente de la conducta prosocial cuando el receptor de la ayuda era un extraño o la urgencia de la situación era baja.

Por sí mismos, los factores ambigüedad de necesidad, la gravedad de necesidad, la apariencia física de la víctima, las condiciones climáticas, similitud con la víctima,

amistad o involucramiento, número de espectadores, cohesión de la comunidad y costo de ayudar –todas variables situacionales- fueron buenos predictores de la conducta prosocial (Batson & Powell, 2003; Galen, 2012; Wenner & Randall, 2016). La centralidad sociocognitiva, es decir el monto de información que comparte un miembro con otros (Vivas, Ricci, & Terroni, 2003) y el medio de comunicación (Vivas, Terroni, & Ricci, 2002) también podría influir en la realización de este tipo de conductas. En un estudio clásico sobre el tema, Darley y Latané (1968) delimitan el efecto del espectador, según el cual cuantas más personas presencien y observen una situación de emergencia, menos probabilidades existen que una de ellas lleve a cabo un comportamiento de ayuda. Dicho efecto, según los autores, se podría deber, entre otras razones, a la dilución de la responsabilidad entre los observadores (Darley & Latané, 1968).

Se ha planteado la posibilidad de que existan variables mediadoras que podrían ayudar a explicar las relaciones entre rasgos de personalidad y ciertos comportamientos prosociales, como el voluntariado. Entre dichas variables mediadoras se encuentran las motivaciones (Carlo, et al., 2005).

### 1.3.1. Variables Sociodemográficas y su Relación con la Conducta Prosocial

#### 1.3.1.1. La Conducta Prosocial y los Ciclos Vitales

La edad se relaciona positivamente con la conducta prosocial durante cierto período ya que, a lo largo de la niñez, el repertorio de respuestas prosociales va aumentando progresivamente (Shaffer, 2002). La conducta prosocial se convierte en relativamente estable durante el final de la infancia y la adolescencia temprana. En la adolescencia media tiende a declinar, para incrementarse en la adolescencia tardía (Carlo, Crockett, Randall, & Roesch, 2007; Inglés et al., 2008).

La posibilidad de realizar distintas conductas prosociales se desarrolla progresivamente durante la niñez (Shaffer, 2002), estabilizándose durante el final de la infancia. Este tipo de conducta surge de complejos procesos del desarrollo general y psicológico. En este desarrollo se implican los procesos de atención y de evaluación, razonamiento moral, competencia social y capacidad de autorregulación (Caprara & Pastorelli, 1993; Eisenberg & Fabes, 1998). Richaud de Minzi (2009) analizó la relación entre la empatía percibida por los niños en sus padres con el razonamiento moral

prosocial, hallando que los niños que percibieron que sus padres eran capaces de entender los puntos de vista de los otros o imaginarse cómo se sienten, tienen un razonamiento menos hedonista u orientado hacia la propia satisfacción. En un estudio realizado con adolescentes tempranos, Richaud de Minzi, Mesurado y Lemos (2012) encontraron una relación positiva entre la aceptación materna, es decir, la actitud cálida y sensible a las necesidades de sus hijos, proximidad en momentos de crisis o tensión, y la preocupación empática materna con el altruismo, indicando que percibir afecto positivo y preocupación por otros por parte de los padres fortalece la motivación de este tipo de conductas.

En la adolescencia, la redefinición de la identidad y los cambios cognitivos consolidan la personalidad (Musitu & Cava, 2003). En esta etapa, la conducta prosocial juega un papel central en el bienestar psíquico y físico, adquiriendo el grupo de pares una relevancia mayor como contexto de socialización (Martínez, Inglés, Piqueras, & Oblitas, 2009). Las interacciones grupales promueven las habilidades interpersonales y la salud psíquica. Los adolescentes con mayor grado de prosocialidad tienen una mejor adaptación escolar y éxito académico, adquiriendo más fácilmente hábitos y estilos de vida saludable (Rodrigo et al., 2004). En la adolescencia media el repertorio de conductas prosociales tiende a declinar, para acrecentarse en la adolescencia tardía (Carlo et al., 2007). Cardozo y Alderete (2009) estudiaron la resiliencia en adolescentes como un proceso dinámico donde las influencias del ambiente y los recursos internos con los que cuenta el individuo interactúan, permitiendo al adolescente una adaptación positiva a pesar de encontrarse en un ambiente adverso, y en situación de riesgo. En este contexto la conducta de adaptación social propia de los adolescentes resilientes fue entendida como una alta prevalencia de conductas prosociales en ausencia de conductas antisociales.

La tendencia hacia una mayor orientación prosocial parece ser una marca de la entrada en la madurez adulta, según consideraron adolescentes y jóvenes de Canadá, Estados Unidos e Israel (Eisenberg, Cumberland, Guthrie, Murphy, & Shepard, 2005). Pero se observó que hay pocos estudios en relación con el significado psicológico de la prosocialidad para el bienestar y el ajuste personal durante la adultez y, especialmente, la vejez (Caprara, et al., 2005). Los comportamientos prosociales se relacionan positivamente en la adultez con la simpatía, la toma de perspectiva, la responsabilidad social y la adscripción de la responsabilidad hacia el propio yo (Carlo, Hausmann, Christiansen, & Randall, 2003; Carlo, Knight, McGinley, & Hayes, 2011). La ayuda humanitaria en adultos se asocia con la preocupación por los derechos y necesidades

humanos, el bajo etnocentrismo, la igual valoración de las vidas de los miembros del endo y exogrupo y el conocimiento de los temas relevantes del humanitarismo (McFarland, Webb, & Brown, 2012).

El envejecimiento supone particulares cambios biológicos, psicológicos y sociales que dan lugar a una crisis normal del envejecimiento, requiriendo el uso de estrategias de afrontamiento y apoyo social (Krzemien, Monchietti, & Urquijo, 2005). Aquí es donde el comportamiento prosocial de ayuda y de cooperación puede cobrar gran importancia. En la vejez los lazos sociales son los que sustentan a la persona en su salud psíquica, es así que el comportamiento prosocial resulta importante en tanto potenciador de dichos lazos. En un estudio realizado con una muestra de adultos mayores de la Ciudad Autónoma de Buenos Aires (Mikulic, Caballero, & Aruanno, 2014) se halló que los participantes contarían con mayor optimismo, capacidad de ejercer o defender sus derechos personales, regulación emocional y autonomía que sus pares más jóvenes. Según Krzemien (2001), en la medida en que el adulto mayor logre reconocimiento a partir de las relaciones sociales significativas, esta etapa será vivida como prolongación del resto del ciclo vital; de lo contrario se vivirá como una fase de declinación funcional y aislamiento social.

Neurológicamente, la actividad social potencia el aprendizaje y la memoria y previene el envejecimiento y las neuropatías. El “enriquecimiento ambiental” protege el cerebro, restaurando funciones que se pierden a causa de la edad, las lesiones o determinadas enfermedades (Bentosela & Mustaca, 2005). Es así que la calidad de vida se halla estrechamente relacionada con la participación en el medio social, y ésta última se ve incrementada por la realización de conductas prosociales. Es por esto que resulta especialmente importante el estudio de la conducta prosocial en esta etapa del ciclo vital, donde menos se ha evaluado e investigado. A mayor edad, los comportamientos prosociales son motivados de forma incremental por valores internalizados o por una autoimagen prosocial (Eisenberg, Hofer, Sulik, & Liew, 2013). En este sentido, en una investigación sobre el perdón como motivador prosocial en la vejez (Morales & Arias, 2014) utilizando la Escala para Evaluar a Capacidad de Perdonar (CAPER; Casullo & Fernández Liporace, 2005) se halló que las personas mayores tienen significados y representaciones más amplias en las creencias y convicciones utilizadas para justificar el perdón.

### 1.3.1.2. La Conducta Prosocial y el Género

Uno de los correlatos más consistentes del comportamiento prosocial es el género. Algunos investigadores han encontrado que el mismo modera la preocupación por los otros. Por otra parte, las diferencias sexuales culturalmente estereotipadas en el comportamiento deseado o esperado pueden llevar a los padres a fomentar, consciente o inconscientemente, resultados divergentes del desarrollo en sus hijas e hijos.

Los roles de género tradicionales implican conductas prosociales diferenciadas para hombres y mujeres. En etapas donde la mujer estaba recluida en el hogar típicamente realizaba sus comportamientos prosociales en el contexto de la familia y otras relaciones sociales cercanas en las que una pareja, un pariente o un amigo solicitaba apoyo, cuidado o empatía. Los hombres tendían a ayudar a otros especialmente cuando la situación lo requería y se necesitaban acciones rápidas y decisivas o cuando alguien estaba en necesidad de ayuda clara o grave peligro. Reflejo de estos prototipos es, por ejemplo, que una carrera como enfermería todavía siga con un estudiantado mayormente femenino, mientras que la profesión de bombero se encuentra repleta de varones.

Actualmente los estereotipos de género continúan mostrando su influencia en la orientación hacia el otro. Sin embargo, se encuentran mediados por expectativas sociales que tienden a una mayor igualdad y disposiciones individuales. Las tendencias tradicionales comienzan a revertirse y a encontrarse resultados contradictorios. Por ejemplo, en un estudio transversal realizado por Plazas et al. (2010) en la ciudad colombiana de Valleduparse halló que las niñas fueron mucho más prosociales durante la primaria que los varones, la diferencia fue menor en la secundaria, y en la universidad la tendencia cambió, siendo los varones más prosociales. En relación con el voluntariado, en Estados Unidos las mujeres tienden a realizar más actividades, pero en roles vinculados a la educación, salud y provisión o preparación de comida mientras que los líderes de los voluntarios son en su mayoría varones (Eagly, 2009).

Se puede observar en las investigaciones actuales que las diferentes conductas prosociales varían en magnitud según el género dependiendo de los comportamientos en sí, su contexto social, las disposiciones individuales y los métodos de estudio, entre otros factores (Eagly, 2009). Las diferencias a favor de las mujeres halladas en diferentes medidas de conducta prosocial podrían deberse al hecho de que los índices utilizados en muchos estudios se corresponden más con el rol femenino que con el masculino. Así,

pues, no existe evidencia consistente sobre esta cuestión (Extebarría et al., 2003). La dificultad de controlar la influencia de los estereotipos de género en la medida de las variables deriva en grandes problemas metodológicos.

#### 1.3.1.3. Conducta Prosocial: Su relación con el Ámbito Académico y la Profesión

Un clima escolar positivo en general contribuye a mejorar la calidad de la educación (Rodríguez, Vivas, Comesaña, Ramírez, & Peña, 2017). Inglés, Martínez-González y García-Fernández (2013) demostraron que la conducta prosocial se relaciona de forma positiva y estadísticamente significativa con puntuaciones altas en diferentes estrategias y habilidades de estudio en adolescentes. Asimismo, el rendimiento es mayor en los alumnos prosociales (Inglés, Benavides, Redondo, García-Fernández, Ruiz-Esteban, Estévez, & Huescar, 2009). También la conducta prosocial se relaciona con autoatribuciones académicas específicas en adolescentes (Redondo, Inglés, & García-Fernández, 2014). Se demostró que existe una moderada y significativa correlación positiva entre responsabilidad y conducta prosocial (Gutiérrez San Martín, Escartí, & Pascual, 2011). Si entendemos por voluntad de trabajo la actitud por la cual una persona se aboca a sus obligaciones realizándolas con responsabilidad, precisión, automotivación y sin dilaciones, motivada por su propio deseo de realizar un buen trabajo sin rehuir el esfuerzo que ello implica (Abal, Lozzia, Aguerri, Galibert, & Attorresi, 2007, 2008; Abal, Lozzia, Blum, Galibert, Aguerri, & Attorresi, 2010; Attorresi, Abal, Galibert, Lozzia & Aguerri, 2011; Galibert, Aguerri, Lozzia, Abal, & Attorresi, 2005), se puede inferir que esta variable también se relaciona positivamente con la conducta prosocial. Tanto la responsabilidad como la conducta prosocial constituyen aspectos que favorecen el desarrollo saludable de la persona (Gutiérrez San Martín et al., 2011).

Inglés, et al. (2013) demostraron que la conducta prosocial es un predictor positivo y estadísticamente significativo de puntuaciones altas en las siguientes estrategias y habilidades de estudio en adolescentes: actitud hacia el éxito académico, motivación, procesamiento de la información, selección de ideas principales, ayudas al estudio, autoevaluación y estrategias de evaluación. Se entiende como participación social académica al conjunto de conductas sociales, tanto con pares como docentes, que realiza el estudiante dentro del ámbito de clase para apoyar, llevar adelante o mejorar su proceso de aprendizaje. Participaciones sociales académicas como la ayuda a pares para

desarrollar una tarea de aprendizaje son conductas prosociales (Molinari, Takaya, Domínguez-Romero, & Sánchez Rosas, 2016). Redondo et al. (2014) estudiaron las autoatribuciones académicas también en adolescentes y hallaron que, respecto a la asignatura de lenguaje, los estudiantes prosociales atribuyen significativamente el éxito a la capacidad, el esfuerzo y, en menor medida, a causas externas. En cuanto a la asignatura de matemáticas, los estudiantes prosociales atribuyeron el éxito significativamente más al esfuerzo y significativamente menos a causas externas, mientras que atribuyeron el fracaso significativamente más a la falta de esfuerzo.

La profesión de una persona también parece relacionarse con la conducta prosocial. Hillygus (2005) halló que los graduados en ciencias sociales tienden a ser más generosos. Sin embargo, el ejercicio de profesiones de ayuda puede ocasionar resultados adversos. Por ejemplo, respecto del ejercicio de la Psicología, se halló que el trabajo continuo con pacientes en situación de trauma puede dar lugar al fenómeno de fatiga de compasión, también llamado estrés traumático secundario. Figley (1995) puntualizó que hay un costo en ayudar, en el sentido de que los profesionales que escuchan los relatos de miedo, dolor y sufrimiento pueden experimentar sentimientos similares a los de sus pacientes. Los síntomas de la fatiga de compasión incluyen desesperanza y confusión y pueden llevar a los profesionales a tomar decisiones clínicas incorrectas (Eastwood & Ecklund, 2008). El profesional de la salud mental en este caso debe, por lo tanto, tomar medidas preventivas para evitar padecer este síndrome.

### 1.3.2. Algunas Variables Disposicionales Relacionadas con la Conducta Prosocial: Empatía, Agresividad, Felicidad y Gratitud

#### 1.3.2.1. La Empatía

El concepto de empatía tiene una historia signada por el desacuerdo. Aunque ha sido estudiada por cientos de años con contribuciones de la filosofía, teología, psicología del desarrollo, social y de personalidad, etología y neurociencia, el campo continúa sufriendo la falta de consenso acerca de la naturaleza el fenómeno (Preston & De Waal, 2002). Las investigaciones experimentales ofrecen resultados muy diversos, dado que las concepciones de las que se parte para definirlo son disímiles e incluso, en algunos casos, antagónicas (Fernández Pinto, López Pérez, & Márquez, 2008; Preston & De Waal,

2002). Esta multiplicidad se transforma en una dificultad cuando surge la necesidad de investigar el concepto en forma precisa (Olivera Ryberg, 2010).

Una de las conceptualizaciones más extendidas sobre la empatía la define como un conjunto de constructos que incluyen los procesos de ubicarse en el lugar del otro, reacciones cognitivas y emocionales del observador ante una situación determinada y respuestas tanto afectivas como no afectivas que resultan de esos procesos (Davis, 1996). Según Batson y Ahmad (2009) en los últimos años el término empatía ha sido aplicado para cuatro estados psicológicos diferentes. Dos de ellos se refieren a formas de la toma de perspectiva y los restantes a formas de respuesta emocional. Más recientemente se ha delimitado a los comportamientos empáticos como aquellas conductas que manifiestan comprensión, refuerzo y soporte emocional (Auné, Abal, & Attorresi, 2017a).

¿Qué es propiamente la empatía dentro de la esfera de la afectividad? ¿Cuáles son los elementos que la forman? ¿Cuál es su efecto real en el desarrollo de la personalidad del sujeto? ¿Es una conducta prosocial (Hay & Cook, 2007) o -como lejanamente sostenía Lipps (1903)- un comportamiento en el que se ignora la alteridad, considerado el otro una invención del yo, al referirse la empatía al sí mismo, considerada sólo como una vivencia que podría no incluirlo?

Esta falta de consenso ha conducido a la proliferación de aproximaciones teóricas dirigidas a determinar si la empatía es un proceso emocional o cognitivo y a distinguir la empatía del contagio emocional, la simpatía y la toma de perspectiva. En un nivel profundo el debate parece surgir de la falta de un mecanismo explicativo (Preston & De Waal, 2002). Citando a Allport (1968 p. 30 en Wispé, 1987) “el proceso de la empatía recuerda a un acertijo en psicología social”.

Mestre Escrivá, Frías Navarro y Samper García (2004) registraron que las dimensiones de empatía del *Interpersonal Reactivity Index* (IRI), en términos generales, correlacionan positivamente con la conducta prosocial y con los estilos de razonamiento prosocial más orientados a una disposición prosocial, mientras que la correlación es negativa con la conducta agresiva y con la inestabilidad emocional. Las correlaciones más fuertes se alcanzaron entre los factores de empatía que indican una disposición empática más madura y la conducta prosocial y entre estos mismos factores y el razonamiento internalizado. Este razonamiento incluyó argumentos orientados a la comprensión del problema del otro, a anticipar consecuencias físicas y emocionales que se puedan derivar de la acción de ayudar o no ayudar y argumentos como la satisfacción personal de actuar

según los propios valores. Por el contrario, las correlaciones son negativas con el razonamiento hedonista, centrado en el beneficio personal que se puede derivar de la acción y con el razonamiento orientado a buscar la aprobación de los demás propio de una menor madurez prosocial (Mestre Escrivá et al., 2004).

Mestre Escrivá et a (2002) plantean a la empatía como factor modulador de la conducta prosocial. Funciona como un mecanismo regulador, favoreciendo la conducta prosocial e inhibiendo la agresividad. Según los resultados que obtuvieron la empatía aparece como el principal motivador de la conducta prosocial, en sus componentes cognitivos -la comprensión del otro-, pero más fuertemente en sus componentes emocionales -la preocupación por el otro-. En este sentido, el actuar prosocialmente supone la utilización de criterios de igualdad y bien social y un razonamiento centrado en una orientación empática abierta (Mestre Escrivá, Samper García, Tur, Cortés, & Nácher, 2006).

Rodríguez y Moreno (2016) investigaron las posturas éticas relativismo -las normas morales tienen validez subjetiva y relativa- y absolutismo -las normas morales tienen una verdad única-, encontrando que influyen sobre la conducta prosocial no directamente sino con la empatía como factor mediador. Mientras que el absolutismo mediado por la empatía correlacionó moderadamente con la conducta prosocial, el relativismo mediado por la empatía lo hizo con valor casi nulo y sin significación estadística.

De esta forma, la empatía sería un prerrequisito para la conducta prosocial pero no parte de la misma. El comprender los sentimientos de otra persona no necesariamente implica que se actuará, o incluso que se sentirá impelido a actuar de forma comprensiva o simpática (Mestre Escrivá et al., 2002).

Caprara et al. (2005) introducen un importante giro conceptual al incorporar a la empatía como una dimensión de la prosocialidad. Destacan la novedad de esta inclusión y afirman que la misma se basó en el convencimiento de que en la adolescencia y adultez los motivos o predisposiciones empáticas no son sólo un correlato de la tendencia a actuar prosocialmente, sino que forman parte de dicha tendencia. La empatía es un componente integral de la prosocialidad en estas etapas ya que el acudir al rescate de los demás implica la capacidad de sentir sus necesidades y estados de ánimo (Eisenberg, & Fabes, 1998; Hoffman, 2000).

Por lo tanto, Caprara et al. (2005) sostienen la hipótesis de que comportamientos diferentes tales como ayudar, confortar y empatizar son expresiones de la misma tendencia o disposición de orden superior denominada prosocialidad (Eisenberg & Fabes, 1998; Hoffman, 2000). En particular la prosocialidad en la adolescencia y la edad adulta va más allá de ayudar, cuidar y compartir, el comportamiento también se puede poner en práctica por simple educación, convención o conveniencia e implica una particular inclinación a participar en las experiencias emocionales de los demás. Básicamente se trata de la capacidad de implicación emocional combinada con plena conciencia de las intenciones y los efectos de las acciones (Caprara et al., 2005).

La discusión entre la empatía como predictor y la empatía como parte de la tendencia prosocial aún no ha sido saldada. Ambas partes encuentran evidencia para su posición, en el primer caso mediante la obtención de correlaciones significativas con criterios externos y en el segundo, hallando que, aún incluyendo a la empatía, en un instrumento para la medición de la prosocialidad en adultos resulta confirmada una estructura interna unidimensional (Caprara et al., 2005). La disputa acerca de la empatía disposicional o situacional ha sido enriquecida con el concepto de interacción entre estos dos factores. El campo de las personalidades, frecuentemente considerado como la organización y estructuración de todas las dimensiones evaluables de las diferencias individuales a través del tiempo y las situaciones (Saklofske, & Zeidner, 1995) tiene a la empatía como parte de dicho campo de estudio.

Uno de los temas más polémicos en la investigación de la empatía es si resulta mejor conceptualizarla como una construcción afectiva, una construcción cognitiva, o ambas. Mientras que la investigación contemporánea parece haber adoptado la opción ambas, no se apoya de forma generalizada y el debate continúa sobre la definición y medición de la empatía. Ambas entidades (emoción y cognición) se combinan para generar empatía y establecer las principales influencias en las conductas de acercamiento y aceptación de personas o de objetos (Newman, Dennis, Wright, & King, 2010). Se puede afirmar por lo tanto que la empatía es no sólo efecto de una respuesta cognitiva sino también una respuesta emocional de un observador al estado afectivo de otro sujeto.

Conducta prosocial y empatía se han relacionado siempre desde el punto de vista conceptual, sin embargo, desde el empírico las investigaciones no han sido concluyentes acerca de la naturaleza de esa relación. Es necesario destacar que estos frentes; enfoque disposicional *versus* enfoque situacional, perspectiva cognitiva *versus* perspectiva

afectiva y empatía como predictor *versus* empatía como dimensión de la conducta prosocial; son parte de una larga serie de desacuerdos: capacidad empática como continuo *versus* visión dicotómica, beneficiosa siempre para el otro *versus* neutral, innata *versus* adquirida, entre otros.

La tendencia actual se dirige hacia el surgimiento de teorías integradoras, dejando atrás los antagonismos en torno al concepto de empatía. Se destaca la necesidad de creación de enfoques multidimensionales que incluyan lo cognitivo, emocional, disposicional y situacional, en la búsqueda por encontrar descripciones de mayor complejidad. De esta forma, los antagonismos representan esfuerzos para sistematizar los conocimientos que van siendo superados. Alcanzar la multidimensionalidad permitirá comprender y definir la naturaleza de la empatía en términos teóricos; pero también concebir, diseñar y construir instrumentos para su medición y formas de abordaje práctico.

#### 1.3.2.2. La Agresividad

El *Homo Sapiens*, hace más de 160000 años, desarrolló no sólo la capacidad de socializar con individuos no emparentados, sino también la de distinguir cuando luchar. Es decir, por ejemplo, la defensa agresiva de fuentes de alimento concentradas y predecibles como así de territorios resultó beneficiosa para la supervivencia (Marean, 2015).

Actualmente se encuentra una relación negativa entre conducta prosocial y agresividad. Representan estrategias de resolución de problemas interpersonales que se aprenden desde muy temprano en la vida. Para Hawley (1999) los individuos acceden con éxito a diferentes tipos de recursos (sociales, informativos o materiales) a través de la utilización de estrategias prosociales y/o coercitivas. Estas estrategias se pueden combinar, dando lugar a cinco patrones: prosocial, coercitivo, biestratégico, típico y no controlador. Específicamente, los sujetos biestratégicos logran un alto grado de éxito social.

Se puede inferir que el patrón de estrategias adecuado dependerá del contexto social en que se encuentre la persona y la ocasión particular. Una estrategia puramente coercitiva será necesaria, incluso para la supervivencia, en un medio violento. La inequidad de *status* de poder y/o económico se asocia con comportamientos de

victimización (Wolke, Woods, & Samara, 2009), acoso escolar (Elgar et al., 2009) y otros comportamientos violentos (Wilkinson & Pickett, 2009). La diferencia de poder es negativa para las relaciones entre pares.

La agresión temprana es el mejor predictor negativo del logro educativo posterior y el mejor predictor positivo del fracaso social (e.g. Cid, Díaz, Pérez, Torruella, & Valderrama, 2008; Fite, Hendrickson, Rubens, Gabrielli, & Evans, 2013). Por el contrario, el comportamiento prosocial precoz es uno de los indicadores más importantes del éxito social, el nivel de instrucción y el desarrollo de la autoestima (e.g. Caprara, Barbaranelli, Pastorelli, Bandura, & Zimbardo, 2000; Zuffiano et al., 2014). En un medio social donde la coerción no es la forma válida de acceder a los recursos, un factor importante en el control del comportamiento agresivo es el aprendizaje de comportamientos no agresivos.

Las estrategias agresivas de resolución de problemas aprendidas en la infancia temprana y media persisten en la edad adulta. La estabilidad de la conducta prosocial y agresiva puede ser considerado bastante alta.

Cuanto más el joven es castigado físicamente, menospreciado y humillado por los padres es más probable que él o ella tenga una agresividad patológica cuando sea adulto. El elemento de disuasión más importante para el desarrollo del comportamiento antisocial y el estímulo de la conducta prosocial parece ser la estrecha identificación entre el niño y sus padres (Eron & Huesmann, 1984). Por lo tanto, el comportamiento prosocial al igual que el comportamiento agresivo se aprende temprano en la vida en el contexto de las interacciones familiares.

### 1.3.2.3. La Felicidad

La felicidad ha sido concebida como estado mental, forma de percibir y experiencia de regocijo y satisfacción (Csikszentmihalyi, 2008; Lyubomirsky, 2008). Se caracteriza por una mayor intensidad de la experiencia que la separa de sensaciones más superficiales como la satisfacción y el bienestar. Típicamente se asocia al amor, la cercanía con la naturaleza y otras dimensiones no racionales (Susniene & Jurkauskas, 2015).

Se puede afirmar que la felicidad es difícil de definir porque es muy personal. Depende de valores propios, inclinaciones, carácter, factores culturales e incluso factores

hereditarios (Hill et al., 2013; Susniene, & Jurkauskas, 2015). Csikszentmihalyi (2013) caracteriza el concepto de felicidad como un *flow*, una experiencia especial, brillante y excitante que acompaña la actividad vital.

La felicidad se encuentra originalmente descrita en la filosofía clásica y en la religión. A pesar de su importancia, la Psicología a lo largo de su historia se ha concentrado mucho más frecuentemente en el estudio de las emociones negativas y las patologías mentales. Sin embargo, en los últimos años los investigadores comenzaron a interesarse en el bienestar subjetivo y en la sensación de felicidad; siendo la llamada Psicología Positiva un área en crecimiento (Donaldson et al, 2015).

Las siguientes cuestiones adquirieron particular relevancia para el estudio de la felicidad en el contexto de la Psicología Positiva: el potencial positivo de la personalidad, el funcionamiento óptimo de la persona y el estudio de los factores que motivan el estado de felicidad (Dzhidaryan, 2013; Gurieva, 2012; Seligman, 2009).

El abordaje psicológico de la felicidad logró un hito importante en 1962, cuando Abraham Maslow publicó su libro “Hacia una Psicología del Ser”. Estableció una teoría acerca de la calidad de vida basada en el concepto de las necesidades humanas y trató de resolver el problema de la felicidad creando un camino universal de desarrollo personal. Para ello, determinó una sucesión de necesidades, donde la siguiente necesidad se revela cuando se satisface la anterior.

Maslow (1962) describió la vida ideal como un largo viaje a través de ocho necesidades, que transitan desde lo concreto y mundano hasta lo abstracto y trascendente. Para que la persona las satisfaga deberá desarrollar cualidades de espontaneidad, independencia y responsabilidad. Posteriormente, el concepto de necesidad fue criticado por muchos académicos por ser difuso e inadecuado para la investigación (Susniene, & Jurkauskas, 2015).

Actualmente, existen diferentes aproximaciones para entender el fenómeno de la felicidad (Dzhidaryan, 2013; Vinichuk, 2008). Básicamente estos enfoques se pueden dividir en objetivistas y subjetivistas. De acuerdo con los enfoques objetivistas, la felicidad se sitúa en el campo de las cuestiones de la personalidad y está conectada con el nivel de bienestar psicológico y social y el nivel de realización personal (Ryff, 2009). En cambio, de acuerdo con el segundo enfoque la felicidad es considerada como una sensación subjetiva, una experiencia psicológica particular (Argyle, 2003).

El fenómeno de la felicidad puede ser abordado desde tres posiciones principales: como satisfacción vital, como valor y como un estado afectivo positivo (Vinichuk, 2008). La felicidad como satisfacción vital se relaciona con los conceptos de calidad de vida y bienestar psicológico (Argyle, 2003). La felicidad como valor puede ser estudiada a través del análisis del sistema de valores humanos, objetivos, expectativas, necesidades básicas y resultados de la actividad (Argyle, 2003; Csikszentmihalyi, 2013). Por último, la felicidad como estado afectivo positivo es considerada como un sentimiento especial de experiencia personal. Implica la ausencia de displacer y emociones negativas simultáneamente con la presencia de emociones placenteras y positivas (Joseph, Linley, Harwood, Lewis, & McCollam, 2004).

El análisis de las categorías de felicidad presentes en la bibliografía permite determinar tres componentes en la estructura del fenómeno: cognitiva, afectiva y conductual (Susniene, & Jurkauskas, 2015). El componente cognitivo incluye la comprensión por parte del sujeto de las razones de su felicidad e incluye una evaluación reflexiva. En este sentido, Andrés, Castañeiras y Richaud de Minzi (2014) hallaron que la reevaluación cognitiva, es decir la construcción de significados alternativos positivos de un evento, media entre la extraversión y la felicidad. El componente afectivo implica un proceso de experiencia subjetiva de la felicidad. Por último, el componente conductual se relaciona con las manifestaciones externas de comportamientos verbales y no verbales.

En cuanto al estudio de los factores que motivan que una persona sea feliz, Argyle (2003) sitúa: las relaciones sociales cercanas, la satisfacción laboral, la salud, tiempo de ocio, rasgos personales (autoestima, extraversión, sentido vital) y emociones positivas. La felicidad se relaciona con la calidad de vida, pero estos conceptos no se superponen completamente. Ambos son conceptos multidimensionales e integradores de las distintas áreas vitales de la persona (Mikulic & Muiños, 2005). Mikulic, Muiños y Crespi (2007) hallaron cuatro factores latentes de la calidad de vida percibida: ambiente, redes sociales, crecimiento personal y necesidad de trascendencia. Esta estructura tetrafactorial se mantuvo en sujetos que afrontan distintos sucesos vitales estresantes tales como encarcelamiento, discapacidad, divorcio, pérdidas de familiares o migración forzada.

Para Seligman et al. (2004) existen tres formas de encontrar felicidad. La primera es el sentimiento positivo hacia el pasado, presente o futuro. La segunda ruta es el perseguimiento de la gratificación. Esta involucra al individuo plenamente y requiere cualidades como creatividad, inteligencia social, sentido del humor y perseverancia. La

tercera ruta se refiere a utilizar las propias fortalezas para contribuir con algo más grande que nosotros mismos, como el conocimiento, la bondad, la familia, la comunidad, la justicia o un poder espiritual. Castro Solano y Sánchez López (2000) hallan que los vínculos personales satisfactorios y la percepción de estar bien físicamente predicen la satisfacción de logros en la vida.

Además, se destaca el papel de la prosocialidad para posibilitar el bienestar y la armonía de la persona (Auné et al., 2014; Caprara et al., 2005). La gratitud y el reconocimiento derivados de cuidar a otros contribuyen a elevar la autoestima y generar relaciones sociales positivas y de apoyo mutuo.

Con respecto a la relación entre felicidad y género, la misma resulta compleja. Para arribar a un resultado más general, Graham y Chattopadhyay (2013) realizaron un estudio global en 160 países, hallando que las mujeres se autorreportan como más felices que los varones consistentemente alrededor del mundo. Cuando los niveles de bienestar general caen o suben, esto afecta a ambos géneros, pero la brecha se mantiene. De hecho, esta diferencia de bienestar a favor de las mujeres es mayor en países desarrollados. La excepción, donde los varones mostraron mayores niveles de felicidad, es la región del África subsahariana.

Varones y mujeres experimentan la felicidad de múltiples formas. Scalco, Araújo y Bastos (2011) hallan en una muestra brasileña del municipio de Pelotas, perteneciente al estado de Rio Grande do Sul, una mayor prevalencia de la felicidad asociada con la juventud en el caso de los varones y con la religiosidad entre las mujeres. Además, la condición de ser separada o viuda se relaciona con los niveles más bajos de felicidad femenina. Más recientemente, Al Nima y Garcia (2015) descubren en una muestra de residentes angloparlantes en Estados Unidos, reclutados a través del sitio web *Amazon's Mechanical Turk*, que las estrategias para aumentar la felicidad varían entre varones y mujeres. Mientras que las mujeres tienden a utilizar las actividades sociales, control mental, ocio pasivo, religión y estrategias proactivas para cultivar el buen humor, los varones prefieren celebrar e ir a fiestas.

En América Latina, la felicidad es uno de los constructos más frecuentemente estudiados por la Psicología Positiva (Castro Solano, 2013). Se destacan los estudios de Alarcón (2006), que concibe a la felicidad como un estado afectivo de satisfacción plena que experimenta subjetivamente el individuo en posesión de un bien anhelado. Para medirla diseña y construye la Escala de Felicidad de Lima (EFL) de la cual resulta una

estructura tetrafactorial, siendo los factores originales de esta escala Sentido Positivo de la Vida, Satisfacción con la Vida, Realización Personal y Alegría de Vivir. En la adaptación argentina de dicho instrumento (Auné, Abal & Attorresi, 2017b) se eliminaron las dimensiones Realización Personal y Alegría de Vivir y se desdobló Sentido Positivo de la Vida. Se hallaron, a través de la evidencia empírica, tres dimensiones: Sentido de la Vida Perspectiva Externa (SVPE), Sentido de la Vida Perspectiva Interna (SVPI) y Satisfacción con la Vida (SCV). Es decir, fue importante para los argentinos distinguir al exterior o al propio yo como causas de felicidad.

#### 1.3.2.4. La Gratitud

Hasta tiempos recientes la gratitud había sido un constructo poco estudiado (Wood, Froh, & Geraghty, 2010). Watkins, Woodward, Stone y Kolts (2003) la consideran una de las virtudes abandonadas por la Psicología, sobre todo teniendo en cuenta que sus equivalentes lingüísticos existen en prácticamente todos los idiomas y que las religiones mayoritarias en el mundo han destacado su importancia.

McCullough, Emmons y Tsang (2002) definieron la disposición a la gratitud como una tendencia generalizada a reconocer y responder con emoción de agradecimiento a la benevolencia de otra persona. Emmons y McCullough (2003) posteriormente relevaron conceptualizaciones más amplias, que caracterizan a la gratitud como una emoción, actitud, virtud moral, rasgo de personalidad o respuesta de afrontamiento. Asimismo, recalcaron que la gratitud posee componentes tanto cognitivos como emocionales.

Watkins et al. (2003) definieron el afecto de gratitud como una apreciación agradecida por los favores recibidos. Watkins et al. (2003) identificaron cuatro características de las personas agradecidas. Los individuos agradecidos tienen una sensación de plenitud, valoran las contribuciones de otros para su bienestar, aprecian los placeres simples de la vida y, por último, reconocen la importancia de experimentar y expresar la gratitud.

Según Wood et al. (2010), a un nivel disposicional, la gratitud es parte de una orientación vital más amplia hacia apreciar y tomar conciencia de lo positivo del mundo. Esta orientación vital debe distinguirse de otras emociones tales como optimismo y esperanza, las cuales se centran en el futuro.

Distintos estudios han relacionado a la gratitud con cada uno de los Cinco Grandes de la Personalidad (*Five Factor Model*; Digman, 1990), hallando que las personas agradecidas son más extrovertidas, agradables, abiertas, conscientes y con menos neuroticismo (e.g. Wood, Maltby, Stewart, Linley et al., 2008). En general, se ha encontrado que las personas agradecidas son menos irascibles, hostiles, depresivas y emocionalmente vulnerables, experimentando emociones positivas más frecuentemente (Auné, Abal & Attorresi, 2016a; Wood et al., 2010).

La gratitud ha correlacionado en forma directa con rasgos asociados al funcionamiento social positivo: calidez emocional, sociabilidad, proactividad y altruismo (Auné et al., 2014). Las personas agradecidas tienen una mayor apertura hacia sus sentimientos, ideas y valores, así como sentido del deber y esfuerzo orientado a logros (Joseph & Wood, 2010; Wood et al., 2010).

La más frecuente asociación de la gratitud en la investigación empírica es con el bienestar subjetivo e incluso lo predice de forma confiable (Gulliford, Morgan & Kristjánsson, 2013). Entre las posibles explicaciones una es que la gratitud es contraria a la envidia y los deseos materialistas, conocidos detractores del bienestar personal. Las personas agradecidas reportaron estar más dispuestas a compartir sus bienes, ser más generosas con ellos y menos envidiosas del éxito ajeno, así como estar menos de acuerdo con las ideas de que el éxito material está relacionado con el éxito vital y de que el bienestar económico conduce a la felicidad (McCullough et al., 2002).

La gratitud podría incluso ayudar a reinterpretar eventos desagradables y mejorar la experiencia de los positivos, crear relaciones personales más fuertes y un crecimiento de las redes sociales (Bartlett, Condon, Cruz, Baumann, & Desteno, 2012; Emmons & Mishra, 2011). Los individuos con una mayor gratitud desarrollan significativamente menos síntomas depresivos a lo largo del tiempo (e.g. Lambert, Fincham & Stillman, 2012) así como una menor tasa de Trastorno de Estrés Postraumático (TEP). La vinculación con el TEP la realizaron Israel-Cohen, Uzefovsky, Kashy-Rosenbaum y Kaplan (2015) a partir de una muestra de adolescentes con bajo a medio nivel socioeconómico de una escuela israelí ubicada en una localidad que estuvo bajo ataque de misiles dos meses y medio antes del estudio.

#### 1.4. Motivaciones de la Conducta Prosocial

Existen perspectivas muy diferentes para entender cuáles son las motivaciones de la conducta prosocial. Un acto prosocial puede estar orientado por el egoísmo, así como por ideales morales o la intención de favorecer al propio grupo (Batson & Powell, 2003). Los motivos del comportamiento prosocial pueden ser incluso, inespecíficos o desconocidos (Beilin, 2013). Es, sin dudas, un fenómeno complejo (Auné, Blum, Abal, Lozzia, & Attorresi, 2014).

Desde una visión biológica, los modelos neo-darwinianos contemporáneos acuerdan en que las tendencias prosociales existen en los humanos por predisposiciones genéticas y porque las personas que actúan de esta forma tienen mayor éxito evolutivo, entendido como la supervivencia de los propios genes a través de diferentes generaciones (Penner et al., 2005).

Ya en 1980, Reykowski y Smolenska contrastan teóricamente la motivación exocéntrica, endocéntrica e intrínseca. La motivación exocéntrica surge de la búsqueda de la ganancia personal o la evitación de pérdida, los motivos endocéntricos se basan en la anticipación de una mejora en la autoestima, mientras que el foco en las necesidades del otro caracteriza la motivación intrínseca.

El enfoque funcional de Clary y Snyder (1991) establece que los individuos se involucran en conductas prosociales cuando tienen motivos definidos, por ejemplo, la expresión de valores personales. El tipo de motivo será esencial para el disfrute y satisfacción que el benefactor sienta posteriormente a la realización de la acción.

Otro enfoque, la teoría del rol social de Grube y Piliavin (2000), sostiene que la motivación para ayudar se basa en influencias externas, incluyendo a la familia y la sociedad en general. En la medida en que el individuo continúa realizando estas actividades, los valores prosociales son adoptados como un componente del *self*. En este sentido, se ha hallado que la fuerza de la identificación predice el compromiso en actividades de voluntariado. Por ejemplo, la identidad como donante de sangre comienza con la percepción de las expectativas de otros, la experiencia pasada de necesidad de recibir sangre y el modelado parental del comportamiento de donación, que llevan a las primeras donaciones. Las primeras experiencias positivas producen una continuidad, que

finalmente resulta en el desarrollo de una faceta del autoconcepto como donante regular de sangre (Piliavin, Grube, & Callero, 2002).

La línea de investigación desarrollada por Omoto, Snyder, Clary y sus colegas (Clary et al., 1998; Omoto & Snyder, 1990), se refiere a la participación y el mantenimiento de las actividades de voluntariado. Distingue seis motivaciones principales que parecen ser su base: valor expresivo (para expresar los valores relacionados con el altruismo y la preocupación humanitaria por los demás), comprensión (para adquirir nuevos conocimientos, y para poder aplicar los conocimientos, destrezas y habilidades), social (estar con amigos y participar en actividades consideradas favorablemente por ellos), de la carrera (para desarrollar actividades que puedan beneficiar la carrera profesional directa o indirectamente), protección (para proteger el ego de uno de los aspectos del yo negativos y obtener ayuda para resolver problemas personales) y mejora (para aumentar los sentimientos positivos hacia uno mismo y el crecimiento y desarrollo personal).

El análisis de las motivaciones del comportamiento de ayuda planeada se ha reducido básicamente a la zona del voluntariado, y las profesiones cuyo rol incluye la conducta de ayuda (por ejemplo, psicología, medicina, enfermería) no han recibido demasiada atención (González Gutiérrez, Peñacoba Puente, Moreno Rodríguez, López López, & Velasco Furlong, 2006). Por el contrario, la perspectiva motivacional que ha despertado más interés en el área de ayuda planeada no coincide con el punto de vista desarrollado por Batson y sus colegas en el ámbito de la actividad no planeada (Batson, 1991; Batson & Powell, 2003). Smith, Kleinbeck, Boyle, Kochinda y Parker (2001) sí han asumido esta perspectiva, aunque en relación con actividades no profesionales. Estos autores desarrollaron tres tipos de motivación para la conducta de ayuda de los cuidadores familiares: altruista, búsqueda de recompensa y motivación castigo-evitación. La motivación altruista implica que el benefactor trata de identificar el receptor de necesidades y ofrece la forma más eficaz de ayudar, aunque esto puede implicar contar con la ayuda de terceros. En la búsqueda de recompensa, el benefactor colabora para obtener algún tipo de beneficio. Por último, quienes están guiados por la motivación castigo-evitación ayudan para evitar sentimientos y pensamientos negativos o las consecuencias negativas del no ayudar.

Los autores de la hipótesis de la empatía-altruismo sostienen que altruismo y egoísmo no son las únicas fuentes de motivación de la conducta prosocial. Batson y

Powell (2003) consideraran además al colectivismo y la motivación moral. El colectivismo supone la intención de beneficiar a un grupo particular en su conjunto. En este sentido, las investigaciones revelan un fuerte sesgo de favoritismo hacia miembros del propio grupo, en contraposición a miembros de otros grupos. Es así que percibir a otro como miembro de un grupo común incrementa la cantidad de acciones de ayuda (Hewstone, Rubin, & Willis, 2002). La motivación moral, en cambio, tiene como objetivo final el sostenimiento de un principio moral universal e imparcial, como por ejemplo la justicia. Estas diferentes motivaciones pueden entrar en conflicto (Batson & Powell, 2003).

### *1.5. Conducta Prosocial y Cultura*

Existe relativamente poca investigación transcultural relevando la generación de diferentes normas de comportamiento prosocial, pero los estudios sugieren que algunos componentes del desarrollo prosocial son compartidos entre las culturas, mientras que otros varían.

Cuando la acción de compartir se examinó entre una diversidad de contextos culturales (por ejemplo, ambientes urbanos ricos y pobres, pequeñas comunidades rurales tradicionales y otras) se replicó la tendencia general de que el autointerés de los tres años de edad se orientó paulatinamente hacia el exterior hasta los cinco (Rochat et al., 2009). Los resultados sugerían una asociación universal entre el desarrollo de la cognición social y el aumento del comportamiento generoso. Sin embargo, existieron diferencias importantes en el nivel de autointerés con el que los niños comenzaron y la magnitud de las diferencias en el desarrollo entre las culturas testeadas.

En otro estudio transcultural (Kärtner, Keller, & Chaudhary, 2010), niños de Berlín y Delhi fueron igualmente capaces de reconocer y responder al estado emocional negativo de un experimentador. No obstante, las dos culturas difirieron en los objetivos de socialización que destacaban y en el papel del desarrollo cognitivo social en la producción del comportamiento reconfortante. En concreto, las madres de Delhi tendían a enfatizar la ventaja de la socialización relacional más fuertemente que las madres de Berlín. Por otra parte, el autorreconocimiento en el espejo correlacionó con el comportamiento reconfortante en los niños de Berlín, pero no en los de Delhi. En

conjunto, estos resultados sugieren que puede haber un número de rutas de desarrollo distintas que conducen a resultados similares de comportamiento.

Aunque la tendencia a realizar conductas prosociales es un universal humano hay una considerable variabilidad cultural en su desarrollo. Asimismo, existe una variación substancial en los niveles de cooperación entre las sociedades que se refleja en la disposición a compartir recompensas económicas con otros y a incurrir en costos económicos para resocializar o castigar al individuo que cometió el comportamiento injusto. La cultura podría ejercer su influencia en el desarrollo del comportamiento prosocial enfatizando diferentes valores y consecuentemente promoviendo oportunidades de socialización particulares (Keller, 2007). Esta variación se encuentra también ligada a variables económicas y demográficas.

Habría una selección de normas e instituciones que mantiene la cooperación incluso en situaciones en que tiene alto costo y con relaciones efímeras. Por ejemplo en grandes urbes, en una situación de violencia entre dos personas, un tercero desconocido puede intervenir para separarlos a pesar de que reciba un golpe.

Existe mucha más variación sistemática a nivel poblacional en la prosocialidad en contextos donde los costos de cooperación son relativamente altos, y menos variación en contextos donde los costos de los actos prosociales son menores. En una sociedad donde los alimentos abundan es más probable que haya más personas dadas con los necesitados que en aquella donde la comida es un bien difícil de conseguir.

Más aún, dependiendo del contexto cultural, es posible que el mismo resultado comportamental pueda emerger a través de diferentes vías (Dunfield, 2014). Mientras que en ciertos contextos el concepto de propiedad privada es resultado de preceptos religiosos, en otros es una norma del Derecho. Es posible concluir que los factores individuales no necesariamente ejercen la misma influencia en todos los tipos de conducta prosocial. Específicamente, es importante considerar el ajuste entre los efectos sociocognitivos o motivacionales de una particular diferencia individual y las demandas de una particular variedad de comportamiento prosocial.

## 1.6. Instrumentos para la Evaluación de la Conducta Prosocial

Existe una diversidad de instrumentos que miden conducta prosocial. Dichos instrumentos pueden tener tres fuentes distintas de información: el propio sujeto (autoevaluación), sus pares, u otros asimétricos; por ejemplo, padres, profesores u experimentador. Los instrumentos más relevantes son los siguientes, ordenados por fecha de aparición en cada apartado:

### 1.6.1. Instrumentos de Autoevaluación

- Batería de Socialización para Adolescentes ([BAS-3] de Silva & Martorell, 1987). Versión autoevaluativa de la BAS para medir dimensiones de orientación interpersonal, aplicable desde los once hasta los diecinueve años. Contiene 75 ítems, siendo la modalidad de respuesta binaria (Sí, No). Se compone de cinco escalas de socialización: consideración hacia los demás, autocontrol en las relaciones sociales, retraimiento social, ansiedad social/timidez y liderazgo. Estas escalas se agrupan en dos factores: conducta prosocial-conducta antisocial, y sociabilidad-insociabilidad. El coeficiente Alfa de Cronbach osciló entre .73 y .82. Fue adaptada en Tucumán (Argentina) con una muestra de adolescentes escolarizados con edades de entre 10 y 15 años (Lacunza, Caballero & Contini, 2013).

- Inventario de Habilidades Sociales para Adolescentes (*Teenage Inventory of Social Skills* [TISS] de Inderbitzen & Foster, 1992). El TISS evalúa la conducta prosocial y agresiva de los adolescentes en las relaciones con sus pares. Consta de 40 ítems agrupados en dos escalas: Conducta Prosocial y Conducta Antisocial. Los ítems se valoran mediante una escala Likert de seis opciones (*No me Describe Nada a Me Describe Totalmente*). El TISS ofrece dos puntuaciones, una para conducta prosocial y otra para conducta antisocial, las cuales se obtienen sumando los valores asignados por los sujetos a los veinte ítems que componen cada dimensión. Puntuaciones altas en cada una de las escalas indican elevada conducta prosocial y agresividad respectivamente (Inglés et al., 2003). La correlación test-retest fue de .90 para la escala positiva y .72 para la escala negativa. Los coeficientes Alfa de Cronbach fueron de .88 para ambas escalas. Por otra

parte, la correlación entre ambas escalas fue de -.26. Las puntuaciones del inventario mostraron evidencias de adecuada validez mediante las relaciones con datos de autoregistro de conductas prosociales y antisociales y con datos sociométricos.

- Escala de Conducta Prosocial (*Prosocial Behavior Scale* [PB] de Caprara & Pastorelli, 1993). Es un test diseñado originalmente para ser aplicado en niños. Es una escala compuesta por 15 ítems que evalúa las conductas de altruismo, de confianza y de agradabilidad a través de tres alternativas de respuesta Tipo Likert en función de la frecuencia con que se den cada una de las conductas descritas. El Alfa de Cronbach para el test completo fue de .71 y mostró evidencias de validez mediante la convergencia con medidas sociométricas y con las puntuaciones en las versiones para padres (Achenbach & Edelbrock, 1983) y para maestros (Achenbach & Edelbrock, 1986) del *Child Behavior Checklist*. Fue adaptada en la Argentina (Urquiza, 2005), con una muestra de adolescentes de doce a quince años de la Ciudad Autónoma de Buenos Aires y la Provincia de Buenos Aires.

- Batería de Personalidad Prosocial (*Prosocial Personality Battery* [PSB] de Penner, Fritzsche, Craige, & Freifeld, 1995). Estos autores desarrollaron la batería de Personalidad Prosocial, de cincuenta y seis ítems, la cual consideran compuesta por dos factores: Empatía Orientada hacia los Otros -*Other Oriented Empathy*, la tendencia a experimentar empatía cognitiva y afectivamente y sentir la responsabilidad y preocupación por el bienestar de los demás- y Tendencia a Ayudar -*Helpfulness*, una autodescripción acerca del involucramiento en actividades de ayuda-. El coeficiente Alfa de Cronbach de ambos factores sobrepasó los .80. La confiabilidad test-retest fue de .77 en el factor de empatía orientada hacia los otros y de .85 en el factor de tendencia a ayudar. Las evidencias de validez fueron: correlaciones positivas y significativas de los puntajes con el *Helping-Orientation Questionnaire* (Romer, Gruder, & Lizzardo, 1986).

- Escala de Tendencias Prosociales (*Prosocial Tendencies Measure* [PTM] de Carlo & Randall, 2002). Se aplicó a una muestra de adolescentes tardíos (Media de edad = 19.9 años). Con 23 ítems en total, está compuesta por seis dimensiones, que reflejan diferentes tipos de comportamiento prosocial: público (cuatro ítems, Alfa de Cronbach = .78), anónimo (cinco ítems, Alfa de Cronbach = .85), en situaciones de urgencia (tres

ítems, Alfa de Cronbach = .63), emocional (cuatro ítems, Alfa de Cronbach = .75), complaciente (dos ítems, Alfa de Cronbach = .80) y altruista (cinco ítems, Alfa de Cronbach = .74). Los ítems se valoran mediante una escala Likert de cinco opciones (1: no me describe en absoluto; 5: me describe en gran medida). Mostró adecuadas evidencias de validez de constructo mediante análisis factorial, convergente y discriminante. En cuanto a las evidencias de validez mediante la convergencia, se obtuvieron correlaciones positivas significativas con medidas de toma de perspectiva (*Multidimensional Measure of Empathy* [MME]; Davis, 1983), razonamiento estereotipado (*Prosocial Moral Reasoning* [PROM-R]; Carlo, Eisenberg, & Knight, 1992; Eisenberg, Carlo, Murphy, & Van Court, 1995) adscripción de la responsabilidad (*Ascription of Responsibility*; Schwartz, 1968) responsabilidad social (*Social Responsibility*, Berkowitz & Lutterman, 1968) y comportamiento prosocial global (*Global Prosocial Behavior*; Rushton, Chrisjohn, & Fekken, 1981). En cuanto las evidencias de validez mediante la divergencia, se obtuvo una correlación negativa y significativa con razonamiento hedonista (PROM-R) y nula con razonamiento orientado a la aprobación (PROM-R), razonamiento orientado a las necesidades (PROM-R), angustia (MME), deseabilidad social (*Social Desirability*; Crowne & Marlowe, 1964) y habilidades de vocabulario (*Quick Word Test*; Borgatta & Corsini, 1960).

- Escala para la Medición de la Prosocialidad en Adultos (*Prosocialness Scale for Adults* [PSA] de Caprara et al., 2005). Es una escala de dieciséis ítems, Tipo Likert y con cinco opciones de respuesta. Cada ítem refleja comportamientos y sentimientos que pueden ser incluidos en una de cuatro clases de acciones: de ayuda, de compartir, de cuidar, y ser empático con los otros. El Alfa de Cronbach para el test completo fue de .91 y la media de la correlación ítem-test corregida, .59. Mediante Teoría de la Respuesta al Ítem se determinó que la mayoría de los ítems requieren un bajo nivel de prosocialidad para ser respondidos en forma favorable y tienen menos poder de discriminación en personas con medio-alto o alto nivel de prosocialidad.

- Cuestionario de Conducta Prosocial (Sánchez Queija, Oliva & Parra, 2006). Es un cuestionario destinado a adolescentes y consta de siete ítems, a responder en una escala Likert de uno a cuatro si se han realizado las siguientes conductas: consolar a un chico o chica (par) que estaba triste o deprimido, quedar al cuidado de algún niño pequeño sin

recibir dinero, dar dinero a instituciones benéficas, realizar tareas sociales de voluntariado en organizaciones no gubernamentales, ayudar a personas con deficiencias físicas o sensoriales, devolver dinero o algún objeto perdido a un desconocido y, por último, cuidar a enfermos o personas mayores. El coeficiente Alfa de Cronbach fue de .67.

- Escala para la Medición de la Conducta Prosocial-Antisocial en el Ámbito Vital y en el Tráfico (López de Cózar et al., 2008). Es un test para adultos. Cuenta con las siguientes dimensiones: prosocialidad vital, prosocialidad vial, altruismo, asertividad y empatía. Se ha elaborado un inventario de comportamientos representativos de las dimensiones, evaluando la frecuencia con que los entrevistados realizan dichas conductas en una escala de 1 (*Nunca*) a 4 (*Siempre*). Las escalas con mayor consistencia interna fueron la de prosocialidad vital (Alfa de Cronbach = .81) y la de empatía (Alfa de Cronbach = .77), aunque todas obtuvieron un Alfa de Cronbach superior a .5.

- Escala de Habilidades Prosociales para Adolescentes ([EHP-A] de Morales Rodríguez & Suárez Pérez, 2011). La escala, destinada a adolescentes, se encuentra conformada por 20 ítems, distribuidos cuatro factores: Toma de Perspectiva; Solidaridad y Respuesta de Ayuda, Altruismo y Asistencia. Los ítems se valoran mediante una escala Likert de cuatro opciones (*No Me Describe*, *Me Describe Poco*, *Me Describe Regular* y *Me Describe Bien*). El coeficiente Alfa de Cronbach de la escala fue de .86. Se obtuvieron evidencias de validez mediante la convergencia con la Escala de Autoeficacia General (EAG), adaptación española de Baessler y Schwarzer (1996). Fue adaptada por Auné, Abal y Attorresi (2014) con una muestra de estudiantes universitarios de la Ciudad de Buenos Aires.

#### 1.6.2. Instrumentos de Heteroevaluación

La evaluación externa al propio sujeto se utiliza principalmente en niños y, en segundo lugar, en adolescentes. En niños las medidas de autoinforme no se consideran válidas por diversos motivos: dificultades de comprensión y lectura, así como escasa capacidad de autoobservación (Guijo Blanco, 2003).

-Evaluación por pares: Se realiza principalmente a través de dos tipos de instrumentos, cuestionarios estructurados y sociogramas. Las técnicas sociométricas son las que tienen la mayor validez predictiva, manifestando una alta correlación entre sí. Se basan en preguntar a todos los miembros del grupo acerca del resto (Guijo Blanco, 2003). El test sociométrico es un instrumento que evalúa las relaciones sociales entre los pares, que pueden ser de atracción o rechazo. Indaga sobre distintos tipos de liderazgo en cinco situaciones sociales: conocimiento, deporte, paseo/fiesta, peligro y tristeza (Maya & Fraile, 1996).

-Evaluación por otros asimétricos: Se utilizan principalmente escalas de apreciación, observación sistemática en contextos naturales y experimentales, entrevistas semiestructuradas e inventarios estructurados para la observación del niño o adolescente (Eisenberg & Fabes, 1998; Guijo Blanco, 2003). Algunos de los inventarios estructurados más importantes son, según orden de aparición:

- Cuestionario de Conducta Prosocial (*Prosocial Behaviour Questionnaire* [PBQ] de Weir & Duveen, 1981). El cuestionario contiene veinte ítems que describen conductas prosociales. El docente informa de las conductas que observa en sus alumnos y los padres valoran la conducta de su hijo en el ambiente familiar. Se aplica a la etapa adolescente. La modalidad de respuesta es una escala Likert de tres opciones (*No Aplica, A Veces Aplica, Ciertamente Aplica*). El coeficiente Alfa de Cronbach fue de .93. Se obtuvieron evidencias de validez mediante el criterio externo de la comparación de juicios independientes de pares, es decir de otros adolescentes.

- Escala de Comportamientos del Niño (*Child Behavior Scale* [CBS] de Ladd & Profilet, 1996). Es una medida de los comportamientos de agresión, retraimiento y conducta prosocial en niños, a completar por el docente. Está compuesta por 17 ítems agrupados en dos escalas, Agresividad con los Pares y Prosocialidad con los Pares. La modalidad de respuesta es una escala Likert de tres opciones (*No Es Verdad, A Veces Es Verdad, Siempre Es Verdad*). El coeficiente Alfa de Cronbach para ambas escalas fue de .93. Se obtuvieron evidencias de validez mediante la correlación positiva y significativa con comportamientos prosociales observados, con las puntuaciones en la versión para

maestros del *Child Behavior Checklist* (Achenbach & Edelbrock, 1986) y mediante la correlación negativa y significativa con comportamientos agresivos observados.

### 1.6.3. Instrumento Mixto (Auto y Heteroevaluación)

- Cuestionario de Conducta Prosocial (*Prosocial Behavior Questionnaire*; [CCP y CCP-H] de Martorell, González, Aloy, & Ferris, 1995). El inventario contiene dos partes: una autoevaluación [CCP], de 55 ítems y una evaluación realizada por padres, docente u otras personas cercanas al sujeto con una formulación similar, pero en tercera persona [CCP-H]. Todos los reactivos son valorados en una escala Likert de cuatro opciones (*Nunca, Alguna Vez, Muchas Veces y Siempre*). Está orientado a dirigido a niños mayores y adolescentes de entre 10 y 17 años. El conjunto de factores hallados -Empatía, Respeto, Relaciones Sociales y Liderazgo) presentan una buena consistencia interna (Alfa de Cronbach entre .76 y .87). La consistencia de cada uno de los factores considerados por separado fue como mínimo de .72.

## 2. Psicometría

### 2.1. Evaluación Psicológica

Evaluación psicológica es la traducción del término anglosajón *psychological assessment*. Comúnmente se utiliza la evaluación psicológica para explorar las características del psiquismo de un individuo, incluidas sus capacidades, aptitudes, actitudes, rasgos y estados, a través de tests, escalas y cuestionarios. Entre los objetivos de evaluación, en su práctica concreta de aplicación, se encuentran: a) describir, que es explicitar las características más sobresalientes del comportamiento de un sujeto a partir de los datos obtenidos a través de las diferentes técnicas de análisis; b) clasificar, que implica la organización de la información en función de criterios previamente fijados, y c) predecir, que es establecer una previsión probabilística de una conducta futura a partir de una conducta actual (Fernández Ballesteros, 1995; Mikulic, 2007). La Psicometría es la disciplina que estudia las condiciones que permiten llevar a cabo el proceso de medición de variables psicológicas estableciendo las bases para que dicho proceso se realice de manera adecuada (Martínez Arias, 1995).

Se define a las pruebas psicológicas como medidas objetivas y estandarizadas de una muestra de conducta (Anastasi & Urbina, 1998). Las técnicas psicométricas sirven precisamente para operacionalizar de manera cuantificable constructos psicológicos (Cohen & Swerlick, 2006). La interpretación de las mismas suele realizarse comparando las cualidades halladas en los individuos de estudio con las usualmente encontradas en la población de referencia. Se trata de verificar si los principios generales establecidos por los cuerpos teóricos se corroboran en un sujeto individual (Fernández Ballesteros, 1995).

Para construir una prueba psicológica, Loevinger (1957) sitúa tres fases: en primer lugar, en la fase *teórico-sustantiva*, se construye un *pool* de ítems, los cuales se depuran mediante juicio experto y pruebas piloto. Posteriormente, en una segunda fase denominada *interna-estructural*, se analiza la estructura interna del instrumento. Las puntuaciones deben presentar las primeras evidencias de validez y confiabilidad de sus resultados (Cohen & Swerlick, 2006) mediante las relaciones entre los ítems y el constructo. Por último, en la fase *externa-validacional* se aportan evidencias adicionales de validez a través de la relación de las puntuaciones con otros constructos o con criterios externos. Es evidente que la validez y la confiabilidad son las características más relevantes de las puntuaciones de una prueba psicológica.

El concepto de validez ha ido variando a lo largo del tiempo. Para este punto se seguirá el desarrollo histórico acerca de la validez realizado por Gómez Benito e Hidalgo (2015). Las autoras sitúan un concepto clásico o estático de validez según el cual lo que se valida es una prueba, que es considerada válida cuando mide lo que pretende medir o sirve para el propósito por el que fue construida. La APA configura en 1974 una estructura tripartita de la validez: contenido, criterio y constructo. En los años '80 se sustenta fuertemente que la validez es un proceso unitario y se refiere al constructo: "La validez es la consideración más importante en la evaluación de un test. El concepto se refiere a la adecuación, significado y utilidad de las inferencias específicas hechas con las puntuaciones de los tests. La validación de un test es el proceso de acumular evidencia para apoyar tales inferencias. Una variedad de evidencias puede obtenerse de las puntuaciones producidas por un test dado, y hay muchas formas de acumular evidencia para apoyar una inferencia específica. La validez, sin embargo, es un proceso unitario. Aunque la evidencia puede ser acumulada de muchas formas, la validez se refiere siempre al grado en que esa evidencia apoya las inferencias que se hacen a partir de las puntuaciones" (APA, AERA, NCME, 1985, p. 8; citado en Gómez Benito & Hidalgo, 2015).

Elosúa Oviden (2003) y Gómez Benito e Hidalgo (2015) sitúan que en la actualidad se considera que lo que se valida no es el test sino sus puntuaciones. La validez se trata de la adecuación de un determinado uso o interpretación de las puntuaciones de un test. La validación es un proceso continuo y dinámico que se asegura mediante la acumulación de evidencia teórica, estadística, empírica y conceptual del uso de las puntuaciones, no pudiéndose resumir en un sólo indicador o índice numérico. Establecer la validez de las puntuaciones test implica también obtener evidencia de que las mismas se encuentran libres de Funcionamiento Diferencial del Ítem. El Análisis Factorial y la Teoría de la Respuesta al Ítem son importantes herramientas que aportan evidencias de validez.

Las evidencias de validez pueden ser aportadas por distintas fuentes. Estas distintas fuentes se dividen globalmente en internas, referidas a un análisis individualizado de los ítems, y externas, que analizan el test en su conjunto (Elosúa Oviden, 2003). Dentro de las fuentes de evidencia interna Elosúa Oviden (2003) sitúa: el contenido, el proceso de respuesta y la estructura interna. Las fuentes de estructura interna se centran en los estudios de dimensionalidad y los estudios de Funcionamiento

Diferencial del Ítem. Las fuentes de validez externas se basan en relacionar las puntuaciones de la escala a validar con otra variable. La evidencia convergente y discriminante, las relaciones de las puntuaciones de la escala con un criterio externo y los estudios de generalización de la validez son las tres fuentes externas de validez que sitúa la autora.

Las evidencias de validez y confiabilidad se enmarcan en una determinada Teoría de los Tests. Las Teorías de los Tests componen una subárea de la Psicometría encargada de proporcionar modelos sobre los que se basa la construcción de tests que luego serán aplicados en la Evaluación Psicológica de las características cognoscitivas (e.g. Cortada de Kohan & Macbeth, 2007), afectivas y conductuales del sujeto. Estas teorías permiten formalizar la relación existente entre la puntuación observada de un sujeto y la cantidad de rasgo que éste posee. Actualmente coexisten dos teorías psicométricas. La Teoría Clásica de Tests (TCT) comprende básicamente el modelo lineal de Spearman (1904) y su extensión, la Teoría de la Generalizabilidad. La Teoría de Respuesta al Ítem (TRI) es la segunda teoría, la cual toma al reactivo como unidad de análisis. Las Teorías de los Test aportan modelos teóricos, donde la elección entre los mismos se encuentra sujeta a las posibilidades de ajuste a los datos. En el siguiente apartado se desarrollan los conceptos básicos de ambas teorías, comenzando por la TCT.

## 2.2. Teoría Clásica de los Test

La TCT comprende un conjunto de conceptos y técnicas que han sido utilizados como base para el desarrollo de diversos instrumentos de medida y como punto de referencia para el abordaje de medidas psicométricas. Se basa en el modelo lineal de puntuaciones formulado por Spearman en 1904 que supone que el puntaje  $X$  obtenido por un sujeto en un test es el resultado de la suma de un puntaje verdadero  $V$  y un error de medida  $\varepsilon$  (Attorresi, Galibert, & Aguerri, 2002; Spearman, 1910).

Su formulación es:

$$X_g = V_g + \varepsilon_g$$

Donde  $X_g$ ,  $V_g$  y  $\varepsilon_g$  son *variables aleatorias* sobre una población de individuos a los que se administra un test  $g$ .

$X_g$  es el puntaje observado de los individuos.

$V_g$  es el puntaje verdadero de los individuos.

$\varepsilon_g$  es el error de medición.

Fijado el sujeto, su puntaje verdadero es un valor fijo pero tanto el puntaje observado como el error de medición son variables aleatorias.

La puntuación verdadera  $V$  se podría considerar como el promedio de los puntajes que el individuo obtendría en un número infinito de test paralelos. Es el valor real del puntaje de la persona en el rasgo evaluado.

Se dice que dos medidas, representadas por las variables  $X_g$  y  $X_h$ , son medidas paralelas si se cumple que:

$$X_g = V_g + \varepsilon_g \text{ y } X_h = V_h + \varepsilon_h$$

$$\text{con } V_g = V_h = V \text{ y } \text{Var}(\varepsilon_g) = \text{Var}(\varepsilon_h)$$

Son supuestos a esta formulación (Muñiz Fernández, 2010): a) Para cada puntaje verdadero, los errores de medición tienen media cero, es decir que con la medición no se tiende sistemáticamente a sobreestimar ni a subestimar el puntaje verdadero, b) La varianza de los errores no depende del puntaje verdadero, c) Los errores no están correlacionados entre medidas paralelas.

Son propiedades derivadas: a) El promedio de todos los puntajes observados de los individuos con un mismo puntaje  $V$  coincide con  $V$ :  $E(X) = V$ , donde  $E(X)$  es la esperanza o el valor esperado de la variable  $X$ , b) Los errores tienen media cero sobre toda la población de individuos:  $E(\varepsilon) = 0$ , c) Los puntajes verdaderos no covarían con los errores:  $\text{Cov}(V, \varepsilon) = 0$ , donde  $\text{Cov}$  es la covarianza entre dos variables, d) Los puntajes observados y los verdaderos son iguales en promedio:  $E(X_g) = E(V_g)$ , e) La varianza de los puntajes observados es la suma de las varianzas de los verdaderos más la de los errores:  $\text{Var}(X) = \text{Var}(V) + \text{Var}(\varepsilon)$  dado que  $\text{Cov}(V, \varepsilon) = 0$ .

Aunque ha sido muy útil y continúa siendo ampliamente utilizada, sus limitaciones han sido señaladas hace ya muchas décadas (e.g. Gulliksen, 1950; Rasch, 1961). Actualmente, diversos trabajos han presentado y discutido estas limitaciones (e.g. De Champlain, 2010; Embretson & Reise, 2000).

La principal limitación de la TCT es que todas las medidas son dependientes de la muestra de individuos que respondieron al instrumento. Esto significa que las estimaciones son válidas solamente para esa muestra u otra con características semejantes (Embretson & Reise, 2000; Hambleton & Slater, 1997). Además, instrumentos con índices de dificultad y discriminación diferentes generan resultados también diversos en los mismos individuos. Por último, otra limitación de la TCT es que, si el mismo constructo es medido por dos o más test diferentes, los resultados no son medidos en la misma escala, dificultando o directamente impidiendo toda comparación directa.

Estas limitaciones encuentran respuesta en un nuevo enfoque psicométrico: la Teoría de la Respuesta al Ítem (TRI).

### 2.2.1. Evidencias de Confiabilidad en el marco de la TCT

En este apartado se desarrollará específicamente la confiabilidad referida a la consistencia interna. La estimación del error de medida  $\varepsilon$  se realiza a través del coeficiente de confiabilidad (Muñiz Fernández, 1996). Desde la TCT, el coeficiente de fiabilidad es una medida global de una determinada prueba y se asume su igualdad para todas las personas a las que se la aplica (Muñiz Fernández, 1998). Dicho coeficiente puede ser, simplifícadamente, una correlación entre mitades del test o la covarianza entre todos los ítems. El método de covarianza más utilizado es el Coeficiente Alfa de Cronbach (1951) que equivale a la media de los coeficientes basados en todas las mitades posibles. Es una función directa de las covarianzas entre los ítems. Su formulación es la siguiente:

$$\alpha = \frac{n}{n-1} \left( \frac{\sum_{j \neq k}^n \sum^k Cov(j, k)}{\sigma^2_X} \right) = \frac{n}{n-1} \left( 1 - \frac{\sum_{k=1}^n \sigma^2_k}{\sigma^2_X} \right)$$

Donde:

$n$  es el número de ítems,

$\sigma^2_X$  es la varianza total, y

$\sigma^2_k$  es la varianza del ítem  $k$ .

Cualquier ítem cuya correlación con el puntaje total de la prueba sea baja, nula o negativa generará una reducción del Alfa de Cronbach. Asimismo, una baja cantidad de ítems y de sujetos generan una reducción en este índice.

### 2.3. Teoría de la Respuesta al Ítem

La construcción y adaptación de instrumentos de evaluación psicológica demanda modelos psicométricos capaces de analizar adecuadamente sus propiedades. La TRI es una familia de modelos novedosos para la medición psicológica y educativa que permite el desarrollo de una nueva tecnología para la construcción y análisis de los tests (Muñiz Fernández, 1997). Sus pioneros son Rasch (1960) y Birnbaum (1968).

La TRI se basa en el concepto de *rasgo latente* (Lazarsfeld, 1950), habitualmente designado con la letra griega theta ( $\theta$ ). Es una variable que puede ser descripta y los expertos pueden ahondar en sus atributos, y sin embargo no puede ser medida directamente (Baker, 2001), por ejemplo, la inteligencia, el altruismo o la conducta prosocial. Las puntuaciones de los sujetos en un test estiman  $\theta$  (Muñiz Fernández, 1997). El objetivo más importante de la TRI es la construcción de instrumentos de medición con propiedades invariantes entre poblaciones. Todo individuo con igual nivel de rasgo tiene la misma probabilidad de acertar o de elegir cierta categoría de respuesta, independientemente de la población de pertenencia (Attorresi et al., 2009). Un importante objetivo de la medición mediante la TRI es caracterizar la relación entre el  $\theta$  de un sujeto y la probabilidad de dicho sujeto de dar una determinada respuesta al ítem. Para ello, se utilizan funciones matemáticas que describen la probabilidad de dar una determinada respuesta al ítem para cada nivel del rasgo medido por este. Estas funciones se denominan Curvas Características del Ítem (CCI) para las pruebas con respuesta dicotómica y Curvas Características de las Categorías de Respuesta al Ítem (CCCRI) para las pruebas con respuesta politómica (Attorresi et al., 2009; Samejima, 1969).

Mientras que la TCT considera las propiedades globales de las escalas, la TRI toma al ítem como unidad de análisis. En la TRI cada ítem se caracteriza por sus parámetros independientemente de la distribución de  $\theta$  en la población y los parámetros de los otros ítems del test. La medida de una persona también es independiente de los parámetros de los ítems y de la medida de los otros participantes (Lord, 1980). Esta

independencia entre el instrumento de medición y los individuos es la diferencia esencial entre los enfoques de la TRI y de la TCT.

Los modelos tradicionales de la TRI se sostienen en dos supuestos básicos: unidimensionalidad e independencia local.

La unidimensionalidad supone que todos los ítems miden el mismo rasgo latente  $\theta$ . Este supuesto nunca se cumplirá totalmente y se considera satisfecho cuando hay un rasgo dominante que explica las respuestas de los individuos y las relaciones entre los ítems. El DIF se opone a la unidimensionalidad ya que si entre sujetos con el mismo nivel de rasgo, uno tiene más probabilidad de responder en forma favorable o correcta a un determinado ítem, esto implica que hay por lo menos un segundo factor subyacente a dichas respuestas. Existen distintas formas de verificar el supuesto de unidimensionalidad, siendo las más utilizadas:

a) Evidencias aportadas por el Análisis Factorial Exploratorio.

- 1- Según el criterio de Kaiser, para que el test sea aceptado como unidimensional debe existir un solo autovalor superior a uno. Sin embargo, es conocido que la regla de Kaiser ha dado lugar a sobredimensionar la solución factorial cuando, como en este caso, las variables son ítems, ya sean dicotómicos o politómicos (Navarro Loli & Merino Soto, 2010; Ruiz & San Martín, 1992).
- 2- Según el criterio de Martínez Arias (1995) la razón del primer autovalor respecto al segundo debe ser igual o mayor a 5.
- 3- El diagrama de autovalores es confiable en general para determinar cuántos factores deben ser interpretados (Pérez & Medrano, 2010). Se basa en localizar gráficamente en el diagrama de autovalores un punto en que los mismos dejan de formar una pendiente pronunciada y comienzan a generar una caída de poca inclinación (Pérez & Medrano, 2010). Según el criterio de caída de Catell la presencia de un salto brusco en el gráfico de autovalores *-scree plot-* entre el primer autovalor y los restantes es indicador de unidimensionalidad. Actualmente se han explorado mejoras

gráficas y computacionales con el fin de hacer más válido e informativo al gráfico de autovalores (Ledesma, Valero-Mora, & Macbeth, 2015).

- 4- Criterios relacionados con el porcentaje de varianza explicada por el modelo. Estos criterios consideran a una escala unidimensional según el porcentaje de varianza que el primer factor explica. Hay una variación entre diversos autores en el porcentaje de varianza mínimo explicado por el primer factor necesario para que la escala sea considerada unifactorial:
- El 60% de la varianza (Hair, Anderson, Tatham, & Black, 1999).
  - El 40% de la varianza (Carmines & Zeller, 1979).
  - El 20% de la varianza (Reckase, 1979).
  - Entre el 17 y 40% de la varianza, si se utilizan matrices de correlaciones phi (Zwick, 1985).
  - Entre el 30% a 40% de la varianza, usando matrices de correlaciones tetracóricas (Zwick, 1985).

b) Evidencias aportadas por el Análisis Factorial Confirmatorio. Mediante Análisis Factorial Confirmatorio, se impone una estructura unifactorial. En el caso de contar con adecuados indicadores de ajuste al modelo, se considera corroborada la unidimensionalidad.

c) Test Estadístico de la Unidimensionalidad Esencial de Stout. El procedimiento DIMTEST fue desarrollado por Stout (1987, 1990) para examinar la existencia de unidimensionalidad mínima y necesaria para satisfacer la exigencia de independencia esencial. Se recomienda realizar este procedimiento con test de más de 25 ítems y con tamaños de muestra de por lo menos 500 participantes (De Champlain & Gessaroli, 1991).

El supuesto de independencia local se refiere a que para un nivel del rasgo  $\theta$  las respuestas a distintos ítems son estadísticamente independientes; es decir, que la probabilidad de dar una determinada respuesta a un ítem no se modifica si se conoce la respuesta de la misma persona a otro ítem (Muñiz Fernández, 2010). Atentan contra la independencia local la dependencia lógica de los ítems. Por ejemplo: ítems muy similares,

paralelos; ítems enganchados -matemática- e ítems referidos al mismo contenido -lengua. Se demuestra matemáticamente que de la unidimensionalidad deriva la independencia local, no así a la inversa (Muñiz Fernández, 1997).

La TRI ofrece modelos específicos para reactivos de respuesta dicotómica (Birnbaum, 1968, Rasch, 1960) y politómica (Bock, 1972; Masters, 1982; Muraki, 1997; Samejima, 1969; Thissen & Steinberg, 1997).

### 2.3.1. Evidencias de Confiabilidad en el Marco de la TRI

Es conocido empíricamente que una misma escala no mide con igual precisión a todas las personas, variando su precisión en función del nivel del individuo en el rasgo evaluado (Muñiz Fernández, 1998). La solución radical a esta cuestión es aportada por la TRI donde la confiabilidad se expresa mediante la Función de Información (FI). En la TRI un ítem es más confiable para determinados niveles de  $\theta$  en los que aporta más información.

Cada ítem aporta a una FI que indica su precisión para estimar los valores de  $\theta$  en los sujetos. La FI es el recíproco del Error Estándar de Medición (EEM) al cuadrado, este último se define como la variabilidad de las estimaciones de  $\theta$  para cada individuo (Montero, 2000). El EEM varía entre personas con diferentes patrones de respuesta, pero el mismo nivel de rasgo tiene igual EEM, invariante por población (Embretson & Hershberger, 1999). La FI de un test se define como la suma de las funciones de información de los ítems que lo componen. También se puede obtener una función de error típico de la estimación del test.

Es posible identificar para qué niveles de rasgo latente las puntuaciones de la escala son más confiables (Attorresi et al., 2009). La TRI permite calcular la confiabilidad marginal, una medida de precisión global del test (Thissen, 1991). Este estimador es un promedio de la fiabilidad a través de todos los niveles del rasgo ponderado por la proporción de individuos que corresponden a cada nivel.

### 2.3.2. Aplicaciones de la TRI a Tests de Comportamiento Típico

Según la clasificación de Cronbach (1972) existirían básicamente dos tipos de comportamientos: aquellos orientados a conducir a un desempeño máximo y otros que

involucran una forma típica. Los comportamientos de desempeño o ejecución máxima se refieren a aquello que los individuos pueden hacer cuando tratan de alcanzar su mejor rendimiento. Las medidas de habilidad cognitiva son un ejemplo de evaluación en el dominio de la ejecución máxima. El comportamiento típico, en cambio, se refiere a lo que los individuos hacen usualmente, aquí el foco está en lo que una persona practica de forma ordinaria, esto es, sus tendencias generales de comportamiento. Las escalas de personalidad, interés vocacional y actitudes son medidas de comportamiento típico.

En los test de ejecución máxima la meta es obtener el mayor puntaje posible para demostrar que se posee la habilidad (Cho, Drasgow, & Cao, 2015). Los participantes tratan de alcanzar este objetivo respondiendo cada ítem correctamente y tienden a hacerlo cuando su nivel de habilidad o conocimiento es suficientemente alto para sobrepasar el nivel de dificultad del reactivo. La estrategia de respuesta para medidas de comportamiento típico es diferente. Los participantes no están habitualmente interesados en obtener un puntaje alto sino en tratar de describirse respondiendo a los ítems de una forma que refleje su nivel percibido en la característica medida (Cho et al., 2015). Sin embargo, existe en este tipo de medidas una mayor complejidad en tanto que pueden existir respuestas distorsionadas por distintos motivos -por ejemplo, en el ámbito laboral, con la intención de obtener un puesto- e influencia de la deseabilidad social.

La aplicación de modelos de la TRI para el análisis de ítems de test de ejecución máxima ha sido mucho más frecuente que para test de comportamiento típico (Embreston & Reise, 2000; Reise & Henson, 2003; Rojas & Pérez, 2001). En este último campo se encuentra atrasada y esto parece deberse a múltiples factores que se conjugan entre sí: problemas asociados al surgimiento y difusión de la TRI, problemas vinculados a la demanda social, barreras técnicas tales como el cumplimiento de los supuestos de la TRI; la complejidad teórica de los constructos medidos con test de ejecución típica y por el hecho de que muchos de los test fueron únicamente diseñados y validados a partir de la TCT (Abal et al., 2010).

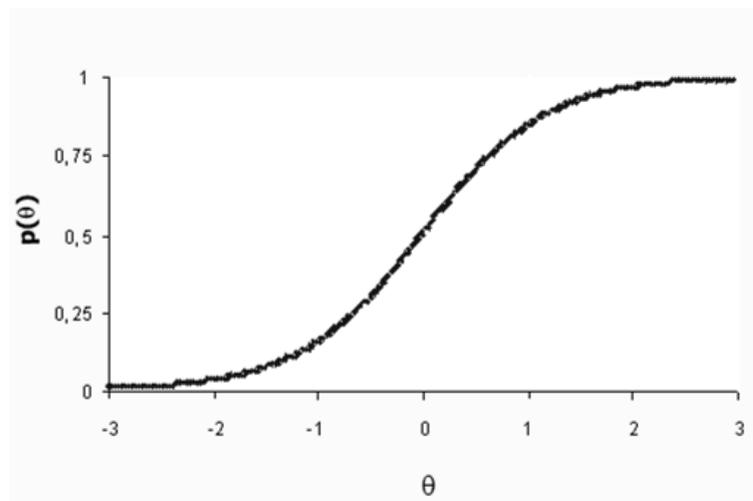
A pesar de estas dificultades en los últimos años se ha incrementado el análisis de los ítems de tests de comportamiento típico mediante los modelos de la TRI. Cobró para ello suma importancia estudiar las características del ítem y cómo se responde. Existen modelos de la TRI específicos para ítems dicotómicos, -aquellos que tienen como formato sólo dos categorías aunque puedan tener varias opciones de respuesta-, y otros adecuados para ítems politómicos -aquellos que tienen tres o más categorías-.

### 2.3.3. Modelos de la TRI para Ítems Dicotómicos

La CCI es una función matemática que describe la probabilidad de dar una determinada respuesta al ítem para cada nivel del rasgo medido por este en las pruebas con respuesta dicotómica (Fraley, Waller, & Brennan, 2000). Específicamente, describe la probabilidad de dar una respuesta favorable al rasgo si es una prueba de comportamiento típico, o correcta en el caso de test de ejecución máxima. Cada ítem tiene su propia CCI (Baker, 2001), de allí su nombre: es característica, típica de cada ítem en particular (Muñiz Fernández, 1997). En ítems de test de ejecución máxima e ítems de pruebas de comportamiento típico con respuesta dicotómica se trata de funciones monótonicamente crecientes. Es decir, la probabilidad de contestar correctamente un ítem se incrementa en forma creciente a lo largo del continuo del rasgo latente. Las curvas son no lineales, sino sigmoideas cuya imagen por ser una probabilidad varía entre 0 y 1. Para determinar la curva se especifican hasta tres parámetros (Muñiz Fernández, 1997). Los modelos utilizan la función logística, más accesible matemáticamente que la curva normal y además adicionando una constante  $D = 1.7$  sus valores dan la mejor aproximación a los de la curva normal acumulada (Muñiz Fernández, 1997).

Figura 1

*Curva Característica del Ítem*



Curva Característica de un Ítem con Parámetro  $b = 0$

Según el tipo de función matemática adoptada para la CCI y el número de parámetros considerados se definen diferentes modelos de TRI para ítems dicotómicos. Se ha utilizado inicialmente la función normal y posteriormente la logística (Santisteban, 1990) dada su mayor simplicidad matemática, de allí su utilización más extendida (Muñiz Fernández, 1997). El Modelo Logístico de Un Parámetro (ML1P) o Modelo de Rasch contempla únicamente al parámetro de localización  $b$ , llamado también parámetro de dificultad en las pruebas de ejecución máxima. El Modelo Logístico de Dos Parámetros (ML2P) o de Birnbaum queda determinado además de por el parámetro de localización  $b$  por el parámetro de discriminación  $a$ , mientras que en el Modelo Logístico de Tres Parámetros (ML3P) la CCI está caracterizada por el parámetro de localización, el de discriminación y el de pseudo-azar  $c$ . Las CCIs delimitadas por cada uno de estos modelos difieren en general en su ubicación, pendiente y asíntota. Cuando  $c = 0$  el ML3P se convierte en un ML2P y cuando  $a = 1$  el ML2P se convierte a ML1P.

El parámetro  $b$ , llamado de localización o dificultad es la primera coordenada (la del rasgo latente  $\theta$ ) del punto de inflexión de la CCI. El punto de inflexión es aquel punto donde cambia la concavidad de la curva. En ML1P y ML2P es el punto en la escala de  $\theta$  donde la probabilidad de dar una respuesta favorable al rasgo o correcta es 0.5. De esta forma, posiciona al ítem en la escala del rasgo: a mayor  $b$ , mayor será el nivel de rasgo necesario para que el sujeto tenga una probabilidad de 0.5 de responder de forma favorable al rasgo o correcta (Cortada de Kohan, 2004; Cortada de Kohan, Macbeth, & López Alonso, 2008). En el ML3P  $b$  es aquél punto en la escala de  $\theta$  donde la probabilidad de dar una respuesta favorable al rasgo o correcta es igual a  $0.5 + c/2$ . A mayor valor del parámetro  $b$ , esto indicará que una respuesta favorable requiere un nivel más alto del rasgo. Cuando una persona responde a un ítem con localización equivalente a su umbral del rasgo, tendrá la misma probabilidad de dar una respuesta favorable al rasgo o correcta como desfavorable al rasgo o incorrecta.

El parámetro  $a$ , denominado índice de discriminación, es proporcional (si  $c$  es constante para todos los ítems en el ML3P) a la pendiente de la recta tangente en el punto de inflexión de la CCI.

En general, no necesariamente hay proporcionalidad entre la derivada de la pendiente de la recta tangente en  $b$  y el parámetro  $a$ :

$$P'(b) = \frac{D}{4} (1 - c_i) a$$

$P'(b)$  es la derivada de la función correspondiente a la CCI en el punto  $b$ .

Si  $c_i$  es constante en todos los ítems, el parámetro  $a$  es proporcional a la derivada de la pendiente de la recta tangente en  $b$ . Por lo tanto:

Si  $c_i = c \forall_i$ , es decir  $c_i$  es constante.

$$P'(b) = \frac{D}{4}(1 - c)a \propto a$$

Si  $c$  es nulo en todos los ítems, resulta:

Si  $c = 0 \forall_i$

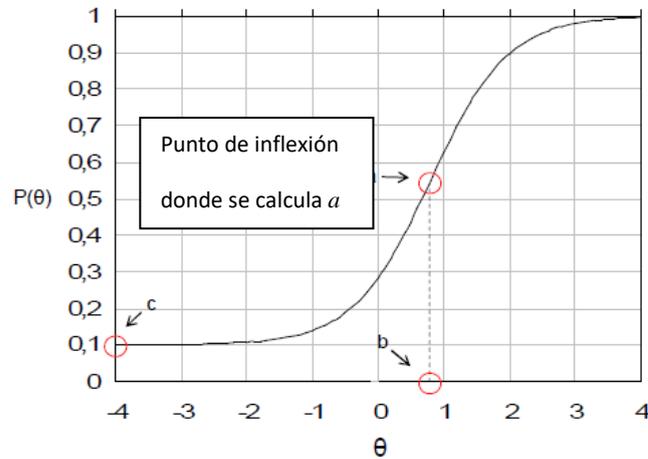
$$P'(b) = \frac{D}{4}a \propto a$$

En un entorno de  $b$  es donde se da la mayor potencia discriminativa del ítem, es decir donde es mayor la pendiente de la recta tangente, que indica máxima variabilidad. Su valor puede oscilar entre 0 y  $+\infty$  donde a mayor valor de  $a$ , mayor será la discriminación del ítem. Baker (2001) proporciona el siguiente criterio para interpretar la correspondencia entre capacidad discriminativa y rango de valores de  $a$ : ninguna (0), muy baja (de 0.01 a 0.34), baja (de 0.35 a 0.64), moderada (de 0.65 a 1.34), alta (de 1.35 a 1.69) y muy alta (mayor a 1.70). Reise y Waller (1990) consideran aceptable para test de comportamiento típico el valor del parámetro  $a$  entre 0.5 y 1.5.

El parámetro  $c$  se define como la probabilidad para un individuo con un nivel muy bajo de  $\theta$  de responder correctamente o responder en forma favorable al rasgo a un ítem. Es el límite de la CCI cuando  $\theta$  tiende a  $-\infty$ , denominándose parámetro de pseudo-azar. No ha recibido una denominación específica para los tests de comportamiento típico (Abal, 2013). Sin embargo, en el contexto de este tipo de test algunos autores han relacionado al parámetro  $c$  con la deseabilidad (e.g. Rouse, Finger, & Butcher, 1999; Zumbo, Pope, Watson, & Hubley, 1997) lo cual no ha sido fehacientemente comprobado. Abal, Auné y Attorresi (2015) estudiaron la relación entre el parámetro  $c$  de la Escala Perseverancia y los resultados en la Escala de Distorsión del *Big Five Questionnaire* (Caprara, Borgogni, Barbaranelli, & Moreno, 1995, adaptación de Bermúdez, 1995) hallando que la respuesta influida por la deseabilidad social no repercute sobre dicho parámetro.

Figura 2

*Ubicación de los Parámetros a, b y c en la Curva Característica del Ítem*



La relación entre la puntuación verdadera  $V$  y el nivel en el rasgo  $\theta$  es una correlación curvilínea perfecta, siendo posible una transformación entre ellos (Lord, 1953). Lord (1980) señala que  $V$  y  $\theta$  son la misma cosa expresada en diferentes escalas de medida. La diferencia entre ellos es que la medición de  $V$  depende de los ítems del test mientras que la escala de  $\theta$  es independiente de los ítems del test. Se define la Curva Característica del Test (CCT) como la suma de las CCI de todos los ítems del test o escala. Se demuestra que es la curva de regresión del puntaje total del test en el rasgo (Lord, 1953). La CCT permite transformar la escala de  $\theta$ , que va de  $-\infty$  a  $+\infty$ , a una escala de puntuaciones verdaderas  $V$  que se encuentra en la escala de 0 a  $n$  (Muñiz Fernández, Fidalgo, García Cueto, Martínez, & Moreno, 2005). La fórmula que describe la relación entre los  $\theta$  y las  $V$  para un  $\theta$  dado es:

$$V = \sum_{i=1}^n P(\theta)$$

Donde:

$V$  es la puntuación verdadera que corresponde a examinados con un nivel dado en el rasgo latente.

$n$  es el número de ítems que componen el test.

$P(\theta)$  es el valor correspondiente a cada CCI para un  $\theta$  dado.

Calculando esta fórmula sobre un número elevado e igualmente espaciado de valores de  $\theta$  se obtiene la CCT.

Lord (1980) expone que  $g^*(\theta)$  es la distribución -densidad de probabilidad- del rasgo en el grupo testeado. Usualmente  $g^*(\theta)$  es desconocido. La distribución observada es de un  $\theta$  estimado y, por lo tanto, una aproximación a  $g^*(\theta)$ . Para poder estimar  $\theta$  a partir de un puntaje  $x$  observado en el test, se realiza una regresión. Un punto estimado de  $\theta$  para un  $x$  dado será obtenido mediante la regresión de  $\theta$  en  $x$ . Esta regresión no queda determinada a menos que conozcamos  $g^*(\theta)$ . La regresión de  $\theta$  en  $x$  es por definición dada por:

$$\mu_{\theta|x} = \frac{1}{\phi(x)} \int_{-\infty}^{\infty} \theta g^*(\theta) \phi(x|\theta) d\theta = \sum_{i=1}^n P(\theta)$$

Donde:

$\mu_{\theta|x}$  es la media poblacional de los puntajes de los examinados en el test dado un nivel específico de rasgo  $\theta$ . Es igual al puntaje verdadero  $V$ .

$\phi(x)$  es la distribución poblacional marginal de los puntajes suponiendo que los niveles en el rasgo  $\theta$  son conocidos.

$\theta$  es el nivel alcanzado en el rasgo.

$g^*(\theta)$  es la distribución poblacional -densidad de probabilidad- del rasgo  $\theta$  en el grupo testeado.

$\phi(x|\theta)$  es la distribución poblacional de frecuencias de los puntajes de los examinados en el test dado un nivel específico de rasgo  $\theta$ .

Esta fórmula puede ser calculada numéricamente sólo si la distribución de  $\theta$  es conocida.  $\mu_{\theta|x}$  es igual al puntaje verdadero  $V$  ya que:

$$\mu_{\theta|x} = \sum_{i=1}^n P(\theta) = V$$

### 2.3.3.1. Modelo Logístico de Un Parámetro o Modelo de Rasch

El modelo de Rasch se denomina de esa forma en honor a su creador, el matemático danés Georg Rasch (Rasch, 1960). El modelo asume que la probabilidad de una respuesta favorable al rasgo o correcta a un ítem es una función logística de la distancia con signo entre la ubicación del individuo en dicha escala y la ubicación del ítem en la escala del

rasgo (Pallant & Tennant, 2007). En otras palabras, la probabilidad de que una persona responda afirmativamente un ítem dado [ $P_i(\theta)$ ] es una función logística de la diferencia entre el nivel de la persona en, por ejemplo, conducta prosocial ( $\theta$ ) y el parámetro de localización ( $b$ ), es decir la posición del ítem en la escala del rasgo conducta prosocial. Se añade la constante  $D = 1.7$  para que los valores de la función tengan la mejor aproximación a la normal acumulada (Muñiz Fernández, 1997). En el ML1P la función logística tiene la expresión para un ítem dado:

$$P(\theta) = \frac{1}{1 + e^{-1.7(\theta-b)}}$$

Donde:

1.7 es el parámetro D especializado, dado que este valor da la mejor aproximación con un decimal de la función logística a la normal acumulada. Originalmente se utilizaba la distribución normal.

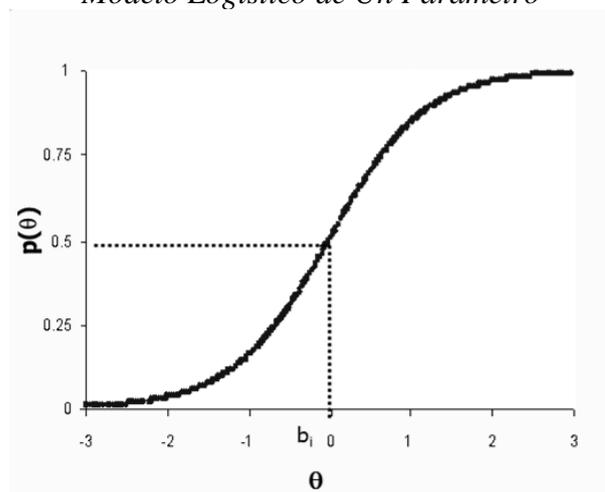
$P(\theta)$  es la probabilidad de respuesta favorable al rasgo o correcta para un nivel dado de  $\theta$ .

$e$  es la base de los logaritmos neperianos: 2.718...

$b$  es el valor  $\theta$  necesario para tener probabilidad 0.5 de contestar favorablemente al rasgo o correctamente. Se deduce que es la primera coordenada del punto de inflexión.

Figura 3

*Modelo Logístico de Un Parámetro*



Modelo Logístico de Un Parámetro Considerando  $D = 1.7$ ,  $b = 0$

El ajuste de los datos al modelo de Rasch sitúa tanto las estimaciones de los parámetros del ítem como de la persona -parámetros independientes- en la misma escala de unidades logarítmicas y es esto lo que permite la transformación lineal de la puntuación bruta (Pallant & Tennant, 2007).

El modelo teórico tiene como parámetro  $a$  el valor 1 y como parámetro  $c$  el valor 0. Por lo tanto, habrá una buena aplicación a una prueba cuando los parámetros  $a$  sean relativamente cercanos a 1 y los  $c$  cercanos a 0.

En el marco de la medición en el modelo de Rasch, la escala debería funcionar de la misma manera independientemente de qué grupo -según el género, nacionalidad u otras variables- está siendo evaluado (Holland & Wainer, 1993). Por ejemplo, en el caso de medir conducta prosocial, varones y mujeres deberían tener la misma probabilidad de responder afirmativamente un ítem para el mismo nivel de conducta prosocial. Es decir, la probabilidad de responder al ítem debe estar condicionada sólo por el rasgo latente. Si por alguna razón un género no tiene la misma probabilidad de responder afirmativamente al ítem, entonces el ítem será considerado con Funcionamiento Diferencial del Ítem (*Differential Item Functioning*, DIF) y se viola el requerimiento de unidimensionalidad (Smith, 2000).

### 2.3.3.2. Modelo Logístico de Dos Parámetros

Este modelo fue propuesto por Birnbaum (1968). En este caso la probabilidad de responder en forma favorable al rasgo o correcta al ítem es función de los parámetros  $b$  y  $a$ , asumiendo que el parámetro  $c$  tiene un valor nulo. En el ML2P la función logística tiene siguiente expresión para un ítem dado:

$$P(\theta) = \frac{1}{1 + e^{-1.7a(\theta-b)}}$$

Donde:

1.7 es el parámetro D especializado, dado que este valor da la mejor aproximación con un decimal de la función logística a la normal acumulada. Originalmente se utilizaba la distribución normal.

$P(\theta)$  es la probabilidad de respuesta favorable al rasgo o correcta para un nivel dado de  $\theta$ .

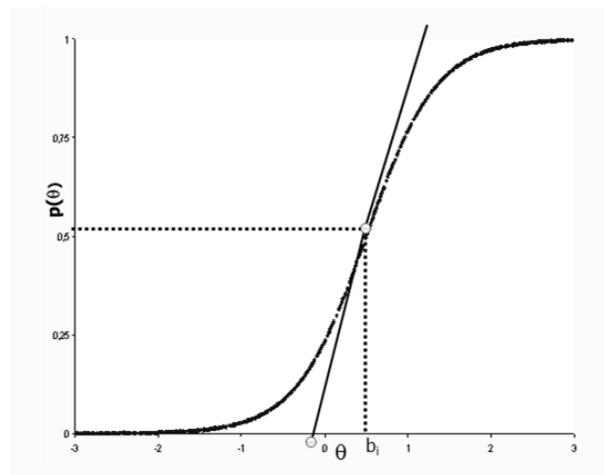
$e$  es la base de los logaritmos neperianos: 2.718...

$a$  es proporcional a la pendiente de la recta tangente en el punto de inflexión de la CCI.

$b$  es el valor  $\theta$  necesario para tener probabilidad 0.5 de contestar favorablemente al rasgo o correctamente. Se deduce que es la primera coordenada del punto de inflexión.

Figura 4

*Modelo Logístico de Dos Parámetros*



Modelo Logístico de Dos Parámetros Considerando  $D = 1.7$ ,  $a = 1.4$ ,  $b = 0.5$

2.3.3.3. Modelo Logístico de Tres Parámetros

Esta transformación lineal del ML2P incluye los parámetros  $a$ ,  $b$  y  $c$ . El parámetro  $c$  será el límite de la CCI cuando  $\theta$  tiende a  $-\infty$ . Su inclusión altera el valor de  $P(b)$  en el ML3P siendo ahora  $0.5 + c/2$ . En el ML3P la función logística tiene la expresión:

$$P(\theta) = c + (1 - c) \frac{1}{1 + e^{-1.7a(\theta-b)}}$$

Donde:

1.7 es el parámetro  $D$  especializado, dado que este valor da la mejor aproximación con un decimal de la función logística a la normal acumulada. Originalmente se utilizaba la distribución normal.

$P(\theta)$  es la probabilidad de respuesta favorable al rasgo o correcta para un nivel dado de  $\theta$ .

$e$  es la base de los logaritmos neperianos: 2.718...

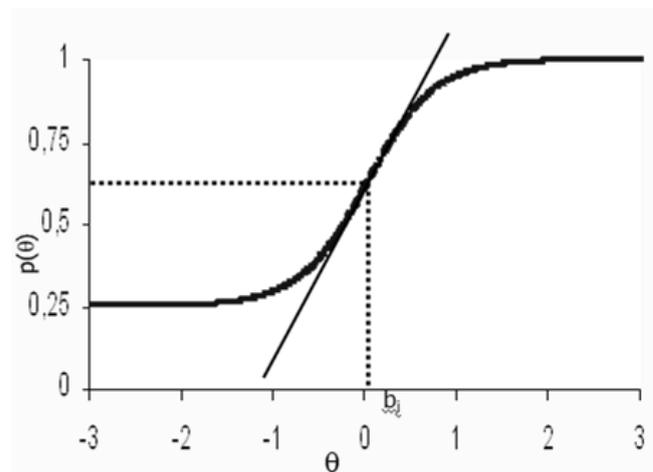
$a$  es proporcional a la pendiente de la recta tangente en el punto de inflexión de la CCI, si  $c$  es constante.

$b$  es el valor  $\theta$  en el punto de inflexión. Ahora el valor de su imagen es  $0.5 + c/2$ .

$c$  es la probabilidad de respuesta favorable al rasgo o correcta cuando el nivel de  $\theta$  tiende a  $-\infty$ .

Figura 5

*Modelo Logístico de Tres Parámetros*



Modelo Logístico de Tres Parámetros Considerando  
 $D = 1.7, a = 1.2, b = 0, c = 0.25$

Este valor de  $c$  corresponde a un ejemplo de un ítem de cuatro opciones y una sola correcta.

Este modelo fue diseñado específicamente para test cognitivos con formato de respuesta múltiple *-multiple choice-* (McDonald, 1999).

2.3.3.4. Comparación entre los Modelos Logísticos de Uno, Dos y Tres Parámetros

El ML1P es el único de los modelos dicotómicos en tener las siguientes propiedades (McDonald, 1999): a) el puntaje total es un estimador suficiente del punto de máxima verosimilitud, y b) las comparaciones entre subpoblaciones son específicamente objetivas, independientes del ítem o ítems usados para la comparación. Un estimador suficiente es aquel donde no se requiere información adicional de los datos para realizar la estimación. Esto es, los ítems específicos que fueron respondidos en forma favorable al rasgo o correcta no proveen información adicional acerca del nivel en el rasgo

en el modelo de Rasch (Embreston & Reise, 2000). Cada puntaje total de la escala tiene asociado un solo nivel en el rasgo. La combinación de estas propiedades ha llevado a ciertos teóricos de la psicometría a preferir este modelo sobre cualquier otro, al punto de recomendar únicamente seleccionar ítems que se ajusten al ML1P (McDonald, 1999). Además, generalmente es estimable (Montero, 2000) y requiere de tamaños de muestra menores que el ML2P y el ML3P. En ML2P y ML3P la configuración específica y la consistencia de los datos son cruciales para poder estimar los modelos (Montero, 2000).

Sin embargo, existen críticos del ML1P. Montero (2000) afirma que los supuestos necesarios para aplicarlo son restrictivos y difíciles de cumplir en situaciones reales. Especialmente resulta difícil asumir que todos los ítems tienen igual poder discriminatorio. Esta autora prefiere al ML2P por ser más poderoso que el de Rasch y resultar estimable en más ocasiones que el ML3P.

Una ventaja de ML2P y ML3P es que las funciones de información del test dan al desarrollador de test un enorme grado de flexibilidad para diseñar un instrumento con una distribución específica del error de medición a lo largo de la dimensión (McDonald, 1999). En ML2P y ML3P pueden existir diferentes estimaciones de  $\theta$  entre diferentes sujetos, aunque su puntuación en el test sea la misma. Esto se debe a que puede existir distintos patrones de respuesta para llegar al mismo valor total y tanto ML2P como ML3P le darán más ponderación en la estimación de  $\theta$  a los ítems con mayor poder de discriminación (Montero, 2000).

Muñiz Fernández et al. (2005) proporcionan un criterio práctico para elegir cuál de estos modelos aplicar: si las correlaciones ítem-test son relativamente homogéneas se puede ajustar el ML1P, si son muy disímiles aconsejan utilizar el ML2P. Si además hay indicios de existencia de aciertos por azar se selecciona el ML3P.

#### 2.3.4. Modelos de la TRI para Ítems Politómicos

Los modelos de la TRI para ítems con respuesta politómica son modelos matemáticos utilizados para entender la interacción entre el nivel de rasgo de los individuos y los ítems, donde los ítems tienen varias categorías de respuesta. Los ítems se puntúan de manera de reflejar la categoría particular que el individuo ha alcanzado, en la que ha sido clasificado, o que ha seleccionado. Los ítems con respuesta politómica proveen una información más rica acerca del constructo medido (Abal, 2013).

Se observa que estos modelos son implementados más frecuentemente en test de comportamiento típico, principalmente en la verificación del ajuste de test diseñados a partir de la TCT, la determinación del número óptimo de alternativas de respuesta (e.g. Abal, Auné, Lozzia, & Attoresi, 2017), el diseño de Test Adaptativos Informatizados (e.g., Lozzia, 2012) y la prueba de nuevos modelos politómicos no paramétricos (Asún & Zuñiga, 2008).

Hasta mediados de la década de los ochenta los desarrolladores de test, psicólogos e investigadores con datos de respuesta politómica tendían a usar TCT. Otros investigadores dicotomizaban sus datos. La falta de programas específicos limitaba el uso de datos de respuesta politómica (Hambleton, van der Linden, & Wells, 2010).

El desarrollo de modelos de la TRI específicamente diseñados para ítems politómicos resultó complejo, involucrando una variedad de perspectivas (Abal, 2013; Abal et al., 2010). Es necesario destacar que los modelos pioneros fueron introducidos mucho antes de poderse encontrar su aplicación en psicología y educación (Hambleton et al. 2010). En 1969 fue introducido por Samejima el primero de estos modelos. Al día de hoy, los modelos politómicos han proliferado, así como se han desarrollado los programas para poder aplicarlos (Martínez Arias, Hernández Lloreda, & Hernández Lloreda, 2006).

Muchos de los modelos politómicos son básicamente una extensión de ciertos modelos dicotómicos. El procedimiento se basa en fraccionar la respuesta múltiple en distintas variables dicotómicas. La respuesta múltiple puede ser nominal u ordinal, en este último caso debe respetarse la jerarquía de las categorías. A diferencia de los modelos dicotómicos, donde a cada ítem le corresponde una CCI, en los modelos politómicos cada opción de respuesta tiene una Curva Característica de las Categorías de Respuesta del Ítem (CCCRI, Samejima, 1969). Estas se distinguen en los modelos politómicos de las Curvas Características Operantes (CCO) (Embreston & Reise, 2000) o Curvas Características de los Pasos del Ítem (CCPI) (Rojas & Pérez, 2001) que indican la probabilidad de realizar el paso de una categoría a otra. La cantidad de pasos siempre es inferior en una unidad a la cantidad de categorías.

De los diversos modelos de la TRI para ítems politómicos, los Modelos de Respuesta Graduada (MRG) y el Modelo de Crédito Parcial (MCP) son que han tenido más relevancia (Abal, 2013).

### 2.3.4.1. Modelos de Respuesta Graduada

Los Modelos de Respuesta Graduada desarrollados por Fumiko Samejima (1969) son una extensión del ML2P para ítems de respuesta dicotómica al caso politómico. Son apropiados cuando las respuestas a un ítem pueden ser clasificadas en más de dos categorías ordenadas para representar distintos grados de logro en la solución a un problema o niveles de acuerdo con un determinado enunciado. Utilizan la ojiva normal o la función logística. Las respuestas a un ítem dado son clasificadas en  $M + 1$  categorías ordenadas, siendo  $M$  el número total de categorías originales del ítem, de tal manera que aquellas con los menores números representen menos del rasgo medido por el ítem que las categorías más altas. Representaremos como  $m$  a una categoría genérica particular del ítem. Los puntajes para el ítem son enteros sucesivos. A pesar de ello, se supone que la respuesta discreta  $s$  de un individuo a un ítem es parte de un continuo de respuesta  $S$ .

Para cada ítem son estimados un parámetro de inclinación  $a$  y un conjunto de parámetros de umbral de categoría  $b_m$ . Los  $b_m$  son los puntos del rasgo latente donde la probabilidad de elegir la categoría  $m$  o superior es 0.5. El parámetro de inclinación  $a$  es constante para cada ítem e informa el grado en que sus categorías de respuesta distinguen entre los niveles del rasgo. El parámetro  $a$  en el MRG es denominado de distintas maneras. Algunos autores conservan la denominación *parámetro de discriminación* del ML2P (e.g. Ecurra Mayaute, & Salas Blas, 2014) mientras que otros lo denominan *parámetro de la pendiente* (e.g. Hernández Baeza, Muñiz Fernández, & García Cueto, 2000) o *de inclinación* (e.g. Hidalgo, López-Pina, Inglés, & Méndez, 2002) debido a que la discriminación del ítem también depende de la diferencia en los umbrales de cada categoría (Cnossen, 2015).

Samejima desarrolló un proceso de dos pasos para obtener la probabilidad de que un individuo elija una determinada categoría en un ítem dado. En el primer paso, la probabilidad de que un individuo con un determinado nivel de rasgo elija una categoría  $m$  o mayor en un ítem dado es expresada por la siguiente fórmula:

$$P_m^*(\theta) = \frac{1}{1 + e^{-1.7a(\theta - b_m)}}$$

Donde:

1.7 es el parámetro D especializado, dado que este valor da la mejor aproximación con un decimal de la función logística a la normal acumulada.

$a$  es el parámetro de inclinación *-slope-*, pendiente o discriminación del ítem.

$b_m$  es el punto del rasgo latente donde la probabilidad de elegir la categoría  $m$  o superior es 0.5.

$e$  es la base de los logaritmos neperianos: 2.718...

$\theta$  es el nivel en el rasgo latente.

$P_m^*(\theta) = 1$  si  $m = 1$  ya que la probabilidad acumulada de responder en la categoría más baja o en todas las mayores que esta es un suceso seguro o cierto.  $P_m^*(\theta) = 0$  si  $m = M + 1$  ya que la probabilidad de responder en una categoría siguiente a la mayor es nula.

Obtener la probabilidad de que un individuo elija una determinada categoría implica abstraer las CCCRI adyacentes. Samejima definió la probabilidad de que un individuo responda en una categoría dada como:

$$P_m(\theta) = P_m^*(\theta) - P_{(m+1)}^*(\theta)$$

La fórmula representa la CCO para el MRG. Casos particulares de esta fórmula son:

$$P_1(\theta) = P_1^*(\theta) - P_2^*(\theta) = 1 - P_2^*(\theta)$$

$$P_m(\theta) = P_m^*(\theta) - P_{(m+1)}^*(\theta) \text{ para } m \text{ tal que } 2 \leq m \leq M-1$$

$$P_M(\theta) = P_M^*(\theta) - P_{M+1}^*(\theta) = P_M^*(\theta) - 0 = P_M^*(\theta)$$

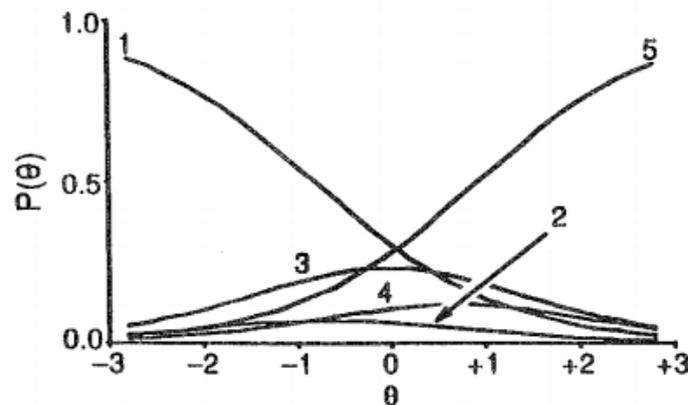
El uso de esta fórmula reduce la respuesta politómica a dicotómica ya que las respuestas graduadas son clasificadas en dos categorías: valores menores que  $m$  y valores iguales o mayores que  $m$ . Así, cada ítem es tratado como una serie de respuestas dicotómicas. En cada una de estas dicotomías se aplica el ML2P considerando, como ya se mencionó,  $a$  constante. Dada la segmentación acumulativa que se realiza, se incluye al MRG en los modelos de probabilidad acumulada (Hemker, van der Ark, & Sijtsma,

2001). También se lo considera un modelo de diferencias, ya que la probabilidad de obtener una categoría determinada se obtiene como diferencia entre dos funciones (Thissen & Steinberg, 1986). Además, es un modelo aditivo, lo que indica que la probabilidad de obtener una categoría que sea unión de otras dos es igual a la suma de las probabilidades de las categorías originales (Ostini & Nering, 2010). La propiedad de aditividad es útil porque permite unir categorías cuando sea necesario, por ejemplo, si no se cuenta con el número mínimo de participantes para estimar los parámetros de cada categoría (Revuelta Menéndez, Abad, & Ponsoda Gil, 2006).

Es importante destacar que, en el MRG, en contraste con el ML2P, las CCCRI no son necesariamente monótonamente crecientes. Además, no tienen la misma forma (Cohen, Kim & Baker, 1993). La CCCRI para la categoría más baja decrece a medida que el nivel de rasgo aumenta, mientras que por el contrario la CCCRI para la categoría más alta crece a medida que el nivel de rasgo aumenta, como se puede observar en la Figura 6.

Figura 6

*CCCRI para un ítem de cinco categorías*



Fuente: Cohen, Kim & Baker (1993)

#### 2.3.4.2. Modelo de Crédito Parcial

El Modelo de Crédito Parcial fue desarrollado por Masters (1982, 2016). Es un modelo de Rasch adecuado para ser utilizado en datos ordinales y cuenta con las propiedades favorables del ML1P. La lógica original del modelo se basa en un proceso de respuesta donde el individuo responde secuencialmente a una serie de subproblemas

en el ítem. El crédito parcial dado iguala la cantidad de pasos completados de forma exitosa. Es un modelo donde la obtención de probabilidad es realizada de manera directa o modelo *divide-by-total* (Thissen & Steinberg, 1986). El Modelo de Rash se aplica a cada uno de los pares de categorías adyacentes de los ítems. Se obtienen así las Curvas Características de cada categoría. Si un ítem dado se responde en  $m$  categorías de 0 a  $m$ , el modelo general para las categorías  $1 \leq h \leq m$  se define como:

$$\frac{P(h|\theta)}{P(h-1|\theta)+P(h|\theta)} = \frac{e^{1.7a(\theta-\delta_h)}}{1+e^{1.7a(\theta-h)}}$$

Donde:

1.7 es el parámetro D especializado, dado que este valor da la mejor aproximación con un decimal de la función logística a la normal acumulada.

$P(h|\theta)$  es la probabilidad de que un sujeto con un nivel dado de  $\theta$  opte por la categoría  $h$ .

$\delta_h$  es el parámetro de umbral que describe la cantidad de  $\theta$  que requiere el ítem para seleccionar la categoría  $h$  en lugar de la categoría  $h - 1$ .

$$P(h|\theta)[1 + e^{1.7a(\theta-h)}] = [P(h - 1|\theta) + P(h|\theta)]e^{1.7a(\theta-\delta_h)}$$

Si se aplica la propiedad distributiva resulta:

$$P(h|\theta) + P(h|\theta)e^{1.7a(\theta-h)} = P(h - 1|\theta)e^{1.7a(\theta-h)} + P(h|\theta)e^{1.7a(\theta-\delta_h)}$$

Si se anulan los  $P(h|\theta)e^{1.7a(\theta-h)}$ , presentes a ambos lados de la igualdad, se obtiene:

$$P(h|\theta) = P(h - 1|\theta)e^{1.7a(\theta-h)}$$

Esta expresión es válida para  $1 \leq h \leq m$

La expresión general del modelo, válida para  $h = 0, \dots, m$  es:

$$P(h|\theta) = \frac{e^{1.7a \sum_{k=1}^h (\theta - \delta_k)}}{\sum_{j=0}^m e^{1.7a \sum_{h=0}^j (\theta - h)}}$$

Con  $1.7a \sum_{h=0}^0 (\theta - \delta_j)$  definido como 0 y donde  $\delta_j$  son las intersecciones entre categorías y  $k$  denota los pasos necesarios para que el ítem se acompletado ( $k = 1, 2, \dots, m$ ) (Masters, 1982). En este modelo, la probabilidad es calculada como el exponente de la suma de todos los términos  $(\theta - \delta_j)$  de cada categoría hasta  $h$ , dividido por la suma de los términos del numerador para todas las categorías posibles.

Los parámetros  $\delta_j$  indican la dificultad del paso  $j$ , es decir cantidad de  $\theta$  que demanda el ítem para que el evaluado tenga más probabilidad de seleccionar la categoría  $j$  en lugar de la categoría  $j - 1$ . Esto puede ser observado cuando el MCP se muestra gráficamente, mostrando las probabilidades de responder en cada una de las categorías en función de  $\theta$ , es decir las CCCRI.

Para ilustrar el MCP, se utilizará un ejemplo de Masters (1982). Supóngase que un ítem dado posee cuatro categorías de respuesta (*Fuertemente en Desacuerdo*, *Desacuerdo*, *Acuerdo*, *Fuertemente de Acuerdo*). Si una persona elije *Acuerdo* puede considerarse que ha elegido *Desacuerdo* sobre *Fuertemente en Desacuerdo* (primer paso realizado) y *Acuerdo* por sobre *Desacuerdo* (segundo paso realizado). Consecuentemente, múltiples pasos se pueden completar en el MCP, lo cual contrasta con el ML1P donde sólo se puede realizar un paso, desde una respuesta favorable al rasgo o correcta a una desfavorable al rasgo o incorrecta.

Denominemos a las categorías como 0, 1, 2 y 3 siendo 0 *Fuertemente en Desacuerdo* y 3 *Fuertemente de Acuerdo*. La probabilidad de que una persona realice el tercer paso en un ítem dado para puntuar 3 en lugar de 2 (si ya ha alcanzado el segundo paso) se escribe como:

$$\frac{P_3(\theta)}{P_2(\theta) + P_3(\theta)} = \frac{e^{1.7a(\theta - \delta_3)}}{1 + e^{1.7a(\theta - \delta_3)}}$$

Donde  $\delta_3$  es el parámetro de umbral que describe la cantidad de  $\theta$  que requiere el ítem para seleccionar la categoría 3 en lugar de la categoría 2.

De forma similar la probabilidad de que una persona realice el segundo paso en el ítem para así puntuar 2 y no 1 -si ya alcanzó el primer paso- puede ser calculada de la siguiente forma:

$$\frac{P_2(\theta)}{P_1(\theta) + P_2(\theta)} = \frac{e^{1.7a(\theta-\delta_2)}}{1 + e^{1.7a(\theta-\delta_2)}}$$

La obtención de la probabilidad de realizar el primer paso para puntuar 1 y no 0 es idéntica al modelo de Rasch excepto porque el parámetro  $b$  es reemplazado por el parámetro  $\delta_1$  del primer paso. Es claro que el MCP se asienta en el *ratio* de las categorías adyacentes y consecuentemente es un modelo de segmentación adyacente del dato politómico. Combinando las ecuaciones en una expresión general para la probabilidad de que una persona obtenga el puntaje  $h$  en un ítem dado resulta la expresión general del MCP.

#### 2.3.4.3. Comparación entre el Modelo de Respuesta Graduada y Modelo de Crédito Parcial

El Modelo de Respuesta Graduada (MRG) de Samejima (1969) y el Modelo de Crédito Parcial (MCP) de Masters (1982, 2016) construidos para ítems politómicos presentan semejanzas y diferencias tanto teóricas como prácticas que justifican su elección para analizar los ítems de la escala de conducta prosocial. Ambos modelos son paramétricos y son utilizados con frecuencia en la modelización de pruebas de actitudes y atributos de la personalidad cuyos ítems se puntúan ordinalmente. Otra virtud es que cuentan con programas estadísticos seguros para llevar a cabo su implementación: calibración de los ítems y obtención de medidas locales de precisión (Revuelta Menéndez et al., 2006). Pero más allá de estas similitudes, el MRG y el MCP provienen de marcos teóricos independientes, aplican distintas formas de segmentación del ítem politómico y utilizan diferentes procedimientos para el cálculo de las probabilidades de respuesta de cada categoría.

La bibliografía psicométrica muestra que el MRG parece relativamente más adecuado para modelizar escalas de comportamiento típico. El MRG fue utilizado más frecuentemente con escalas de respuesta tipo Likert que el MCP. El MCP ha sido implementado, en general, para considerar puntuaciones parciales o la cantidad de etapas sorteadas en la resolución de una tarea (Aguado, Rubio, Hontangas, & Hernández, 2005). La mayor calidad del MRG es avalada por la experiencia obtenida en investigaciones que compararon ambos modelos con diferentes constructos (Abal, 2013; Asún & Zuñiga,

2008; Baker, Rounds, & Zevon, 2000). Además, el supuesto de homogeneidad de los parámetros de discriminación del MCP puede resultar de difícil cumplimiento en ítems de escalas de comportamiento típico (Boomsma, Van Duijn, & Snidjers, 2001). Por último, el MRG presenta una propiedad que no se cumple para el MCP consistente en la invarianza de los parámetros tras la unión de dos categorías adyacentes (Andrich, 1995). Esta propiedad puede ser útil cuando no existe una cantidad suficiente de individuos como para estimar los parámetros de una categoría.

Sin embargo, es de destacar a favor del MCP que este, al ser una generalización del Modelo de Rasch, cuenta con estadísticos suficientes. En cambio, en el MRG, la puntuación en el test no es un estadístico suficiente para la obtención de los parámetros (Embretson & Reise, 2000). Además, también por pertenecer el MCP a la familia de los modelos de Rasch, los parámetros de los ítems son separables de los parámetros de las personas. Es decir, las estimaciones de los ítems no dependen de las estimaciones de las personas, y viceversa.

De todas maneras, se recomienda aplicar diferentes modelos a los datos y evaluar las estimaciones de parámetros por su viabilidad y significación a lo largo de los modelos comparados (Whittaker, & Worthington, 2016) y eso es lo que se hizo en este caso.

### 2.3.5. Estimación de Parámetros

En la elección del modelo de la TRI a aplicar, se debe considerar tanto el formato de respuesta de los ítems, el tipo de escala -si mide un comportamiento típico o la ejecución máxima-, los índices de ajuste del modelo, como así también las consideraciones que se han expuesto en los apartados de comparación entre ML1P, ML2P y ML3P y de comparación entre MRG y MCP.

Posteriormente a la elección del modelo y a la administración de la prueba, se realiza el proceso de calibración. En este, se estiman los parámetros de la CCI o CCCRI de cada ítem y los  $\theta$  por individuo a partir de las respuestas observadas de los participantes. Para ello, habitualmente se utiliza el método de Máxima Verosimilitud (MV), que consiste en estimar los  $\theta$ , es decir determinar los estimadores de  $\theta$  dado un  $\theta$  con los que el conjunto de respuestas  $u_1, u_2, \dots, u_n$  de los sujetos alcanza la máxima probabilidad de la función de verosimilitud. Los estimadores de  $\theta$  así hallados se denominan estimadores de Máxima Verosimilitud de  $\theta$ .

$$L(u_1, u_2, \dots, u_n / \theta) = \prod_{i=1}^n P_i^{u_i}(\theta)(1 - P_i(\theta))^{1-u_i}$$

En la práctica resulta más amigable maximizar el logaritmo de la función de verosimilitud. Como la función logaritmo natural  $\ln$  -es decir el logaritmo de base  $e$ - es una transformación monótona creciente, maximizar  $\ln$  es equivalente a maximizar la función de verosimilitud.

$$\ln L(u_1, u_2, \dots, u_n / \theta) = \sum_{i=1}^n u_i \ln P_i + (1 - u_i) \ln(1 - P_i)$$

Por ejemplo, para estimar el nivel de rasgo latente de un sujeto, se consideran sus respuestas y se determina el nivel de rasgo para el cual es más probable que se presenten dichas respuestas. De manera similar se procede con los parámetros de cada uno de los ítems. Así se pueden hallar los parámetros que hacen más probable que a ese ítem le correspondan las configuraciones de respuestas que, de hecho, se han observado para él.

Para encontrar el máximo, se deriva el  $\ln$  de la función de verosimilitud y se iguala a 0. A esta expresión se la denomina ecuación de verosimilitud:

$$\frac{d}{d\theta} \ln L(\mathbf{u} / \theta) = 0$$

Donde  $\mathbf{u} = (u_1, u_2, \dots, u_n)$

Esta ecuación no tiene en general solución explícita sino aproximada por métodos numéricos. Se utilizan métodos iterativos. Esto es, algoritmos que en cada paso – iteración- van dando una solución provisoria que se aproxima mejor a la dada en el paso anterior. Cuando la diferencia entre dos soluciones consecutivas es lo *suficientemente pequeña* el algoritmo culmina y se dice que *converge*. Por *suficientemente pequeña* se entiende una distancia inferior a la establecida por el criterio de convergencia previamente establecido, por ejemplo 0.01.

Dicho estimador de Máxima Verosimilitud bajo ciertas condiciones tiene las siguientes propiedades:

- a) Es consistente. Se aproxima al verdadero valor del parámetro  $\theta$  conforme aumenta el tamaño muestral y el número de ítems, hasta coincidir en el infinito.

- b) Función de estadísticos suficientes cuando éstos existen. Un estadístico es suficiente cuando contiene toda la información acerca del parámetro.
- c) Es eficiente. Tiene asintóticamente menor varianza.
- d) Se distribuye asintóticamente normal con media  $\theta$  (insesgado) y varianza:

$$Var(\hat{\theta} / \theta) = \frac{1}{E \left[ \left( \frac{d \ln L}{d \theta} \right)^2 / \theta \right]}$$

Uno de los métodos que ha resultado de mucha utilidad especialmente cuando hay que estimar los parámetros de los ítems y de los sujetos con la misma muestra es el método de Máxima Verosimilitud Marginal (MVM) introducido por Bock y Liberman (1970) y mejorado por Bock y Aitkin (1981) a través de la implementación del algoritmo Esperanza y Maximización. El MVM modela la probabilidad de hallar una secuencia de respuesta específica en una población, que puede ser especificada para cada patrón de respuesta singular. MULTILOG es un programa específico (Thissen, Chen, & Bock, 2003) que utiliza MVM.

### 2.3.6. Evaluación del Ajuste de los Datos a los Modelos

Una vez que se han estimado los parámetros del modelo elegido, se realiza el estudio de la bondad de ajuste de dicho modelo. Si existe un ajuste adecuado se pueden realizar inferencias y predicciones acertadas. Si la función base del modelo no demuestra ajuste, se corre el riesgo de tomar decisiones inadecuadas a partir del mismo. El nivel en el rasgo y los parámetros estimados del ítem pueden, en ese caso, estar sesgados y así ponerse en peligro la aplicación adecuada de los modelos de la TRI (Kang & Chen, 2008). La consideración del ajuste del modelo es un paso sumamente importante en el desarrollo y adaptación de test. Los ítems que no alcanzan niveles satisfactorios de ajuste son modificados y sino, descartados.

Para evaluar el ajuste existen distintos procedimientos tanto gráficos como estadísticos (Abad, Olea, Ponsoda, & García, 2011). Actualmente hay controversias con respecto a algunos de estos procedimientos, por ejemplo, ha sido cuestionado el supuesto de que los estadísticos de bondad de ajuste sigan la distribución  $\chi^2$  y la objetividad de los

métodos gráficos (Sinharay, Haberman, & Jia, 2011). Las limitaciones de los métodos de evaluación del ajuste pueden conducir a evaluar como desajustados ítems que en realidad presentan buen ajuste debido a elevadas tasas de error de Tipo I de los estadísticos, es decir falsos positivos, se rechaza la hipótesis nula siendo esta verdadera (Demars, 2005; Sinharay et al., 2011). Es por esto que los métodos bayesianos de evaluación del ajuste están cobrando progresiva popularidad frente a los métodos tradicionales (Ames & Pendfield, 2015).

#### 2.4. *Funcionamiento Diferencial del Ítem*

Mallinckrodt, Miles y Recabarren (2016) sugirieron el estudio del Funcionamiento Diferencial del Ítem (*Differential Item Functioning*, DIF) para ayudar a identificar reactivos que se comportan de forma diferente entre grupos. La detección del DIF es un paso importante en el proceso de desarrollo de escalas para asegurar que los ítems midan el constructo de interés de forma similar entre subpoblaciones. Si la probabilidad de contestar favorable o correctamente aun ítem para un nivel dado de rasgo depende de alguna otra característica que el rasgo en cuestión, dicha probabilidad podrá variar entre las poblaciones que difieran en tal característica, con lo que el ítem tendrá un funcionamiento diferencial. Es decir, se afirma que un ítem presenta funcionamiento diferencial del ítem cuando sujetos pertenecientes a distintos grupos, pero de un mismo nivel de rasgo a medir -sujetos comparables- tienen probabilidad disímil de dar la misma respuesta a un ítem (Holland, 1985; Holland & Thayer, 1988). La presencia de DIF atenta seriamente contra la validez e implica la falta de unidimensionalidad.

La literatura distingue las diferencias reales en el rasgo medido entre grupos, que se denominan impacto, de aquéllas producidas artificialmente y que favorecen o perjudican a uno de los grupos (Camilli & Shepard, 1994). Sólo en estos casos se considera que se presenta DIF. Así, uno de los problemas centrales de la Psicometría es el estudio del DIF y la obtención de métodos que permitan su detección (Hambleton & Swaminathan, 1985; Holland & Wainer, 1993). Por su importancia, el análisis del DIF forma parte de los procesos de construcción y validación de una escala. Cuando se detecta DIF es de interés hallar una razón que lo explique, en este caso se afirma que el ítem está sesgado.

Cuando se analizan dos grupos éstos se denominan como Grupo de Referencia ( $G_R$ ) y Grupo Focal ( $G_F$ ) (Zumbo, 2007). Generalmente integran el  $G_R$  los individuos que pueden resultar favorecidos por el test mientras que el  $G_F$  se compone de individuos pertenecientes a un grupo social en desventaja que puede resultar desfavorecido.

El DIF suele clasificarse en Uniforme y no Uniforme, según si para todos los niveles del puntaje total, el cociente de las posibilidades (*Odds Ratio*) es constante y diferente de 1 o no lo es. Es decir que si existe DIF Uniforme hay una diferencia invariante en el desempeño en el ítem que favorece a uno de los grupos a lo largo de los niveles del rasgo. Si la diferencia en el desempeño en el ítem varía a lo largo de los niveles del rasgo se trata de DIF no Uniforme (Aguerri, 2012).

Distintos estudios han caracterizado al DIF uniforme por la posición relativa de las dos CCIs en el sistema de coordenadas (e.g. Osterlind & Everson, 2009). Clauser y Mazor (1998) señalaron que el DIF Uniforme se observa mediante los modelos de la TRI cuando las CCIs difieren sólo en el parámetro de localización o dificultad  $b$ . Sin embargo, existe una definición más exhaustiva, propuesta por Hanson (1998), en la cual se distingue entre DIF Paralelo, DIF uniforme y DIF unidireccional (las CCIs no se cortan). Mostró que: (1) todo DIF paralelo es unidireccional, (2) todo DIF uniforme es unidireccional, (3) no todo DIF paralelo es uniforme, y (4) no todo DIF uniforme es paralelo (Aguerri, Prieto-Marañón, Galibert, & Attorresi, 2011).

Existe DIF Paralelo si las CCIs son paralelas (Hanson, 1998), es decir si existe  $k$ ,  $k \neq 0$ , tal que  $P_1(\theta) = P_2(\theta + k)$ . Para definir al DIF Uniforme, Hanson (1998) utilizó la función de los *Odds Ratio* (*Odds Ratio Function*, *ORF*) siendo su expresión:

$$f(\theta) = \frac{P_1(\theta)[1 - P_2(\theta)]}{P_2(\theta)[1 - P_1(\theta)]}$$

Hanson (1998) afirmó que un ítem presenta DIF Uniforme si la ORF es una función constante distinta de la unidad, es decir si existe  $\alpha \in \mathfrak{R}^+$ ,  $\alpha \neq 1$ , tal que  $f(\theta) = \alpha \forall \theta \in \mathfrak{R}$ .

Camilli y Shepard (1994) presentan diversos procedimientos, medidas y test de hipótesis, para el estudio del DIF que pueden clasificarse entre los que aplican la TRI y los llamados de Tablas de Contingencia.

#### 2.4.1. Métodos que Aplican la TRI

Genéricamente, cuando se estudia el DIF en el marco de la TRI, se considera que un ítem tiene funcionamiento diferencial cuando las CCI no coinciden entre los grupos (Aguerri, 2012). Los métodos de análisis del DIF basados en la aplicación de modelos de la TRI consideran como sujetos comparables a aquellos que tienen igual nivel en el rasgo, requieren tamaños grandes de muestra para poder estimar los parámetros del ítem y el valor de rasgo de cada individuo. A su vez, necesitan para su aplicación programas estadísticos específicos (MULTILOG, LOGIST, BILOG, BILOG-MG) dado el trabajo computacional intensivo que demandan.

Desde la perspectiva de los parámetros de los ítems y si el análisis se presenta en el marco del ML1P o del ML2P, se dice que un ítem presenta DIF uniforme cuando las CCIs difieren sólo en el parámetro de localización  $b$ . En el caso del ML1P, todo DIF existente es uniforme. En cambio, para el ML2P el DIF es no Uniforme cuando las CCIs difieren al menos en el parámetro de discriminación  $a$ . Se clasifica en DIF no Uniforme Simétrico si difieren únicamente en el parámetro  $a$  y DIF no Uniforme Asimétrico o Mixto si difieren en los parámetros  $a$  y  $b$ .

Hay procedimientos estadísticos para la detección del DIF basados en la TRI que comparan los parámetros de los ítems. Es el caso de la prueba normal para la diferencia de los  $b$  (e.g. Schulz, 1990) y también de la prueba  $\chi^2$  de Lord (1977, 1980). También se pueden comparar las CCI midiendo el área abarcada entre las CCI correspondientes a cada grupo. Se destacan los métodos de comparación de áreas de Raju (1988,1990). Por último, también se puede comparar el ajuste de los modelos. Es el caso de la prueba de la razón de verosimilitud (Thissen, Steinberg, & Wainer, 1993).

En el ML3P, el DIF Uniforme no implica DIF Paralelo (Aguerri, 2012). Necesariamente, se corresponde con CCIs con diferente parámetro de pseudoazar  $c$  e igual parámetro de discriminación  $a$ . Si hay DIF Paralelo, no existe DIF Uniforme, o sea el DIF presente es no Uniforme (Aguerri et al., 2011).

Siguiendo el desarrollo matemático realizado por Aguerri et al. (2011) se produce la disociación entre DIF Paralelo y DIF Uniforme cuando el modelo es el ML3P. Según Hanson (1998) El DIF es Uniforme en tanto exista  $\alpha \in \mathfrak{R}^+$ ,  $\alpha \neq 1$ , tal que resulte:

$$P(\theta) = c + (1 - c) \frac{1}{1 + e^{-1.7a(\theta-b)}}$$

$$f(\theta) = \frac{P_1(\theta)[1 - P_2(\theta)]}{P_2(\theta)[1 - P_1(\theta)]}$$

$$f(\theta) = \alpha = \frac{(1-c_2)[c_1+e^{1.7a_1(\theta-b_1)}]}{(1-c_1)[c_2+e^{1.7a_2(\theta-b_2)}]} \quad \forall \theta \in \mathfrak{R}$$

De donde debe ser:

$$\frac{\alpha(1-c_1)[c_2+e^{1.7a_2(\theta-b_2)}]}{1-c_2} = c_1+e^{1.7a_1(\theta-b_1)} \quad \forall \theta \in \mathfrak{R}$$

$$1 - c_2 = 1 - \frac{c_1}{\alpha+(1-\alpha)c_1} = \frac{\alpha(1-c_1)}{\alpha+(1-\alpha)c_1}$$

Operando resulta:

$$[\alpha(1-\alpha)c_1][c_2+e^{1.7a_2(\theta-b_2)}] = c_1+e^{1.7a_1(\theta-b_1)} \quad \forall \theta \in \mathfrak{R}$$

Luego:

$$[\alpha+(1-\alpha)c_1]e^{1.7a_2(\theta-b_2)} = e^{1.7a_1(\theta-b_1)} \quad \forall \theta \in \mathfrak{R}$$

De donde se deduce que (Aguerri et al., 2011):

$$\ln[\alpha+(1-\alpha)c_1]1.7a_2(\theta-b_2) = 1.7a_1(\theta-b_1) \quad \forall \theta \in \mathfrak{R}$$

Luego:

$$\ln[\alpha+(1-\alpha)c_1] = 1.7\theta(a_1-a_2) - 1.7a_1b_1 + 1.7a_2b_2 \quad \forall \theta \in \mathfrak{R}$$

Para que la expresión anterior no dependa de  $\theta$ , debe ser  $a_1 = a_2 = a$

De donde resulta:

$$\ln[\alpha+(1-\alpha)c_1] = 1.7a(b_2-b_1)$$

O sea (Aguerri et al., 2011):

$$b_2 = b_1 + \frac{1}{1.7a} \ln[\alpha + (1 - \alpha)c_1]$$

Luego para el ML3P un ítem exhibe DIF Uniforme con  $f(\theta) = \alpha$ , para todo  $\theta \in \mathfrak{R}$  siendo  $\alpha \in \mathfrak{R}^+$  y  $\alpha \neq 1$ , si:

$$a_1 = a_2 = a, b_2 = b_1 + \frac{1}{1.7a} \ln[\alpha + (1 - \alpha)c_1] \text{ y } c_2 = \frac{c_1}{\alpha + (1 - \alpha)c_1}$$

O bien

$$a_1 = a_2 = a, b_2 = b_1 + \frac{1}{1.7a} \ln \left[ \frac{c_1}{c_2} \right] \text{ y } c_2 = \frac{c_1}{\alpha + (1 - \alpha)c_1}$$

Despejando  $\alpha$  de  $c_2 = \frac{c_1}{\alpha + (1 - \alpha)c_1}$  resulta:

$$f(\theta) = \alpha = \frac{(1 - c_2) \cdot c_1}{(1 - c_1) \cdot c_2}$$

#### 2.4.2. Métodos de Tablas de Contingencia

De acuerdo a los métodos de Tablas de Contingencia, el nivel en el rasgo es estimado por el puntaje total. Si la performance de los individuos es similar entre los grupos para cada uno de los niveles de puntaje, se considera que no existe DIF. Cuando la ventaja de uno de los grupos permanece constante a lo largo de todos los niveles de puntaje, existe DIF Uniforme. Por el contrario, si varía, existe interacción entre el grupo y el rasgo. En ese caso el DIF es denominado No Uniforme.

La ventaja de los métodos de análisis del DIF de Tablas de Contingencia es que pueden aplicarse con muestras pequeñas y su costo computacional es más bajo, por lo que incluso pueden implementarse en Excel.

##### 2.4.2.1. Prueba de Mantel-Haenszel

En 1959 Nathan Mantel y William Haenszel presentan por primera vez en el marco de estudios sobre el cáncer la que actualmente es conocida como prueba de Mantel-

Haenszel (MH). Posteriormente Holland y Thayer (1988) la aplican al estudio del DIF. Es fácil de aplicar y no requiere cálculo iterativo. La prueba de Mantel-Haenszel es una de las más potentes para la detección del DIF Uniforme (Aguerri, Galibert, Zanelli, & Attorresi, 2005; Aguerri, 2012). Supone en la hipótesis nula que los *Odds Ratio* se mantienen constantes a lo largo de los  $k$  niveles de la variable estratificadora, es decir ausencia de DIF, mientras que la hipótesis alternativa es compatible con la presencia de DIF Uniforme o no Uniforme. Los ítems con presencia de DIF que requieran un nivel muy bajo o muy alto en el rasgo para ser respondidos favorablemente pueden no ser detectados por la prueba de Mantel-Haenszel (Hambleton, Clauser, Mazor, & Jones, 1993).

La razón entre la proporción de individuos que responden favorablemente al ítem y las que no lo hacen debe ser igual en los grupos comparados en todos los niveles  $k$  de puntuación total si no existe DIF. Las posibilidades del Grupo de Referencia ( $G_R$ ) y del Grupo Focal ( $G_F$ ) de responder favorablemente al ítem  $j$  son  $W_{Rk} = A_k/B_k$  siendo  $A_k$  la cantidad de individuos del  $G_R$  que responden favorablemente al ítem y  $B_k$  la cantidad de individuos del  $G_R$  que responden desfavorablemente al ítem y  $W_{Fk} = D_k/C_k$  siendo  $D_k$  la cantidad de individuos del  $G_F$  que responden favorablemente al ítem y  $C_k$  la cantidad de individuos del  $G_F$  que responden desfavorablemente al ítem. La razón de las posibilidades u *Odds Ratio* es  $\alpha_k = W_{Rk}/W_{Fk}$ , existiendo un *Odds Ratio* para cada uno de los  $k$  niveles. Tanto  $W_{Rk}$  como  $W_{Fk}$  son estimadores de sus parámetros  $\Omega$ . La prueba Mantel-Haenszel decide si la razón  $\alpha_k$  para todo nivel  $k$  del puntaje total es igual a 1 o difiere significativamente. Su estadístico sigue la distribución ji-cuadrado ( $\chi^2$ ) con un grado de libertad si la hipótesis nula es verdadera y se define como:

$$\chi_{MH}^2 = \left( \left| \sum_{k=1}^m A_k - \sum_{k=1}^m E(A_k) \right| - 0,5 \right)^2 / \sum_{k=1}^m Var(A_k)$$

Donde  $E(A_k)$  es la esperanza de  $A_k$ .

$$E(A_k) = (N_{Rk}N_{1k}) / N_k.$$

$Var(A_k)$  es la varianza de  $A_k$ .

$$Var(A_k) = N_{Rk}N_{Fk}N_{1k}N_{0k} / N_k^2(N_k-1).$$

Además de la prueba de hipótesis, el método MH proporciona una estimación de la magnitud del DIF: el cociente de razones común MH ( $\hat{\alpha}_{MH}$ ):

$$\hat{\alpha}_{MH} = \frac{\sum_{k=1}^m A_k D_k / N_k}{\sum_{k=1}^m B_k C_k / N_k}$$

Donde:

$N_k$  es el total de individuos con puntaje  $k$ .

$A_k$  es la cantidad de individuos del  $G_R$  que responden favorablemente al ítem.

$D_k$  es la cantidad de individuos del  $G_F$  que responden favorablemente al ítem.

$B_k$  es la cantidad de individuos del  $G_R$  que responden desfavorablemente al ítem.

$C_k$  es la cantidad de individuos del  $G_F$  que responden desfavorablemente al ítem.

Si el cociente de razones común MH es igual a la unidad debe interpretarse como ausencia de DIF Unidireccional. Los procedimientos de Mantel-Haenszel no son sensibles al DIF no Unidireccional cuando el punto de corte está en niveles medios del puntaje total. En ese caso conviene inspeccionar los gráficos. Si el cociente de razones común MH es mayor a la unidad el DIF favorece al  $G_R$ , por el contrario, si es menor a la unidad favorece al  $G_F$  (Dorans & Holland, 1993).

Holland y Thayer (1988) construyeron una medida del DIF en base a una transformación logarítmica del cociente de razones común MH. Denominaron a este estimador Delta-DIF de Mantel-Haenszel (MH D-DIF) y lo formularon como  $MH\ D-DIF = -2.35 \ln \hat{\alpha}_{MH}$ . Si su valor es nulo indica ausencia de DIF unidireccional (podría existir DIF no unidireccional con punto de corte en valores medios del puntaje total), si su valor es negativo el DIF favorece al  $G_R$  y si es positivo favorece al  $G_F$ .

El *Educational Testing Service* (ETS) utilizando pruebas de hipótesis y el valor absoluto del MH D-DIF para clasificar a los ítems en tres clases según la magnitud del DIF (Zieky, 1993):

- Tipo C: con DIF grande. Si MH D-DIF es significativamente mayor a 1 con  $\alpha = .05$  y es en valor absoluto mayor o igual que 1.
- Tipo B: con DIF leve a moderado. Si MH D-DIF presenta diferencias significativas de 0 con  $\alpha = .05$  y es en valor absoluto mayor que 1 pero menor que 1.5.
- Tipo A: son considerados como con DIF libre o despreciable. No presentan diferencias significativas de 0 con  $\alpha = .05$  o son en valor absoluto menores que 1.

Mazor, Clauser y Hambleton (1994) desarrollaron el procedimiento de Mantel-Haenszel modificado. Se basa en realizar la prueba de MH para tres grupos: la muestra completa, la submuestra de individuos con media inferior o igual a la mediatotal y la submuestra de individuos con media superior a la media total. Es suficiente que se rechace la hipótesis nula en alguna de las tres pruebas para que se considere que el ítem tiene DIF. El procedimiento de Mantel-Haenszel modificado incrementa la potencia para detectar DIF no Uniforme en comparación con el procedimiento estándar.

Ferreres Traver, González Romá y Gómez Benito (2002) aplicaron el criterio de Bonferroni al nivel de significación del estadístico  $\chi_{MH}^2$ . O sea, propusieron considerar al ítem con DIF cuando el estadístico  $\chi_{MH}^2$  resulte significativo a un nivel  $\alpha$  de significación de 0.05 dividido el número de ítems de la escala, con el fin de evitar cometer el error de Tipo I.

Además, consideraron al ítem con presencia de DIF si MH D-DIF es en valor absoluto mayor que 1.5. A este último criterio Aguerri, Zanelli, Galibert y Attorresi (2002) lo denominaron método empírico para la detección del DIF ya que reduce los requerimientos de la ETS ya que no se realiza la prueba de hipótesis.

#### 2.4.2.2. Pruebas de Breslow-Day

Las pruebas presentadas por Breslow y Day son específicamente adecuadas para detectar el DIF no Uniforme (Camilli & Shepard, 1994; Penfield, 2003).

La Prueba Global de Breslow-Day (BD, Breslow & Day, 1980) testea la hipótesis nula de homogeneidad de los *Odds Ratio* en todos los  $k$  niveles de puntaje total con

respecto a un determinado ítem. Su estadístico sigue la distribución ji-cuadrado ( $\chi^2$ ) con grados de libertad número de tablas de contingencia analizadas-1 si la hipótesis nula es verdadera y se define como:

$$BD = \sum_{k=1}^m \frac{(A_k - E(A_k))^2}{Var(A_k)}$$

Donde:

$A_k$  es la cantidad de individuos del  $G_R$  de puntaje total  $k$  que responden favorablemente al ítem.

$E(A_k)$  es la esperanza de  $A_k$ . Para su cálculo se utiliza el supuesto de que los odds ratio para la totalidad de los niveles de puntaje total tienen igual valor. Se deduce que:

$$E(A_k) = (\hat{\alpha}(N_{Rk} + N_{1k}) + (N_{Fk} - N_{1k}) \pm \pm \left\{ [\hat{\alpha}(N_{Rk} + N_{1k}) + (N_{Fk} - N_{1k})]^2 - [4(\hat{\alpha} - 1)\hat{\alpha}(N_{Rk}N_{1k})] \right\}^{1/2}) / 2(\hat{\alpha} - 1)$$

Donde:

$N_{Rk}$  es la frecuencia marginal que indica la cantidad de individuos del Grupo de Referencia con puntaje total  $k$ .

$N_{1k}$  es la frecuencia marginal que indica la cantidad de individuos que respondieron favorablemente al ítem o que lo acertaron con puntaje total  $k$ .

$N_{Fk}$  es la frecuencia marginal que indica la cantidad de individuos del Grupo Focal con puntaje total  $k$ .

Se deduce que  $Var(A_k)$  es la varianza de  $A_k$ :

$$Var(A_k) = \left( \frac{1}{E(A_k)} + \frac{1}{N_{Rk} - E(A_k)} + \frac{1}{N_{1k} - E(A_k)} + \frac{1}{N_{Fk} - N_{1k} + E(A_k)} \right)^{-1}$$

Si la razón de las posibilidades u *Odds Ratio* es igual a la unidad indica ausencia de DIF, si es diferente de uno hay DIF Uniforme, mientras que la hipótesis alternativa afirma la presencia de DIF no Uniforme.

La prueba de Breslow-Day de la Tendencia en la Heterogeneidad de los *Odds Ratio* (BDT, Breslow & Day, 1980) en su hipótesis nula sostiene la homogeneidad de los *Odds Ratio*, lo cual puede indicar ausencia de DIF o presencia de DIF Uniforme. La hipótesis alternativa afirma que existe una relación lineal entre los *Odds Ratio* y los puntajes, es decir presencia de DIF no Uniforme. Su estadístico sigue la distribución ji-cuadrado ( $\chi^2$ ) con un grado de libertad si la hipótesis nula es verdadera y se define como:

$$BDT = \frac{\left[ \sum_{k=1}^m x_k (A_k - E(A_k)) \right]^2 - 0.5}{\sum_{k=1}^m x_k^2 Var(A_k) - \left[ \sum_{k=1}^m x_k Var(A_k) \right]^2 / \sum_{k=1}^m Var(A_k)}$$

Penfield (2003) construyó el método de Regla de Decisión Combinada (*Combined Decision Rule, CDR*) que combina BDT con MH. El ítem se considera con DIF si BDT o MH así lo indican. Aplicó el ajuste de Bonferroni en el nivel de significación de ambas pruebas, es decir dividió el alfa por la cantidad de pruebas de significación, en este caso dos.

#### 2.4.2.3. Diferencia de Proporciones Estandarizada (DPE)

La Diferencia de Proporciones Estandarizada (DPE) fue propuesta por Dorans y Kulick (1983, 1986) y desarrollada por Dorans y Holland (1993). Si un ítem fue detectado con DIF, la DPE permite describir el sentido de dicho funcionamiento diferencial. Se trata de un estadístico que estandariza y combina la diferencia de proporciones de respuestas favorables al ítem entre  $G_R$  y  $G_F$  en todos los niveles  $k$  de puntuación total. Si toma un valor positivo el DIF favorece al  $G_F$ , por el contrario, si es negativo favorece al  $G_R$ . Si está comprendida entre -0.05 y 0.05 indica ausencia de DIF, entre 0.05 y 0.10 o entre -0.10 y -0.05 se aconseja revisión del ítem. Si está fuera del rango [-0.10, 0.10] se considera al ítem altamente sospechoso.

$$DPE = \frac{\sum_{k=1}^s n_{Fk} \Delta p_k}{\sum_{k=1}^s n_{Fk}}$$

Donde:

$n_{Fk}$  es el total de examinados del  $G_F$  en el nivel  $k$  de puntuación total.

$\Delta p_k$  es la proporción de respuesta favorable al rasgo o correcta del  $G_F$  menos la proporción de respuesta favorable al rasgo o correcta del  $G_R$ .

#### 2.4.3. TCT, TRI y Ausencia de DIF

En este recorrido, se pudo delimitar cómo desde la TCT y la TRI es posible aportar evidencias de validez y confiabilidad que indican si la prueba tiene una calidad psicométrica aceptable. Mientras que los Análisis Factoriales Exploratorio y Confirmatorio, propios de la TCT, examinan la estructura interna de la prueba, los modelos de la TRI brindan evidencias acerca de si los ítems dan información acerca del rasgo latente y en qué niveles del rasgo es mayor dicha información. Por otra parte, el análisis del DIF atraviesa tanto la TCT como la TRI y constituye un requisito ineludible para la construcción de un test válido y equitativo.

## SEGUNDA PARTE

### ESTUDIO PSICOMÉTRICO

#### 3. *Objetivos e Hipótesis*

##### 3.1. *Objetivos*

El estudio de la relación de los sujetos con la sociedad se ha basado típicamente en la exploración de comportamientos antisociales. Sólo más recientemente se ha focalizado el comportamiento prosocial, su antítesis, como objeto de estudio. Dentro de este contexto, resulta contrastante la gran cantidad de investigaciones acerca de la conducta prosocial en la infancia y en la adolescencia con los escasos desarrollos enmarcados en etapas posteriores. Se podría afirmar, así, que el estudio de dicho constructo en la adultez y la vejez es un campo incipiente, con una gran variedad de aristas por explorar.

En el plano internacional, se destaca para la medición de la conducta prosocial en adultos, la PSA de Caprara et al. (2005) por su brevedad, sencillez y alcance de los estudios psicométricos realizados, que permiten concluir no sólo acerca de su consistencia interna, sino también acerca de su capacidad de discriminación entre sujetos, ya que los autores utilizan TRI. Sin embargo, la mayoría de los parámetros  $b$  hallados fueron negativos, mostrando que los ítems requirieron bajos niveles de prosocialidad para ser respondidos en forma favorable al rasgo.

Es decir que aún internacionalmente, a la escasez de instrumentos para la medición de la conducta prosocial en la etapa adulta, se suma la desactualización de la mayoría de ellos y que el más novedoso, la PSA de Caprara et al. (2005), brinda un bajo nivel de información en niveles moderadamente altos y altos del rasgo, lo cual constituye una importante falencia.

Además, con respecto a la validez de los test de conducta prosocial para adultos, las evidencias brindadas son sólo mediante la correlación con otros instrumentos sin aportar una relación con criterios externos no instrumentales. En general, no se elimina

de la muestra a los sujetos con alta deseabilidad social o distorsión ni se descartan los datos anómalos de la muestra. Tampoco, en la mayoría de los instrumentos, se evalúa la existencia de DIF.

Estos factores denotan un evidente vacío local en lo que respecta a la medición de la conducta prosocial en la etapa adulta. Además, podrían existir cuestiones culturales (Cortada de Kohan & Macbeth, 2005) que hagan inconveniente realizar una adaptación ya que, como se mencionó anteriormente, la cultura ejercería influencia en el desarrollo del comportamiento prosocial enfatizando ciertos valores particulares (Keller, 2007).

Por lo tanto, sería de utilidad la construcción de un instrumento de medición de la conducta prosocial en la etapa adulta realizado en la Argentina que cubra este vacío local, contando con una adecuada depuración de los datos anómalos y de sujetos con alta deseabilidad social. Una escala de conducta prosocial que aporte evidencias de validez mediante análisis factorial exploratorio y confirmatorio, análisis de DIF, modelización mediante la TRI, y que, además de la correlación con instrumentos, cuente con un criterio externo para su validación. Un criterio externo posible no instrumental es estar cursando una carrera universitaria que implique directamente a las conductas prosociales. Para hacer un estudio diferencial se tomarán estudiantes universitarios de carreras que impliquen y que no impliquen directamente a las conductas prosociales.

Por lo tanto, se formulan los siguientes objetivos:

### 3.1.1. Objetivo General

Construir una Escala diseñada *ad hoc* para la medición de la conducta prosocial aplicada a poblaciones de estudiantes terciarios y universitarios en general y modelizarla con la Teoría de Respuesta al Ítem.

### 3.1.2. Objetivos Específicos

1.- Elaborar un conjunto de ítems para la medición de la conducta prosocial en población terciaria y universitaria.

2.- Realizar una depuración primaria y secundaria de los ítems utilizando indicadores de calidad de la Teoría Clásica de los Test para seleccionar los reactivos.

3.- Establecer la estructura dimensional de la Escala mediante Análisis Factorial Exploratorio y Confirmatorio.

- 4.- Aplicar los Modelos Logísticos de Uno, Dos y Tres Parámetros.
- 5.- Aplicar el Modelo de Respuesta Graduada.
- 6.- Aplicar el Modelo de Crédito Parcial.
- 7.- Aportar evidencias de validez de los reactivos mediante:
  - a- Fuentes Internas de Validez
    - Análisis de contenido a través de la crítica de jueces expertos.
    - Evaluación de la dimensionalidad mediante Análisis Factorial Exploratorio y Confirmatorio de la estructura interna de la Escala.
    - Estudios de detección del Funcionamiento Diferencial del Ítem según el género de los participantes.
    - La modelización de las respuestas de los participantes mediante diferentes modelos de la TRI: ML1P, ML2P, ML3P, MRG y MCP.
  - b- Fuentes Externas de Validez
    - Evaluación de la validez convergente mediante:
      - La correlación entre las puntuaciones en la escala de conducta prosocial construida y las puntuaciones en las dimensiones de la Escala de Habilidades Sociales de Morales Rodríguez y Suárez Pérez (2011) adaptada a la población argentina.
      - La correlación entre las puntuaciones en la escala de conducta prosocial construida y enunciados contruidos acerca de la Autopercepción de los participantes en Empatía, Conducta Prosocial, Datividad y Altruismo.
    - Análisis de las relaciones test - criterio mediante:
      - La evaluación de relaciones estadísticas entre género y los puntajes en la escala de conducta prosocial.
      - La evaluación de relaciones estadísticas entre el tipo de carrera - relacionada directamente con la conducta prosocial o no- que el participante está cursando y los puntajes en la escala de conducta prosocial.

8.- Obtener evidencias de confiabilidad mediante el índice Alfa de Cronbach y los índices de confiabilidad marginal y las Funciones de Información de los Ítems y del Test calculados mediante ML1P, ML2P, ML3P, MRG y MCP.

9.- Establecer puntos de corte para determinar el nivel de conducta prosocial de los participantes, obteniendo la proporción de sujetos con muy bajo nivel, bajo nivel, nivel intermedio, alto nivel y muy alto nivel de conducta prosocial.

### *3.2. Hipótesis*

Al tratarse de una investigación de desarrollo instrumental (Ato, López & Benavente, 2013; Montero & León, 2007) y por tanto no experimental, no pueden plantearse las tradicionales hipótesis. Sin embargo, es posible bosquejar los siguientes resultados que razonablemente podrían esperarse:

1.- Dados los antecedentes teóricos y de medición recolectados, se prevee que la Escala de Conducta Prosocial tenga una estructura multidimensional.

2.- Con respecto a los modelos de la TRI desarrollados para ítems dicotómicos, ya que la escala no se trata de una prueba de ejecución máxima donde ante el desconocimiento existen intentos de acertar por azar, se descarta que podría ser adecuado el ML3P. Como también se hipotetiza que los ítems pueden tener distinto grado de discriminación, por ende, el Modelo Logístico de Dos Parámetros sería más adecuado que el ML1P.

3.- Con respecto a los modelos de la TRI desarrollados para ítems politómicos, se hipotetiza que el Modelo Respuesta Graduada sea el más adecuado ya que como se ha explicado previamente, presenta mayor calidad psicométrica y contempla que los ítems puedan tener distinto grado de discriminación.

4.- Se prevee, dados los antecedentes teóricos, que las mujeres obtengan puntuaciones significativamente más altas que los varones.

5.- Se prevee que los estudiantes terciarios y universitarios de profesiones que implican solidaridad, ayuda o conductas empáticas hacia otros puntuarán significativamente más alto en la escala que aquellos estudiantes de carreras que no involucren en forma directa una conducta prosocial.

6.- Se espera que los puntajes en la Escala de Conducta Prosocial correlacionen en forma positiva, significativa y, por lo menos, moderada, con los puntajes en las dimensiones de la escala de Habilidades Sociales de Morales Rodríguez y Suárez Pérez adaptada a la argentina y con los enunciados construidos de Autopercepción del participante en Empatía, Conducta Prosocial, Datividad y Altruismo.

#### 4. *El Proceso de Construcción de la Escala de Conducta Prosocial*

Para construir una escala diseñada ad-hoc que evalúe la conducta prosocial los reactivos elaborados y ensayados fueron analizados bajo los supuestos de calidad psicométrica. Previamente, se demarcó teóricamente la definición conceptual del constructo conducta prosocial. Luego de delimitar conductas representativas de la conducta prosocial, se determinó el formato de los ítems y se redactaron reactivos acordes a las mismas. Cuando se define conceptualmente el constructo que se pretende medir y cuando se elaboran los ítems que permiten su operacionalización es recomendable tener presentes los supuestos de las teorías psicométricas que posteriormente se aplicarán Padilla y González (2001) a fin de poder satisfacerlos. Posteriormente se ensayó la prueba administrando los ítems a diferentes muestras de participantes (Cohen & Swerdlik, 2006; Martínez Arias, 1995). Simultáneamente, se decidió además de la TCT qué modelos psicométricos de la TRI se utilizarán.

En los siguientes apartados se brindarán los detalles del procedimiento que se llevó a cabo para cumplir con los objetivos previstos.

Teniendo en cuenta el Estado del Arte y considerando los requerimientos que supone la TRI en relación a la unidimensionalidad del rasgo latente se presentará una definición operacional para el constructo que será medido.

A los fines del trabajo propuesto en la presente Tesis se define a la conducta prosocial como un fenómeno complejo que involucra acciones de los individuos basadas en creencias y sentimientos y que describe la forma en que éstos se orientan hacia los otros al realizar conductas solidarias. Se adopta la definición de González Portal (2000), que demarca el término en su acepción más abarcativa, entendiendo por tal a toda conducta social positiva con o sin motivación altruista. Esta definición pone el acento en la acción en sí misma, y deja en un segundo plano el análisis acerca de cuál es su motivación.

Dados los antecedentes expuestos en la bibliografía, esta variable se considera *a priori* como multidimensional. Bajo esta condición, el supuesto de unidimensionalidad de los modelos de la TRI consignados obliga a efectuar estudios independientes para cada una de las dimensiones resultantes.

## 5. Metodología

### 5.1. Diseño General

#### 5.1.1. Tipo de Estudio

Se trata de un tipo de investigación no-experimental dada la ausencia de manipulación de variables independientes y de aleatorización (Anguera, Arnau, Ato, Martínez, Pascual, & Vallejo, 1995). Según la clasificación de los tipos de investigación propuesta por Montero y León (2007) se trataría de un estudio instrumental. Pertenecen a esta categoría os los estudios encaminados al desarrollo de pruebas y aparatos, incluyendo tanto el diseño -o adaptación- como el estudio de las propiedades psicométricas de los mismos (Montero & León, 2007).

#### 5.1.2. Instrumentos

Se administraron los siguientes instrumentos a la muestra:

*Escala de Conducta Prosocial.* Sobre la que se realizó la depuración final del instrumento. Se trata de 29 ítems (ver Apéndice D), donde la modalidad de respuesta se especificó con una escala Likert de seis opciones que refleja la frecuencia con que se realiza la acción (1= *Nunca*, 2 = *Casi Nunca*, 3 = *A Veces*, 4 = *Con Frecuencia*, 5 = *Casi Siempre*, 6 = *Siempre*).

*Escala Distorsión del Cuestionario de los Cinco Grandes* (Big Five Questionnaire, [BFQ] de Caprara et al., 1995). Está formada por 12 ítems que se contestan en una escala Likert de cinco opciones de respuesta (desde *Completamente falso para mí* a *Completamente verdadero para mí*). La escala de Distorsión mide la deseabilidad social entendida como un estilo de respuesta que el sujeto puede poner en juego para mostrar aspectos favorables de su personalidad y eludir los menos favorables. Previamente a su aplicación, un grupo de expertos verificó la equivalencia lingüística y conceptual. En el manual de la adaptación, Bermúdez (1995) refiere un índice de consistencia interna Alfa

de Cronbach de .77. En la presente investigación se halló un índice más bajo (Alfa de Cronbach de .64) pero aceptable.

*Versión Argentina de la Escala de Habilidades Prosociales (EHP)* de Morales Rodríguez y Suárez Pérez (2011). Los ítems son planteados en términos de la evaluación de una situación particular y valorados mediante una escala Likert de cuatro opciones (*No me describe, Me describe poco, Me describe regular y Me describe bien*). Dicha Escala fue adaptada a la Argentina por Auné, Abal y Attorresi (2014). La EHP adaptada consta de 12 ítems divididos en tres dimensiones: a) Toma de Perspectiva (4 ítems, Alfa de Cronbach de .66), b) Asistencia (3 ítems, Alfa de Cronbach de .73) y c) Altruismo (5 ítems, Alfa de Cronbach de .68). Para la muestra utilizada en la presente Tesis ( $n = 1383$ ), los índices Alfa de Cronbach tuvieron los siguientes valores: Toma de Perspectiva: .65, Asistencia .72, Altruismo: .63. El protocolo de la Versión Argentina de la EHP se puede encontrar en el Apéndice E.

*Autopercepción de la Empatía, Conducta Prosocial, Dadivosidad y Altruismo.* Consta de cuatro enunciados, contruidos para la presente Tesis, que permiten obtener una autovaloración global del individuo (en una escala de 0 a 10) respecto de su nivel de empatía, conducta prosocial, dadivosidad y altruismo. Con el fin de reducir ambigüedades en la comprensión de cada aspecto, los sujetos recibieron una caracterización de cada constructo considerando que: a) una persona empática se pone en el lugar de otras personas, las comprende y experimenta sentimientos similares, apoyándolas emocionalmente; b) una persona prosocial realiza acciones que benefician a los demás; c) una persona dadivosa da, presta o comparte objetos, tiempo, dinero, experiencias, u otro aspecto de valor y d) una persona altruista busca beneficiar a los demás aunque esto la perjudique. Posteriormente se le solicita al participante que se autovalore en dichas características. El protocolo con estos enunciados se puede encontrar en el Apéndice F.

### 5.1.3. Pautas Éticas

El diseño de investigación propuesto contempla las pautas éticas locales propuestas por la Federación de Psicólogos de la República Argentina (FePRA, 2013) y la Asociación Argentina de Estudio e Investigación en Psicodiagnóstico (ADEIP, 1999).

Asimismo, sigue las normativas internacionales planteadas por la Asociación Americana de Psicología (*American Psychological Association, APA, 2010*) y el Consejo Nacional de Medición en Educación (*National Council on Measurement in Education, 2013*) para la adaptación y la validación de tests. La escala es anónima y se les informó a los participantes que los datos derivados de esta investigación se utilizarán con fines exclusivamente científicos bajo la Ley Nacional 25.326 de Protección de Datos Personales, sin ser expuestos de forma individual en ningún momento.

## 6. *Procedimientos y resultados para la Concreción de los Objetivos Propuestos*

### 6.1. *Fase I: Construcción de Ítems, Estudio Piloto y Crítica de Jueces Expertos*

#### *Procedimiento*

Se realizó una revisión de las tipologías más recientes de conducta prosocial con el fin de definir adecuadamente el constructo (Caprara et al., 2005; Hay & Cook, 2007; López de Cózar, et al., 2008; Warneken & Tomasello, 2009), observándose las dimensiones más frecuentes y relevantes postuladas en la literatura específica (Auné & Attorresi, 2014; Auné et al., 2014). Posteriormente se definió a la conducta prosocial como un fenómeno complejo que involucra acciones de los individuos basadas en creencias y sentimientos y que describe la forma en que éstos se orientan hacia los otros al realizar conductas solidarias. Dicha labor permitió delimitar una clasificación de cuatro dimensiones de la conducta prosocial: Confortar (CF), Altruismo (AL), Ayuda (AY) y Compartir y Donar (CD).

Se definió operacionalmente Confortar como aquellos comportamientos que manifiestan empatía, comprensión, refuerzo y soporte emocional. Se consideró como Altruistas a aquellas conductas que impliquen beneficiar a otros, aunque perjudiquen el bienestar personal. La dimensión de Ayuda se operacionalizó como comportamientos de asistencia, cuidado y compromiso con los otros. Implica necesariamente un benefactor y un beneficiario claramente diferenciados. Por último, se entendió por Compartir y Donar: dar, donar, prestar o compartir objetos, tiempo, dinero, experiencias u otro aspecto de valor para los usos y costumbres (Auné et al., 2014).

Se elaboró un conjunto inicial de 114 ítems (ver Apéndice B) representativos de las cuatro dimensiones identificadas mediante revisión teórica (Auné, Blum, Abal, & Attorresi, 2014). Se planteó al evaluado un conjunto de situaciones habituales con el fin de obtener información acerca de cómo se comporta usualmente, lo que es característico de los tests de comportamiento típico (Cronbach, 1972). Se han utilizado, entonces, en los ítems, conductas generales propias de la vida cotidiana (e.g. *Intento “levantar” la autoestima a mis amigos*) así como comportamientos específicos del contexto de educación (e.g. *Comparto con mis compañeros las experiencias que tuve en las materias que ya cursé*). En ningún caso se incluyó en el ítem la motivación con la cual se realizó la conducta, ya que el concepto de comportamiento prosocial pone el acento en la acción

en sí misma, es decir en el acto de beneficiar a otros, sea la intención primaria egoísta o desinteresada (Gonzalez Portal, 2000).

Los ítems se redactaron en dos direcciones para prevenir la tendencia a elegir según la dirección -aquiescencia-. Además, se cuidó de redactar una cantidad equilibrada de reactivos por cada dimensión. La redacción se efectuó utilizando un criterio donde cada reactivo hizo referencia a un comportamiento prosocial específico. Se incluyeron formas homólogas de los mismos reactivos para poder realizar una selección de las más adecuadas.

Por ejemplo, algunos de los ítems construidos fueron: *Cuando siento que alguien está mal, le demuestro que lo entiendo* (CF), *Aunque pudiera cobrar mucho dinero por él, compartiría gratuitamente un descubrimiento que contribuyera a un mundo mejor* (AL), *Ayudo a ancianos a cruzar la calle o cargar objetos* (AY), *Comparto las malas experiencias de la vida con amigos para que no cometan los mismos errores* (CD). En el caso de que el sujeto no haya tenido la oportunidad de estar en una situación que requiera la conducta se aclaró que debían responder en base a cómo suponen que actuarían, con sus valores, personalidad y mentalidad, imaginándose la situación. Se pidió expresamente sinceridad al responder, explicando que no hay respuestas correctas o incorrectas y que la prueba es anónima. El anonimato tendería a disminuir la deseabilidad social (Fisher, 1993).

Con respecto a la modalidad de respuesta, la misma se construyó con formato de escala Likert de seis opciones que refleja la frecuencia con que se realiza la acción (1= *Nunca*, 2 = *Casi Nunca*, 3 = *A Veces*, 4 = *Con Frecuencia*, 5 = *Casi Siempre*, 6 = *Siempre*). No se ha incluido una categoría central.

Una revisión preliminar por parte de un grupo de jueces expertos en Psicometría permitió disminuir la cantidad de ítems a 62 (Ver Apéndice C). En esta etapa se seleccionaron los reactivos más adecuados dentro de las formas homólogas redactadas previamente y además los jueces realizaron ciertas modificaciones en algunos de los ítems. Se procuró especialmente los reactivos seleccionados abarcaran distintas intensidades de beneficio hacia el otro -es decir que potencialmente discriminen en todos los niveles de conducta prosocial incluidos los extremos-, así como diferentes receptores de la acción -compañeros, amigos, conocidos, desconocidos-. Se buscó que fueran claros, de fácil comprensión con palabras sociales, relevantes y representativos de cada categoría (Auné, et al., 2014).

Los 62 reactivos fueron sometidos a un proceso de depuración primaria más exhaustivo. En primer lugar, fueron evaluados por cinco jueces expertos en base a los siguientes criterios, aportando evidencias de validez por fuente interna y de contenido:

- a) Grado de relevancia como indicador del constructo,
- b) Congruencia con la definición operacional de la dimensión medida,
- c) Adecuación respecto de los aspectos sintácticos y semánticos y al nivel de comprensión de la población muestreada (Cohen & Swerdlik, 2006; Tornimbeni, Pérez & Olaz, 2008).

El análisis de los conceptos emitidos por los jueces se realizó siguiendo los criterios de Compas, Davis, Forsythe & Wagner (1987), es decir el porcentaje de acuerdo entre jueces y el coeficiente de concordancia W de Kendall. Para el criterio de Compas et al., un porcentaje mayor a 90% se considera acuerdo mientras que el coeficiente W de Kendall indica concordancia cuando se rechaza la hipótesis de acuerdo nulo. La concordancia es mayor cuanto más se aproxima a 1 el valor absoluto del W de Kendall.

Posteriormente, se realizó una prueba piloto con los reactivos conservados en un grupo de 56 sujetos con el fin de detectar aquellos términos o expresiones que resulten confusos para los examinados tanto en los ítems como en las instrucciones o perjudiquen la validez aparente de la prueba.

### *Participantes Muestra Piloto*

Conformada por estudiantes de distintas carreras de la Universidad de Buenos Aires (UBA) de Argentina. La muestra estuvo compuesta por 56 sujetos, de los cuales 82% fueron mujeres y 18% varones. La edad varió entre 18 y 56 años con una media de 21.5, una mediana de 21 y un desvío estándar de 5.1.

### Resultados

#### *Crítica de Jueces Expertos*

Para a) el porcentaje de acuerdo interjueces fue de 92% y el índice de concordancia W de Kendall fue de .78 ( $p < .01$ ). Para b) de 93% y de .81 ( $p < .01$ ) respectivamente. Para c) de 95% y de .83 ( $p < .01$ ) respectivamente.

Un total de 29 ítems (ver Apéndice D) fueron los que cumplieron con los criterios de adecuación mediante la depuración por jueces. Los jueces expertos aportaron

sugerencias y modificaciones sobre los ítems y sobre las consignas y opciones de respuesta.

### *Muestra Piloto*

Los participantes manifestaron una adecuada comprensión de los reactivos y de las consignas depurados previamente por los jueces expertos.

## *6.2. Fase II: Análisis Factorial Exploratorio y Confirmatorio. Depuración Secundaria de Ítems.*

### *Participantes*

Se solicitó la colaboración voluntaria de estudiantes universitarios de Argentina. Luego de depurar los protocolos con más de 10% de respuestas incompletas, la muestra estuvo compuesta por 1540 sujetos, de los cuales 80% fueron mujeres y 20% varones. La media de edad fue de 22.4 años, la mediana de 20 y el desvío estándar de 5.8. La edad varió entre 18 y 60 años. Las carreras con mayor peso en la muestra fueron Psicología (29%), Medicina (10%), Abogacía (7%), Contador Público (7%), Lic. en Administración de Empresas (4.2%) y Lic. en Artes Audiovisuales (3.6%). El resto de los sujetos se repartieron en más de 50 carreras diferentes (ver Apéndice A). El 73.5% de los estudiantes relevados cursa su carrera en la Universidad de Buenos Aires, el 18.9% en otras universidades nacionales y el 7.6% en universidades privadas.

### *Análisis de Datos*

El protocolo de 29 ítems (ver Apéndice D) se administró sin tiempo límite a la muestra. Las respuestas fueron volcadas en una base de datos aplicando el formato de codificación tradicional. Con respecto a ítems no respondidos, los mismos fueron reemplazados por la Media en dicho ítem. Se homogeneizó la dirección de los ítems para que en todos los reactivos las opciones se encuentren ordenadas de menor a mayor nivel del constructo medido. Es decir, por ejemplo, que responder *Siempre* en un ítem redactado en sentido positivo equivale a responder *Nunca* a un ítem de sentido negativo.

En relación al análisis de datos anómalos multivariados, se calculó la distancia de Mahalanobis utilizando un punto de corte de  $p \leq .001$  (Tabachnick & Fidell, 2001). Los datos anómalos u *outliers* multivariantes son observaciones que se consideran extrañas no por el valor que toman en una determinada variable, sino en un conjunto de variables (Muñoz García & Anón Uribe, 2013). En este caso se trata de un patrón de respuestas a los ítems muy poco probable. Su presencia puede perjudicar los análisis estadísticos por lo que es deseable su detección y exclusión de la muestra. Se dice que un *outlier* enmascara a un segundo, si el segundo *outlier* puede ser detectado como un valor extremo sólo si primero es eliminado (Muñoz García & Anón Uribe, 2013). Por lo tanto, se debe eliminar ambos. En forma progresiva, para evitar el efecto de enmascaramiento, se eliminaron 51 casos de la muestra. Además, se descartaron 106 casos más de la muestra por puntuar más de un desvío y medio por encima de la media utilizando el criterio de Inglés et al. (2008) en la Escala Distorsión del BFQ. La cantidad total de individuos analizados, entonces, fue de 1383. El tamaño de la muestra original, como ya se indicó, era de 1540 participantes.

### *Procedimiento General*

Además, se obtuvieron los valores del índice de Mardia (1970) para la asimetría y curtosis multivariante con el macro para SPSS de DeCarlo (1997). Los valores del índice de Mardia (1970) para la asimetría y curtosis multivariante obtenidos con el macro para SPSS de DeCarlo (1997) (coeficiente de asimetría de 33.38,  $p < .00001$ ; coeficiente de curtosis de 21.62,  $p < .00001$ ) mostraron que los datos no se distribuían de forma normal multivariante. Como ya se ha mencionado, cuando la distribución no es normal existe una tendencia del estadístico  $\chi^2$  a rechazar la hipótesis nula (Bearden, Sharma, & Teel, 1982). Por lo tanto, el ajuste del modelo se evaluó con base en indicadores que no dependen directamente del  $\chi^2$  (Ferrando & Lorenzo Seva, 2014): el índice de bondad de ajuste (*Goodness of Fit Index*, GFI) y la raíz media cuadrática residual (*Root Mean Square Residual*, RMSR).

Se separó la muestra en dos partes en forma aleatoria: las respuestas de 761 participantes (55%) se utilizaron en un AFE mientras que las respuestas de 622 participantes (45%) se utilizaron para realizar un AFC. De esta forma, la estructura hallada en el AFE se corroboró en el AFC. Se obtuvieron distribuciones de frecuencias y

estadísticos descriptivos de cada ítem conservado: media (M), mediana (Mdn), desvío estándar (DE), asimetría univariante (As) y curtosis univariante (K).

### *Procedimiento del Análisis Factorial Exploratorio*

De estos 1383 sujetos, se seleccionaron aleatoriamente 761 participantes para realizar un Análisis Factorial Exploratorio (AFE). Se utilizaron la medida de adecuación muestral de Kaiser-Meyer-Olkin (KMO) y la prueba de esfericidad de Bartlett para analizar si la realización del AFE era pertinente. Una vez verificada dicha adecuación (KMO = .90; Bartlett:  $\chi^2(406) = 7441; p = .00001$ ) se procedió a realizar el AFE mediante el programa Factor 9.2. (Lorenzo Seva & Ferrando, 2013) considerando una estructura, en primer lugar, tetrafactorial. Posteriormente, en sucesivos análisis se eliminaron los reactivos asociados a los mayores residuos estandarizados hasta que el RMSR satisfizo el criterio de Kelley (1935). Como el tamaño de muestra es 761, se consideró un buen ajuste cuando RMSR fue .0331 que es menor que  $.0363 = 1/\sqrt{761}$ . Una vez que el modelo ajustó, se conservaron los ítems con cargas factoriales mayores a .30 y que, en cada dimensión, cumpliesen el criterio de la TCT de no aumentar el coeficiente Alfa de Cronbach al quitar el reactivo.

Actualmente se afirma que cuando la validez de constructo es analizada con datos ordinales provenientes de Escalas Likert, los factores resultantes muestran un mejor ajuste al modelo teórico si la factorización es realizada con la matriz de correlaciones policóricas antes que la de Pearson (Holgado Tello, Chacón Moscoso, Barbero García, & Vila Abad, 2010). La utilización de la correlación de Pearson implica suponer distribución normal, y un nivel de medición de la variable, como mínimo, intervalar.

El hallazgo de correlaciones interfactor mayores a .32 motivó la utilización del método Promax para la rotación (Tabachnick & Fidell, 2001) con un parámetro kappa de 4 (Hendrickson, & White, 1964). Esta rotación oblicua no asume incorrelación entre los factores y es por ello que resulta adecuada para estos casos.

El método de extracción de factores fue el de mínimos cuadrados no ponderados (*Unweighted Least Squares*, ULS), que, según las recomendaciones actuales, resulta más adecuado que el de máxima verosimilitud cuando la factorización de reactivos se realiza a través de la matriz de correlaciones policóricas (Lloret-Segura, Ferreres-Traver, Hernández-Baeza, & Tomás-Marco, 2014).

Se utilizó el análisis paralelo (Horn, 1965) con implementación óptima (Timmerman & Lorenzo Seva, 2011) para determinar el número de factores. Dicho

método posee cierta robustez frente a desviaciones de la normalidad (Buja & Eyuboglu, 1992), especialmente cuando se basa en la matriz de correlaciones policóricas (Garrido, Abad, & Ponsoda, 2013), como en este caso. Asimismo, cada factor debió poseer al menos cuatro ítems con correlaciones iguales o superiores a .40, para considerarlo como tal. Se interpretó al factor desde el punto de vista del constructo atendiendo a las correlaciones ítem-factor más elevadas (Glutting, Monaghan, Adams, & Sheslow, 2002).

### *Resultados del Análisis Factorial Exploratorio y Estadísticos Descriptivos de los Ítems Conservados*

Previamente a la realización del Análisis Factorial Exploratorio, se calculó el coeficiente KMO que indicó la pertinencia de realizar el análisis (KMO = .83) y la prueba de esfericidad de Bartlett:  $\chi^2(465) = 1627$ ;  $p = .00001$ . La estructura de cuatro factores no pudo sostenerse ya que la implementación óptima del análisis paralelo recomendó la reducción a tres dimensiones. Consecuentemente, se realizó un nuevo AFE imponiendo una estructura trifactorial. Posteriormente, en sucesivos análisis se eliminaron los reactivos asociados a los mayores residuos estandarizados, hasta que el RMSR satisfizo el criterio de Kelley. Al quitarse los reactivos con mayores residuos la implementación óptima del análisis paralelo sugirió una estructura bifactorial. Se continuó con el mismo criterio para eliminar ítems considerando dos factores.

Se calcularon nuevamente la medida de adecuación muestral KMO y la prueba de esfericidad de Bartlett para analizar si la realización del AFE era pertinente. Una vez verificada dicha adecuación (KMO = .90; Bartlett:  $\chi^2(406) = 7441$ ;  $p = .00001$ ) se procedió a realizar un nuevo AFE considerando una estructura bifactorial. Posteriormente, en sucesivos análisis se eliminaron los reactivos asociados a los mayores residuos estandarizados hasta que el RMSR satisfizo el criterio de Kelley. Como el tamaño de muestra es 761, se consideró un buen ajuste cuando RMSR fue .0331 que es menor que  $.0363 = 1/\sqrt{761}$ . Una vez que el modelo ajustó, se conservaron los ítems con cargas factoriales mayores a .30 y que, en cada dimensión, cumplieren el criterio de la TCT de no aumentar el coeficiente Alfa de Cronbach al quitar el reactivo. Con estos criterios, se eliminaron 13 ítems de los 29 originales.

Así, un modelo de 16 ítems y dos factores logró adecuados índices de ajuste (RMSR = .0331, Cr. Kelley < .0363; GFI = .99, mayor a .95; Hu & Bentler, 1999). La

varianza explicada por los dos factores fue de 44.67%. La implementación óptima del análisis paralelo (Timmerman & Lorenzo Seva, 2011) indicó la pertinencia de la estructura bifactorial. Ambos factores obtuvieron un coeficiente alfa de Cronbach de .79. Según el criterio de DeVellis (2012) este valor puede considerarse como aceptable. La correlación entre ambos factores fue de .505.

La interpretación de cada uno de los dos factores de este modelo se realizó observando las correlaciones ítem-factor más elevadas (Glutting, et al. 2002). El primer factor agrupó 9 ítems relacionados con lo que se denominó Ayuda: 1, 2, 6, 7, 8, 9, 10, 11 y 13. Estos ítems pertenecían anteriormente a las dimensiones Ayuda y Altruismo. El segundo factor reunió 7 ítems que implican Confortar al otro y esa fue su denominación. Los ítems correspondientes a la dimensión Confortar fueron: 3, 4, 5, 12, 14, 15 y 16.

La Tabla 1 muestra las matrices de configuración y de estructura de este modelo de 16 ítems. Son dos matrices rotadas. Como se puede observar dicha Tabla, interpretando ambas matrices, tanto las cargas factoriales como las correlaciones fueron las mayores en la dimensión correspondiente al ítem. En consecuencia, se puede inferir un adecuado ajuste de la escala al modelo bifactorial (Vélez & Martínez Lugo, 1995).

Tabla 1

*Matrices de Configuración y de Estructura de los ítems conservados*

*luego del AFE*

Ítems	M. de Configuración		M. de Estructura	
	1° f	2° f	1° f	2° f
1.- Dedico tiempo para facilitar la cursada a mis compañeros.	<b>.36</b>	.05	<b>.39</b>	.23
2.- Si tuviera un único plato de comida, lo compartiría con un desconocido.	<b>.49</b>	.07	<b>.53</b>	.32
3.- Cuando siento que alguien está mal, le demuestro que lo entiendo.	-.03	<b>.69</b>	.32	<b>.67</b>
4.- Actúo como bastón de los demás.	.20	<b>.45</b>	.42	<b>.55</b>
5.- Siento todo dolor ajeno como propio.	.17	<b>.37</b>	.35	<b>.46</b>
6.- Me comprometo con causas nobles.	<b>.70</b>	.06	<b>.73</b>	.41
7.- Dedico un tiempo importante de mi vida a mejorar el mundo.	<b>.70</b>	-.03	<b>.69</b>	.32
8.- Expreso abiertamente mi apoyo a minorías discriminadas aunque no pertenezca a las mismas.	<b>.54</b>	.10	<b>.59</b>	.37
9.- Doy dinero a cualquiera que lo necesite.	<b>.55</b>	-.07	<b>.52</b>	.20
10.- Me ocupo del bienestar de cualquier individuo, grupo o comunidad.	<b>.66</b>	.08	<b>.70</b>	.41
11.- Me quedo con lo justo y necesario para vivir y reparto todo lo demás.	<b>.60</b>	-.12	<b>.54</b>	.18
12.- Les digo a los demás lo mucho que valen.	.04	<b>.68</b>	.39	<b>.70</b>
13.- Hago donaciones a organizaciones benéficas.	<b>.52</b>	-.06	<b>.49</b>	.20

14.- Comparto las malas experiencias de la vida con amigos para que no cometan los mismos errores.	-.05	<b>.56</b>	.24	<b>.54</b>
15.- Intento “levantar” la autoestima a mis amigos.	-.13	<b>.87</b>	.31	<b>.81</b>
16.- Acompaño a conocidos en momentos dolorosos.	-.01	<b>.66</b>	.33	<b>.66</b>

Nota. 1º f: Ayuda - 2º f: Confortar

Los estadísticos descriptivos para los ítems conservados se pueden observar en la Tabla 2:

Tabla 2  
*Estadísticos Descriptivos de los Ítems Conservados  
luego del AFE*

Ítems	M	DE	As	K
1.- Dedico tiempo para facilitar la cursada a mis compañeros.	3.0	1.1	0.3	0.1
2.- Si tuviera un único plato de comida, lo compartiría con un desconocido.	3.5	1.6	0.2	-1.1
3.- Cuando siento que alguien está mal, le demuestro que lo entiendo.	4.9	1.1	-0.7	-0.5
4.- Actúo como bastón de los demás.	3.9	1.3	-0.1	-0.4
5.- Siento todo dolor ajeno como propio.	3.4	1.3	0.3	-0.4
6.- Me comprometo con causas nobles.	3.6	1.3	0.2	-0.5
7.- Dedico un tiempo importante de mi vida a mejorar el mundo.	2.9	1.3	0.7	0.0
8.- Expreso abiertamente mi apoyo a minorías discriminadas, aunque no pertenezca a las mismas.	3.9	1.5	-0.1	-1.0
9.- Doy dinero a cualquiera que lo necesite.	2.9	1.2	0.6	0.0
10.- Me ocupo del bienestar de cualquier individuo, grupo o comunidad.	2.9	1.2	0.5	0.0
11.- Me quedo con lo justo y necesario para vivir y reparto todo lo demás.	2.2	1.1	1.0	0.8
12.- Les digo a los demás lo mucho que valen.	4.0	1.4	-0.1	-0.9
13.- Hago donaciones a organizaciones benéficas.	2.4	1.3	0.8	0.1
14.- Comparto las malas experiencias de la vida con amigos para que no cometan los mismos errores.	4.4	1.3	-0.4	-0.6
15.- Intento “levantar” la autoestima a mis amigos.	5.1	1.1	-0.9	-0.3
16.- Acompaño a conocidos en momentos dolorosos.	4.8	1.2	-0.8	-0.3

Nota. M: Media aritmética - DE: Desviación Estándar - As: índice de Asimetría - K: índice de curtosis

### *Procedimiento del Análisis Factorial Confirmatorio*

Como se indicó previamente las respuestas de 622 participantes, seleccionados en forma aleatoria de la muestra, se utilizaron para realizar un Análisis Factorial Confirmatorio (AFC) del modelo bifactorial de 16 ítems descrito en la Tabla 1. El AFC

se implementó mediante el programa Mplus (Muthén & Muthén, 2010). Para respetar el carácter ordinal de los datos el análisis se realizó en base a las matrices de correlaciones policóricas. Se estimaron los parámetros con el método de mínimos cuadrados ponderados robustos (*Weighted Least Squares Mean and Variance Adjusted*, WLSMV). Siguiendo las recomendaciones de Abad, Olea, Ponsoda y García (2011), el ajuste del modelo se examinó mediante el índice de ajuste comparativo (*Comparative Fit Index*, CFI), el índice de Tucker-Lewis (*Tucker-Lewis Index*, TLI) y el error medio cuadrático de aproximación (*Root Mean Square Error of Approximation*, RMSEA). Estos resultan los índices de bondad de ajuste más utilizados al aplicar AFC con datos categóricos (Byrne, 2012). Si bien no existe un consenso acabado sobre los puntos de cortes de estos índices de ajuste es posible aceptar razonablemente el ajuste del modelo para CFI y TLI  $\geq .90$  y RMSEA  $\leq .08$  (Abad, et al., 2011; Brown, 2015; Byrne, 2012).

Se sometió a prueba la estructura obtenida con el AFE, es decir el modelo de dos factores correlacionados (se encontró en el AFE una relación moderada), pero también se propusieron estructuras alternativas que podrían eventualmente resultar interpretables. Se examinó el ajuste de los datos a los siguientes modelos alternativos al hallado en el AFE: a) un modelo unidimensional que establece que los ítems se relacionan directamente con el constructo global conducta prosocial, b) un modelo en el que los dos factores aislados (Ayuda y Confortar) son incorrelados y c) un modelo jerárquico en el que los dos factores Ayuda y Confortar de primer orden definen un factor de segundo orden.

### *Resultados del Análisis Factorial Confirmatorio*

La Tabla 3 exhibe los índices de ajuste obtenidos en los AFC para los cuatro modelos propuestos. Los resultados reflejan que ni el modelo unidimensional ni el modelo de dos factores incorrelados se ajustaron a los datos (TLI y CFI  $< .90$ ; RMSEA  $> .08$ ). Aun cuando el ajuste del modelo que incluye un factor de segundo orden es idéntico al ajuste del modelo con dos factores correlacionados, el modelo jerárquico presenta una estimación de parámetros con valores inaceptables, "casos *Heywood*" (Brown, 2015). En este caso, la varianza residual del factor Ayuda adoptó un valor negativo durante la estimación, lo que resulta inadmisibles. Este indicador puede estar reflejando que los factores de primer orden presentan cierto grado de heterogeneidad que no está en consonancia con la hipótesis de un factor de segundo orden (Geiser, 2013). Por lo tanto,

se descarta también la aplicación del modelo jerárquico, y en consecuencia, se considera más adecuado el modelo de dos factores correlacionados.

Finalmente, el modelo que contempla los dos factores correlacionados es el único que mostró indicadores de ajuste adecuados y parámetros estimados significativos al 5%. Los factores Ayuda y Confortar presentaron entre sí una relación moderada ( $r = .63$ ;  $p = <.0001$ ).

Tabla 3

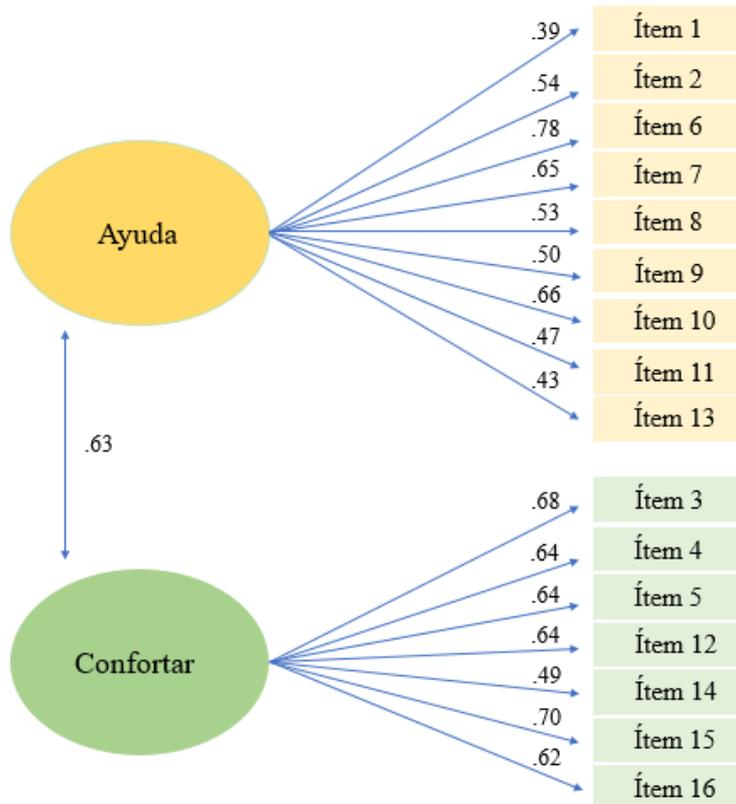
*Indicadores de ajuste del Análisis Factorial Confirmatorio*

	CFI	TLI	RMSEA (IC 90%)
Dos dimensiones correlacionadas	.921	.908	.070 (.063-.077)
Unifactorial	.804	.774	.110 (.103-.117)
Dos dimensiones incorreladas	.610	.550	.155 (.149-.162)
Jerárquico	.921	.908	.0720 (.063-.077)

En la Figura 7 se representan los pesos estandarizados de cada uno de los ítems en sus respectivos factores junto con los valores de interrelación entre las variables.

Figura 7

*Modelo Factorial Confirmatorio de la Escala de Conducta Prosocial*



*Comentarios de la Fase II*

Los Análisis Factoriales Exploratorio y Confirmatorio permitieron determinar una Estructura Bifactorial para la escala de Conducta Prosocial. Los factores hallados tienen interpretación psicológica y son coherentes con la definición del constructo, por lo tanto, la estructura hallada es una fuente de validez interna.

Para que un test sea válido, debe comprobarse que no favorezca a un grupo en forma artificial, es decir por factores extraños al constructo medido. Es por esto que en el siguiente apartado se realiza el estudio del Funcionamiento Diferencial del Ítem según el género.

### 6.3. Fase III: Estudios de Funcionamiento Diferencial del Ítem Según el Género

#### *Justificación de la Elección del Género para el estudio del DIF*

La evaluación en conducta prosocial, según algunos autores, incluye un número desproporcionado de ítems sesgados a favor del género femenino (Pursell, Laursen, Rubin, Booth-LaForce, & Rose-Krasnor, 2008). Por lo tanto, se hace discutible si a nivel global existe un género más prosocial que otro. Es por esto que se ha decidido incluir un análisis del funcionamiento diferencial de los ítems para determinar si la diferencias intergéneros que se estudiarán en un apartado posterior son genuinas o generadas artificialmente por el instrumento.

#### *Procedimiento*

Este análisis del DIF según el género se realizó por separado para cada dimensión. Si bien en esta escala de conducta prosocial no hay respuestas correctas, sí las respuestas están en la dirección de un mayor o menor nivel del rasgo. Éstas se dicotomizaron asignando 1 y 0 respectivamente y así se pudieron utilizar los modelos del análisis del DIF para ítems dicotómicos. El 1 indica entonces el conjunto de respuestas asociadas con los mayores niveles en el rasgo y el 0 las asociadas a los menores niveles (Galibert, Aguerri, Lozzia, & Abal, 2006).

Se calcularon mediante Excel los siguientes estadísticos e indicadores necesarios para el estudio del DIF: Estadístico  $\chi^2$  de la prueba Mantel-Haenszel y su  $p$  valor asociado, cociente de razones común MH ( $\hat{\alpha}_{MH}$ ), la Diferencia de Proporciones Estandarizada (DPE) y el Delta-DIF de Mantel-Haenszel (MH D-DIF). Se usó el método empírico como criterio de eliminación del ítem. Es decir, la detección de presencia de DIF severo según lo planteado por Ferreres Traver et al. (2002), o sea cuando el valor absoluto de MH D-DIF supera 1.5.

## Participantes

Se consideró el total depurado de participantes de la muestra ( $n = 1383$ ) menos dos individuos que reportaron un género diferente al femenino o masculino. Se utilizó por lo tanto  $n = 1381$ .

## Resultados del Estudio del DIF en la Dimensión Ayuda

Según la evidencia muestral, existiría un posible DIF según el género en los ítems 2 (*Si tuviera un único plato de comida, lo compartiría con un desconocido.*), 7 (*Dedico un tiempo importante de mi vida a mejorar el mundo*), 9 (*Doy dinero a cualquiera que lo necesite*) y 13 (*Hago donaciones a organizaciones benéficas*). El resto de los ítems de la dimensión Ayuda estarían libres de DIF, tal como se muestra en la Tabla 4.

Tabla 4

### Estadísticos e Indicadores de DIF en la dimensión Ayuda

Ítem	MHn	$p_{MHn}$	$\hat{\alpha}_{MH}$	MH D-DIF	DPE
1	0.25	.62	0.91	0.23	0.02
<b>2</b>	<b>7.99</b>	<b>.004**</b>	<b>1.67</b>	<b>-1.20</b>	<b>-0.08</b>
6	0.29	.58	0.88	0.29	0.01
<b>7</b>	<b>4.22</b>	<b>.04*</b>	<b>0.64</b>	<b>1.03</b>	<b>0.05</b>
8	1.19	.27	0.81	0.48	0.03
9	3.33	.07	0.70	0.82	0.05
10	0.24	.62	1.12	-0.27	-0.01
11	1.98	.16	0.68	0.90	0.03
<b>13</b>	<b>16.21</b>	<b>.00005**</b>	<b>3.07</b>	<b>-2.64</b>	<b>-0.09</b>

Nota: MHn = estadístico  $\chi^2$  Mantel-Haenszel -  $p_{MHn}$  = valor  $p$  asociado al estadístico  $\chi^2$  Mantel-Haenszel -  $\hat{\alpha}_{MH}$  = cociente de razones común MH - MH D-DIF = Delta-DIF de Mantel-Haensze - DPE = Diferencia de Proporciones Estandarizada.

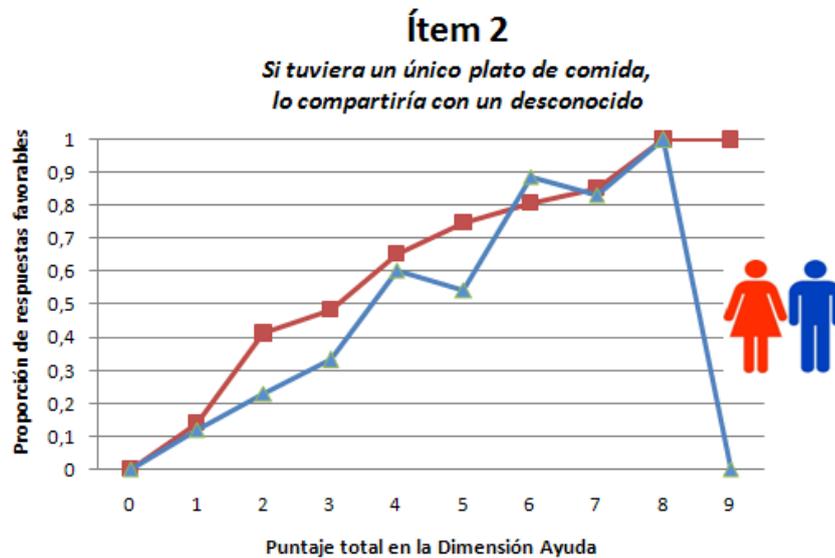
[\*\*]  $p < .01$ ; [\*]  $.01 \leq p < .05$

En el caso del ítem 2 se rechaza la hipótesis nula de ausencia de DIF de la prueba de Mantel-Haenszel con un  $p$  de 0.004. Según el criterio empírico de Ferreres Traver et al. (2002) no presenta DIF severo y por lo tanto puede conservarse ya que su MH D DIF es -1.20, menor que 1.5 en valor absoluto. Se puede observar las mujeres presenta más

chance que los varones de estar de acuerdo con el ítem cuando presentan puntajes entre 2 y el 5, posteriormente las curvas se cruzan y en el puntaje total 6 el ítem favorece levemente a los varones. Para los puntajes totales 7 y 8 la proporción de respuestas favorables de varones y mujeres se igualan. Finalmente, en el puntaje 9, el ítem vuelve a favorecer a las mujeres.

Figura 8

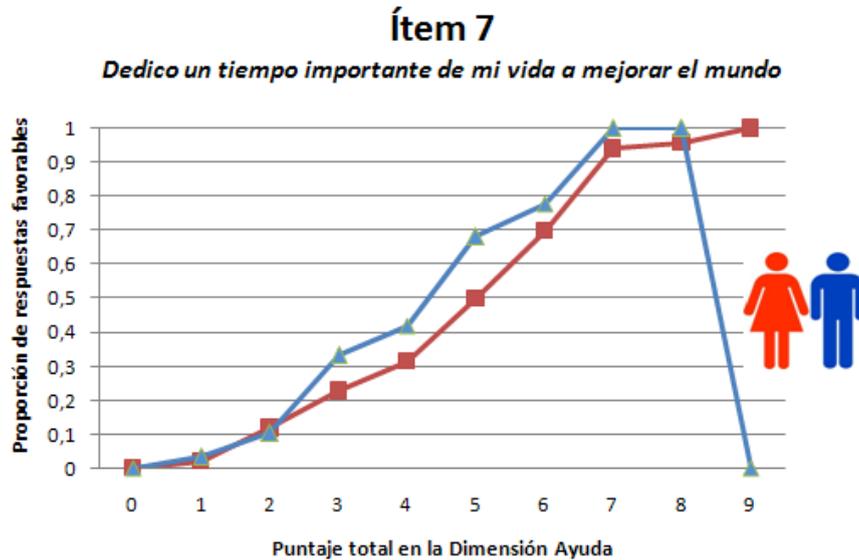
*Proporción de respuestas favorables de mujeres y varones según nivel de puntaje total del ítem 2*



En el caso del ítem 7 se rechaza la hipótesis nula de ausencia de DIF de la prueba de Mantel-Haenszel con un  $p$  de 0.04. Según el criterio empírico de Ferreres Traver et al. (2002) no presenta DIF severo y por lo tanto puede conservarse ya que su MH D DIF es 1.03, menor que 1.5. Se puede observar que las proporciones de respuestas favorables son iguales para ambos géneros para los puntajes totales desde 0 hasta 2. El ítem favorece a los varones desde el puntaje total 3 hasta el 8, posteriormente las curvas se cruzan y en el puntaje total 6 el ítem favorece levemente a los varones. En el puntaje 9, el ítem vuelve a favorecer a las mujeres.

Figura 9

*Proporción de respuestas favorables de mujeres y varones según nivel de puntaje total del ítem 7*



En el caso del ítem 13 se ha rechazado la hipótesis nula de ausencia de DIF en la prueba  $\chi^2$  de Mantel Haenszel con  $p = .00005$ . Según el criterio empírico de Ferreres Traver et al. (2002) presenta DIF severo por ser su MH D-DIF -2.64, en valor absoluto mucho mayor a 1.5 y por lo tanto debe eliminarse de la Escala. Este DIF es a favor de las mujeres quienes tienen mayor proporción de respuestas favorables al ítem a lo largo de casi todos los puntajes totales de la dimensión Ayuda, excepto para los puntajes totales 0 donde la proporción de respuestas favorables para varones y mujeres es igual y para el puntaje 8 donde es mayor la proporción de varones. Gráficamente se puede observar una mayor distancia entre las curvas características empíricas con respecto a las de los ítems 2 y 7.

Figura 10

Proporción de respuestas favorables de mujeres y varones según nivel de puntaje total del ítem 13



#### Análisis del DIF de la Dimensión Confortar

Según la evidencia muestral, existiría un posible DIF según el género sólo en el ítem 12 (*Les digo a los demás lo mucho que valen*). El resto de los ítems de la dimensión Confortar estarían libres de DIF.

Tabla 5

#### Estadísticos e Indicadores de DIF en la dimensión Confortar

Item	MHn	pMHn	$\hat{\alpha}_{MH}$	MH D-DIF	DPE
3	0.54	.46	0.81	0.47	0.02
4	0.09	.75	0.83	0.17	0.01
5	2.33	.13	1.37	-0.74	-0.04
<b>12</b>	<b>6.67</b>	<b>.009**</b>	<b>1.66</b>	<b>-1.19</b>	<b>-0.07</b>
14	3.55	.06	0.71	0.81	0.06
15	0.54	.46	1.25	-0.53	-0.02
16	3.19	.07	0.64	1.03	0.04

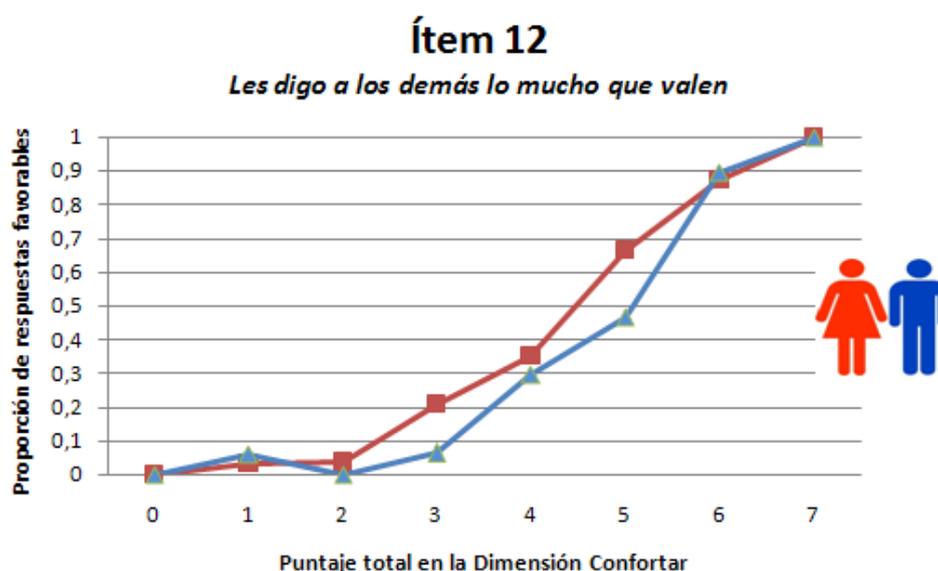
Nota: MHn = estadístico  $\chi^2$  Mantel-Haenszel - pMHn = valor p asociado al estadístico  $\chi^2$  Mantel-Haenszel -  $\hat{\alpha}_{MH}$  = cociente de razones común MH - MH D-DIF = Delta-DIF de Mantel-Haenszel - DPE = Diferencia de Proporciones Estandarizadas.

[\*\*]  $p < .01$

En el ítem 12 se ha rechazado la hipótesis nula de ausencia de DIF en la prueba  $\chi^2$  de Mantel Haenszel con  $p = .009$ . Según el criterio empírico de Ferreres Traver et al. (2002) no presenta DIF severo y por lo tanto puede conservarse ya que su MH D DIF es -1.19, menor que 1.5 en valor absoluto. Este posible DIF sería a favor de las mujeres quienes tienen mayor proporción de respuestas favorables al rasgo en el ítem entre los puntajes totales 2 al 5. Para los puntajes 0, 1, 6 y 7 la proporción de respuestas favorables al rasgo en el ítem es similar para varones y mujeres.

Figura 11

*Proporción de respuestas favorables de mujeres y varones según nivel de puntaje total del ítem 12*



### *Comentarios de la Fase III*

En esta Fase se ha eliminado un Ítem que presentaba DIF severo según el género. La existencia de DIF atenta contra la unidimensionalidad, siendo ésta un supuesto de los modelos de la TRI que se van a utilizar. Al asegurarse que el resto de los ítems no presentan DIF severo, se ha dado una nueva evidencia de validez de las puntuaciones del instrumento a través de una Fuente Interna.

En los tres ítems analizados de la dimensión Ayuda se evidencia una caída abrupta en la proporción de ítems con puntaje total 9. Esta caída se debe a que ningún varón tuvo

este puntaje total. Esto podría deberse a la presencia de impacto, es decir, de diferencias reales existentes entre varones y mujeres con respecto a la variable. También la cantidad de varones fue 267 y la de mujeres 1114. Siendo el puntaje 9 una puntuación extrema, hay menos chances de encontrar personas con puntuaciones extremas con un tamaño de muestra mucho menor.

#### *6.4. Fase IV Modelización de la Escala de Conducta Prosocial Con la Teoría de la Respuesta al Ítem*

##### *Participantes*

La Fase IV se realizó con la totalidad de participantes de la muestra (n =1383).

##### *Procedimiento General*

Whittaker y Worthington (2016) recomiendan aplicar diferentes modelos a los datos y evaluar las estimaciones de parámetros por su viabilidad y significación a lo largo de los modelos comparados. En línea con esta sugerencia, se aplicó el ML1P, ML2P, ML3P, MRG y MCP a las dimensiones Ayuda y Confortar mediante el programa MULTILOG. Previamente a la aplicación de los modelos se verificó el supuesto de unidimensionalidad para cada dimensión mediante Análisis Factorial Confirmatorio.

Los parámetros de los modelos fueron estimados mediante Máxima Verosimilitud Marginal. El programa asume por *default* que el nivel en el rasgo se distribuye normal estándar y es por esto que tanto los gráficos como el nivel de información se detallan para el rango -3 a 3.

Para evaluar el ajuste de los datos a los modelos se observó especialmente la similitud entre las proporciones observadas y las esperadas por el modelo para cada categoría de respuesta (Revuelta et al., 2006; Thissen, 1991). En ML1P, ML2P, ML3P se realizó una prueba de hipótesis bilateral al 5% de diferencia de proporciones de respuestas favorables al rasgo entre las obtenidas empíricamente (PO) y las esperadas (PE) según dichos modelos. En MRG y MCP dichas proporciones fueron transformadas a frecuencias absolutas (n = 1383) para poder aplicar la prueba ji cuadrado donde su fórmula es:

$$\chi^2 = \sum \left[ \frac{(f_0 - f_e)^2}{f_e} \right]$$

Otros indicadores de ajuste fueron la efectiva convergencia de estimación de los parámetros y la existencia de errores estándar relativamente bajos.

Se analizaron los parámetros obtenidos para cada uno de los modelos. Se graficaron las CCI para los modelos dicotómicos y las CCCRI para los modelos politómicos y las Funciones de Información para cada ítem y para el test.

### *Resultados*

#### Aplicación de los modelos de la TRI a la Dimensión Ayuda

##### *Aplicación de los Modelos Logísticos de Uno, Dos y Tres Parámetros a la Dimensión Ayuda*

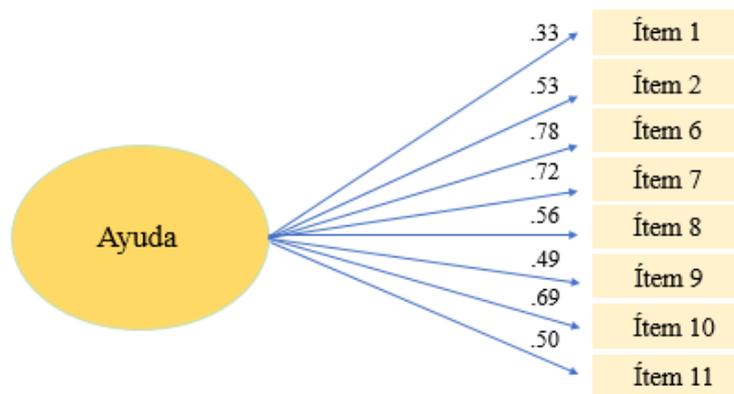
##### *Cumplimiento de los supuestos de los Modelos Logísticos de Uno, Dos y Tres Parámetros en la Dimensión Ayuda*

Se implementó un AFC con el programa Mplus (Muthén & Muthén, 2010) sobre los ocho ítems que componen la dimensión Ayuda para verificar el supuesto de unidimensionalidad requerido por los modelos de la TRI. Las respuestas fueron previamente dicotomizadas. Se respetó el carácter categórico/binario de los datos por lo que se estimaron los parámetros con el método WLSMV sobre la base de las matrices de correlaciones tetracóricas. Se analizó el ajuste al modelo considerando los índices CFI, TLI y RMSEA. Este análisis fue realizado sobre la totalidad de los participantes de la muestra (n = 1383)

Los indicadores de ajuste comparativos muestran un resultado muy adecuado para la verificación del modelo unidimensional (CFI = .983; TLI = .976; RMSEA = .032, IC 90% .020-.043) por lo cual se acepta el cumplimiento del supuesto de unidimensionalidad para la dimensión Ayuda. Tal como se observa en la Figura 12, las cargas factoriales son elevadas (Abad et al., 2011).

Figura 12

*Modelo Factorial Confirmatorio de la Dimensión Ayuda*



El supuesto de independencia local se dio por satisfecho al comprobarse la unidimensionalidad de la escala (Lord & Novick, 1968).

*Ajuste de los datos a los Modelos Logísticos de Uno, Dos y Tres Parámetros en la Dimensión Ayuda*

Para evaluar el ajuste de los datos a los modelos logísticos, se realizó una prueba de hipótesis bilateral al 5% de diferencia de proporciones de respuestas favorables al rasgo entre las obtenidas empíricamente (PO) y las esperadas (PE) según dichos modelos. Los valores  $p$  fueron en su totalidad mayores a .44 indicando el no rechazo de la hipótesis nula de igualdad de proporciones observadas y esperadas, es decir, un adecuado ajuste a los modelos ML1P, ML2P y ML3P.

Tabla 6

*Índices de ajuste de los ítems de la dimensión Ayuda  
a los modelos ML1P, ML2P y ML3P*

Ítem	PO	ML1P		ML2P		ML3P	
		PE <sub>ML1P</sub>	<i>p</i>	PE <sub>ML2P</sub>	<i>p</i>	PE <sub>ML3P</sub>	<i>p</i>
1	.2813	.2818	.4836	.2814	.4968	.2814	.4968
2	.4136	.4141	.4852	.4142	.4820	.4136	.5000
6	.4432	.4439	.4792	.4452	.4407	.4435	.4912
7	.2473	.2477	.4864	.2486	.4558	.2479	.4796
8	.5785	.5788	.4912	.5789	.4880	.5783	.5059
9	.2393	.2398	.4828	.2397	.4860	.2395	.4932
10	.2690	.2695	.4832	.2702	.4601	.2697	.4768
11	.0947	.0947	.5000	.0947	.5000	.0948	.4952

*Estimación de Parámetros de los Modelos Logísticos de Uno, Dos y Tres Parámetros  
para los ítems de la Dimensión Ayuda*

Los parámetros de cada ítem fueron estimados, para ML1P, ML2P y ML3P, con un criterio de convergencia de .001. En la Tabla 9 se exponen la cantidad de ciclos necesarios para alcanzar la convergencia *-finished cycle-* así como la máxima variación interciclo de las estimaciones *-maximum intercycle parameter change-* para cada uno de dichos modelos. El ML3P fue el modelo que requirió más ciclos para alcanzar su convergencia, lo cual es coherente con la cantidad de parámetros a estimar.

Tabla 7

*Procesos de estimación de los ítems de la Dimensión Ayuda con los los modelos ML1P,  
ML2P Y ML3P*

	ML1P	ML2P	ML3P
Cantidad de parámetros estimados	9	16	24
Cantidad de ciclos	17	19	100
Mayor cambio registrado en la estimación	0.00056	0.00069	0.00190

En la calibración de los ítems del ML1P se obtuvieron ocho parámetros de localización  $b$ , uno por cada ítem, que expresan el mínimo nivel de  $\theta$  requerido para tener una probabilidad de .50 o mayor de dar una respuesta que implique Ayuda. El parámetro  $b$  promedio fue de 0.87 (DE = 0.79) siendo el ítem 8 (*Expreso abiertamente mi apoyo a minorías discriminadas, aunque no pertenezca a las mismas*) aquél con el mínimo valor, y el ítem 11 (*Me quedo con lo justo y necesario para vivir y reparto todo lo demás*), el reactivo con el máximo  $b$ . Además, se obtuvo un único parámetro  $a$ , constante para los ocho ítems cuyo valor fue de 1.21 con una DE de 0.04.

En el ML2P el promedio de los  $b$  fue de 0.96 (DE = 0.91) con un mínimo en el ítem 8 (*Expreso abiertamente mi apoyo a minorías discriminadas, aunque no pertenezca a las mismas*) y un máximo valor en el ítem 11 (*Me quedo con lo justo y necesario para vivir y reparto todo lo demás*). El parámetro  $a$  indica en el entorno de  $b$  dónde el ítem posee mayor potencia discriminatoria con respecto al nivel de Ayuda de los sujetos. El valor promedio de los parámetros  $a$  en el ML2P fue de 1.29 (DE = 0.52) con un mínimo en el ítem 1 (*Dedico tiempo para facilitar la cursada a mis compañeros*) y un máximo valor en el ítem 6 (*Me comprometo con causas nobles*).

En el ML3P se estima además de  $a$  y  $b$ , el parámetro  $c$ . Es decir que en el ML3P se estiman tres parámetros por ítem. La utilidad del parámetro  $c$  es incierta en test de comportamiento típico. MULTILOG no informa acerca de los errores de estimación de los parámetros de este modelo. Los parámetros  $c$  fueron muy cercanos a cero, con un valor medio de 0.06 (DE 0.06). Para el ML3P el parámetro  $a$  promedio estimado fue de 0.93 (DE = 0.42) y el parámetro  $b$  promedio estimado de 1.06 (DE = 0.84).

En la Tabla 8 se encuentran los parámetros estimados obtenidos para los modelos en los ítems de la dimensión Ayuda.

Tabla 8

*Parámetros estimados los Modelos Logísticos de Uno, Dos y Tres Parámetros  
para los ítems de la Dimensión Ayuda*

Ítem	ML1P	ML2P		ML3P		
	<i>b</i> (DE)	<i>b</i> (DE)	<i>a</i> (DE)	<i>b</i>	<i>a</i>	<i>c</i>
1	0.99 (0.07)	1.70 (0.25)	0.59 (0.08)	1.85	0.63	0.14
2	0.37 (0.07)	0.41 (0.08)	1.04 (0.10)	0.41	0.62	0.00
6	0.24 (0.07)	0.18 (0.04)	2.17 (0.16)	0.34	1.82	0.10
7	1.17 (0.08)	0.94 (0.07)	1.79 (0.14)	0.94	1.05	0.00
8	-0.34 (0.07)	-0.36 (0.07)	1.13 (0.10)	-0.01	0.83	0.15
9	1.21 (0.08)	1.47 (0.16)	0.92 (0.10)	1.51	0.63	0.04
10	1.05 (0.08)	0.89 (0.07)	1.62 (0.13)	0.95	1.22	0.05
11	2.30 (0.12)	2.47 (0.26)	1.09 (0.15)	2.48	0.64	0.00
Media	0.87	0.96	1.29	1.06	0.93	0.06
DE	0.79	0.91	0.52	0.84	0.42	0.06
Mínimo	-0.34	-0.36	0.59	-0.01	0.62	0.00
Máximo	2.30	2.47	2.17	2.48	1.82	0.15

Los *b* promedio estimados son positivos para los tres modelos (excepto para el ítem 8, *Expreso abiertamente mi apoyo a minorías discriminadas, aunque no pertenezca a las mismas*), indicando que se encuentran en valores relativamente elevados de  $\theta$ . Es necesario un nivel medio o alto en el rasgo para responder los ítems de la dimensión Ayuda de manera favorable. Sin embargo, existe una importante dispersión promedio entre los *b*. La elevada dispersión indicaría que, a pesar de ser los *b* positivos, esta

dimensión tendría capacidad discriminativa en los diferentes niveles del rasgo con una tendencia a discriminar mejor en los niveles moderados y altos del mismo.

Los valores estimados promedio del parámetro  $a$  para los tres modelos muestran una capacidad discriminativa alta según el criterio de Reise y Waller (1990), con mayor capacidad discriminativa estimada en el ML2P y la menor en el ML3P.

La cercanía a cero de los parámetros  $c$  indicaría la poca relevancia de su inclusión dado que se trata de una escala de comportamiento típico.

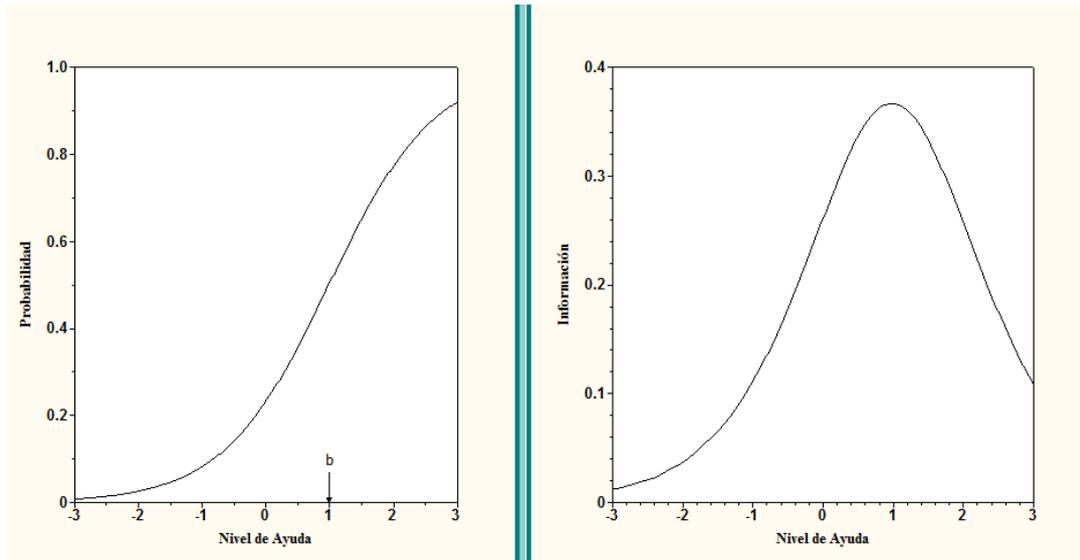
Figura 13

CCIs y FIs de los ítems de la dimensión Ayuda para ML1P, ML2P y ML3P

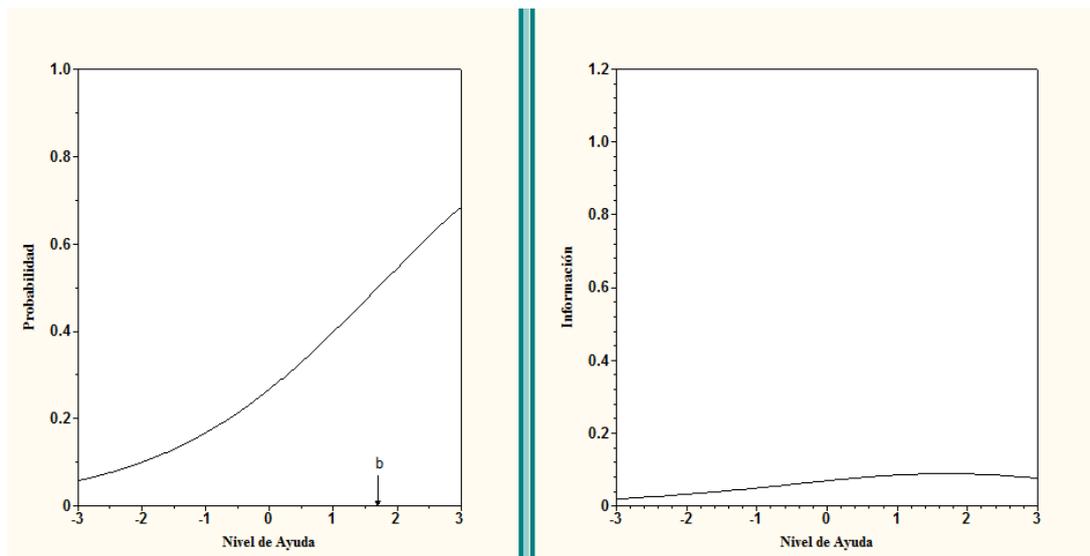
CCI

FI

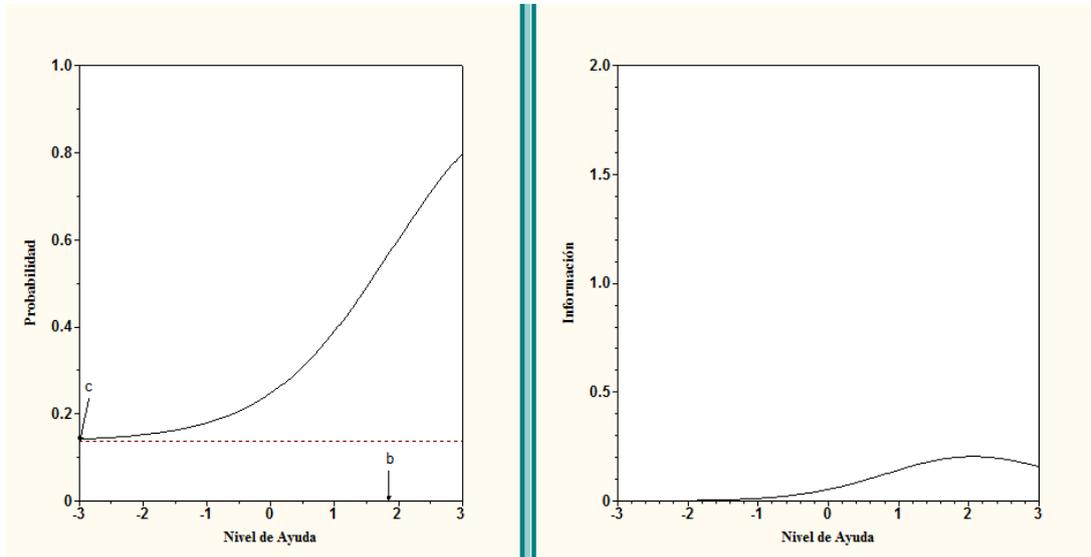
ML1P Ítem 1



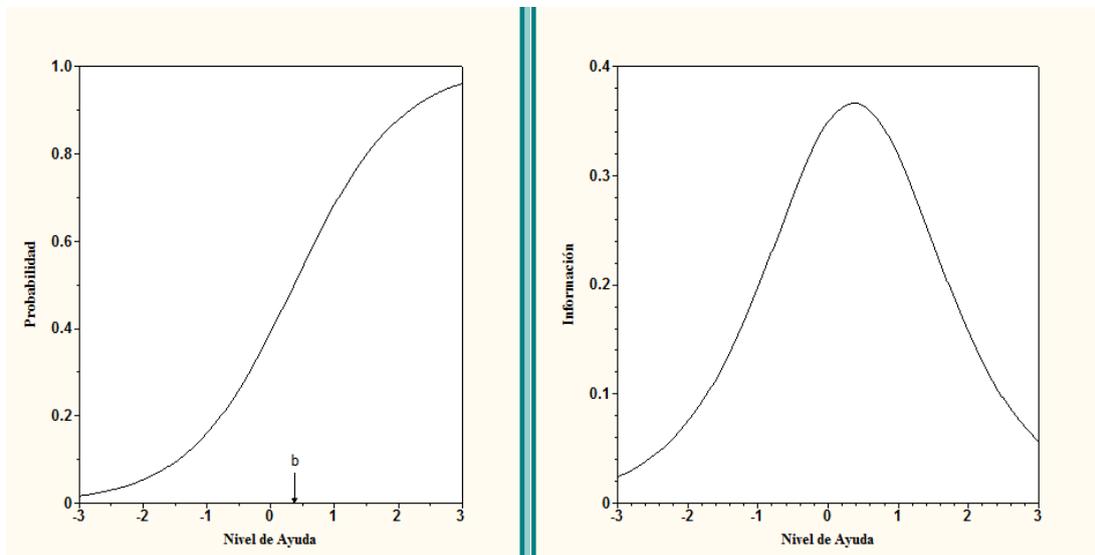
ML2P Ítem 1



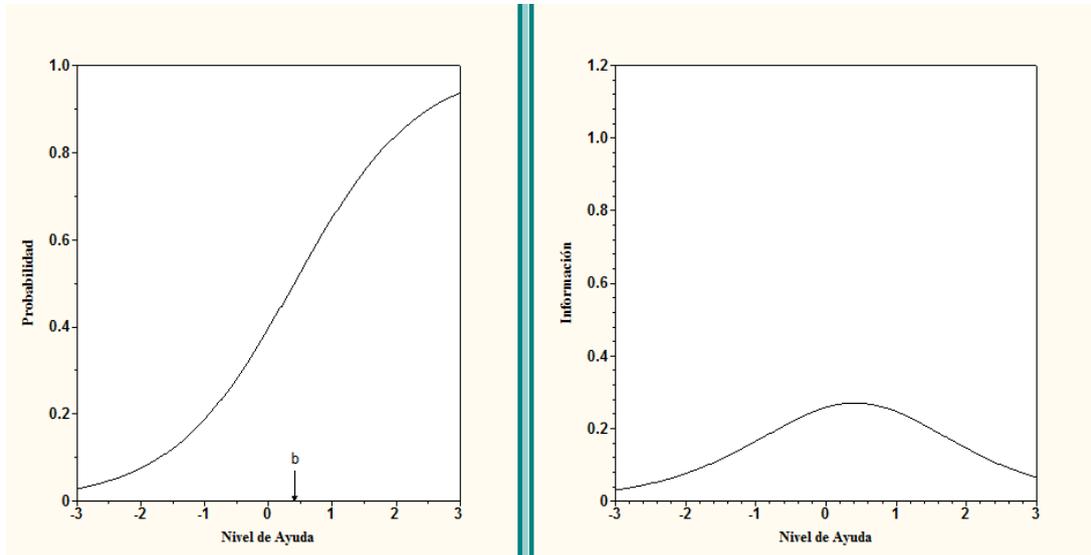
## ML3P Ítem 1



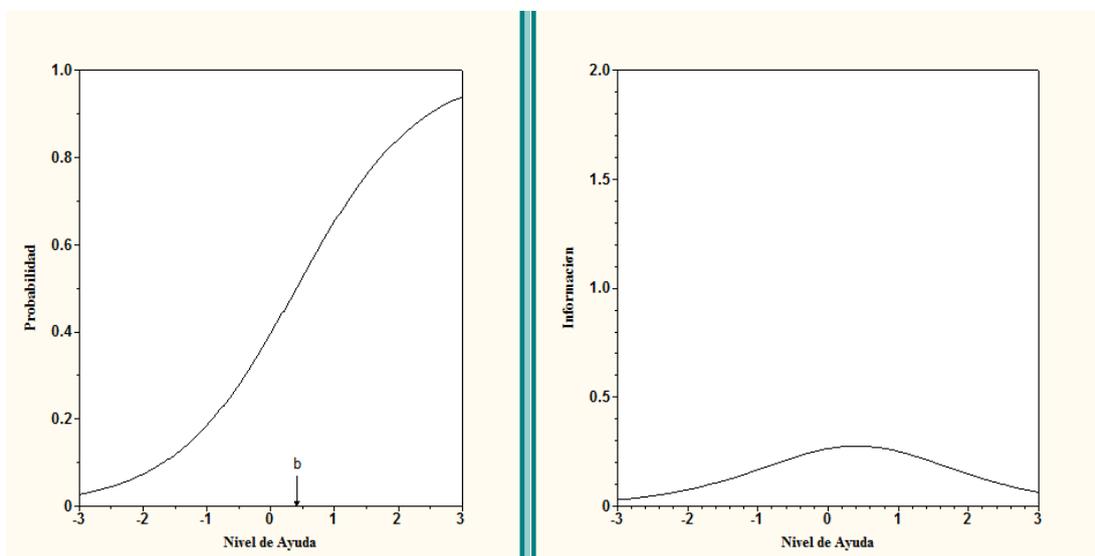
## ML1P Ítem 2



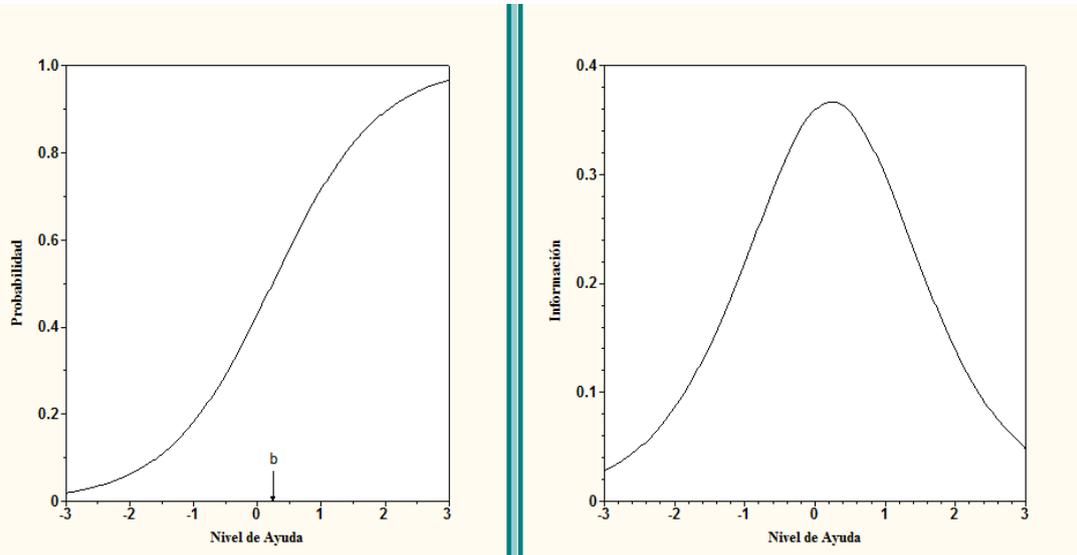
## ML2P Ítem 2



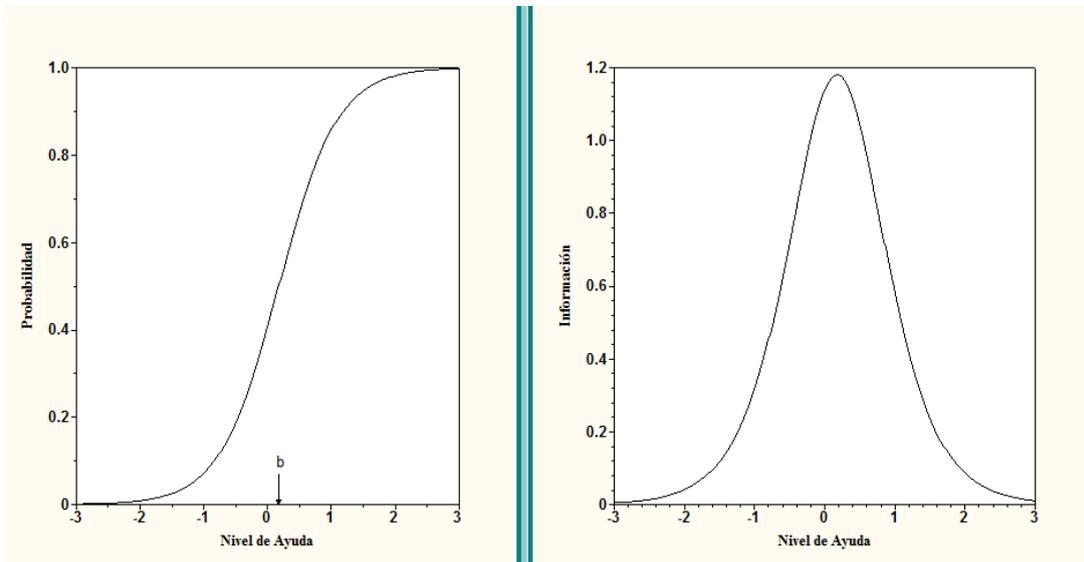
## ML3P Ítem 2



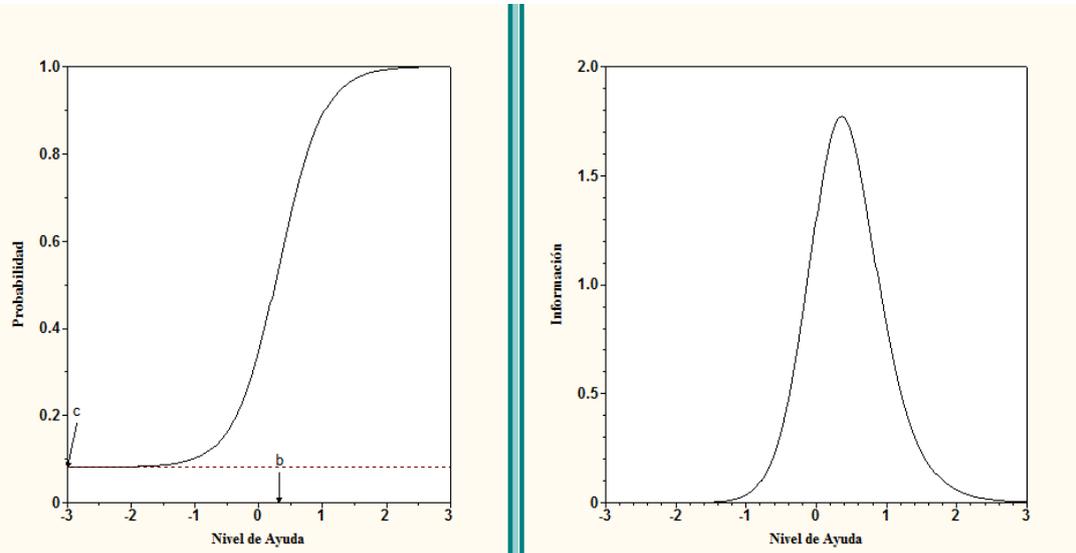
## ML1P Ítem 6



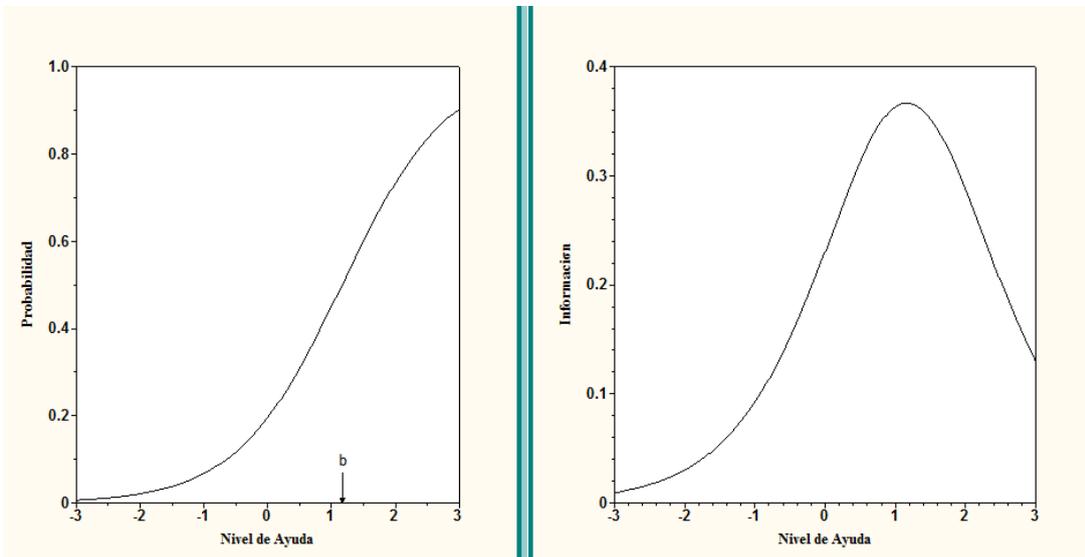
## ML2P Ítem 6



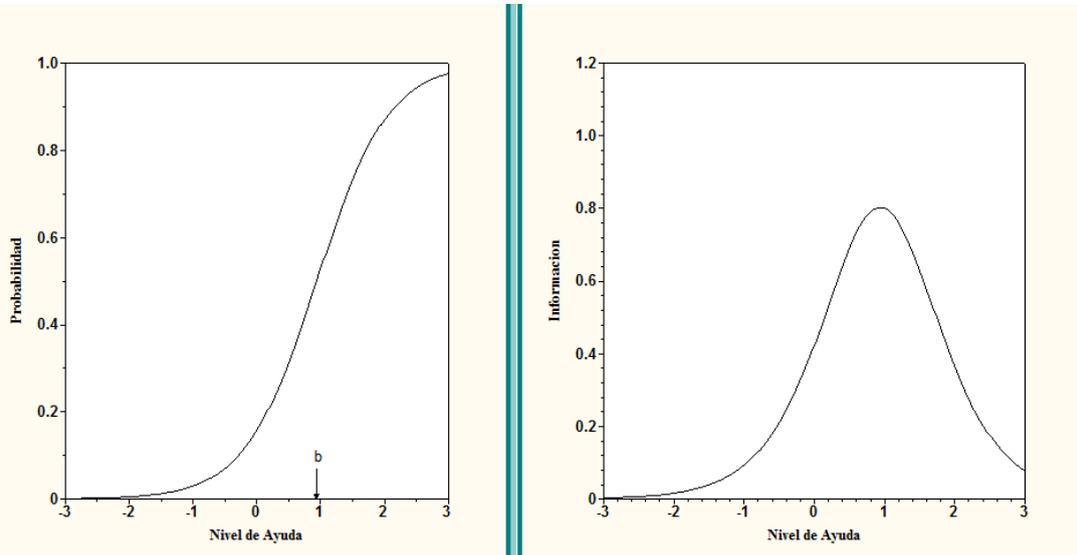
## ML3P Ítem 6



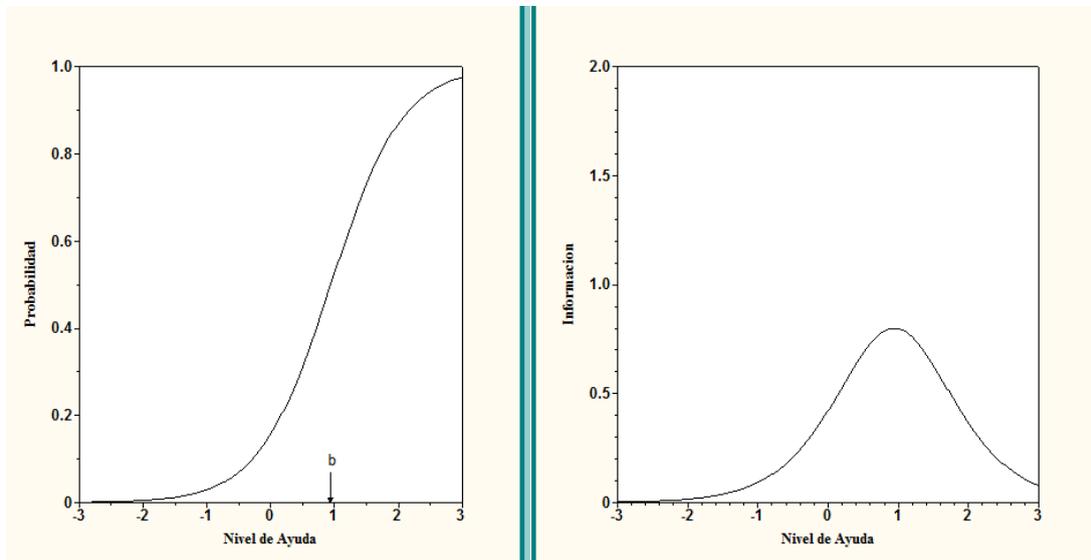
## ML1P Ítem 7



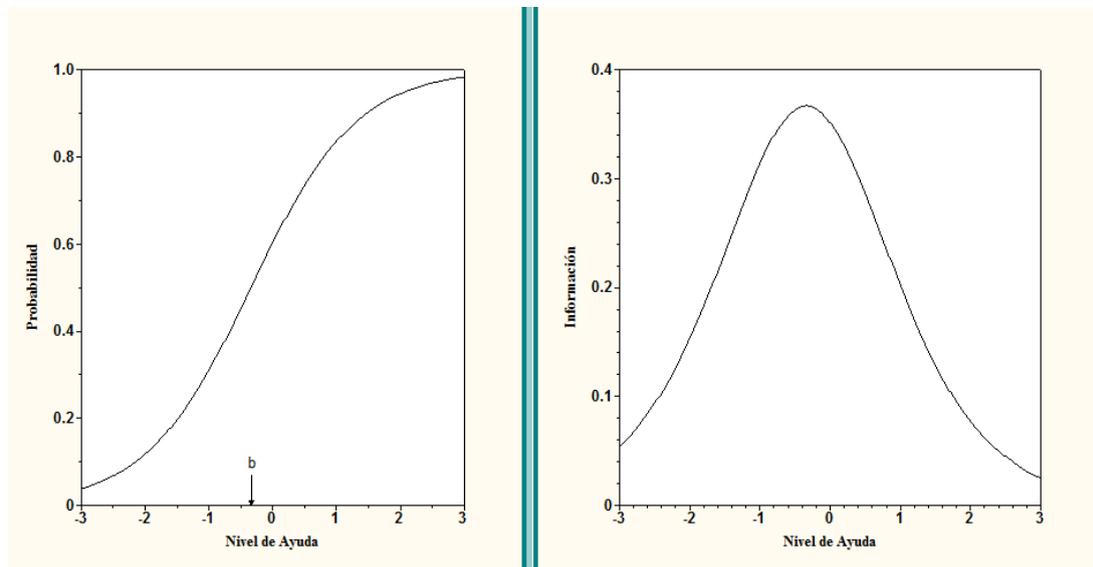
## ML2P Ítem 7



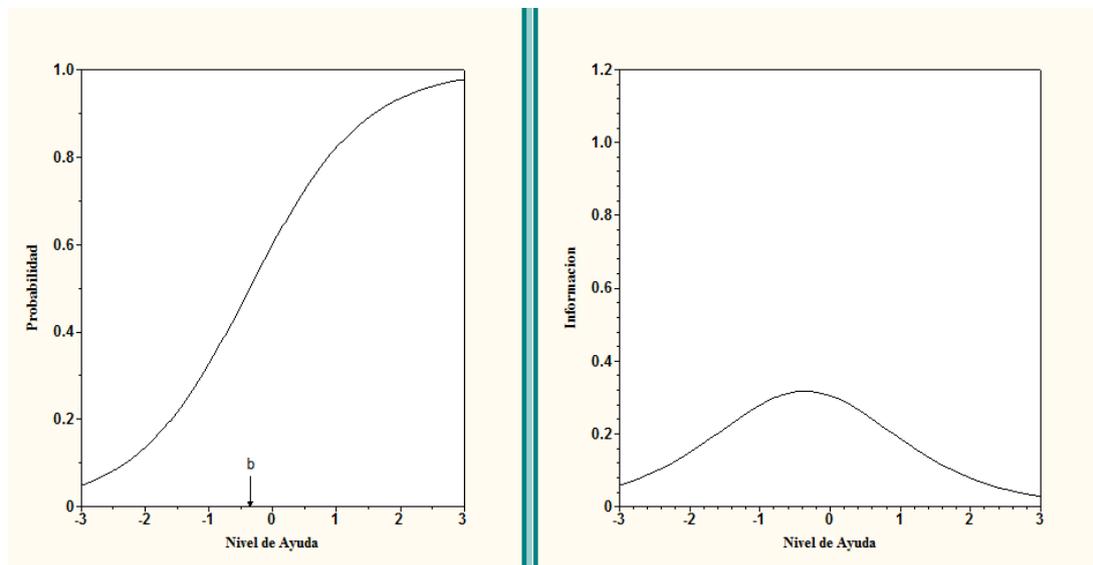
## ML3P Ítem 7



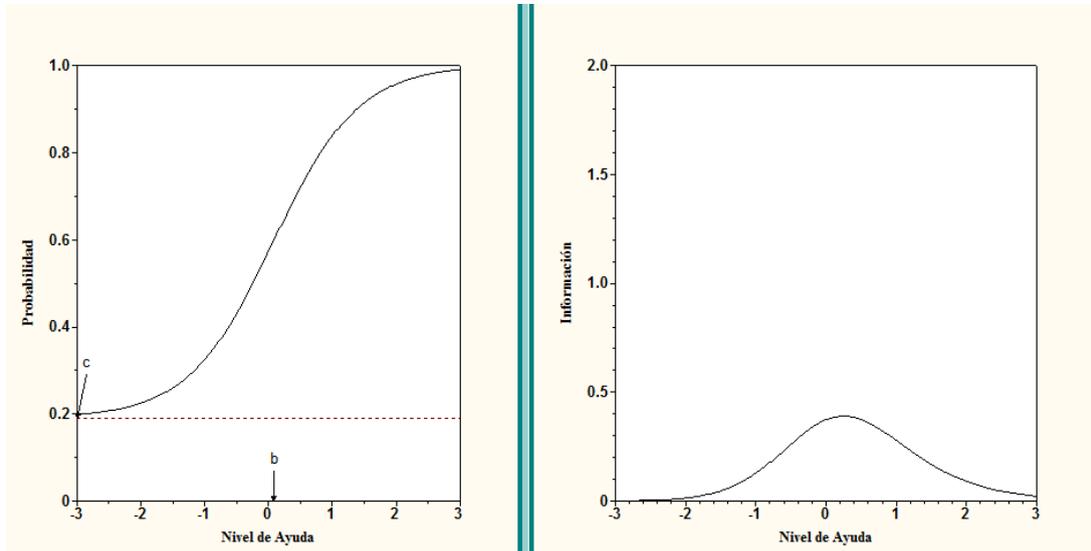
## ML1P Ítem 8



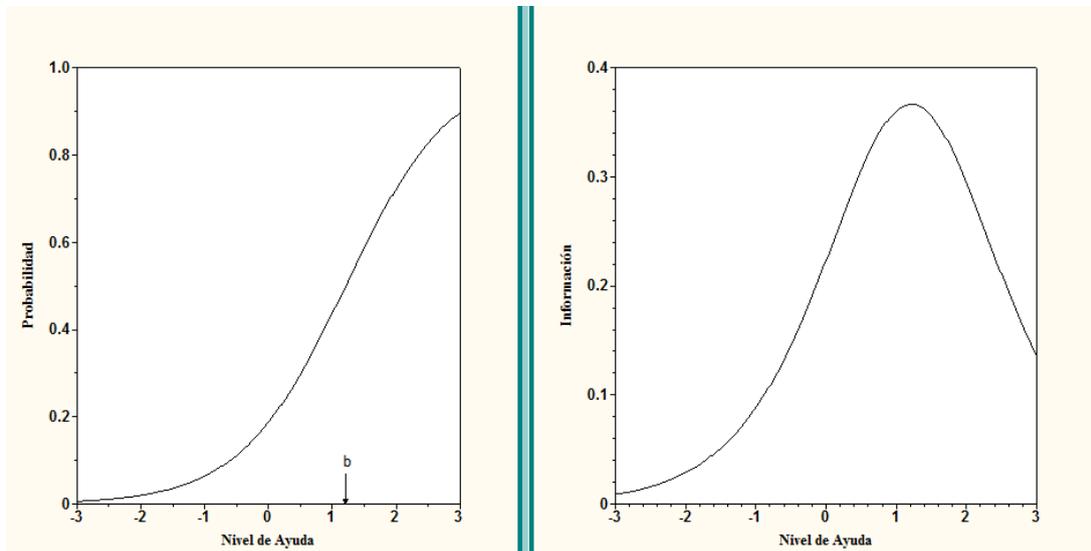
## ML2P Ítem 8



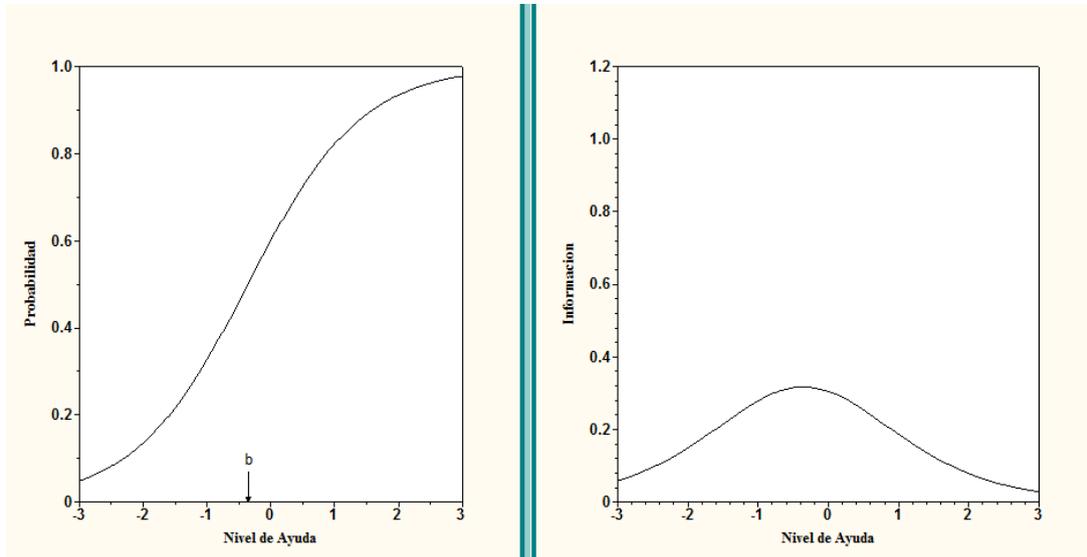
## ML3P Ítem 8



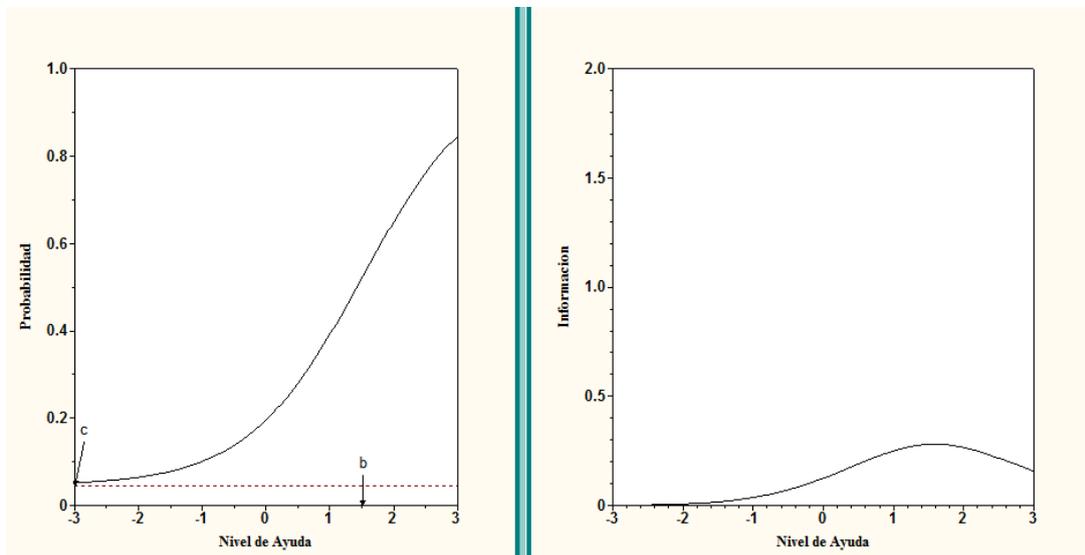
## ML1P Ítem 9



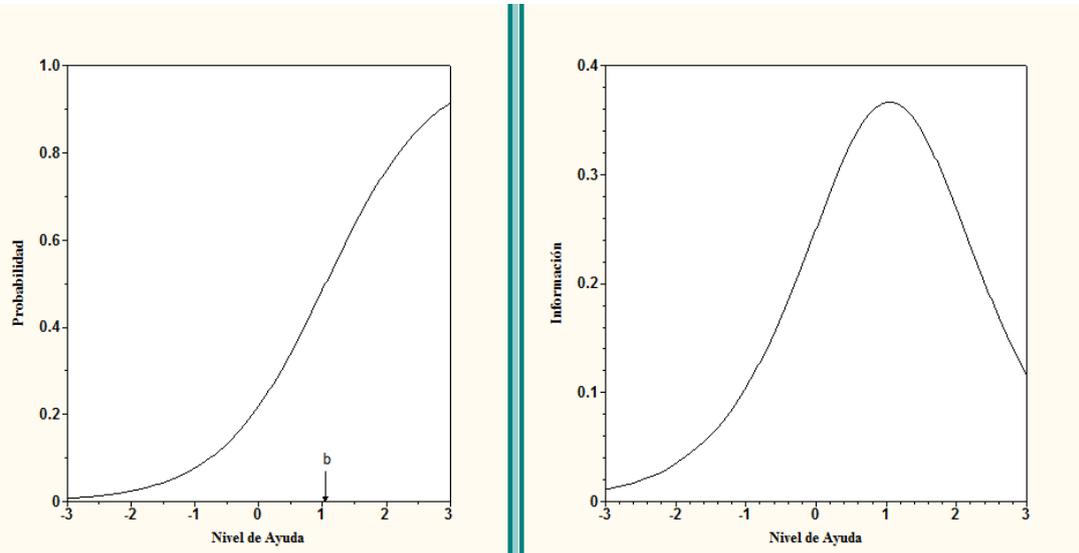
## ML2P Ítem 9



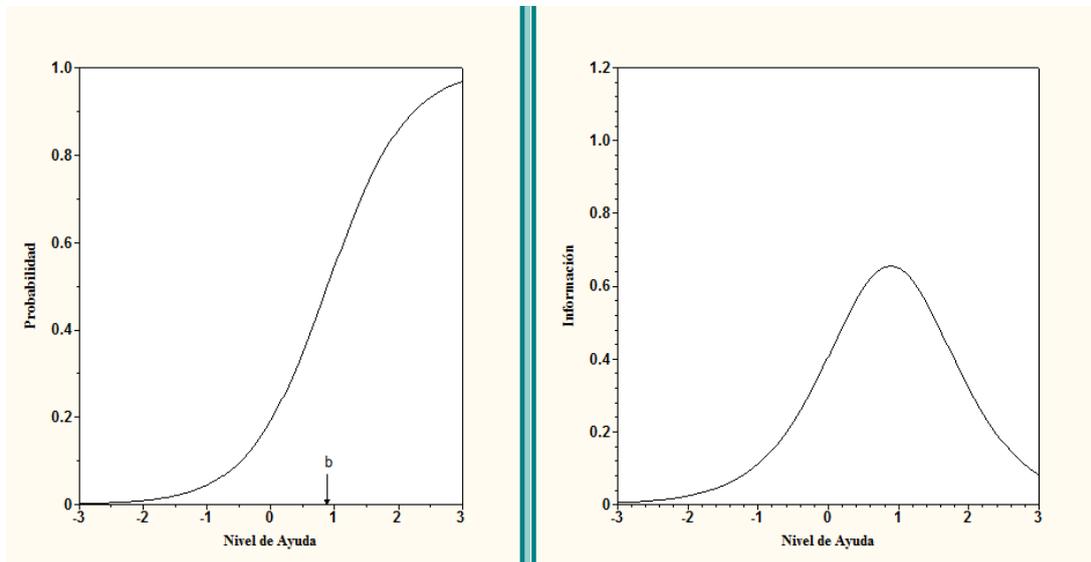
## ML3P Ítem 9



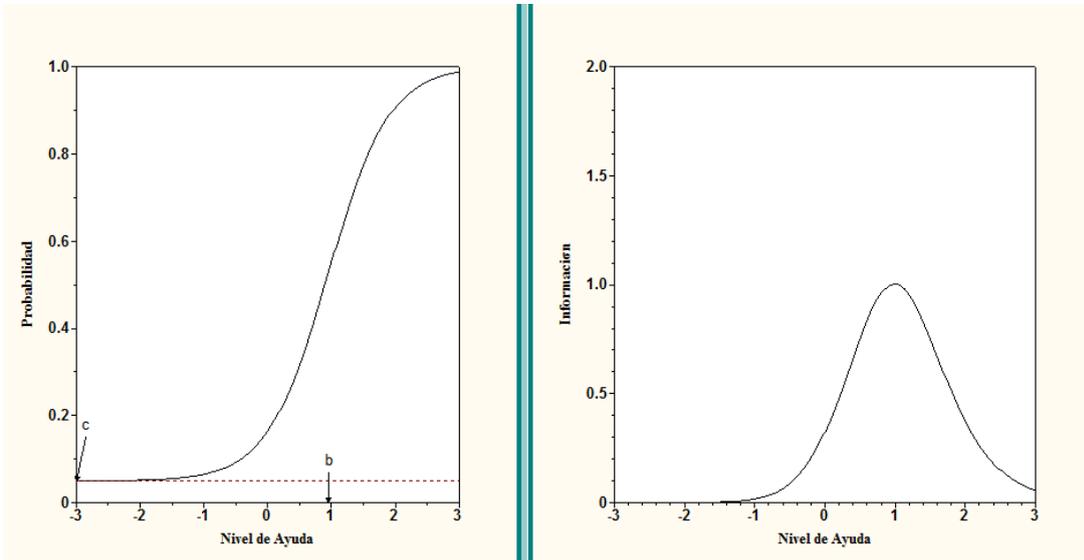
## ML1P Ítem 10



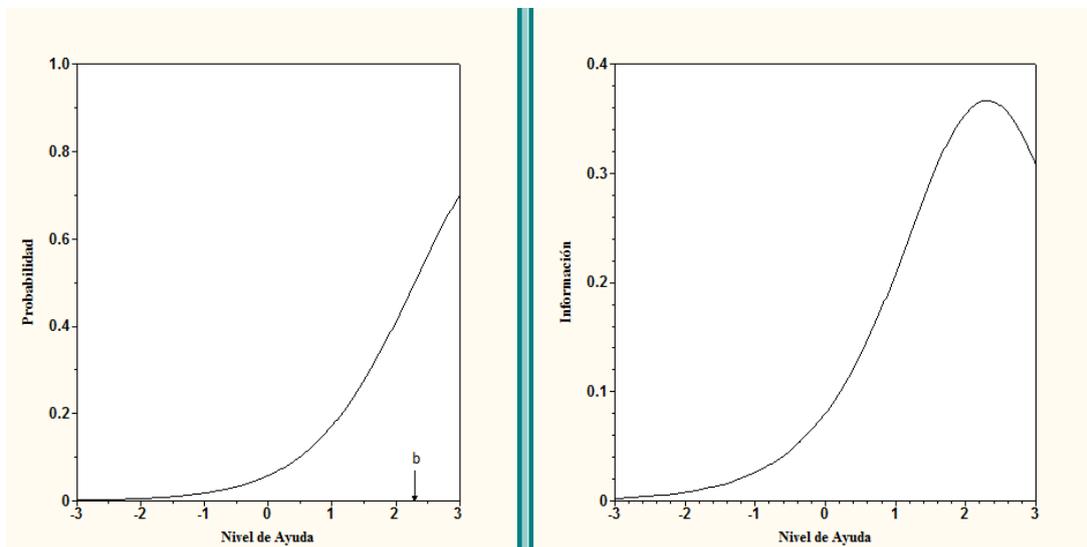
## ML2P Ítem 10



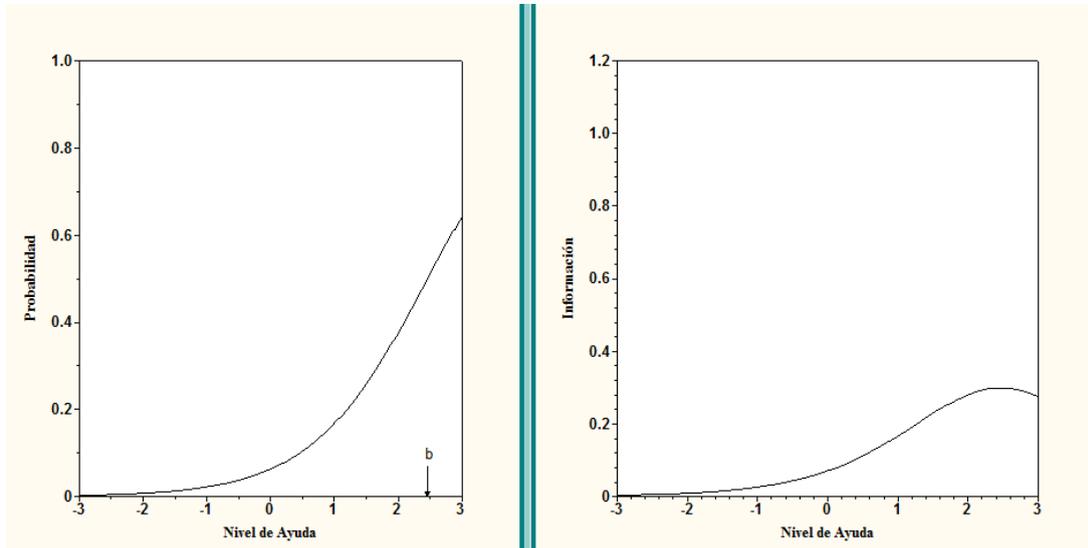
## ML3P Ítem 10



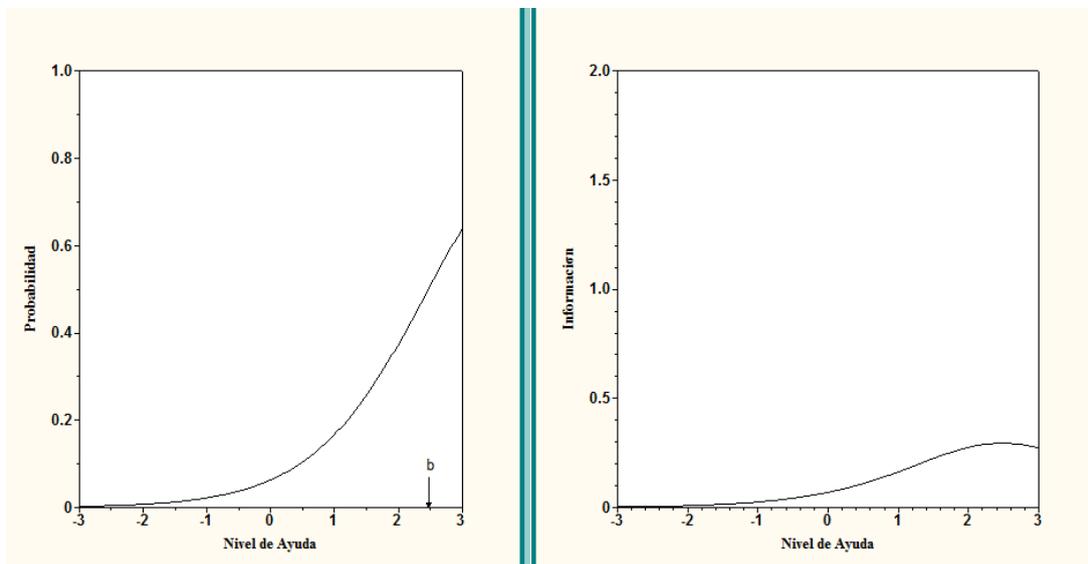
## ML1P Ítem 11



## ML2P Ítem 11



## ML3P Ítem 11



Las Tablas 9, 10 y 11 exhiben los valores que alcanzan las FIs de los ítems y del test para algunos niveles del rasgo Ayuda en los modelos dicotómicos analizados.

Tabla 9

*Funciones de Información de los Ítems y del Test obtenidas mediante MLIP*

*en la Dimensión Ayuda*

Ítem \ $\theta$	$\theta$						
	$\theta = -3$	$\theta = -2$	$\theta = -1$	$\theta = 0$	$\theta = 1$	$\theta = 2$	$\theta = 3$
1	0.01	0.04	0.11	0.26	0.37	0.26	0.11
2	0.02	0.07	0.20	0.35	0.32	0.16	0.06
6	0.03	0.09	0.22	0.36	0.30	0.14	0.05
7	0.01	0.03	0.09	0.23	0.36	0.29	0.13
8	0.05	0.15	0.31	0.35	0.20	0.08	0.03
9	0.01	0.03	0.09	0.22	0.36	0.29	0.14
10	0.01	0.04	0.10	0.25	0.37	0.27	0.12
11	0.00+	0.01	0.03	0.08	0.21	0.35	0.31
Test	1.15	1.45	2.15	3.10	3.48	2.84	1.93
DE	0.93	0.83	0.68	0.57	0.54	0.59	0.72

Tabla 10

*Funciones de Información de los Ítems y del Test obtenidas mediante ML2P  
en la Dimensión Ayuda*

Ítem	$\theta$						
	$\theta = -3$	$\theta = -2$	$\theta = -1$	$\theta = 0$	$\theta = 1$	$\theta = 2$	$\theta = 3$
1	0.02	0.03	0.05	0.07	0.09	0.09	0.08
2	0.03	0.08	0.17	0.26	0.25	0.15	0.06
6	0.00+	0.04	0.32	1.14	0.58	0.09	0.01
7	0.00+	0.02	0.09	0.42	0.80	0.36	0.08
8	0.06	0.15	0.28	0.30	0.19	0.08	0.03
9	0.01	0.03	0.07	0.14	0.20	0.20	0.13
10	0.00+	0.02	0.11	0.41	0.65	0.32	0.08
11	0.00+	0.01	0.03	0.07	0.17	0.28	0.28
Test	1.14	1.38	2.11	3.81	3.91	2.56	1.74
DE	0.94	0.85	0.69	0.51	0.51	0.63	0.76

Tabla 11

*Funciones de Información de los Ítems y del Test obtenidas mediante ML3P  
en la Dimensión Ayuda*

$\theta$	$\theta = -3$	$\theta = -2$	$\theta = -1$	$\theta = 0$	$\theta = 1$	$\theta = 2$	$\theta = 3$
Ítem							
1	0.00+	0.00+	0.01	0.05	0.15	0.22	0.16
2	0.03	0.08	0.17	0.26	0.25	0.15	0.06
6	0.00+	0.00+	0.02	1.29	0.87	0.05	0.00+
7	0.00+	0.02	0.09	0.42	0.80	0.36	0.08
8	0.00+	0.03	0.17	0.37	0.25	0.09	0.02
9	0.00+	0.01	0.04	0.13	0.24	0.25	0.15
10	0.00+	0.00+	0.02	0.33	0.97	0.37	0.06
11	0.00+	0.00+	0.02	0.07	0.16	0.28	0.27
Test	1.04	1.14	1.54	3.92	4.70	2.76	1.81
DE	0.98	0.94	0.80	0.51	0.46	0.60	0.74

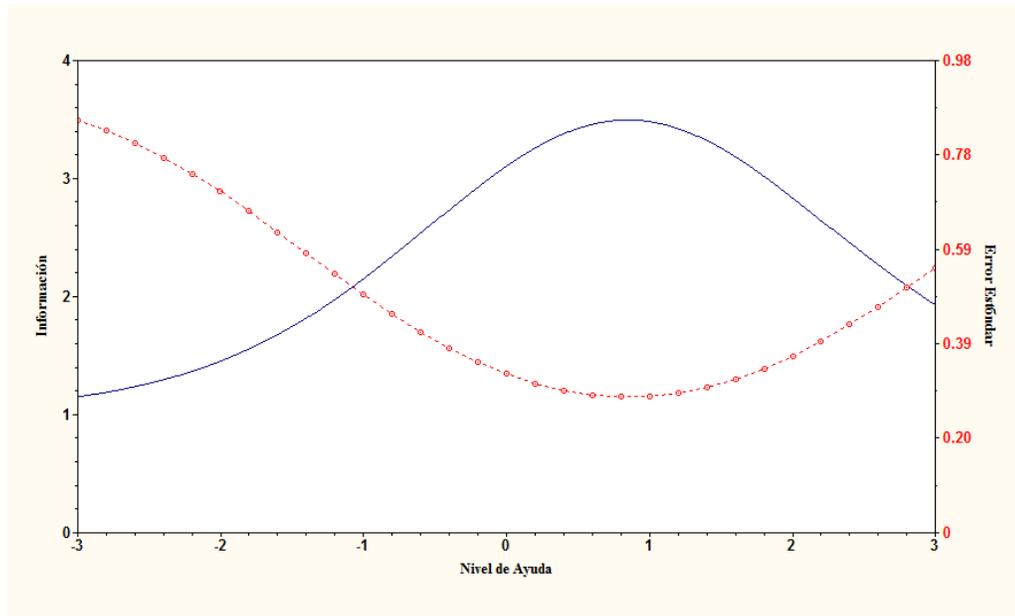
La Figura 14 muestra que el gráfico de la FI del test alcanzó para el ML1P su valor máximo de 3.49 en  $\theta = 0.80$  con un valor mínimo del error estándar de medida (EE) en ese punto de 0.53. Para el ML2P el máximo punto de la FI de 4.20 se ubicó en  $\theta = 0.60$  donde el EE fue de 0.49. En el caso del ML3P el valor máximo de la FI de 5.35 fue en  $\theta = 0.60$  con un EE en dicho punto de 0.43.

Las FI de ML1P, ML2P y ML3P muestran que el test en su versión de respuesta dicotómica proporciona estimaciones de Ayuda más precisas en un rango de Ayuda de 0 a 2, mientras que el error en la medición crece hacia los niveles de rasgo menores a 0 y mayores a 2.

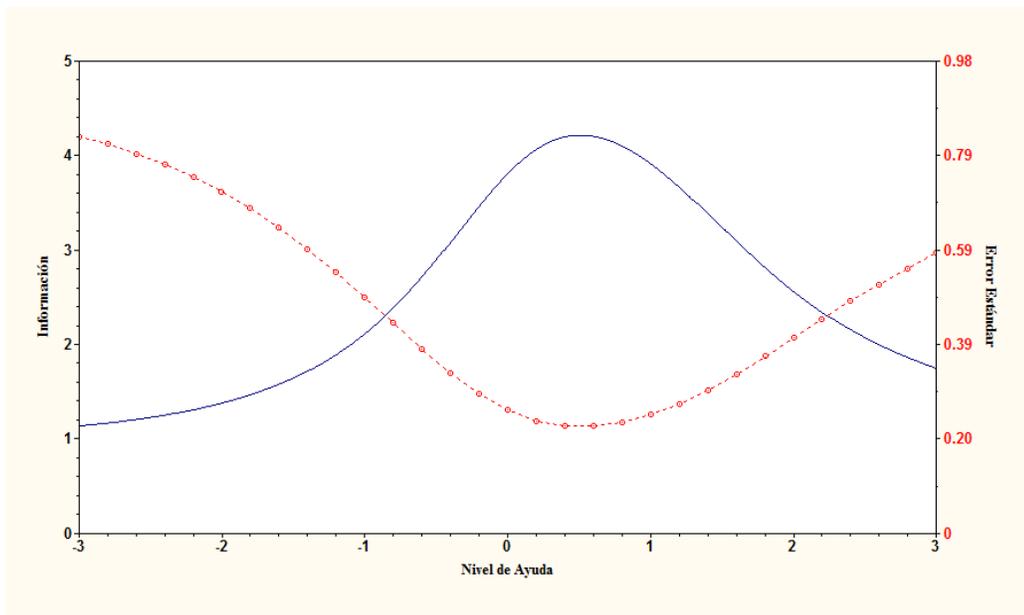
Figura 14

*Función de Información del Test y Error Estándar de Medida  
para ML1P, ML2P y ML3P en la dimensión Ayuda*

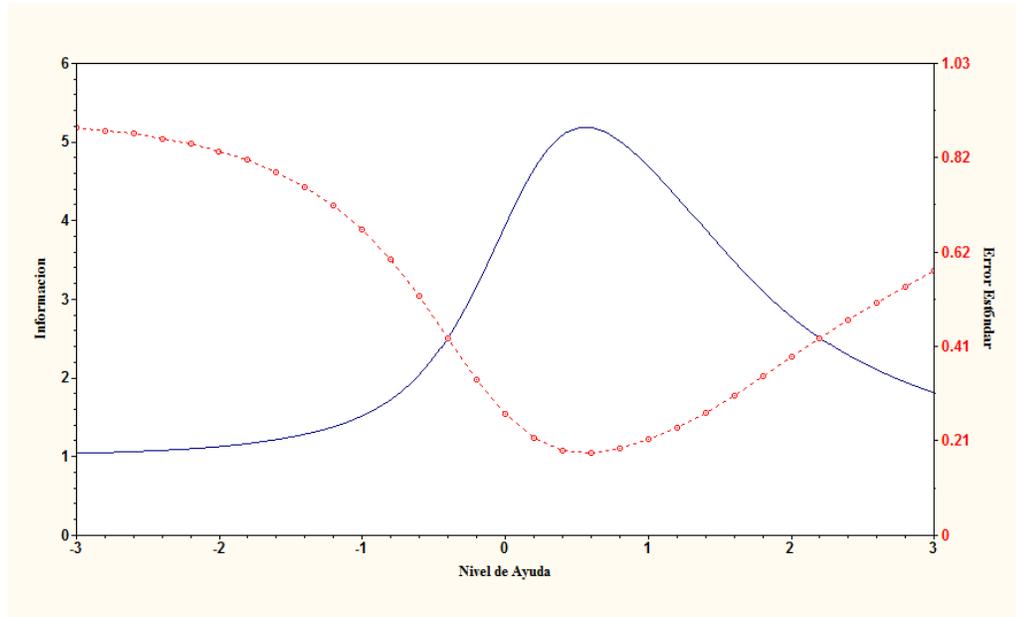
FI del Test y EE de Medida ML1P



FI del Test y EE de Medida ML2P



## FI del Test y EE de Medida ML3P



Notación

FI — EE - - -

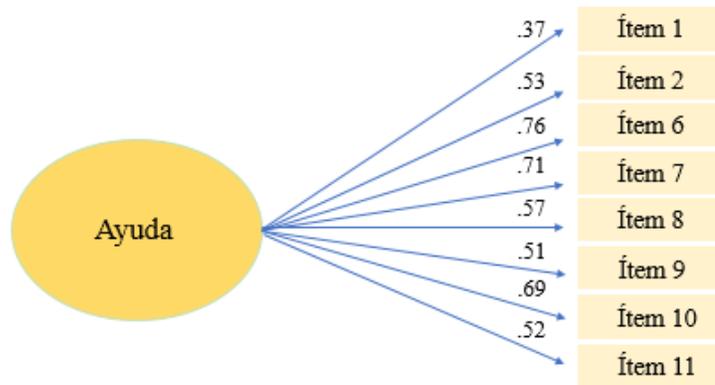
### *Cumplimiento de los supuestos de los Modelos de Respuesta Graduada y Crédito Parcial en la Dimensión Ayuda*

Se implementó un AFC con el programa Mplus (Muthén & Muthén, 2010) sobre los ocho ítems que componen la dimensión Ayuda para verificar el supuesto de unidimensionalidad requerido por los modelos de la TRI. Se respetó el carácter ordinal de los datos por lo que se estimaron los parámetros con el método WLSMV sobre la base de las matrices de correlaciones policóricas. Se analizó el ajuste al modelo considerando los índices CFI, TLI y RMSEA. Este análisis fue realizado sobre la totalidad de los participantes de la muestra ( $n = 1383$ )

Los indicadores de ajuste comparativos muestran un resultado muy adecuado para la verificación del modelo unidimensional (CFI = .972; TLI = .961; RMSEA = .067, IC 90% .056-.077) por lo cual se acepta el cumplimiento del supuesto de unidimensionalidad para la dimensión Ayuda. Tal como se observa en la Figura 15, las cargas factoriales son elevadas (Abad et al., 2011).

Figura 15

*Modelo Factorial Confirmatorio de la Dimensión Ayuda*



El supuesto de independencia local se dio por satisfecho al comprobarse la unidimensionalidad de la escala (Lord & Novick, 1968).

*Aplicación del Modelo de Respuesta Graduada a la Dimensión Ayuda*

*Ajuste de los datos al Modelo de Respuesta Graduada*

Las proporciones esperadas fueron transformadas a frecuencias absolutas ( $n = 1383$ ) para poder aplicar la prueba ji cuadrado. El modelo presentó un adecuado ajuste ya no se rechazó la hipótesis nula de igualdad entre frecuencias absolutas observadas y esperadas según el MRG para ninguno de los ítems.

Tabla 12

*Índices de Ajuste de los Ítems de la dimensión Ayuda  
al Modelo de Respuesta Graduada*

Ítem	Frecuencias Absolutas	<i>Nunca</i>	<i>Casi Nunca</i>	<i>A Veces</i>	<i>Con Frecuencia</i>	<i>Casi Siempre</i>	<i>Siempre</i>	$\chi^2$	<i>p</i>
1	Observadas	156	290	548	264	92	33	0.15	.9995
	Esperadas	157	294	544	262	93	34		
2	Observadas	155	271	385	163	136	273	0.08	.9999
	Esperadas	156	274	384	162	134	273		
6	Observadas	59	186	524	291	192	131	0.31	.9974
	Esperadas	57	192	521	288	193	131		
7	Observadas	152	435	454	182	101	59	0.21	.9989
	Esperadas	155	439	447	182	102	59		
8	Observadas	103	171	309	265	221	314	1.07	.9567
	Esperadas	106	181	308	256	216	315		
9	Observadas	158	390	504	178	103	50	0.37	.9960
	Esperadas	160	394	494	178	106	51		
10	Observadas	131	334	546	239	88	45	0.28	.9980
	Esperadas	132	338	538	240	91	45		
11	Observadas	398	516	338	76	42	13	0.11	.9998
	Esperadas	399	511	339	78	42	13		

*Estimación de Parámetros del Modelo de Respuesta Graduada para los ítems  
de la Dimensión Ayuda*

Los 48 parámetros fueron estimados en 24 ciclos con un criterio de convergencia de .001 y un máximo cambio interciclo de .00074. Como se observa en la Tabla 13, para uno de los ítems se estimó un parámetro de inclinación  $a$ , así como cinco parámetros de umbral  $b_1$ ,  $b_2$ ,  $b_3$ ,  $b_4$  y  $b_5$ .

El parámetro  $b_1$  indica el valor de umbral que separa las categorías de respuesta *Nunca* de *Casi Nunca*. Todos los parámetros  $b_1$  de los ítems de la dimensión Ayuda son negativos, indicando que es suficiente un nivel bajo en el rasgo para tener una probabilidad mayor a 0.50 de elegir la categoría *Casi Nunca* o una superior.

El parámetro  $b_2$  indica el valor de umbral que separa las categorías de respuesta *Casi Nunca* de *A Veces*. Es el mínimo  $\theta$  necesario para tener una probabilidad de 0.50 de elegir la categoría *A Veces* o una superior. En los ítems 9 y 11 los  $b_2$  fueron positivos y para el resto de los ítems fueron

negativos, pero con menor valor absoluto con respecto a los  $b_1$ . Esto denota que para tener una probabilidad de 0.50 de elegir la categoría *A Veces* o una superior es necesario un nivel en el rasgo mayor que para tener una probabilidad de 0.50 de elegir la categoría *Casi Nunca* o una superior, pero en la mayoría de los ítems -menos el 9 y el 11- fue suficiente un nivel relativamente bajo en el rasgo alcanzar este valor de umbral.

El parámetro  $b_3$  indica el valor de umbral que separa las categorías de respuesta *A Veces* de *Con Frecuencia*. Denota el mínimo  $\theta$  necesario para tener una probabilidad de 0.50 de elegir la categoría *Con Frecuencia* o una superior. Su valor es positivo en casi todos los casos excepto para el ítem 8. En esta categoría, además de existir una distancia alta con  $b_2$  es donde el signo de los  $b$  vira de negativo a positivo (excepto para los ítems 9 y 11). Es decir que para obtener este valor de umbral, en general, no es suficiente con un nivel bajo en el rasgo, sino que será necesario desde un nivel medio a un nivel alto de rasgo para ubicarse en  $b_3$  dependiendo del ítem.

El parámetro  $b_4$  indica el valor de umbral que separa las categorías de respuesta *Con Frecuencia* de *Casi Siempre*. O sea que es el mínimo  $\theta$  necesario para tener una probabilidad de 0.50 de elegir las categorías *Casi Siempre* o *Siempre*. Es en todos los casos positivo, señalando que para tener este valor de umbral es necesario un nivel alto en el rasgo en la mayoría de los ítems.

El parámetro  $b_5$  indica el valor de umbral que separa las categorías de respuesta *Casi Siempre* de *Siempre*, el mínimo  $\theta$  necesario para tener una probabilidad de 0.50 de elegir la categoría *Siempre*. Obtener este valor de umbral es de “difícil” a “muy difícil” ya es necesario un nivel alto a extremadamente alto en el rasgo para tener una probabilidad de 0.50 de elegir la categoría *Siempre*.

Dada la amplia variabilidad de los valores asumidos por los  $b_m$  los reactivos recorren efectivamente el rasgo latente, desde -3.19 ( $b_1$  ítem 1, *Dedico tiempo para facilitar la cursada a mis compañeros*) hasta 5.6 ( $b_5$  ítem 1). Esta tendencia no es equilibrada por encima y por debajo de la media de  $\theta$  ya que los parámetros  $b_m$  son positivos (a excepción del  $b_3$  del ítem 8) a partir del paso 3. Se puede observar que las distancias entre los  $b_m$  son amplias indicando la importante discriminación de los ítems en los distintos niveles del rasgo por lo que es pertinente una escala Likert de seis categorías para la respuesta. Aunque, si bien la escala puede discriminar en un amplio espectro, es más precisa para aquellas personas con un nivel de Ayuda por encima de la media.

Los parámetros de inclinación  $a$  tuvieron un valor promedio estimado de 1.33 (DE = 0.47) que representa una capacidad alta de las categorías de respuesta para distinguir entre los diferentes

niveles de rasgo según el criterio de Reise y Waller (1990). El ítem 6 (*Me comprometo con causas nobles*) tuvo el parámetro de inclinación  $a$  más elevado con un valor de 2.07 (DE = 0.09). Los ítems 7 (*Dedico un tiempo importante de mi vida a mejorar el mundo*) y 10 (*Me ocupo del bienestar de cualquier individuo, grupo o comunidad*) también alcanzaron valores considerables en este parámetro, de 1.79 (DE = 0.08) y 1.74 (DE = 0.09) respectivamente.

Tabla 13

*Parámetros estimados mediante MRG para los ítems de la dimensión Ayuda*

Ítem	$b_1$ (DE)	$b_2$ (DE)	$b_3$ (DE)	$b_4$ (DE)	$b_5$ (DE)	$b_{prom}$	$a$ (DE)
1	-3.19 (0.32)	-1.15 (0.15)	1.48 (0.18)	3.55 (0.36)	5.60 (0.59)	1.06	0.70 (0.06)
2	-2.32 (0.17)	-0.92 (0.09)	0.42 (0.08)	1.01 (0.10)	1.61 (0.13)	-0.04	1.05 (0.07)
6	-2.30 (0.11)	-1.20 (0.06)	0.19 (0.04)	0.94 (0.06)	1.72 (0.08)	-0.13	2.07 (0.09)
7	-1.70 (0.09)	-0.25 (0.05)	0.94 (0.06)	1.67 (0.09)	2.44 (0.13)	0.62	1.79 (0.08)
8	-2.52 (0.17)	-1.40 (0.10)	-0.30 (0.07)	0.50 (0.07)	1.28 (0.10)	-0.48	1.22 (0.07)
9	-2.35 (0.18)	0.48 (0.08)	1.35 (0.12)	2.38 (0.18)	3.66 (0.28)	1.10	1.02 (0.07)
10	-1.86 (0.10)	-0.58 (0.06)	0.86 (0.06)	1.84 (0.10)	2.67 (0.15)	0.59	1.74 (0.09)
11	-1.03 (0.10)	0.75 (0.09)	2.48 (0.18)	3.44 (0.27)	4.88 (0.48)	2.10	1.07 (0.07)
Media	-2.16	-0.39	0.93	1.92	2.98		1.33
DE	0.64	0.89	0.86	1.14	1.59		0.47
Mínimo	-3.19	-1.40	-0.30	0.50	1.28		0.70
Máximo	-1.03	0.75	2.48	3.55	5.60		2.07

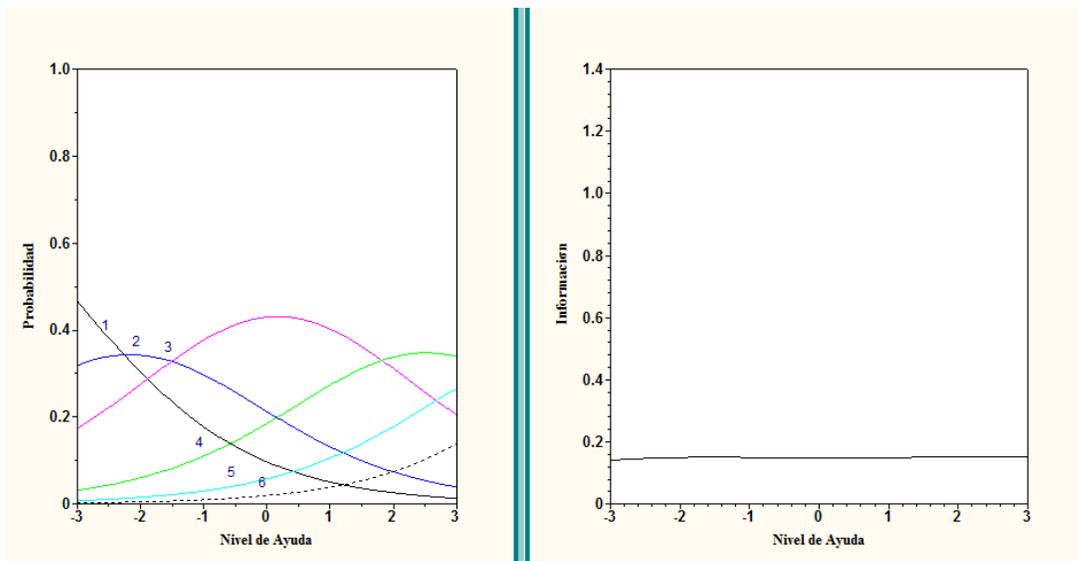
Figura 16

CCCRI y FIs de los ítems de la dimensión Ayuda para el MRG

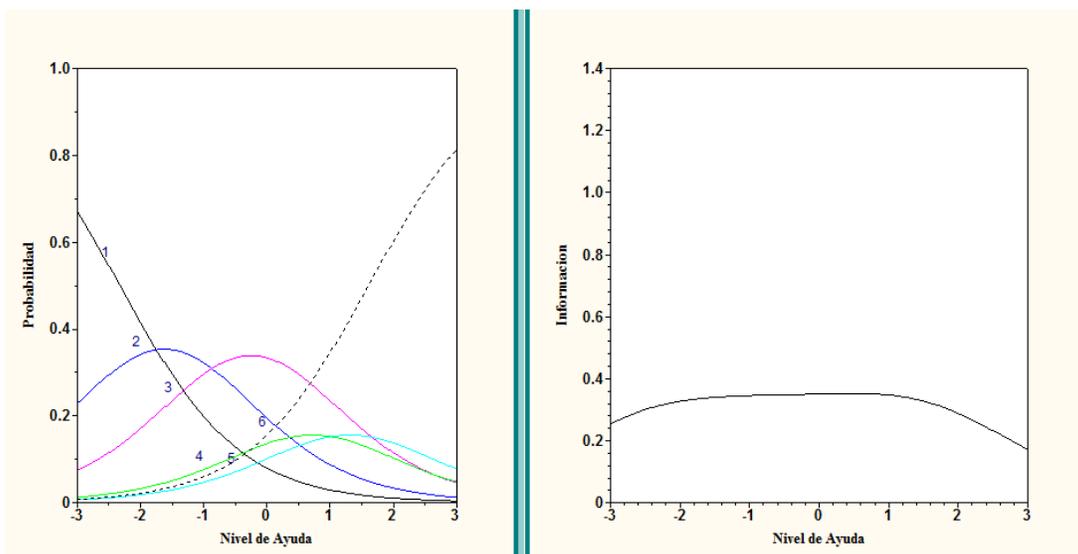
CCCRI

FI

Ítem 1



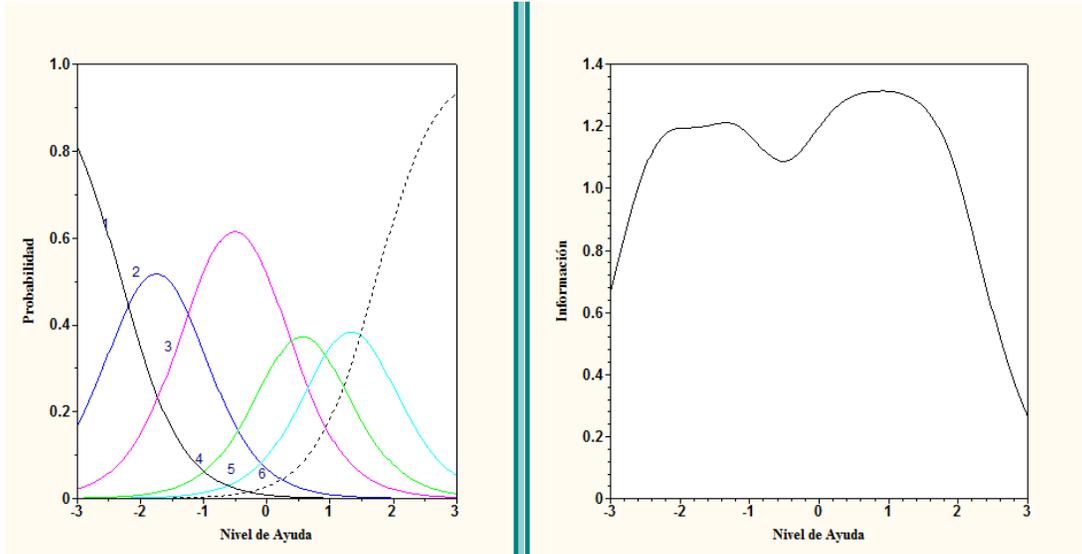
Ítem 2



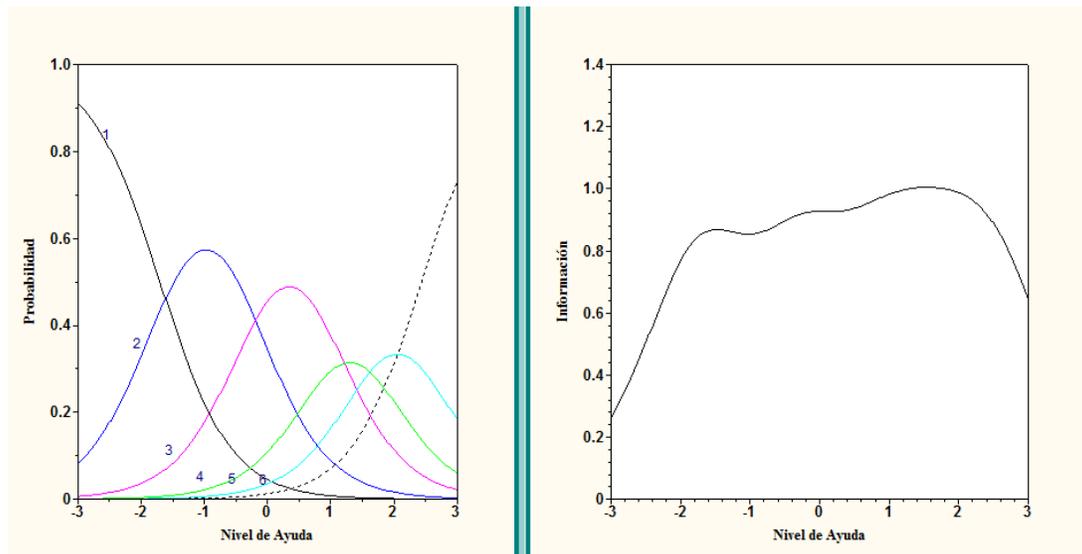
CCCRI

FI

Ítem 6



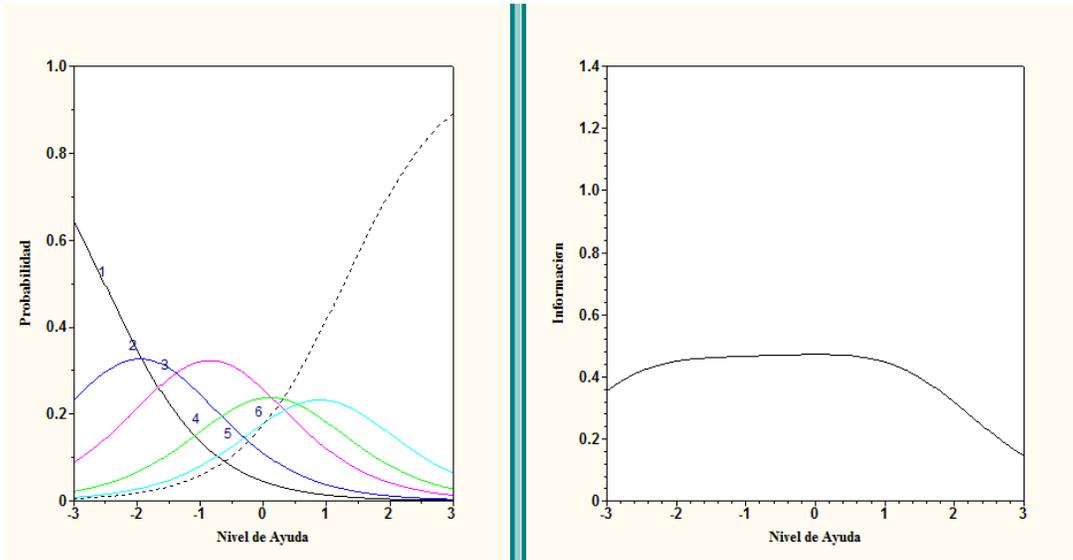
Ítem 7



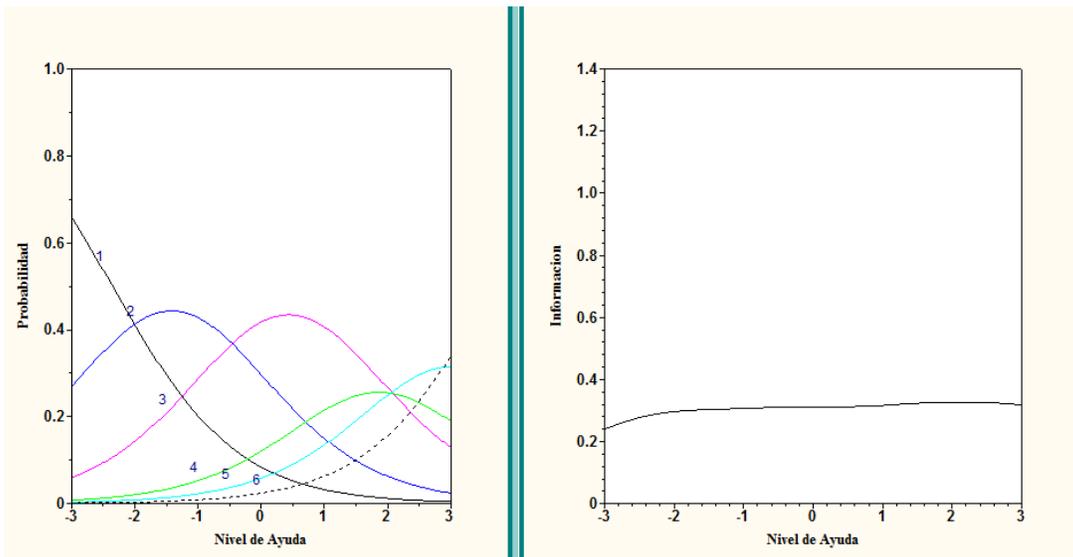
CCCRI

FI

Ítem 8



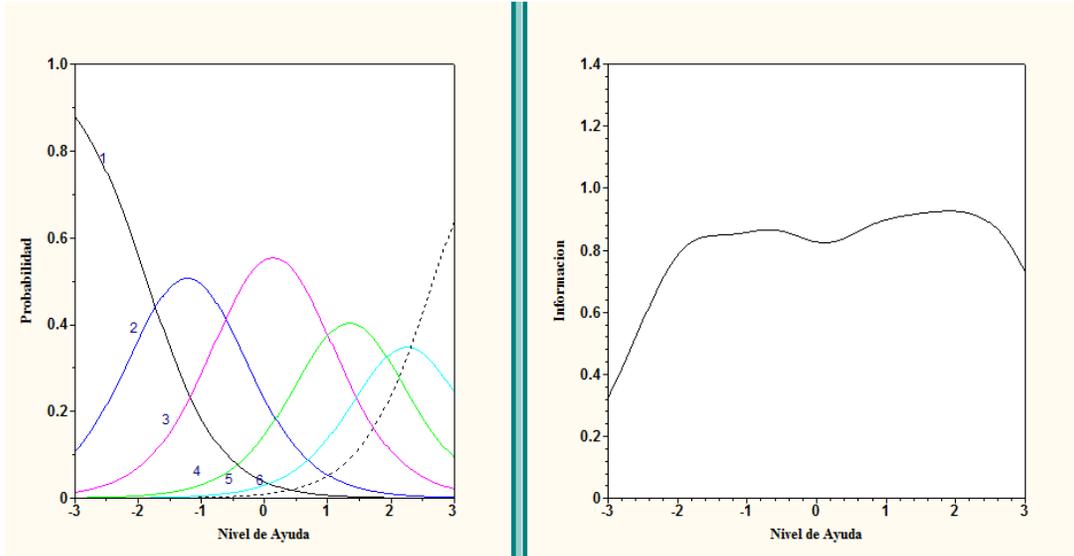
Ítem 9



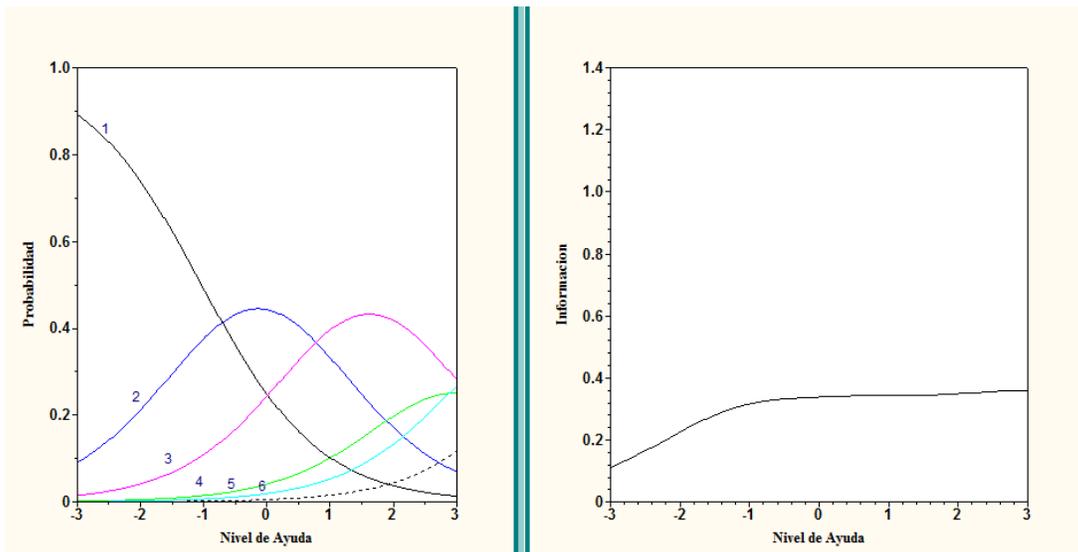
CCCRI

FI

Ítem 10



Ítem 11



## Notación

CCCRI Categoría 1 = <i>Nunca</i>	—
CCCRI Categoría 2 = <i>Casi Nunca</i>	—
CCCRI Categoría 3 = <i>A veces</i>	—
CCCRI Categoría 4 = <i>Con Frecuencia</i>	—
CCCRI Categoría 5 = <i>Casi Siempre</i>	—
CCCRI Categoría 6 = <i>Siempre</i>	---

La Tabla 14 exhibe los valores que alcanzan las FIs de los ítems y del test para algunos niveles del rasgo Ayuda. Los ítems 1 (*Dedico tiempo para facilitar la cursada a mis compañeros*), 2 (*Si tuviera un único plato de comida, lo compartiría con un desconocido*), 8 (*Expreso abiertamente mi apoyo a minorías discriminadas, aunque no pertenezca a las mismas*), 9 (*Doy dinero a cualquiera que lo necesite*) y 11 (*Me quedo con lo justo y necesario para vivir y reparto todo lo demás*) presentaron una FI con valores relativamente bajos y uniformes a lo largo de todos los niveles del rasgo. El ítem 1 presentó la curva más plana y el nivel de información más bajo. Este bajo nivel de información es explicado por tener el parámetro de discriminación más pequeño. Los ítems 6 (*Me comprometo con causas nobles*), 7 (*Dedico un tiempo importante de mi vida a mejorar el mundo*) y 10 (*Me ocupo del bienestar de cualquier individuo, grupo o comunidad*) tienen FIs notablemente más elevadas y aportan mayor información desde los niveles bajos hasta los niveles altos del rasgo.

Tabla 14

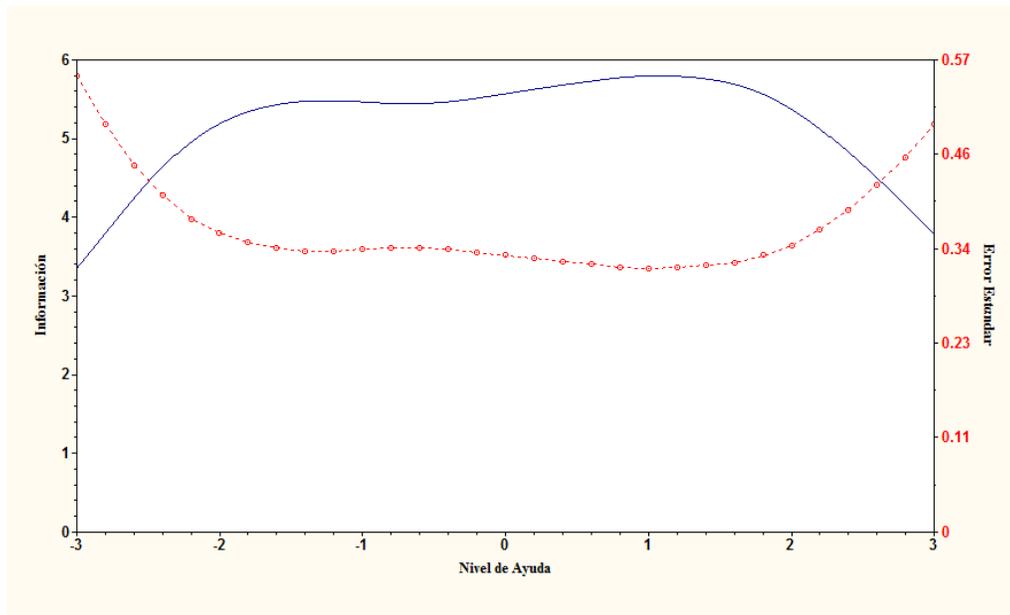
*Funciones de Información de los Ítems y del Test obtenidas mediante MRG  
en la Dimensión Ayuda*

Ítem \ $\theta$	$\theta$						
	$\theta = -3$	$\theta = -2$	$\theta = -1$	$\theta = 0$	$\theta = 1$	$\theta = 2$	$\theta = 3$
1	0.14	0.15	0.15	0.15	0.15	0.15	0.15
2	0.25	0.33	0.35	0.35	0.35	0.29	0.17
6	0.66	1.19	1.17	1.20	1.31	1.03	0.27
7	0.26	0.67	0.85	0.93	0.98	0.99	0.65
8	0.36	0.45	0.47	0.47	0.45	0.32	0.15
9	0.24	0.30	0.31	0.31	0.32	0.33	0.32
10	0.32	0.78	0.86	0.83	0.91	0.93	0.73
11	0.11	0.22	0.32	0.34	0.34	0.35	0.36
Test	3.35	5.20	5.47	5.57	5.79	5.37	3.79
DE	0.55	0.44	0.43	0.43	0.42	0.43	0.51

La Figura 17 muestra que el gráfico de la FI del test tuvo forma de meseta y alcanzó su valor máximo de 5.80 en  $\theta = 1$ . Como es de esperar dada su relación inversa, el error de medida alcanzó en este punto su valor mínimo de .42. Esta función muestra que el test proporciona estimaciones de Ayuda relativamente precisas en un rango amplio de rasgo de -3 a 3, mientras que el error en la medición crece hacia los niveles de Ayuda menores a -3 y mayores a 3.

Figura 17

*Función de Información del Test y Error Estándar  
en la dimensión Ayuda para el MRG*



Notación

FI — EE - - -

*Aplicación del Modelo de Crédito Parcial a la Dimensión Ayuda*

*Ajuste de los Datos al Modelo de Crédito Parcial*

Las proporciones esperadas fueron transformadas a frecuencias absolutas ( $n = 1383$ ) para poder aplicar la prueba ji cuadrado. El modelo presentó un adecuado ajuste ya que no se rechazó la hipótesis nula de igualdad entre frecuencias absolutas observadas y esperadas según el MCP para ninguno de los ítems.

Tabla 15

*Índices de Ajuste de los Ítems de la Dimensión Ayuda  
al Modelo de Crédito Parcial*

Ítem	Frecuencias Absolutas	<i>Nunca</i>	<i>Casi Nunca</i>	<i>A Veces</i>	<i>Con Frecuencia</i>	<i>Casi Siempre</i>	<i>Siempre</i>	$\chi^2$	<i>p</i>
1	Observadas	156	290	548	264	92	33	0.08	.9999
	Esperadas	156	292	546	262	93	34		
2	Observadas	155	271	385	163	136	273	0.10	.9998
	Esperadas	155	274	384	161	134	275		
6	Observadas	59	186	524	291	192	131	0.11	.9998
	Esperadas	58	188	525	288	191	133		
7	Observadas	152	435	454	182	101	59	0.11	.9998
	Esperadas	152	437	451	181	102	61		
8	Observadas	103	171	309	265	221	314	0.13	.9996
	Esperadas	102	174	311	263	218	314		
9	Observadas	158	390	504	178	103	50	0.12	.9997
	Esperadas	158	392	501	177	104	52		
10	Observadas	131	334	546	239	88	45	0.14	.9996
	Esperadas	131	336	543	237	89	47		
11	Observadas	398	516	338	76	42	13	0.11	.9998
	Esperadas	399	515	336	76	43	14		

*Estimación de Parámetros del Modelo de Crédito Parcial para los ítems  
de la Dimensión Ayuda*

La estimación de los 41 parámetros para el MCP en la dimensión Ayuda convergió en el ciclo 19, con un criterio de convergencia de .001. La variación máxima de los parámetros estimados entre los ciclos fue de .00078. Este modelo se aplicó imponiendo restricciones al Modelo de Respuesta Nominal tal como lo exige el programa MULTLOG.

Los parámetros  $\delta_m$  del MCP no tienen en lo absoluto el mismo significado que los parámetros  $b_m$  del MRG ya que el MCP no parametriza umbrales entre regiones de un continuo (Masters, 1982). Cuando las respuestas a un ítem se encuentran ordenadas en  $m$  categorías ordenadas es posible interpretar que  $m - 1$  pasos deben ser realizados para elegir la categoría más alta. Con respecto a una persona que selecciona la categoría *Siempre*, se considera que ha realizado cinco pasos: en el primer paso elige *Casi Nunca* frente a *Nunca*, en el segundo paso elige *A Veces*

frente a *Casi Nunca*, en el tercer paso elige *Con Frecuencia* frente a *A Veces*, en el cuarto paso elige *Casi Siempre* frente a *Con Frecuencia* y, por último, en el quinto paso elige *Siempre* frente a *Casi Siempre*. Idéntico razonamiento se sigue si la persona selecciona otra categoría de respuesta.

Los parámetros  $\delta_m$  representan una dificultad relativa e individual de cada paso. El parámetro  $\delta_5$ , por ejemplo, indica cuan probable es para una persona que ya ha alcanzado la categoría *Casi Siempre* completar el quinto paso y seleccionar la categoría *Siempre*. Los parámetros  $\delta_m$  indican en qué parte del continuo del rasgo latente las CCCRI se intersectan, y, consiguientemente para esta dimensión en qué nivel de Ayuda elegir una determinada categoría se vuelve más probable que elegir la previa (Embretson & Reise, 2000).

El MCP no impone restricciones a los valores de los  $\delta_m$ . Los pasos no necesariamente están ordenados con respecto a la cantidad de rasgo que requieran (Masters, 1982, 2016). Mientras que ciertos pasos de una categoría a otra podrían requerir un alto nivel del rasgo, en otros alcanzaría uno bajo, sin importar el orden en el cual los pasos fueron realizados.

En la Tabla 16 se muestran los valores de los parámetros  $\delta_m$  estimados para cada uno de los ítems de la dimensión Ayuda, así como sus estadísticos descriptivos. El parámetro de inclinación  $a$  constante estimado fue de 0.62 (DE = 0.01), lo cual indica una capacidad discriminativa aceptable para un test de comportamiento típico (Reise & Waller, 1990). MULTILOG estimó los errores estándar de los  $\delta_m$  a través de una transformación de los  $\delta_m$  denominada  $C_{\text{coeff}}$ . Dichos errores no alcanzaron una magnitud considerable.

Tabla 16

*Parámetros estimados mediante MCP para los ítems de la dimensión Ayuda*

Ítem	$\delta_1$	$\delta_2$	$\delta_3$	$\delta_4$	$\delta_5$	$\delta_{prom}$
1	-1.67	-1.21	1.43	2.37	2.74	0.73
2	-1.68	-0.90	<b>1.42</b>	<b>0.68</b>	<b>-0.39</b>	-0.17
6	-2.80	-2.11	0.95	1.08	1.42	-0.29
7	-2.32	-0.21	<b>1.76</b>	<b>1.63</b>	1.95	0.56
8	-1.79	-1.48	0.11	<b>0.52</b>	<b>0.03</b>	-0.52
9	-2.08	-0.56	<b>1.97</b>	<b>1.56</b>	2.28	0.63
10	-2.19	-0.98	1.58	<b>2.27</b>	<b>2.16</b>	0.57
11	-0.65	0.84	<b>3.00</b>	<b>1.97</b>	3.31	1.69
Media	-2.08	-1.06	1.32	1.44	1.46	
DE	0.41	0.62	0.62	0.73	1.19	
Mínimo	-2.80	-2.11	0.11	0.52	-0.39	
Máximo	-1.67	-0.21	2.37	2.37	2.74	

*Nota:* Los parámetros  $\delta_m$  reversados se muestran en negrita.

Se puede observar que seis de los ocho ítems presentan inversiones en el orden de sus parámetros  $\delta_m$ : en el ítem 2  $\delta_4 < \delta_3$  y  $\delta_5 < \delta_4$ , en el ítem 7  $\delta_4 < \delta_3$ , en el ítem 8  $\delta_5 < \delta_4$ , en el ítem 9  $\delta_4 < \delta_3$ , en el ítem 10  $\delta_5 < \delta_4$  y finalmente en el ítem 11  $\delta_4 < \delta_3$ . Este fenómeno se denomina *reversal* en la literatura inglesa (Dodd & Koch, 1987; Embretson & Reise, 2000) y, como ya se anticipó, su aparición es una posibilidad en el MCP. No indica un desajuste del modelo o una pobre calidad psicométrica de los ítems. Sin embargo, dado que la escala Likert implica una monotonía creciente podría ser deseable un reagrupamiento de las categorías de respuesta en las que existe reverso (Sam, Li, & Lo, 2016; Stump, Husman, & Brem, 2012). En los ítems con *reversal* existe al menos una opción de respuesta que no es seleccionada frecuentemente por los sujetos con un nivel particular

en el rasgo (Andrich, 1988). Dodd y Koch (1987) señalaron posteriormente a un estudio de simulación que cuando la distancia entre los pasos es relativamente constante, la curva de información será más apuntada cuantos más  $\delta_m$  se encuentren fuera del orden creciente y cuanto mayor sea la distancia entre los  $\delta_m$  reversados. En concordancia con los hallazgos de Dodd y Koch (1987) el ítem 2, que tiene dos reversos y con importante distancia entre ellos, es el que presenta un mayor apuntamiento en su Función de Información, lo cual se puede apreciar en la Figura 18.

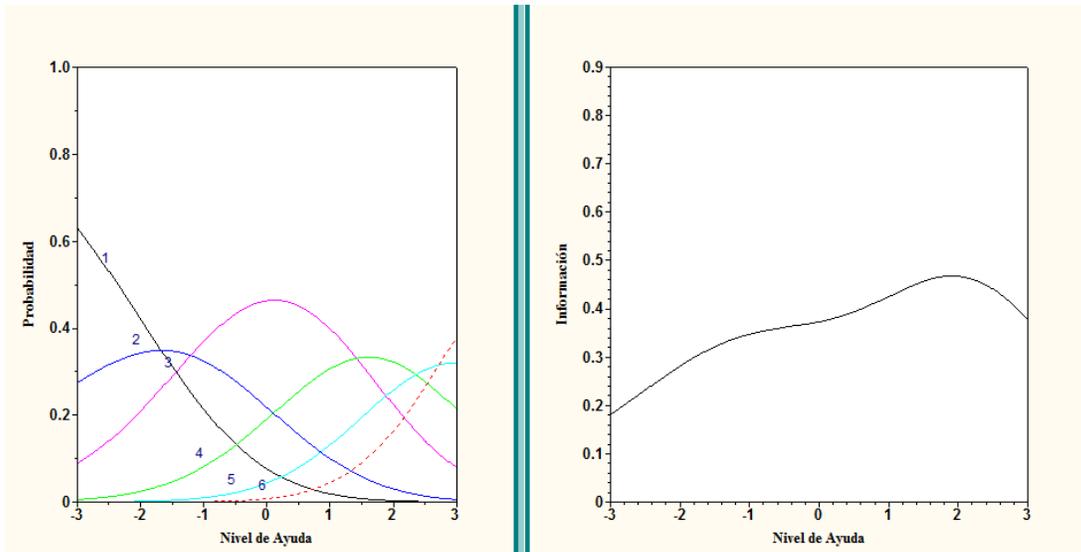
Figura 18

CCCRI y FIs de los Ítems de la Dimensión Ayuda para el MCP

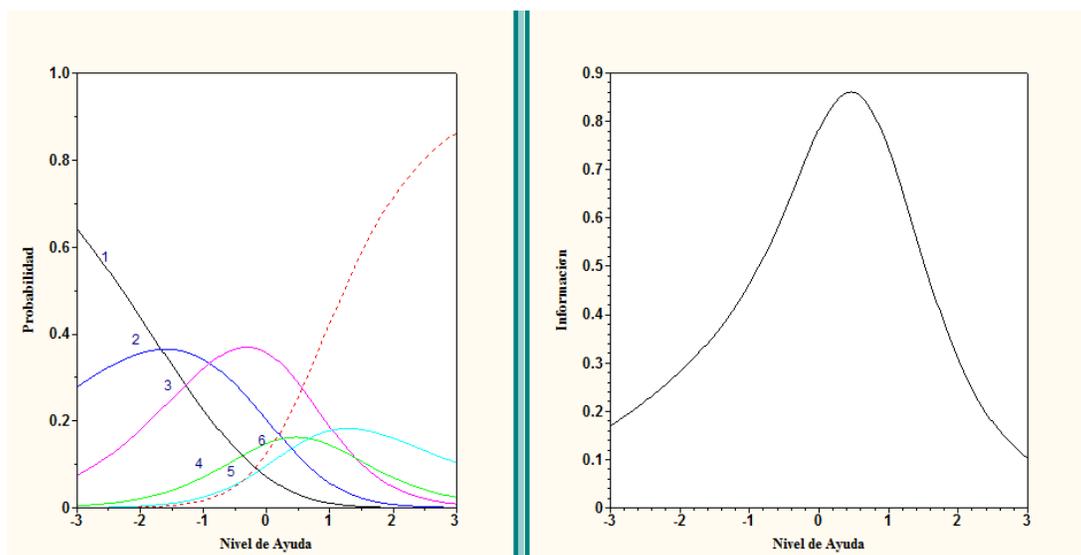
CCCRI

FI

Ítem 1



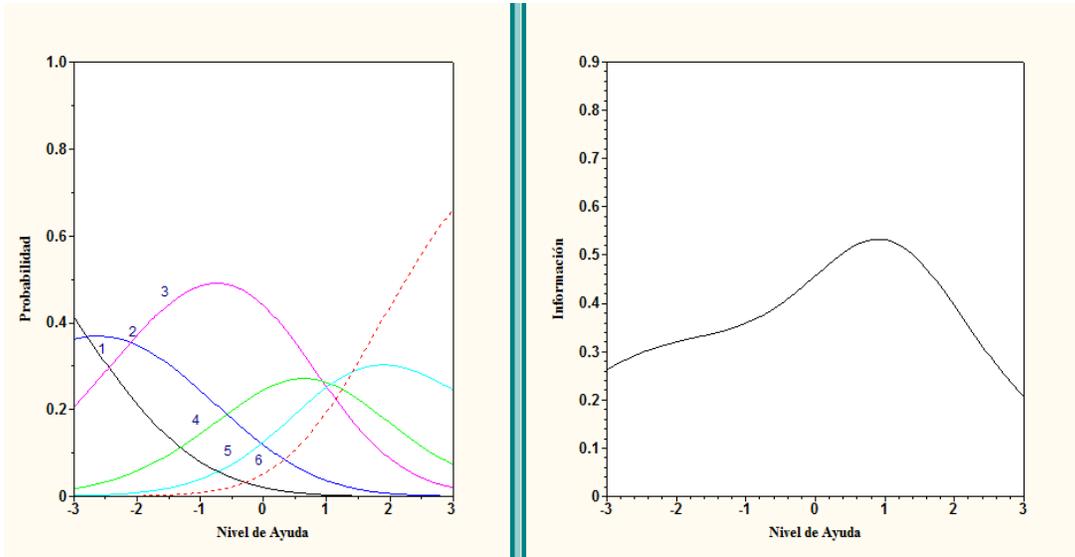
Ítem 2



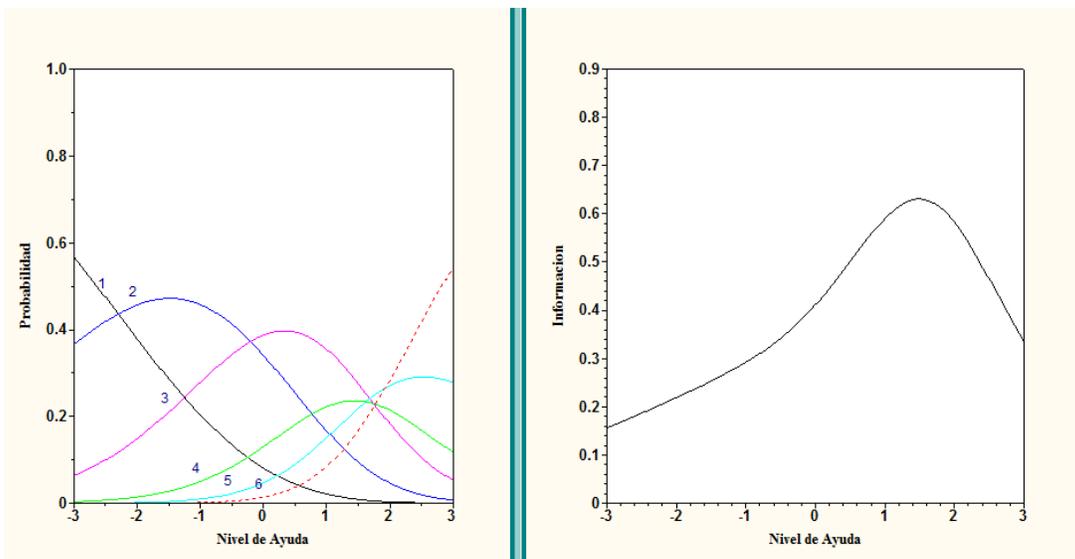
CCCRI

FI

Ítem 6



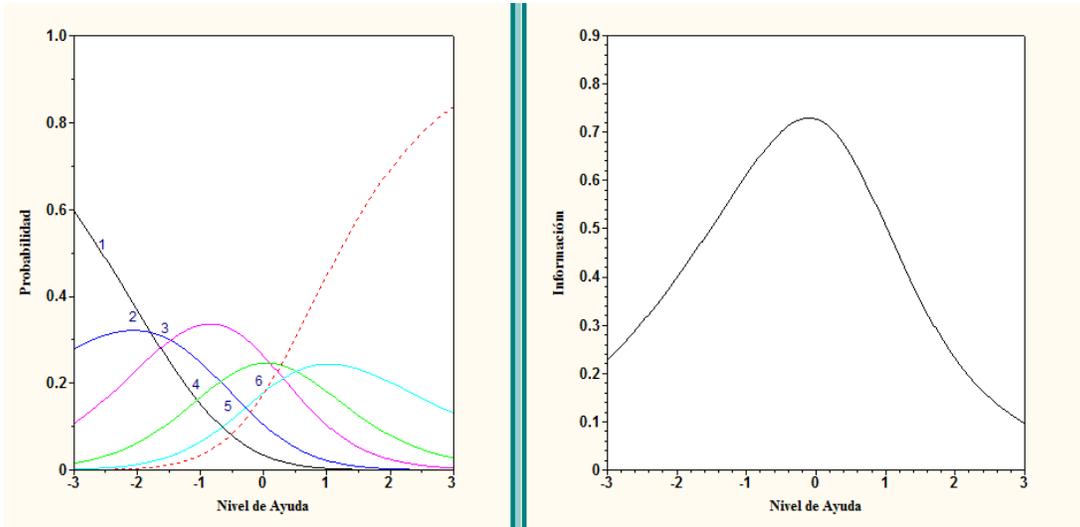
Ítem 7



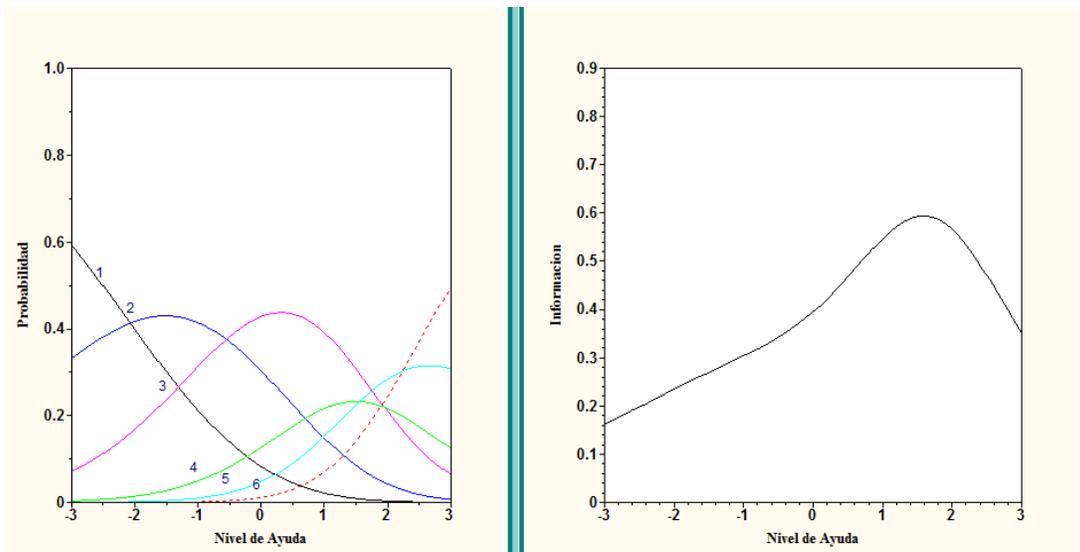
CCCRI

FI

Ítem 8



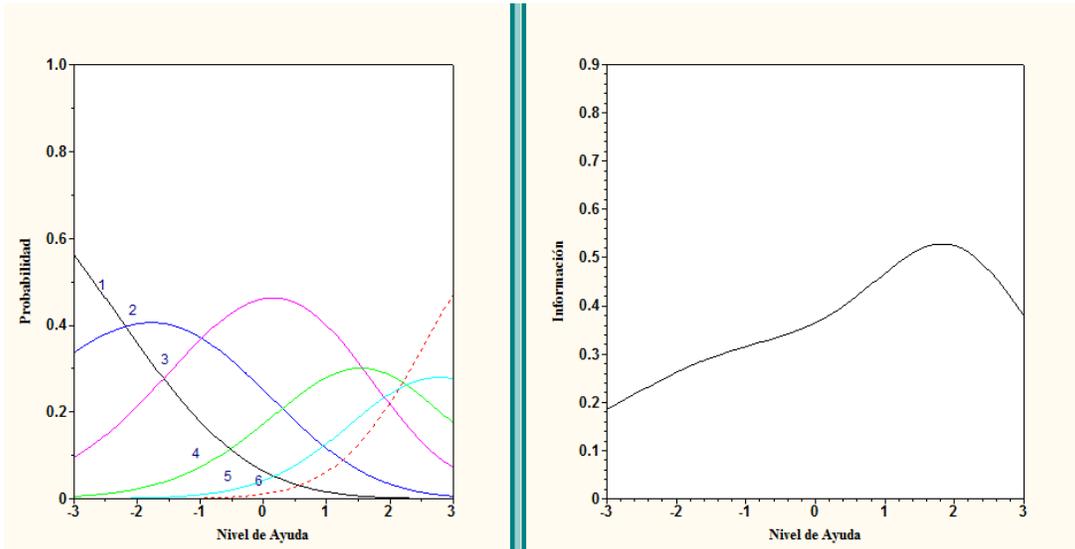
Ítem 9



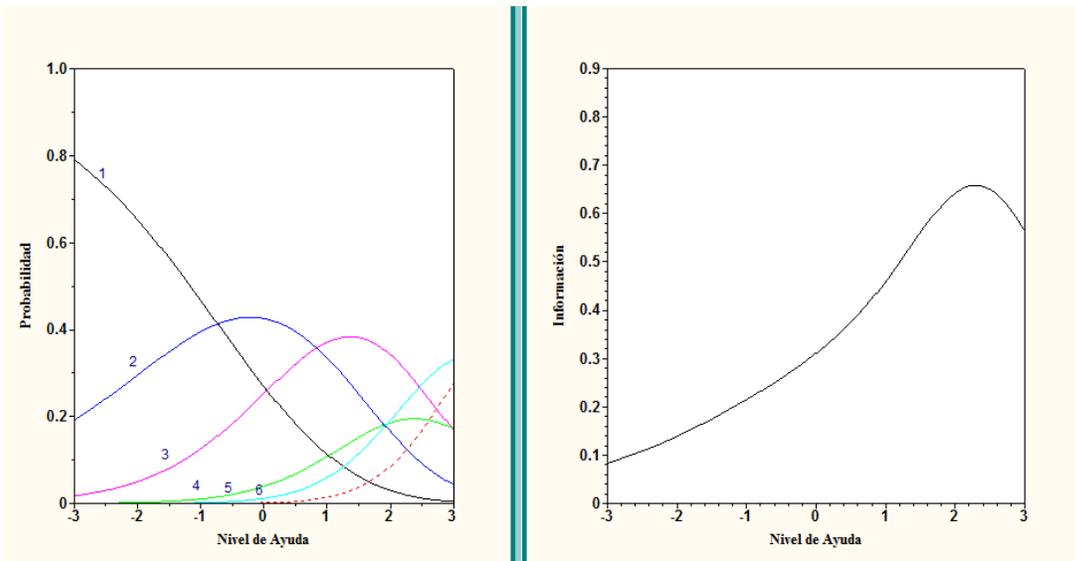
CCCRI

FI

Ítem 10



Ítem 11



Notación

- CCCRI Categoría 1 = *Nunca* —
- CCCRI Categoría 2 = *Casi Nunca* —
- CCCRI Categoría 3 = *A veces* —
- CCCRI Categoría 4 = *Con Frecuencia* —
- CCCRI Categoría 5 = *Casi Siempre* —
- CCCRI Categoría 6 = *Siempre* - - -

La Tabla 17 exhibe los valores que alcanzan las FIs de los ítems y del test para algunos niveles del rasgo Ayuda.

Tabla 17

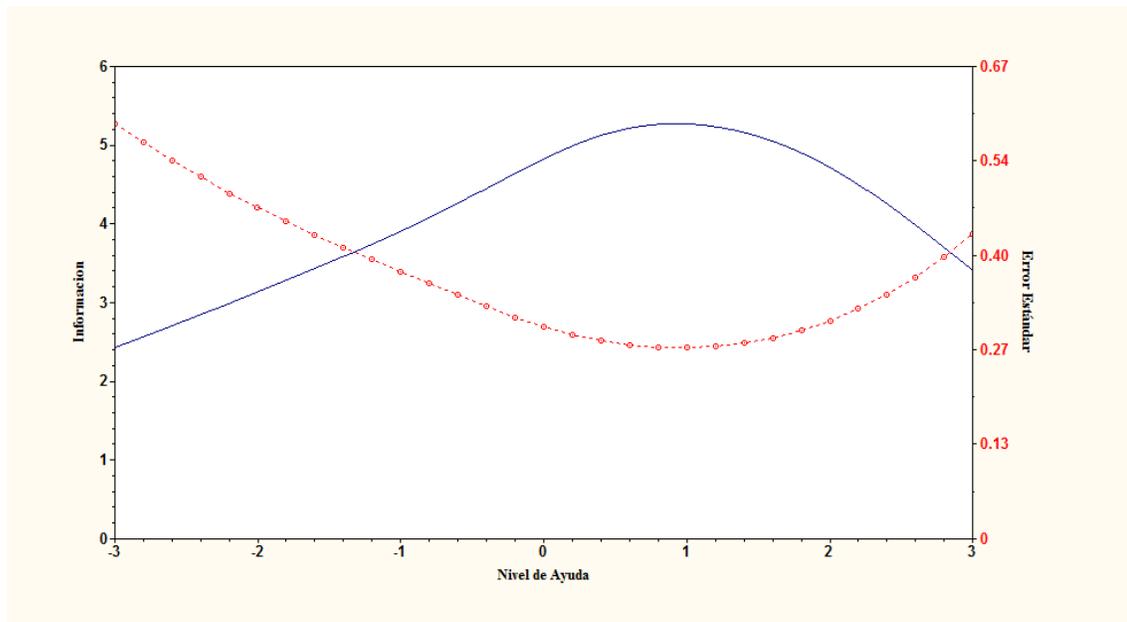
*Funciones de Información de los Ítems y del Test obtenidas mediante MCP  
en la Dimensión Ayuda*

$\theta$							
Ítem	$\theta = -3$	$\theta = -2$	$\theta = -1$	$\theta = 0$	$\theta = 1$	$\theta = 2$	$\theta = 3$
1	0.18	0.28	0.35	0.37	0.43	0.47	0.38
2	0.17	0.28	0.46	0.78	0.74	0.31	0.10
6	0.26	0.32	0.36	0.46	0.53	0.40	0.21
7	0.16	0.21	0.29	0.41	0.59	0.58	0.34
8	0.23	0.40	0.61	0.73	0.51	0.23	0.10
9	0.16	0.24	0.30	0.40	0.55	0.57	0.35
10	0.19	0.26	0.32	0.37	0.47	0.53	0.38
11	0.08	0.14	0.22	0.31	0.50	0.64	0.56
Test	2.43	3.14	3.90	4.82	5.27	4.72	3.41
DE	0.64	0.56	0.51	0.45	0.44	0.46	0.54

La Figura 19 muestra que el gráfico de la FI del test alcanzó su valor máximo de 5.25 en  $\theta = 1$ . Como es de esperar dada su relación inversa, el error de medida alcanzó en este punto su valor mínimo de .44. Esta función muestra que el test proporciona estimaciones de Ayuda más precisas en un rango de rasgo theta de -1 a 3, mientras que el error en la medición crece hacia los niveles de rasgo Ayuda menores -1.

Figura 19

*Función de Información del Test y Error Estándar  
en la Dimensión Ayuda para el MCP*



Notación

FI — EE - - -

## Aplicación de los Modelos de la TRI a la Dimensión Confortar

### *Aplicación de los Modelos Logísticos de Uno, Dos y Tres Parámetros a la Dimensión Confortar*

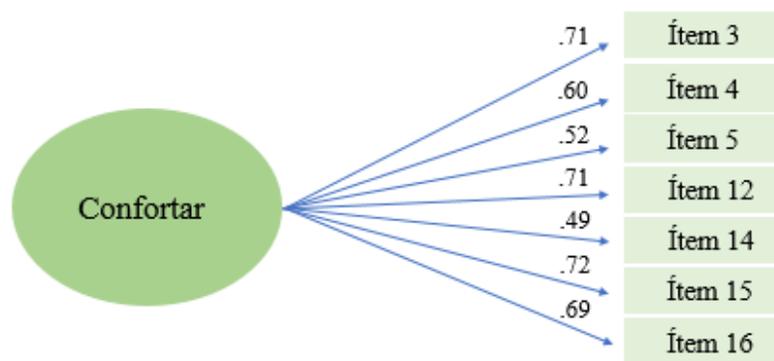
#### *Cumplimiento de los supuestos de los Modelos Logísticos de Uno, Dos y Tres Parámetros en la Dimensión Confortar*

Se implementó un AFC con el programa Mplus (Muthén & Muthén, 2010) sobre los ocho ítems que componen la dimensión Confortar para verificar el supuesto de unidimensionalidad requerido por los modelos de la TRI. Las respuestas fueron previamente dicotomizadas. Se respetó el carácter categórico/binario de los datos por lo que se estimaron los parámetros con el método WLSMV sobre la base de las matrices de correlaciones tetracóricas. Se analizó el ajuste al modelo considerando los índices CFI, TLI y RMSEA. Este análisis fue realizado sobre la totalidad de los participantes de la muestra (n = 1383)

Los indicadores de ajuste comparativos muestran un resultado muy adecuado para la verificación del modelo unidimensional (CFI = .987; TLI = .981; RMSEA = .031, IC 90% .017-.045) por lo cual se acepta el cumplimiento del supuesto de unidimensionalidad para la dimensión Confortar. Tal como se observa en la Figura 20, las cargas factoriales son elevadas (Abad et al., 2011).

Figura 20

*Modelo Factorial Confirmatorio de la Dimensión Confortar*



El supuesto de independencia local se dio por satisfecho al comprobarse la unidimensionalidad de la escala (Lord & Novick, 1968).

*Ajuste de los datos a los Modelos Logísticos de Uno, Dos y Tres Parámetros  
en la Dimensión Confortar*

Para evaluar el ajuste de los datos a los modelos logísticos, se realizó una prueba de hipótesis de diferencia de proporciones de respuestas favorables entre las obtenidas empíricamente y las esperadas según dichos modelos. Los valores  $p$  fueron en su totalidad mayores a .48 indicando el no rechazo de la hipótesis nula de igualdad de proporciones observadas y esperadas, es decir, un adecuado ajuste a los modelos ML1P, ML2P y ML3P.

Tabla 18

*Índices de ajuste de los ítems de la dimensión Confortar  
a los modelos ML1P, ML2P y ML3P*

Item	ML1P			ML2P		ML3P	
	PO	PE <sub>ML1P</sub>	$p$	PE <sub>ML2P</sub>	$p$	PE <sub>ML3P</sub>	$p$
3	.8568	.8567	.5039	.8566	.5083	.8555	.5544
4	.5824	.5806	.5536	.5817	.5207	.5816	.5239
5	.3812	.3801	.5334	.3807	.5151	.3808	.5119
12	.5987	.5967	.5600	.5972	.5449	.5970	.5509
14	.7028	.7011	.5548	.7024	.5131	.7024	.5131
15	.9017	.9020	.4852	.9019	.4900	.9020	.4852
16	.8265	.8259	.5231	.8260	.5195	.8260	.5195

*Estimación de Parámetros de los Modelos Logísticos de Uno, Dos y Tres Parámetros para los  
ítems de la Dimensión Confortar*

Los parámetros de cada ítem fueron estimados, para ML1P, ML2P y ML3P, con un criterio de convergencia de .001. En la Tabla 19 se exponen la cantidad de ciclos necesarios para alcanzar la convergencia *-finished cycle-* así como la máxima variación interciclo de las estimaciones -

*maximum intercycle parameter change*- para cada uno de dichos modelos. Tal como en la dimensión Ayuda, el ML3P fue el modelo que requirió más ciclos para alcanzar su convergencia, lo cual es coherente con la cantidad de parámetros a estimar.

Tabla 19

*Procesos de estimación de los ítems de la Dimensión Confortar  
con los modelos ML1P, ML2P y ML3P*

	ML1P	ML2P	ML3P
Cantidad de parámetros estimados	8	14	21
Cantidad de ciclos	17	18	42
Mayor cambio registrado en la estimación	0.00051	0.00047	0.00002

En la Tabla 20 se encuentran los parámetros obtenidos para los modelos ML1P, ML2P y ML3P en los ítems de la dimensión Confortar.

Para el ML1P el parámetro  $b$  promedio fue de -0.91 (DE = 0.91) siendo el ítem 15 (*Intento “levantar” la autoestima a mis amigos*) aquél con el mínimo valor y el ítem 5 (*Siento todo dolor ajeno como propio*) el reactivo con el máximo  $b$ . Se obtuvo asimismo el parámetro constante  $a$  cuyo valor fue de 1.37 con un DE de 0.05.

En el ML2P en cuanto a los  $b$  su promedio fue de -0.84 (DE = 0.83) con un mínimo en el ítem 15 (*Intento “levantar” la autoestima a mis amigos*) y un máximo valor en el ítem 5 (*Siento todo dolor ajeno como propio*). El parámetro  $a$  indica en el entorno de  $b$  dónde el ítem posee mayor potencia discriminatoria con respecto al nivel del comportamiento Confortar de los sujetos. El valor promedio de los parámetros  $a$  estimados en el ML2P fue de 1.45 (DE = 0.35) con un mínimo en el ítem 14 (*Comparto las malas experiencias de la vida con amigos para que no cometan los mismos errores*) y un máximo valor en el ítem 15 (*Intento “levantar” la autoestima a mis amigos*).

Para el ML3P parámetro  $b$  promedio estimado de -0.74 (DE = 0.80), el parámetro  $a$  promedio estimado fue de 0.94 (DE = 0.30) y los parámetros  $c$  fueron muy cercanos a cero, con un valor medio de 0.06 (DE = 0.13).

Tabla 20

*Parámetros estimados de ML1P, ML2P y ML3P  
para los ítems de la dimensión Confortar*

Ítem	ML1P	ML2P		ML3P		
	<i>b</i> (DE)	<i>b</i> (DE)	<i>a</i> (DE)	<i>b</i>	<i>a</i>	<i>c</i>
3	-1.72 (0.09)	-1.52 (0.10)	1.71 (0.17)	-0.91	1.47	0.36
4	-0.32 (0.06)	-0.34 (0.07)	1.26 (0.11)	-0.34	0.74	0.00
5	0.48 (0.06)	0.56 (0.08)	1.08 (0.10)	0.64	0.73	0.04
12	-0.39 (0.06)	-0.34 (0.05)	1.75 (0.12)	-0.35	1.01	0.00
14	-0.84 (0.06)	-1.09 (0.12)	0.93 (0.10)	-1.10	0.54	0.00
15	-2.10 (0.10)	-1.80 (0.11)	1.80 (0.25)	-1.78	1.09	0.00
16	-1.50 (0.08)	-1.36 (0.09)	1.64 (0.15)	-1.36	0.97	0.00
Media	-0.91	-0.84	1.45	-0.74	0.94	0.06
DE	0.91	0.83	0.35	0.80	0.30	0.13
Mínimo	-2.10	-1.80	0.93	-1.78	0.54	0.00
Máximo	0.48	0.56	1.80	0.64	1.47	0.36

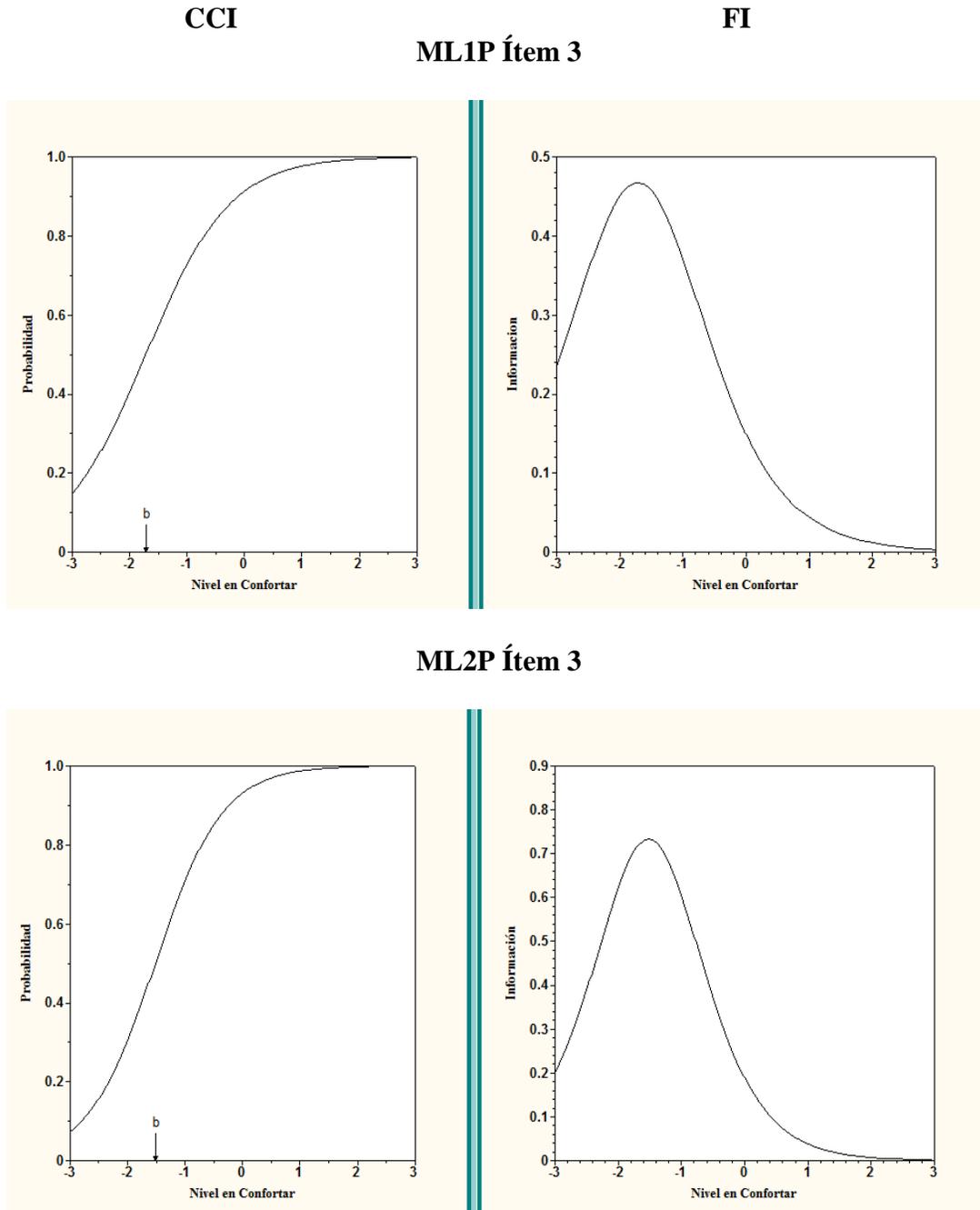
Los *b* promedio estimados son negativos para los tres modelos, indicando que se encuentran en valores relativamente bajos de  $\theta$ . Es necesario apenas un nivel bajo en el rasgo para responder los ítems de la dimensión Confortar de manera favorable. La elevada dispersión entre los *b* indicaría que los ítems de esta dimensión pueden tener capacidad discriminativa dentro de los niveles bajos del rasgo.

Los valores estimados promedio del parámetro *a* para los tres modelos muestran una capacidad discriminativa muy alta en el entorno de *b* (Reise & Waller, 1990), con mayor capacidad discriminativa estimada en el ML2P y la menor en el ML3P. Esta capacidad es relativamente homogénea entre los diferentes ítems.

La cercanía a cero de los parámetros  $c$  indicaría la poca relevancia de su inclusión dado que se trata de una escala de comportamiento típico. Es sólo en el ítem 3 (*Cuando siento que alguien está mal, le demuestro que lo entiendo*) donde se presenta el valor más elevado: 0.36. Para descartar la influencia de la deseabilidad social, la cual es coherente con el contenido del ítem que indica una capacidad básica de empatía, se correlacionaron los puntajes en este ítem con el puntaje en la Escala de Distorsión del BFQ. Dicha correlación no resultó significativa ( $r = .026, p = .33$ ). También se descartó mediante regresiones la existencia de relaciones no lineales. Por lo cual, se concluye que el valor del parámetro  $c$  para este ítem podría deberse a un efecto aleatorio.

Figura 21

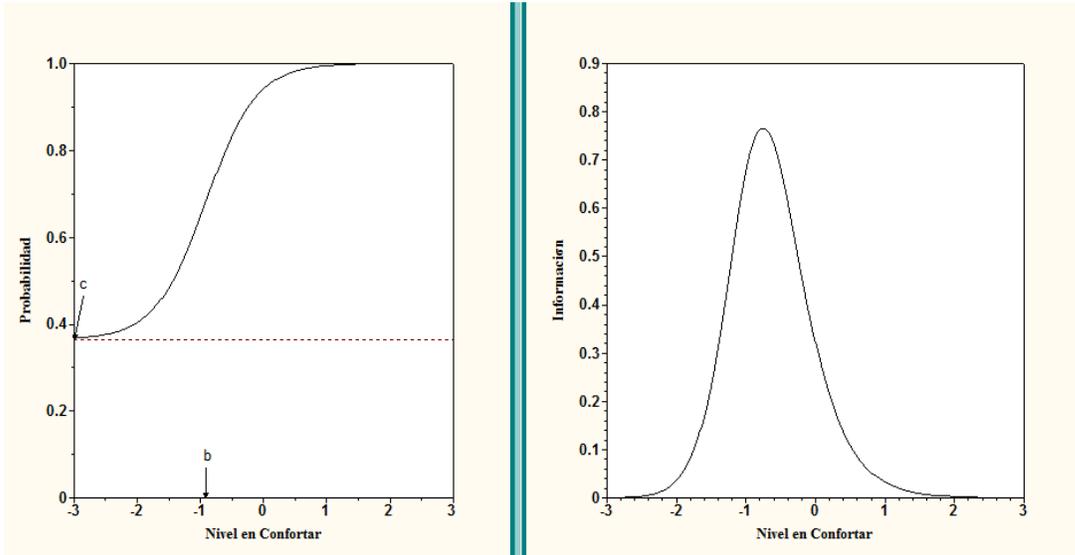
CCIs y FIs de los ítems de la dimensión Confortar para ML1P, ML2P y ML3P



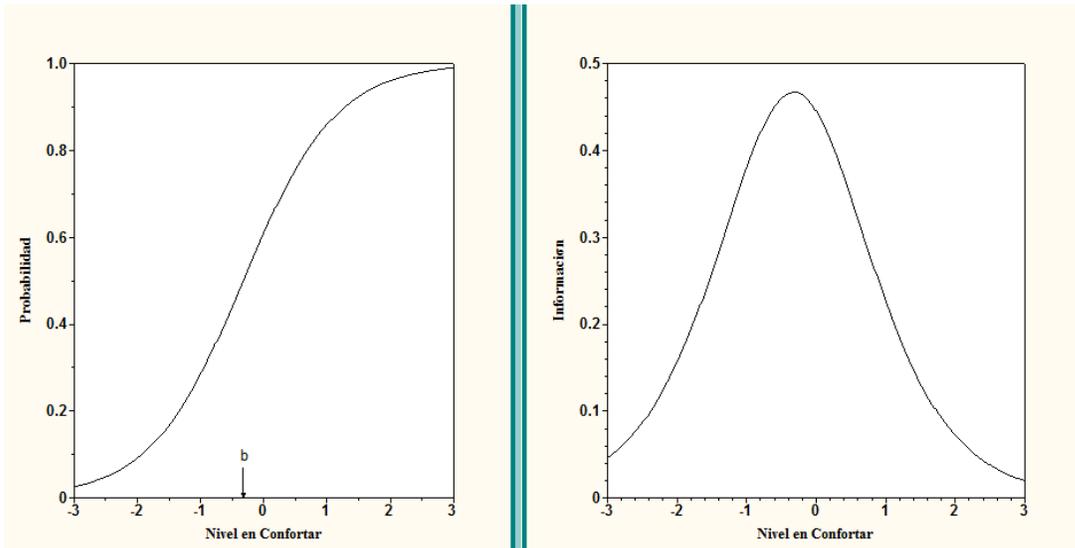
CCI

FI

ML3P Ítem 3



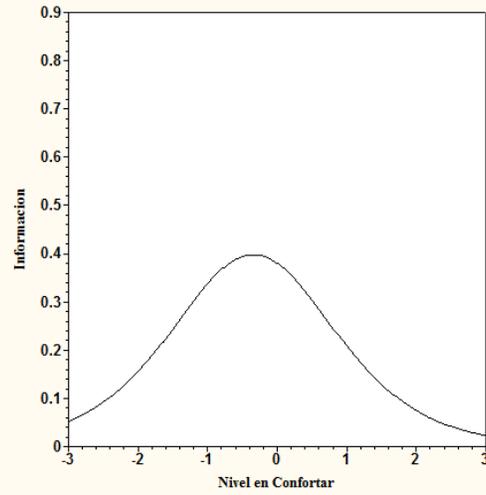
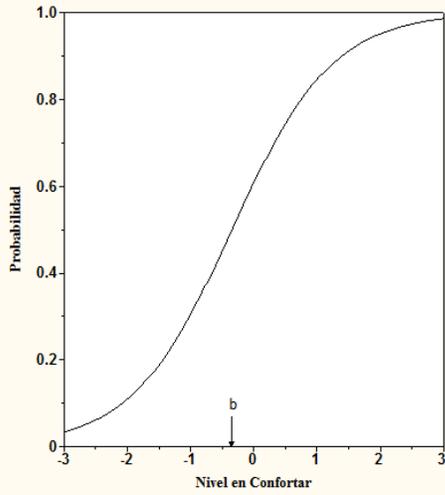
ML1P Ítem 4



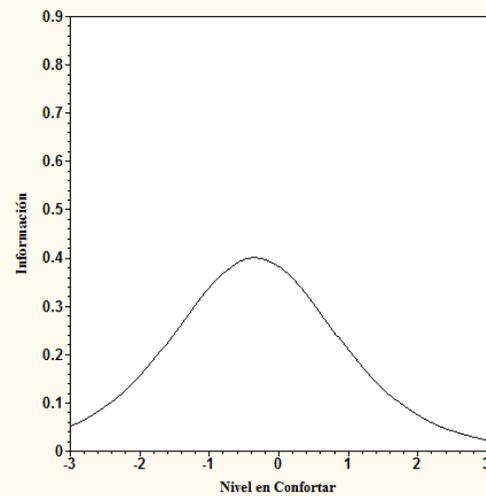
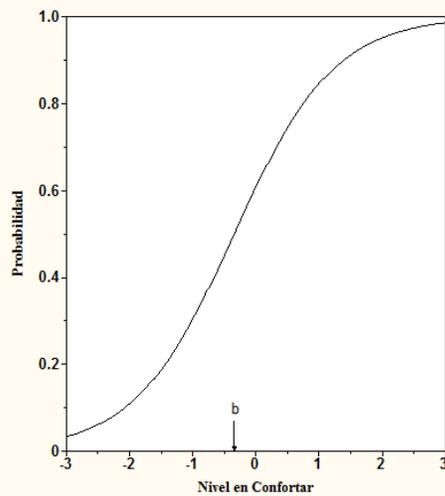
CCI

FI

ML2P Ítem 4



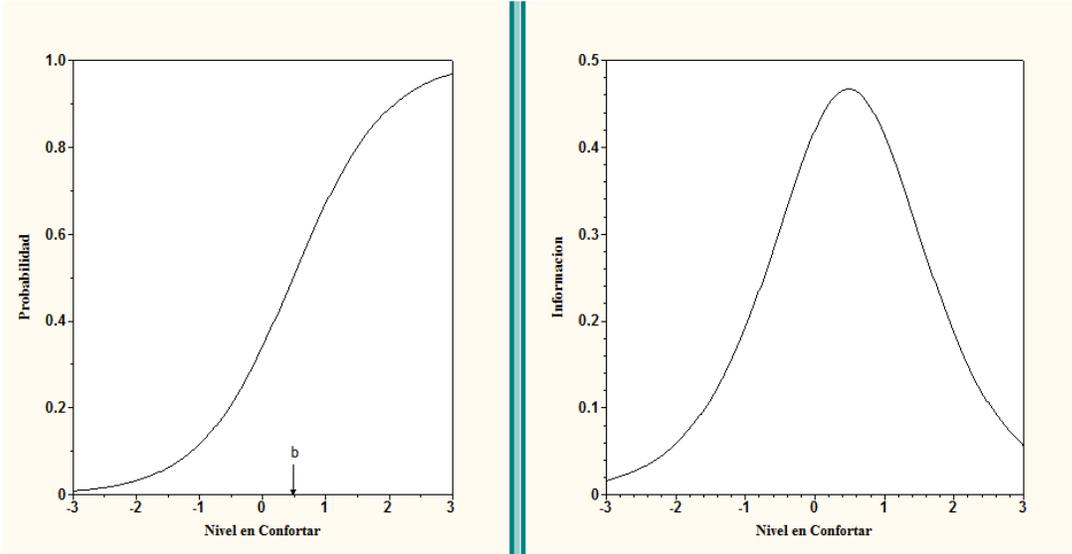
ML3P Ítem 4



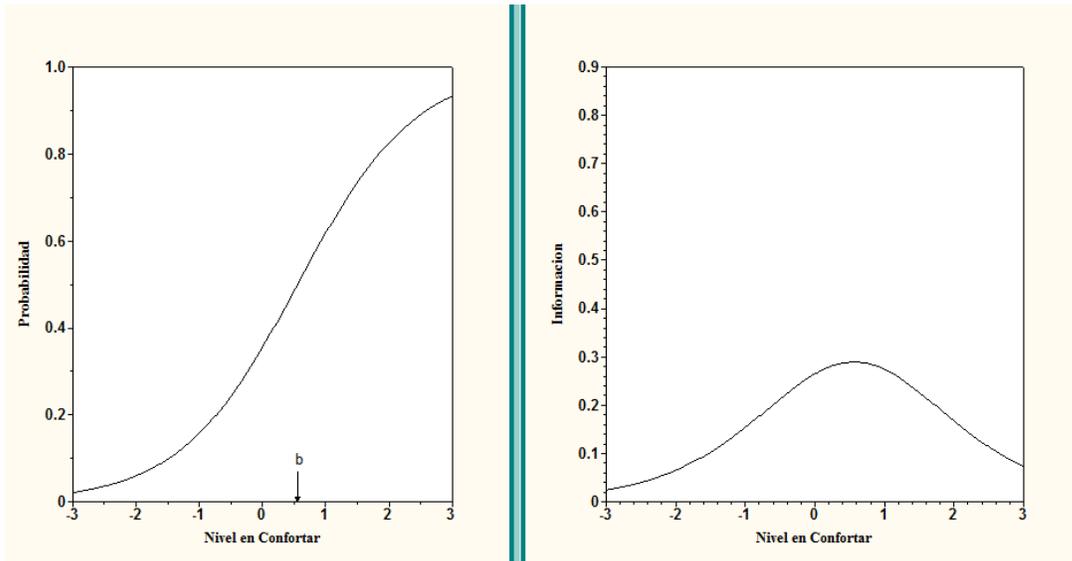
CCI

ML1P Ítem 5

FI



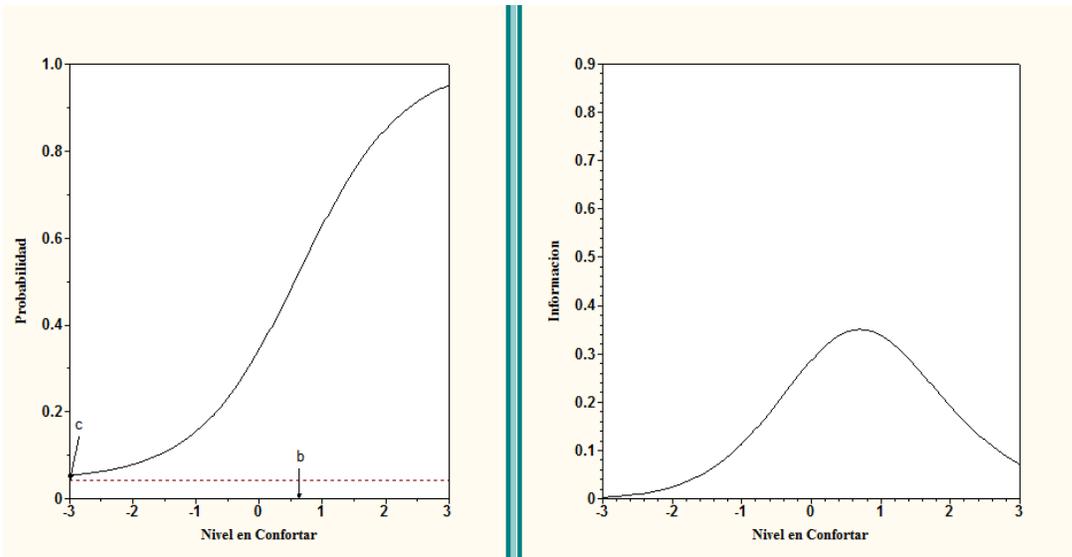
ML2P Ítem 5



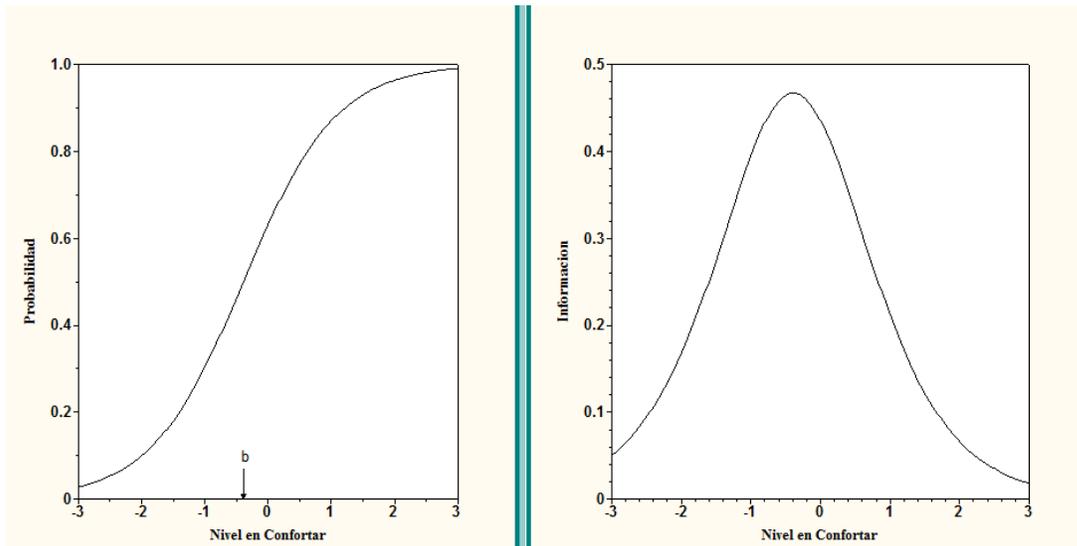
CCI

FI

ML3P Ítem 5



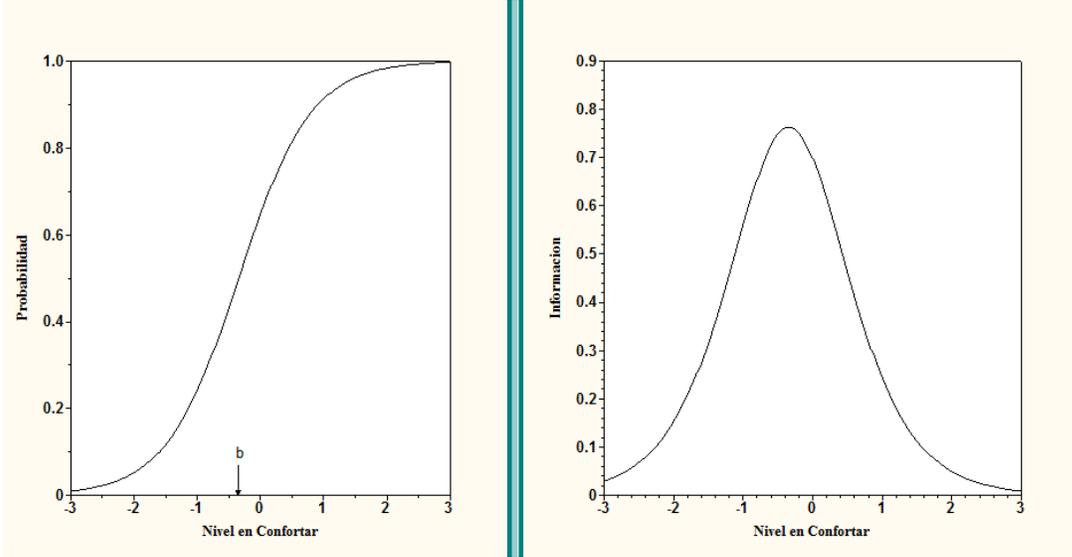
ML1P Ítem 12



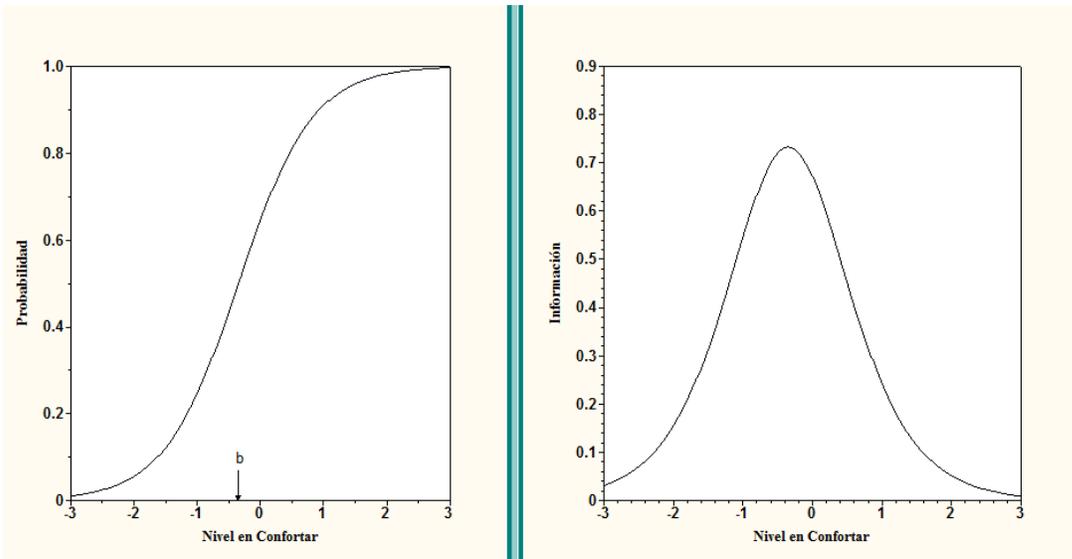
CCI

ML2P Ítem 12

FI



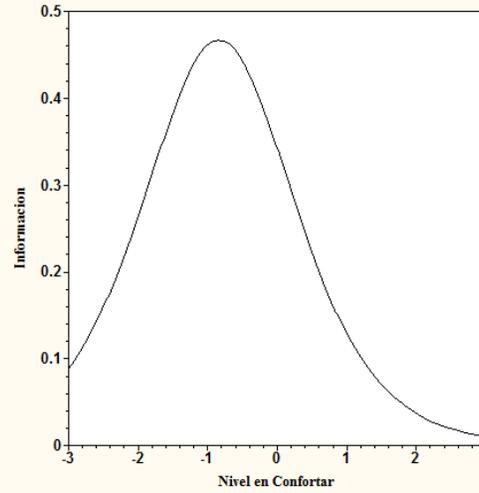
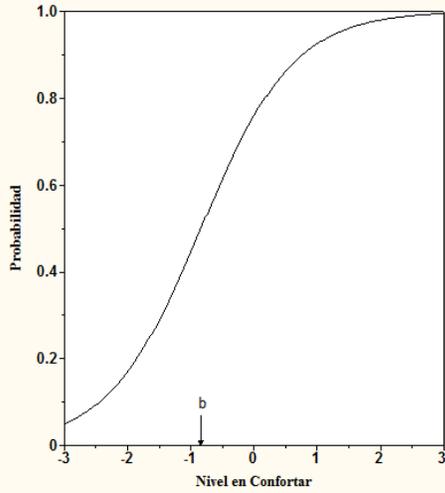
ML3P Ítem 12



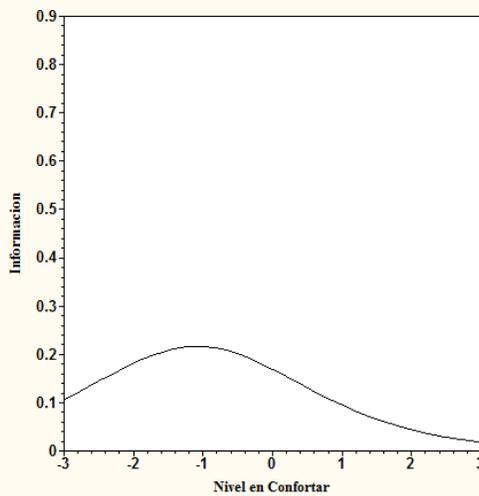
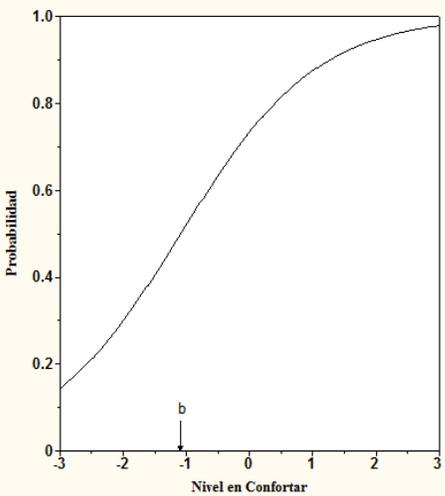
CCI

ML1P Ítem 14

FI



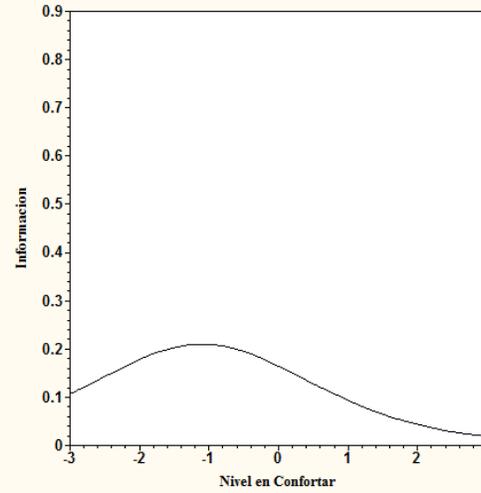
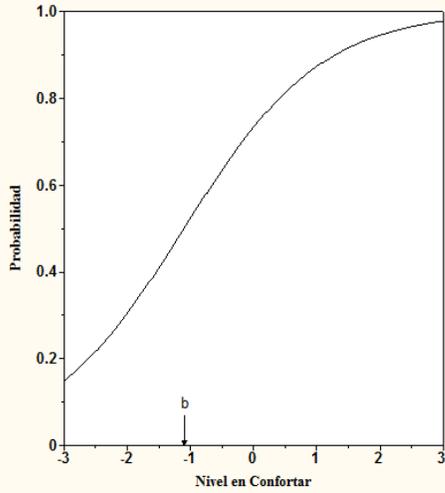
ML2P Ítem 14



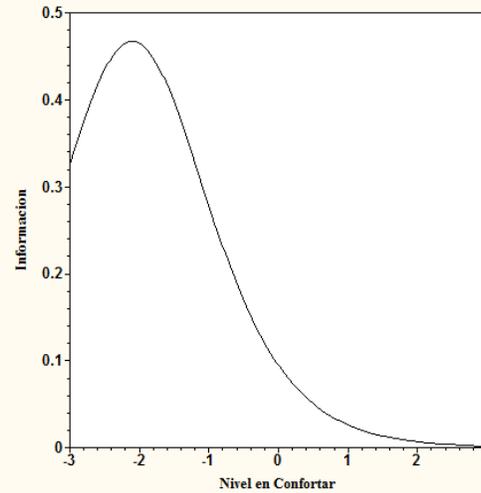
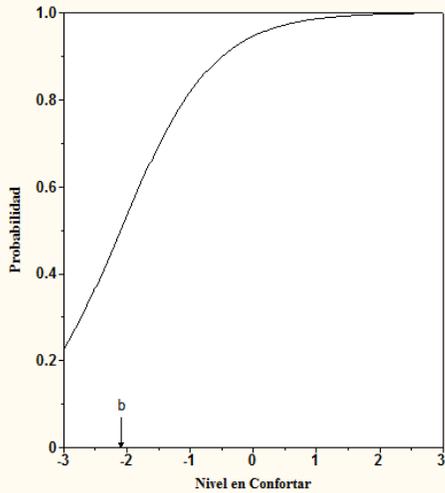
CCI

FI

ML3P Ítem 14



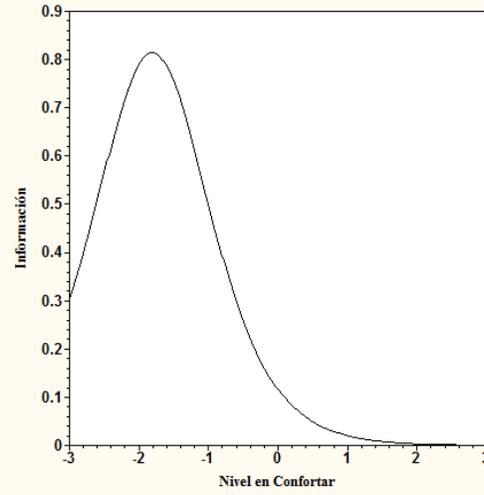
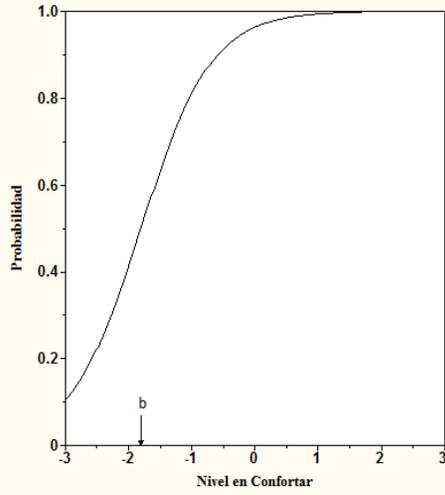
ML1P Ítem 15



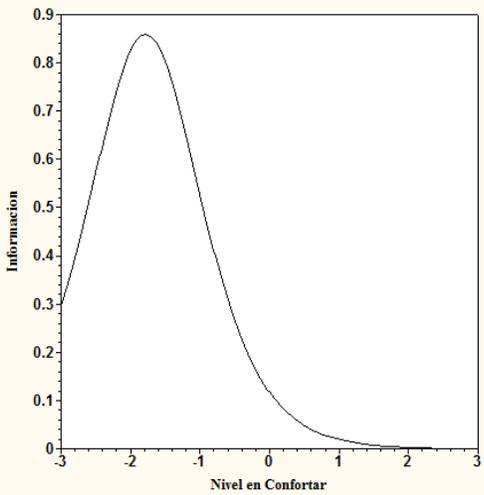
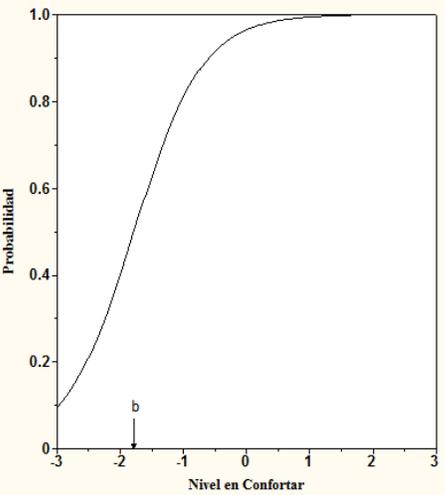
CCI

ML2P Ítem 15

FI



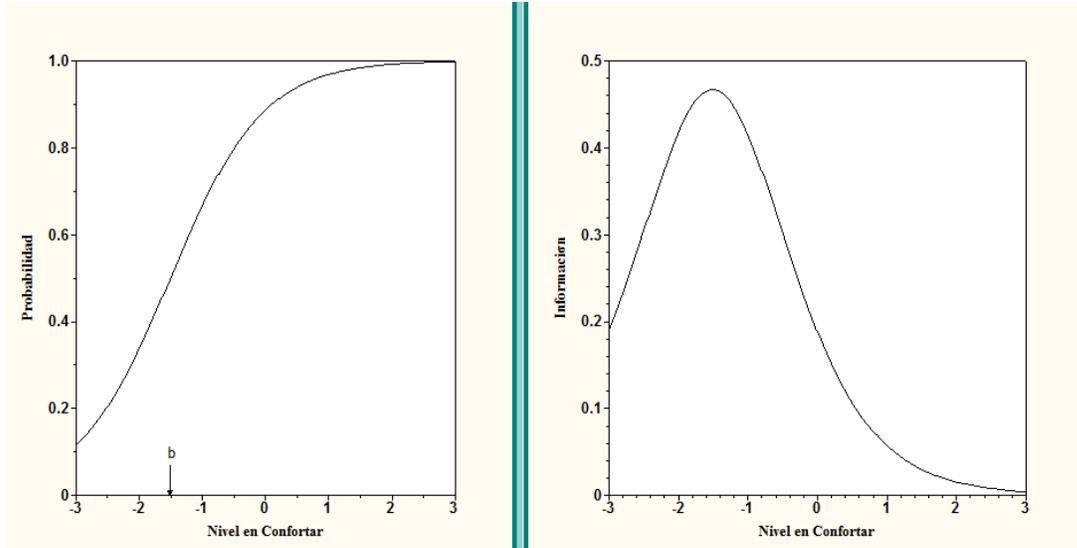
ML3P Ítem 15



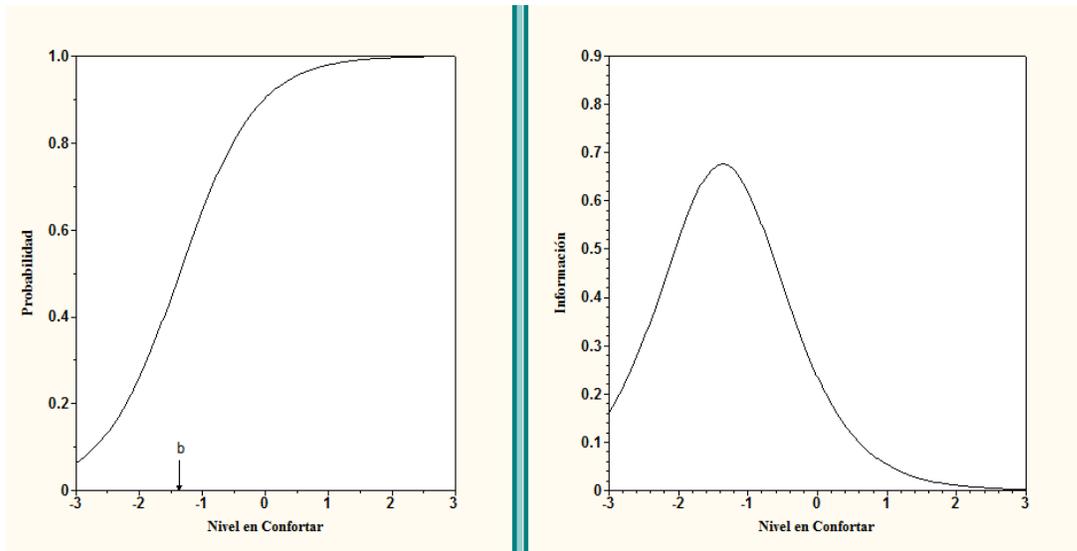
CCI

FI

ML1P Ítem 16



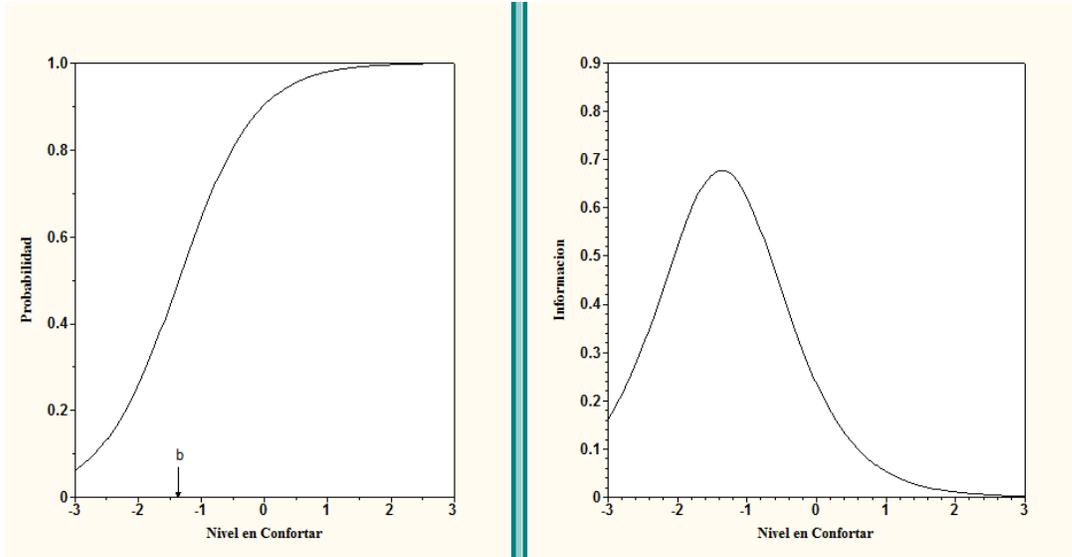
ML2P Ítem 16



CCI

FI

ML3P Ítem 16



Las Tablas 21, 22 y 23 exhiben los valores que alcanzan las FIs de los ítems y del test para algunos niveles del rasgo Confortar en los modelos dicotómicos analizados.

Tabla 21

*Funciones de Información de los Ítems y del Test obtenidas mediante ML1P*

*en la Dimensión Confortar*

Ítem	$\theta$						
	$\theta = -3$	$\theta = -2$	$\theta = -1$	$\theta = 0$	$\theta = 1$	$\theta = 2$	$\theta = 3$
1	0.24	0.45	0.37	0.15	0.04	0.02	0.00+
2	0.05	0.16	0.38	0.45	0.23	0.07	0.02
6	0.02	0.06	0.19	0.42	0.37	0.19	0.06
7	0.05	0.17	0.39	0.44	0.21	0.07	0.02

8	0.09	0.26	0.46	0.34	0.13	0.04	0.01
9	0.33	0.47	0.28	0.09	0.03	0.01	0.00+
10	0.19	0.42	0.42	0.19	0.06	0.02	0.00+
Test	1.95	2.98	3.50	3.07	2.11	1.40	1.11
DE	0.72	0.58	0.54	0.57	0.69	0.85	0.95

Tabla 22

*Funciones de Información de los Ítems y del Test obtenidas mediante ML2P  
en la Dimensión Confortar*

Ítem \ $\theta$	$\theta$						
	$\theta = -3$	$\theta = -2$	$\theta = -1$	$\theta = 0$	$\theta = 1$	$\theta = 2$	$\theta = 3$
1	0.20	0.62	0.60	0.19	0.04	0.01	0.00+
2	0.05	0.16	0.34	0.38	0.21	0.07	0.02
6	0.02	0.06	0.15	0.27	0.27	0.17	0.07
7	0.03	0.15	0.56	0.70	0.24	0.05	0.01
8	0.11	0.18	0.22	0.17	0.09	0.04	0.02
9	0.30	0.79	0.50	0.12	0.02	0.00+	0.00+
10	0.16	0.52	0.62	0.24	0.05	0.01	0.00+
Test	1.87	3.48	3.99	3.05	1.93	1.36	1.13
DE	0.73	0.54	0.50	0.57	0.72	0.86	0.94

Tabla 23

*Funciones de Información de los Ítems y del Test obtenidas mediante ML3P  
en la Dimensión Confortar*

Ítem \ $\theta$	$\theta$						
	$\theta = -3$	$\theta = -2$	$\theta = -1$	$\theta = 0$	$\theta = 1$	$\theta = 2$	$\theta = 3$
1	0.00	0.04	0.67	0.32	0.03	0.00+	0.00+
2	0.05	0.16	0.34	0.38	0.21	0.08	0.02
6	0.00+	0.02	0.11	0.29	0.34	0.19	0.07
7	0.03	0.15	0.54	0.67	0.24	0.05	0.01
8	0.11	0.18	0.21	0.16	0.09	0.04	0.02
9	0.30	0.82	0.53	0.12	0.02	0.00+	0.00+
10	0.16	0.52	0.62	0.24	0.05	0.01	0.00+
Test	1.65	2.89	4.03	3.18	1.99	1.38	1.12
DE	0.78	0.59	0.49	0.56	0.71	0.85	0.94

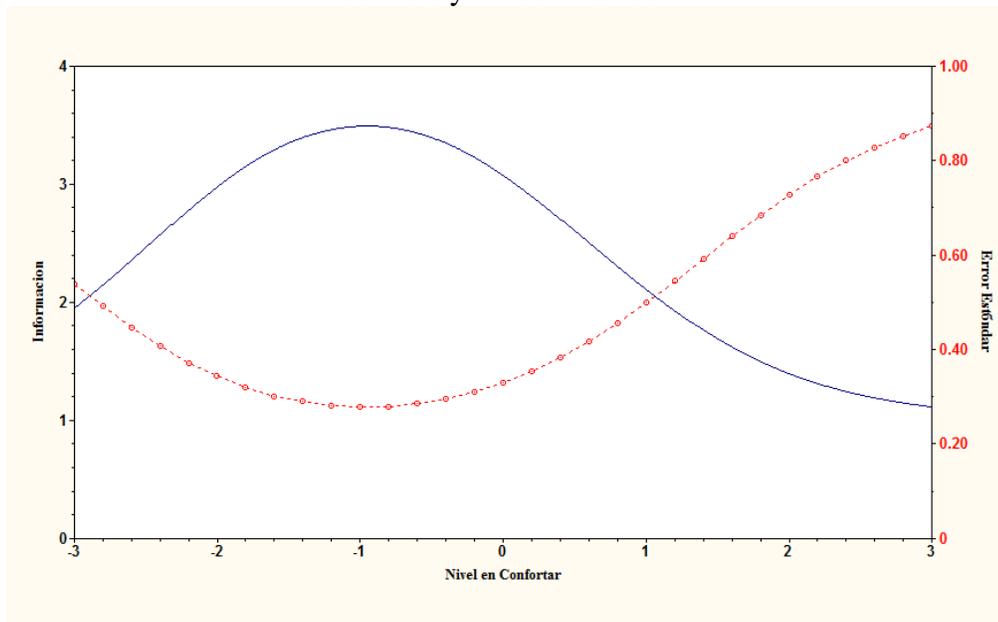
La Figura 22 muestra que el gráfico de la FI del test alcanzó para el ML1P su valor máximo de 3.49 en  $\theta = -1$  con un valor mínimo del error estándar de medida (EE) en ese punto de 0.54. Para el ML2P el máximo punto de la FI de 4.06 se ubicó en  $\theta = -1.2$  donde el EE fue de 0.50. En el caso del ML3P el valor máximo de la FI de 4.08 se alcanzó en  $\theta = -0.8$  con un EE en dicho punto de 0.49.

Las FI de ML1P, ML2P y ML3P muestran que el test en su versión de respuesta dicotómica proporciona estimaciones en Confortar más precisas en un rango de rasgo de -2 a 0, mientras que el error en la medición crece hacia los niveles de rasgo menores a -2 y mayores a 0.

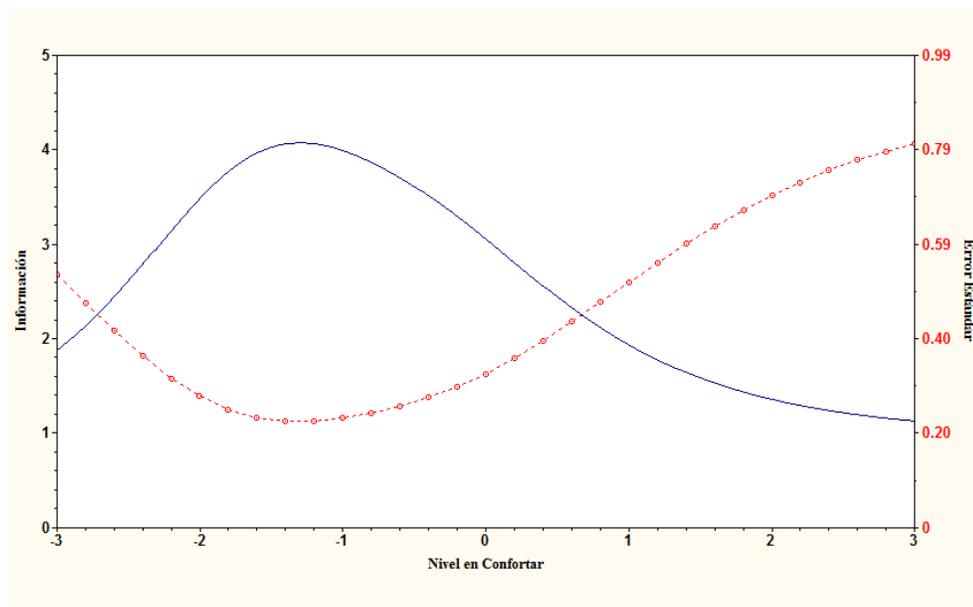
Figura 22

*Función de Información del Test y Error Estándar de Medida  
para ML1P, ML2P y ML3P en la dimensión Confortar*

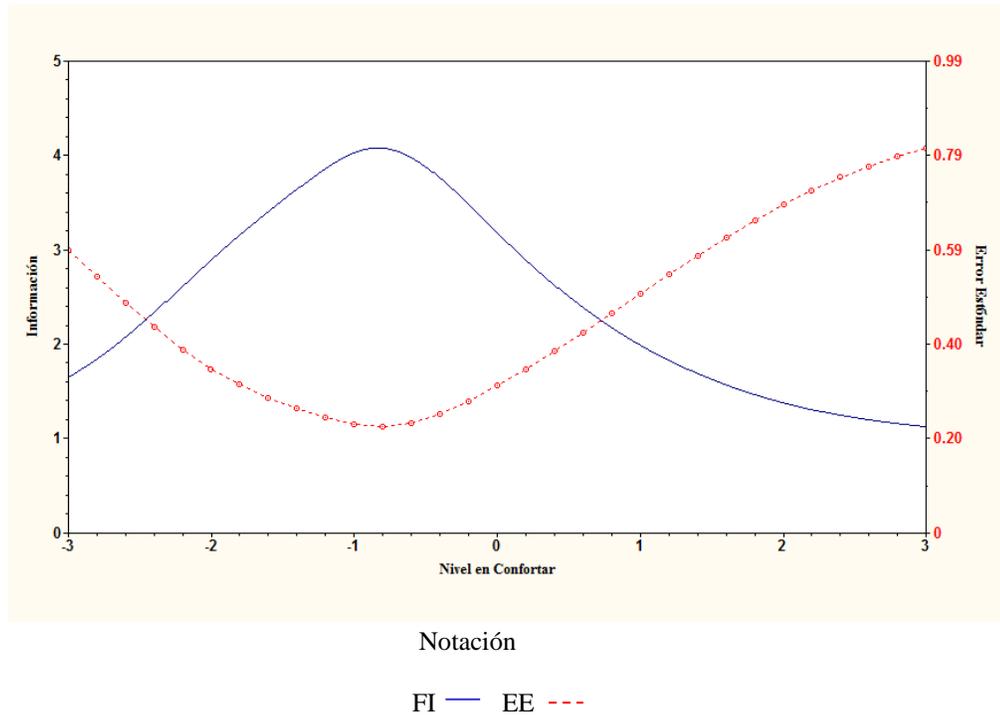
FI del Test y EE de Medida ML1P



FI del Test y EE de Medida ML2P



## FI del Test y EE de Medida ML3P



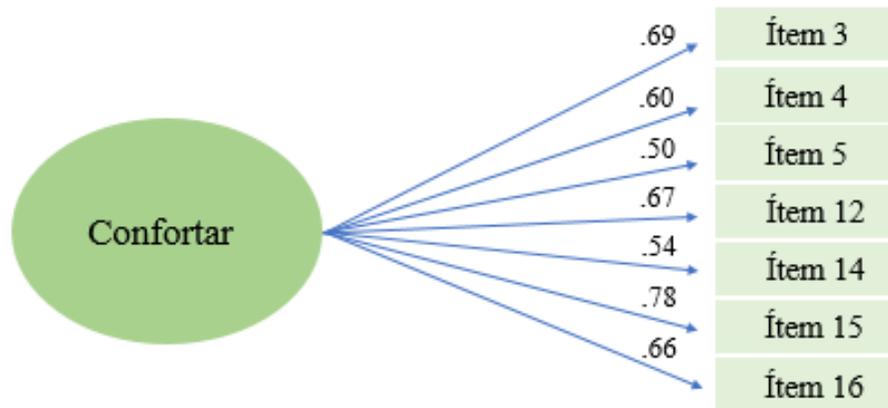
### *Cumplimiento de los supuestos de los Modelos de Respuesta Graduada y Crédito Parcial en la Dimensión Confortar*

Se implementó un AFC con el programa Mplus (Muthén & Muthén, 2010) sobre los siete ítems que componen la dimensión Confortar para verificar el supuesto de unidimensionalidad requerido por los modelos de la TRI. Se respetó el carácter ordinal de los datos por lo que se estimaron los parámetros con el método WLSMV sobre la base de las matrices de correlaciones policóricas. Se analizó el ajuste al modelo considerando los índices CFI, TLI y RMSEA. Este análisis fue realizado sobre la totalidad de los participantes de la muestra ( $n = 1383$ ).

Los indicadores de ajuste comparativos muestran un resultado muy adecuado para la verificación del modelo unidimensional (CFI = .977; TLI = .966; RMSEA = .069, IC 90% .057-.081), por lo cual se acepta el cumplimiento del supuesto de unidimensionalidad para la dimensión Confortar. Tal como se observa en la Figura 23, las cargas factoriales son elevadas (Abad et al., 2011).

Figura 23

*Modelo Factorial Confirmatorio de la Dimensión Confortar*



El supuesto de independencia local se dio por satisfecho al comprobarse la unidimensionalidad de la escala (Lord & Novick, 1968).

*Aplicación del Modelo de Respuesta Graduada a la Dimensión Confortar*

*Ajuste de los datos al Modelo de Respuesta Graduada*

*en la Dimensión Confortar*

Las proporciones esperadas fueron transformadas a frecuencias absolutas ( $n = 1383$ ) para poder aplicar la prueba ji cuadrado. El modelo presentó un adecuado ajuste ya que no se rechazó la hipótesis nula de igualdad entre frecuencias absolutas observadas y esperadas según el MRG para ninguno de los ítems.

Tabla 24

*Índices de Ajuste de los Ítems de la Dimensión Confortar  
al Modelo de Respuesta Graduada*

Ítem	Frecuencias Absolutas	<i>Nunca</i>	<i>Casi Nunca</i>	<i>A Veces</i>	<i>Con Frecuencia</i>	<i>Casi Siempre</i>	<i>Siempre</i>	$\chi^2$	<i>p</i>
3	Observadas	2	29	167	264	368	553	0.31	.9974
	Esperadas	2	29	173	264	362	553		
4	Observadas	60	102	413	398	255	155	0.34	.9968
	Esperadas	59	103	409	393	262	157		
5	Observadas	86	191	580	236	172	118	0.15	.9995
	Esperadas	86	192	573	235	175	119		
12	Observadas	40	137	378	323	246	259	0.30	.9976
	Esperadas	40	139	375	317	246	265		
14	Observadas	31	68	312	321	276	375	0.43	.9944
	Esperadas	30	67	304	320	282	379		
15	Observadas	3	8	125	234	310	703	0.57	.9891
	Esperadas	2	8	127	236	311	699		
16	Observadas	12	38	190	245	329	569	0.21	.9989
	Esperadas	11	39	193	245	325	570		

*Estimación de Parámetros del Modelo de Respuesta Graduada para los ítems  
de la Dimensión Confortar*

Los 42 parámetros fueron estimados en 30 ciclos con un criterio de convergencia de .001 y un máximo cambio interciclo de .00075. Como se observa en la Tabla 25, para cada uno de los ítems se estimó un parámetro de inclinación  $a$ , así como cinco parámetros de localización  $b_1$ ,  $b_2$ ,  $b_3$ ,  $b_4$  y  $b_5$ .

El parámetro  $b_1$  indica el valor de umbral que separa las categorías de respuesta *Nunca* de *Casi Nunca*. Todos los parámetros  $b_1$  son en extremo negativos, indicando que es suficiente un nivel bajísimo en el rasgo para tener una probabilidad mayor a 0.50 de elegir la categoría *Casi Nunca* o una superior. Es decir, que resulta una elección superlativamente fácil: es realmente improbable que una persona no la realice. Son conductas muy frecuentes en la población donde se ha tomado la prueba.

El parámetro  $b_2$  indica el valor de umbral que separa las categorías de respuesta *Casi Nunca* de *A Veces*. Es el mínimo  $\theta$  necesario para tener una probabilidad de 0.50 de elegir la categoría *A Veces* o una superior. En todos los ítems los  $b_2$  fueron negativos, pero con una distancia importante con respecto a los  $b_1$ . Esto denota que para obtener  $b_2$  es necesario un nivel en el rasgo mayor que para obtener  $b_1$ , pero alcanza con un nivel muy bajo del rasgo.

El parámetro  $b_3$  indica el valor de umbral que separa las categorías de respuesta *A Veces* de *Con Frecuencia*. Es el mínimo  $\theta$  necesario para tener una probabilidad de 0.50 de elegir la categoría *Con Frecuencia* o una superior. Se observa una importante variabilidad en los ítems de la dimensión Confortar con respecto a los valores de los  $b_3$ , aunque la mayoría son valores negativos, con la excepción del  $b_3$  del ítem 5 (*Siento todo dolor ajeno como propio*). Para la mayoría de los ítems, continúa siendo una elección fácil, aunque se necesita un nivel de rasgo sustantivamente más alto para seleccionar la categoría *Con Frecuencia* o una mayor con respecto al necesario para seleccionar la categoría *A Veces* o una mayor.

El parámetro  $b_4$  indica el valor de umbral que separa las categorías de respuesta *Con Frecuencia* de *Casi Siempre*. O sea que es el mínimo  $\theta$  necesario para tener una probabilidad de 0.50 de elegir las categorías *Casi Siempre* o *Siempre*. Se observan valores muy diferentes entre los ítems en los  $b_4$  obtenidos que indican desde bajos niveles en el rasgo hasta relativamente altos.

El parámetro  $b_5$  indica el valor de umbral que separa las categorías de respuesta *Casi Siempre* de *Siempre*, el mínimo  $\theta$  necesario para tener una probabilidad de 0.50 de elegir la categoría *Siempre*. Para obtener este valor de umbral es necesario un nivel medio a alto en el rasgo para tener una probabilidad de 0.50 de elegir la categoría *Siempre*.

Los reactivos recorren efectivamente el rasgo latente en un amplio espectro, desde -4.82 ( $b_1$  ítem 3, *Cuando siento que alguien está mal, le demuestro que lo entiendo*) hasta 2.69 ( $b_5$  ítem 5, *Siento todo dolor ajeno como propio*). Esta tendencia no es equilibrada por encima y por debajo de la media de  $\theta$  ya que los parámetros  $b_m$  son negativos (a excepción del  $b_3$  del ítem 5) hasta el paso 3 inclusive. Se puede observar que las distancias entre los  $b_m$  son amplias indicando la discriminación de los ítems en los distintos niveles del rasgo y la pertinencia de la escala Likert de seis categorías para la respuesta.

Los parámetros de inclinación  $a$  tuvieron un valor promedio estimado de 1.51 (DE = 0.38) que representa una capacidad muy alta de las categorías de respuesta para distinguir entre los

diferentes niveles de rasgo (Reise & Waller, 1990). El ítem 15 (*Intento “levantar” la autoestima a mis amigos*) tuvo el parámetro de inclinación  $a$  más elevado con un valor de 2.16 (DE = 0.12).

Tabla 25

*Parámetros estimados mediante MRG para los ítems de la dimensión Confortar*

Ítem	$b_1$ (DE)	$b_2$ (DE)	$b_3$ (DE)	$b_4$ (DE)	$b_5$ (DE)	$b_{prom}$	$a$ (DE)
3	-4.82 (0.60)	-2.97 (0.19)	-1.50 (0.08)	-0.59 (0.06)	0.36 (0.06)	-1.90	1.69 (0.10)
4	-2.95 (0.20)	-1.99 (0.13)	-0.36 (0.06)	0.84 (0.08)	2.00 (0.13)	-0.49	1.30 (0.08)
5	-3.06 (0.24)	-1.60 (0.13)	0.57 (0.08)	1.53 (0.12)	2.69 (0.20)	0.03	1.03 (0.07)
12	-2.84 (0.17)	-1.64 (0.09)	-0.36 (0.05)	0.48 (0.05)	1.26 (0.08)	-0.62	1.64 (0.09)
14	-3.84 (0.31)	-2.70 (0.20)	-0.98 (0.09)	0.10 (0.07)	1.06 (0.10)	-1.26	1.15 (0.08)
15	-4.04 (0.39)	-3.27 (0.25)	-1.66 (0.08)	-0.79 (0.05)	-0.02 (0.04)	-1.96	2.16 (0.12)
16	-3.80 (0.31)	-2.74 (0.17)	-1.38 (0.08)	-0.55 (0.06)	0.33 (0.06)	-1.31	1.58 (0.09)
Media	-3.62	-2.42	-0.81	0.15	1.10		1.51
DE	0.71	0.66	0.80	0.86	0.98		0.38
Mínimo	-4.82	-3.27	-1.66	-0.79	-0.02		1.03
Máximo	-2.84	1.60	0.57	1.53	2.69		2.16

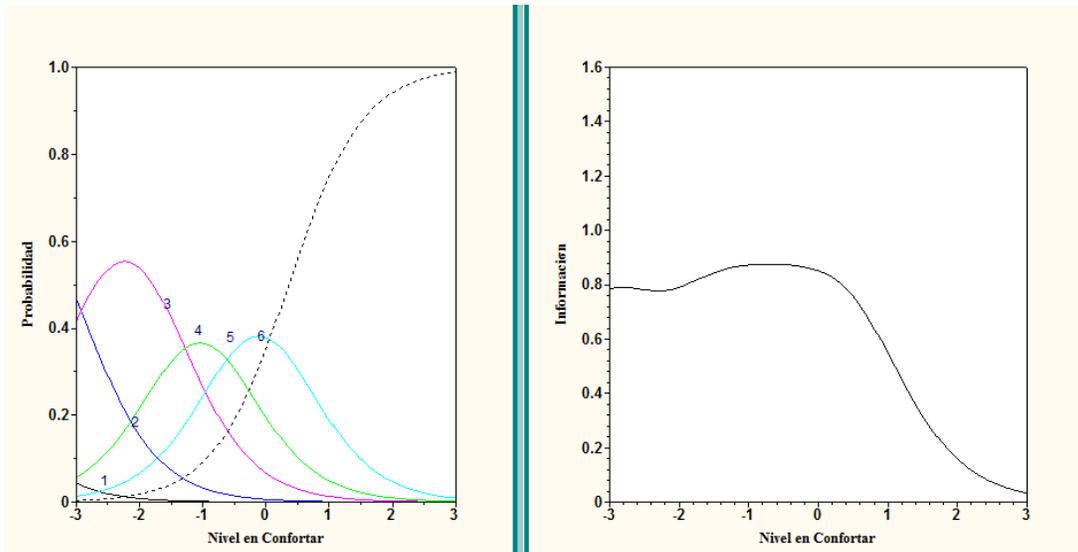
Figura 24

CCCRIs y FIs de los Ítems de la Dimensión Confortar para el MRG

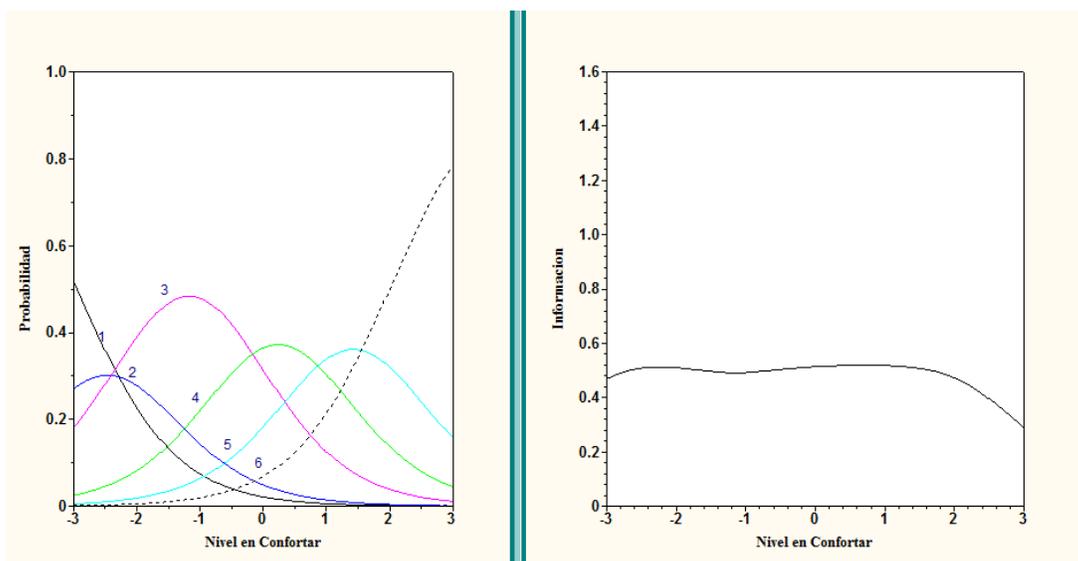
CCCRI

FI

Ítem 3



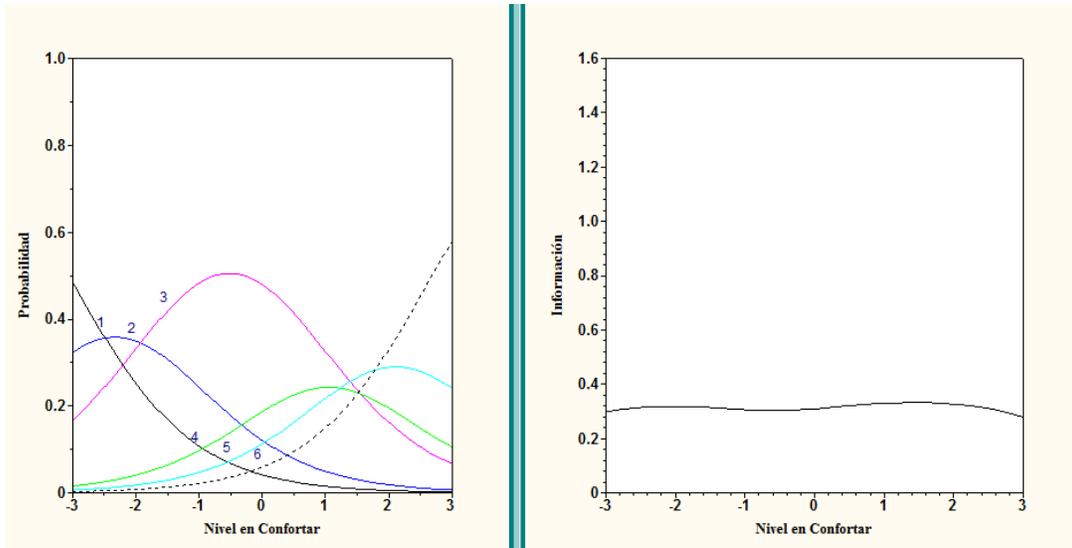
Ítem 4



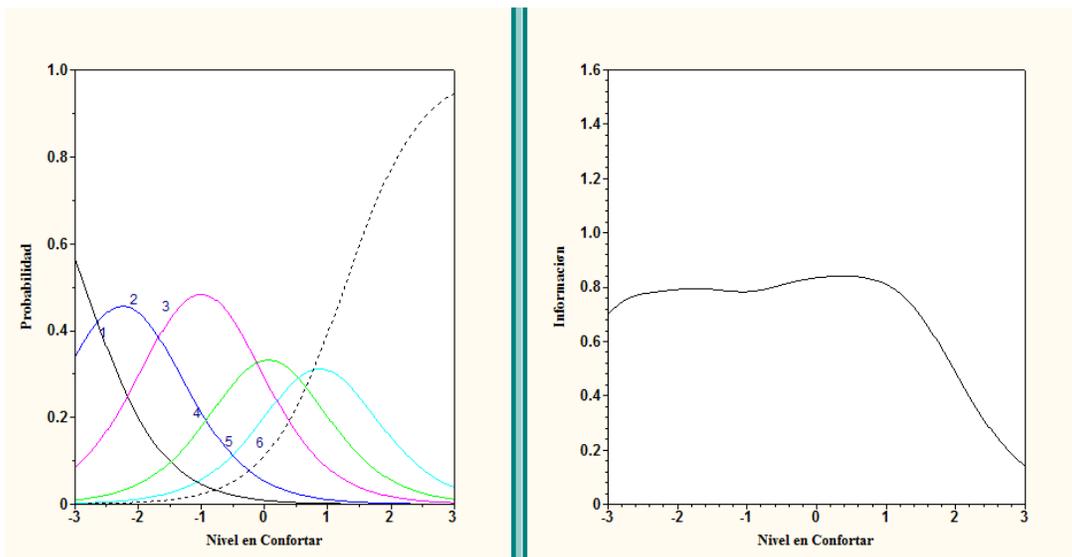
CCCRI

FI

Ítem 5



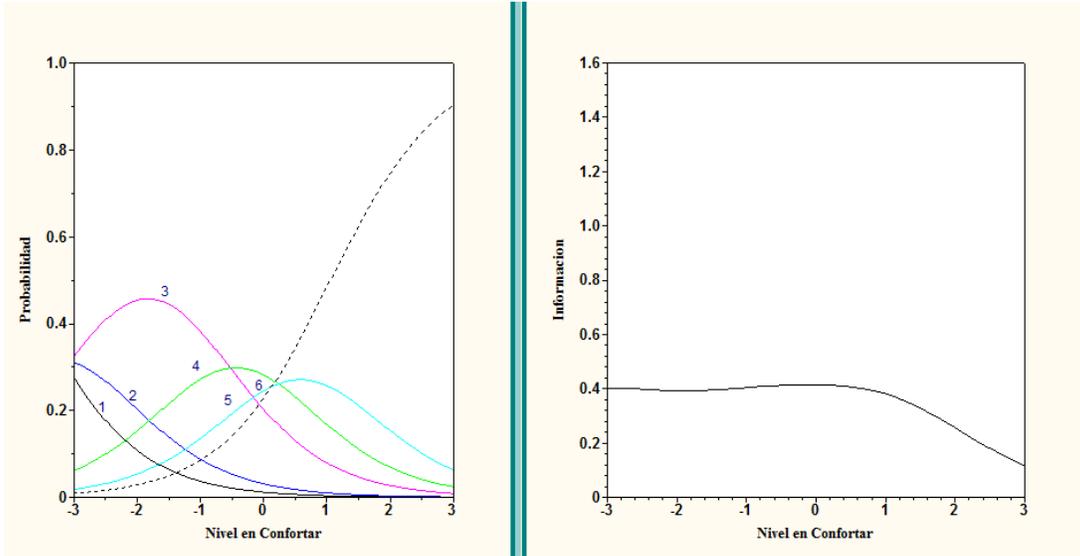
Ítem 12



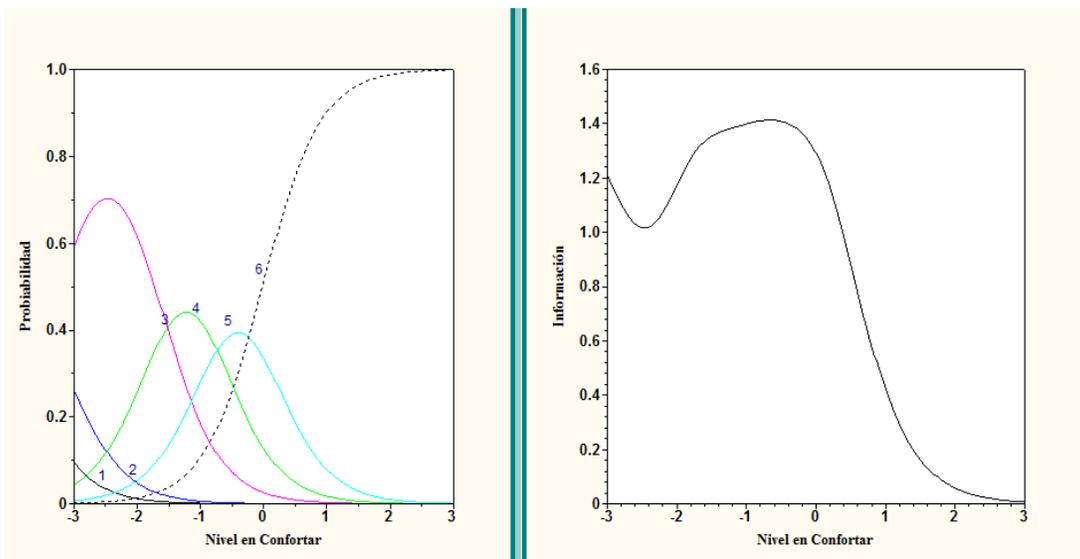
CCCRI

FI

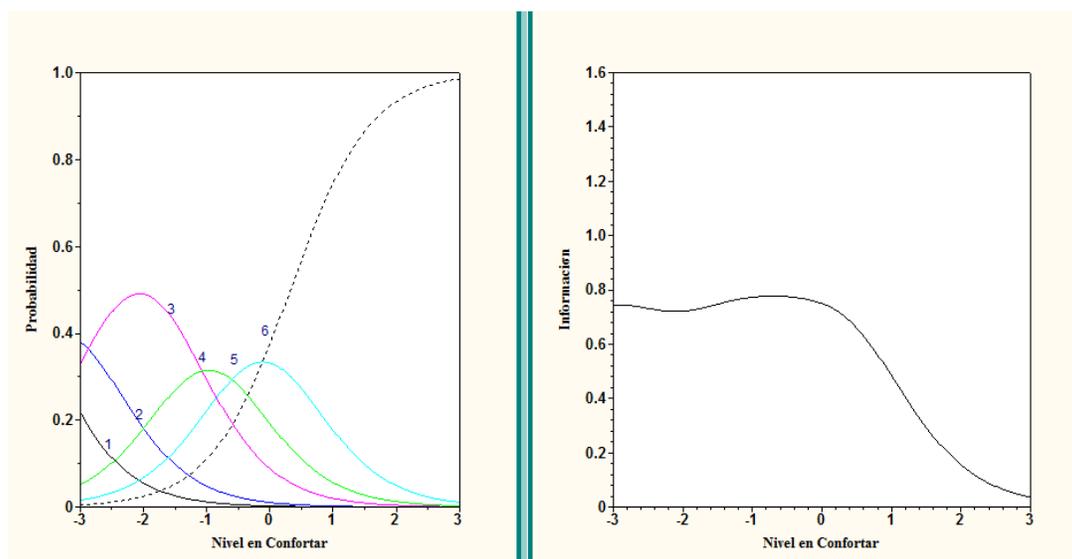
Ítem 14



Ítem 15



## Ítem 16



## Notación

CCCRI Categoría 1 = <i>Nunca</i>	—
CCCRI Categoría 2 = <i>Casi Nunca</i>	—
CCCRI Categoría 3 = <i>A veces</i>	—
CCCRI Categoría 4 = <i>Con Frecuencia</i>	—
CCCRI Categoría 5 = <i>Casi Siempre</i>	—
CCCRI Categoría 6 = <i>Siempre</i>	---

La Tabla 26 exhibe los valores que alcanzan las FIs de los ítems y del test para algunos niveles del rasgo Confortar. Los ítems 5 (*Siento todo dolor ajeno como propio*) y 14 (*Comparto las malas experiencias de la vida con amigos para que no cometan los mismos errores*) presentaron una FI con valores relativamente bajos y uniformes a lo largo de todos los niveles del rasgo. El ítem 5 presentó la curva más plana y el nivel de información más bajo. Este bajo nivel de información es explicado por tener el parámetro de discriminación más pequeño. El ítem 15 tiene

una FI elevada, aportando un alto grado de información desde los niveles bajos hasta los niveles medios del rasgo.

Tabla 26

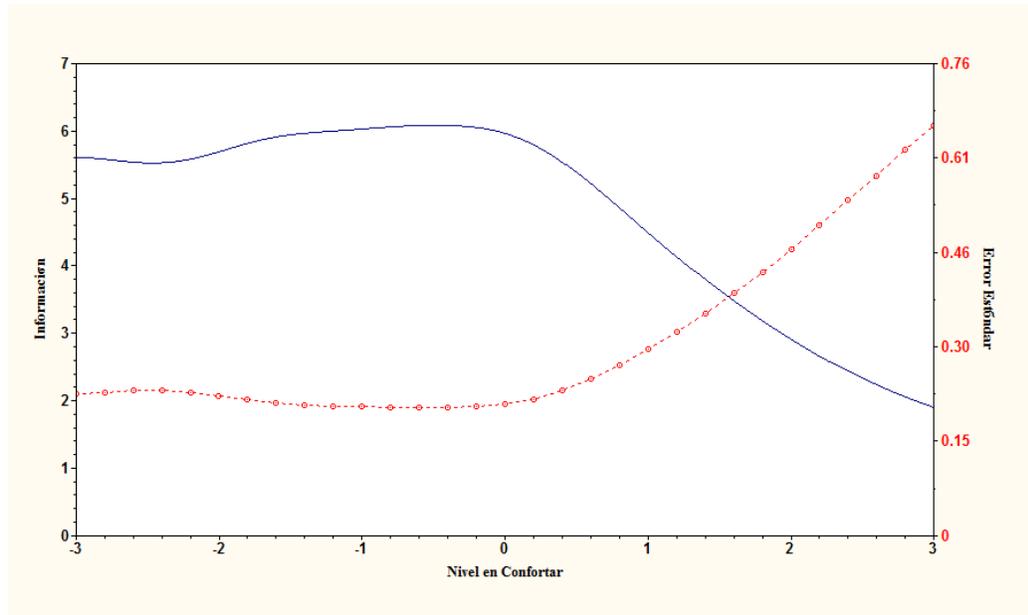
*Funciones de Información de los Ítems y del Test en la Dimensión Confortar*

Ítem	$\theta$						
	$\theta = -3$	$\theta = -2$	$\theta = -1$	$\theta = 0$	$\theta = 1$	$\theta = 2$	$\theta = 3$
3	0.79	0.79	0.87	0.85	0.55	0.16	0.03
4	0.47	0.51	0.49	0.51	0.52	0.47	0.29
5	0.30	0.32	0.33	0.31	0.33	0.33	0.28
12	0.70	0.79	0.82	0.84	0.78	0.49	0.14
14	0.40	0.39	0.40	0.41	0.38	0.26	0.12
15	1.21	1.17	1.40	1.29	0.42	0.06	0.01
16	0.75	0.72	0.77	0.75	0.48	0.16	0.04
Test	5.61	5.69	6.03	5.96	4.49	2.91	1.90
DE	0.42	0.42	0.41	0.41	0.47	0.58	0.73

La Figura 25 muestra que la FI del test resultó ligeramente decreciente y alcanzó su valor máximo de 6.08 en  $\theta = -0.4$ . Como es de esperar dada su relación inversa, el error de medida alcanzó en este punto su valor mínimo de 0.40. Esta función muestra que el test proporciona estimaciones de Confortar más precisas en los niveles de rasgo theta de -3 a 1, mientras que el error en la medición crece hacia los niveles mayores a una del rasgo Confortar. Aunque la escala politómica puede aportar información en un amplio espectro es mucho más precisa para aquellas personas con un nivel en Confortar bajo o intermedio.

Figura 25

*Función de Información del Test y Error Estándar en la Dimensión Confortar*



Notación

FI — EE - - -

*Aplicación del Modelo de Crédito Parcial a la Dimensión Confortar*

*Ajuste de los datos al Modelo de Crédito Parcial en la Dimensión Confortar*

Las proporciones esperadas fueron transformadas a frecuencias absolutas ( $n = 1383$ ) para poder aplicar la prueba ji cuadrado. El modelo presentó un adecuado ajuste ya que no se rechazó la hipótesis nula de igualdad entre proporciones observadas y esperadas según el MCP para ninguno de los ítems.

Tabla 27

*Índices de Ajuste de los Ítems de la Dimensión Confortar  
para el Modelo de Crédito Parcial*

Ítem	Frecuencias Absolutas	<i>Nunca</i>	<i>Casi Nunca</i>	<i>A Veces</i>	<i>Con Frecuencia</i>	<i>Casi Siempre</i>	<i>Siempre</i>	$\chi^2$	<i>p</i>
3	Observadas	2	29	167	264	368	553	0.04	.9999
	Esperadas	2	29	169	265	367	551		
4	Observadas	60	102	413	398	255	155	0.03	.9999
	Esperadas	60	103	414	396	254	155		
5	Observadas	86	191	580	236	172	118	0.05	.9999
	Esperadas	87	193	579	235	172	119		
12	Observadas	40	137	378	323	246	259	0.05	.9999
	Esperadas	40	139	379	321	245	259		
14	Observadas	31	68	312	321	276	375	0.05	.9999
	Esperadas	31	69	314	320	274	374		
15	Observadas	3	8	125	234	310	703	0.06	.9999
	Esperadas	3	8	127	236	309	701		
16	Observadas	12	38	190	245	329	569	0.15	.9995
	Esperadas	11	38	193	246	328	567		

*Estimación de Parámetros del Modelo de Crédito Parcial para los ítems  
de la Dimensión Confortar*

La estimación de los 36 parámetros para el MCP en la dimensión Confortar convergió en el ciclo 22, con un criterio de convergencia de .001. La variación máxima de los parámetros estimados entre los ciclos fue de .00079. Este modelo se aplicó imponiendo restricciones al Modelo de Respuesta Nominal tal como lo impone el programa MULTILOG.

En la Tabla 28 se muestran los valores de los parámetros  $\delta_m$  estimados para cada uno de los ítems de la dimensión Confortar, así como sus estadísticos descriptivos. El parámetro de inclinación  $a$  constante estimado fue de 0.73 (DE = 0.02), lo cual indica una capacidad discriminativa aceptable para un test de comportamiento típico (Reise & Waller, 1990). MULTILOG estimó los errores estándar de los  $\delta_m$  a través de una transformación de los  $\delta_m$  denominada  $C_{\text{coeff}}$ . Dichos errores no alcanzaron una magnitud considerable.

Tabla 28

*Parámetros estimados mediante MCP para los ítems de la dimensión Confortar*

Ítem	$\delta_1$	$\delta_2$	$\delta_3$	$\delta_4$	$\delta_5$	$\delta_{prom}$
3	-5.42	-3.58	-1.30	-0.65	-0.26	-2.24
4	<b>-1.85</b>	<b>-2.53</b>	-0.10	0.92	1.47	-0.42
5	-2.01	-1.93	<b>1.29</b>	<b>0.95</b>	1.48	-0.04
12	-2.85	-2.05	0.00	0.60	0.60	-0.74
14	<b>-2.40</b>	<b>-2.92</b>	-0.41	<b>0.27</b>	<b>0.10</b>	-1.07
15	<b>-3.29</b>	<b>-5.10</b>	-1.67	<b>-0.71</b>	<b>-0.95</b>	-2.34
16	<b>-3.18</b>	<b>-3.29</b>	-0.97	-0.57	-0.44	-1.69
Media	-3.00	-3.06	-0.45	0.12	0.29	
DE	1.20	1.09	0.99	0.75	0.94	
Mínimo	-5.42	-5.10	-1.67	-0.71	-0.95	
Máximo	-1.85	-1.93	1.29	0.95	1.48	

*Nota:* Los parámetros  $\delta_m$  reversados se muestran en negrita.

Se puede observar que cinco de los siete ítems presentan el fenómeno *reversal* (Dodd & Koch, 1987; Embretson & Reise, 2000), es decir inversiones en el orden de sus parámetros  $\delta_m$ : en el ítem 4  $\delta_2 < \delta_1$ , en el ítem 5  $\delta_4 < \delta_3$ , en el ítem 14  $\delta_2 < \delta_1$  y  $\delta_5 < \delta_4$ , en el ítem 15  $\delta_2 < \delta_1$  y  $\delta_5 < \delta_4$  y finalmente en el ítem 16  $\delta_2 < \delta_1$ . Como ya se mencionó, la existencia del fenómeno *reversal* no implica un mal funcionamiento de los ítems o del modelo. Sin embargo, dado que la escala Likert implica una monotonía creciente podría ser deseable un agrupamiento de las categorías de respuesta intercambiadas (Sam, et al., 2016; Stump et al, 2012) en los ítems con *reversal*.

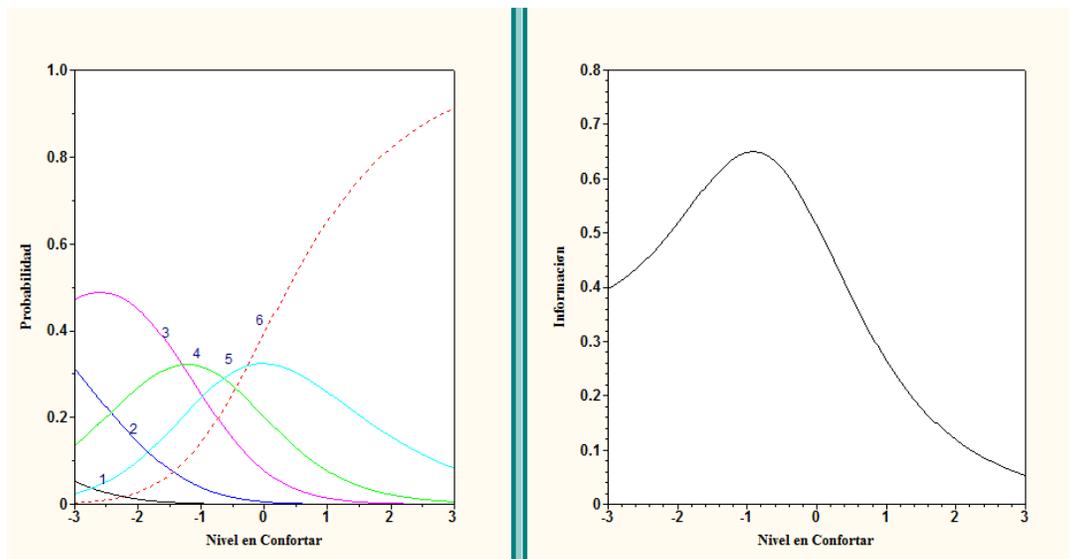
Figura 26

CCCRIs y FIs de los ítems de la dimensión Confortar para el MCP

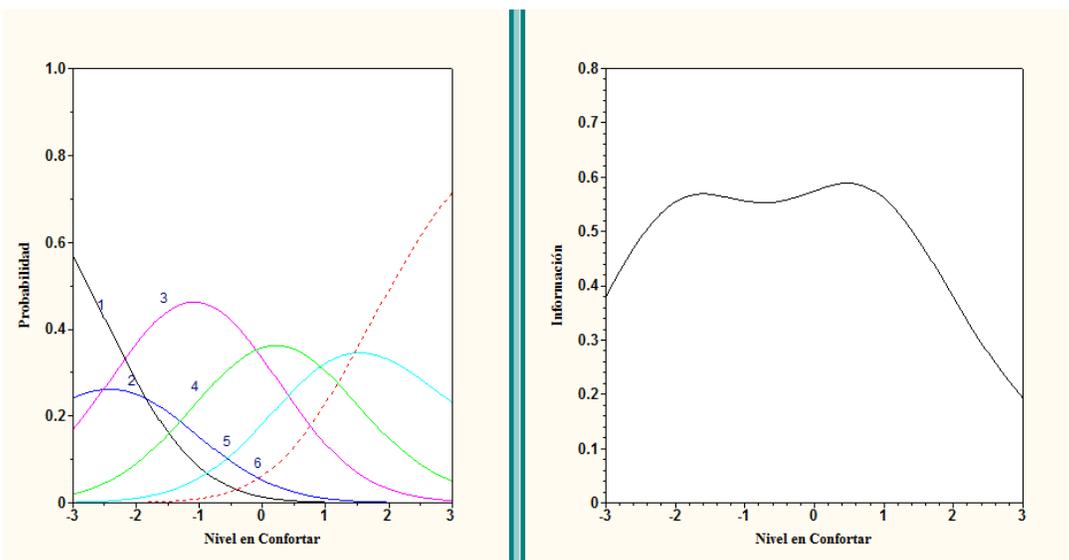
CCCRI

FI

Ítem 3



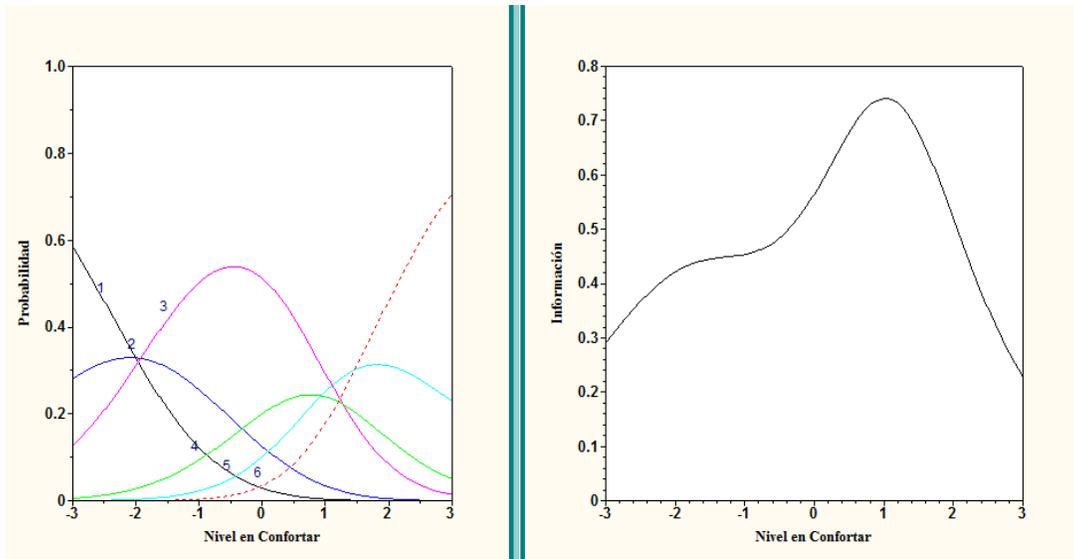
Ítem 4



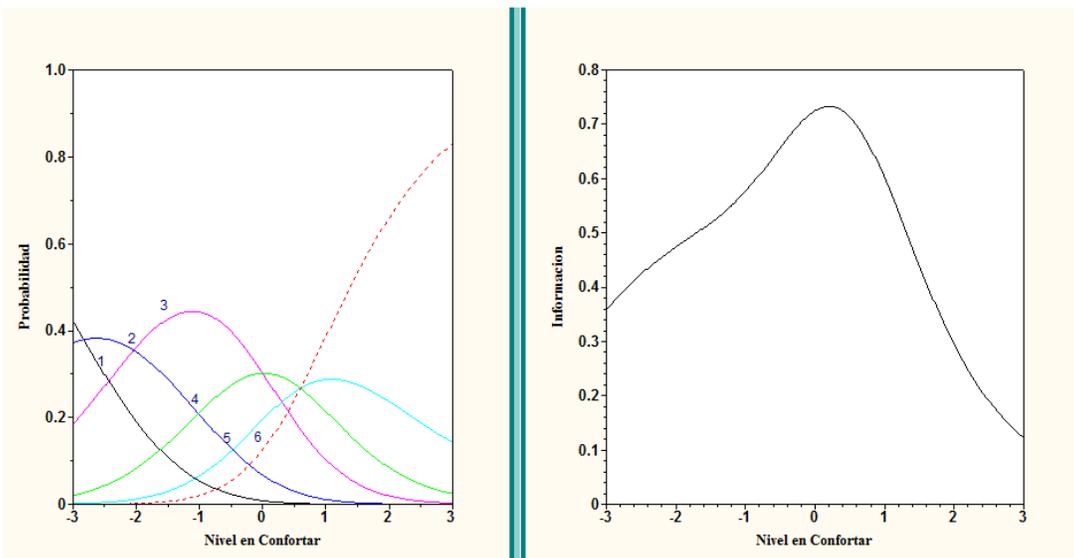
CCCRI

FI

Ítem 5



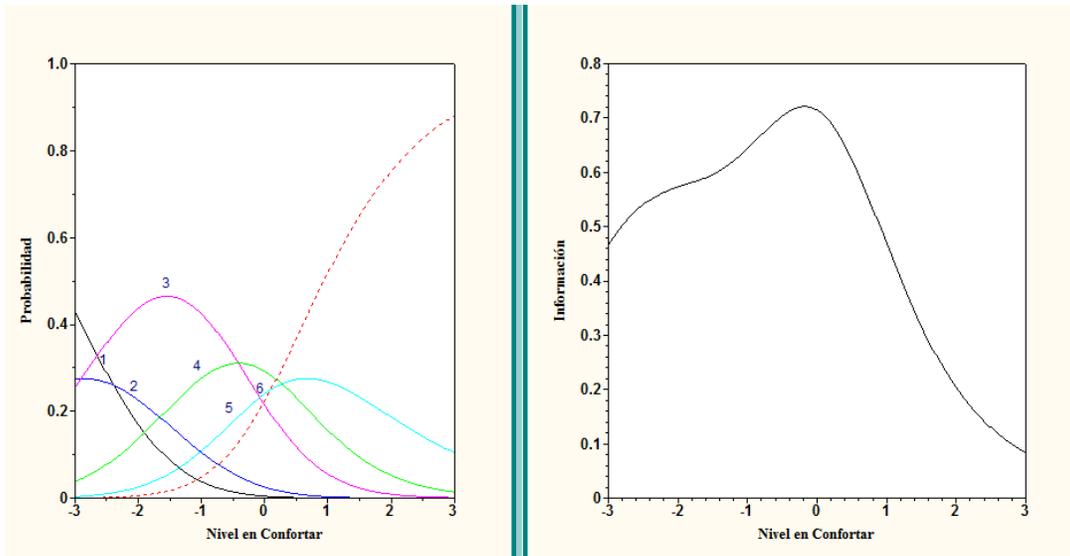
Ítem 12



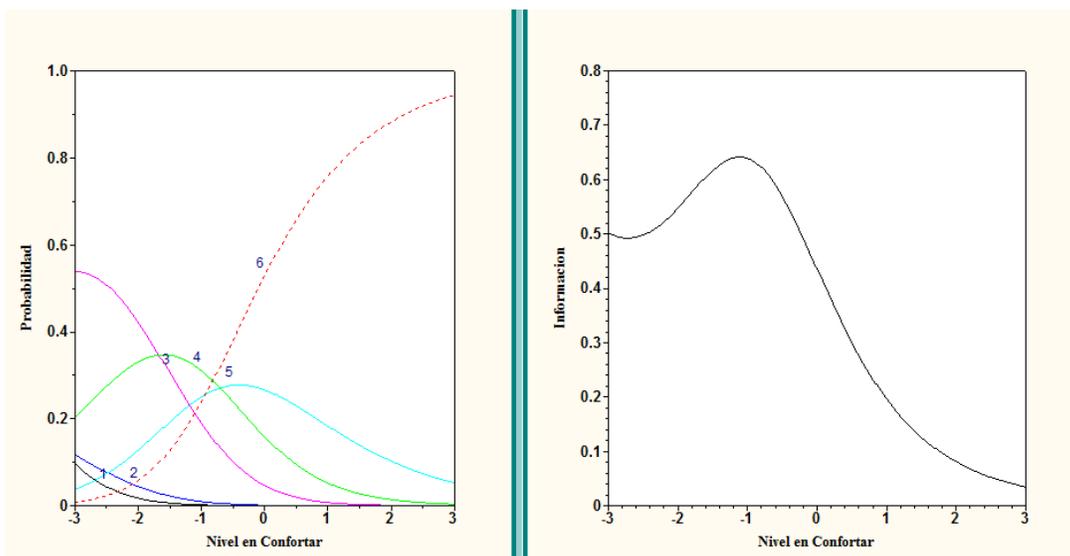
CCCRI

FI

Ítem 14



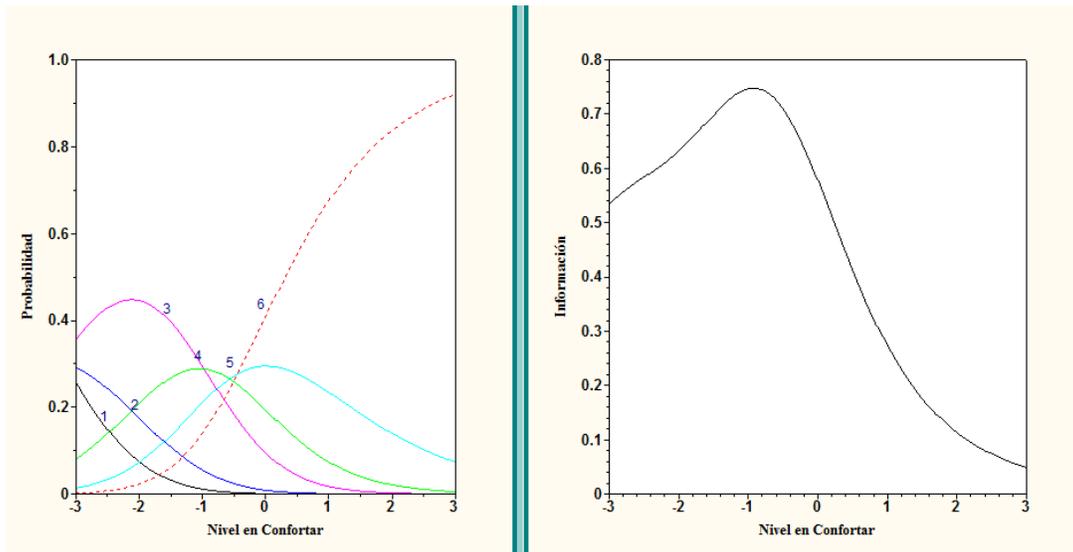
Ítem 15



CCCRI

FI

Ítem 16



Notación

- CCCRI Categoría 1 = *Nunca* —
- CCCRI Categoría 2 = *Casi Nunca* —
- CCCRI Categoría 3 = *A veces* —
- CCCRI Categoría 4 = *Con Frecuencia* —
- CCCRI Categoría 5 = *Casi Siempre* —
- CCCRI Categoría 6 = *Siempre* - - -

La Tabla 29 exhibe los valores que alcanzan las FIs de los ítems y del test para algunos niveles del rasgo Confortar.

Tabla 29

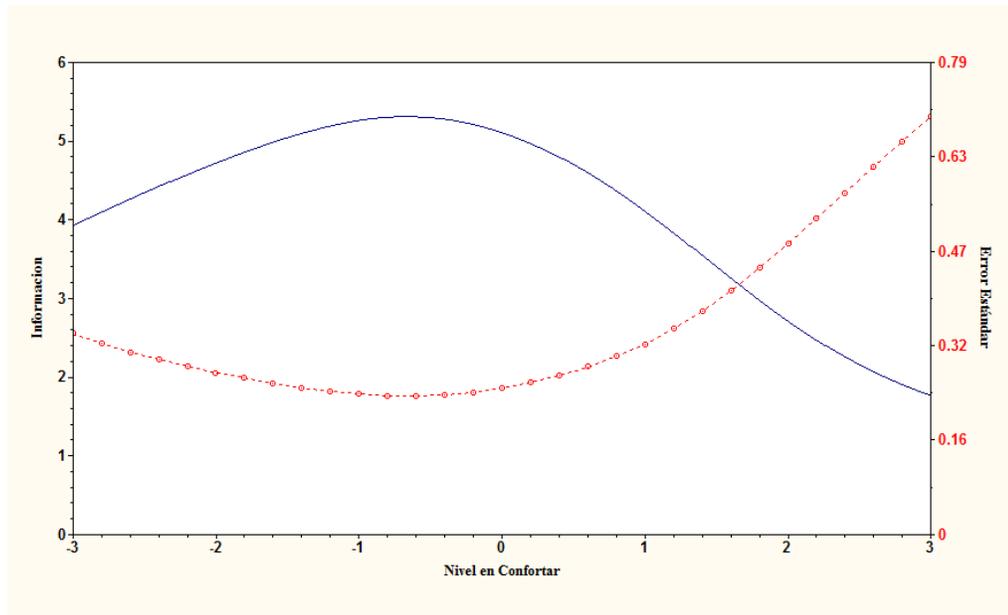
*Funciones de Información de los Ítems y del Test obtenidas mediante MCP  
en la Dimensión Confortar*

Ítem \ $\theta$	$\theta$						
	$\theta = -3$	$\theta = -2$	$\theta = -1$	$\theta = 0$	$\theta = 1$	$\theta = 2$	$\theta = 3$
1	0.40	0.52	0.65	0.51	0.27	0.12	0.05
2	0.38	0.55	0.56	0.57	0.56	0.38	0.19
6	0.29	0.42	0.45	0.56	0.74	0.52	0.23
7	0.39	0.47	0.58	0.72	0.60	0.30	0.12
8	0.47	0.57	0.64	0.71	0.47	0.20	0.08
9	0.50	0.55	0.64	0.44	0.20	0.08	0.03
10	0.54	0.63	0.75	0.58	0.27	0.11	0.05
Test	3.93	4.72	5.26	5.11	4.11	2.71	1.76
DE	0.51	0.46	0.44	0.44	0.49	0.61	0.75

La Figura 27 muestra que el gráfico de la FI del test alcanzó su valor máximo (5.307) en  $\theta = -0.6$ . Como es de esperar dada su relación inversa, el error de medida alcanzó en este punto su valor mínimo de .43. Esta función muestra que el test proporciona estimaciones de Ayuda más precisas en un rango de rasgo theta de -3 a 1, mientras que el error en la medición crece hacia los niveles de rasgo Confortar mayores a 1.

Figura 27

*Función de Información del Test y Error Estándar  
en la Dimensión Confortar para el MCP*



*Obtención de Evidencias de Confiabilidad Global*

Mediante cada uno de los modelos de la TRI aplicados se obtuvo un índice de confiabilidad marginal. La confiabilidad marginal es una medida de precisión global del test propia de la TRI (Thissen, 1991). Su cálculo se basa en un promedio de la confiabilidad a lo largo de todos los niveles del rasgo ponderado por la proporción de individuos que corresponden a cada nivel (Attorresi et al., 2011). Es un estimador cercano al Alfa de Cronbach, por lo que su representatividad aumenta con la uniformidad de la FI del test.

En la Tabla 30 se expone el índice Alfa de Cronbach de la TCT, así como la confiabilidad marginal para cada una de las dimensiones según los modelos de la TRI aplicados.

Tabla 30

*Medidas de Confiabilidad de las dimensiones Ayuda y Confortar*

	Alfa de Cronbach	Conf. Marg. ML1P	Conf. Marg. ML2P	Conf. Marg. ML3P	Conf. Marg. MRG	Conf. Marg. MCP
Ayuda	.77	.63	.65	.61	.82	.78
Confortar	.79	.62	.62	.63	.81	.78

Siguiendo a DeVellis (2012) y a Gempp, Denegri, Caripán, Catalán, Hermosilla y Caprile (2007) las dos dimensiones presentaron un nivel relativamente adecuado de confiabilidad. Se debe tener en cuenta que cuando el número de ítems es bajo, los índices de confiabilidad tienden a disminuir.

Dado que existen diferencias significativas en las puntuaciones de mujeres y varones, se calculó además el Alfa de Cronbach por separado para cada género, el cual resultó similar.

Tabla 31

*Alfa de Cronbach para cada Dimensión en función del Género*

	Mujeres n = 1114	Varones n = 267
Ayuda	.76	.77
Confortar	.76	.82

*6.5. Fase V: Fuentes Externas de Validez*

En esta fase, se obtuvieron evidencias de validez de las puntuaciones de la Escala de Conducta Prosocial mediante fuentes externas. Las fuentes externas utilizadas fueron: los puntajes

en las dimensiones de la versión argentina de la EHP, el nivel de Autopercepción en Empatía, Conducta Prosocial, Dadivosidad y Altruismo, el género, y el tipo de carrera que estudia el participante.

### *Procedimiento General*

Se aportaron evidencias de validez convergente con otros instrumentos mediante la correlación entre las puntuaciones en las dimensiones Ayuda y Confortar y las puntuaciones en las dimensiones de la versión argentina de la EHP -la versión argentina fue obtenida mediante una adaptación preliminar de dicho instrumento-. Asimismo, se correlacionaron las puntuaciones en las dimensiones Ayuda y Confortar con los niveles de Autopercepción de los participantes en Empatía, Conducta Prosocial, Dadivosidad y Altruismo.

Se aportaron evidencias de validez convergente de las puntuaciones en las dimensiones Ayuda y Confortar con los criterios externos no instrumentales género y tipo de carrera - relacionada directamente con la conducta prosocial o no- que el participante está cursando. Para ello se emplearon las pruebas no paramétricas *U* de Mann–Whitney y *H* de Kruskal-Wallis.

Relación de las Dimensiones Ayuda y Confortar con la Escala de Habilidades Prosociales

### *Caracterización de la Escala de Habilidades Prosociales (EHP)*

*de Morales Rodríguez y Suárez Pérez (2011)*

Existe una diversidad de instrumentos psicométricos destinados a la medición de la conducta prosocial. Entre ellos se destaca la Escala de Habilidades Prosociales (EHP) construida por Morales Rodríguez y Suárez Pérez (2011) de la Universidad Michoacana de San Nicolás de Hidalgo, México. Este inventario fue desarrollado para evaluar las habilidades sociales positivas avanzadas, entendidas como conducta prosocial. Se encuentra conformada por 20 ítems, distribuidos en cuatro subescalas. Morales Rodríguez y Suárez Pérez (2011) caracterizan a las subescalas de la siguiente forma:

1. Toma de Perspectiva (5 ítems, Alfa de Cronbach = .70). Se refiere a la capacidad que tiene un sujeto para ponerse en el lugar del otro, comprender sus pensamientos, sentimientos, motivos y conductas.
2. Solidaridad y Respuesta de Ayuda (6 ítems, Alfa de Cronbach = .71). La *solidaridad* se refiere a la virtud y deber social con las personas más necesitadas, beneficiando a un máximo de personas. La *respuesta de ayuda* es la acción del sujeto ante una carencia de otro.
3. Altruismo (5 ítems, Alfa de Cronbach = .67). Es toda conducta prosocial que se realiza por motivos o valores internos, sin buscar recompensas externas.
4. Asistencia (4 ítems, Alfa de Cronbach = .71). Es la ayuda física que se le proporciona a alguien ante una situación de emergencia.

Morales Rodríguez y Suárez Pérez (2011) elaboraron el inventario en base a la discusión teórica y aspectos empíricos relacionados con propiedades de la conducta prosocial. En su versión original la escala fue desarrollada y validada con una muestra de 880 participantes mexicanos con una edad promedio de 18 años y donde 49% eran mujeres y el 51% hombres.

Los ítems fueron planteados en términos de la evaluación de una situación particular, valorados mediante una escala Likert de cuatro opciones (*No me describe, Me describe poco, Me describe regular y Me describe bien*). Sus creadoras concordaron en que la EHP es un inventario que puede ser útil para identificar actos positivos realizados con o sin motivación altruista, encaminados al bienestar del otro (Morales Rodríguez & Suárez Pérez, 2011).

Las autoras reportaron evidencias de validez basadas en la estructura interna de la prueba a partir de la aplicación de un Análisis Factorial Exploratorio basado en la matriz de correlaciones de Pearson. En consecuencia, para la aplicación de este estudio se consideraron como intervalares los códigos numéricos impuestos a las categorías de respuesta de la escala Likert (Gil, Rodríguez & García, 2000). Se empleó el método de Componentes Principales para la extracción de factores y el método de rotación ortogonal (Varimax). Este último procedimiento implica suponer que los factores identificados son incorrelados entre sí. Según sus resultados, los cuatro factores aislados describieron el 47% de la varianza y los ítems presentaron pesajes adecuados (Morales Rodríguez & Suárez Pérez, 2011).

El estudio de la dimensionalidad del instrumento se complementó con el aporte de evidencias de validez basadas en la relación de la EPH con otras variables, entre ellas las

subdimensiones de la Escala de Autoeficacia General (EAG) adaptación española de Baessler y Schwarzer (1996) y la diferencia de medias según el género.

Adaptación de la Escala de Habilidades Prosociales de Morales Rodríguez y Suárez Pérez (2011)

### *Procedimiento*

Con el propósito de realizar una adaptación de la Escala de Habilidades Prosociales (EHP) de Morales Rodríguez y Suárez Pérez (2011) en población argentina adulta se solicitó y obtuvo la autorización correspondiente a las autoras de la prueba. Luego, se efectuó una aplicación piloto de la prueba en 30 estudiantes, a fin de evaluar las expresiones lingüísticas utilizadas en los ítems y la comprensión de la consigna. Fueron recurrentes las dificultades para la comprensión de algunos términos comunes en la población mexicana pero poco frecuentes en el contexto argentino. A partir de estos resultados se realizaron adecuaciones lingüísticas a los ítems 6, 7, 11, 13, 15 y 17 de la versión original, tratando de mantener la equivalencia de los términos. Las adecuaciones realizadas se muestran en la Tabla 32.

Tabla 32  
*Adaptación lingüística de la EHP*

Ítem	Ítem Original	Versión adaptada
6	Trabajaría como voluntario(a) en grupos como cruz roja o bomberos.	Trabajaría como voluntario/a en instituciones como cruz roja o bomberos.
7	Acostumbro a donar dinero o en especie, para apoyar a las personas necesitadas.	Acostumbro realizar donaciones para ayudar a las personas necesitadas.
11	Brindaría alimento a personas que vienen de fuera.	Brindaría alimento a personas necesitadas.
13	Cuando veo un accidente en carretera inmediatamente actúo, llamo a emergencias.	Cuando veo un accidente en la ruta actúo inmediatamente llamando a emergencias.

15	Me gustaría ser paramédico aunque no recibiera un sueldo.	Me gustaría ayudar en situaciones de emergencia aunque no recibiera un sueldo.
17	Ofrezco de mis alimentos cuando hay otras personas.	Convido de mi comida cuando hay otras personas.

Posteriormente, el protocolo fue validado por el sistema de jueces, a fin de identificar la concordancia entre el ítem original y la modificación lingüística, así como la adecuación de los ítems a la población adulta. El análisis de los conceptos emitidos por los jueces se realizó siguiendo los criterios de Compas et al. (1987), es decir el porcentaje de acuerdo entre jueces. A cinco profesionales con experiencia en Psicología Social y Psicometría se les solicitó una valoración a ciegas de los ítems. Los jueces evidenciaron preferencia por las versiones locales de los reactivos, considerándolas más adecuadas que las originales. Este nuevo protocolo fue ensayado en una muestra piloto de 35 participantes, los que no mostraron dificultades de comprensión de la prueba.

Luego el instrumento fue administrado a la muestra final de 404 estudiantes de la Universidad de Buenos Aires en forma anónima de los cuales 322 (80%) fueron mujeres y 82 (20%) varones. La media de edad fue de 23 años con un desvío estándar de 6.3. Para la selección de los individuos se aplicó un diseño muestral no probabilístico por accesibilidad (Gil-Escudero & Martínez-Arias, 2001). Se enfatizó el carácter voluntario de la participación y se destacó que la información recabada solo sería utilizada con fines de investigación. La ubicación de los reactivos conservó la distribución propuesta en la versión original de la EHP.

Se llevó adelante un AFE con el fin de estudiar la estructura interna de la prueba. Previamente se verificó la aplicabilidad de este análisis multivariante mediante el cálculo del coeficiente KMO y la prueba de esfericidad de Bartlett. Dada la naturaleza ordinal de las respuestas a los ítems (García-Cueto y Fidalgo, 2005), se efectuó un AFE a partir de la matriz de correlaciones policóricas (Brown, 2006) mediante el programa Factor 9.2 (Lorenzo-Seva & Ferrando, 2006). Se utilizó el método de mínimos cuadrados no ponderados (ULS) para la extracción de factores y el método Promax para la Rotación. La cantidad de factores retenidos se determinó a partir de la implementación óptima del análisis paralelo (Timmerman & Lorenzo Seva, 2011). También se calculó el RMSR, relacionándolo con el criterio de Kelley (1935).

La decisión de adoptar un método de rotación oblicua obedeció a la aparición de factores con correlaciones superiores a .32 (Tabachnick & Fidell, 2001). Conforme lo recomendaron Martínez Arias et al. (2006), el método Promax se implementó usando un parámetro kappa de 4

para limitar la rotación. Se llevó a cabo una depuración de los ítems que no respondían a lo esperado a nivel teórico considerando la información suministrada por la matriz de configuración. El uso de esta matriz en lugar de la matriz de estructura responde a que los elementos de esta matriz de configuración son los que más se aproximan a la estructura simple y, en consecuencia, se facilita la interpretación de los factores identificados (Martínez Arias et al., 2006).

A fin de evaluar la confiabilidad se calculó el coeficiente Alfa de Cronbach para cada factor.

### *Resultados*

El porcentaje de acuerdo interjueces permitió corroborar la pertinencia de la adaptación lingüística pues alcanzó un 75% de concordancia entre los evaluadores en su primera revisión, finalizando con un 95% posterior una vez realizadas las sugerencias indicadas.

Posteriormente se analizó la dimensionalidad de instrumento. El valor correspondiente a la prueba de KMO de la escala fue de .82 y la prueba de esfericidad de Bartlett fue significativa ( $\chi^2=1715$ ;  $gl=190$ ;  $p<.0001$ ) mostrando condiciones adecuadas para la realización del AFE. Basado en un criterio a priori que establece el modelo teórico de Morales Rodríguez y Suárez Pérez (2011) se forzó la extracción de cuatro factores. Sin embargo, no se logró sostener esta estructura ya que en uno de los factores los contenidos de los ítems no resultaron coherentes entre sí como para aceptar la identificación de una dimensión teórica.

Dado que no se cumplió el modelo propuesto para la versión original de la EHP, se realizó una depuración del instrumento a fin de alcanzar una estructura más parsimoniosa para analizar las conductas prosociales en población argentina. Fueron descartados progresivamente los ítems 1, (*Si alguien perdió algún objeto personal, le ayudo a buscarlo*), 3 (*Doy mi asiento a personas que lo requieren*) y 5 (*Ante una catástrofe, pienso que unidos podemos salir adelante*) de la EHP original ya que, según la información recogida de la matriz de configuración, estos ítems presentaron coeficientes de regresión estandarizados insuficientes en todos los factores (menores a .30). Los tres reactivos correspondían a la dimensión *Solidaridad y Respuesta de Ayuda*.

Asimismo, fueron eliminados los ítems 2 (*Cuando ofendo a una persona, procuro disculparme*), 4 (*Colaboraría con instituciones de beneficencia*), 14 (*Si una persona olvida algún objeto, lo recojo y se lo entrego*) y 17 (*Convido de mi comida cuando hay otras personas*) por cargar en un factor que no les correspondía según los hallazgos de Morales Rodríguez y Suárez

Pérez (2011) y el ítem 7 (*Acostumbro realizar donaciones para ayudar a las personas necesitadas*) por presentar doble pesaje.

Se procedió entonces a realizar un nuevo AFE donde se forzó la extracción a tres factores. La solución final ( $KMO = .79$ ; prueba de esfericidad de Bartlett,  $\chi^2 = 1009.1$ ,  $gl = 66$ ,  $p < .00001$ ) permitió identificar un conjunto de tres factores intercorrelacionados que explicaron el 60% de la varianza. La implementación óptima del análisis paralelo indicó que la estructura trifactorial era adecuada. El valor del RMSR fue de .038, por debajo de .05 lo que indica que la discrepancia entre la matriz de varianzas y covarianzas observada y la reproducida por el modelo es baja. Se cumplió el criterio de Kelley de que el RMSR sea menor que  $1/\sqrt{n}$  con  $.038 < .049$ . El RMSEA fue de .059, indicando un ajuste óptimo del modelo (Hu & Bentler, 1995).

En la Tabla 33 se encuentran las matrices de configuración y de estructura definitivas posteriores a la depuración. En la primera se presentan los coeficientes de regresión estandarizados que permiten apreciar el aporte relativo de cada factor para explicar el puntaje individual del ítem. En la segunda se exhiben las correlaciones obtenidas por cada uno de los ítems con el correspondiente factor. Los ítems que saturaron en el primer factor (ítems 8, 10, 11, 12 y 13) son indicativos de la dimensión Altruismo en la versión mexicana de la prueba. El segundo factor estuvo compuesto por los ítems 9, 16, 18 y 19, que fueron planteados originalmente para la dimensión Toma de Perspectiva. Por último, los elementos que componen el tercer factor (ítems 6, 15 y 20) pertenecían a la dimensión Asistencia.

Al comparar ambas matrices rotadas se puede observar que la de configuración resultó más próxima a una estructura simple. En cambio, la matriz de estructura permite evidenciar que los ítems de un factor también alcanzaron correlaciones significativas con los otros factores. Esto concuerda ampliamente con las expectativas basadas en la naturaleza del constructo estudiado. En efecto, al estudiar la correlación entre los factores se encontró que una asociación moderada de Altruismo tanto con Toma de Perspectiva ( $r = .45$ ;  $p < .01$ ) como con Asistencia ( $r = .45$ ;  $p < .01$ ). No obstante, la correlación entre estas dos últimas resultó positiva pero más débil ( $r = .27$ ;  $p < .01$ ).

Tabla 33  
*Matrices de Configuración y de Estructura*  
*Alfa de Cronbach por Dimensión*

Ítems	M. de configuración			M. de estructura		
	1*	2*	3*	1*	2*	3*
6.- Trabajaría como voluntario/a en instituciones como cruz roja o bomberos	.05	-.06	<b>.73</b>	.35	.15	<b>.74</b>
8.- Ayudo a personas de la tercera edad y con capacidades diferentes a cruzar la calle	<b>.52</b>	-.00	.22	<b>.62</b>	.29	.45
9.- Si una persona es ofendida, imagino lo mal que se siente	.01	<b>.76</b>	-.05	.33	<b>.76</b>	.15
10.- Cuando me percató de un robo doy aviso a las autoridades	<b>.76</b>	.03	-.16	<b>.70</b>	.33	.19
11.- Brindaría alimento a personas necesitadas	<b>.37</b>	-.02	.28	<b>.56</b>	.32	.40
12.- Acostumbro reunir ropa en buen estado, que no uso, para regalarla	<b>.35</b>	-.04	.14	<b>.40</b>	.15	.29
13.- Cuando veo un accidente en la ruta actúo inmediatamente llamando a emergencias	<b>.71</b>	-.02	.02	<b>.70</b>	.30	.33
15.- Me gustaría ayudar en situaciones de emergencia aunque no recibiera un sueldo	-.02	-.01	<b>.80</b>	.32	.19	<b>.79</b>
16.- Comprendo cómo se siente una persona cuando es rechazada	.11	<b>.75</b>	-.08	.42	<b>.78</b>	.17
18.- Comprendo cuando alguien está pasando por un mal momento	-.01	<b>.70</b>	.03	.32	<b>.71</b>	.22
19.- Antes de criticar a alguien me imagino cómo se sentiría si estuviera en su lugar	-.16	<b>.53</b>	.19	.16	<b>.50</b>	.26
20.- Cuidaría enfermos sin recibir un pago	-.07	.10	<b>.71</b>	.30	.27	<b>.71</b>
<b>Alfa de Cronbach por Dimensión</b>	<b>.68</b>	<b>.66</b>	<b>.73</b>			

\* 1 = Altruismo, 2 = Toma de Perspectiva, 3 = Asistencia

Como puede observarse las dimensiones Altruismo y Toma de Perspectiva presentaron un nivel aceptable de confiabilidad (.68 y .66 respectivamente), mientras que la dimensión Asistencia, con .73, mostró un nivel adecuado (DeVellis, 2012).

*Correlación de las Dimensiones de la Escala de Habilidades Prosociales (EHP) con las Dimensiones Ayuda y Confortar de la Escala de Conducta Prosocial*

La versión argentina de la EHP (ver Apéndice E) se aplicó a la muestra con el total de participantes (n = 1383). Así, se pudieron obtener las correlaciones de las dimensiones Ayuda y Confortar de la ECP con las dimensiones Altruismo, Toma de Perspectiva y Asistencia de la EHP. Todas las correlaciones fueron de, como mínimo .32 y significativas con  $p < .01$  tal como se muestra en la Tabla 34. Además, la correlación más alta de Ayuda fue con Asistencia y la correlación más alta de Confortar fue con Toma de Perspectiva, de forma coherente con la coincidencia a nivel teórico de los contenidos de estas dimensiones.

Tabla 34  
*Análisis de validez convergente*

Dimensiones de la ECP	Dimensiones de la Versión Argentina de la EHP		
	Altruismo	Toma de Perspectiva	Asistencia
Ayuda	.46**	.32**	.49**
Confortar	.33**	.42**	.32**

\*\*  $p < .01$

*Relación de las Dimensiones Ayuda y Confortar de la ECP con la Autopercepción en Empatía,  
Conducta Prosocial, Dadivosidad y Altruismo*

Se estudió la validez convergente mediante análisis correlacional bivariado ( $r$  de Pearson) encontrándose correlaciones significativas con  $p < .01$  entre las dimensiones Ayuda y Confortar y las Autopercepciones de los participantes con respecto a su Empatía (AE), Conducta Prosocial (ACP), Dadivosidad (AD) y Altruismo (AA) en la muestra (Tabla 35).

La dimensión Ayuda tuvo la correlación más fuerte con la Autopercepción en Conducta Prosocial ( $r = .53, p < .01$ ) y la más débil con la Autopercepción en Empatía ( $r = .36, p < .01$ ). La correlación de la dimensión Confortar con la Autopercepción en Empatía ( $r = .49, p < .01$ ), por el contrario, fue la más alta para ese factor. Mientras que la correlación más baja de la dimensión Confortar fue con la Autopercepción en Altruismo ( $r = .28, p < .01$ ). Las correlaciones obtenidas son coherentes con lo esperable de forma teórica y aportan evidencias de validez convergente.

Tabla 35

*Correlación entre las Dimensiones Ayuda y Confortar y la Autopercepción en Empatía,  
Conducta Prosocial, Dadivosidad y Altruismo*

	Autopercepción			
	AE	ACP	AD	AA
Ayuda	.36**	.53**	.43**	.42**
Confortar	.49**	.31**	.32**	.28**

*Nota.* AE = Autopercepción de Empatía; ACP = Autopercepción en Conducta Prosocial; AD = Autopercepción en Dadivosidad; AA = Autopercepción en Altruismo

\*\*  $p < .01$

## Relación de las Dimensiones Ayuda y Confortar con el Género

Se exploró la existencia de un efecto del género de los participantes en los puntajes de las dimensiones Ayuda y Confortar. Para este análisis se utilizó el total depurado de participantes de la muestra ( $n = 1383$ ) menos dos individuos que reportaron un género diferente al femenino o masculino. Es decir, se utilizó  $n = 1381$ . El análisis de un efecto de género fue posterior a la depuración de ítems que presentasen DIF por género. En este caso como se mencionó previamente se descartó el ítem 13 (*Hago donaciones a organizaciones benéficas*) por presentar DIF a favor de las mujeres.

Dado que los puntajes en Ayuda y Confortar no se encuentran normalmente distribuidos (prueba de Kolmogorov-Smirnov,  $p < .0001$  para ambas dimensiones) se aplicó la prueba no paramétrica  $U$  de Mann–Whitney por separado para cada una de las dos dimensiones de la escala de conducta prosocial en la muestra ( $n = 1381$ ) empleando el programa SPSS versión 21. El puntaje total de cada dimensión se adoptó como variable dependiente, mientras que el género (Femenino,  $n = 1114$  o Masculino,  $n = 267$ ) se configuró como factor fijo.

En cuanto al análisis descriptivo, en la Tabla 36 pueden apreciarse las medias y desviaciones estándar para cada dimensión según el género.

Tabla 36

*Media y Desviación Estándar de Puntaje en las Dimensiones Ayuda y Confortar según Género de los Participantes*

Dimensión	Género	
	Femenino	Masculino
Ayuda	25.18 (6.37)	23.30 (6.52)
Confortar	31.06 (5.50)	27.97 (6.07)

Los resultados de la prueba  $U$  de Mann–Whitney para las dimensiones Ayuda y Confortar indican que el género tiene efectos significativos sobre el puntaje en dichos factores con  $p < .0001$

para ambos. En las dos dimensiones las mujeres son quienes tienen significativamente mayor puntaje. Sin embargo, las diferencias halladas podrían ser consecuencias del tamaño muestral.

Si se considera como muy prosociales a aquellos que tienen más de un desvío por sobre la media, las proporciones para cada género son:

Tabla 37

*Proporciones de Personas Muy Prosociales según Género*

Dimensión	Género	
	Femenino	Masculino
Ayuda	.161	.161
Confortar	.171	.146

Las diferencias entre proporciones de personas muy prosociales no son estadísticamente significativas para ninguna de las dos dimensiones (Confortar,  $p = .85$ ).

#### Relación de las Dimensiones Ayuda y Confortar con el Tipo de Carrera

Con el fin de obtener evidencias de validez relacionada con un criterio externo, se verificó la hipótesis indicada oportunamente de la existencia de una relación entre el tipo de carrera que cursan los participantes y los puntajes de las dimensiones Ayuda y Confortar. Para ello se dividieron las carreras en tres grupos:

Grupo 1 (G1): Lic. En Psicología, Psicopedagogía, Trabajo Social, Terapia Ocupacional, Musicoterapia y Minoridad. El tamaño de este grupo ( $n_1$ ) fue de 469 participantes.

Grupo 2 (G2): Medicina, Enfermería, Kinesiología, Fonoaudiología, Nutrición, Obstetricia y Odontología. El tamaño de este grupo ( $n_2$ ) fue de 192 participantes.

Grupo 3 (G3): Carreras no incluidas en los Grupos 1 y 2, las cuales no involucran en forma directa a las conductas prosociales. El tamaño de este grupo ( $n_3$ ) fue de 722 participantes. De estos, las carreras con más proporción de estudiantes fueron Abogacía (13.4 %), Contador Público (12.7

%), Licenciatura en Administración (7.9 %), Licenciatura en Artes Audiovisuales (7.1%), Recursos Humanos (5.0%) e Ingeniería (4.7 %).

Lo esperable sería que los estudiantes de carreras que involucran directamente a las conductas prosociales (Grupo 1 y Grupo 2) puntúen más alto en alguna o ambas dimensiones Ayuda y Confortar. De todas maneras, en las carreras incluidas en el Grupo 3 es necesaria una conducta ética.

Dado que los puntajes en Ayuda y Confortar no se encuentran normalmente distribuidos (prueba de Kolmogorov-Smirnov,  $p < .0001$  para ambas dimensiones) se aplicó la prueba no paramétrica  $H$  de Kruskal-Wallis por separado para cada una de las dos dimensiones de la escala de conducta prosocial en la muestra ( $n = 1383$ ) empleando el programa SPSS versión 21. Para el contraste post-hoc se utilizó la prueba no paramétrica  $U$  de Mann-Whitney. El puntaje total de cada dimensión se adoptó como variable dependiente, mientras que el tipo de carrera se configuró como factor fijo.

En cuanto al análisis descriptivo, en la Tabla 38 pueden apreciarse las medias y desviaciones estándar para cada dimensión según el tipo de carrera (G1, G2 o G3).

Tabla 38

*Media y Desviación Estándar de Puntaje en las Dimensiones Ayuda y Confortar según Tipo de Carrera*

	Tipo de Carrera		
	G1	G2	G3
	$n = 1383$		
Dimensión			
Ayuda	24.90 (0.30)	26.29 (0.46)	24.39 (0.24)
Confortar	30.76 (5.15)	31.57 (6.06)	29.97 (5.96)

*Nota.* G1: Estudiantes de Lic. En Psicología, Psicopedagogía, Trabajo Social, Terapia Ocupacional, Musicoterapia y Minoridad, G2: Medicina, Enfermería, Kinesiología, Fonoaudiología, Nutrición, Obstetricia y Odontología, G3: Estudiantes de carreras que no involucran en forma directa a las conductas prosociales.

Los resultados de la prueba  $H$  de Kruskal-Wallis para la dimensión Ayuda indican que el tipo de carrera tiene efectos significativos sobre el puntaje en dicho factor con  $p = .001$ .

Las comparaciones post-hoc sobre el tipo de carrera mediante la prueba  $U$  de Mann-Whitney señalan que los estudiantes de carreras pertenecientes a G2 obtienen puntuaciones significativamente superiores que los estudiantes de carreras pertenecientes a G1 ( $p = .016$ ) y G3 ( $p < .0001$ ). No existieron diferencias estadísticamente significativas entre los estudiantes de carreras de tipo G1 y G3 ( $p = .133$ ).

Los resultados de la prueba  $H$  de Kruskal-Wallis para la dimensión Confortar indican que el tipo de carrera tiene efectos significativos sobre el puntaje en dicho factor con  $p = .002$ .

Las comparaciones post-hoc sobre el tipo de carrera mediante la prueba  $U$  de Mann-Whitney señalan que los estudiantes de carreras pertenecientes a G1 y G2 no obtienen puntuaciones significativamente diferentes ( $p = .059$ ). Los estudiantes de carreras pertenecientes a G1 y G2 obtienen puntuaciones significativamente superiores que los estudiantes de carreras pertenecientes a G3 con  $p = .030$  y  $p = .001$  respectivamente.

## 6.6. Fase VI: Caracterización de los Participantes en su Conducta Prosocial

### *Procedimiento General*

Para caracterizar a los participantes en su conducta prosocial, se establecieron puntos de corte con base en el criterio de Inglés et al. (2008) para determinar su nivel, obteniendo la proporción de sujetos con muy bajo nivel, bajo nivel, nivel intermedio, alto nivel y muy alto nivel de conducta prosocial según cada uno de los modelos utilizados. Se consideró la puntuación directa con dicotomización ( $TCT_{DICO}$ ) y sin dicomización ( $TCT_{POLI}$ ) de las respuestas según la TCT y los rasgos latentes  $\theta$  estimados por ML1P, ML2P, ML3P, MRG y MCP. Este procedimiento se realizó

para las dimensiones Ayuda y Confortar. Previamente, se calcularon estadísticos descriptivos y la prueba de Kolmogorov-Smirnov de dichas puntuaciones directas y rasgos latentes  $\theta$  estimados.

## *Resultados*

### Caracterización de los Participantes en la Dimensión Ayuda

Los resúmenes descriptivos de los valores –que pueden ser puntajes directos o  $\theta$  - obtenidos por los participantes en la dimensión Ayuda se muestran en la Tabla 39. Como puntos en común entre los diferentes modelos, se observa que tanto la Media como la Mediana ( $Q_2$ ) se ubican alrededor de valores centrales. Además, los  $\theta$  estimados mediante los modelos de la TRI presentan correlaciones muy altas con los puntajes correspondientes de la TCT. Las distribuciones de los valores de los modelos de la TCT y dicotómicos de la TRI tuvieron índices de asimetría positivos y moderados evidenciando que la mayoría de los participantes se ubicó por debajo de la media de los valores. Para las distribuciones de frecuencias de los modelos politómicos de la TRI el índice de asimetría fue negativo pero muy leve, o sea que existió una aproximación a una distribución simétrica. Con respecto a la curtosis, la misma fue negativa para todos los modelos dicotómicos tanto de la TCT como de la TRI, muy próxima a 0 para el puntaje politómico de la TCT y positiva para los modelos politómicos de la TRI. Mientras que la curtosis negativa supone mayor heterogeneidad de los valores y por lo tanto menor representatividad de la Media y la Mediana, la curtosis positiva indica mayor concentración de los valores en torno a las medidas de tendencia central. La distribución de los  $\theta$  obtenidos mediante MRG es la única que ajusta a la distribución normal, desde que en la prueba de Kolmogorov-Smirnov no se rechazó la hipótesis nula que sostiene dicha distribución con  $p = 0.92$ .

Tabla 39

*Distribución de los puntajes y de los rasgos latentes estimados en la dimensión Ayuda y prueba de normalidad*

		TCT <sub>DICO</sub>	TCT <sub>POLI</sub>	ML1P	ML2P	ML3P	MRG	MCP
Estadísticos Descriptivos	Media	2.57	24.83	0.04	0.05	0.12	0.00	0.01
	DE	1.92	6.44	0.76	0.75	0.72	0.90	0.87
	Mín	0	8	-1.08	-1.03	-0.88	-2.82	-2.81
	Máx	8	47	2.10	2.02	2.04	3.11	2.96
	Q <sub>1</sub>	1	21	-0.56	-0.56	-0.44	-0.60	-0.45
	Q <sub>2</sub>	2	25	-0.12	-0.02	0.04	0.01	0.09
	Q <sub>3</sub>	4	29	0.62	0.60	0.61	0.58	0.58
	Asimetría	0.61	0.19	0.37	0.39	0.43	-0.05	-0.17
	Curtosis	-0.25	-0.02	-0.42	-0.53	-0.51	0.20	0.31
<i>p</i> valor Prueba K-S		<i>p</i> < .0001	<i>p</i> = 0.001	<i>p</i> < .0001	<i>p</i> < .0001	<i>p</i> < .0001	<i>p</i> = 0.92	<i>p</i> = 0.001
<i>r</i> con puntuación total TCT		1	1	.997**	.997**	.973**	.969**	.996**

\*\**p* < .0001

Siguiendo el criterio de Inglés et al. (2008) que caracteriza como prosociales a los participantes con más de un desvío y medio por sobre la media, la prevalencia de sujetos prosociales en la dimensión Ayuda para la muestra es de entre un 5.8% y un 8.7%, dependiendo del modelo (ver Tabla 40). Si el criterio se amplía y se considera como prosociales a aquellos participantes con más de un desvío por sobre la media, la prevalencia aumenta a valores de entre 15.0% y 16.4%.

Utilizando el criterio de Inglés et al. (2008) de manera inversa, se puede considerar como sujetos muy poco prosociales a aquellos que tienen valores menores a un desvío y medio por debajo de la media. Mientras que los modelos dicotómicos no detectaron a participantes con estas características, los modelos politómicos coincidieron en señalar muy bajo nivel de conducta prosocial de ayuda para el 6.6% o 6.7% de los sujetos. El pobre nivel de información que brindan los modelos dicotómicos en los niveles bajos del rasgo Ayuda podría explicar esta falta de

detección. Con el criterio negativo ampliado, se halla que entre 14.5% y 17.5% de los participantes reporta un bajo nivel de Ayuda.

La mayoría de los sujetos se ubica en valores intermedios -entre menos de un desvío por debajo y por encima de la Media- en los puntajes y los rasgos latentes estimados de la conducta prosocial de Ayuda. En estos niveles medios se encuentran entre el 66.1% y el 70.2% de los participantes, siempre según el modelo.

Tabla 40

*Prevalencia de las conductas prosociales de Ayuda*

	TCT <sub>DICO</sub>	TCT <sub>POLI</sub>	ML1P	ML2P	ML3P	MRG	MCP
% Media + 1.5 DE	8.4	7.7	8.4	8.7	7.3	7.2	5.8
% Media + 1.0 DE	16.4	15.0	16.4	16.4	15.7	14.8	15.0
% Media - 1.0 DE	14.5	16.1	14.5	17.5	17.5	15.0	16.1
% Media - 1.5 DE	0	6.6	0	0	0	6.7	6.6
% Intermedios	69.1	68.9	69.1	66.1	66.8	70.2	68.9

Los sujetos altamente prosociales en Ayuda (puntaje mayor que Media + 1.5 DE) según el modelo TCT politómico (n =106) tienen una media de edad de 22.15 años (DE = 5.32) y 88.7% son mujeres. Las carreras que estos 106 participantes más eligen son Lic. en Psicología: 23.6%, Medicina: 20.8%, Abogacía: 13.2%) y Trabajo Social: 5.7%. En la muestra las proporciones de participantes que cursan estas carreras son: Lic. en Psicología: 30.4%, Medicina: 10.1%, Abogacía: 7.0% y Trabajo Social: 1.5%.

Los participantes con muy bajo nivel de Ayuda (puntaje menor que Media - 1.5 DE) según el modelo TCT politómico (n = 91) tienen una media de edad de 22.40 años (DE 6.10) y 73.6% son mujeres. Las carreras más elegidas por estos 91 participantes fueron Lic. en Psicología: 30.8%, Contador Público: 8.8% y Lic. en Administración: 8.8%. En la muestra las proporciones de participantes que estudian estas carreras son Lic. en Psicología: 30.4%, Contador Público: 6.7% y Lic. en Administración: 4.1%.

## Caracterización de los Participantes en la Dimensión Confortar

Los resúmenes descriptivos de los valores –que pueden ser puntajes brutos o  $\theta$  - obtenidos por los participantes en la dimensión Confortar se muestran en la Tabla 41. Se observa que tanto la Media como la Mediana ( $Q_2$ ) se ubican alrededor de valores centrales del rasgo para los modelos de la TRI y de valores relativamente altos para los modelos de la TCT. Los  $\theta$  estimados mediante los modelos de la TRI presentan correlaciones muy altas con los puntajes correspondientes de la TCT. Todas las distribuciones de los valores de los modelos de la tuvieron índices de asimetría negativos evidenciando que la mayoría de los participantes se ubicó por sobre la media. De manera similar a lo ocurrido en la dimensión Ayuda, la curtosis fue negativa para todos los modelos dicotómicos tanto de la TCT como de la TRI, muy próxima a 0 para el puntaje politómico de la TCT y positiva para los modelos politómicos de la TRI. La distribución de los  $\theta$  obtenidos mediante MRG es la única que ajusta a la distribución normal, desde que en la prueba de Kolmogorov-Smirnov no se rechazó la hipótesis nula que sostiene dicha distribución con  $p = 0.396$ .

Tabla 41

*Distribución de los puntajes y de los rasgos latentes estimados  
en la dimensión Confortar y prueba de normalidad*

		TCT <sub>DICO</sub>	TCT <sub>POLI</sub>	ML1P	ML2P	ML3P	MRG	MCP
Estadísticos Descriptivos	Media	4.87	30.46	-0.05	-0.06	-0.05	-0.02	-0.01
	DE	1.72	5.74	0.76	0.76	0.75	0.88	0.87
	Mín	0	9	-2.10	-2.08	-1.97	-3.61	-3.37
	Máx	12	42	1.01	0.97	0.99	2.09	2.05
	Q <sub>1</sub>	4	27	-0.46	-0.50	-0.47	-0.58	-0.54
	Q <sub>2</sub>	5	31	-0.04	0.02	0.02	0.02	0.03
	Q <sub>3</sub>	6	35	0.44	0.48	0.46	0.57	0.63
	Asimetría	-0.66	-0.41	-0.51	-0.51	-0.47	-0.24	-0.21
	Curtosis	0.27	-0.04	-0.21	-0.32	-0.37	0.26	0.16
<i>p</i> valor		<i>p</i> <	<i>p</i> <	<i>p</i> <	<i>p</i> <	<i>p</i> <	<i>p</i> =	<i>p</i> =
Prueba K-S		.0001	.0001	.0001	.0001	.0001	0.396	0.001
<i>r</i> con puntuación total TCT		1	1	.984**	.977**	.943**	.983**	.998**

\*\* $p < .0001$

Siguiendo el criterio de Inglés et al. (2008) la prevalencia de sujetos prosociales en la dimensión Confortar para la muestra es de entre un 0% y un 6.1%, siendo los modelos dicotómicos los que brindan valores nulos o casi nulos. El pobre nivel de información que brindan los modelos dicotómicos en los niveles altos del rasgo podría explicar esta falta de detección, tal como se puede observar ut supra en la Función de Información de la dimensión Confortar para ML1P, ML2P y ML3P.

Si el criterio se amplía y se considera como prosociales a aquellos participantes con más de un desvío por sobre la media, la prevalencia aumenta muy sustancialmente, a valores de entre 15.1% y 17.4%.

Utilizando el criterio de Inglés et al. (2008) de manera inversa, se puede considerar como sujetos muy poco prosociales a aquellos que tienen valores menores a un desvío y medio por debajo de la media. Así, se señaló con muy bajo nivel de conducta prosocial de confortar a entre 7.2% y 11.0% de los sujetos. Con el criterio negativo ampliado, se halla que entre 15.0% y 19.5% de los participantes reporta un bajo nivel en Confortar.

La mayoría de los sujetos se ubica en valores intermedios -entre menos de un desvío por debajo y por encima de la Media- en los puntajes y los rasgos latentes estimados de las conductas prosociales incluidas en la dimensión Confortar. En estos niveles medios se encuentran entre el 63.1% y el 69.1% de los participantes, siempre según el modelo.

Tabla 42

*Prevalencia de las conductas prosociales en Confortar*

	TCT <sub>DICO</sub>	TCT <sub>POLI</sub>	ML1P	ML2P	ML3P	MRG	MCP
% Media + 1.5 DE	0.4	4	0	0	0	6.1	6.1
% Media + 1.0 DE	17.4	15.1	17.0	17.0	17.0	15.9	15.1
% Media - 1.0 DE	19.5	15.9	19.6	16.6	16.8	15.0	15.9
% Media - 1.5 DE	10.9	7.2	11.0	8.4	8.5	7.2	7.2
% Intermedios	63.1	69.0	63.4	66.4	66.2	69.1	69.0

Los participantes con muy alto nivel en Confortar según el modelo TCT politómico (n = 56) tienen una media de edad de 20.59 años (DE 3.28) y 94.6% son mujeres. Las carreras más elegidas por estos participantes fueron Medicina: 28.6%, Lic. en Psicología: 21.4%, Abogacía: 12.5% y Lic. en Administración: 7.1%. En la muestra las proporciones de participantes que eligen estas carreras son Medicina: 10.1%, Lic. en Psicología: 30.4%, Abogacía: 7.0% y Lic. en Administración: 4.1%.

Los sujetos con muy baja prosocialidad en Confortar según el modelo TCT politómico (n = 99) tienen una media de edad de 22.61 años (DE = 6.11) y 57.0 % son mujeres. Las carreras que estos participantes más eligen son Lic. en Psicología: 22.2%, Contador Público: 16.2% y Medicina: 13.1%. Como ya se indicó, en muestra las proporciones de participantes que estudian estas carreras son Lic. en Psicología: 30.4%, Contador Público: 6.7% y Medicina: 10.1%.

## TERCERA PARTE

### DISCUSIÓN

#### *7.1. La Conducta Prosocial, Su Evaluación y Su Dimensionalidad*

Se ha generado en los últimos años una importante producción científica referida a las conductas prosociales. Sin embargo, no existe aún un consenso acabado entre los diferentes investigadores respecto de cuántas y cuáles son las dimensiones que componen este constructo (Auné et al., 2014). En consecuencia, la medición de las conductas prosociales se topa principalmente con los problemas que atañen a una variedad de definiciones teóricas y operacionales. Se ha observado que las tipologías más recientes han tendido a reducir el número de categorizaciones y a incluir en las mismas los sentimientos empáticos como un componente de la conducta prosocial para los adultos (Auné et al., 2016a; Auné, Abal, & Attorresi, 2017a). Se hace necesario, en este sentido, un consenso acerca del significado de la conducta prosocial en lo que respecta a beneficiar a otros y la especificación más clara de su relación con constructos relacionados, tales como empatía y altruismo.

La conducta prosocial es importante en la edad adulta por distintos motivos: se espera del adulto que pueda tener cierto grado de prosocialidad, ya que esto contribuye a una mejor convivencia creando lazos solidarios, y, así, a un mayor bienestar social, personal y colectivo. El comportamiento prosocial es considerado valioso ya que favorece los vínculos sociales y modera la agresividad (Carlo, Mestre, Samper, Tur, & Armenta, 2010) lo que es un recurso valorado por la sociedad. Las conductas prosociales pueden reforzar la autoestima de otros y la del que la realiza por actuar conforme a determinados valores. Existen conductas prosociales que son específicas de la adultez, como ser la donación voluntaria de sangre o desempeñarse en una profesión que implica ayuda aun poniendo en riesgo el bienestar personal. Es en la adultez donde la conducta prosocial tiene mayor importancia y valor social. Durante la edad adulta, de hecho, es donde la tendencia a actuar prosocialmente podría ser amenazada y contrarrestada por experiencias interpersonales que se basan en valores alternativos u objetivos tales como, por ejemplo, perseguir intereses propios o familiares, la competencia, el éxito y el poder individual. En todas las etapas etáreas los predictores

de la conducta prosocial son específicos (Caprara et al., 2005; Wenner & Randall, 2016) pero eventualmente diferentes.

En relación con la evaluación de la conducta prosocial, son escasos los instrumentos que cuentan con adecuadas evidencias de validez, así como estudios con base en la Teoría de la Respuesta al Ítem. Además, en concordancia con lo expuesto por González Portal (2000), se halló que no existe un cuerpo de trabajos referente a la correlación entre instrumentos de medida de conducta prosocial y tampoco se ha arribado a conclusiones claras acerca de las posibles relaciones entre los diferentes procedimientos de medida empleados (Auné et al., 2014). Asimismo, la mayoría de los instrumentos existentes tienden a especializarse en las etapas vitales de niñez y adolescencia. En niños la evaluación se realiza a través de escalas de estimación, realizadas por los adultos que rodean al niño y pruebas situacionales (Tremblay, Vitaro, Gagnon, Piché, & Reyer, 1992). Sin embargo, en la edad adulta -y también en la vejez-, como ya se ha mencionado, también la conducta prosocial es específicamente relevante.

Con el fin de obtener un instrumento específico para la adultez modelizado con la TRI que cuente con evidencias suficientes de validez y confiabilidad, se ha construido esta escala de conducta prosocial para adultos de población terciaria y universitaria argentina. El instrumento resultó de una aproximación al estudio de este complejo constructo, luego de una profunda revisión teórica acerca del mismo. Se construyó un *pool* de ítems que fue progresivamente depurado a través de sucesivos análisis factoriales y del análisis del DIF según el género.

A través de las sucesivas depuraciones, una estructura de cuatro factores (Confortar - Altruismo - Ayuda - Compartir y Donar) planteada originalmente fue dando lugar a una estructura bifactorial Ayuda - Confortar, más simple y parsimoniosa. En este trabajo se presenta una diferenciación bifactorial básica entre los comportamientos que suponen un apoyo emocional y aquellos que representan una asistencia y un compromiso solidario con un otro.

Los resultados fueron coherentes con la bibliografía existente. De manera amplia, los factores hallados coinciden con las dimensiones de la PSB de Penner et al., 1995. Se han hallado incluso distintas rutas de desarrollo y correlatos para diferentes tipos de acción prosocial (Dunfield, Kuhlmeier, O'Connell, & Kelley, 2011). Con respecto a Ayuda y Confortar, hallazgos recientes realizados mediante electroencefalograma indican que ambos comportamientos se sustentan en bases neurofisiológicas específicas y diferentes, con patrones separados de activación cerebral al

realizar cada uno de ellos (Paulus, Kühn-Popp, Licata, Sodian, & Meinhardt, 2013). Las definiciones de Ayuda y Confortar que utilizan Paulus et al. no se condicen exactamente con las utilizadas en la presente Tesis, tienen una vinculación en rasgos generales. Mientras que Confortar se vinculó con el lóbulo frontal izquierdo, un marcador de proceso emocional; la Ayuda se relacionó con el lóbulo temporal, probablemente reflejando la comprensión relacionada con objetivos. Estos resultados son coherentes con la correlación moderada hallada entre los factores. Por lo cual, se concluye que operacionaliza de manera adecuada a la conducta prosocial.

#### 7.1.1. La Dimensión Confortar

La dimensión Confortar incluye comportamientos de apoyo emocional (e.g. Ítem 12, *Les digo a los demás lo mucho que valen*) y percepciones empáticas (e.g. Ítem 5, *Siento todo dolor ajeno como propio*). Estos comportamientos se desarrollan alrededor del segundo año de vida y tienen como base la posibilidad de distinguir entre el sí mismo y los otros (Decety & Meyer, 2008) y la capacidad de regular las propias emociones negativas (Paulus, et al., 2013). Esta faceta se corresponde con la dimensión confortar de Dunfield (2014) y de Warneken y Tomasello (2009). Hay y Cook (2007) la describen como sentimientos por el otro y Caprara et al. (2005) la denominan ser empático. Es decir, que esta dimensión ha sido hallada reiteradamente (Auné, et al., 2016a; 2017a).

Como se ha mencionado, Caprara et al. (2005) destacan a la empatía como dimensión de la conducta prosocial en la adultez y proponen incluirla en las escalas para este ciclo vital. Para ellos, la capacidad de sentir las necesidades y estados de ánimo ajenos sería una parte integral de la conducta prosocial (Caprara, Capanna, Steca, & Paciello, 2005). Se ha resaltado la relevancia de la empatía en la disposición prosocial de las personas y su función inhibidora de la agresividad (Moya Albiol, Herrero, & Bernal, 2010). Cuanto mayor es la capacidad empática, es menos probable que se responda agresivamente y de forma vengativa ante un ofensor y mayor la probabilidad de una respuesta prosocial (Moreno & Fernández, 2011). Sin embargo, aunque la relación entre empatía y prosocialidad viene siendo estudiada desde hace años prevalece aún una importante indefinición conceptual e incluso de resultados (Auné et al., 2014; Sánchez-Queija, et al., 2006). Aunque es un tópico controvertido si la empatía en sí es parte de la prosocialidad (Auné et al., 2016a), existe un acuerdo generalizado acerca de que la acción de confortar sí lo es.

Parece ser aceptable que, aunque la mayoría de los seres humanos no son en extremo prosociales ni antisociales, anhelan de los demás una conducta empática. En cualquier cultura, la misantropía siempre es una excepción, nunca la norma (Rifkin, 2010). Actualmente, la expresión empática y su comprensión es considerada necesaria para un desarrollo saludable del ser humano (Rifkin, 2010). Prueba de ello son las extensas investigaciones sobre el autismo, afección donde justamente habría un déficit en estas capacidades. Actualmente, se ha considerado que la empatía en su aspecto cognitivo se relaciona con el concepto de Teoría de la Mente (e.g. Baron Cohen & Wheelwright, 2004). La Teoría de la Mente es definida como habilidad de explicar y predecir el comportamiento propio y ajeno mediante la atribución de estados mentales independientes (Gallagher & Frith, 2003). Los trastornos en la capacidad empática y el autismo serían consecuencias directas de fallas en esta habilidad la cual estaría asentada en mecanismos biológicos específicos (Elliott et al., 2006). Se observa como un desorden en principio cognitivo conduce a consecuencias sociales (Baron Cohen, et al. 1997).

La empatía incluida en los ítems de Confortar, es, sobre todo, afectiva. Ha resultado conformar una sola dimensión con consolar, acompañar y fortalecer emocionalmente a otro. Las acciones incluidas en Confortar forman parte de la vida cotidiana de la mayoría de las personas siendo realizadas, se infiere, generalmente en un círculo de apoyo social mutuo donde ya se conocen cognitivamente las necesidades emocionales de los demás.

La modelización de la dimensión Confortar con los modelos Logísticos de Uno, Dos y Tres Parámetros indicó que resulta más informativa para niveles bajos del rasgo, discriminando aceptablemente hasta niveles medios. El análisis con los Modelos de Respuesta Graduada y Crédito Parcial, evidenció que la dimensión en su versión politómica aumenta su discriminación niveles medio-altos del rasgo. El error de medición crece sustancialmente hacia niveles altos o muy altos del rasgo, según el modelo. Por lo tanto, si se incorporaran nuevos ítems a esta dimensión, sería aconsejable que los mismos requieran un nivel alto de Confortar para ser respondidos en forma favorable. Los ítems de esta dimensión pueden discriminar en forma alta según los modelos dicotómicos, alta según MRG y en forma moderada según MCP. Esta capacidad discriminativa así como el nivel de información alcanzado tuvieron valores adecuados para la totalidad de los ítems del factor y especialmente elevados para el ítem 15 (*Intento “levantar” la autoestima a mis amigos*). Este ítem aporta importante información desde niveles muy bajos a moderados del rasgo.

### 7.1.2. La Dimensión Ayuda

La dimensión Ayuda incluye comportamientos solidarios, así como una orientación general a dar, por ejemplo, dinero, esfuerzo o tiempo para mejorar la situación de otro. El comportamiento de ayuda se comienza a observar ya en niños de un año y medio (Warneken & Tomasello, 2006). Tanto confortar como ayudar son conductas positivas hacia un otro. Sin embargo, la realización de conductas incluidas en la dimensión Ayuda requiere, en general, un esfuerzo más alto y planeado que aquellas incluidas en la dimensión Confortar. De hecho, algunos de los ítems habían sido originalmente planteados para la dimensión Altruismo y suponen un nivel muy elevado de sacrificio (e.g. *Si tuviera un único plato de comida, lo compartiría con un desconocido*). Conviene recordar que la existencia misma del altruismo es un tópico controvertido, con defensores y detractores. Dados los elevados índices de ajuste de la estructura unidimensionalidad, se puede plantear una hipótesis a corroborar en posteriores estudios: considerar que el Altruismo es la forma más alta de la dimensión Ayuda. Su existencia no es aislada, sino que formaría parte de un continuo. En este sentido, se ha hallado una asociación entre actitudes altruistas y comportamientos de ayuda (Kahana, Bhatta, Lovegreen, Kahana, & Midlarsky, 2013). Además, este postulado sería coherente con la Hipótesis Social Heurística, una teoría que postula que las estrategias beneficiosas en la vida cotidiana se manifiestan también en escenarios menos frecuentes y que requieren una respuesta rápida, como son las situaciones de emergencia (Rand & Epstein, 2014).

La modelización de la dimensión Ayuda con los modelos Logísticos de Uno, Dos y Tres Parámetros, así como los Modelos de Respuesta Graduada y Crédito Parcial, evidenció que la dimensión resulta más informativa para niveles medio-altos -en los modelos dicotómicos- o altos -en los modelos politómicos- del rasgo. El análisis con el MRG muestra una precisión aceptable en un rango muy amplio de niveles del rasgo, donde la mayoría de las personas podría encuadrarse. El error de medición en el MRG sólo es muy elevado en personas con niveles extremos de Ayuda. Sin embargo, según el resto de los modelos de la TRI utilizados, el error de medición crece sustancialmente hacia niveles bajos del rasgo, por más que no sean extremos. Por lo tanto, si se incorporaran nuevos ítems a esta dimensión, sería aconsejable que los mismos requieran un nivel bajo de Ayuda para ser respondidos en forma favorable. Los ítems de esta dimensión pueden discriminar (según el parámetro  $a$ ) en forma alta según MRG y los modelos dicotómicos, y en forma aceptable según el MCP. Esta capacidad discriminatoria, así como el nivel de información

alcanzado fue mucho mayor para los ítems 6, 7 y 10, es decir que comprometerse con causas nobles, dedicar tiempo a que el mundo sea mejor y ocuparse del bienestar de otros -reactivos que denotan acciones concretas- fueron contenidos efectivos en detectar pequeñas variaciones en el nivel de Ayuda.

### *7.2. Características Generales de los Participantes*

Los resultados son coherentes con la hipótesis planteada de que los estudiantes terciarios y universitarios de profesiones que implican solidaridad, ayuda o conductas empáticas hacia otros puntuarán significativamente más alto en la escala que aquellos estudiantes de carreras que no involucren en forma directa una conducta prosocial y aportan evidencias de validez relacionada con un criterio externo, como lo es la carrera universitaria o terciaria que el participante está estudiando. Los estudiantes de Medicina y carreras afines manifestaron un alto nivel de conductas prosociales tanto en Ayuda como en Confortar con diferencias muy significativas con respecto a los estudiantes de carreras muestreadas que no involucran directamente las conductas prosociales. Los estudiantes de Lic. en Psicología y carreras afines, tal como era esperable, tuvieron significativamente mayores puntajes en la dimensión Confortar que los estudiantes de carreras muestreadas que no involucran directamente las conductas prosociales.

Otra coincidencia para ambas dimensiones, fue que la mayoría de los participantes se ubicó en niveles intermedios, con una minoría en niveles extremos. Si se utiliza el criterio de Inglés et al (2008) de considerar como prosociales a las personas con más de un desvío y medio de la media, se hallan prevalencias tanto en Ayuda como en Confortar mucho menores al 17.35 % hallado por este en una muestra de estudiantes adolescentes españoles con la escala TISS. Al tratarse de poblaciones, edades e instrumentos diferentes, sería forzada una comparación. Pero, es un punto importante a desarrollar en posteriores trabajos: la obtención de un instrumento con ítems etáreamente comunes que se pueda aplicar en la adolescencia, adultez y vejez y que permitiera efectuar comparaciones de la conducta prosocial entre estas etapas vitales.

En cuanto a la relación entre conducta prosocial y género, los resultados obtenidos concuerdan con las predicciones de teorías sociales y biológicas que sostienen una diferencia intergénero en la realización de estas conductas prosociales (Beutel & Johnson, 2004; Zahn-Waxler, Radke-Yarrow, Wagner, & Chapman, 1992). Sin embargo, realizando una revisión más amplia, la literatura no registra resultados consistentes. Los estudios realizados con los inventarios

NEO-PI-R (Costa & McCrae, 1999; Costa, Terracciano & McCrae, 2001) y BFQ (Caprara et al., 1995) aseguran que las mujeres presentan un nivel de altruismo significativamente mayor que los varones. Por otra parte, Smith (2003) encontró mayores puntajes para las mujeres en la escala de valores altruistas y la inexistencia de diferencias intergénero al evaluar altruismo a través de conductas concretas. Attorresi et al. (2008), utilizando una muestra de estudiantes universitarios argentinos, hallaron que varones y mujeres no reportaron diferencias significativas al comparar las puntuaciones obtenidas en la Escala de Altruismo Situacional (EAS).

El mayor puntaje obtenido por el género femenino en las dimensiones Ayuda y Confortar podría ser resultado de la vigencia de los estereotipos acerca de cómo es y debe ser un varón y una mujer. Por atribuir a las mujeres cualidades de amabilidad, calidez y simpatía los estereotipos de género suponen que las mujeres tienen una propensión para la unión con los demás en relaciones estrechas y diádicas. Las cualidades de expresividad y afectividad facilitarían las amistades, las relaciones románticas, y las relaciones familiares y de interdependencia cooperativa con otros (Fiske, Cuddy, Glick, & Xu, 2002). Por el contrario, al atribuirles cualidades de asertividad, ambición y competitividad a los hombres, los estereotipos de género crean un contexto social en el que las personas difieren en *status* y los varones se esfuerzan por mejorar su posición jerárquica (Gardner & Gabriel, 2004).

Las diferencias de género surgirían en la medida en que las percepciones de la conducta prosocial se confunden con las percepciones de la personalidad. No todas las mujeres que se perciben como prosociales, incluso por pares, necesariamente lo son, sino que podrían estar exhibiendo rasgos de conducta social exigidos por los estereotipos sociales consistentes con comportamientos prosociales. Es así que podría existir un solapamiento de la conducta prosocial y de dichos rasgos. La percepción del comportamiento prosocial podría estar más ligada al comportamiento general de la mujer que del hombre (Pursell et al., 2008). Por lo tanto, se hace discutible si a nivel global existe un género más prosocial que otro.

El hecho de haber alcanzado predicciones similares a partir de distintos modelos de la TRI constituye una importante evidencia de validez de constructo. Además de hallar una relación esperable con el género y la carrera que estudia el participante, los puntajes en Ayuda y Confortar han correlacionado de forma significativa con las dimensiones de la EHP y con las preguntas de Autopercepción de la Empatía, Conducta Prosocial, Dadivosidad y Altruismo. De esta forma, se

han brindado una diversidad de evidencias de validez. En resumen, respecto a las propiedades psicométricas de la escala de conducta prosocial, el análisis resulta coherente con el modelo planteado, acompañándose por una adecuada confiabilidad. Se destaca la coherencia de los resultados obtenidos con la bibliografía psicológica existente en el campo.

### *7.3. Limitaciones y Futuras Líneas de Investigación*

Para las dimensiones Ayuda y Confortar se halló en el MCP una alta proporción de ítems con *reversal*. Como ya se mencionó, esto puede estar señalando que sería deseable un agrupamiento de ciertas categorías de respuesta (Sam, et al., 2016; Stump et al., 2012). El comportamiento de los Modelos Politémicos de la TRI, además, puede verse modulado por la cantidad de categorías de la escala Likert que se utiliza. Por lo tanto, se realizará un nuevo estudio reduciendo las categorías a cuatro y cinco opciones para evaluar si existe una mejora en los indicadores psicométricos de los ítems.

Al estar la muestra circunscripta a estudiantes universitarios, los resultados obtenidos deberían ser corroborados en estudiantes de otros niveles educativos. Una muestra más heterogénea permitirá potenciar notablemente las posibilidades de generalización de los resultados, además se podrán realizar comparaciones de interés. Por otra parte, sería conveniente estudiar las condiciones para la existencia de estabilidad a lo largo de los ciclos vitales en este rasgo, así como incorporar otros criterios externos a la escala como evidencia de validez y evidencias adicionales de ajuste de los ítems a los modelos de la TRI. Sería especialmente importante estudiar la relación entre los niveles de conducta prosocial y los rasgos de personalidad. En futuros desarrollos sería asimismo útil la construcción de instrumentos de medición específicos de la conducta prosocial en la adultez para distintos ámbitos, tales como el clínico, jurídico, educacional y laboral.

Otro aspecto que sería relevante estudiar es la construcción de modelos de redes semánticas (Vivas, 2004; Vivas, Comesaña, & Vivas, 2007) relacionadas con las conductas prosociales. Este aspecto podría contribuir especialmente dada la relativamente baja prevalencia de este tipo de conductas en la muestra utilizada en la presente Tesis a la formulación de programas para aumentar las conductas prosociales.

El hecho de contar con un instrumento adecuado para la evaluación de la conducta prosocial puede contribuir al diseño e implementación de planes integrales de mejoramiento conducentes a potenciar habilidades sociales y ambientes de convivencia más armoniosos y productivos. La temática expuesta, brinda información útil para continuar con la profundización de una problemática actual.

## REFERENCIAS BIBLIOGRÁFICAS

- Abad, F., Olea, J., Ponsoda, V., & García, C. (2011). *Medición en ciencias sociales y de la salud*. Madrid: Síntesis.
- Abal, F. J. P. (2013). *Comparación de modelos politómicos y dicotómicos de la Teoría de la Respuesta al Ítem aplicados a un test de Comportamiento Típico*. Tesis de Doctorado, Facultad de Psicología de la Universidad de Buenos Aires.
- Abal, F. J. P., Auné, S. E., & Attorresi, H. F. (2014). Comparación del Modelo de Respuesta Graduada y la Teoría Clásica de Tests en una Escala de Confianza para la Matemática. *Summa Psicológica UST*, 11(2), 101-113.
- Abal, F. J. P., Auné, S. E., & Attorresi, H. F. (2015). Relación entre la Deseabilidad Social y los parámetros del Modelo Logístico de tres parámetros. *Psicología, Conocimiento y Sociedad*, 5(1), 5-24.
- Abal, F. J. P., Auné, S. E., Lozzia, G. S., & Attorresi, H. F. (2015). Modelización de una prueba de afecto hacia la matemática con la teoría de respuesta al ítem. [*Modelling affect towards mathematics scale using item response theory*]. *Revista de Psicología*, 11(21), 23-34.
- Abal, F. J. P., Auné, S. E., Lozzia, G. S., & Attorresi, H. F. (2017). Funcionamiento de la categoría central en ítems de Confianza para la Matemática. *Revista Evaluar. En prensa*.
- Abal, F. J. P., Lozzia, G., Aguerri, M. E., Galibert, M. S., & Attorresi, H. F. (2007). Construcción de una escala de voluntad de trabajo. *Investigaciones en Psicología*, 12(1), 7-16.
- Abal, F. J. P., Lozzia, G., Aguerri, M. E., Galibert, M. S., & Attorresi, H. F. (2008). Evidencias acerca de la validez de constructo de la escala Voluntad de Trabajo. *Anuario de Investigaciones de la Facultad de Psicología de la Universidad de Buenos Aires*, 15(1), 283-289.
- Abal, F. J. P., Lozzia, G., Aguerri, M. E., Galibert, M. S., & Attorresi, H. F. (2010). La escasa aplicación de la Teoría de Respuesta al Ítem en Tests de Ejecución Típica. *Revista Colombiana de Psicología*, 19(1), 111-122.
- Abal, F. J. P., Lozzia, G., Blum, G. D., Galibert, M. S., Aguerri, M. E., & Attorresi, H. F. (2010). Análisis de ítem de un test de Altruismo a partir de un Modelo Logístico de un parámetro. *Perspectivas en Psicología*, 7(1), 16-23.
- Abal, F. J. P., Lozzia, G., Blum, G. D., Galibert, M. S., Aguerri, M. E., & Attorresi, H. F. (2010). Comparación de varones y mujeres respecto de la voluntad de trabajo. Póster. *II Encuentro de Docentes e Investigadores de Estadística en Psicología Facultad de Psicología, UBA*. 18 al 20 de noviembre de 2010. Memorias en el CD del Congreso ISSN 1852-3560. Eje Temático I, pp. 6-7.
- Abal, F. J. P., Lozzia, G. S., & Galibert, M. S. (2008). Aplicación del modelo logístico de dos parámetros en una escala de altruismo [*Application of the two-parameter logistic model to the ítem analysis of an altruism scale*]. *Memorias de las XV Jornadas de Investigación y Cuarto Encuentro de Investigadores en Psicología del Mercosur*, 2, 453-454.
- Achenbach, T. M. & Edelbrock, C. A. (1983). *Manual for the Child Behavior Checklist and Revised Child Behavior Profile*. Burlington, Author, VM.
- Achenbach, T. M. & Edelbrock, C. A. (1986). *Manual for the Teacher's Reports and Teacher Version of the Child Behavior Profile*, Burlington, Author, VM.
- Aguado, D., Rubio, V. J., Hontangas, P. M., & Hernández, J. M. (2005). Propiedades psicométricas de un test adaptativo informatizado para la medición del ajuste emocional. *Psicothema*, 17, 484-491.
- Aguerri, M. E. (2012). *Detección del Funcionamiento Diferencial del Ítem mediante la Prueba de Breslow-Day*. Tesis de Doctorado, Facultad de Psicología de la Universidad de Buenos Aires.
- Aguerri, M. E., Galibert, M. S., Zanelli, M. L., & Attorresi, H. F. (2005). Detección errónea del funcionamiento diferencial del ítem. Una comparación de métodos. *Psicothema*, 17(2), 350-355.
- Aguerri, M. E., Prieto Marañón, P., Galibert, M. S., & Attorresi, H. F. (2011). DIF paralelo vs. DIF uniforme. Simposio Funcionamiento Diferencial de los Ítems. *Actas del XII Congreso de Metodología de las Ciencias Sociales y de la Salud. Asociación Española de Metodología de las Ciencias del Comportamiento (AEMCCO)*. San Sebastián, España. 19-22 de Julio de 2011. pp. 244-252.
- Aguerri, M. E., Prieto Marañón, P., Galibert, M. S., & Attorresi, H. F. (2013). Evaluación del DIF Unidireccional mediante las pruebas de Pruebas de Breslow-Day. Presentación realizada en el Symposium: Funcionamiento Diferencial del Ítem. *XIII Congreso de Metodología de las Ciencias Sociales y de la Salud*. Tenerife, España 3-7 de Septiembre de 2013.
- Aguerri, M. E.; Zanelli, M. L.; Galibert, M. S., & Attorresi, H. F. (2002). Evaluación de un método empírico para detectar el funcionamiento diferencial del ítem. *Interdisciplinaria. Revista de Psicología y Ciencias Afines. Journal of Psychology and Related Sciences*, 19(2), 185-203.

- Aknin, L. B., Broesch, T., Hamlin, J. K. & van de Vondervoort, J. W. (2015). Prosocial behavior leads to happiness in a small-scale rural society. *Journal of Experimental Psychology: General*, 144(4), 788-795. doi: 10.1037/xge0000082
- Alarcón, R. (2006). Desarrollo de una escala factorial para medir la felicidad. *Revista Interamericana de Psicología*, 40(1), 95-102.
- Al Nima, A., & Garcia, D. (2015). Factor Structure of the Happiness-Increasing Strategies Scales (H-ISS): activities and coping strategies in relation to positive and negative affect. *PeerJ*, 3, e1059. doi:10.7717/peerj.1059.
- Amato, P. R. (1983). Helping behavior in urban and rural environments: Field studies based on a taxonomic organization of helping episodes. *Journal of Personality and Social Psychology*, 45(3), 571-586 doi: 10.1037/0022-3514.45.3.571.
- American Psychological Association (2010). *Ethical Principles for Psychologists and Code of Conduct*. Washington, D.C.: APA.
- Ames, A. J., & Penfield, R. D. (2015). An NCME Instructional Module on Item-Fit Statistics for Item Response Theory Models. *Educational Measurement: Issues and Practice*, 34(3), 39-48. doi:10.1111/emip.12067
- Anastasi, A., & Urbina, S. (1998). *Tests psicológicos*. México: Pearson Educación.
- Andrés, M. L., Castañeiras, C. E., & Richaud de Minzi, M. C. (2014). Relaciones entre la personalidad y el bienestar emocional en niños. El rol de la regulación emocional. *Cuadernos de Neuropsicología/Panamerican Journal of Neuropsychology*, 8(2), 217-241. doi: 10.7714/cnps/8.2.205
- Andrich, D. (1988). A general form of Rasch's extended logistic model for partial credit scoring. *Applied Measurement in Education*, 1, 363-378.
- Andrich, D. (1995). Distinctive and incompatible properties of two common classes of IRT model for graded responses. *Applied Psychological Measurement*, 19, 101-119.
- Anguera, M. T., Arnau, J., Ato, M., Martínez, M. R., Pascual, J., & Vallejo, G. (1995). *Métodos de investigación en Psicología*. Madrid: Síntesis.
- Argyle, M. (2003). *The psychology of happiness* (2nd ed.). St. Petersburg: Peter.
- Ashton, M., Paunonen, S. V., Helmes, E., & Douglas, N. (1998). Kin altruism, reciprocal altruism, and the Big Five personality factors. *Evolution and Human Behavior*, 19(4), 243-255.
- Asociación Argentina de Estudio e Investigación en Psicodiagnóstico. (1999). *Código de Ética del Psicodiagnostador*.
- Asún, R. & Zuñiga, C. (2008). Ventaja de los modelos politómicos de teoría de respuesta al ítem en la medición de actitudes sociales. El análisis de un caso. *Psykhé*, 17(2), 103-115.
- Ato, M., López, J. J., & Benavente, A. (2013). Un sistema de clasificación de los diseños de investigación en psicología. *Anales de Psicología*, 29 (3), 1038 - 1059. doi: 10.6018/analesps.29.3.178511.
- Attorresi, H. F., Abal, F., Lozzia, G., Picón Janeiro, J., Aguerri, M. E., & Galibert, M. S. (2008). Construcción de una escala para la medición del Altruismo. *Investigando en Psicología*, 10(10), 7-24.
- Attorresi, H. F., Abal, F., Galibert, M. S., Lozzia, G., & Aguerri, M. E. (2011). Aplicación del Modelo de Respuesta Graduada a una Escala de Voluntad de Trabajo. *Interdisciplinaria*, 28(2), 231-244.
- Attorresi, H. F., Galibert, M. S., & Aguerri, M. E. (2002). Modelo de medición de componente no observable. Una presentación formal de la axiomatización de la teoría clásica de tests. *Psicothema*, 14(3), 665-668.
- Attorresi, H. F., Lozzia, G. S., Abal, F. J. P., Galibert, M. S., & Aguerri, M. E. (2009). Teoría de Respuesta al Ítem. Conceptos básicos y aplicaciones para la medición de constructos psicológicos. *Revista Argentina de Clínica Psicológica*, 18(2), 179-188.
- Auné, S. E., Abal, F. J. P., & Attorresi, H. F. (2014). Versión argentina de la Escala de Habilidades Prosociales de Morales Rodríguez y Suárez Pérez (2011). *Praxis. Revista de Psicología*, 16(26), 31-48.
- Auné, S. E., Abal, F. J. P., & Attorresi, H. F. (2016a). Antagonismos entre concepciones de empatía y su relación con la conducta prosocial. *Revista de Psicología*, 17(2), 137-149.
- Auné, S., Abal, F. J. P., & Attorresi, H. F. (2016b). Diseño y construcción de una escala de conducta prosocial para adultos. *Revista Iberoamericana de Diagnóstico y Evaluación Psicológica*, 42(2), 15-25. doi: 10.21865/RIDEP42\_15
- Auné, S. E., Abal, F. J. P., & Attorresi, H. F. (2017a). Propiedades psicométricas de una prueba de conducta empática. *Revista Iberoamericana de Diagnóstico y Evaluación Psicológica. En Prensa*.
- Auné, S. E., Abal, F. J. P., & Attorresi, H. F. (2017b). Versión argentina de la Escala de Felicidad de Lima. Manuscrito enviado para su publicación.
- Auné, S. E. & Attorresi, H. F. (2014). Dimensiones de la conducta prosocial en la construcción de un test. VI Congreso Internacional de Investigación y Práctica Profesional en Psicología. XXI Jornadas de Investigación. X

- Encuentro de Investigadores en Psicología del Mercosur. Facultad de Psicología, UBA. "Adicciones: desafíos y perspectivas para la investigación científica y la práctica profesional".
- Auné, S. E., Blum, G. D., Abal, F. J. P., & Attorresi, H. F. (2014). Redacción de ítems de un test de conducta prosocial. I Congreso Internacional de Psicología. IV Congreso Nacional de Psicología "Ciencia y Profesión". 23, 24 y 25 de abril de 2014. Universidad Nacional de Córdoba, Argentina.
- Auné, S. E., Blum, G. D., Abal, F. J. P., Lozzia, G., & Attorresi, H. F. (2014). La Conducta Prosocial: estado actual de la investigación. *Perspectivas en Psicología*, 11(2), 21-33.
- Aydinli, A., Bender, M., & Chasiotis, A. (2013). Helping and volunteering across cultures: Determinants of prosocial behavior. *Online Readings in Psychology and Culture*, 5(3), 1-27. doi: 10.9707/2307-0919.1118
- Badaracco, M., & Martínez, L. (2010). Sistema Tutor Inteligente con Arquitectura inspirada en la Enseñanza Basada en Competencias (STI-C). En *XVI Congreso Argentino de Ciencias de la Computación*.
- Baessler, J. & Schwarzer, R. (1996). Evaluación de la autoeficacia: Adaptación española de la escala de Autoeficacia General. *Ansiedad y Estrés*, 2, 1-8.
- Baker, F. B. (2001). *The basics of item response theory*. En ERIC clearinghouse on assessment and evaluation. College Park, MD: University of Maryland. Disponible en: <http://ericae.net/irt/baker>.
- Baker, J., Rounds, J., & Zevon, M. (2000). A Comparison of Graded Response and Rasch Partial Credit Models with Subjective Well-Being. *Journal of Educational and Behavioral Statistics*, 25, 253-270.
- Bandura, A. (1991). Social cognitive theory of moral thought and action. En W. M. Kurtines & J. L. Gewirtz (Eds.), *Handbook of moral behavior and development: Theory, research and applications* (Vol. 1, pp. 71-129). Hillsdale, NJ: Erlbaum.
- Bandura, A. (1999). Moral Disengagement in the Perpetration of Inhumanities. *Personality and Social Psychology Review*, 3(3), 193-209.
- Baron Cohen, S., Jolliffe, T., Mortimore, C., & Robertson, M. (1997). Another advanced test of Theory of mind: Evidence from very High functioning adults with autism or Asperger Syndrome. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 38(7), 813-822.
- Baron Cohen, S., & Wheelwright, S. (2004). The empathy quotient: an investigation of adults with Asperger syndrome or high functioning autism, and normal sex differences. *Journal of autism and developmental disorders*, 34(2), 163-175.
- Bartlett, M. Y., Condon, P., Cruz, J., Baumann, J., & Desteno, D. (2012). Gratitude: Prompting behaviours that build relationships. *Cognition & Emotion*, 26(1), 2-13. doi: 10.1080/02699931.2011.561297
- Batson, C. D. (1991). *The altruism question: Toward a social psychological answer*. Hillsdale, NJ: Erlbaum.
- Batson, C. D., & Ahmad, N. Y. (2009). Using empathy to improve intergroup attitudes and relations. *Social Issues and Policy Review*, 3(1), 141-177.
- Batson, C. D., & Powell, A. (2003). Altruism and prosocial behavior. En M. Theodore (Ed.) y L. Melvin (Ed.) *Handbook of Psychology: Personality and Social Psychology* (pp 282-316). Nueva York: John Wiley y Sons, Inc. XIX.
- Batson, C. D., Van Lange, P. A., Ahmad, N., & Lishner, D. A. (2007). *Altruism and helping behavior*. The SAGE Handbook of Social Psychology: Concise Student Edition, 241.
- Baumeister, R. F., & Leary, M. R. (1995). The need to belong: Desire for interpersonal attachments as a fundamental human motivation. *Psychological Bulletin*, 117(3), 497-529.
- Bearden, W. O., Sharma, S., & Teel, J. E. (1982). Sample size effects on chi square and other statistics used in evaluating causal models. *Journal of Marketing Research*, 19(4), 425-430. doi:10.2307/3151716
- Beilin, H. (2013). *The development of prosocial behavior*. N. Eisenberg (Ed.). Academic Press.
- Bekkers, R. (2006). Traditional and Health Related Philanthropy: The Role of Resources and Personality. *Social Psychology Quarterly*, 68(4), 349-366.
- Bentosela, M., & Mustaca, A. (2005). Efectos cognitivos y emocionales del envejecimiento: aportes de investigaciones básicas para las estrategias de rehabilitación. *Interdisciplinaria*, 22, 211-235.
- Berkowitz, L., & Lutterman, K. G. (1968). The traditional socially responsible personality. *Public Opinion Q.*, 32, 169-185.
- Bermúdez, J. (1995). *Manual del Cuestionario "Big Five" (BFQ)*. Madrid: TEA.
- Berra, S., Tebé, C., Esandi, M. E., & Carignano, C. (2013). Fiabilidad y validez del cuestionario KIDSCREEN-52 para medir calidad de vida relacionada con la salud para población argentina de 8 a 18 años. *Archivos Argentinos de Pediatría*, 111(1), 29-35. doi: 10.5546/aap.2013.29
- Besser, A., Amir, M., & Barkan, S. (2004). Who signs an organ transplant donor card? A study of personality and individual differences in a sample of Israeli university students. *Personality and Individual Differences*, 36(7), 1709-1723.

- Beutel, A. M., & Johnson, M. K. (2004). Gender and prosocial values during adolescence: a research note. *The Sociological Quarterly*, 45, 379-393.
- Birnbaum, A. (1968). Some latent trait models and their use in inferring an examinee's ability. In F. M. Lord & M. R. Novick (Eds.), *Statistical Theories of Mental Test Scores*. Reading, MA: Addison Wesley.
- Bock, R. D. (1972). Estimating item parameters and latent ability when responses are scored in two or more latent categories. *Psychometrika*, 37(1), 29-51.
- Bock, R. D., & Aitkin, M. (1981). Marginal maximum likelihood estimation of item parameters: Application of an EM algorithm. *Psychometrika*, 46(4), 443-459.
- Bock, R., & Lieberman, M. (1970). Fitting a response model for dichotomously scored items. *Psychometrika*, 35(2), 179-197. doi: 10.1007/BF02291262
- Böckler, A., Tusche, A. & Singer, T. (2016). The Structure of Human Prosociality Differentiating Altruistically Motivated, Norm Motivated, Strategically Motivated, and Self-Reported Prosocial Behavior. *Social Psychological and Personality Science*, 1-12. doi: 10.1177/1948550616639650
- Boomsma, A., Van Duijn, M. A. J., & Snijders, T. A. B. (2001). *Essays on item response theory*. New York: Springer-Verlag.
- Borgatta, E. P., & Corsini, R. J. (1960). The Quick Word Test (QWT). *J. Educ. Res.* 54, 15-20.
- Breslow, N. E., & Day, N. E. (1980). *Statistical Methods in Cancer Research*. Volume I. The Analysis of Case-Control Studies. Lyon, France. International Agency for Research on Cancer (IARC Scientific Publication No. 32).
- Brown, T. (2015). *Confirmatory factor analysis for applied research* (2nd Ed.). New York: Guilford Press.
- Brown, E., & Ferris, J. M. (2007). Social Capital and Philanthropy: An Analysis of the Impact of Social Capital on Individual Giving and Volunteering. *Non profit and Voluntary Sector Quarterly*, 36(1), 85-99.
- Buja, A., & Eyuboglu, N. (1992). Remarks on parallel analysis. *Multivariate Behavioral Research*, 27(4), 509-540. doi:10.1207/s15327906mbr2704\_2
- Byrne, B. M. (2012). *Structural equation modeling with Mplus: Basics, concepts, applications, and programming*. New York: Routledge.
- Camilli, G., & Shepard, L. A. (1994). *Methods for identifying biased test items*. Thousand Oaks: Sage.
- Caprara, G. V., Alessandri, G., Di Giunta, L., Panerai, L., & Eisenberg, N. (2010). The contribution of agreeableness and self-efficacy beliefs to prosociality. *European Journal of Personality*, 24(1), 36-55.
- Caprara, G. V., Alessandri, G., & Eisenberg, N. (2012). Prosociality: The contribution of traits, values, and self-efficacy beliefs. *Journal of Personality and Social Psychology*, 102(6), 1289-1303.
- Caprara, G. V., Barbaranelli, C., Pastorelli, C., Bandura, A., & Zimbardo, P. (2000). Prosocial foundations of children's academic achievement. *Psychological Science*, 11, 302-306. doi:10.1111/1467-9280.00260
- Caprara, G. V., Borgogni, L., Barbaranelli, C., & Moreno, J. B. (1995). *BFQ: cuestionario "Big Five": manual*. TEA Ediciones.
- Caprara, G. V., Capanna, C., Steca, P., & Paciello, M. (2005). Misura e determinanti personali della prosocialità. Un approccio sociale cognitivo. *Giornale italiano di psicologia*, 32(2), 287-308.
- Caprara, G. V., & Pastorelli, C. (1993). Early emotional instability, prosocial behavior, and aggression: some methodological aspects. *European Journal of Personality*, 7(1), 19-36.
- Caprara, G. V., & Steca, P. (2005). Affective and social self-regulatory efficacy beliefs as determinants of positive thinking and happiness. *European Psychologist*, 10(4), 275-286. doi: 10.1027/1016-9040.10.4.275
- Caprara, G. V., Steca, P., Zelli, A., & Capanna, C. (2005). A new scale for measuring adults' prosocialness. *European Journal of Psychological Assessment*, 21(2), 77-89. doi: 10.1027/1015-5759.21.2.77
- Cardozo, G. & Alderete, A. (2009). Adolescentes en riesgo psicosocial y resiliencia. *Psicología desde el Caribe*, 23, 148-182.
- Carlo, G., Crockett, L. J., Randall, B. A., & Roesch, S. C. (2007). A latent growth curve analysis of prosocial behavior among rural adolescents. *Journal of Research on Adolescence*, 17(2), 301-324.
- Carlo, G., Eisenberg, N., & Knight, G. P. (1992). An objective measure of adolescents' prosocial moral reasoning. *J. Res. Adolesc.*, 2, 331-349.
- Carlo, G., Hausmann, A., Christiansen, S., & Randall, B. A. (2003). Sociocognitive and behavioral correlates of a measure of prosocial tendencies for adolescents. *The journal of early adolescence*, 23(1), 107-134. doi: 10.1177/0272431602239132
- Carlo, G., Knight, G. P., McGinley, M., & Hayes, R. (2011). The roles of parental inductions, moral emotions, and moral cognitions in prosocial tendencies among Mexican American and European American early adolescents. *The Journal of Early Adolescence*, 31(6), 757-781. doi: 10.1177/0272431610373100

- Carlo, G., Mestre Escrivá, M. V., Samper, P., Tur, A. M., & Armenta, B. E. (2010). Feelings or cognitions? Moral cognitions and emotions as longitudinal predictors of prosocial and aggressive behaviors. *Personality and Individual Differences, 48*, 872-877.
- Carlo, G., Okun, M. A., Knight, G. P., & de Guzman, M. R. T. (2005). The interplay of traits and motives on volunteering: agreeableness, extraversion and prosocial value motivation. *Personality and Individual Differences, 38*(6), 1293-1305.
- Carlo, G., & Randall, B. A. (2002). The Development of a Measure of Prosocial Behaviors for Late Adolescents. *Journal of Youth and Adolescence, 31*(1), 31-44. doi: 10.1023/A:1014033032440
- Carmine, E. G., & Zeller, R. A. (1979). *Reliability and validity assessment*. Londres: Sage.
- Castro Solano, A. (2013). La psicología positiva en América Latina. Desarrollos y perspectivas. *Psiencia. Revista Latinoamericana de Ciencia Psicológica, 4*(2), 108-116.
- Castro Solano, A. & Sánchez-López, M. P. (2000). Objetivos de vida y satisfacción autopercebida en estudiantes universitarios. *Psicothema, 12*(1), 87-92.
- Casullo, M. M. & Fernández Liporace, M. (2005). Evaluación de la capacidad de perdonar. Desarrollo y validación de una escala. *Acta Psiquiátrica y Psicológica de América Latina, 51*(1), 14-20.
- Casullo, M. M., & Fernández Liporace, M. (2007). Las propuestas de la Psicología Positiva: ¿Universales psicológicos o particulares de una visión cultural? *Anuario de investigaciones, 14*, 261-268.
- Cervantes, V. H., Cepeda, E., & Camargo, S. L. (2008). Una propuesta para la obtención de niveles de desempeño en los modelos de Teoría de Respuesta al Ítem. *Avances en Medición, 6*(1), 49-58.
- Chacón, F. (1986). Una aproximación al concepto psicosocial de altruismo. *Boletín de Psicología, 11*, 41-62.
- Cho, S., Drasgow, F., & Cao, M. (2015). An investigation of emotional intelligence measures using item response theory. *Psychological assessment, 27*(4), 1241-1252. doi: 10.1037/pas0000132
- Cialdini, R. B., Schaller, M., Houlihan, D., Arps, K., Fultz, J., & Beaman, A. L. (1987). Empathy-based helping: Is it selflessly or selfishly motivated?. *Journal of personality and social psychology, 52*(4), 749-758.
- Cid, P., Díaz, A., Pérez, M. V., Torruella, M., & Valderrama, M. (2008). Agresión y violencia en la escuela como factor de riesgo del aprendizaje escolar. *Ciencia y Enfermería, 14*(2), 21-30.
- Clary, E. G., & Snyder, M. (1991). A functional analysis of altruism and prosocial behavior: The case of volunteerism. In M. Clark (Ed.), *Review of personality and social psychology* (Vol. 12, pp. 119-148). Newbury Park, CA: Sage.
- Clary, E. G., Snyder, M., Ridge, R. D., Copeland, J., Stukas, A. A., Haugen, J., & Miene, P. K. (1998). Understanding and assessing the motivations of volunteers: A functional approach. *Journal of Personality and Social Psychology, 74*(6), 1516-1530.
- Clauser, B. E., & Mazor, K. M. (1998). Using statistical procedures to identify differential functioning test items. *Educational Measurement: Issues and Practice, 17*, 31-44.
- Cnossen, S. (2015). *Statistical models of children's strategy change in analogical reasoning*. Tesis de Maestría. Facultad de Ciencias Sociales, Universidad de Leiden.
- Cohen, A. S., Kim, S. H., & Baker, F. B. (1993). Detection of differential item functioning in the graded response model. *Applied psychological measurement, 17*(4), 335-350.
- Cohen, R. J., & Swerdlik, M. E. (2006). *Pruebas y evaluación psicológicas*. México: McGraw-Hill.
- Compas, B., Davis, D., Forsythe, B., & Warner, T. (1987). Assessment of major and daily stress full events during adolescence. The adolescent Perceived Events Scale. *Journal of consulting and Clinical Psychology, 55*(4), 534-541.
- Cortada de Kohan, N. (2004). Teoría de respuesta al ítem: supuestos básicos. *Evaluar, 4*, 95-110.
- Cortada de Kohan, N., & Cupani, M. (2008). Teoría de respuesta al ítem. En S. Tornimbeni, E. Pérez & F. Olaz (Eds.), *Introducción a la Psicometría*, (pp. 217-242). Buenos Aires: Paidós.
- Cortada de Kohan, N. & Macbeth, G. (2005). Características psicosociales de la familia argentina. En: Congreso Interamericano de Psicología, Sociedad Interamericana de Psicología.
- Cortada de Kohan, N. & Macbeth, G. (2007). Construction of a mathematical test for adolescents and adults, *Interdisciplinaria, 24*(1), 43-64.
- Cortada de Kohan, N., Macbeth, G., & López Alonso, A. O. (2008). *Técnicas de investigación científica con aplicaciones en Psicología, Ciencias Sociales y de la Educación*. Lugar: Buenos Aires.
- Cortada de Kohan, N., & Tornimbeni, S. (2004). Paradigmas de la Psicometría. En S. Tornimbeni, E. Pérez, F. Olaz & A. Fernández (Eds.), *Introducción a los Tests Psicológicos*, (pp. 23-44). Córdoba: Editorial Brujas.
- Costa, P. T. & McCrae, R. R. (1999). *Manual, NEO PI-R, Inventario de Personalidad NEO Revisado y NEO-FFI, Inventario NEO reducido de Cinco Factores*. Madrid: TEA Ediciones, S.A.

- Costa, P. T., Terracciano, A. & McCrae, R. R. (2001). Gender Differences in Personality Traits across Cultures. Robust and surprising findings. *Journal of Personality and Social Psychology*, 81, 322-331.
- Crespi, M. C., & Mikulic, I. M. (2009). Reinserción social de liberados condicionales: análisis de la dimensión relacional desde el enfoque de redes de apoyo social. *Anuario de investigaciones*, 16, 211-221.
- Cronbach, L. J. (1972). *Fundamentos de la exploración psicológica*. Madrid: Biblioteca Nueva.
- Crowne, D. P., & Marlowe, D. (1964). *The Approval Motive*. New York: Wiley.
- Csikszentmihalyi, M. (2008). *El yo evolutivo. Una psicología para un mundo globalizado*. Barcelona: Kairos.
- Csikszentmihalyi, M. (2013). *The Flow: The Psychology of Optimal experience* (3rd ed.). Moscow: Alpina non-fiction.
- Cuadrado, E., Taberner, C., & Steinel, W. (2016). Prosocial behaviour, inclusion and exclusion: why and when do we behave prosocially?/ Comportamiento prosocial, inclusión y exclusión: ¿cuándo y por qué adoptamos comportamientos prosociales?. *Revista de Psicología Social*, 31(3), 463-499. doi: 10.1080/02134748.2016.1190126
- Darley, S., & Latané, B. (1968). Bystander intervention in emergencies: Diffusion of responsibility. *Journal of Personality and Social Psychology*, 8, 377-388.
- Davis, M. H. (1983). Measuring individual differences in empathy: Evidence for a multidimensional approach. *J. Pers. Soc. Psychol.*, 44, 113-126.
- DeCarlo, L. T. (1997). On the meaning and use of kurtosis. *Psychological Methods*, 2(3), 292-307. doi:10.1037/1082-989X.2.3.292
- Decety, J., & Meyer, M. (2008). From emotion resonance to empathic understanding: A social developmental neuroscience account. *Development and Psychopathology*, 20(04), 1053-1080. doi:10.1017/S0954579408000503
- De Champlain, A. F. (2010). A primer on Classical Test Theory and Item Response Theory for Assessments in Medical Education. *Medical Education*, 44, 109-117.
- De Champlain, A., & Gessaroli, M. E. (1991). *Assessing Test Dimensionality Using an Index Based on Nonlinear Factor Analysis*. Paper presented at the Annual Meeting of the American Educational Research Association, Chicago, IL.
- Demars, C. E. (2005). Type I errors for PARSCALE's fit index. *Educational and Psychological Measurement*, 65, 42-50. doi: 10.1177/0013164404264849
- DeVellis, R. (2012). *Scale development. Theory and applications*. New York: SAGE
- Digman, J. M. (1990). Personality structure: Emergence of the five-factor model. *Annual Review of Psychology*, 41(1), 417-440.
- Dodd, B. G., & Koch, W. R. (1987). Effects of variations in item step values on item and test information in the partial credit model. *Applied Psychological Measurement*, 11(4), 371-384.
- Donaldson, S. I., Dollwet, M., & Rao, M. A. (2015). Happiness, excellence, and optimal human functioning revisited: Examining the peer-reviewed literature linked to positive psychology. *The Journal of Positive Psychology*, 10(3), 185-195. doi: 10.1080/17439760.2014.943801
- Dorans, M. J., & Holland, P. W. (1993). DIF detection and description: Mantel-Haenszel and Standardization. En P. W. Holland y H. Wainer (Eds.). *Differential Item Functioning* (pp. 35-66). Hillsdale, NJ: Erlbaum.
- Dorans, M. J., & Kulick, E. M. (1983). *Assessing unexpected differential item performance of female candidates on SAT and TSWE forms administered in December 1977. An application of the standardization approach*. (ETS Research report RR-83.9). Princeton, NJ: Educational Testing Service.
- Dorans, M. J., & Kulick, E. M. (1986). Demonstrating the utility of the standardization approach to assessing unexpected differential item performance on the Scholastic Aptitude Test. *Journal of Educational Measurement*, 23(4), 355-368.
- Dunfield, K. A. (2014). A construct divided: prosocial behavior as helping, sharing, and comforting subtypes. *Front. Psychol.* 5:958. doi: 10.3389/fpsyg.2014.00958
- Dunfield, K. A., Kuhlmeier, V. A., O'Connell, L., & Kelley, E. (2011). Examining the diversity of prosocial behavior: Helping, sharing, and comforting in infancy. *Infancy*, 16(3), 227-247.
- Dzhidaryan, I. A. (2013). *Psychology of happiness and optimism*. Moscow: Publishing House of the "Institute of Psychology, Russian Academy of Sciences".
- Eagly, A. H. (2009). The his and hers of prosocial behavior: an examination of the social psychology of gender. *American Psychologist*, 64(8), 644-658. doi:10.1037/0003-066X.64.8.644
- Eastwood, C. D., & Ecklund, K. (2008). Compassion fatigue risk and self-care practices among residential treatment center childcare workers. *Residential Treatment for Children & Youth*, 25(2), 103-122.
- Eisenberg, N., Carlo, G., Murphy, B., & Van Court, P. (1995). Prosocial development in late adolescence: A longitudinal study. *Child Dev.* 66, 1179-1197.

- Eisenberg, N., Cumberland, A., Guthrie, I. K., Murphy, B. C., & Shepard, S. A. (2005). Age Changes in Prosocial Responding and Moral Reasoning in Adolescence and Early Adulthood. *Journal of Research in Adolescence*, *15*(3), 235-260.
- Eisenberg, N., Eggum, N. D. & Di Giunta, L. (2010). Empathy-related responding: Associations with prosocial behavior, aggression, and intergroup relations. *Social Issues and Policy Review*, *4*, 143-180. doi:10.1111/j.1751-2409.2010.01020.x
- Eisenberg, N., & Fabes, R. (1998). Prosocial development. En W. Damon y N. Eisenberg (Ed.), *Handbook of child psychology: vol. 3. Social, emotional and personality development* (pp. 701-778). New York: Wiley.
- Eisenberg, N., Fabes, R., Guthrie, I., & Reiser, M. (2000). Dispositional emotionality and regulation: Their role in predicting quality of social functioning. *Journal Personality and Social Psychology*, *78*(1), 136-157.
- Eisenberg, N., Hofer, C., Sulik, M. J., & Liew, J. (2014). The development of prosocial moral reasoning and a prosocial orientation in young adulthood: Concurrent and longitudinal correlates. *Developmental psychology*, *50*(1), 1-13. doi: 10.1037/a0032990
- Elliott, R., Völlm, B., Drury, A., McKie, S., Richardson, P., & Deakin, J. F. W. (2006). Cooperation with another player in a financially rewarded guessing game activates regions implicated in theory of mind. *Social Neuroscience, Special Issue: Theory of Mind*, *1*(3-4), 385-395.
- Elosua Oliden, P. (2003). Sobre la validez de los tests. *Psicothema*, *15*(2), 315-321.
- Embretson, S. E., & Reise, S. (2000). *Item Response Theory for Psychologists*. Mahwah, NJ: Erlbaum Publishers.
- Emmons, R. A., & McCullough, M. E. (2003). Counting blessings versus burdens: an experimental investigation of gratitude and subjective well-being in daily life. *Journal of personality and social psychology*, *84*(2), 377-389. doi: 10.1037/0022-3514.84.2.377
- Emmons, R. A., & Mishra, A. (2011). Why gratitude enhances well-being: What we know, what we need to know. In M. Kennon, T. Sheldon, T. Kashdan, & M. F. Steger (Eds), *Designing positive psychology: Taking stock and moving forward* (pp. 248-264). New York: Oxford Press.
- Eron, L. D., & Huesmann, L. R. (1984). The relation of prosocial behavior to the development of aggression and psychopathology. *Aggressive Behavior*, *10*, 201-211.
- Escurre Mayaute, L. M. (2011). Análisis psicométrico del Cuestionario de Honey y Alonso de Estilos de Aprendizaje (CHAEA) con los modelos de la Teoría Clásica de los Tests y de Rasch. *Persona*, *14*, 71-109.
- Escurre Mayaute, L. M., & Delgado Vásquez, A. E. (2010). Análisis psicométrico del test de matrices progresivas avanzadas de Raven mediante el modelo de tres parámetros de la teoría de la respuesta al ítem. *Persona*, *13*, 71-97.
- Escurre Mayaute, M., & Salas Blas, E. (2014). Construcción y validación del cuestionario de adicción a redes sociales (ARS). *Liberabit*, *20*(1), 73-91.
- Federación de Psicólogos de la República Argentina. (2013). *Código de Ética de la Federación de Psicólogos de la República Argentina (Fe.P.R.A.)*.
- Fernández Ballesteros, R. (1995). *Introducción a la Evaluación Psicológica I*. Madrid: Ediciones Pirámide.
- Fernández Pinto, I. López Pérez, B., & Márquez, M. (2008). Empatía: Medidas, teorías y aplicaciones en revisión. *Anales de Psicología*, *24*(2), 284-298.
- Ferrando, P. J., & Lorenzo-Seva, U. (2014). El análisis factorial exploratorio de los ítems: algunas consideraciones adicionales. *Anales de psicología*, *30*(3), 1170-1175.
- Ferreres Traver, D. F., González Romá, V., & Gómez Benito, J. (2002). Funcionamiento diferencial de los ítems en una situación de contacto de lenguas. *Psicothema*, *14*(2), 483-490.
- Figley, C. (1995). *Compassion Fatigue: Coping with Secondary Traumatic Stress Disorders in Those Who Treat the Traumatized*. London: Brunner-Routledge.
- Fisher, R. J. (1993). Social desirability bias and the validity of indirect questioning. *Journal of consumer research*, *20*(2), 303-315.
- Fiske, S. T., Cuddy, A. J. C., Glick, P., & Xu, J. (2002). A model of (often mixed) stereotype content: Competence and warmth respectively follow from perceived status and competition. *Journal of Personality and Social Psychology*, *82*, 878-902.
- Fite, P. J., Hendrickson, M., Rubens, S. L., Gabrielli, J., & Evans, S. (2013). The role of peer rejection in the link between reactive aggression and academic performance. In *Child & Youth Care Forum* (Vol. 42, No. 3, pp. 193-205). Springer US. doi: 10.1007/s10566-013-9199-9
- Fortaleza, D., Luzardo, M., & Padula, D. (2010). Estimación no paramétrica monótona de curvas características de ítems. *Memorias del II Encuentro de Docentes e Investigadores de Estadística en Psicología*. Buenos Aires, Argentina.

- Fortaleza, D., Padula, D., & Luzardo, M. (2010). Estimación no paramétrica de curvas características de ítems mediante egresión no paramétrica. *Memorias del II Encuentro de Docentes e Investigadores de Estadística en Psicología*. Buenos Aires, Argentina.
- Fraley, R. C., Waller, N. G., & Brennan, K. A. (2000). An item response theory analysis of self-report measures of adult attachment. *Journal of personality and social psychology*, 78(2), 350-365. doi: 10.1037/0022-3514.78.2.350
- Gallagher, H. L., & Frith, C. D. (2003). Functional imaging of “theory of mind”. *Trends in Cognitive Sciences*, 7(2), 77-83.
- Galen, L. W. (2012). Does religious belief promote prosociality? A critical examination. *Psychological Bulletin*, 138(5), 876-906.
- Galibert, M. S., Aguerri, M. E., Lozzia, G., & Abal, F., (2006). Análisis del DIF en una escala de voluntad de trabajo mediante los procedimientos de Mantel-Haenszel y Breslow-Day, regresión logística y el criterio ETS. *XIII Jornadas de Investigación y Segundo Encuentro de Investigadores en Psicología del Mercosur*. Facultad de Psicología -Universidad de Buenos Aires, Buenos Aires.
- Galibert, M. S., Aguerri, M. E., Lozzia, G., Abal, F., & Attorresi, H. F. (2005). Voluntad de Trabajo. Construcción de una Escala para su Medición. *Memorias del II Congreso Marplatense de Psicología de Alcance Nacional e Internacional*. Fac. Psicología, Univ. Nac. Mar del Plata. Soporte CD.
- Gardner, W. L., & Gabriel, S. (2004). Gender differences in relational and collective interdependence: Implications for self-views, social behavior, and subjective well-being. In A. H. Eagly, A. Beall, & R. J. Sternberg (Eds.), *The psychology of gender* (2nd ed., pp. 169–191). New York: Guilford Press.
- Garrido, L. E., Abad, F. J., & Ponsoda, V. (2013). A new look at Horn’s parallel analysis with ordinal variables. *Psychological Methods*, 18(4), 454-474. doi:10.1037/a0030005
- Geiser, C. (2013). *Data Analysis with Mplus*. New York, NY: The Guilford Press.
- Gempp, R., Denegri, M., Caripán, N., Catalán, V., Hermosilla, S., & Caprile, C. (2007). Desarrollo del test de alfabetización económica para adultos. *Interamerican Journal of Psychology*, 41(3), 275-284.
- Gil, J., Rodríguez, G., & García, E. (2000). *Análisis Factorial*. Madrid: La Muralla.
- Glutting, J. J., Monaghan, M. C., Adams, W., & Sheslow, D. (2002). Some psychometric properties of a system to measure ADHD among college students: Factor pattern, reliability, and one-year predictive validity. *Measurement and Evaluation in Counseling and Development*, 34, 194 - 208.
- Gómez Benito, J. & Hidalgo, M. D. (2015). *La validez en los Tests, Escalas y Cuestionarios*. Disponible en: <http://aprendeenlinea.udea.edu.co/revistas/index.php/ceo/article/download/1750/1370>.
- Gómez Jiménez, A., & Gaviria Stewart, E. (2010). Conducta de ayuda. En: Gaviria Stewart E., Cuadrado Guirado I., López Sáez M. (Coords). *Introducción a la Psicología Social* (pp. 285-317). Madrid: Sanz y Torres.
- Gómez Rada, C. A. (2004). Diseño, construcción y validación de un instrumento que evalúa clima organizacional en empresas colombianas, desde la teoría de respuesta al ítem. *Acta Colombiana de Psicología*, 11, 97-113.
- González Gutiérrez, J., Peñacoba Puente, C., Moreno Rodríguez, R., López López, A., & Velasco Furlong, L. (2006). Nursing Motives for Helping Scale (N-MHS): Reliability and Validity. *The Spanish Journal of Psychology*, 9(1), 103-112. doi: 10.1017/S1138741600006028
- González Portal, M. D. (2000). *Conducta prosocial: Evaluación e Intervención*. Madrid, España: Morata.
- Graham, C., & Chattopadhyay, S. (2013). Gender and well-being around the world. *International Journal of Happiness and Development*, 1(2), 212-232. doi: 10.1504/IJHD.2013.055648
- Graziano, W. G., Bruce, J., Sheese, B. E., & Tobin, R. M. (2007). Attraction, personality, and prejudice: Liking none of the people most of the time. *Journal of Personality and Social Psychology*, 93(4), 565-582. doi: 10.1037/0022-3514.93.4.565
- Graziano, W. G., Habashi, M. M., Sheese, B. E., & Tobin, R. M. (2007). Agreeableness, empathy, and helping: A person × situation perspective. *Journal of Personality and Social Psychology*, 93, 583-599.
- Grube, J. A., & Piliavin, J. A. (2000). Role identity, organizational experiences and volunteer performance. *Personality and Social Psychology Bulletin*, 26, 1108-1119.
- Guijo Blanco, V. (2003). *Estudio multifactorial de la conducta prosocial en niños de cinco y seis años*. Burgos: Universidad de Burgos.
- Gulliford, L., Morgan, B., & Kristjánsson, K. (2013). Recent work on the concept of gratitude in philosophy and psychology. *Journal of value inquiry*, 47(3), 285-317. doi: 10.1007/s10790-013-9387-8
- Gulliksen, H. (1950). *Theory of mental tests*. New York: Wiley.
- Gurieva, L. P. (2012). Positive and traditional psychology (Studies of Seligman M. and Russian psychologists). Bulletin of the Moscow State Humanitarian University of Mikhail Sholokhov. *Pedagogy and Psychology*, 3, 83-90.

- Gutiérrez San Martín, M., Escartí, A., & Pascual, C. (2011). Relaciones entre empatía, conducta prosocial, agresividad, autoeficacia y responsabilidad personal y social de los escolares. *Psicothema*, 23(1), 13-19.
- Hair, J. F., Anderson, R. E., Tatham, R. L., & Black, W. C. (1999). *Análisis Multivariante*. Madrid: Prentice Hall.
- Hambleton, R. K., Clauser, B. E., Mazor, K. M., & Jones, R. W. (1993). Advances in the detection functioning of differentially functioning test items. *European Journal of Psychological Assessment*, 9(1), 1-18.
- Hambleton, R. K., & Slater, S. C. (1997). Item Response Theory Models and Testing Practices: Current international status and future directions. *European Journal of Psychological Assessment*, 13(1), 21-28.
- Hambleton R. K., & Swaminathan H. (1985). *Item Response Theory: Principles and applications*. Boston: Kluwer.
- Hambleton, R. K., van der Linden, W. J., & Wells, C. S. (2010). IRT models for the analysis of polytomously scored data: Brief and selected history of model building advances. In M. L. Nering & R. Ostini (Eds.), *Handbook of polytomous item response theory models* (pp. 21-42). New York: Routledge.
- Hanson, B. A. (1998). Uniform DIF and DIF defined by differences in item response functions. *Journal of Educational and Behavioral Statistics*, 23(3), 244-253.
- Hawley, P. H. (1999). The ontogenesis of social dominance: A strategy-based evolutionary perspective. *Developmental Review*, 19(1), 97-132.
- Hay, D. F., & Cook, K. V. (2007). The transformation of prosocial behavior from infancy to childhood. En C. A. Brownell & C. B. Kopp (Eds.), *Socioemotional development in the toddler years: Transitions and transformations* (pp. 100-131). New York: The Guilford Press.
- Hemker, B. T., van der Ark, L. A., & Sijtsma, K. (2001). On measurement properties of continuation ratio models. *Psychometrika*, 66(4), 487-506.
- Hendrickson, A. E., & White, P. O. (1964). Promax: A quick method for rotation to oblique simple structure. *British Journal of Statistical Psychology*, 17(1), 65-70. doi:10.1111/j.2044-8317.1964.tb00244.x
- Hernández Baeza, A., Muñoz Fernández, J., & García Cueto, E. (2000). Comportamiento del modelo de respuesta graduada en función del número de categorías de la escala. *Psicothema*, 12(2), 288-291.
- Hewstone, M., Rubin, M., & Willis, H. (2002). Intergroup bias. *Annual review of psychology*, 53(1), 575-604.
- Hidalgo, M. D., López-Pina, J. A., Inglés, C. J., & Méndez, F. X. (2002). Análisis psicométrico del Cuestionario de confianza para hablar en público, usando la teoría de respuesta al ítem. *Anales de Psicología*, 18(2), 333-349.
- Hill, C.T.; Ivan, L.; González, R.C.; Brumbaugh, C.; Barros, M.; Rodriguez, J.C.Y.; Kito, M.; Macbeth, G.; Razumiejczyk, E.; Mari, S. (2013). Happiness and dimensions of relationship intimacy across gender, sex of partner, and cultures. In *Acti di European Congress of Psychology*, Stockholm, Sweden, July 10-12.
- Hillygus, D. S. (2005). The Missing Link: Exploring the Relationship Between Higher Education and Political Engagement. *Political Behavior*, 27(1), 25-48.
- Hinde, R. A., & Groebel, J. (1991). *Cooperación y conducta prosocial*. Madrid: Visor.
- Hirschberger, G., Ein-Dor, T., & Almakias, S. (2008). The self-protective altruist: Terror management and the ambivalent nature of prosocial behavior. *Personality and Social Psychology Bulletin*, 34(5), 666-678.
- Hoffman M. L. (2000). *Empathy and moral development: Implications for caring and justice*. New York: Cambridge University Press.
- Hogan, T. (2004). *Pruebas psicológicas. Una introducción práctica*. México: Manual Moderno.
- Holgado Tello, F. P., Chacón Moscoso, S., Barbero García, I., & Vila-Abad, E. (2010). Polychoric versus Pearson correlations in exploratory and confirmatory factor analysis of ordinal variables. *Quality & Quantity*, 44(1), 153-166. doi:10.1007/s11135-008-9190-y
- Holland, P. W. (1985). *On the study of differential item performance without IRT*. Comunicación presentada en Proceedings of the Militar Testing Association.
- Holland, P. W., & Thayer, D.T. (1988). Differential item functioning and the Mantel-Haenszel procedure. En H. Wainer & H. I. Braun (Eds.), *Test Validity* (pp. 129-145). Hillsdale, NJ: Lawrence Erlbaum.
- Holland, P. W., & Wainer, H. (1993). *Differential item functioning*. Hillsdale, NJ: Lawrence Erlbaum.
- Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika*, 30(2), 179-185. doi:10.1007/BF02289447
- Howard, J. A., & Piliavin, J. A. (2000). *Altruism. Encyclopedia of Sociology*. New York: Macmillian.
- Hu, L., & Bentler, P. M. (1999). Cut off criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6(1), 1-55. doi:10.1080/10705519909540118
- Inglés, C., Benavides, G., Redondo, J., García-Fernández, J. M., Ruiz-Esteban, C. Estévez, C., & Huescar, E. (2009). Conducta prosocial y rendimiento académico en estudiantes españoles de Educación Secundaria Obligatoria. *Anales de Psicología*, 25(1), 93-10.
- Inderbitzen, H. M., & Foster, S. L. (1992). The Teenage Inventory of Social Skills: Development, reliability, and validity. *Psychological Assessment*, 4(4), 451-459.

- Inglés, C., Hidalgo, M. D., Méndez, F. X., & Inderbitzen, H. M. (2003). The Teenage Inventory of Social Skills: reliability and validity of the Spanish translation. *Journal of Adolescence*, 26(4), 505-510.
- Inglés, C. J., Martínez-González, A., & García-Fernández, J. M. (2013). Conducta prosocial y estrategias de aprendizaje en una muestra de estudiantes españoles de Educación Secundaria Obligatoria. *European Journal of Education and Psychology*, 6(1), 33-53.
- Inglés, C. J., Martínez-González, A., García-Fernández, J. M., Torregrosa, M. S., & Ruiz-Esteban, C. (2012). Prosocial behavior and self-concept of Spanish students of Compulsory Secondary Education. *Revista de Psicodidáctica*, 17(1), 135-156.
- Inglés, C. J., Martínez-Monteagudo, M. C., Delgado, B., Torregrosa, M. S., Redondo, J., Benavides, G., García-Fernández, J. M., & García-López, L. J. (2008). Prevalencia de la conducta agresiva, conducta prosocial y ansiedad social en una muestra de adolescentes españoles: Un estudio comparativo. *Infancia y Aprendizaje*, 31(4), 449-461.
- Israel-Cohen, Y., Uzefovsky, F., Kashy-Rosenbaum, G., & Kaplan, O. (2015). Gratitude and PTSD symptoms among Israeli youth exposed to missile attacks: examining the mediation of positive and negative affect and life satisfaction. *The Journal of Positive Psychology*, 10(2), 99-106. doi: 10.1080/17439760.2014.92791
- Joseph, S., Linley, P. A., Harwood, J., Lewis, C. A., & McCollam, P. (2004). Rapid assessment of well-being: The Short Depression-Happiness Scale (SDHS). *Psychology and psychotherapy: Theory, research and practice*, 77(4), 463-478. doi: 10.1348/1476083042555406
- Joseph, S., & Wood, A. M. (2010). Assessment of positive functioning in clinical psychology: Theoretical and practical issues. *Clinical Psychology Review*, 30(7), 830-838. doi: 10.1016/j.cpr.2010.01.002
- Kahana, E., Bhatta, T., Lovegreen, L. D., Kahana, B., & Midlarsky, E. (2013). Altruism, helping, and volunteering: pathways to well-being in late life. *Journal of Aging and Health*, 25(1), 159-187.
- Kang, T., & Chen, T.T. (2008). Performance of the generalized S - X<sup>2</sup> model-data fit index for polytomous IRT models. *Journal of Educational Measurement*, 45, 391-406. doi: 10.1111/j.1745-3984.2008.00071.x
- Kärtner, J., Keller, H., & Chaudhary, N. (2010). Cognitive and social influences on early prosocial behavior in two sociocultural contexts. *Dev. Psychol.* 46:905. doi: 10.1037/a0019718
- Keller, H. (2007). *Cultures of Infancy*. Mahwah, NJ: Psychology Press.
- Kelley, T. L. (1935). *Essential traits of mental life*. Cambridge, MA: Harvard University Press.
- Krzemien, D. (2001). Calidad de vida y participación social en la vejez femenina. En: *Temas de Psicología del Desarrollo*. Mar de Plata: Universidad Nacional de Mar del Plata.
- Krzemien, D., Monchetti, A., & Urquijo, S. (2005). Afrontamiento activo y adaptación al envejecimiento en mujeres de la ciudad de Mar del Plata: una revisión de la estrategia de autodistracción. *Interdisciplinaria*, 22(2), 183-210.
- Lacunza, A. B., Caballero, S. V., & Contini, E. N. (2013). Adaptación y evaluación de las propiedades psicométricas de la BAS-3 para población adolescente de Tucumán (Argentina). *Diversitas*, 9(1), 29-44.
- Ladd, G. W., & Profilet, S. M. (1996). The Child Behavior Scale: A teacher-report measure of young children's aggressive, withdrawn, and prosocial behaviors. *Developmental Psychology*, 32(6), 1008-1024.
- Lambert, N. M., Fincham, F. D., & Stillman, T. F. (2012). Gratitude and depressive symptoms: The role of positive reframing and positive emotion. *Cogn. Emot.* 26(4), 615-633. doi: 10.1080/02699931.2011.595393
- Lazarsfeld, P. F. (1950). The logical and mathematical foundation of latent structure analysis. En S. A. Stouffer et al. (Eds.). *Measurement and Prediction*. Princeton: Princeton University Press.
- Ledesma, R. D., Poó, F. M., & Montes, S. A. (2011). Psicología del tránsito: logros y desafíos de la investigación. *Psiciencia: Revista Latinoamericana de Ciencia Psicológica*, 3(2), 108-119.
- Ledesma, R. D., Valero-Mora, P., & Macbeth, G. (2015). The scree test and the number of factors: a dynamic graphics approach. *The Spanish Journal of Psychology*, 18, 1-10. doi: 10.1017/sjp.2015.13
- Lemos, V. N., & Richaud de Minzi, M. C. (2010). Construcción de un instrumento para evaluar el razonamiento prosocial en niños de 7 y 8 años: una versión pictórica. *Universitas Psychologica*, 9(3), 879-891.
- Lipps, T. (1903). Einfühlung, innere Nachahmung und Organ empfindung. *Archiv für die Gesamte Psychologie* 1, 465-519.
- Lloret-Segura, S., Ferreres-Traver, A., Hernández-Baeza, A., & Tomás-Marco, I. (2014). El análisis factorial exploratorio de los ítems: una guía práctica, revisada y actualizada. *Anales de psicología*, 30(3), 1151-1169. doi:10.6018/analesps.30.3.199361
- Loevinger, J. (1957). Objective tests as instruments of psychological theory. *Psychological Reports*, 3, 635-694.
- López, F. (1994). *Para Comprender la Conducta Altruista*. Navarra: Verbo Divino.

- López de Cózar, E., Alonso Plá, F., Esteban Martínez, C., Calatayud Miñana, C., & Alamar Rocati, B. (2008). *Diseño y validación de una escala para la medición de la conducta prosocial - antisocial en el ámbito vital y en el tráfico*. Instituto Universitario de Tráfico y Seguridad Vial. Universidad de Valencia.
- Lord, F. M. (1953). The relation of test score to the trait underlying the test. *Educational and Psychological Measurement, 13*, 517-548.
- Lord, F. M. (1977). An study of item bias using item characteristic curve theory. En Y. H. Poortinga (Ed.), *Basic problems in cross-cultural psychology* (pp. 19-29). Amsterdam, The Netherlands: Swets & Zeitlinger.
- Lord, F. M. (1980). *Applications of item response theory to practical testing problems*. Hillsdale, N. J.: Lawrence Erlbaum.
- Lord, F. M., & Novick, M. R. (1968). *Statistical theories of mental tests scores*. Reading, MA: Addison-Wesley.
- Lorenzo Seva, U., & Ferrando, P. J. (2013). FACTOR 9.2: A comprehensive program for fitting exploratory and semi confirmatory factor analysis and IRT models. *Applied Psychological Measurement, 37*(6), 497-498. doi:10.1177/0146621613487794
- Loudin, J. L., Loukas, A., & Robinson, S. (2003). Relational aggression in college students: Examining the roles of social anxiety and empathy. *Aggressive Behavior, 29*(5), 430-439.
- Lozzia, G. (2012). *Construcción de un Banco de Ítems de Analogías Verbales y su Aplicación a la Elaboración de un Test Adaptativo Informatizado*. Tesis de Doctorado, Facultad de Psicología de la Universidad de Buenos Aires.
- Lyubomirsky, S. (2008). *La ciencia de la felicidad. Un método para conseguir el bienestar*. España: Urano.
- Macbeth, G., Cortada de Kohan, N. C., & Razumiejczyk, E. (2007). Perspectivas en epistemología de la psicología. *Psicología y Psicopedagogía, 16*(1), 1-15.
- Macbeth, G., Cortada de Kohan, N. C., Razumiejczyk, E., & López Alonso, A. O. (2006). La Investigación Científica en Psicología: un Desarrollo Histórico. *Psicología y Psicopedagogía, 5*(15), 1-17.
- Mallinckrodt, B., Miles, J. R., & Recabarren, D. A. (2016). Using focus groups and Rasch item response theory to improve instrument development. *The Counseling Psychologist, 44*(2), 146-194. doi:10.1177/0011000015596437
- Mantel, N., & Haenszel, W. (1959). Statistical aspects of the analysis of data from retrospective studies of disease. *Journal of the National Cancer Institute, 22*(4), 719-748.
- Mardia, K. V. (1970). Measures of multivariate skewness and kurtosis with applications. *Biometrika, 57*(3), 519-530. doi:10.1093/biomet/57.3.519
- Marean, C. W. (2015). La especie más invasora. *Investigación y ciencia, 469*, 14-21.
- Martela, F. & Ryan, R. M. (2016). Prosocial behavior increases well-being and vitality even without contact with the beneficiary: Causal and behavioral evidence. *Motivation and Emotion, 40*(3), 351-357. doi: 10.1007/s11031-016-9552-z
- Martínez, A., Inglés, C., Piqueras, J. & Oblitas, L. (2009). Papel de la conducta prosocial y de las relaciones sociales en el bienestar psíquico y físico del adolescente. *Avances en Psicología Latinoamericana, 28*(1), 74-84.
- Martínez Arias, M. R. (1995). *Psicometría: Teoría de los Tests Psicológicos y Educativos*. Madrid: Síntesis.
- Martínez Arias, M. R., Hernández Lloreda, M. V., & Hernández Lloreda, M. J. (2006). *Psicometría*. Madrid: Alianza Editorial.
- Martorell, M. C., González, R., Aloy, M., & Ferris, M. C. (1995). Socialización y conducta prosocial. *Revista Iberoamericana de Diagnóstico y Evaluación Psicológica, 1*(1), 73-102.
- Maslow, A. (1962). *Toward a Psychology of Being*. New York: Van Nostrand.
- Masters, G. N. (1982). A Rasch model for partial credit scoring. *Psychometrika, 47*(2), 149-174.
- Masters, G. N. (2016). Partial Credit Model. En W. J. van der Linden (Ed.). *Handbook of Item Response Theory, Volume 1: Models* (pp. 109-126). Boca Raton: Chapman & Hall/CRC.
- Maya, V. & Fraile, E.A. (1996). El test sociométrico: teoría y aplicación. Cuadernos de realidades sociales, 47-48, 111-129.
- Mazor, K. M., Clauser, B. E., & Hambleton, R. K. (1994). Identification of nonuniform differential item functioning using a variation of the Mantel-Haenszel procedure. *Educational and Psychological Measurement, 54*(2), 284-291.
- McCrae, R. R. & Costa, P. T. (1999). A five-factor theory of personality. En L. A. Pervin & O. P. John (Eds.), *Handbook of personality* (2nd ed., pp. 139-153). New York: Guilford Press.
- McCullough, M. E., Emmons, R. A., & Tsang, J. A. (2002). The grateful disposition: a conceptual and empirical topography. *Journal of personality and social psychology, 82*(1), 112-127. doi: 10.1037/0022-3514.82.1.112
- McDonald, R. P. (1999). *Test theory: A unified treatment*. Mahwah, NJ: Erlbaum.

- McFarland, S., Brown, D., & Webb, M. (2013). Identification with all humanity as a moral concept and psychological construct. *Current Directions in Psychological Science*, 22(3), 194-198.
- Mestre Escrivá, M. V., Frías Navarro, M. D., & Samper García, P. (2004). La medida de la empatía: análisis del Interpersonal Reactivity Index. *Psicothema*, 16(2), 255-260.
- Mestre Escrivá, M. V., Frías Navarro, M. D., Samper García, P., & Nácher, M. J. (2003). Estilos de crianza y variables personales como factores de riesgo de la conducta agresiva. *Revista Mexicana de Psicología*, 20(2), 189-199.
- Mestre Escrivá, M. V., Samper García, P., & Frías Navarro, M. D. (2002). Procesos cognitivos y emocionales predictores de la conducta prosocial y agresiva: La empatía como factor modulador. *Psicothema*, 14(2), 227-232.
- Mestre Escrivá, M. V., Samper García, P., Tur, A. M., Cortés, M. T., & Nácher, M. J. (2006). Conducta prosocial y procesos psicológicos implicados: un estudio longitudinal en la adolescencia. *Revista Mexicana de Psicología*, 23(2), 203-215.
- Mikulic, I. M. (2007). *La Evaluación Psicológica y el Análisis Ecoevaluativo*. Ficha de la Cátedra I Teoría y Técnica de la Evaluación y Diagnóstico Psicológico Módulo I. Departamento de Publicaciones. Facultad de Psicología.
- Mikulic, I. M., Caballero, R., & Aruanno, Y. (2014). Competencias socio-emocionales en adultos mayores de la Ciudad de Buenos Aires. *Anuario de investigaciones*, 21(1), 277-284.
- Mikulic, I. M., & Muiños, R. (2005). La construcción y uso de instrumentos de evaluación en la investigación e intervención psicológica: el Inventario de Calidad de Vida Percibida (ICV). *Anuario de Investigaciones de la Facultad de Psicología UBA*, 12, 193-202.
- Mikulic, I. M., Muiños, R., & Crespi, M. (2007). Crecimiento personal, necesidad de trascendencia, redes sociales y ambiente: análisis de la calidad de vida percibida en diferentes sucesos estresantes desde un enfoque integrativo y ecológico. *Anuario de investigaciones*, 14, 277-285.
- Mikulic, I. M., Muiños, R., Crespi, M., Fernández, G., Monges, M. E., Lorenzut, V., & Ojeda, G. (2006). "Transiciones ecológicas": Estudio del afrontamiento de las crisis vitales migración, encarcelamiento y divorcio desde el enfoque de la Psicología Positiva. *Anuario de Investigaciones*, 13, 175-184.
- Mola, D. J., Saavedra, B. A., Reyna, C. E., & Balaus, A. (2013). Valoración psicométrica de la Psychological Entitlement Scale desde la teoría clásica de los tests y la teoría de respuesta al ítem. *Pensamiento Psicológico*, 11, 19-38.
- Molinari, A. V., Takaya, P. B., Domínguez-Romero, N. C., & Sánchez Rosas, J. (2016). Desarrollo y estructura factorial de una Escala de Participación Social Académica en Clase: Datos Preliminares. II Congreso Internacional de Psicología - V de Congreso Nacional de Psicología: "CIENCIA Y PROFESIÓN". *Desafíos para la construcción de una psicología regional*, Facultad de Psicología, UNC., At Córdoba. Argentina.
- Montero, E. (2000). La teoría de respuesta a los ítemes: una moderna alternativa para el análisis psicométrico de instrumentos de medición. *Revista de Matemática: Teoría y Aplicaciones*, 7(1,2), 217-228.
- Montero, I., & León, O. G. (2007). Sistema de clasificación del método en los informes de investigación en Psicología. *International Journal of Clinical and Health Psychology*, 5(1), 115-127.
- Morales, F., & Arias, C. J. (2014). El perdón como motivador prosocial en la vejez. *Neurama revista electrónica de Psicogerontología*, 1(2), 31-43.
- Morales Rodríguez, M. & Suárez Pérez, C. D. (2011). Construcción y validación de una escala para evaluar habilidades prosociales en adolescentes. Ponencia. *XI Congreso Nacional de Investigación Educativa*. Consejo Mexicano de Investigación Educativa, Delegación Coyoacán, México, Distrito Federal.
- Moreno, J. E., & Fernández, C. (2011). Empatía y flexibilidad yoica, su relación con la agresividad y la prosocialidad. *Límite: revista de filosofía y psicología*, 23, 41-56.
- Moya Albiol, L., Herrero, N., & Bernal, M. C. (2010). Bases neuronales de la empatía. *Rev Neurol*, 50(2), 89-100.
- Muñiz Fernández, J. (1996). Fiabilidad. En J. Muñiz (Coor.), *Psicometría*. Madrid: Universitat
- Muñiz Fernández, J. (1997). *Introducción a la teoría de respuesta a los ítems*. Madrid: Ediciones Pirámide.
- Muñiz Fernández, J. (1998). La medición de lo psicológico. *Psicothema*, 10(1), 1-21.
- Muñiz Fernández, J. (2010). Las teorías de los tests: Teoría Clásica y Teoría de Respuesta a los Ítems. *Papeles del Psicólogo*, 31(1), 57-66.
- Muñiz Fernández J., Fidalgo, A. M., García Cueto, E., Martínez, R., & Moreno, R. (2005). *Análisis de los ítems*. Madrid: La Muralla.
- Muñoz García, J. A. & Anón Uribe, I. (2013). Técnicas para detección de outliers multivariantes. *Revista en telecomunicaciones e informática*, 3(5), 11-25.
- Muraki, E. (1997). A Generalized Partial Credit Model. En W. J. Van der Linden & R. K. Hambleton (Eds.), *Handbook of Modern Item Response Theory*, (pp. 153-164). New York: Springer.

- Musitu, G. & Cava, M. J. (2003). El rol del apoyo social en el ajuste de los adolescentes. *Intervención psicosocial*, 12(2), 179-192.
- Muthén, L. & Muthén, B. (2010). *Mplus User's Guide, 6th Edn*. Los Angeles, CA: Muthén & Muthén.
- National Council on Measurement in Education (2013). *Standards for educational and psychological testing*. Recuperado de: [https://www.ncme.org/ncme/NCME/NCME/Resource\\_Center/Standards.aspx](https://www.ncme.org/ncme/NCME/NCME/Resource_Center/Standards.aspx)
- Navarro Loli, J., & Merino Soto, C. (2010). Un réquiem para la regla de Kaiser (eigen > 1) en la retención del número de factores. *Fractal Revista de Psicología*, 22(3), 641-642.
- Newman, A., Dennis, C., Wright, L. T., & King, T. (2010). Shoppers' experiences of digital signage-a cross-national qualitative study. *JDCTA*, 4(7), 50-57.
- Olivar, R. R. (1998). El uso educativo de la televisión como optimizadora de la prosocialidad. *Psychosocial Intervention*, 7(3), 363-378.
- Olivera Ryberg, J. M. (2010). *Acerca del concepto de empatía. Su rol y evaluación en psicoterapia*. Tesis para obtener el título de Licenciado en Psicología, Universidad de Belgrano, Argentina.
- Omoto, A. M., & Snyder, M. (1990). Basic research in action: Volunteerism and society's response to AIDS. *Personality and Social Psychology Bulletin*, 16(1), 152-166.
- Osterlind, S. J., & Everson, H. T. (2009). *Differential item functioning*. Thousand Oaks, California: Sage Publications, Inc.
- Ostini, R., & Nering, M. L. (2005). *Polytomous item response theory models*. Newbury Park, CA: Sage.
- Padilla, J. L., & González, A. (2001). Fundamentos de medición psicológica para los modelos basados en la Teoría de Respuesta a los Ítems. En Rojas Tejada, A. & Pérez Meléndez, C. *Nuevos Modelos para la Medición de Actitudes*. (pp. 19-30). Valencia: Promolibro.
- Pallant, J. F., & Tennant, A. (2007). An introduction to the Rasch measurement model: an example using the Hospital Anxiety and Depression Scale (HADS). *British Journal of Clinical Psychology*, 46(1), 1-18.
- Paulus, M., Kühn-Popp, N., Licata, M., Sodian, B., & Meinhardt, J. (2013). Neural correlates of prosocial behavior in infancy: different neurophysiological mechanisms support the emergence of helping and comforting. *Neuroimage*, 66, 522-530. doi:10.1016/j.neuroimage.2012.10.041
- Pearce, P. L. & Amato, P. R. (1980). A taxonomy of helping: A multidimensional scalling analysis. *Social Psychology Quarterly*, 43(4), 363-371.
- Penfield, R. D. (2003). Applying the Breslow-Day test of trend in odds ratio heterogeneity to the analysis of nonuniform DIF. *Alberta Journal of Educational Research*, 49(3), 231-243.
- Penner, L. (2002). Dispositional and organizational influences on sustained volunteerism: An interactionist perspective. *Journal of Social Issues*, 58(3), 447-467.
- Penner, L., Dovidio, J., Piliavin, J., & Schroeder, D. (2005). Prosocial behavior: Multilevel perspectives. *Annual Review of Psychology*, 56(1), 365-392.
- Penner, L., Fritzsche, B., Craiger, J., & Freifeld, T. (1995). Measuring the prosocial personality. En J. Butcher, & C. Spielberger, (Eds.), *Advances in personality assessment* (Vol. 10). Hillsdale, NJ: Erlbaum.
- Pérez, E., & Medrano, L. A. (2010). Análisis factorial exploratorio: bases conceptuales y metodológicas. *Revista Argentina de Ciencias del Comportamiento*, 2(1), 58-66.
- Piliavin, J. A., Grube, J. A., & Callero, P. L. (2002). Role as resource for action in public service. *Journal of Social Issues*, 58, 469-485. doi: 10.1111/0022-4537.t01-1-00027
- Plazas, E. A., Morón, M. L., Santiago, A., Sarmiento, H., Ariza, S. E., & Patiño, C. (2010). Relaciones entre iguales, conducta prosocial y género desde la educación primaria hasta la universitaria en Colombia. *Universitas Psychologica*, 9(2), 357-369.
- Preston, S. D., & De Waal, F. (2002). Empathy: Its ultimate and proximate bases. *Behavioral and brain sciences*, 25(1), 1-20.
- Pursell, G. R., Laursen, B., Rubin, K. H., Booth-LaForce, C., & Rose-Krasnor, L. (2008). Gender differences in patterns of association between prosocial behavior, personality, and externalizing problems. *Journal of Research in Personality*, 42(2), 472-481.
- Raju, N. S. (1988). The area between two item characteristic curves. *Psychometrika*, 53, 284-291.
- Raju, N. S. (1990). Determining the significance of estimated signed and unsigned areas between two item response functions. *Applied Psychological Measurement*, 14, 197-207.
- Rand, D. G. & Epstein, Z. G. (2014). Risking Your Life without a Second Thought: Intuitive Decision-Making and Extreme Altruism. *PLoS ONE* 9(10): e109687. doi:10.1371/journal.pone.0109687
- Rasch, G. (1960). *Probabilistic Models for Some Intelligence and Attainment Tests*. Copenhagen: The Danish Institute for Educational Research.

- Rasch, G. (1961). On general laws and the meaning of measurement in psychology. In *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability* (Vol. 1, pp. 321-334). Berkeley, CA: University of Chicago Press.
- Reckase, M. (1979). Unifactor latent trait models applied to multifactor tests: Results and implications. *Journal of Educational Statistics*, 4(3), 207-230.
- Redondo Pacheco, J. & Guevara Melo, E. (2012). Diferencias de género en la prevalencia de la conducta prosocial y agresiva en adolescentes de dos colegios de la ciudad de Pasto-Colombia. *Revista Virtual Universidad Católica del Norte*, 1(36), 173-192.
- Redondo Pacheco, J., Inglés, C. J., & García-Fernández, J. M. (2014). Conducta prosocial y autoatribuciones académicas en Educación Secundaria Obligatoria. *Anales de Psicología*, 30(2), 482-489.
- Reise, S. P., & Henson, J. M. (2003). A discussion of modern versus traditional psychometrics as applied to personality assessment scales. *Journal of Personality Assessment*, 81, 93-103.
- Reise, S. P., & Waller, N. G. (1990). Fitting the two-parameter model to personality data. *Applied Psychological Measurement*, 14, 45-58.
- Revuelta Menéndez, J., Abad, F. J. & Ponsoda Gil, V. (2006). *Modelos Políticos de respuesta al ítem*. Madrid: La Muralla.
- Reykowski, J., & Smolenska, Z. (1980). Personality mechanisms of prosocial behavior. *Polish Psychological Bulletin*, 11, 219-230.
- Richaud de Minzi, M. C. (2005). Desarrollos del análisis factorial para el estudio de ítem dicotómicos y ordinales. *Interdisciplinaria*, 22, 237-251.
- Richaud de Minzi, M. C. (2008). Nuevas tendencias en psicometría. *Evaluar*, 8, 1-19.
- Richaud de Minzi, M. C. (2009). Influencia del modelado de los padres sobre el desarrollo del razonamiento prosocial en los/las niños/as. *Interamerican Journal of Psychology*, 43(1), 187-198.
- Richaud de Minzi, M. C. (2013). Nuevos métodos estadísticos para la investigación en evaluación de la personalidad. *Revista de Psicología*, 21(2), 291-310.
- Richaud de Minzi, M. C., Mesurado, B., & Lemos, V. (2012). Links between perception of parental actions and prosocial behavior in early adolescence. *Journal of Child and Family Studies*, 22(5), 637-646. doi: 10.1007/s10826-012-9617-x
- Rifkin, J. (2010). *La civilización empática. La carrera hacia una conciencia global en un mundo en crisis*. Ed. Paidós, Barcelona.
- Rochat, P., Dias, M. D., Liping, G., Broesch, T., Passos-Ferreira, C., Winning, A., et al. (2009). Fairness in distributive justice by 3- and 5-year-olds across seven cultures. *J. Cross Cult. Psychol.* 40, 416-442. doi: 10.1177/0022022109332844
- Rodríguez, M. C., Vivas, J. R., Comesaña, A., Ramírez, L. M., & Peña, J. A. (2017). Perceived Social Support in Middle School Students. *International Review of Social Sciences*, 5(1), 81-90.
- Rodríguez, L. M., & Moreno, J. E. (2016). Posturas éticas y empatía, predictores de prosocialidad y de penalización de faltas y delitos [*Ethical position and empathy, predictors of prosociality and penalization of offenses and crimes*]. *Acción Psicológica*, 13(2), 43-56. doi: 10.5944/ap.13.2.17809
- Rojas, M., Gatica, Y., Manríquez, G., & Salcedo Lagos, P. (2004). *Curso de UML Multiplataforma Adaptativo Basado en la Teoría de Respuesta al Ítem*. Concepción, Chile: Universidad de Concepción.
- Rojas, A. J., & Pérez, C. (2001). *Nuevos Modelos para la Medición de Actitudes*. Valencia: Promolibro.
- Romer, D., Gruder, C. L., & Lizzardo, T. (1986). A person-situation approach to altruistic behavior. *Journal of Personality and Social Psychology*, 51, 1001-1012.
- Rossel, G. (2006). Usando IRT y Agentes para Educación Distancia Adaptativa. *I Congreso de Tecnología en Educación y Educación en Tecnología*. La Plata. Buenos Aires. Argentina.
- Rouse, S. V., Finger, M. S., & Butcher, J. N. (1999). Advances in clinical personality measurement: An item response theory analysis of the MMPI-2 PSY-5 scales. *Journal of Personality Assessment*, 72(2), 282-307.
- Rushton, J. P., Chrisjohn, R. D., & Fekken, G. C. (1981). The altruistic personality and the self-report altruism scale. *Pers. Individ. Differ.* 2, 1-11.
- Ryff, C. (2009). The Structure of psychological well-being revisited. *Journal of Personality and Social Psychology*, 69, 719-727. doi: 10.1037/0022-3514.69.4.719
- Saklofske, D. H., & Zeidner, M. (1995). *International handbook of personality and intelligence*. Springer Science & Business Media.
- Salovey, P., Mayer, J. D., & Rosenhan, D. L. (1991). Mood and helping: Mood as a motivator of helping and helping as a regulator of mood. En M. S. Clark (Ed.), *Prosocial behaviour: Review of personality and social psychology*. Vol 12 (pp. 215-237). Newbury Park: Sage.

- Sam, K. L., Li, C., & Lo, S. K. (2016). Validation of the mental retardation attitude inventory-revised (MRAI-R): a multidimensional rasch analysis. *International Journal of Social Science and Humanity*, 6(7), 519-524. doi: 10.7763/IJSSH.2016.V6.703
- Samejima, F. (1969). Estimation of latent ability using a response pattern of graded scores. *Psychometrika monograph supplement*, 17(4), 2.
- Sánchez Queija, I., Oliva, A., & Parra, Á. (2006). Empatía y conducta prosocial durante la adolescencia. *Revista de Psicología Social*, 21(3), 259-271.
- Santisteban, C. (1990). *Psicometría. Teoría y práctica en la construcción de test*. Madrid: Norma.
- Scalco, D. L., Araújo, C. L., & Bastos, J. L. (2011). Autopercepção de felicidade e fatores associados em adultos de uma cidade do sul do Brasil: estudo de base populacional. *Psicologia: Reflexão e Crítica*, 24, 648-657.
- Schroeder, D. A., Penner, L. A., Dovidio, J. F., & Piliavin, J. A. (1995). *The psychology of helping and altruism: Problems and puzzle*. New York: McGraw-Hill.
- Schulz, E. M. (1990). DIF detection: Rasch vs. Mantel-Haenszel. *Rasch Measurement Transactions*, 4(2), 107.
- Scourfield, J., John, B., Martin, N., & McGuffin, P. (2004). The development of prosocial behavior in children and adolescents: A twin study. *Journal of Child Psychology and Psychiatry*, 45(5), 927-935.
- Seligman, M. (2009). *In a search of happiness. How to enjoy life every day*. Moscow: Mann, Ivanov and Ferber.
- Seligman, M. E., Parks, A. C., & Steen, T. (2004). A balanced psychology and a full life. *Philosophical Transactions-Royal Society of London Series B Biological Sciences*, 1379-1382. doi: 10.1098/rstb.2004.1513
- Shaffer, D. (2002). *Desarrollo Social y de la Personalidad*. Madrid: Thomson.
- Silva, F., & Martorell, M. C. (1987). *BAS-3. Bateria de Socialización*. Madrid: TEA.
- Smith, R. M. (2000). Fit analysis in latent trait measurement models. *Journal of Applied Measurement*, 2, 199-218.
- Smith, T. W. (2003). *Altruism in contemporary America: A report from the National Altruism Study*. Chicago, IL: National Opinion Research Center.
- Smith, C. E., Kleinbeck, S. V., Boyle, D., Kochinda, C., & Parker S. (2001). Family Caregivers' Motives for Helping Scale derived from motivation-to-help theory. *Journal of Nursing Measurement*, 9(3), 239-257.
- Solomon, S., Greenberg, J., & Pyszczynski, T. (1991). A terror management theory of social behavior: The psychological functions of self-esteem and cultural worldviews. *Advances in experimental social psychology*, 24, 93-159.
- Spearman, C. (1910). Correlation calculated from faulty data. *British Journal of Psychology*, 3, 271-295.
- Staub, E. (1974). Helping a distressed person: Social, personality, and stimulus determinants. *Advances in experimental social psychology*, 7, 293-341.
- Stout, W. (1987). A nonparametric approach for assessing latent trait unidimensionality. *Psychometrika*, 52(4), 589-617.
- Stout, W. (1990). A new item response theory modeling approach with applications to unidimensionality assessment and ability estimation. *Psychometrika*, 55(2), 293-325.
- Stump, G. S., Husman, J., & Brem, S. K. (2012). The nursing student self-efficacy scale: development using item response theory. *Nursing Research*, 61(3), 149-158. doi: 10.1097/NNR.0b013e318253a750
- Susniene, D., & Jurkauskas, A. (2015). The concepts of quality of life and happiness—correlation and differences. *Engineering Economics*, 63(4), 58-66.
- Tabachnick, B. G., & Fidell, L. S. (2001). *Using multivariate statistics*. Boston: Ally and Bacon.
- Thissen, D. (1991). *MULTILOGTM. User's guide. Multiple, categorical item analysis and test scoring using item response theory*. Chicago: Scientific Software.
- Thissen D., Chen, W. H., & Bock, R. D. (2003). *MULTILOG 7*. Chicago: Scientific Software International.
- Thissen, D., & Steinberg, L. (1986). A taxonomy of item response models. *Psychometrika*, 51, 567-577.
- Thissen, D., & Steinberg, L. (1997). A Response Model for Multiple-Choice Items. En W. J. Van der Linden & R. K. Hambleton (Eds.), *Handbook of Modern Item Response Theory*, (pp. 51-66). New York, Springer.
- Thissen, D., Steinberg, L., & Wainer, H. (1993). Detection of differential item functioning using the parameters of item response models. En P. W. Holland & H. Waines (Eds.), *Differential Item Functioning* (pp. 67-113). Hillsdale, NJ: Erlbaum.
- Timmerman, M. E., & Lorenzo-Seva, U. (2011). Dimensionality Assessment of Ordered Polytomous Items with Parallel Analysis. *Psychological Methods*, 16, 209-220.
- Tornimbeni, S., Pérez, E., & Olaz, F. (2008). *Introducción a la psicometría*. Buenos Aires: Paidós.
- Tremblay, R., Vitaro, F., Gagnon, C., Piché, C., & Reyer, N. (1992) A prosocial scale for the preschool behavior questionnaire concurrent and predictive correlates. *International Journal of Behavioral Development*, 15(2), 227-245.

- Urquiza, V. (2005). *Empatía y Razonamiento Moral Prosocial: Adaptación y Validación de instrumentos para su evaluación psicológica*. Informe Final de Beca Estímulo. Facultad de Psicología, U.B.A.
- Uslaner, E. M. (2002). *The Moral Foundations of Trust*. Cambridge, UK: Cambridge University Press.
- Van Willigen, M. (2000). Differential benefits of volunteering across the life course. *Journals of Gerontology: Series B: Psychological Sciences and Social Sciences*, 55(5), 308-318.
- Vélez, L., & Martínez Lugo, M. (1995). El mundo del trabajo en Puerto Rico. Perfil motivacional de un grupo de trabajadores. *Revista Latinoamericana de Psicología*, 27(2), 283-304.
- Vinichuk, N. V. (2008). *Psychosemantic analysis of Chinese and Russian students' vision of happiness*. Ph.D.Thesis, Far Eastern State Transport University.
- Vivas, J. (2004). Método Distsem: Procedimiento para la evaluación de distancias semánticas [*Distsem Method: Procedure for assessment of semantic distance*]. *Perspectivas en Psicología. Revista de Psicología y Ciencias Afines*, 1(1), 56-61.
- Vivas, J., Comesaña, A. & Vivas, L. (2007). Evaluación de las redes semánticas de conceptos académicos en estudiantes universitarios [*Semantic networks assessment of academic concepts in undergraduate students*]. *PsicoUSF*, 12(1), 111-119.
- Vivas, J., Ricci, L., & Terroni, N. (2003). Centralidad sociocognitiva e influencia social en la construcción de consensos. Conocimiento compartido modulado por el canal de comunicación. *Interdisciplinaria*, 20(2), 147-171.
- Vivas, J. R., Terroni, N., & Ricci, L. (2002). Retricciones del canal de comunicación efectos psicosociales en entornos educativos. *Psicología Escolar e Educativa*, 6(1), 21-29.
- Warneken, F., & Tomasello, M. (2009). The roots of human altruism. *British Journal of Psychology*, 100(3), 455-471.
- Watkins, P. C., Woodward, K., Stone, T., & Kolts, R. L. (2003). *Social Behavior and Personality: an international journal*, 31(21), 431-451. doi: 10.2224/sbp.2003.31.5.431
- Weir, K., & Duveen, G. (1981). Further development and validation of the prosocial behaviour questionnaire for use by teachers. *Journal of Child Psychology and Psychiatry*, 22(4), 357-374.
- Wenner, J. R., & Randall, B. A. (2016). Predictors of prosocial behavior: Differences in middle aged and older adults. *Personality and Individual Differences*, 101, 322-326. doi: 10.1016/j.paid.2016.05.367
- Whittaker, T. A., & Worthington, R. L. (2016). Item Response Theory in Scale Development Research A Critical Analysis. *The Counseling Psychologist*, 44(2), 216-225.
- Wilkinson, R. G., & Pickett, K. (2009). *The spirit level: Why more equal societies almost always do better*. London: Allen Lane.
- Wispé, L. (1987). History of the concept of empathy. In: *Empathy and its development*, ed. N. Eisenberg & J. Strayer. Cambridge University Press.
- Wolke, D., Woods, S., & Samara, M. (2009). Who escapes or remains a victim of bullying in primary school? *British Journal of Developmental Psychology*, 27, 835-851. doi:10.1348/026151008X383003
- Wood, A. M., Froh, J. J., & Geraghty, A. W. A. (2010). Gratitude and well-being: A review and theoretical integration. *Clinical Psychology Review*, 30(7), 890-905. doi: 10.1016/j.cpr.2010.03.005
- Wood, A. M., Maltby, J., Stewart, N., Linley, P. A., & Joseph, S. (2008). A social-cognitive model of trait and state levels of gratitude. *Emotion*, 8, 281-290. doi: 10.1037/1528-3542.8.2.281
- Zahn-Waxler, C., Radke-Yarrow, M., Wagner, E., & Chapman, M. (1992). Development of concern for others. *Developmental Psychology*, 28, 126-136.
- Zieky, M. (1993). Practical questions in the use of DIF statistics in ítem development. In P. W. Holland & H. Wainer (Eds.) *Differential Item Functioning* (pp. 337-347). Hillsdale, NJ: Erlbaum.
- Zuffiano, A., Alessandri, G., Luengo Kanacri, B. P., Pastorelli, C., Milioni, M., Ceravolo, R., et al. (2014). The relation between prosociality and self-esteem from middle adolescence to young adulthood. *Personality and Individual Differences*, 63, 24-29. doi:10.1016/j.paid.2014.01.041.
- Zumbo, B. D. (2007). Three generations of DIF analysis: considering where it has been, where it is now, and where it is going. *Language Assessment Quarterly*, 4(2), 223-233.
- Zumbo, B., Pope, G. A., Watson, J. E., & Hubley, A. M. (1997). An empirical test of Roskam's conjecture about the interpretation of an ICC parameter in personality inventories. *Educational and Psychological Measurement*, 57, 963-969.
- Zwick, R. (1985). *Assessment of the dimensionality of NAEP year 15 reading data* (ETS Res. Rep. N0. 86-4). Princeton: Educational Testing Service.

# APÉNDICES

## APÉNDICE A

### CARRERAS A LAS QUE PERTENECEN LOS PARTICIPANTES DE LA MUESTRA

ABOGACÍA	111	7,2
ACTUACIÓN	9	,6
ACTUARIO	22	1,4
AGRONOMÍA	1	,1
ANTROPOLOGÍA	10	,6
ARQUITECTURA	15	1,0
ARTES AUDIOVISUALES	56	3,6
ARTES ELECTRÓNICAS	2	,2
BIBLIOTECOLOGÍA	2	,1
BIOLOGÍA	6	,4
BIOQUÍMICA	11	,7
BIOTECNOLOGÍA	1	,1
CALÍGRAFO PÚBLICO	1	,1
CINE	1	,1
COMUNICACIÓN SOCIAL	28	1,8
COMERCIO INTERNACIONAL	15	1,0
COREOGRAFÍA	5	,3
CONTADOR PÚBLICO	106	6,9
CIENCIAS AMBIENTALES	6	,4
CIENCIAS DE LA ATMÓSFERA	3	,2
CIENCIAS DE LA COMPUTACIÓN	10	,6
CIENCIAS DE LA EDUCACIÓN	14	,9
CIENCIAS POLÍTICAS	3	,2
CIENCIAS SOCIALES	1	,1
DISEÑO GRÁFICO	11	,7
DISEÑO DE IMAGEN Y SONIDO	9	,6
DISEÑO DE INDUMENTARIA	19	1,3
DISEÑO INDUSTRIAL	10	,6
DISEÑO MULTIMEDIA	1	,1
EDICIÓN	14	,9
ENFERMERÍA	26	1,7

ESCENOGRAFÍA	4	.3
EXPRESIÓN CORPORAL	1	.1
FARMACIA	7	.5
FILOSOFÍA	15	1,0
FÍSICA	2	.1
FONOAUDIOLOGÍA	3	.2
GASTRONOMÍA	1	.1
GEOGRAFÍA	2	.1
GEOLOGÍA	2	.1
HISTORIA	9	.6
INGENIERÍA	43	2,8
KINESIOLOGÍA	10	.6
LETRAS	21	1,4
LICENCIATURA EN ADMINISTRACIÓN	64	4,2
LICENCIATURA EN ECONOMÍA	18	1,2
LIC. EN NIVEL INICIAL	1	.1
MATEMÁTICA	2	.1
MEDICINA	156	10,1
MINORIDAD	1	.1
MUSICOTERAPIA	17	1,1
NUTRICIÓN	11	.7
OBSTETRICIA	8	.5
ODONTOLOGÍA	6	.4
ÓPTICA	1	.1
PALEONTOLOGÍA	1	.1
PERIODISMO	3	.2
PROFESORADO DE LITERATURA	1	.1
PSICOLOGÍA	445	28,9
PSICOPEDAGOGÍA	11	.7
PUBLICIDAD	10	.6
QUÍMICA	4	.3
RADIOLOGÍA	4	.3
RESTAURACIÓN	8	.5
RECURSOS HUMANOS	37	2,4
RELACIONES PÚBLICAS	11	.7
SISTEMAS	9	.6
SOCIOLOGÍA	6	.4
TERAPIA OCUPACIONAL	3	.2

TRADUCTOR PÚBLICO	4	,3
TRABAJO SOCIAL	28	1,8
TURISMO	5	,3
VETERINARIA	16	1,0
Total	1540	100,0

## APÉNDICE B

### CONJUNTO INICIAL DE ÍTEMS

ÍTEM 1	Colaboro en el cuidado del medio ambiente.
ÍTEM 2	Cuido el lugar en el que vivo.
ÍTEM 3	Intento mejorar mi barrio.
ÍTEM 4	Me organizo con otros vecinos para embellecer cuidar el barrio.
ÍTEM 5	Participo en actividades políticas o comunitarias.
ÍTEM 6	Realizo tareas domésticas o mandados.
ÍTEM 7	Voy a buscar apuntes para mis compañeras/os.
ÍTEM 8	Realizo trámites y pagos para una persona anciana o con dificultades físicas.
ÍTEM 9	Dedico un tiempo importante a mejorar el mundo en que vivimos.
ÍTEM 10	Leo a no videntes.
ÍTEM 11	Ayudo a minusválidos a cruzar la calle, cargar objetos, u otra actividad física de corta duración.
ÍTEM 12	Acompaño a personas solas que no pueden movilizarse.
ÍTEM 13	Baño a niños o familiares enfermos.
ÍTEM 14	Cuido de mis familiares cuando se enferman.
ÍTEM 15	Cuido de niños o inválidos.
ÍTEM 16	Ayudo a mis padres o familiares con las tareas diarias.
ÍTEM 17	En una situación de emergencia, pongo en riesgo mi vida por la de un ser querido.
ÍTEM 18	En una situación de emergencia, pongo en riesgo mi vida por la de un desconocido.
ÍTEM 19	Defiendo físicamente a una persona que está siendo agredida.

ÍTEM 20	Comparto las malas experiencias de la vida con amigos para que no cometan los mismos errores.
ÍTEM 21	Comparto todo lo que me pasa con mis amigos.
ÍTEM 22	Comparto mis objetos con familiares y amigos.
ÍTEM 23	Me quedo con lo justo y necesario para vivir, y reparto todo lo demás.
ÍTEM 24	Hago con gusto tareas de más en un trabajo de grupo.
ÍTEM 25	Si tengo un único plato de comida, lo comparto con un desconocido.
ÍTEM 26	Aunque lo podría vender, dono lo que no uso o no me sirve.
ÍTEM 27	Estoy de acuerdo con donar mis órganos si falleciera.
ÍTEM 28	Estoy de acuerdo con que mis familiares donaran sus órganos si fallecieran.
ÍTEM 29	En una situación de emergencia, donaría un órgano a un ser querido.
ÍTEM 30	Dono sangre.
ÍTEM 31	Doy dinero a los mendigos.
ÍTEM 32	Doy dinero a quien me lo pida.
ÍTEM 33	Dono comida, dinero, ropa, juguetes u otros insumos físicos a comedores, centros comunitarios u otras instituciones benéficas.
ÍTEM 34	Doy hasta lo que no tengo.
ÍTEM 35	Organizo donaciones.
ÍTEM 36	Intento que otros donen también.
ÍTEM 37	Cuando veo una persona desconsolada, la aliento y escucho.
ÍTEM 38	Me alejo de los que están mal.
ÍTEM 39	Intento comprender a la otra persona cuando está en una situación problemática.
ÍTEM 40	Siempre me pongo en el lugar del otro.
ÍTEM 41	Intento no ponerme en el lugar del otro todo el tiempo.
ÍTEM 42	Cuando siento que alguien está mal, me acerco y lo abrazo.

ÍTEM 43	Cuando siento que alguien está mal, le demuestro mi cariño.
ÍTEM 44	Cuando siento que alguien está mal, le demuestro que lo entiendo.
ÍTEM 45	Escucho los problemas de los otros.
ÍTEM 46	Escucho con atención a toda persona que necesita contar su problema.
ÍTEM 47	Soy la oreja de todos mis amigos.
ÍTEM 48	Intento comprender qué sienten los otros a cada instante.
ÍTEM 49	Trato de hacer sentir bien a quienes se sienten mal.
ÍTEM 50	Intento que las personas en conflicto comprendan el punto de vista de la otra parte.
ÍTEM 51	Trato de apaciguar las peleas.
ÍTEM 52	Intervengo a favor de cualquiera que esté siendo tratado injustamente.
ÍTEM 53	Si me equivoco, pido perdón
ÍTEM 54	Elogio a quien se lo merece.
ÍTEM 55	Felicito a quienes se esfuerzan para conseguir sus metas.
ÍTEM 56	Les digo a los demás lo mucho que valen.
ÍTEM 57	Le “levanto” la autoestima a mis amigos.
ÍTEM 58	Destaco frente a los demás a quienes hacen las cosas bien.
ÍTEM 59	Expreso afecto y cariño a mis seres queridos.
ÍTEM 60	Expreso a mis seres queridos que los acepto tal cual son.
ÍTEM 61	Si entiendo un tema, lo explico a mis compañeros con gusto.
ÍTEM 62	Explico a mis compañeros cómo hacer los trámites de la facultad.
ÍTEM 63	Si tengo un compañero extranjero, le explico cómo funciona nuestra facultad.
ÍTEM 64	Comparto con mis compañeros las experiencias que tuve en las materias que ya cursé.
ÍTEM 65	Doy información acerca de ubicación de calles y transportes a desconocidos.
ÍTEM 66	Corrijo a las personas cuando se equivocan.

ÍTEM 67	Comparto todo lo que sé para que otros tengan éxito.
ÍTEM 68	Comparto mis conocimientos.
ÍTEM 69	Enseño a los demás cómo estudiar.
ÍTEM 70	Muevo cielo y tierra para resolver los problemas de familiares y amigos.
ÍTEM 71	Afronto junto a mis seres queridos sus situaciones conflictivas y estresantes.
ÍTEM 72	Me sacrifico por mis seres queridos.
ÍTEM 73	No resigno metas personales para beneficiar a mi familia o grupo.
ÍTEM 74	Actúo como bastón de los demás.
ÍTEM 75	Hago compañía a personas solas.
ÍTEM 76	Invito a mi casa a personas que están solas, para compartir un momento agradable.
ÍTEM 77	Colaboro sólo con mi grupo de pertenencia.
ÍTEM 78	Si eso significa un peligro para mí, no alerto a un desconocido de que le pueden hacer un daño.
ÍTEM 79	Hago un favor a otros siempre y cuando esto no me perjudique.
ÍTEM 80	Si tengo un compañero de clases con alguna discapacidad, lo ayudo en todo lo posible.
ÍTEM 81	Intento mantener a la gente con problemas alejada de mi vista.
ÍTEM 82	No realizo acción alguna por las desgracias ajenas.
ÍTEM 83	Hago todo lo que puedo por ayudar las víctimas de catástrofes.
ÍTEM 84	Lloro junto a las víctimas de catástrofes.
ÍTEM 85	Reclamo por los derechos vulnerados, sea de quien sea.
ÍTEM 86	Voy a marchas por reivindicaciones que no me involucran directamente.
ÍTEM 87	Expreso mi apoyo abiertamente a las minorías discriminadas (ej: homosexuales, minorías religiosas, indígenas, inmigrantes de países limítrofes).
ÍTEM 88	Apoyo los paros de transporte por subas salariales aunque los mismos me perjudiquen.

ÍTEM 89	Muestro respeto a los que tienen una opinión diferente acerca de un tema importante para mí.
ÍTEM 90	Aunque pudiera cobrar mucho dinero por él, compartiría gratuitamente un descubrimiento que contribuyera a un mundo mejor.
ÍTEM 91	Cuido los espacios públicos.
ÍTEM 92	No tiro basura en otro lugar que no sea el cesto o contenedor.
ÍTEM 93	Cedo el asiento a ancianos, embarazadas o discapacitados.
ÍTEM 94	Ayudo a ancianos a cruzar la calle o cargar objetos
ÍTEM 95	Me cuelo en las filas.
ÍTEM 96	Pongo la música muy alta a pesar de que a otros les moleste.
ÍTEM 97	Cuido los objetos que me prestan.
ÍTEM 98	Como alimentos calóricos frente a personas que están haciendo dieta.
ÍTEM 99	Fumo en espacios cerrados.
ÍTEM 100	Intento cuidar la energía en mi hogar (p.ej., apagando las luces cuando no las uso).
ÍTEM 101	Respeto las normas de tránsito.
ÍTEM 102	Cuido la seguridad de mi hogar.
ÍTEM 103	Me ofrezco o soy voluntaria/o en instituciones de beneficencia.
ÍTEM 104	Me ofrezco o soy voluntaria/o en instituciones de beneficencia, realizando tareas de limpieza y organización.
ÍTEM 105	Realizo todo tipo de tareas de voluntariado.
ÍTEM 106	Me ofrezco o soy voluntaria/o en actividades que se relacionan con mi futura profesión.
ÍTEM 107	Estudio esta carrera para poder ayudar mejor a los demás en el futuro.
ÍTEM 108	Contribuyo a un mundo mejor estudiando esta carrera.
ÍTEM 109	Gano experiencia en mi profesión futura realizando tareas que benefician a otros.

ÍTEM 110 Genero ambientes más amigables.

---

ÍTEM 111 Me preocupo por el bienestar de cualquier individuo, grupo o comunidad.

---

ÍTEM 112 Trabajo en equipo y coopero con los demás.

---

ÍTEM 113 No realizo ninguna actividad diferente a la de un estudiante de cualquier otra carrera.

---

ÍTEM 114 Estoy comprometido/a de forma importante con la realidad social.

---

## APÉNDICE C

### CONJUNTO DE ÍTEMS RESULTADO DE LA PRIMERA DEPURACIÓN POR JUECES

ÍTEM 1	Me alejo de los que están mal.
ÍTEM 2	Hago donaciones a organizaciones benéficas.
ÍTEM 3	Apoyo los paros por causas que me parecen justas aunque los mismos me perjudiquen.
ÍTEM 4	Escucho con atención al que necesita contar sus problemas.
ÍTEM 5	Me quedo con lo justo y necesario para vivir, y reparto todo lo demás.
ÍTEM 6	Comparto las malas experiencias de la vida con amigos para que no cometan los mismos errores.
ÍTEM 7	Hago grandes sacrificios sin recibir nada a cambio.
ÍTEM 8	Me satisface ver los defectos ajenos.
ÍTEM 9	Destaco a quienes hacen las cosas bien.
ÍTEM 10	Digo mentiras.
ÍTEM 11	Participo en actividades solidarias.
ÍTEM 12	Me pongo en el lugar del otro.
ÍTEM 13	Comparto con mis compañeros las experiencias que tuve en las materias que ya cursé.
ÍTEM 14	Ayudo a los conocidos a enfrentar sus problemas.
ÍTEM 15	Intento “levantar” la autoestima a mis amigos.
ÍTEM 16	Si una persona me cuenta un conflicto, intento que comprenda el punto de vista de la otra parte.
ÍTEM 17	Reconforto en momentos dolorosos a personas que conozco poco.

ÍTEM 18	Si mis compañeros están con dificultades, no me molesta hacer tareas de más en un trabajo de grupo.
ÍTEM 19	Dedico un tiempo importante de mi vida a mejorar el mundo.
ÍTEM 20	Expreso abiertamente mi apoyo a minorías discriminadas, aunque no pertenezca a las mismas.
ÍTEM 21	Evito involucrarme demasiado en los problemas ajenos.
ÍTEM 22	Les digo a los demás lo mucho que valen.
ÍTEM 23	Presto objetos de valor a cualquiera.
ÍTEM 24	Me molesta escuchar a un amigo cuando se pone insistente con sus problemas.
ÍTEM 25	Aunque lo podría vender, dono lo que no uso o no me sirve.
ÍTEM 26	Si entiendo un tema, lo explico a mis compañeros con gusto.
ÍTEM 27	Ofrezco dinero a mis amigos cuando pienso que lo necesitan.
ÍTEM 28	Elogio a quien se lo merece.
ÍTEM 29	No me arriesgo para defender a alguien que corre peligro.
ÍTEM 30	Siento todo dolor ajeno como propio.
ÍTEM 31	Realizo trámites y pagos para una persona anciana o con dificultades físicas.
ÍTEM 32	Evito a los desconocidos que empiezan a comentarme sus problemas.
ÍTEM 33	Aunque pudiera cobrar mucho dinero por él, compartiría gratuitamente un descubrimiento que contribuyera a un mundo mejor.
ÍTEM 34	Felicito a quienes se esfuerzan para conseguir sus metas.
ÍTEM 35	Dono dinero o comida a comedores comunitarios.
ÍTEM 36	Me ocupo sólo de mí.
ÍTEM 37	Comparto mis conocimientos con compañeros.

- ÍTEM 38 Intervengo a favor de cualquiera que esté siendo tratado injustamente.
- 
- ÍTEM 39 Realizo actividades como voluntario.
- 
- ÍTEM 40 Actúo como bastón de los demás.
- 
- ÍTEM 41 Evito prestar mis objetos a amigos.
- 
- ÍTEM 42 Ayudo a los conocidos a enfrentar situaciones conflictivas o estresantes.
- 
- ÍTEM 43 Si tuviera un único plato de comida, lo compartiría con un desconocido.
- 
- ÍTEM 44 Apoyo los reclamos justos, aunque me perjudiquen de forma significativa.
- 
- ÍTEM 45 Acompaño a conocidos en momentos dolorosos.
- 
- ÍTEM 46 Explico a mis compañeros cómo hacer los trámites de la facultad.
- 
- ÍTEM 47 Intento mejorar la autoestima de mis amigos.
- 
- ÍTEM 48 No me comprometo en la defensa de alguien que está siendo atacado.
- 
- ÍTEM 49 Doy hasta lo que no tengo.
- 
- ÍTEM 50 A pesar de que no me guste, si es necesario hago tareas de más.
- 
- ÍTEM 51 Presto apuntes o libros a compañeros.
- 
- ÍTEM 52 Ayudo a ancianos a cruzar la calle o cargar objetos.
- 
- ÍTEM 53 Cuando siento que alguien está mal, le demuestro que lo entiendo.
- 
- ÍTEM 54 Doy dinero a cualquiera que lo necesite.
- 
- ÍTEM 55 Resalto los defectos de los demás.
- 
- ÍTEM 56 Me molesta cuando mis amigos intentan cargarme con sus problemas.
- 
- ÍTEM 57 Intento que las personas en conflicto comprendan el punto de vista de la otra parte.

ÍTEM 58 Si puedo hacerlo, dono sangre para desconocidos.

---

ÍTEM 59 No miento.

---

ÍTEM 60 Me desentiendo de los problemas ajenos.

---

ÍTEM 61 Me ocupo del bienestar de cualquier individuo, grupo o comunidad.

---

ÍTEM 62 Presto dinero a mis amigos aún sabiendo que quizás no puedan devolvérmelo.

## APÉNDICE D

### PROTOCOLO ADMINISTRADO A LA MUESTRA

#### INSTRUCCIONES

A continuación te presentaremos varios enunciados. Rodeá con un círculo el número de la opción que mejor refleje la FRECUENCIA con la cual pensás que realizarías la conducta descrita.

Cada frase debe responderse eligiendo SÓLO UNA de estas respuestas:

Nunca	Casi nunca	A veces	Con frecuencia	Casi siempre	Siempre
1	2	3	4	5	6

Por favor, contestá sinceramente. Este test es **COMPLETAMENTE ANÓNIMO** y no hay respuestas correctas o incorrectas.

Es muy importante que **RESPONDAS A TODAS LAS PREGUNTAS**, por favor no dejes ninguna sin contestar.

En el caso de que no hayas tenido la oportunidad de estar en una situación que te requiera alguna de estas conductas, respondé en base a cómo creés que actuarías vos mismo –con tus valores, personalidad y mentalidad-, imaginándote la situación.

Esta administración es para realizar un ensayo de un conjunto de ítems con el fin de estudiar su calidad y posteriormente construir la primera versión de un test. Por este motivo no podemos brindarte todavía un resultado válido. Su uso está restringido a fines de investigación.

**¡MUCHAS GRACIAS POR TU COLABORACIÓN!**

Género:..... Edad: ..... años

Carrera .....

Universidad/ Instituto .....

<b>Rodeá con un círculo el número de la opción que mejor refleje la frecuencia con la cual pensás que realizarías la conducta descrita</b>	Nunca	Casi Nunca	A veces	Con frecuencia	Casi siempre	Siempre
1. Relego mi beneficio personal para ayudar a otros.	1	2	3	4	5	6
2. Dedico tiempo para facilitar la cursada a mis compañeros.	1	2	3	4	5	6
3. Si tuviera un único plato de comida, lo compartiría con un desconocido.	1	2	3	4	5	6
4. Realizo actividades como voluntario.	1	2	3	4	5	6
5. Cuando siento que alguien está mal, le demuestro que lo entiendo.	1	2	3	4	5	6
6. Actúo como bastón de los demás.	1	2	3	4	5	6
7. Regalo mi material de cursada a compañeros que lo necesitan.	1	2	3	4	5	6
8. Siento todo dolor ajeno como propio.	1	2	3	4	5	6
9. Me comprometo con causas nobles.	1	2	3	4	5	6
10. Dedico un tiempo importante de mi vida a mejorar el mundo.	1	2	3	4	5	6
11. Expreso abiertamente mi apoyo a minorías discriminadas aunque no pertenezca a las mismas.	1	2	3	4	5	6
12. Doy dinero a cualquiera que lo necesite.	1	2	3	4	5	6
13. Me ocupo del bienestar de cualquier individuo, grupo o comunidad.	1	2	3	4	5	6
14. Presto apuntes o libros a compañeros.	1	2	3	4	5	6
15. Explico a mis compañeros cómo hacer los trámites de la facultad.	1	2	3	4	5	6
16. Ofrezco dinero a mis amigos cuando pienso que lo necesitan.	1	2	3	4	5	6
17. Comparto mis conocimientos con compañeros.	1	2	3	4	5	6
18. Evito involucrarme en los problemas ajenos.	1	2	3	4	5	6
19. Dono lo que no uso o no me sirve aunque lo pueda vender.	1	2	3	4	5	6
20. Me quedo con lo justo y necesario para vivir, y reparto todo lo demás.	1	2	3	4	5	6

21. Me pongo en el lugar del otro.	1	2	3	4	5	6
22. Les digo a los demás lo mucho que valen.	1	2	3	4	5	6
23. Participo en actividades solidarias.	1	2	3	4	5	6
24. Hago donaciones a organizaciones benéficas.	1	2	3	4	5	6
25. Comparto las malas experiencias de la vida con amigos para que no cometan los mismos errores.	1	2	3	4	5	6
26. Si una persona me cuenta un conflicto, intento que comprenda el punto de vista de la otra parte.	1	2	3	4	5	6
27. Intento "levantar" la autoestima a mis amigos.	1	2	3	4	5	6
28. Acompaño a conocidos en momentos dolorosos.	1	2	3	4	5	6
29. Aconsejo a mis compañeros sobre cuestiones de la facultad.	1	2	3	4	5	6

## APÉNDICE E

### VERSIÓN ARGENTINA DE LA ESCALA EHP DE MORALES RODRÍGUEZ Y SUÁREZ PÉREZ (2011)

#### INSTRUCCIONES

Las siguientes preguntas se refieren a lo que hacés cotidianamente, o qué harías en una variedad de situaciones. No hay respuestas buenas ni malas, lo más importante es que respondas con sinceridad. Lee cuidadosamente cada frase y elegí la opción que más te describa. Cuando hayas elegido tu respuesta, marcá con una cruz la casilla correspondiente, en una de las cuatro opciones.

#### NO DEJES NINGUNA PREGUNTA SIN CONTESTAR.

		No me describe	Me describe poco	Me describe regular	Me describe bien
1	Trabajaría como voluntario/a en instituciones como cruz roja o bomberos.	1	2	3	4
2	Ayudo a personas de la tercera edad y con capacidades diferentes a cruzar la calle.	1	2	3	4
3	Si una persona es ofendida, imagino lo mal que se siente.	1	2	3	4
4	Cuando me percató de un robo doy aviso a las autoridades.	1	2	3	4
5	Brindaría alimento a personas necesitadas.	1	2	3	4
6	Acostumbro reunir ropa en buen estado, que no uso, para regalarla.	1	2	3	4
7	Cuando veo un accidente en la ruta actúo inmediatamente llamando a emergencias.	1	2	3	4
8	Me gustaría ayudar en situaciones de <u>emergencia</u> aunque no recibiera un sueldo.	1	2	3	4
9	Comprendo cómo se siente una persona cuando es rechazada.	1	2	3	4
10	Comprendo cuando alguien está pasando por un mal momento.	1	2	3	4
11	Antes de criticar a alguien me imagino cómo se sentiría si estuviera en su lugar.	1	2	3	4
12	Cuidaría enfermos sin recibir un pago.	1	2	3	4

\* Versión argentina de la Escala EHP de Mtra. Morales Rodríguez y Lic. Suárez Pérez

**APÉNDICE F**  
**AUTOPERCEPCIÓN DE LA EMPATÍA, CONDUCTA PROSOCIAL,**  
**DADIVOSIDAD Y ALTRUISMO**

**INSTRUCCIONES**

Por favor, completá con números las casillas en blanco, respondiendo cada pregunta:

**Una persona *empática* se pone en el lugar de otras personas, las comprende y experimenta sentimientos similares, apoyándolas emocionalmente.**

a - En una escala de 0 a 10 (0 significa NADA y 10 TOTALMENTE) ¿en qué medida te considerarás una persona *empática*?

.....(\_\_\_\_)

**Una persona *prosocial* realiza acciones que benefician a los demás.**

b - En una escala de 0 a 10 (0 significa NADA y 10 TOTALMENTE) ¿en qué medida te considerarás una persona *prosocial*?

.....(\_\_\_\_)

**Una persona *dadivosa* da, presta o comparte objetos, tiempo, dinero, experiencias, u otro aspecto de valor.**

c - En una escala de 0 a 10 (0 significa NADA y 10 TOTALMENTE) ¿en qué medida te considerarás una persona *dadivosa*?

.....(\_\_\_\_)

**Una persona *altruista* busca beneficiar a los demás aunque esto la perjudique.**

d - En una escala de 0 a 10 (0 significa NADA y 10 TOTALMENTE) ¿en qué medida te considerarás una persona *altruista*?

.....(\_\_\_\_)