



Universidad Latina
Campus Cuernavaca

Licenciatura en Psicología

Incorporada a la Universidad Nacional Autónoma de México
con Clave de Registro No. 8344-25

TEMA:

*Revisión de la Validez y Confiabilidad de las Matrices Progresivas de Raven
Escala General en Alumnos de 15 a 19 Años*

QUE PRESENTA:

Gerardo Nájera Figueroa

PARA OBTENER EL TÍTULO DE:

Licenciado en Psicología

ASESOR DE TESIS:

Mtro. Amador Ocampo Flores

Cuernavaca, Morelos

Marzo, 2015



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

AGRADECIMIENTOS:

A mis abuelos por todo el apoyo incondicional que me dieron a lo largo de mi vida y por inculcarme valores que me han definido como persona.

A mi madre y mi hermana por ayudarme siempre que tuve problemas de salud.

A Coronel y Cerberus por acompañarme siempre y ser parte importante de mi vida, pero sobre todo por aliviar mi soledad en los momento más complicados.

A mi familia que siempre estuvo ahí para ayudarme de manera incondicional.

A mi nana Toña por haberme cuidado, aconsejarme y por haberme brindado apoyo en los momentos de luto.

A Mauricio por todo el conocimiento compartido y por haberme formado como músico.

A Dagon y Arcano por todo el apoyo que me dieron como banda y como amigos, por todo lo que aprendí con ustedes y por hacerme crecer como persona y como músico.

A mis amigos por todos los momentos compartidos, las palabras de apoyo y por ayudarme siempre que lo necesité.

A Sarahí por su amistad, por sus conocimientos compartidos, por la música compartida y por haberme ayudado en gran parte de mi proyecto de tesis.

A Tony por su gran ayuda para realizar este proyecto de tesis y por todos los conocimientos sobre psicometría y estadística compartidos.

A Amador por haber sido mi asesor de tesis y por ayudarme con las correcciones y dudas que me surgieron en las últimas etapas de la consolidación de este proyecto.

A la Directora Olga por su disposición y apoyo brindados durante todo el proceso de revisión de tesis.

Al *Laboratorio de Dexter y Beakman* por detonar mi interés por la ciencia y la investigación.

A *Scooby Doo* por enseñarme que todo fenómeno tiene una explicación lógica.

A *Sepultura, Emperor, Immortal, Burzum, Behemoth, Dodheimsgard, Storm, Vintersorg* y *Opeth* por haberme inspirado, por acompañarme en todo momento y por ser gran parte de mi influencia musical.

A Jean Paul Sartre por enseñarme que en la vida siempre hay elecciones y que elegir no decidir, también es una elección.



Título original: *Revisión de la Validez y Confiabilidad de las Matrices Progresivas de Raven Escala General en Alumnos de 15 a 19 Años*
Escrito por: Gerardo Nájera Figueroa

Diseño editorial por: ©Thésika · Diseño de tesis
contacto@thesika.com.mx | www.thesika.mx
Impreso en México DF durante 2015

Composición & Diseño editorial: J. Martín Rejón (Thésika)
Diseño de cubierta & Encuadernación: J. Martín Rejón (Thésika)
Corrección ortográfica: Rebeca Plata (Thésika)

• GERARDO
NÁJERA
FIGUEROA

Revisión de la Validez y Confiabilidad
de las Matrices Progresivas de Raven
Escala General en Alumnos
de 15 a 19 Años

RESUMEN | 9

INTRODUCCIÓN | 11

CAPÍTULO 1 Generalidades acerca de las pruebas psicológicas | 14

1.1 Antecedentes de las pruebas psicológicas | 16

1.2 Clasificación de las pruebas psicológicas | 18

1.3 Propiedades psicométricas de las pruebas | 20

1.3.1 Confiabilidad | 20

1.3.2 Validez | 22

CAPÍTULO 2 Pruebas de inteligencia | 26

2.1 Definición de inteligencia | 29

2.2 Modelos teóricos que explican la inteligencia | 32

2.3 Inteligencia: heredabilidad, modificabilidad y la interpretación del CI | 43

2.4 Tests de inteligencia más utilizados | 46

CAPÍTULO 3 Matrices progresivas de Raven | 52

3.1 Antecedentes de las matrices progresivas de Raven | 54

3.2 Bases teóricas del test Matrices Progresivas de Raven | 56

3.3 Descripción del test matrices progresivas de Raven | 57

3.3.1 Material de la prueba | 59

3.3.2 Aplicación y calificación | 59

3.4 Propiedades psicométricas del test | 63

3.4.1 Estudios de confiabilidad | 64

3.4.2 Estudios de validez | 65

3.5 Estudios posteriores sobre las propiedades psicométricas del test | 66

3.6 Aplicación del test en diversos grupos | 73

CAPÍTULO 4 Método | 76

4.1 Justificación | 78

4.2 Diseño del estudio | 80

4.3 Muestra | 81

4.4 Objetivos | 81

4.5 Hipótesis | 81

4.6 Definición de variables | 81

4.7 Instrumento | 82

4.8 Procedimiento | 83

4.9 Consideraciones éticas | 84

CAPÍTULO 5 Resultados | 86

CAPÍTULO 6 Conclusiones | 104

6.1 Limitaciones del estudio | 110

BIBLIOGRAFÍA | 113

ÍNDICE DE TABLAS | 117

ÍNDICE DE FIGURAS | 119

La presente investigación tuvo como objetivo revisar la confiabilidad y validez de las Matrices Progresivas de Raven (MPR) en una población mexicana, la cual estuvo conformada por 123 alumnos, de entre los cuales el 53.7% fueron de sexo femenino y el 46.3% de sexo masculino. La media de edad de los sujetos es de 16.32 años, con una desviación estándar de 1.051. Por otro lado, el 39.8% de los participantes pertenecían al 1^{er} grado de preparatoria, el 24.4% pertenecían al 2^o grado y el 35.8% al 3^{er} grado. La mayoría de los participantes radican en el municipio de Cuernavaca (97.7%), el resto pertenece a Jiutepec (4.1%), Temixco (1.6%), Tepoztlán (.8%) y Yauhtepec (.8%). La investigación tiene un enfoque cuantitativo, con un diseño no experimental transeccional exploratorio. El lugar en el que se llevaron a cabo las aplicaciones de las MPR fue la Preparatoria N° 1 de Cuernavaca, Morelos.

Palabras clave: Matrices Progresivas de Raven, coeficiente intelectual, factor g, validez, confiabilidad.



Actualmente existe una gran variedad de instrumentos de medición en psicología. Estas pruebas son una herramienta de apoyo del psicólogo que le permiten recolectar información de uno o varios sujetos. No obstante, cada instrumento debe ser diseñado según lo establecido por la psicometría y atravesar por una serie de procedimientos que permitan constatar si dicho instrumento mide aquello que pretende medir o si es consistente a través del tiempo.

Las pruebas se clasifican de acuerdo con el constructo que pretenden medir, uno de ellos es la inteligencia. Las pruebas de inteligencia se han desarrollado desde hace mucho tiempo y obtuvieron gran popularidad después de que Estados Unidos de América utilizó de manera masiva los test mentales para reclutar rápidamente a una gran cantidad de soldados durante la Primera Guerra Mundial. Después de este acontecimiento, su uso se transportó rápidamente a los ámbitos educativo, industrial y médico, y se reforzó con la aparición de la Segunda Guerra Mundial (PICHOT, 1986). Los tests de inteligencia han sido instrumentos en los que se ha centrado un gran interés. Actualmente siguen siendo utilizados en los diferentes escenarios de actuación de la psicología, y por ello se consideran de gran relevancia.

La presente investigación se centra en el test de MPR, una prueba desarrollada a mediados de la década de 1930 y publicada en 1938 (RAVEN, 2004). Este test ha sido empleado en diversas investigaciones (p.e. GONZÁLEZ & HERNÁNDEZ, 2013; ELISONDO & DONOLO, 2010; BAUSELA & OROZCO, 2009) y es uno de los test más utilizados dentro de la psicología. Es común encontrar esta prueba como parte de los instrumentos a enseñar en la carrera de psicología en diversas universidades en México. El instrumento es también utilizado con frecuencia en escenarios laborales para obtener una medida aproximada de la inteligencia de los sujetos, tanto por su sencilla aplicación como por su fácil calificación y obtención de los resultados. Sin embargo, a pesar de que en el manual de las MPR se reportan estudios sobre su confiabilidad y validez, ninguno de ellos se realizó en población mexicana o en otras partes de Latinoamérica. Lo más cercano a esta población son los estudios realizados en Estados Unidos donde se incluyeron niños de habla hispana (RAVEN, COURT & RAVEN, 2004). La importancia de evaluar la confiabilidad y validez de la prueba en población mexicana motivó la presente investigación. Los resultados permitirán iniciar la adecuación del test para su aplicación y correcta calificación e interpretación en México.

Dentro del capítulo 1 se abordan las generalidades de las pruebas psicológicas, con la finalidad de brindar un panorama más amplio de la construcción de instrumentos a aquellos lectores que estén interesados en el conocimiento de las pruebas psicométricas, por lo que este primer capítulo cuenta con una breve historia de las pruebas o tests psicológicos y las propiedades psicométricas con las que debería contar todo instrumento de evaluación psicológica, además de la clasificación de estos tests.

Para el capítulo 2, llamado pruebas psicológicas de inteligencia, se ofrece la definición de inteligencia, así como los distintos modelos teóricos que explican este concepto. Asimismo, se ofrece un apartado en el que se menciona la intervención de la genética y el ambiente sobre la inteligencia y algunas consideraciones al momento de interpretar un coeficiente intelectual (CI). Además se adhiere un apartado en el que se mencionan algunos de los tests de inteligencia más utilizados en diversos países.

Dentro del capítulo 3 se incluye información específica acerca de la prueba MPR: los antecedentes del test, una descripción general del test MPR, que incorpora su definición según su propósito, el material con el que se aplica, así como las normas que deben tenerse en cuenta al momento de aplicarse y calificarse. Por otro lado, se incluye información sobre las bases teóricas del test y las propiedades psicométricas con las que éste cuenta. Asimismo, se reportan estudios realizados, en distintas poblaciones, con el test de MPR para comprobar si este cuenta o no con una adecuada confiabilidad y validez. A demás, se mencionan algunos estudios actuales en los que se han realizado mediciones con el test MPR.

En el capítulo 4 el lector encontrará los datos acerca del método del presente estudio. Con un enfoque cuantitativo y un diseño transeccional exploratorio, en el cual se tiene como objetivo principal comprobar la confiabilidad y validez del test MPR en un grupo de estudiantes de la preparatoria N° 1 de Cuernavaca, Morelos. Se aplicó la prueba a una muestra de 123 alumnos del turno matutino. Se tomó como variable independiente el test MPR y como variables dependientes la confiabilidad y validez de las MPR. En cuanto a la hipótesis de investigación, se plantea que el test de MPR tendrá validez y confiabilidad al aplicarse en la muestra seleccionada para el presente estudio, al igual que en la población en la que fue validada y confiabilizada. Además, se han incluido las consideraciones éticas para llevar a cabo la aplicación del test.

En el capítulo 5 se exponen los resultados obtenidos, para lo cual se realizó la aplicación de coeficientes Alfa de Cronbach y análisis factoriales.

Finalmente, en el capítulo 6 se incluyen la discusión y conclusiones de los resultados en los cuales se pueden observar ítems con varianza cero y un número de factores más alto del esperado; además se encontró que la eliminación de ciertos ítems contribuiría significativamente al coeficiente

de confiabilidad. Por otra parte, dentro de éste capítulo se incluyen las limitaciones de la muestra y aplicación a las que estuvo sometida la presente investigación.





• CAPÍTULO 1

Generalidades
acerca de las pruebas psicológicas

Las pruebas psicológicas o tests mentales se pueden definir como «las técnicas mediante las cuales se recogen los datos referentes a las características psicológicas de las personas estudiadas» (GONZÁLEZ, 2007, pág. 3). Los instrumentos de medición psicológica se han desarrollado desde hace varios años y su construcción se realiza con base en teorías psicológicas que intentan explicar la conducta humana.

El surgimiento de los tests es importante para la psicología en general. Los instrumentos psicométricos constituyen para el psicólogo una herramienta que permite tener una medida aproximada de un fenómeno, establecer un diagnóstico adecuado, realizar investigación con un enfoque cuantitativo y tomar decisiones acertadas. Sin los tests, probablemente el carácter científico de la psicología estaría limitado.

En la actualidad, se cuenta con una amplia variedad de instrumentos psicológicos, clasificados en diversas categorías. Para conocer mejor los tests psicológicos, en este capítulo se aborda desde la historia de las pruebas, hasta los aspectos relacionados con su construcción.

1.1 ANTECEDENTES DE LAS PRUEBAS PSICOLÓGICAS

La medición psicológica apareció antes de que la psicología se constituyera como una ciencia. Uno de los primeros registros se sitúa en España, en 1575, con el Examen de Ingenios para las Ciencias de Juan Huarte de San Juan. La prueba tenía como propósito clasificar a las personas de acuerdo con sus talentos, lo que permitía adecuar las características de las personas a las diversas ocupaciones (GONZÁLEZ, 2007). Otro de los instrumentos construidos de antaño es el de Gydeón, en 1622, aplicado con la finalidad de llevar a cabo una selección en dos niveles de los combatientes en la Guerra de Israel contra Midia (LÓPEZ, et al, 2010).

A finales del siglo XIX Wilhelm Wundt desarrolló el primer laboratorio de psicología experimental (hecho que separó a la psicología de la filosofía) en el que se encargaba del estudio de los procesos psíquicos (GONZÁLEZ, 2007). Algunos psicólogos de la época se interesaron en los procesos psíquicos descritos por Wundt; otros encontraban más interesante el estudio de las diferencias entre éstos (GONZÁLEZ, 2007), lo que derivó en la psicología diferencial.

Francis Galton es considerado como el precursor de la psicología diferencial; Galton fundó en 1884 un laboratorio antropométrico, donde realizaba mediciones físicas, de funciones sensoriales,

perceptivas y motoras, por ejemplo: estatura, peso, capacidad auditiva, agudeza visual, capacidad sensorial discriminativa, etc. (GONZÁLEZ, 2007). Posteriormente, James McKeen Cattell acuñó el término Test Mental, concepto que se refería a una serie de pruebas psicológicas utilizadas para estudiar las diferencias individuales de estudiantes universitarios. Cattell rechazó la introspección como método de exploración y abogaba por la utilización de los tests y la necesidad de que estos fueran objetivos, para lo cual plantea el uso de baterías de pruebas (PICHOT, 1986).

En 1905, el psicólogo Alfred Binet desarrolló el primer test práctico publicado en un artículo titulado Nuevos Métodos para el Diagnóstico del Nivel Intelectual de los Anormales (Pichot, 1986). Binet, no estaba interesado en la medición de las diferencias individuales, sino que pretendía ocuparse de la medición de las funciones psíquicas superiores, para ello propuso el uso de los tests mentales. Pensando en la objetividad de las medidas, también sugirió que las pruebas debían ser sencillas, que en su aplicación no se invirtiera mucho tiempo, que fueran independientes del examinador y que los resultados obtenidos pudieran contrastarse (GONZÁLEZ, 2007).

Binet llevó a cabo un estudio con niños, en el que evaluó las diferencias entre el niño normal y el retrasado mental. En esta evaluación, Binet tomó en cuenta el examen médico, el examen escolar y el examen psicológico, centrado en evaluar las funciones psicológicas superiores de los niños a través de la ejecución de una prueba; a partir de este estudio aparece la primera prueba de inteligencia en 1905, creada por Binet y su colega Theodore Simon (GONZÁLEZ, 2007).

Durante la Primera Guerra Mundial el desarrollo de los tests mentales se centraba únicamente en la medición de la inteligencia y las aptitudes, por ejemplo, en 1917 los Estados Unidos de América se vieron forzados a reclutar rápidamente miembros para su ejército y asignarles un grado militar; para ello se tomó la decisión de emplear tests mentales de manera masiva. Debido al éxito obtenido, algunos creyeron que los psicólogos habían fabricado un método cuasi perfecto para medir la inteligencia, tanto así que llegó a emplearse en la industria, la educación y la medicina (PICHOT, 1986).

La Segunda Guerra Mundial reafirmó la importancia de los tests de inteligencia y aptitudes e impulsó el desarrollo de tests de personalidad. Los tests de personalidad se desarrollaron a la par que los tests de inteligencia, sólo que con mayor lentitud, ya que el interés se centraba en los tests psicométricos (PICHOT, 1986). Los primeros intentos por explorar la personalidad se centraron en la obra de Freud La Interpretación de los Sueños, publicada en 1900, de la cual se desprenden las

denominada Técnicas Projectivas. Ejemplos de estas técnicas projectivas son: el Test de Asociación de Palabras de C. Jung, El Rorschach de Herman Rorschach, Test de Dibujo de la Figura Humana de Goodenough, Test de Apreciación Temática (TAT) de Murray y el Test de Tolerancia a la Frustración de Rozensweig, publicados en 1906, 1921, 1926, 1938 y 1944, respectivamente (González, 2007).

Desde que la psicología se consolidó como una ciencia, los tests psicológicos la han acompañado. A partir de la fabricación de pruebas, la psicología tuvo la necesidad de desarrollar las teorías que sustentan su construcción, una parte de suma importancia tanto para la recolección de datos como para su tratamiento estadístico. Las pruebas psicométricas, a diferencia de las técnicas projectivas, cuentan con un sustento teórico para su construcción, lo que les confiere su carácter científico. El conocimiento de la psicometría es imprescindible para comprender un instrumento, es por ello que en el siguiente apartado se aborda aspectos que se deben tener en cuenta a la hora de construir o utilizar un instrumento.

1.2 CLASIFICACIÓN DE LAS PRUEBAS PSICOLÓGICAS

Existen diversas clasificaciones en las que se pueden encuadrar las pruebas psicológicas, según el material que se utiliza, de acuerdo a su(s) objetivos(s), por el número de sujetos a los que se les puede administrar, según el tipo de estímulos presentes en el test, etc. En la tabla N° 1 se presentan algunas clasificaciones:

Tabla N° 1. Clasificaciones de las pruebas psicológicas

CLASIFICACIÓN	COMPRENDE
De acuerdo con sus propósitos	Conocimientos adquiridos. Aptitudes (habilidades, destrezas). Inteligencia. Actitudes. Personalidad. Projectivas.
Según la forma de aplicación	Heteroaplicadas. Autoaplicadas. Autoevaluación asistida.
Según el objetivo de la prueba	Tests de eficiencia. Tests de personalidad.
Según el estímulo presentado	Verbales. No verbales.
De acuerdo con el número de sujetos	Individuales. Colectivas.

De acuerdo con la construcción del test	Estandarizados (cuestionario, escala). No estandarizados (Elaborado racionalmente o empíricamente, mnemotécnico o ideográfico).
Según el tipo de respuesta	Seleccionada o cerrada. Libre o abierta. Entre dos alternativas. Elección múltiple. Reacciones complementarias sin guía.
Según evidencia-disimulación	Projectivas. No projectivas.
Según el modo de interpretación	Detallado y fijo. Abierto y libre. Cuantitativo / no cuantitativo. Limitado a una característica u holístico.
Según el principio básico del método	Subjetivo-introspectivo Objetivos Projectivos
Según las condiciones de administración	Laboratorio Consulta Docente

Nota. Esta tabla se realizó de acuerdo a la clasificación de los tests psicológicos ofrecida por López et al. (2010).

Las clasificaciones de los test pueden resultar de gran ayuda para conocer la variedad de pruebas existentes. También ayudan a saber qué tipo de prueba sería conveniente aplicar si se quiere obtener determinada información sobre ciertos sujetos; no obstante, algunas clasificaciones como la anterior resultan muy extensas. Una clasificación más concreta podría ser la presentada en la tabla N° 2 propuesta por (GONZÁLEZ, 2007):

Tabla N° 2. Clasificaciones de las pruebas psicológicas

CLASIFICACIÓN	COMPRENDE
Según su construcción	Psicométricos. Projectivos.
Según el modo de aplicación	Individual Colectiva
Según la conducta que mide	Inteligencia. Personalidad. Atención. Ansiedad, etc.
Según el material que utilizan	Papel y lápiz. Ejecución. Verbales.

Nota. La presente tabla se realizó con base en la clasificación realizada por González (2007).

Existen otras clasificaciones realizadas por otros autores, sin embargo, en esta investigación no se pretende profundizar al respecto, sino sólo proporcionar algunas clasificaciones comúnmente observadas y así conocer, a grandes rasgos, las variadas maneras en las que se diversifican los tests. No obstante, las pruebas de inteligencia constituyen un tema relevante para este trabajo, por ello en el siguiente capítulo se trata de manera más extensa y detallada.

1.3 PROPIEDADES PSICOMÉTRICAS DE LAS PRUEBAS

La psicometría es la rama de la psicología encargada de la medida de los fenómenos psicológicos a través del uso de la estadística, la cual permite elaborar nuevas técnicas de medición. La psicometría comprende tres núcleos temáticos: 1) Teorías de la medición, 2) Escalamiento y 3) Teorías de los tests. Estos núcleos se interesan en el establecimiento de las condiciones de calificación, asignación de escalas de medición centradas en fenómenos estrictamente psicológicos y apego a las teorías que sustenten la adecuada elaboración de un test como son la Teoría Clásica de los Test (TCT) y la Teoría de Respuesta al Ítem (TRI), respectivamente (ABAD, GARRIDO, OLEA & PONSODA, 2006).

La evaluación en psicología no sólo se lleva a cabo con instrumentos psicométricos, sino que se puede llevar a cabo a partir de diversas técnicas, por ejemplo, entrevistas, cuestionarios, observación, psicofisiología, índices clínicos y registros médicos (JOHNSTON, FRENCH, BONETTI & JOHNSTON, 2004). No obstante, el uso de instrumentos psicométricos en determinados casos, pueden proporcionar al psicólogo mayor información acerca del fenómeno que pretenda estudiar. Sin embargo, cuando se utilizan tests psicométricos, el psicólogo debe tener presente los conceptos confiabilidad y validez.

1.3.1 CONFIABILIDAD.

Si a un grupo de estudiantes hoy se les aplicara un test para medir determinado atributo y después de un par de años se decidiera volverlo aplicar, con las mismas condiciones, el test debería arrojar resultados similares. Todos los tests psicométricos deben presentar esta propiedad denominada confiabilidad. La confiabilidad se puede definir como «el grado de estabilidad, previsibilidad, exactitud y consistencia con la que se realiza una evaluación» (BARRON, BROWN, EGAN, GESUALDI & MARCHUK, 2008, pág. 29). Lo anterior puede traducirse como el grado en que las puntuaciones obtenidas por un grupo determinado de personas son consistentes a través del tiempo si se aplica el mismo test o una forma equivalente, con excepción de los tests que miden estados emocionales, debido a que éstos si pueden presentar cambios constantes (LÓPEZ, et al, 2010).

La confiabilidad interna de un instrumento, por otro lado, depende del grado en el que los reactivos de un test se encuentran o no relacionados, es decir, los reactivos de los que está compuesto un test deben medir un mismo constructo. Si las puntuaciones de un test no son consistentes, quiere decir que los reactivos no miden adecuadamente el constructo que se pretende medir (GARCÍA, DEL CASTILLO, GUZMÁN & MARTÍNEZ, 2010). Existen diversos métodos a través de los cuales es posible conocer la confiabilidad de un test, a continuación se mencionan.

Una forma de saber si un instrumento es confiable consiste en aplicar en dos momentos distintos un mismo instrumento a N número de sujetos, a esto se le conoce como método test-retest. El tratamiento estadístico para estas puntuaciones debe ser una correlación de Pearson; si la correlación entre los puntajes obtenidos es cercana a uno, se considera que el instrumento es confiable. Mientras las correlaciones se acerquen más a uno significa que el instrumento es más consistente (ABAD, GARRIDO, OLEA & PONSODA, 2006); además, mientras mayor sea el tiempo que se deje transcurrir entre cada aplicación del test, mejor será la consistencia de la prueba, siempre y cuando se haya obtenido una correlación cercana a uno.

Otra manera de proceder para saber si un determinado instrumento es confiable es el método de división por mitades, en el cual, como su nombre lo indica, el test se parte en mitades equivalentes para poder aplicarse a un grupo de participantes. Por ejemplo, se toma el 50% de los ítems del test A y se aplica a los participantes, cuando finalizan se les aplica el otro 50% de los ítems, después de correlacionar los resultados obtenidos y si la correlación es cercana a uno, se considera que la prueba es confiable. Para evitar que los participantes se desesperen o se aburran al calificar muchos ítems, se elige de manera aleatoria la división de los ítems, se debe recordar que mientras el número de reactivos sea menor, es probable que la confiabilidad sea menor (GARCÍA, DEL CASTILLO, GUZMÁN & MARTÍNEZ, 2010).

El método de formas paralelas, es similar al método de división por mitades, la diferencia consiste en que se diseñan dos instrumentos con el propósito de medir el mismo constructo, sólo que con reactivos diferentes. La aplicación de los dos instrumentos se realiza en la misma muestra. El uso de este método no es muy frecuente debido a la cantidad de tiempo requerido para la construcción de ambos instrumentos, lo cual resulta poco práctico, por ello las formas paralelas se utilizan en situaciones convenientes para el investigador, por ejemplo, estudios en los cuales se pretenda realizar dos mediciones evitando el recuerdo (ABAD, GARRIDO, OLEA & PONSODA, 2006).

Por último, el método más sencillo y con frecuencia utilizado es el coeficiente α de Cronbach. Este coeficiente es un indicador de la consistencia interna del instrumento, es decir, el grado en el que un constructo se encuentra presente entre los ítems de un test. El coeficiente α de Cronbach no es sino una correlación entre los ítems del test y requiere de una sola aplicación, por ello es uno de

los métodos más prácticos para revisar la correlación de un instrumento. Un adecuado coeficiente sería aquel que se encuentra entre .70 y .90 (FITZPATRICK, DAVEY, BUXTON & JONES, 2001). La fórmula a través de la que se obtiene el coeficiente α de Cronbach se expresa de la siguiente manera:

$$\alpha = \frac{k}{k-1} \left(1 - \sum \frac{s_i^2}{s_x^2} \right)$$

En las ciencias sociales no existen reglas exactas acerca de la medición, por ello es que no se expresa a partir de qué número se considera que un instrumento presenta un coeficiente confiable o no. Sin embargo, Hernández, Fernández y Baptista (2006) ofrecen una manera sencilla de interpretar los coeficientes de confiabilidad: si se obtiene un coeficiente de 0.25 o menor, se dice que la confiabilidad es baja; si se obtiene un coeficiente de 0.50, se dice que el coeficiente de confiabilidad es regular; si el coeficiente es de 0.75, la confiabilidad es aceptable; por último, si el coeficiente es de 0.90 o mayor, la confiabilidad es alta, de significativa importancia.

1.3.2 VALIDEZ.

El concepto de validez se refiere al grado en que un instrumento cumple con el propósito con el que fue diseñado. Se puede decir que un instrumento es válido si éste mide aquello que pretende medir (NUNNALLY, 2009). López (2010, pág. 13) ofrece la siguiente definición: «la validez es la estimación de la utilidad práctica de la prueba, de su fuerza predictiva, inferencial y diagnóstica».

Contar con validez es de suma importancia para un instrumento. La validación, por otro lado, es una tarea a la que se debe someter todo instrumento que se pretenda utilizar para obtener una medida. Cuando se habla de conductas, la medición resulta complicada y supone una constante revisión de aquellas técnicas utilizadas para la medición, al contrario de otro tipo de medidas más exactas, por ejemplo, una regla. Una regla ofrece siempre la misma medida y no varía según las condiciones en las que se mida. No obstante, los fenómenos psicológicos son proclives de cambio y, por lo tanto, los instrumentos con los que se miden se deben someter a constantes revisiones (NUNNALLY, 2009).

Cuando un investigador pretende desarrollar un instrumento, el primer requisito es que el investigador cuente con un conocimiento amplio del constructo que pretende medir. Si la información teórica del constructo con la que se cuenta no es adecuada o suficiente, es probable que los reactivos se encuentren mal formulados y, en consecuencia, no midan el constructo que se tenía pensado medir. Por ello, la identificación y conocimiento de los síntomas, signos o conductas que caracterizan un fenómeno son de vital importancia a la hora de fabricar un test válido (GARCÍA, DEL CASTILLO, GUZMÁN & MARTÍNEZ, 2010).

La validez se puede diferenciar en un primer momento como validez interna y validez externa. La validez interna tiene que ver con los reactivos de los que está compuesto el test, aquellos ítems incluidos en el instrumento deben componer una matriz suficientemente representativa del constructo. En una investigación experimental, se dice que se cuenta con validez interna cuando se tiene la certeza de que la manipulación de la variable independiente explica los cambios en la variable dependiente (presencia-absencia) y no otras fuentes denominadas explicaciones rivales. La validez externa, por otra parte, se refiere al grado en el que los resultados de un experimento se pueden extrapolar a situaciones no experimentales, otros sujetos u otras poblaciones (HERNÁNDEZ, FERNÁNDEZ & BAPTISTA, 2006).

Además de la validez interna y externa, existen otros tres tipos de validez, que son la validez predictiva, de contenido y de constructo. La validez predictiva, como su nombre lo indica, tiene que ver con el grado en el que un instrumento puede predecir una conducta, a la cual se le denomina criterio. Cuando se dice que un instrumento predice adecuadamente un criterio, se refiere a la relación entre el instrumento y los eventos ocurridos antes, durante y después de su aplicación, es decir, un instrumento no sólo puede predecir conductas que sucederán en un futuro, sino también conductas que ya sucedieron o que están sucediendo (NUNNALLY, 2009). Un instrumento es capaz de predecir una conducta ocurrida en la infancia de un sujeto, por ejemplo. Asimismo, si se aplicara un test para medir inteligencia se obtendría un CI con el que cuenta un sujeto actualmente, pero a partir de esto se podría predecir el futuro rendimiento escolar del sujeto.

Para constar si un test predice adecuadamente un criterio, se debe saber el nivel en el que el instrumento se correlaciona con el criterio, por ejemplo, si en un estudio se encuentra que el uso de drogas está altamente correlacionado con el rendimiento escolar, entonces consumir o no sustancias consideradas drogas sería una medida adecuada para predecir el éxito o fracaso escolar (ABAD, GARRIDO, OLEA & PONSODA, 2006).

La validez de contenido se conoce a través del conjunto de reactivos de los que está compuesto un instrumento. Los ítems que se han elegido para conformar un test deben representar adecuadamente una teoría, un cúmulo de conocimientos o una serie de habilidades. Al igual que una investigación pretende generalizar los hallazgos extraídos de una muestra de individuos al universo al que pertenecen, los tests deben ser capaces de generalizar los resultados de un sujeto al universo de conocimiento al que pertenecen. Por ejemplo, si se aplica una prueba de rendimiento escolar a un grupo de estudiantes, el instrumento seguramente estaría compuesto por varias secciones de conocimientos en diversas áreas. Si una de las áreas se llamara Historia, los ítems seleccionados (muestra) deberían representar adecuadamente el nivel de conocimientos en la materia demostrados por alumnos con determinada edad y grado de educación escolar, y, al mismo tiempo, los

resultados obtenidos por los alumnos deben ser capaces de generalizarse como parte del universo de conocimientos pertenecientes a la materia de Historia (ARAGÓN, 2004).

García et al (2010) y Abad et al (2006) consideran que la validez de contenido se encuentra altamente relacionada con la validez de *facie* o apariencia, es decir, dice que un instrumento debe dar la impresión de ser un instrumento y que al usuario le debe proporcionar, en apariencia, cierta información acerca de lo que el instrumento pretende medir.

La validez de constructo surge a partir de la necesidad de medir conceptos abstractos y complejos, para los que la validez predictiva y la validez de contenido no pueden sugerir una adecuada manera de medirlos (NUNNALLY, 2009). Un constructo es un concepto abstracto o variable relevante que forma parte de una teoría psicológica, por lo tanto, la validez de constructo tiene que ver con la adecuada capacidad de un instrumento para evaluar un rasgo psicológico derivado de una determinada postura teórica (ARAGÓN, 2004).

Hogan (2004) considera a la validez de contenido y a la validez de criterio como parte de la validez de constructo, de tal forma que las dos anteriores funcionan como fuentes de información sobre un determinado instrumento, que a su vez constituyen casos a través de los cuales se puede comprobar el grado en que un test mide un determinado constructo.

Para Gregory (2001) la validez de constructo se puede determinar de acuerdo a los siguientes métodos:

1. Análisis para determinar si las subpruebas o reactivos de una prueba son homogéneos y, por tanto, miden un solo constructo.
2. Estudio de los cambios relacionados con el desarrollo para determinar si son consistentes con la teoría del constructo.
3. Investigación para evaluar si las diferencias grupales en las puntuaciones de una prueba son consistentes con la teoría.
4. Análisis para determinar si los efectos de intervención sobre las puntuaciones de prueba son consistentes con la teoría.
5. Correlación de la prueba con otras pruebas, medidas relacionadas y no relacionadas.
6. Análisis factorial de las puntuaciones de la prueba en relación con otras fuentes de información.

Uno de los métodos más frecuentemente utilizado para determinar la validez de constructo es el análisis factorial. Este método consiste en reducir el número de dimensiones, también conocidas como factores o rasgos comunes, a través de los cuales se puede llegar a explicar la relación entre distintas variables. El análisis factorial es una técnica estadística que hace más simple la explicación de un constructo (ANASTASI & URBINA, 1998).

El análisis factorial cuenta con dos formas distintas, análisis factorial exploratorio (AFE)

y análisis factorial confirmatorio (AFC). En el AFE, no se tiene un conocimiento claro acerca del número total de factores en los que se encuentra agrupado el constructo ni las variables con más saturación dentro de cada factor, mientras que en el AFC el investigador, a partir del compendio de estudios que ha ido recabando, plantea previamente hipótesis acerca del número de factores en los que cree que se va a agrupar el constructo y hacia dónde se van a inclinar las cargas de las variables (ABAD, GARRIDO, OLEA & PONSODA, 2006).

El motivo por el cual se ha hecho hincapié en las propiedades psicométricas de un test se debe a que el instrumento en el que se centra la presente investigación es un test psicométrico y por tanto se rige por los lineamientos teóricos que permiten la construcción de instrumentos psicológicos. No obstante, las MPR también pertenecen a la clasificación de pruebas de inteligencia, es por ello que el siguiente capítulo ha sido destinado a las mismas.





• CAPÍTULO 2

Pruebas de inteligencia

De acuerdo con Hogan (2004), las investigaciones sobre inteligencia realizadas por los psicólogos, se pueden agrupar en cuatro áreas diferentes:

1. las teorías sobre la naturaleza de la inteligencia,
2. la metodología, ya sea teórica o aplicada, para la medición de la inteligencia,
3. el ámbito de las diferencias grupales (por edad, género, condición socioeconómica, grupo racial o étnico, etc.), y
4. las influencias hereditarias y ambientales en el desarrollo de la inteligencia.

Sin embargo, la investigación acerca del término inteligencia es sumamente extensa; incluso revistas como *Intelligence* y *The Journal of Psychoeducational assessment* han surgido debido al gran interés centrado en el estudio del constructo (GREGORY, 2001).

A pesar de toda investigación existente, la inteligencia sigue siendo un reto tanto para quien pretende conocer sobre el tema, como para los estudiosos. Las definiciones sobre inteligencia son diversas y varían según la postura científica o teórica; la biología, la filosofía, la educación, e incluso los psicólogos con especialidad en áreas diferentes cuentan con definiciones distintas sobre el término (ANASTASI & URBINA, 1998).

No obstante, los diversos estudios que se han llevado a cabo sobre la inteligencia han dado paso a las teorías psicológicas que intentan explicarla. Es importante mencionar este hecho dado que los psicómetros, a lo largo de la historia de la psicología como ciencia, han llevado a cabo la construcción de instrumentos capaces de medir este constructo, al menos relativamente, partiendo siempre de una determinada teoría (HOGAN, 2004).

En un principio, las tres clasificaciones más comunes para estos instrumentos de medición psicológica fueron: pruebas de inteligencia, pruebas de capacidad mental y pruebas de aptitud. Sin embargo, como la definición exacta de inteligencia es complicada y no existe, los investigadores han optado por utilizar términos más acordes con una específica finalidad (Por ejemplo, el *Otis Intelligence Scale*, que pasó a ser el *Otis-Lennon Ability Test* y finalmente el *Otis-Lennon School Ability Test*). Por lo tanto, actualmente es común encontrar tests con nombres que hacen referencia a la medición de aspectos específicos, pero que están relacionados con la capacidad mental o la inteligencia (HOGAN, 2004).

Los psicómetros dedicados a desarrollar instrumentos para medir el constructo inteligencia o parte de él, suelen desarrollar una definición de inteligencia como punto de partida (GREGORY, 2001). Además, Anastasi & Urbina (1998) hacen hincapié en que, tanto la selección de un instrumento apropiado, como la interpretación de sus resultados, exigen el conocimiento de la investigación conductual que se pretenda estudiar. Por consiguiente, la definición de inteligencia es el siguiente paso.

2.1 DEFINICIÓN DE INTELIGENCIA

En 1971 (COLOM & PUEYO, 1999), Weschler afirmó que la inteligencia era el concepto del cual se tenía más conocimiento en psicología, en comparación con otros fenómenos psicológicos, aunque muchos otros psicólogos opinaban lo contrario. Actualmente, los psicólogos también se opondrían a dicha afirmación, dado que el constructo inteligencia sigue generando discusiones entre los estudiosos y no se cuenta con una definición exacta del término.

Antes de comenzar a exponer algunas de las definiciones realizadas por los expertos, es importante mencionar que la naturaleza de la definición es la que se debe tener en cuenta en primer lugar. Sternberg (1986) diferencia entre definiciones reales y operacionales. Esta última se refiere a un concepto definido en términos de cómo se mide. La distinción entre estos dos tipos de definiciones es relevante debido a las dificultades que supone construir una definición concreta de inteligencia. Uno de los casos más extremos, con respecto a las definiciones operacionales, es el de Boring (1923, pág. 170), quien definió la inteligencia como «aquello que miden las pruebas». Esta propuesta para definir la inteligencia, por extraña que parezca, pretendía evitar los grandes desacuerdos relacionados con el mismo propósito.

Cuando se habla de una definición operacional de inteligencia, es posible encontrar dos situaciones desventajosas. En primer lugar, se dice que las definiciones operacionales son circulares, es decir, que suponen una comprensión del término anteriormente definido. En ese sentido, las pruebas de inteligencia fueron diseñadas para poder medir la inteligencia, mas no para definirla. En segundo lugar, las definiciones operacionales son un impedimento para una mejor comprensión de

la verdadera naturaleza de la inteligencia, debido a que se anula la posibilidad de un análisis acerca de lo adecuado de las teorías que la sustentan (STERNBERG, 1986).

Si se depende en gran medida de las definiciones operacionales de inteligencia, es posible evidenciar cierta dificultad a la hora de establecer la validez predictiva de un instrumento a través de la correlación con otras pruebas construidas con anterioridad. Esta práctica sugiere que las nuevas pruebas serían válidas siempre y cuando los resultados presenten una alta correlación con otras más viejas, que sirven de criterio contra el que se comparan. Esto deriva en una evidente dificultad para adoptar conceptos nuevos que pudieran contribuir al conocimiento de la inteligencia, debido a que las pruebas se contrastan con otros tests sustentados en conceptos que quizá carezcan de vigencia (GREGORY, 2001).

Al contrario de las definiciones operacionales, las definiciones teóricas pretenden proporcionar una explicación acerca de la verdadera naturaleza de aquello que se pretende definir (Sternberg, 1986). Una de las maneras más comunes de construir una definición real de inteligencia es pidiendo a los expertos en la materia que la definan (GREGORY, 2001).

En 1921 tuvo lugar un simposio titulado *Intelligence and its Measurement*, en el cual se dieron cita catorce expertos para discutir acerca del término inteligencia y su significado. De este simposio se desprendieron varias definiciones, algunas de las cuales siguen vigentes actualmente y representan teorías completas (MOLERO, SAIZ, & ESTEBAN, 1998); a continuación se presentan de manera textual las definiciones obtenidas del simposio *Intelligence and its measurement* y de una actualización de este primer simposio: *What is intelligence?* (THORNDIKE, 1921; STERNBERG & DETTERMAN, 1921; MOLERO, SAIZ & ESTEBAN, 1998):

- Spearman (1904, 1923): una capacidad general que implica principalmente la deducción de relaciones y correlatos.
- Binet & Simon (1905): la capacidad para juzgar bien, para comprender bien y para razonar bien.
- Terman (1916): La capacidad para formar conceptos y para comprender su importancia.
- Pintner (1921): la capacidad del individuo para adaptarse de manera adecuada a las situaciones relativamente nuevas de la vida.
- Thorndike (1921): el poder de dar buenas respuestas, desde el punto de vista de la verdad o el hecho.
- Thurstone (1921): la capacidad para inhibir las adaptaciones instintivas, para imaginar de manera flexible diferentes respuestas y para realizar adaptaciones instintivas modificadas en conducta manifiesta.
- Wechsler (1939): el conjunto o capacidad global del individuo para actuar de manera propositiva, pensar de manera racional y enfrentarse de manera efectiva con el ambiente.
- Humphreys (1971): el repertorio completo de habilidades, conocimiento, inclinación para el

aprendizaje y tendencias a la generalización adquiridas, considerados de naturaleza intelectual y que están disponibles en un momento determinado.

- Piaget (1972): un término genérico que indica las formas superiores de organización o equilibrio de la estructura cognitiva que se utilizan para la adaptación al ambiente físico y social.
- Sternberg (1985a, 1986): la capacidad mental para automatizar el procesamiento de información y para emitir en un contexto la conducta adecuada en respuesta a situaciones novedosas; la inteligencia también incluye metacomponentes, componentes de ejecución y de adquisición del conocimiento (los cuales se analizan después).
- Eysenck (1986): transmisión correcta de la información a través de la corteza.
- Gardner (1986): la capacidad o habilidad para resolver problemas o para crear productos que se consideran valiosos dentro de uno o más entornos.
- Ceci (1994): capacidades innatas múltiples que dan un rango de posibilidades; éstas se desarrollan (o dejan de desarrollarse o se desarrollan y después se atrofian) dependiendo de la motivación y exposición a experiencias educativas relevantes (pág. 171; pág.171; pág. 19)

La presente investigación se inclina por la definición proporcionada por Spearman, ya que es en ésta en la que el creador de las MPR se ha centrado para el desarrollo de las mismas.

Más recientemente y a raíz de la controversia provocada por la publicación de *The Bell Curve* en 1994 por Herrnstein y Murray, obra que fue tachada de pseudocientífica gracias a las críticas realizadas por los medios de comunicación, se publicó en el *Wall Street Journal*, el 13 de diciembre del mismo año, un artículo con 25 puntos sobre el estudio científico de la inteligencia y apoyado por 52 expertos a nivel internacional, con la finalidad de esclarecer y corregir los errores cometidos por los medios a través de las críticas. Uno de los puntos principales de este artículo fue establecer una definición (COLOM & PUEYO, 1999):

La inteligencia es una capacidad mental muy general que permite razonar, planificar, resolver problemas, pensar de modo abstracto, comprender ideas complejas, aprender con rapidez y aprender de la experiencia. No constituye un simple conocimiento enciclopédico, una habilidad académica particular o una pericia para resolver tests, sino que refleja una capacidad más amplia y profunda para comprender el ambiente –darse cuenta, dar sentido a las cosas o imaginar qué se debe hacer (pág. 58). |

Posterior a esta declaración, la *American Psychological Association* (APA) tomó la decisión de realizar un informe más completo y detallado acerca de la inteligencia y la investigación científica actual de la misma, para lo cual convocó un grupo de reconocidos expertos en la materia. Según el informe oficial de la APA, publicado en 1996, la polémica suscitada por la publicación de *The Bell Curve* estaba repleta de declaraciones que iban más allá de la evidencia empírica con la que se contaba. Además, el informe aborda la concepción científica de la inteligencia a partir de la aproxi-

mación psicométrica y los tests psicológicos, presentando una explicación de lo que significan las puntuaciones en los tests (CI), así como una aclaración de lo que predicen estos tests y la precisión con la que lo hacen, sin mencionar la notable estabilidad de las puntuaciones en los tests de inteligencia durante el desarrollo individual y la utilidad de los tests para predecir adecuadamente el rendimiento escolar u ocupacional, además de la relación que tiene la inteligencia con situaciones sociales tales como el estatus socioeconómico y la delincuencia (COLOM & PUEYO, 1999).

Ya se han expuesto algunas de las definiciones más comunes del término inteligencia y el debate que este concepto ha generado y sigue generando actualmente, sin embargo, es importante comprender que estas definiciones provienen de teorías completas, fruto de largos años de investigación, por lo tanto en el siguiente apartado se abordan algunos de los modelos teóricos más representativos en lo que respecta al estudio científico de la inteligencia.

2.2 MODELOS TEÓRICOS QUE EXPLICAN LA INTELIGENCIA

La investigación acerca de la naturaleza de la inteligencia es muy extensa y, actualmente, existen diversos modelos teóricos que la explican; sin embargo, todas las teorías sobre inteligencia que se han conocido y se conocen en la actualidad, están relacionadas o tienen sus raíces en dos principales teorías clásicas: La g de Spearman y la Teoría de las Capacidades Mentales Primarias propuesta por Thurstone.

Charles Spearman, a partir de sus estudios, en los cuales determinó las correlaciones entre diversas pruebas de capacidad intelectual y de funciones sensoriales simples, propuso una teoría de la inteligencia, que él mismo denominó como: Teoría bifactorial; aunque es más comúnmente conocida como Teoría de la g de Spearman. Spearman consideraba que, debido a las elevadas correlaciones encontradas, se podía concluir que la ejecución de estas pruebas era producto de una capacidad mental general, a la cual denominó factor g. No obstante, las correlaciones entre las diferentes pruebas no eran perfectas, sino que cada una de ellas presentaba una varianza específica independiente de g, conocida como factor s. Así, cada conjunto de pruebas podía presentar una serie de factores s y un solo factor g (HOGAN, 2004). La teoría de la g de Spearman será detallada más adelante.

La g de Spearman ha sido una teoría medular en el estudio de la inteligencia humana y en la comprensión que los psicólogos tienen de la inteligencia, incluso actualmente sigue vigente en el campo de la investigación (JENSEN, 1979; HOGAN, 2004). Sin embargo, no fue la única teoría clásica relevante.

La Teoría de las Capacidades Mentales Primarias fue propuesta por el psicólogo Louis Leon Thurstone, quien consideraba, en contraposición a Spearman, que las correlaciones observadas

entre diversas pruebas en realidad no eran lo suficientemente altas como para considerar que se trataba de la medición de un solo factor; sino por el contrario, se trataba de correlaciones lo bastante bajas como para dar cuenta de varios factores independientes (THURSTONE, 1938).

Thurstone también contribuyó en gran medida al método del análisis factorial, desarrollando procedimientos de análisis que ayudaban a encontrar matrices de correlación para factores grupales; dos de sus obras, *The vectors of the mind* y *Multiple-factor analysis*, son especialmente importantes por sus aportaciones para definir el análisis factorial moderno (THURSTONE, 1938).

Una de sus más representativas investigaciones, donde incluyó 240 sujetos y a los cuales les aplicó una batería de pruebas compuesta por 60 tests, tuvo como resultado la extracción de 12 factores; de los anteriores, Thurstone consideró sólo 9 factores como interpretables y a los que denominó Capacidades Mentales Primarias o CMP (THURSTONE, 1938); a continuación se presentan en la tabla N° 3:

Tabla N° 3. 9 factores que componen las CMP

SÍMBOLO	FACTOR	DEFINICIÓN
E	Espacial	Capacidad espacial, especialmente visual, como al hacer rotaciones mentales de figuras geométricas o al contar bloques ocultos.
P	Perceptual	Capacidad perceptual, sobre todo velocidad de percepción visual, como al explorar una página impresa para identificar letras o comparar columnas de números.
N	Numérica	Capacidad numérica, especialmente velocidad y precisión de cálculo.
V	Verbal	Capacidad verbal, incluidas analogías verbales, antónimos, comprensión de lectura.
M	Memoria	Capacidad de memorización de corto plazo, como en el aprendizaje de asociaciones pareadas.
Pa	Palabras	Facilidad de palabra, sobre todo manejo de palabras aisladas, como en una prueba de palabras desordenadas o en una prueba de afluencia.
I	Inducción	Capacidad para hallar una regla o principio para resolver un problema, como en series de números, clasificación de figuras o analogías de patrones.
R	Razonamiento	Capacidad de razonamiento, en especial al enfrentar un problema con solución cerrada como en el razonamiento aritmético.
D	Deducción	Factor definido en forma deficiente en varias pruebas que exige la aplicación de una regla.

Nota. Esta tabla fue tomada de las CMP de Thurstone (HOGAN, 2004, pág. 199).

A partir de estos 9 factores, Thurstone desarrolló una prueba que tuvo varias versiones y en las que nunca se incluyeron los nueve factores que componían las CMP de Thurstone, sino que por

lo general comprendían sólo cinco factores. No obstante, los factores de inducción, razonamiento y deducción se presentaban como un solo factor; mientras que los factores espacial, numérico, verbal y razonamiento se encontraban presentes en casi todas las versiones de la prueba CMP. Los factores perceptual, memoria y facilidad de palabra sólo aparecieron en algunas versiones (HOGAN, 2004).

Como se ha visto, los trabajos de Spearman y Thurstone componen dos de las teorías clásicas más influyentes en la historia de la inteligencia. Aunque las teorías de estos dos autores se contraponían la una con la otra, Spearman terminaría por reconocer la existencia de factores grupales que podían explicar capacidades mentales especiales, mientras que Thurstone también reconoció la existencia del factor g como factor de primer orden (BRODY & BRODY, 1976).

Como ya se mencionó, tanto Spearman como Thurstone realizaron contribuciones al análisis factorial que explicaron en gran medida sus teorías, sin embargo, sus contribuciones también dieron origen al desarrollo de nuevas metodologías y, con ellas, la aparición de los modelos jerárquicos, que a continuación se presentan.

Uno de los modelos jerárquicos que produjo bastante influencia en el campo del estudio de la inteligencia fue la teoría propuesta por Cattell: Teoría de la inteligencia fluida y cristalizada. Esta teoría se basa en dos conceptos principales: Fg y Cg. Fg se refiere a la inteligencia fluida general y tiene que ver con el potencial inherente que tiene un individuo para enfrentarse a situaciones novedosas y, por lo tanto, cuenta con un componente adaptativo; además no es verbal y no está influida (al menos relativamente) por la cultura. En contraste, Cg se refiere a la inteligencia cristalizada general y se puede entender como el conjunto de aprendizajes de un individuo; depende en gran medida de la cultura y representa tanto las habilidades mentales, como la información adquirida a partir de la experiencia, la práctica y la instrucción escolar. La inteligencia cristalizada es utilizada en situaciones cotidianas que demandan una respuesta aprendida (CATTELL & HORN, 1966).

Las principales diferencias entre Fg y Cg radican en lo que se conoce como habilidades potenciales y reales; en otras palabras, influencias hereditarias y ambientales sobre la inteligencia (CATTELL & HORN, 1966).

A diferencia de las teorías de Spearman y Thurstone, que se debaten entre un único factor general o varios factores específicos, la Teoría Fg-Cg contempla que tanto la inteligencia fluida general, como la inteligencia cristalizada general, están compuestas por diversos factores específicos. No obstante, Fg ha cobrado especial interés, debido a que se ha cuestionado la posibilidad de medir aquello que no guarda relación alguna con la cultura, la experiencia o la instrucción escolar (HOGAN, 2004).

Otra de las teorías jerárquicas más frecuentemente citadas es la propuesta por Philip Vernon, a quien se reconoce la contribución de llevar a cabo un esquema de organización de factores construido a partir del resumen de un gran número de investigaciones realizadas por otros investiga-

dores (HOGAN, 2004). En el esquema de Vernon, el factor g constituye la cima de la jerarquía y está compuesto por dos factores de grupo: factores verbal-educativo (v:ed) y espacial-mecánico (k:m) o también conocido como práctico-mecánico (p:m). En el siguiente nivel se encuentran los factores de grupo menores, donde el factor verbal-educativo se divide en los factores verbal y numérico, mientras que el factor espacial-mecánico se puede dividir en los factores de información mecánica, espacial y psicomotora. El siguiente nivel se conoce como factores de grupo específicos y se pueden obtener a partir del análisis de los factores de grupo menores (VERNON, 1960). A continuación, en la figura N° 1, se presenta el esquema de Vernon.

Figura N° 1

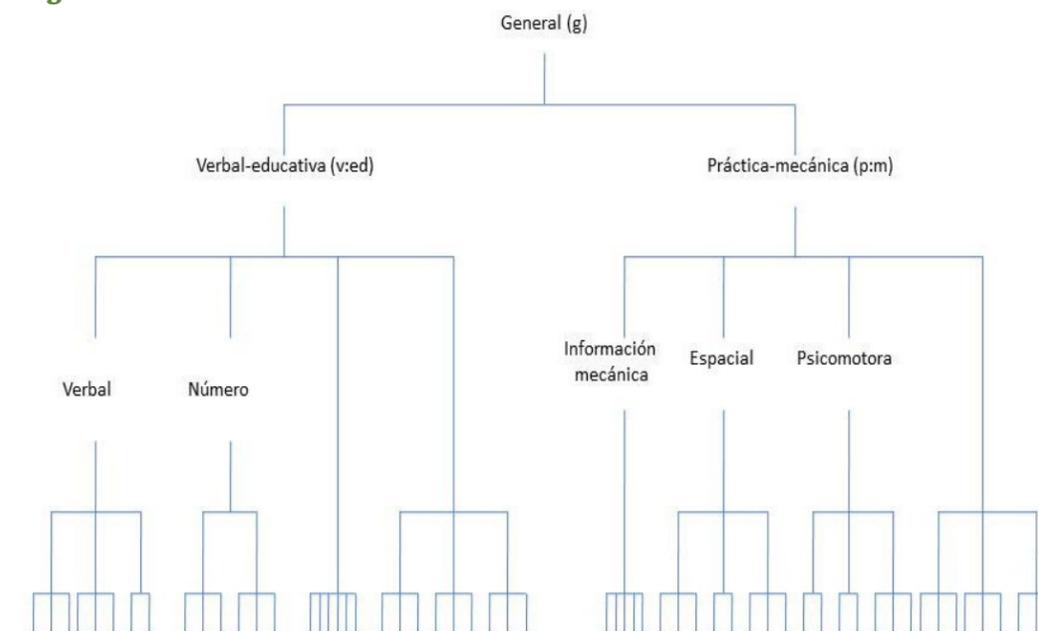


Figura N° 1: Modelo de organización jerárquica de factores de Vernon tomado de Anastasi & Urbina (1998, pág. 316).

La obra de John Carroll, *Human Cognitive Abilities: a Survey of Factor Analytical Studies*, constituye un resumen de numerosos estudios sobre las habilidades humanas en los cuales se utilizaron análisis factoriales. La conclusión de esta obra es el modelo jerárquico de Carroll, al cual se le denomina teoría de los tres estratos, el tercer estrato del esquema de Carroll es el factor g de Spearman y corresponde al nivel más alto; el segundo estrato incluye las Fg y Cg de la teoría de Cattell y otros factores que se relacionan en gran medida con las capacidades mentales primarias de Thurstone; por último, el primer estrato se encuentra integrado por diversas capacidades específicas, las cuales pueden llegar a estar relacionadas con más de un solo factor de grupo (CARROLL, 1993). En la figura N° 2 se puede observar el esquema de Carroll.

Figura N° 2

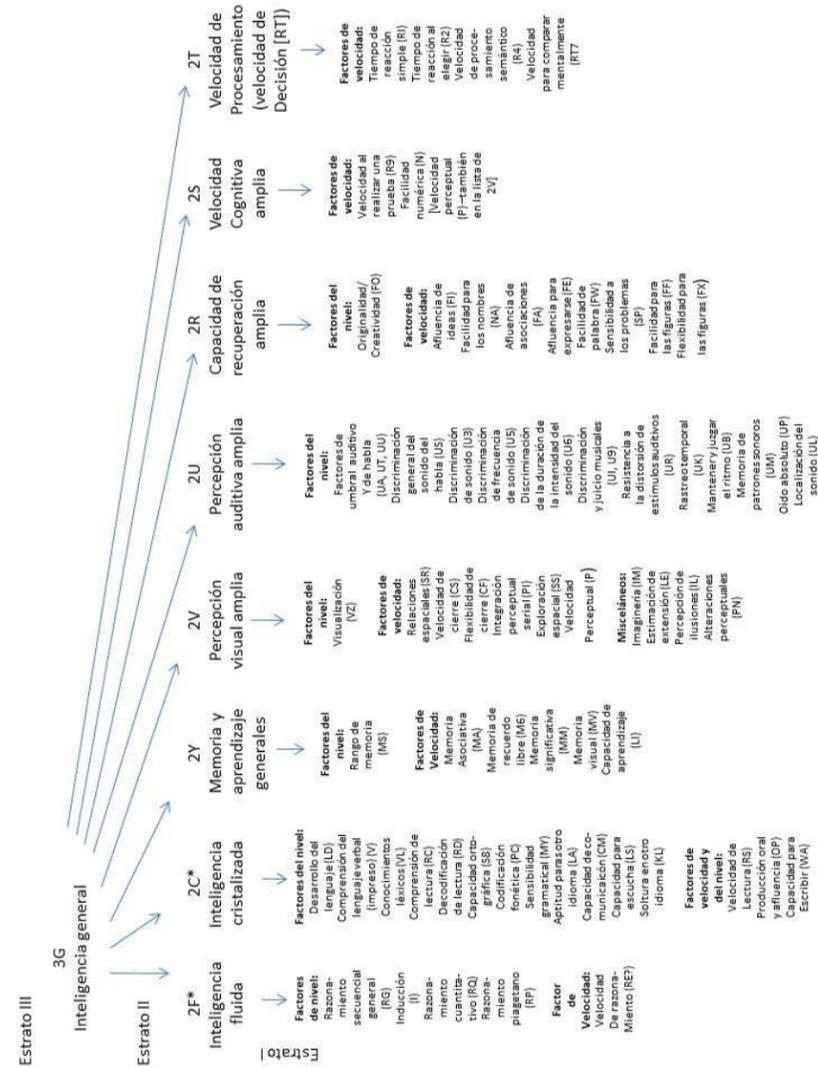


Figura N° 2: Modelo jerárquico de Carroll de los tres estratos tomado de Hogan (2004, pág. 203).

Tanto las teorías clásicas, como las jerárquicas que hasta el momento se han mencionado, pertenecen al campo de las teorías psicométricas de la inteligencia; sin embargo, existen otras teorías que parten de consideraciones teóricas distintas y con métodos de exploración que no implican el tratamiento estadísticos de datos obtenidos a través de la aplicación de pruebas psicométricas, algunos de estos modelos son las teorías del desarrollo.

Las teorías del desarrollo se encuentran cimentadas en el conocimiento del desarrollo mental humano a través del tiempo y de las experiencias. Básicamente, suelen presentar cuatro principales características:

1. Para describir el desarrollo mental se valen de una serie de etapas, que se diferencian entre sí porque cada etapa se compone de ciertas particularidades por las que se debe atravesar;
2. el paso por las etapas es secuencialmente invariable, es decir, todos los individuos atraviesan por ellas en el mismo orden, aunque probablemente no a la misma edad;
3. las etapas no funcionan de manera inversa, por lo tanto, cuando un individuo atraviesa por una de las etapas, la transición es irreversible;
4. la transición de una etapa a otra contempla un promedio de edad, sin embargo, pueden existir variaciones (HOGAN, 2004).

Jean Piaget es autor de una de las teorías del desarrollo con mayor influencia en el campo de la psicopedagogía y la psicología del desarrollo. El estudio del desarrollo cognitivo, Piaget lo llevó a cabo a través de una serie de observaciones, aplicó entrevistas y realizó algunas pruebas poco ortodoxas, pero que lo llevaron a plantear conceptos revolucionarios para su época, según los cuales: el desarrollo cognitivo cursa por cuatro etapas, el pensamiento del infante difiere cualitativamente en comparación con el adulto y la base de todo aprendizaje parte de esquemas mentales (Piaget, 1926,1952,1972). En la tabla N° 4 se presentan las etapas del desarrollo cognitivo de Piaget.

Tabla N° 4. Etapas del desarrollo cognitivo de Piaget

ETAPA Y RANGO DE EDADES	CARACTERÍSTICAS DEL PENSAMIENTO
Sensoriomotora: nacimiento a los 2 años	Los lactantes experimentan al mundo principalmente a través de sus sentidos y capacidades motoras, actúan como si un objeto dejara de existir cuando no está ante su vista, pero desarrollan permanencia de éste al final de este estadio.
Preoperacional: 2 a 6 años	No se desarrollan aún los conceptos de conservación, pero estos niños comprenden la idea de una relación funcional –por ejemplo, se tira de un cordel para abrir una cortina y entre más se tira de él, más se abre la cortina. También se desarrolla la capacidad para simbolizar mentalmente las cosas con palabras e imágenes.
Operaciones concretas: 7 a 12 años	Por lo común, los niños desarrollan conservación y demuestran capacidades limitadas de razonamiento lógico. Por ejemplo, se desarrolla el concepto de reversibilidad –el conocimiento de que una acción puede revertir o negar a otra.
Operaciones formales: 12 años en adelante	En general en esta etapa se desarrolla la solución sistemática de los problemas que se asocian con el pensamiento adulto. Existe mayor capacidad para generar hipótesis y someterlas a prueba.

Nota: Tabla de las etapas del desarrollo cognitivo de Piaget, tomada de Gregory (2001).

La teoría de Jean Piaget es bastante conocida entre los psicólogos, sin embargo, existen otras teorías del desarrollo. Kohlberg, por ejemplo, es autor de la teoría del desarrollo moral, un modelo que aborda un aspecto más concreto del razonamiento; éste se encuentra relacionado con la manera en la que el razonamiento distingue entre lo correcto e incorrecto (KOHLBERG, 1984).

La teoría de Kohlberg se basa en tres niveles: preconventional, convencional y posconventional; cada nivel se constituye por dos estadios (GRIMALDO, 2007). Aunque la teoría de Kohlberg se interesa por el juicio moral y no meramente en la inteligencia, es importante mencionarla debido a que esta teoría aborda un aspecto particular de la cognición humana y, por tanto, no deja de tener relación con ésta. A continuación, en la tabla N° 5 se mencionan y explican brevemente los niveles y estadios de la teoría del desarrollo moral de Kohlberg.

Tabla N° 5. Teoría del desarrollo moral de Kohlberg

NIVEL	ESTADIOS	CARACTERÍSTICAS DEL NIVEL
Preconventional	Castigo y orientación a la obediencia	Las acciones del sujeto están determinadas por la consecuencia del castigo, es decir, una mala conducta es aquella que trae consigo un castigo.
	La orientación instrumental relativista	
Convencional	La orientación de la concordia interpersonal	El comportamiento aceptado socialmente y legalmente es el considerado correcto, porque en esta etapa se busca la aprobación de los demás.
	La orientación de la ley y el orden	
Posconventional	La orientación legalista del contrato	El sujeto elige de manera consciente los principios éticos por los que se ha de regir. Se buscan principios que van más allá de la autoridad social.
	La orientación de los principios éticos universales	

Nota: Tabla de la teoría del desarrollo moral de Kohlberg, realizada con base en Grimaldo (2007).

Otros modelos teóricos sobre la inteligencia humana son las teorías del procesamiento de la información. En este caso, las teorías no se interesan por el estudio del cúmulo de conocimientos adquiridos por un sujeto, sino por la manera en la que ha sido procesado ese conocimiento. Por lo general, estos modelos utilizan las computadoras como analogía para explicar el procesamiento de información en humanos; además, se considera que guardan estrecha relación con los modelos biológicos de la inteligencia, debido a que están interesados en conocer el funcionamiento del cerebro y a que las redes neuronales suelen compararse con el funcionamiento de un ordenador (HOGAN, 2004).

Es importante mencionar que gran parte de las teorías del procesamiento de la información utilizan en sus estudios un elemento llamado tarea cognitiva elemental (TCE), la cual consiste en

llevar a cabo una actividad relativamente sencilla y que demanda algún tipo de procesamiento cognitivo. La razón por la cual se utilizan estas tareas, es que se cree que están libres de influencias tales como la instrucción escolar y la cultura y, por lo tanto, suponen ser indicadores exentos de sesgo. Algunas de estas TCE's son: el tiempo de reacción simple, que consiste en presionar un botón en cuanto el sujeto observa la aparición de una luz; el tiempo de reacción de elección, donde se lleva una tarea similar a la anterior, sólo que en este caso se le pide al sujeto presionar el interruptor únicamente cuando encienda determinada luz y no otra; identificación de letras, que consiste en reconocer si los pares de letras mostrados son semejantes o si los pares de letras designan la misma letra; la verificación semántica, donde se utilizan las letras A, B y C en distintos ordenes y se incluye un enunciado que el sujeto debe determinar si es correcto o incorrecto; por último, el tiempo de inspección, donde al sujeto se le presentan dos líneas paralelas que aparecen de manera intermitente y éste debe indicar si son del mismo tamaño (HOGAN, 2004).

Uno de los modelos de procesamiento de la información es el propuesto por Arthur Jensen, el cual parte del desempeño observado en las TCE para llegar a g. Jensen centra su interés en la inteligencia fluida, dado que ésta explica el desempeño en las TCE. De manera sencilla, la teoría de Jensen se puede explicar de la siguiente manera: el desempeño en las tareas (T1, T2, T3...) está determinado por varios procesos de información (P1, P2, P3...), que a su vez dependen tanto de un factor general de procesamiento de información (PI), como de un factor general de tiempo de reacción (TR); la inteligencia fluida determina el factor general de procesamiento de información, mas no el factor general de tiempo de reacción. Finalmente, la inteligencia fluida también explica el desempeño en las pruebas psicométricas (JENSEN, 1998). En la figura N° 3 se puede observar un esquema de la teoría de Jensen.

Figura N° 3

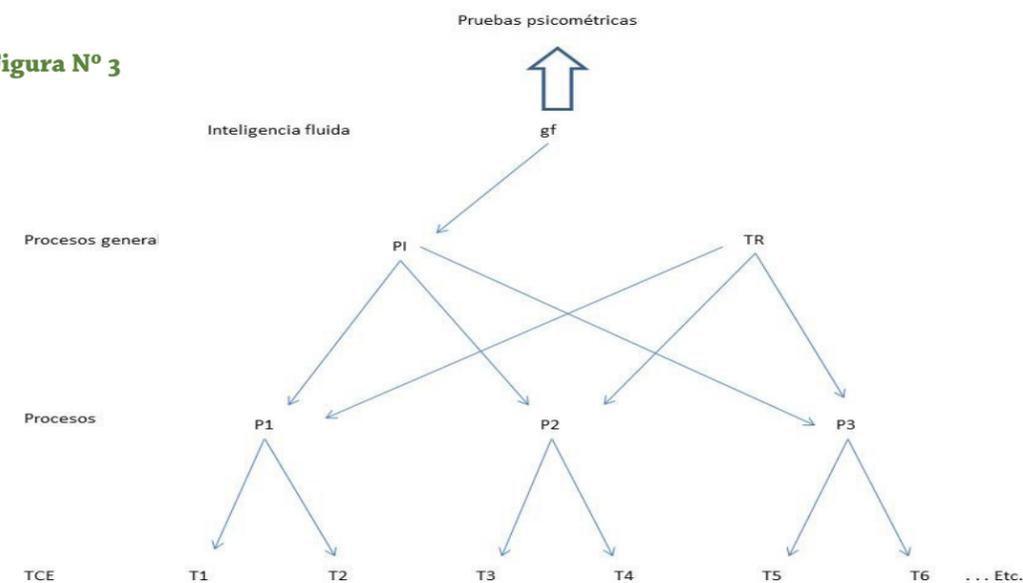


Figura N° 3: Teoría del procesamiento de la información tomado de Jensen (1998, pág. 207).

La teoría PASS (*Planning, Attention-arousal, Simultaneous and Successive Processing*) se encuentra dentro de los modelos de procesamiento de la información, sin embargo, esta teoría utiliza planteamientos de tipo biológico que hacen referencia a distintas áreas encefálicas. El modelo parte de tres unidades funcionales principales: atención o excitación, la cual requiere de un estado de vigilia y la voluntad de prestar atención; procesamiento de la información, que depende de dos tipos de procesamiento que son el secuencial (encargado de la información que se relaciona entre sí o que está ordenada de manera temporal) y el simultáneo (puede encargarse de todo el material disponible de manera simultánea); finalmente, la planeación que implica las funciones ejecutivas, tales como supervisión, evaluación, etc (DAS, NAGLIERI, & KIRBY, 1994). El esquema de la teoría PASS se puede observar en la figura N° 4.

Figura N° 4

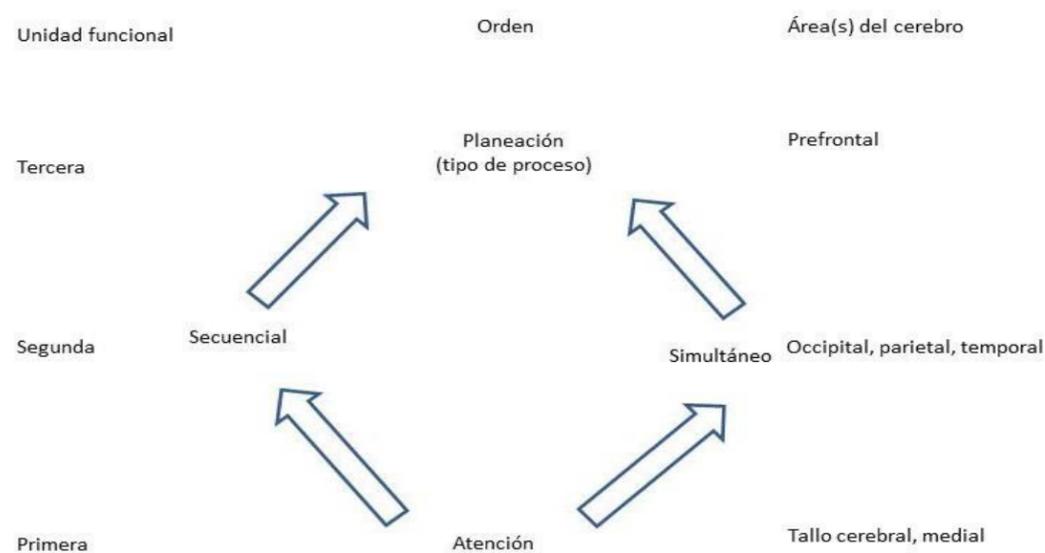


Figura N° 4: Esquema de la teoría PASS tomado de Hogan (2004, pág. 209).

Una teoría que se puede clasificar dentro de los modelos biológicos (debido al uso frecuente de términos relacionados con el funcionamiento cerebral), es la teoría de Howard Gardner. Su teoría es mejor conocida como teoría de las inteligencias múltiples o Teoría IM y plantea la existencia de diversas inteligencias que actúan, parcialmente, de manera independiente. En un primer momento, Gardner propuso siete inteligencias: lingüística, musical, lógico-matemática, espacial, cinestésico-corporal, intrapersonal e interpersonal (GARDNER, 1983, 1993). No obstante, las inteligencias lingüística, lógico-matemática y espacial han sido ampliamente reconocidas en otras teorías, por lo que los conceptos restantes (musical, cinestésico-corporal e intrapersonal/interpersonal) son la parte novedosa de la teoría de Gardner.

Más recientemente, Gardner (1998) introdujo tres nuevos tipos de inteligencia: naturalista, espiritual y existencial. En la tabla N° 6 se ofrece una breve descripción de las inteligencias musical, cinestésico-corporal, intrapersonal/interpersonal, naturalista, espiritual y existencial.

Tabla N° 6. Inteligencias múltiples de Gardner

INTELIGENCIAS PROPUESTAS POR GARDNER	DESCRIPCIÓN
Musical	Se refiere a las personas que cuentan con una habilidad innata para aprender a tocar un instrumento, ejecutar y/o componer piezas musicales.
Cinestésico-corporal	Son habilidades relacionadas con destrezas motrices.
Intrapersonal-interpersonal	Tiene que ver con la consciencia del plano emocional interno y el plano emocional de los otros.
Naturalista	Se evidencia en personas con conocimiento acerca del mundo natural.
Espiritual	Interés por cuestiones espirituales, relacionadas con el cosmos y el desarrollo interior.
Existencial	Interés por cuestiones relativas al significado de la existencia.

Nota: Descripción de los conceptos introducidos por Gardner con base en Gardner (1983, 1993) y Gardner (1998).

Uno de los criterios utilizado por Gardner para explicar la posibilidad de que existan diversas inteligencias funcionando con independencia relativa, son los llamados Autistas sabios; individuos que padecen discapacidad intelectual, mas son capaces de desarrollar de manera excepcional una determinada habilidad. De hecho este criterio es el que más fuerza le ha dado a la teoría IM. Sin embargo, la evidencia empírica acerca del tema es escasa y, por lo tanto, la teoría de las inteligencias múltiples no cuenta con un sustento sólido que le otorgue peso en el estudio de la inteligencia. (GREGORY, 2001).

La teoría de la inteligencia más extensa que se ha planteado hasta la actualidad, es la propuesta por Robert J. Sternberg y se conoce como teoría triárquica de la inteligencia. Se le denomina triárquica porque se basa en tres conceptos sobre la inteligencia: componencial, experiencial y contextual, que a su vez se constituyen de dos o más componentes. El elemento componencial es aquella en la que se incluyen los procesos mentales, que están integrados por los metacomponentes o también llamados procesos ejecutivos (encargados de evaluar la magnitud de un determinado problema que se presenta, planear una estrategia para darle solución y supervisar la ejecución de la estrategia), por los componentes de ejecución (son los que llevan a cabo la resolución real de un problema) y por los componentes de adquisición del conocimiento (incluyen los mecanismos

a través de los cuales es posible el aprendizaje: codificación, combinación y comparación). Con respecto al elemento experiencial, ésta tiene que ver con el enfrentamiento de actividades novedosas y aquellas con las que se está familiarizado hasta el punto de la automatización. Por último, el elemento contextual es la que provee tres maneras de enfrentarse al ambiente: la adaptación (desarrollo de capacidades demandadas por el entorno), moldeamiento (modificación del entorno para que éste se adapte mejor a las habilidades propias) y selección (cambiar de entorno en beneficio de las habilidades potenciales) (STERNBERG, 1986). En la figura N° 5 se puede apreciar un esquema de la teoría triárquica de Sternberg.

Figura N° 5

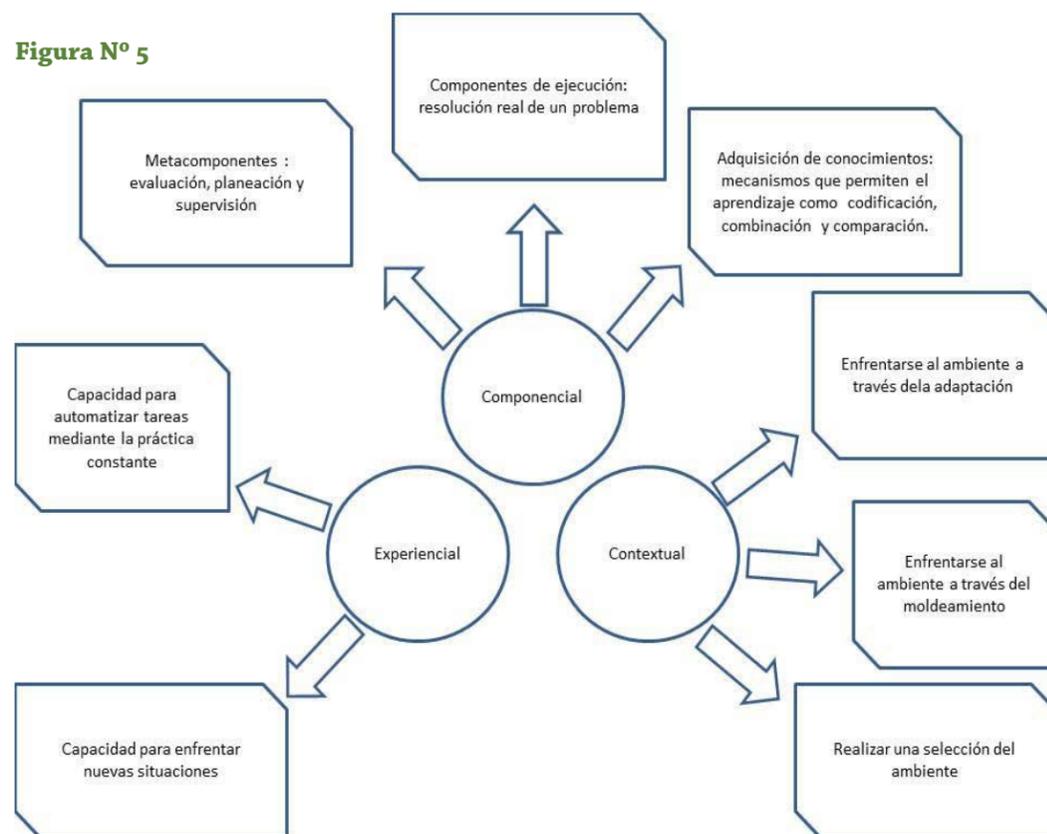


Figura N° 5: Esquema de la teoría triárquica de Sternberg adaptado de Sternberg (1986).

Las teorías de la inteligencia mencionadas hasta el momento no son las únicas que existen, pero representan algunas de las posturas teóricas más influyentes en el estudio de la inteligencia. No obstante, parte del interés del presente estudio se centra en los conceptos emanados de las teorías psicométricas; el CI y el índice de heredabilidad son especialmente importantes, dado que las teorías psicométricas utilizan estos conceptos para representar un promedio de la inteligencia y el grado en que se explica a través de la herencia, respectivamente.

2.3 INTELIGENCIA: HEREDABILIDAD, MODIFICABILIDAD Y LA INTERPRETACIÓN DEL CI

El concepto de heredabilidad, que se relaciona con los rasgos o características psicológicas heredadas por los progenitores, proviene de los estudios de la genética conductual, los cuales se han centrado en la investigación de características como el CI, la extroversión, el neuroticismo, entre otros. La heredabilidad se puede definir como (OLIVA, 1997):

Un parámetro estadístico que da información sobre la relación entre la varianza genética y la varianza fenotípica de una determinada población. Nos indicará hasta qué punto las diferencias genéticas existentes entre los individuos son responsables de las diferencias que presentan en una característica o rasgo conductual. La heredabilidad es, por lo tanto, un parámetro poblacional que no es aplicable a un individuo aislado, no nos dice nada acerca de los determinantes de las características de un sujeto concreto. Tampoco nos da información sobre las diferencias entre grupos: hombres y mujeres, blancos y de color, etc. Por ejemplo, si decimos que el CI tiene una heredabilidad de 0.6, ello significa que el 60% de las variaciones en CI de una determinada población, en un momento determinado, se deben a diferencias genéticas. Esto no quiere decir que en un sujeto que presente un CI de 110, el 60% de su inteligencia esté determinada genéticamente y el ambiente explique el 40% restante. Sino que si la media poblacional del CI es de 100, seis de los diez puntos que separan a este sujeto de la media se deberían a influencias genéticas y cuatro a efectos ambientales. De los 100 puntos restantes no podemos decir nada (págs. 24-25).

La creencia de que la predisposición genética es un determinante en el desarrollo de ciertas conductas es inexacta; sin embargo, los genes pueden llegar a hacer más o menos probable la aparición de una conducta (KAIL & CAVANAUGH, 2006).

Las investigaciones cuantitativas sobre la herencia genética se basan en estudios sobre gemelos homocigóticos (que comparten el 100% de los genes) y gemelos heterocigóticos (que comparten el 50% de los genes), ya sea que vivan juntos o separados. La razón es porque las correlaciones entre los puntajes obtenidos en las pruebas determinaron el grado en el que los genes influyen, por ejemplo, en el CI. Las correlaciones pueden ir de 0 a 1, donde cero representaría ausencia total de influencia y uno el 100% de influencia (Lejarraga, 2010).

La heredabilidad de la inteligencia es uno de los rasgos más estudiados y se ha determinado que la evolución en las puntuaciones de CI se relacionan con la herencia (WILSON, 1983). En las investigaciones realizadas con gemelos monocigóticos, se ha encontrado que las puntuaciones de CI son proclives a elevarse con el transcurso del tiempo, es decir, si uno de los gemelos obtiene un CI más alto que el otro, con la edad el CI de ambos incrementará su correlación; al contrario de lo que

pasa con gemelos dicigóticos. Por otro lado, en el caso de los hermanos adoptivos las diferencias se hacen más evidentes conforme la edad se eleva, un efecto que pone de manifiesto la influencia genética (OLIVA, 1997).

De acuerdo con Yela (1996), a través de un número importante de estudios se ha encontrado que el parentesco es el que determina las diferencias de correlación entre los índices de heredabilidad:

Las correlaciones medias de decenas de miles de casos son: cero, entre personas sin parentesco; 0.20 entre primeros hermanos; 0.50 entre padres e hijos y entre hermanos; 0.65 entre gemelos dicigóticos, y 0.90 entre gemelos monocigóticos. La concordancia entre las relaciones obtenidas y las pronosticadas por el modelo poligénico es, en el promedio, casi perfecta (pág. 193).

Algunas críticas realizadas al estudio de la heredabilidad son:

1. se ha comprobado que el entorno compartido por gemelos heterocigóticos difiere más en comparación con el ambiente compartido por los gemelos homocigóticos, y
2. los gemelos que han sido separados y enviados a hogares distintos, de hecho comparten entornos similares porque los sistemas de adopción buscan hogares para niños de acuerdo a las características de éstos y de las familias adoptivas.

Esto último sugiere que los sujetos, utilizados en los estudios de heredabilidad, en realidad no fueron asignados a hogares al azar y probablemente compartan entornos muy similares (ANASTASI & URBINA, 1998).

Los genes no son los únicos que influyen en el desarrollo de habilidades o rasgos. Los individuos también están en contacto con el medio que les rodea y son influenciados por el mismo; por tanto, la aparición de rasgos depende de la interacción entre genes y ambiente. Kail & Cavanaugh (2006) lo ejemplifican con el caso de personas que cuentan con la habilidad para realizar varias canastas con facilidad. Estos sujetos han desarrollado una altura que les facilita esa actividad; sin embargo, estos mismos sujetos podrían desarrollar una altura superior a la del promedio sin la habilidad para encestar con facilidad y esto se debe a que el ambiente influye de manera importante en el desarrollo de características.

Un hecho indiscutible es que, a pesar de se pudiera comprobar que la inteligencia humana está influida por un alto índice de heredabilidad, el ambiente seguiría siendo un factor relevante en el desarrollo de ésta. Además, el hecho de que un rasgo psicológico, como el CI, se encuentre influenciado por los factores genéticos, no significa que éste sea constante a través del tiempo e inmodificable; por el contrario, existen datos empíricos que apoyan la idea de que las puntuaciones de CI pueden sufrir cambios mediante la intervención del entorno. Estas intervenciones del entorno pueden aparecer tanto en etapas tempranas como en etapas posteriores, ya sea de manera incidental o planeada y producir cambios, que pueden ser positivos o negativos; es decir, así como el CI

puede incrementar por cambios fortuitos, también puede sufrir un decremento debido a cambios de carácter nocivo (ANASTASI & URBINA, 1998).

De acuerdo con algunos autores (SCARR, 1992,1993; SACARR & MCCARTNEY, 1983), existen tres diferentes clases de relación entre los genes y el ambiente: relación pasiva entre genes y ambiente, que consiste en un ambiente creado por los progenitores, estrechamente ligado con los genes que se transmiten; relación evocativa entre genes y ambiente, la cual tiene que ver con la manera en la que los sujetos responden a su entorno de acuerdo a sus genes, y relación activa entre genes y ambiente, donde el sujeto busca de manera consciente entornos vinculados con su herencia genética. Un ejemplo de cómo funcionarían estas etapas, en el caso de la inteligencia, se puede apreciar en la figura N° 7.

Figura N° 6

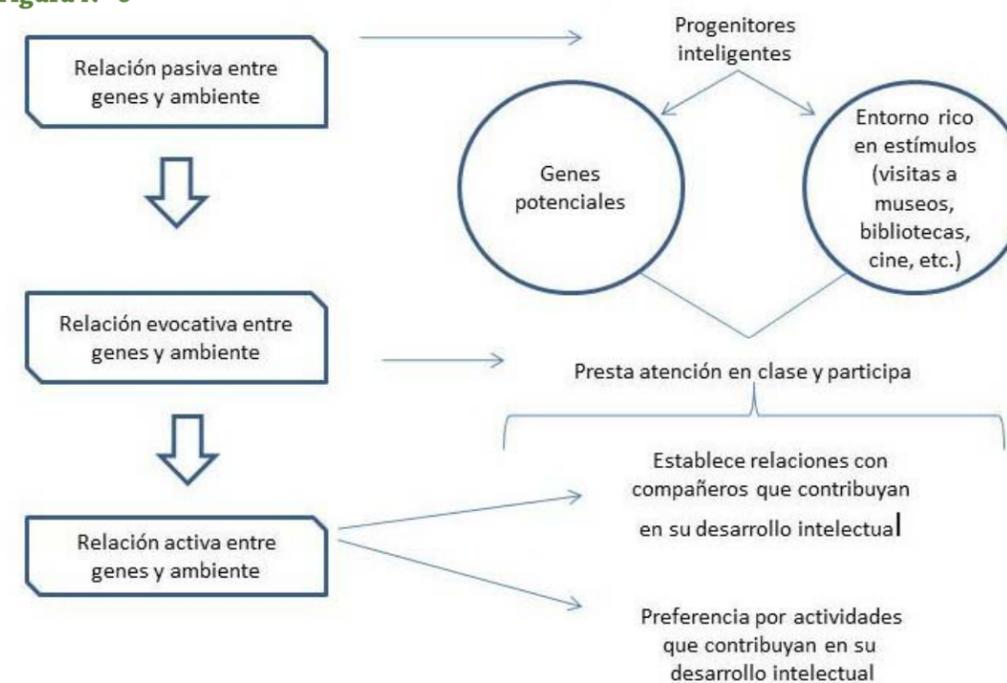


Figura N° 6: Esquema de los tres tipos de relación entre genes y ambiente, con base en Kail & Cavanaugh (2006).

En un caso similar al ejemplo de la figura 6, es muy probable que el individuo sufra cambios positivos en relación a la inteligencia. Si a este sujeto se le aplicaran mediciones en diferentes etapas de su vida, seguramente se podrían apreciar las contribuciones del entorno favorable a través de su CI.

Es importante hacer hincapié en que los términos de inteligencia y CI, que en muchas ocasiones se utilizan como sinónimos, difieren entre sí. La interpretación adecuada del CI es importante por sus implicaciones. En primer lugar, cuando se habla de CI, se hace referencia al puntaje

obtenido por un sujeto en una prueba y este puntaje, a su vez, se puede interpretar como el grado en el que una capacidad se hace presente en un sujeto en un momento determinado, tomando en cuenta las normas de edad contempladas por el test administrado. A este respecto, se debe agregar que los puntajes obtenidos en las pruebas de inteligencia, se deben analizar de acuerdo con el test utilizado, debido a que los instrumentos varían tanto en contenido como en las normas para la interpretación de resultados (ANASTASI & URBINA, 1998).

Diversos instrumentos psicométricos tradicionales, utilizados para medir la inteligencia, han sido desarrollados para medir aspectos relacionados con el desempeño escolar, en consecuencia, estos tests se pueden valorar como instrumentos para medir el rendimiento académico y, por lo tanto, en este tipo de pruebas, el CI se puede interpretar como el nivel de aprovechamiento escolar y puede servir para realizar predicciones sobre éste (ANASTASI & URBINA, 1998).

2.4 TESTS DE INTELIGENCIA MÁS UTILIZADOS

La cantidad de instrumentos utilizados en la actualidad para medir inteligencia es muy extensa. La investigación sobre inteligencia humana ha propiciado y sigue propiciando la creación de nuevos instrumentos. No obstante, existen pruebas de inteligencia que se han popularizado tanto que se aplican en diversas poblaciones.

Antes de presentar un listado de pruebas psicométricas de inteligencia, utilizadas con frecuencia en diversos países, es necesario conocer algunas características generales sobre las pruebas de inteligencia de aplicación individual y grupal, debido a que los diseños de estas dos variantes de instrumentos difieren por sus propósitos y circunstancias en las que se aplican. En la tabla N° 7 se presentan las características de las pruebas de inteligencia para aplicación individual y grupal.

Tabla N° 7. Características de las pruebas de inteligencia de aplicación individual y grupal

PRUEBAS DE APLICACIÓN INDIVIDUAL	PRUEBAS DE APLICACIÓN GRUPAL
1) Los reactivos se pueden presentar como preguntas o estímulos, a los cuales el examinado responde de manera escrita, oral o a través de la manipulación de material.	1) El número de participantes en una aplicación es relativo. No existen reglas que definan el número de participantes en un grupo de aplicación.
2) Requieren del conocimiento exhaustivo de sus reactivos, material y normas de calificación.	2) Los reactivos son de opción múltiple.
3) Por lo general cuentan con reglas que determinan los reactivos adecuados para distintos rangos de edad y reglas para suspender o continuar con la prueba.	3) El contenido es similar al de las pruebas individuales.

4) Para su aplicación, es necesario que el examinador propicie un ambiente agradable para el examinado.	4) No utilizan reactivos que requieran la manipulación de material.
5) Los reactivos que solicitan respuestas libres son más comunes en este tipo de instrumentos.	5) Algunas pruebas utilizan límite de tiempo.
6) En ocasiones es necesario que el examinador califique los reactivos mientras el examinado los resuelve.	6) El tiempo promedio de aplicación es de una hora.
7) El tiempo promedio para completar una prueba es de una hora.	7) Arrojan un puntaje general.
8) La observación del examinado se convierte en un factor relevante dado que en este tipo de pruebas suele hacerse observaciones en los reportes.	8) Su utilización es más frecuente en la investigación.

Nota: Tabla sobre las principales características de las pruebas psicológicas para medir inteligencia de aplicación individual y grupal con base en Hogan (2004, págs. 220-258).

A continuación se presenta una lista de pruebas de inteligencia comúnmente utilizadas en múltiples países. Sin embargo, debido a la extensión de la lista y a que el interés del presente estudio no se centra en la compilación de instrumentos para medir inteligencia, únicamente se proporcionan los nombres de los instrumentos y de sus respectivos autores. En la tabla N° 8 se puede observar un listado de instrumentos para medir inteligencia.

Tabla N° 8. Lista de pruebas de inteligencia

TEST	AUTOR
Progressive Matrices for Children (Coloured)	J.C. Raven
Standard Progressive Matrices	J.C. Raven
Wechsler Memory Scale Form I & II	David Wechsler
Culture Fair (or Free) Intelligence Test Scale-1	R.B. Cattell
Culture Fair Intelligence Test Scale-2 (Form A & B)	R.B. Cattell
Culture Fair Intelligence Tests Scale-2 & 3 (Form A & B)	R.B. Cattell
Henmon-Nelson Tests of Mental Ability	Tom A. Lamke & M.J. Nelson
Kahn Test of Symbol Arrangement	T.C. Kahn
Kahn Intelligence Test	T.C. Kahn
Wechsler Adult Intelligence Scale (WAIS)	D. Wechsler
Wechsler Intelligence Scale for Children (WISC)	D. Wechsler
Wechsler Intelligence Scale for Children (WPPSI)	D. Wechsler
Columbia Mental Maturity Scale	B.B. Burgemeister
Raven's Progressive matrices and Vocabulary Scales	J.C. Raven
Revised Beta Examination	C.E. Kellogg & N.W. Morton

Wesman Personnel Classification Test Form A, B, & C	Alexander G. Wesman
Daneshill Intelligence test	R. Macdonald
Advanced Progressive Matrices Set I, II	J.C. Raven
The Chicago Non-Verbal Examination	A.W. Brown
Goldstein-Schreerer Tests of abstract and concrete Behaviour	K. Goldstein
Intelligence Testing for Visually Impaired Persons	C.J.V. Kolk
Slosson Intelligence Test and Oral Reading Test for Children and Adults	R.L. Slosson
Group Embedded Figures Test	P.K. Oltman
The D48 Test (A Non-Verbal Intelligence Test)	J.D. Black
Pictorial Test of Intelligence	J.L. French
Canadian Intelligence Examination	H. Amoss
Otis Self-Administering test of Mental Ability	A.S. Otis
Social Intelligence Test	F.A. Moss
A Point Scale of Performance test	G. Arthur
Non-Language Multi-Mental test	E.L. Terman
SRA Verbal & Non-Verbal Forms	R.N. McMurry
Moray House Adult Intelligence Test (Verbal Reasoning)	M.H.T.
The Merrill-Palmer Scale of Mental Tests	R. Stutsman
California test of Mental Maturity	E.T. Sullivan
Leiter Adult Intelligence Scale	R.G. Leiter
Pintner-Paterson Performance Tests	H. Pintner
Chicago Non-Verbal Examination	A.Q. Brown
Test of Mental Efficiency	B. Babook
Modified Alpha Examination	F.L. Wells
Engle Psychology Test	T.L. Engle
Revised Beta Examination	C.E. Kellogg
Essential Junior Intelligence Test	F.J. Schonell
Metropolitan Readiness Tests	G.H. Hildreth, L. Nellie M.A. Griffiths
Thurstone Test of Mental Alertness	L.L. Thurstone
Ohio State University Psychological Test	H.A. Toops
The Henmon-Nelson Tests of Mental Ability	V.A.C. Womon
Minnesota Pre-School Scale	E.T.S
Picture Intelligence Test	M.S. Mellone
Army General Classification Test	S.R.A.
The Simple Group Intelligence Scale	C.A. Richardson
Revised Alpha Examination Form	F.L. Wells
Terman McNemar Test of Mental ability	L.M. Terman
Detroit Advanced Intelligence Test	H.J. Baker
ACER Advanced Test B-40	ACER

General Ability Test (Intermediate) D.Test)	ACER
General Ability Test (Advanced M)	ACER
ACER Intermediate Test-D	ACER
ACER Intermediate Test-A	ACER
Otis Quick Scoring Mental Ability Test	A.S. Otis
Group Test of Intelligence	T.P. Tomilson
The Quick Picture Vocabulary Test (Monograph Supplement I-VII)	R.B. Ammons & C.H. Ammons
Elementary Reading General Ability Test	W. Reed
A Measure of Intelligence	C. Spearman
Methods and Experiments in Mental Tests	C.A. Richardson
Goodenough Intelligence Test	F.L. Goodenough
Pintner General Ability Test-Form A	R. Pintner
Pintner General Ability Test-Form L	R. Pintner
Kuhlmann-Anderson Test	F. Kuhlmann & R.G. Anderson
A Culture-Free Test	R.B. Cattell
Lorge-Thorndike Intelligence Test	I. Lorge
Holzinger-Crowder Uni-Factor Test	K.J. Holzinger
Test of General Intelligence or Problem-Solving Ability	A. Davis & Kenneth Eells
Simplex GNV2 Intelligence Test	C.A. Richardson
Simplex Jr. 'A' Intelligence Test	C.A. Richardson
Simplex Jr. 'A' Intelligence Test	C.A. Richardson
ACER Junior 'A' Test	ACER
Northumberland Mental Test	G.H. Thomson
Simple GNV Intelligence Test	C.A. Richardson
General test N and Advanced Test N	ACER
Group Test of Intelligence	ACER
Otis Self Administering Test-Form A, B, C, D	ACER
Intermediate Test C	ACER
ACER Junior B Test	ACER
Army Alpha Examination	E.C. Bregman
Carnegie Mental Ability Test	G.U. Cleeton
Pintner General Ability Test Form B	R. Pintner
Cornell-Cox Performance Ability Scale	E.L. Cornell
H.T.P. Test	Issac Jalles
Public School Primary Intelligence Test	H.J. Baker
Series of Emergency Scales	G.H. Kent & A. Vermont
The Measurement of Efficiency of Mental Functioning	C.H. Staelting Co., Chicago
The Merrill Palmer Scale of Mental Tests	C.H. Staelting Co., Chicago
Leiter International Performance Scale	R. C. Leiter
The Leiter Adult Intelligence	C.H. Staltin

Mental Status Checklist for Adolescents	E. H. Dougherty; J.A.Schinka
Mental Status Checklist for Children	E. H. Dougherty; J.A.Schinka
Naglieri Nonverbal Ability Test Indian Administration	Jack A.Naglieri
The MIDAS (Multiple Intelligence Developmental Assessment Scale)	C.Branton Shearer
Wechsler Memory Scale (Third edition)	David Wechsler
Kaufman Brief Intelligence Test (KBIT2) Second Edition	A. S.Kaufman & N. L.Kaufman

Nota: con excepción de algunas repeticiones encontradas, esta lista de pruebas fue tomada de la lista presentada en Department of Educational Psychology and Foundations of Education (2013, págs. 58-64).

Como se observa en la tabla N° 8, las listas de pruebas para medir inteligencia son bastante numerosas y por esa razón, si se pretende indagar con más detalle sobre las compilaciones, materiales y normas de calificación de diversos instrumentos de medición, el lector puede consultar las siguientes fuentes: González (2007), López, et al (2010), *Department of Educational Psychology and Foundations of Education* (2013), *Pensions Section* (2013).

Uno de los instrumentos de medición de inteligencia más frecuentemente utilizado, como se puede corroborar en las fuentes anteriormente citadas, es el test MPR y es por ello que el presente estudio se interesa por esta prueba. El capítulo siguiente está dedicado exclusivamente al test MPR y se incluyen estudios acerca de sus propiedades psicométricas, así como algunos ejemplos de estudios en los cuales se utilizó esta prueba como instrumento de medición.





• CAPÍTULO 3

Matrices progresivas de Raven

El test MPR cuenta con tres variantes dirigidas a personas de distintas edades y características: las matrices progresivas escala coloreada (MPC), dirigidas a niños con edades entre los 5 y 11 años (aproximadamente) o para adultos con discapacidad intelectual; las matrices progresivas escala general (MPG), para personas de casi todas las edades, exceptuando niños muy pequeños, y las matrices progresivas escala avanzada (MPA), diseñadas para determinar de manera rápida si una persona es torpe, término medio o brillante y para medir la eficiencia mental de personas superiores al promedio (RAVEN, 2004).

El test MPR es un instrumento que, a través del uso de figuras para representar un problema, intenta evitar cualquier sesgo cultural o influencia de la educación escolar en su aplicación. A diferencia de otros tests, las MPR no utilizan el lenguaje para plantear un problema, sino que intentan valerse de instrucciones sencillas que se dan de manera verbal antes de la aplicación; así mismo, se intenta relacionar al examinado con la forma de trabajo, primordialmente, a través de la solución de ciertos problemas iniciales que se utilizan como ejemplos e indicadores para que el examinador compruebe si se ha comprendido la esencia de los problemas.

Las MPR son un instrumento desarrollado hace más de medio siglo; desde su publicación hasta la actualidad, el instrumento ha sufrido cambios en algunos aspectos, mientras que en otros se ha mantenido. Conocer los antecedentes de las MPR constituye el paso principal para la comprensión completa del test, es por ello que el siguiente apartado aborda los antecedentes de este instrumento.

3.1 ANTECEDENTES DE LAS MATRICES PROGRESIVAS DE RAVEN

La primer escala desarrollada de las MPR fueron las MPG. Ésta se publicó en 1938 y su finalidad era servir como un instrumento capaz de englobar una variedad extensa de habilidades mentales, lo cual no significa que el test mida otras habilidades a parte de la inteligencia, sino que su aplicación se puede llevar a cabo en sujetos con distintas edades, con diferentes niveles educativos, diferente nacionalidad y diferentes características físicas. El manual de las MPR no especifican rangos de edad límite para su aplicación, pero su aplicación se puede realizar en niños, jóvenes, adultos y ancianos (RAVEN, 2004).

Sin embargo, posteriormente se desarrollaron las MPC y MPA. Las MPC fueron desarrollada, específicamente, para dar atención a niños, personas con discapacidad intelectual y ancianos. Mientras que la escala avanzada se creó para medir la capacidad mental de personas con promedios por encima de la norma (un 25% superiores). Según el manual de las MPR, las MPA pueden medir con mayor exactitud la velocidad requerida en el desempeño intelectual (RAVEN, 2004).

Durante la etapa de fabricación de las MPR, las escalas estaban constituidas principalmente por nueve ítems. Los ítems eran figuras similares a un rompecabezas, que se resolvían encontrando la pieza ausente (en esta primera fase, el instrumento era muy parecido a otro desarrollado por Spearman, aunque el procedimiento para resolver los ítems de cada uno difería). Los primeros hallazgos respecto al método de trabajo que exigía el instrumento fueron que, a diferencia de los adultos, los niños de corta edad tenían dificultades para comprender la esencia de algunos problemas planteados. Por un lado, si se trataba de figuras en las cuales se representaban patrones continuos, los niños no tenían problemas para resolverlos; por otro lado, si se trataba de un ítem con cuatro figuras relacionadas entre sí, la tarea se tornaba complicada para los niños, incluso si antes se les aplicaban varios patrones continuos. Por lo que se concluyó que 1) para resolver problemas donde se incluían distintas figuras relacionadas entre sí, la madurez intelectual era un factor importante, y que 2) la resolución de patrones de cierta naturaleza, no ayudaba en el adiestramiento de los niños para resolver otros patrones distintos (RAVEN, 2004).

En relación a la manera en que los examinados responden a los ítems, se han encontrado algunas variaciones. En lo que concierne a los niños, se determinó que la respuesta seleccionada está relacionada estrechamente con el campo de pensamiento utilizado por el niño a través del aprendizaje previo de un problema (MILLER & RAVEN, 1939). En ese sentido, la repetición activa es la manera más común en la que responden los niños de corta edad. Ésta consiste en que el niño percibe el patrón como una sucesión repetida de formas o símbolos. La manera menos común de responder, por niños no tan pequeños, es la denominada percepción pasiva; aquí el examinado no alcanza a vislumbrar una tarea a resolver y es observada, principalmente, en ancianos de edad avanzada y niños con discapacidad intelectual (RAVEN, 2004). Además, las respuestas a los ítems también se encuentran influenciadas por la posición en la que se ubican: las personas tienden a

explorar toda la gama de respuestas si aquellas que evidentemente son equívocas son colocadas en posiciones más atractivas para la percepción visual (al centro) (MILLER & RAVEN, 1939). Por lo tanto, se llegó a la conclusión de que las mejores posiciones para las respuestas correctas eran los extremos, mientras que para las soluciones erróneas es mejor una posición central.

De acuerdo con Raven (2004), se ha encontrado que la posición de una gama de posibles respuestas no tiene una gran relevancia en tanto se trate de ítems de dificultad moderada, debido a que las probabilidades de acertar no se disparan. No obstante, si se trata de ítems de gran dificultad, una posición central de la respuesta correcta incrementaría las posibilidades de acertar por azar.

Además, el test MPR fue diseñado con la finalidad de evitar sesgos culturales, así como sesgos por instrucción escolar y a eso se debe la naturaleza de sus ítems. Su diseño también intenta evitar que los puntajes obtenidos se vean alterados por una experiencia previa con instrumentos similares (RAVEN, 2004).

A continuación se detallará más sobre las series y los ítems que componen las MPR, así como los materiales, instrucciones y normas de calificación necesarios para realizar una correcta aplicación.

3.2 BASES TEÓRICAS DEL TEST MATRICES PROGRESIVAS DE RAVEN

La teoría clásica en la que se fundamentan las MPR es la *g* de Spearman. El autor de la teoría observó que, a partir del análisis factorial de correlaciones entre los puntajes obtenidos en diferentes pruebas, ciertas tareas tendían a presentar una alta correlación entre ellas. Esta correlación fue interpretada como un factor general (*g*) de la capacidad mental, mientras que las correlaciones más bajas en la matriz factorial fueron interpretadas como un factor específico (*s*). Para Spearman, el *g* era el factor que le interesaba, porque éste era una representación de una capacidad mental general (SPEARMAN, 1927).

Spearman mostró menos interés en el estudio del factor específico o *s*, debido a su baja influencia en la determinación del nivel de desempeño, concentrándose únicamente en la descripción de la naturaleza de *g*; por ello es que, aunque Spearman utilizó el término bifactorial, su teoría es más comúnmente denominada teoría unifactorial o teoría de un solo factor (GREGORY, 2001).

Spearman relacionaba el factor *g* con una energía que se extendía a toda la corteza, mientras que a *s* lo relacionaba con grupos neuronales focalizados en llevar a cabo la operación mental específica que requiere una prueba (KANE & BRAND, 1905): «Así, estos grupos neuronales pueden funcionar como maquinarias opcionales a las que el suministro común de energía puede distribuirse de manera alternativa» (SPEARMAN, 1923, pág. 175).

Durante el desarrollo de su teoría sobre la inteligencia humana, Spearman contribuyó con el invento del análisis factorial, técnica estadística que utilizó con la finalidad de determinar el número de factores subyacentes que explicaban las correlaciones entre diversas pruebas (AMADOR, FORNS, & KIRCHNER, 2005).

Gran parte de los estudios de Spearman se concentraron en demostrar la proposición de que las pruebas psicológicas con cargas elevadas de *g* debían estar correlacionadas, mientras que las pruebas donde no se apreciaba correlación alguna, las cargas de *g* debían ser bajas; en especial, Spearman consideró que las pruebas meramente sensoriales eran un ejemplo de medidas con saturaciones de *g* bajas y por lo tanto podrían representar un factor *s* (GREGORY, 2001).

Para explicar el factor *g* y las diferencias observables de este factor en las pruebas de capacidad mental entre individuos, Spearman consideró que el uso de tres principios cognitivos era clave: la *aprehensión de la experiencia*, entendida como la comprensión de conceptos a partir de una experiencia previa; la *edución de relaciones*, que se refiere a la capacidad de inferir o establecer relaciones entre dos objetos y, por último, la *edución de correlaciones*, comprendida como la capacidad de extrapolar el anterior principio a un nuevo dominio (KANE & BRAND, 1905). Para entender de manera más sencilla los tres principios anteriores, basta con recordar la manera en la que se resuelve una analogía, por ejemplo: A:B::C:D. La cual sería imposible resolver si no se tiene conocimiento anticipado de los términos A, B y C.

Las MPR fueron diseñadas para medir la capacidad eductiva. Ya que esta capacidad está relacionada con las representaciones mentales no verbales (abstracciones) y la habilidad de percibir relaciones que no son en absoluto evidentes, las MPR utilizan imágenes en lugar de ítems donde se solicite una respuesta influida por el lenguaje y, por lo tanto, por la cultura y la instrucción escolar. A esto último se le conoce como capacidad reproductiva, ésta se relaciona con el aprendizaje de información explícita (RAVEN, 2004).

No obstante, para determinar si las MPR cuentan con una adecuada consistencia de medición del constructo para el que fueron diseñadas (que en este caso es la inteligencia) y que realmente están midiendo ese determinado constructo y no otro, fue necesaria la realización de diversos estudios de confiabilidad y validez que precisaran aquello.

3.3 DESCRIPCIÓN DEL TEST MATRICES PROGRESIVAS DE RAVEN

El manual de las MPR facilita un fragmento introductorio donde se ofrecen una serie de descripciones del test de acuerdo a sus distintas propiedades (RAVEN, 2004):

- Por su objetivo: Un test de capacidad intelectual (habilidad mental general). Es un instrumento destinado a medir la capacidad intelectual... para comparar formas y razonar por analogía, con

independencia de los conocimientos adquiridos. Informa acerca de la capacidad presente del examinado para la actividad intelectual en el sentido de su más alta claridad de pensamiento en condiciones de disponer de tiempo ilimitado.

- Por su material: Un test de lagunas geométricas abstractas. Como material de prueba, este test utiliza una serie de figuras abstractas (geométricas) incompletas. Es un test no-verbal tanto por la índole del material como por la respuesta que demanda del examinado.
- Por la índole de sus problemas: Un test de matrices. Cada figura geométrica implica un patrón de pensamiento (incompleto, trunco), una matriz (de la palabra latina *mater*=madre). Cada figura representa la madre o fuente de un sistema de pensamiento. La serie completa integra una escala de matrices en orden de complejidad creciente construida con la finalidad de identificar en la forma más completa posible los sistemas de pensamiento del desarrollo intelectual.
- Por la tarea interna que debe realizar el examinado: Un test perceptual, de observación, comparación y razonamiento analógico.
- Por la tarea externa que debe realizar el examinado: Un test de complemento. El examinado suministra su respuesta a los problemas planteados, completando las figuras lacunarias con el trozo que corresponda.
- Por la forma de obtener la respuesta: Un test de selección múltiple. Para que el sujeto brinde su solución, se le ofrece, para cada una de las figuras lacunarias, varios (6 u 8) trozos entre los cuales se encuentra el único adecuado, a fin de que indique cuál de ellos es el que considera apropiado.
- Por su administración: Es un test autoadministrable, y por tanto puede administrarse, indistintamente, en forma individual o colectiva.
- Por sus empleos: Un test de clasificación y clínico.
- Por sus fundamentos: Un test factorial (pág. 9).

Se advierte que se dará una descripción más específica de las MPG, en comparación con las MPC y MPA, debido a que las MPG son primordiales para la presente investigación; no obstante, tanto las MPC como las MPA se describirán de manera breve, con la finalidad de aclarar las diferencias entre las distintas escalas de las MPR.

Las MPC, como ya se ha dicho, están destinadas a medir la capacidad de niños, personas con discapacidad intelectual o ancianos de edad avanzada. Las MPC son las únicas que cuentan con dos tipos de materiales a elegir por el examinador:

1. Un cuadernillo constituido por tres series (A, Ab y B), cada una con 12 ítems. Las figuras contenidas en este cuadernillo están coloreadas, con el propósito de hacerlo más atractivo a los niños.
2. Un tablero con piezas movibles, que funcionan como un rompecabezas (de acuerdo con el

manual, el tablero no se comercializa en todos los países en los que la prueba está a la venta, pero se proporcionan instrucciones para que el profesional con intención de utilizarlo lo pueda fabricar) (RAVEN, 2004).

En el caso de las MPA, el material que se utiliza son dos cuadernillos que contienen la serie I (con 12 ítems que representan toda la gama de dificultades contempladas en las MPG) y la serie II (con 36 ítems que tienen como propósito discriminar a personas con capacidades sobresalientes). La serie II se puede aplicar sin límite de tiempo o con límite de tiempo.

Las MPG están integradas por 5 series (A, B, C, D y E), cada una con 12 problemas distintos y ordenados de menor a mayor dificultad. La aplicación de las MPG no requieren de establecer un tiempo límite; se deja trabajar al examinado de manera libre. El tiempo promedio de una aplicación de las MPG es de una hora y, al igual que las MPC y MPA, se pueden aplicar tanto individual como colectivamente. Para la aplicación, el manual especifica las instrucciones textuales que hay que proporcionar al examinado, además de sugerencias sobre las condiciones físicas del lugar en las que se ha de utilizar. En seguida se aborda el material necesario para llevar a cabo la correcta aplicación del instrumento.

3.3.1 MATERIAL DE LA PRUEBA

Las MPR cuentan con instrucciones de aplicación dependiendo de la situación a la que se enfrente el examinador, es decir, proporciona pasos a seguir tanto para las aplicaciones individuales como para las aplicaciones grupales. En el caso de las aplicaciones individuales, se ha de contar con un cuadernillo de matrices, que contiene las cinco series de las MPG (A, B, C, D y E). Los primeros ejercicios de la serie A se utilizan como ejemplos para ilustrar la manera en la que se debe trabajar; además, se necesita proporcionar al examinado un lápiz, una goma y una hoja de respuestas para registrar las opciones elegidas. En el caso de las aplicaciones grupales, el material sólo difiere en cantidad. Aunque, cuando se trata de grupos, se puede realizar una ampliación del ítem A1 y del protocolo de respuestas para explicar la forma de trabajo.

3.3.2 APLICACIÓN Y CALIFICACIÓN

Como ya se ha mencionado, las MPR contemplan aplicaciones individuales y aplicaciones grupales. Las primeras que se describen a continuación son las instrucciones para aplicaciones individuales:

*Durante una conversación preliminar, anote los datos del niño en la hoja de respuestas.
Abra el cuaderno de matrices en la primera ilustración y diga: «observa esto [señale la figura de arriba]. Es un dibujo en el que se ha recortado una parte. Cada una de estas*

piezas [señale una tras otra] tiene una forma que ajusta bien en el espacio en blanco, pero solamente una completa bien el dibujo. La N° 2 ni siquiera tiene un dibujo. La N° 3 es totalmente equivocada. La N° 6 está casi bien, pero no concuerda aquí [señale la parte en blanco]. Sólo una de estas piezas es la correcta. Indica cuál es». Si el niño no elige la figura correcta, continúe explicando hasta que haya captado con claridad la naturaleza del problema que hay que resolver. A continuación pase al problema A2 y diga: «ahora señala la pieza que salió de este dibujo». Si el examinado fracasa, hay que redemostrar el problema A1, y volver a pedir la solución del problema A2. Si el niño elige la figura correcta, pase al problema A3 y proceda como antes. En la matriz A4, antes que el niño tenga tiempo de señalar una de las figuras, vuelva a hablarle: «Mira con cuidado el dibujo [recórralo con el dedo]. Sólo una de estas piezas es totalmente correcta. Presta atención, primero míralas a todas una por una [señale cada una de las 6 piezas]. Ahora señala la que corresponde poner aquí [señale el espacio en blanco de la matriz]». Cuando el niño ha elegido una de las piezas, sea ella correcta o errónea, pregunte: «¿Es ésta la que va bien aquí? [señale el dibujo y el espacio a llenar]». Si el examinado responde «sí», acepte esa elección con aprobación, sea ella correcta o errónea. Si el examinado quiere revisar la decisión, dígame: «bien, señala entonces la correcta». Después que el niño haya vuelto a optar, bien o mal, repita la pregunta anterior: «¿Es esta la figura correcta?» Si el niño quedó satisfecho, acepte esa elección. Si todavía queda duda, insista: «Bien, ¿cuál es la figura correcta?»; después acepte como decisión final la pieza señalada por el niño. El problema A5 se demuestra del mismo modo que el A4. En cualquier paso entre A1 y A5 se puede volver al problema A1 para ilustrar el planteamiento si resulta necesario, y después pedirle al niño un nuevo intento. Si el niño no logra resolver correctamente los problemas 1 a 5, no tiene sentido continuar con el test. Si los problemas 1 a 5 son resueltos con facilidad, pase al A6 y diga solamente: «Mira con cuidado el dibujo. ¿Cuál de estas piezas [señálelas por turno] va bien aquí? [señale el espacio en blanco]. Presta atención, sólo una es la correcta. ¿Cuál es? Trata de estar seguro antes de señalarla». Se repiten estas instrucciones con cada problema mientras resulte útil. Registre las respuestas del niño en el lugar previsto de la hoja de respuestas (RAVEN, 2004, pág. MPG 23).

En lo que respecta a la aplicación grupal, el test de MPR también proporciona algunas sugerencias al examinador sobre la distribución de los examinados, con la finalidad de realizar una adecuada aplicación:

El test puede administrarse a cualquier número de sujetos, según las comodidades disponibles. A cada grupo examinado se le debe dar más o menos una hora. Cada sujeto estará sentado cómodamente a una mesa o pupitre con lugar para el cuaderno y la hoja de res-

puestas; los pupitres deben estar lo suficientemente distanciados entre sí como para evitar la copia. Entre ellos debe quedar un espacio que permita al examinador desplazarse sin perturbar la tarea. Los sujetos estarán de frente al examinador.

Se distribuyen los lápices y las hojas de respuestas o protocolos, y se les pide a los examinados que anoten sus datos en ellas. Hecho esto se reparten los cuadernos de matrices. Se les dice a los niños que no los abran hasta que todos estén listos.

El examinador dice: «Abran los cuadernos en la primera página. Van a ver algo como esto». Se le presenta al grupo un cuaderno abierto en el problema A1, o la ampliación para la demostración. «Arriba dice serie A, y en sus hojas de respuestas tienen una columna, aquí, destinada a esa serie A. Ésta es la prueba A1. Ya ven de qué se trata. Arriba hay un dibujo en el que se ha suprimido una parte. Cada una de estas figuras de abajo» (señálelas, una tras otra) «tiene una forma que ajusta en el espacio en blanco, pero no todos completan bien el dibujo. La N° 1 (señale la pieza y después la matriz) «es totalmente errónea. Las número 2 y 3 tampoco sirven –ajustan en el espacio, pero no completan bien el dibujo-. ¿Y la número 6? Ésta tiene el dibujo correcto» (mostrar que el dibujo es el mismo de la matriz) «pero con una parte en blanco»: Señale con el dedo la figura correcta. Observe si todos saben hacerlo. En caso necesario agregue nuevas explicaciones, y añada: «Sí, la N° 4 es la figura correcta. De modo que la respuesta a A1 es 4. Escriban el 4 aquí, junto al número uno en la columna A de la hoja de respuestas. No den vuelta todavía la hoja del cuaderno».

Aguarde a que todos hayan terminado para continuar. «En cada página del cuaderno hay un dibujo con un espacio en blanco. En cada caso hay que descubrir cuál de las figuras de la parte inferior es la que completa el dibujo de arriba. Cuando hayan encontrado la figura correcta, escriban su número en la hoja de respuestas al lado del número del dibujo. No escriban nada en los cuadernos. Los problemas son simples al principio y se van volviendo más difíciles a medida que ustedes avanzan. No hay ninguna trampa. Si prestan atención al modo como resuelven los problemas fáciles, los que siguen les resultarán menos difíciles. Traten de resolver un problema tras otro, desde el principio hasta el final del cuaderno. Trabajen a su propia velocidad. No salten ningún problema ni vuelvan atrás. Veamos cuántos pueden resolver. Tienen todo el tiempo que deseen. Den vuelta a la hoja y traten de encontrar la solución del problema que sigue».

Cuando ha transcurrido el tiempo suficiente como para que todos hayan anotado la respuesta A2, diga: «Desde luego, la solución es la figura 5. ¿Han escrito el número 5 al lado del número 2 en la columna A de la hoja de respuestas? Ahora sigan solos hasta el final del cuaderno. Yo me voy a acercar para ver si están procediendo bien».

Es preciso que el examinador verifique si cada sujeto ha anotado sus soluciones a los cinco primeros problemas. En cuanto el examinado ha captado la naturaleza de los problemas iniciales, el examinador deja de ayudar con el método de razonamiento, pero observa si los sujetos anotan correctamente sus elecciones.

Al cabo de más o menos media hora, se pide a los examinados que avisen cuando han terminado. Cuando lo hacen, el examinador controla que el protocolo haya sido llenado correctamente y que se hayan abordado todos los problemas (RAVEN, 2004, págs. MPG 24-25).

Llevar a cabo la calificación de las MPR no es una tarea difícil, salvo que uno o varios sujetos hayan anotado dos respuestas para un único problema. En ese caso, el examinador deberá preguntar al examinado por la respuesta hacia la que se inclina más. Para realizar una adecuada calificación del test es necesario seguir las siguientes indicaciones contenidas en el manual del test:

El puntaje del sujeto con la escala es el número total de problemas que ha resuelto bien al permitirle trabajar con tranquilidad a lo largo del todo el test, desde el principio hasta el fin.

La consistencia o regularidad del trabajo puede evaluarse restando del puntaje parcial en cada una de las cinco series el puntaje esperado en cada una de ellas para el mismo puntaje total. La diferencia entre el puntaje en cada serie y el normalmente esperado por el puntaje total presenta, por ejemplo, el siguiente aspecto numérico: Discrepancias: 0, -1, +2, -2, +1.

Si en algunas de las series el puntaje del sujeto se desvía más de 2, el puntaje total que obtuvo con la escala no puede aceptarse en su valor nominal como medición consistente de la capacidad general para la actividad intelectual. A los fines generales, el puntaje total parece ser relativamente válido incluso cuando al comienzo aparezcan discrepancias mayores de 2 puntos.

En una cierta cantidad de casos, el sujeto acierta por azar con la respuesta correcta. Cuando se le permite que complete todo el test, el número de las elecciones por azar será proporcional al número de problemas que no pudo resolver. Los sujetos con puntajes bajos presentan una cantidad proporcionalmente mayor de éxitos por azar, de modo que los puntajes bajos son siempre menos consistentes y confiables que los puntajes altos (RAVEN, 2004, MPG, pág. 26).

Por supuesto, el manual de las MPR le brinda al examinador las tablas necesarias para la calificación del test. Hay una tabla con todas las soluciones correctas de las series, una tabla con los puntajes esperados y una serie de tablas con percentiles que se aplican según la edad de los examinados. Además, se proporciona una interpretación de acuerdo con el percentil obtenido. En los anexos se pueden observar los problemas de las cinco series del test MPG, así como el protocolo de respuestas y las tablas anteriormente mencionadas.

Ya se ha mencionado con anterioridad algunos antecedentes del test MPR, sobre todo aquellos relacionados con el desarrollo de sus diferentes escalas y la construcción de sus ítems; mas es preciso abordar sus antecedentes teóricos, es decir, las bases teóricas que sustentan las MPR, por ello el siguiente apartado se ha dedicado a la descripción de éstas.

3.4 PROPIEDADES PSICOMÉTRICAS DEL TEST

Antes de hacer mención de los estudios de confiabilidad y validez de las MPR, es importante tener en cuenta algunos sucesos anteriores, tales como los estudios de estandarización. La estandarización de una prueba está relacionada con un proceso al que se someten los instrumentos psicométricos para establecer normas de aplicación, calificación e interpretación de las puntuaciones arrojadas por éstos. Las estandarizaciones deben realizarse mediante la aplicación de un determinado instrumento psicométrico a una población que represente aquella población a la que está dirigida. Las puntuaciones crudas arrojadas por el instrumento se transforman en puntuaciones normalizadas. Generalmente las muestras poblacionales presentan una proporción similar en relación a la edad y al género de los examinados (SIERRA & CASTRO, 2013).

A continuación se presentan algunos de los primeros estudios llevados a cabo para la normalización de las MPR. Es preciso mencionar que se ha citado de manera textual, no obstante, el siguiente párrafo está basado en otras fuentes referidas por el autor, que aquí se han omitido a propósito, con la idea de no crear confusión sobre las fuentes de este trabajo:

Las matrices progresivas escala general fueron desarrolladas originalmente a mediados de la década de 1930, y revisadas y estandarizadas en Ipswich en 1938. Durante la guerra se recogieron normas amplias para adultos, y el test fue nuevamente estandarizado, junto con la Escala de Vocabulario Mill Hill (VMH), con escolares de Colchester en 1943. En la década de 1940, se acumularon datos sobre personas ancianas y otras. Durante las décadas de 1950 y 1960 se llevaron a cabo varios chequeos de la exactitud de las normas. En 1972, se examinó con hojas de respuesta de puntuación mecánica una muestra representativa de 3700 escolares irlandeses de 6 a 12 años. En 1979, se publicó una estandarización alemana en gran escala de las MPG. En 1979 se tomó el test a una muestra representativa nacional de 3500 escolares británicos de 6 a 16 años que no incluía alumnos de las escuelas especiales. Los resultados detallados de este estudio fueron protocolizados con la British Lending Library. Entre 1984 y 1986 se realizaron una serie de estudios normativos locales en distritos escolares de todo Estados Unidos (RAVEN, 2004, MPG, pág. 6)

De los estudios realizados para una nueva estandarización en Inglaterra, se desprendió que a pesar de la aparición de un incremento en los puntajes de las MPR observado a través de los años

(razón por la cual se actualizaron las normas en 1979), parecía que culturalmente las normas no eran afectadas. Además, se encontró que los puntajes no presentaban una diferencia significativa relacionada con el género de los examinados, que el test se ajustaba correctamente a los examinados sin importar el lugar al que pertenecían (es decir, que todos los examinados comprendían la naturaleza del test) y una probabilidad ínfima de observar respuestas debidas al azar. En contraste con lo anterior, en Estados Unidos las investigaciones permitieron advertir variaciones de distrito escolar a distrito escolar. Estas variaciones estaban sujetas a variables étnicas, socioeconómicas y de localización geográfica. Sin embargo, las variaciones en las normas encontradas en un mismo distrito no se relacionaban con las variables anteriores. Así mismo, se pudo observar un adecuado coeficiente de validez de criterio y una comprensión adecuada de la forma de trabajo, sin importar el origen de los examinados, por lo que se concluyó que, pese a las variaciones de las normas observadas, el test funciona adecuadamente (RAVEN, 2004).

3.4.1 ESTUDIOS DE CONFIABILIDAD

Acerca de los estudios sobre la estabilidad del test, Raven (2004) respalda la confiabilidad de las MPR en una serie de estudios psicométricos (además de los realizados durante las estandarizaciones de 1938, 1943 y 1979 en Inglaterra), publicados desde finales de la década de los 40's hasta casi la mitad de la década de los 80's, entre los que destacan los llevados a cabo por autores como Banks & Sinha, Elley & MacArthur, King, Laroche, Loranger & Misiak, Georgas, Keir, Burke, Gittins, Evans, Jensen & Munro, Morgan, Kuhnlein, Stinissen, Swinnen, Baraheni, Johnson, Sorokin, Rocco, Ganguly, Dolke & Sharma.

La mayoría de los estudios, anteriormente mencionados, utilizaron el método de división por mitades para determinar la confiabilidad de las MPR. El resto se trata de estudios que utilizaron la técnica de Kuder-Richardson. Estos estudios se realizaron tanto en niños como en adultos. Entre los resultados obtenidos por estos estudios, se encontraron coeficientes de confiabilidad que oscilan entre 0.60 y 0.98 (RAVEN, 2004), los cuales se pueden interpretar como coeficientes de confiabilidad moderados y altos, respectivamente. No obstante, de la totalidad de los estudios citados en el manual de las MPR, una gran parte arroja coeficientes altos.

Es preciso mencionar que gran parte de los estudios sobre la confiabilidad de las MPR se realizaron en Gran Bretaña y Estados Unidos, mientras que el resto se llevó a cabo en Alemania, Bélgica, Irán, Corea, China, Yugoslavia, Uruguay e India. En algunos de aquellos estudios se utilizaron sujetos de distintos grupos étnicos, tales como anglos, asiáticos, negros, navajos e hispanos); sin embargo, ninguno de ellos se realizó en México u otros países de América Latina.

En relación a la consistencia de la prueba a través del tiempo, comprobada a través de un test-retest, algunos estudios reportaron valores de confiabilidad bastante bajos. Entre los datos recogidos se hallaron valores entre 0.88 y 0.55, 0.55 y 0.84 y de 0.61 a 0.46. Los intervalos de tiempo de aplicación van de semanas a varios años, pero se considera que las bajas correlaciones se pudieron deber a muestras de sujetos poco representativas (RAVEN, 2004).

3.4.2 ESTUDIOS DE VALIDEZ

Sobre la validez de las MPR, se han contemplado los resultados de numerosos estudios que han analizado la validez de contenido, de criterio y de constructo, respectivamente. En lo que respecta a las revisiones de validez de criterio, se realizaron contrastes con las escalas de Wechsler, el test de Binet y la Escala de Vocabulario Mill Hill; los coeficientes de correlaciones que se desprenden de las investigaciones van de 0.30 a 0.95. Por supuesto, los estudios varían en relación a las características de las muestras utilizadas, tales como la edad, el grupo étnico y el sexo. Por otro lado, las MPR también se contrastaron con pruebas de rendimiento escolar y con evaluaciones directas de los profesores, en estos estudios las correlaciones más bajas observadas son de 0.20, 0.22 y 0.26, mientras que las correlaciones más altas son de 0.69, 0.76 y 0.82. A partir de estas correlaciones, se concluyó que al utilizar tests de rendimiento escolar o evaluaciones directas (calificaciones de trabajos y exámenes, por ejemplo) como criterio para determinar la validez de las MPR, los coeficientes tienden a demostrar coeficientes más inferiores, en comparación con los obtenidos al utilizar pruebas de inteligencia como criterio (específicamente las escalas de Wechsler, Binet y VMH) (RAVEN, 2004).

De otros estudios sobre la validez de contenido de las MPR, se desprenden coeficientes de correlación biserial promedio que van de 0.45 a 0.52, mientras que las correlaciones entre diferentes reactivos fueron de muy bajas (0.2) a bastante altas (no se especifica concretamente). En uno de los estudios donde se analizó la consistencia interna del test MPR, se observó que los índices de dificultad entre los ítems no variaban adecuadamente, sino que se observaba una predominancia de ítems con dificultad media, lo cual sugería un reordenamiento de los ítems (RAVEN, 2004). Generalmente la validez de contenido se determina a través de un jueceo realizado por expertos, mas, en este caso, sólo se han mencionado estudios cuantitativos.

En los análisis factoriales realizados en las investigaciones sobre la validez de constructo de las MPR, se observan saturaciones en g que van de 0.36 a 0.91. De los análisis factoriales se desprende que las MPR comprenden otros factores además de g. Por lo anterior, algunos investigadores han sugerido que existen cargas moderadas relacionadas con un factor espacial, mas no se reportan cargas significativas en otros factores (RAVEN, 2004).

Los estudios mencionados en este apartado son aquellos en los que el manual de las MPR se ha respaldado para dar razón de sus coeficientes de confiabilidad y validez, así como datos importantes sobre sus estandarizaciones; sin embargo, a continuación se muestran diversos estudios (no mencionados por el manual de las MPR) donde se analizan las propiedades psicométricas del test en cuestión.

3.5 ESTUDIOS POSTERIORES SOBRE LAS PROPIEDADES PSICOMÉTRICAS DEL TEST

Los estudios realizados para conocer y comprender el funcionamiento del test MPR son escasos, no obstante, es importante abordarlos para adquirir un panorama que nos ayude a comprender hacia dónde se inclinan los hallazgos previos a la presente investigación.

Torres & Cuesta (1992) llevaron a cabo una investigación con 1028 estudiantes universitarios de primer curso de la Universidad del País Vasco, en la que pretendían realizar una revisión de las propiedades psicométricas del test MPR, partiendo tanto de la TCT (confiabilidad, índice de dificultad de los ítems e índice de discriminación), como de la TRI (función de información y curva característica del test parámetros a y b). En cuanto a los índices de dificultad y discriminación, encontraron que los ítems del test resultaron muy fáciles y con un índice de discriminación muy bajo (sobre todo en aquellos ítems de mayor dificultad) para la muestra utilizada en el estudio. Al respecto, la media del índice de dificultad del ítem (b) fue de -1,798. Además, en los ítems 1 y 2 se pueden observar valores de b extremos de -7,583 y -5,083 respectivamente. Por otro lado, los valores del índice de discriminación (a) se encuentran por debajo de 1, por lo tanto, de acuerdo a la TRI, no se consideran discriminativos, es decir, sin importar el valor de Θ (nivel de rasgo), los ítems serán respondidos de manera correcta por muchos sujetos, aunque también serán respondidos de manera incorrecta por muchos otros sujetos. A decir verdad, sólo 4 ítems muestran un índice de discriminación superior a uno: 4, 16, 22 y 23. Por otro lado, Torres & Cuesta (1992) obtuvieron un índice α de 0,83, lo cual hace confiable a la prueba. Por último, se constató que sólo en 14 de los 48 ítems se refleja el aporte de información máxima con valores entre ± 1 , no obstante, la mayor cantidad de información se observa en los ítems ubicados en los niveles más bajos de Θ .

En otra investigación (CHICO, 1997), en la cual se pretendía determinar si la interpretación de las puntuaciones en el test MPR es igualmente válida para grupos de personas antisociales (942 internos del centro penitenciario de Tarragona, con un nivel de estudios primarios y buena comprensión de lectura) que para no-delinquentes (400 soldados que cumplían servicio militar en Tarragona, con nivel de estudios similares a los reclusos), con un nivel educativo similar, hallaron un incremento sucesivo en las series, excepto en las series 3 (C) y 4 (D), en las cuales los valores no

presentan grandes diferencias, respectivamente para los delinquentes la media de la serie C= 7,53, mientras que en la serie D= 7,90. En cuanto a los no-delinquentes, la media para la serie C= 9,26 y para la serie D= 9,25. Como se puede observar, estos datos sugieren que en realidad las series 3 y 4 mantienen una dificultad muy similar e, incluso, en el grupo de delinquentes se observa ligeramente una menor dificultad de la serie D en comparación con la C. Los grupos parecen presentar diferencias en cuanto al rendimiento general durante la prueba, y va incrementando conforme la dificultad es mayor, a pesar de que ambos grupos cuentan con un nivel educativo similar. También se encontró que, a medida que se eleva la dificultad de las series, las diferencias de comportamiento entre los grupos delinquentes y no delinquentes aumentan, lo cual significa que pudieran estar interviniendo otros factores como la desmotivación, cansancio, aburrimiento, etc. Y con ello, aparecer situaciones de respuestas azarosas. No obstante, al correlacionar la serie 5 con las otras 4 series se observa, para el grupo de delinquentes un coeficiente de 0,638, mientras que para el grupo de no-delinquentes obtuvieron un coeficiente de 0,628, lo que significa que las puntuaciones obtenidas en la serie E por los grupos, tanto de delinquentes como de no-delinquentes, son bastante predecibles a partir del comportamiento observado en las 4 series anteriores, sin mencionar el hecho de que los coeficientes son muy similares para los dos grupos, por lo tanto Chico (1997) descarta la probabilidad de que el grupo de delinquentes (que suelen mostrar puntuaciones más bajas en pruebas que miden el mismo constructo) presenten respuestas azarosas o negligentes, además de rechazar la hipótesis de que la prueba no muestra funcionamiento diferencial en absoluto.

Por su parte, Cairo, Cairo, Bouza & Ponce (2000), llevaron a cabo un estudio acerca de las MPR, con la finalidad de establecer el índice de dificultad y discriminación de los ítems, demostrar alternativas acerca del empleo de las dos formas básicas de calificación (por aciertos y por errores), proponer una nueva alternativa de formato de la prueba a partir de los hallazgos durante el estudio, así como presentar las normas de los percentiles obtenidos. Dicha investigación se llevó a cabo con una muestra de 589 sujetos (estudiantes pre-universitarios, universitarios y profesionales), 303 mujeres y 286 hombres con edades que van desde los 11 años hasta los 68 años. En relación al análisis de dificultad de los ítems, lo determinaron a través de la resta de las respuestas erróneas y los ítems sin contestar, a partir de lo cual obtuvieron que el ítem más fácil es el A1 con un 100% de respuestas correctas, y el ítem más difícil es el E12 con un 12,6% de aciertos. Con respecto a los porcentajes de dificultad de las series, la serie A aparece como la más fácil con un 95,1% de aciertos, mientras que la serie E cuenta con un 43,8% de aciertos, lo que la hace la más difícil. Sin embargo, aunque aparentemente (a partir de estos resultados) se pueda pensar que el orden de dificultad de los ítems es el óptimo, Cairo, Cairo, Bouza & Ponce (2000) mencionan que estos porcentajes no siguen en realidad un orden de dificultad que va de menor a mayor, sino que algunos ítems dentro

de las series, que se supone deberían representar un índice de dificultad mayor, de hecho, son de menor dificultad. Además, en las series C y D se puede observar que la primera de ellas cuenta con un índice de dificultad mayor que la segunda, observándose a la serie C con un 73.6% de éxitos, mientras que la serie D cuenta con un 77.6%, con lo cual podemos decir que la serie C es más difícil con respecto de la D. Posteriormente, al clasificarse a los ítems en tres categorías a) fáciles (acertados por el 60% de los sujetos) b) moderados (resueltos entre el 40% y el 60% de los sujetos) y c) difíciles (contempla el 30% restante de sujetos que respondieron correctamente) se observan que los ítems de menor dificultad son 46, representando el 76.7% del total de ítems; los ítems de moderada dificultad son 6 representando el 10% del total; y finalmente, los ítems de mayor dificultad son 8 representando el 13% del total. Con lo anterior se puede deducir que en general la prueba ha resultado muy fácil para los participantes. Para contrarrestar los problemas anteriores, los autores proponen tres posibilidades, la primera consiste en eliminar la organización serial de los ítems y presentarlos de manera continua en orden de dificultad ascendente, con ello se reduce el número de ítems y por lo tanto el tiempo de aplicación; la segunda consiste en la reordenación completa de los ítems, lo que conllevaría pasar ítems de una serie a otra para obtener un orden de dificultad más aceptable; en cuanto a la tercera alternativa, se sugiere una sustitución, reordenamiento y completamiento de un grupo más o menos numeroso de ítems, además de adecuar las normas de los percentiles para la población que se pretenda evaluar.

En una muestra de 4258 estudiantes, con edades de entre 5 y 22 años de Chile, se aplicó las MPR para conocer el CI de los alumnos y a través de estos datos poder hacer una revisión de los baremos propuestos por Raven, así como realizar una revisión tanto de la validez (a través de una asociación con el test Goodenough) como de la confiabilidad (llevando a cabo un test-retest). A través de los resultados obtenidos, se hace notar que un 44.5% de los sujetos se posiciona en el rango III) término medio, el 21.4% se ubicó en el rango IV) inferior al término medio, el 20.6% en el rango II) superior al término medio y por último el 6.4% y 6.8% se ubicaron en los rangos V) deficiente y I) intelectualmente superior; lo anterior con respecto a la muestra completa (4258 sujetos) y, de acuerdo con Ivanovic, Forno, Durán, Hazbún, Castro & Ivanovic (2000), se puede calificar como una distribución normal de los datos. No obstante, al observar los resultados obtenidos en las MPC, por niños con edades de entre 5 y 11 años (2210 sujetos), se puede apreciar en las gráficas de frecuencia cómo varía la dificultad de los ítems; tanto en la serie A, Ab y B se ve cómo en algunos de los ítems, el porcentaje de respuestas correctas cae precipitadamente y después vuelve a repuntar. Los autores realizaron una categorización de acuerdo a los rangos en los que se ubicaban los sujetos estudiados, para graficar de una manera clara. Como un dato que se considera relevante para tomarse en cuenta, se menciona que en las series Ab y B se nota cómo los rangos intercambian

posiciones, cuando no debería pasar, en caso de que la prueba contara con un orden de dificultad ascendente. Un ejemplo extremo se puede observar en la serie Ab, donde los sujetos ubicados en los rangos III y IV son superados en porcentaje de acierto en el ítem 12. Con respecto a los 2203 sujetos a los que se les aplicaron las MPG, se observa el mismo caso de ítems con mayores porcentajes de aciertos, cuando de hecho se supone que cuentan con una dificultad superior. Al correlacionar el test de Goodenough con el test de MPR, se puede observar un bajo coeficiente de correlación, presentando los siguientes puntajes I año básico $r=0.214$, II año básico $r=0.326$, VI año básico $r=0.297$, VI año básico $r=0.349$, VIII año básico $r=0.248$ y I año medio $r=0.335$. En relación al test-retest para revisar la confiabilidad, se volvió a aplicar las MPR a un grupo de 468 estudiantes, encontrando que el 76.3% de los estudiantes mantuvieron su rango percentil e incluso lo elevaron; por otro lado, el 23.7% disminuyó su percentil (IVANOVIC, FORNO, DURÁN, HAZBÚN, CASTRO, & IVANOVIC, 2000).

En un estudio realizado por Delgado, Escurra, Bulnes & Quesada (2001), se procuraba evaluar la adecuación psicométrica de los ítems que conformaban el test de Matrices Progresivas Avanzadas de Raven, en una muestra de 501 alumnos de nuevo ingreso de la Universidad Nacional Mayor de San Marcos (UNMSM). Los autores realizaron la siguiente categorización de los alumnos: área A (Química, Ciencias Biológicas, Física, Matemática, Estadística, Investigación Operativa), área B (Medicina Humana, Farmacia y Bioquímica, Odontología, Medicina Veterinaria, Obstetricia, Enfermería, Laboratorio Clínico y Anatomía Patológica, Terapia Física y Rehabilitación, Radiología, Terapia Ocupacional, Nutrición, Psicología), área C (Ingeniería Mecánica de Fluidos, Ingeniería Geológica, Ingeniería Geográfica, Ingeniería de Minas, Ingeniería Metalúrgica, Ingeniería Industrial, Ingeniería Electrónica, Ingeniería Química, Ingeniería de Sistemas, Ingeniería Eléctrica), área D (Contabilidad, Economía, Administración, Turismo, Negocios Internacionales) y por último área E (Derecho y Ciencias Políticas, Filosofía, Lingüística, Comunicación Social, Arte, Bibliotecología y Ciencias de la Información, Historia, Sociología, Antropología, Arqueología, Trabajo Social, Geografía, Literatura, Educación y Educación Física). Primeramente, un coeficiente $\alpha=0.75$ permite confirmar que la prueba, efectivamente es confiable. Para evaluar la validez de constructo, se aplicó un análisis factorial, del cual se desprende que los ítems se encuentran agrupados en 12 factores, mostrando una explicación del 64% de la varianza total. Además, se puede observar una mayor agrupación de ítems en los factores 1 y 2. Por otro lado, Delgado, Escurra, Bulnes & Quesada (2001) investigaron la existencia de diferencias significativas en los niveles de inteligencia de los alumnos de nuevo ingreso de cada categoría, encontrando niveles de inteligencia general estadísticamente significativos ($=137.09, p < 0.05$), sin embargo, es importante mencionar que la técnica estadística utilizada para llegar a dicha conclusión fue una H de Kruskal-Wallis, utilizada en estadística no paramétrica, es decir, para muestras menores a 30 casos. Asimismo, se compararon los niveles de

inteligencia de cada área contrastando una con otra. De los resultados obtenidos, obtuvieron que sólo entre las áreas A - E y C - D no existen diferencias significativas entre los puntajes; no obstante, para realizar los cálculos anteriormente mencionados, se utilizó la prueba estadística U de Mann Whitney para dos muestras independientes menores a 30 casos.

Casé, Neer & Lopetegui (2002) llevaron a cabo un estudio con una muestra de 988 sujetos, con edades de entre 13 y 18 años, de diferentes planteles educativos de enseñanza media de la ciudad de La Plata, Argentina. Esta investigación tuvo como propósitos construir nuevas normas percentiles adecuadas para la población de La Plata, comparar las puntuaciones obtenidas por los sujetos de 1997-1998 con los anteriormente publicados de 1964, comparar las medias obtenidas tanto en las MPG, como en las MPA y, por último, comparar puntuaciones medias entre los sujetos que cuentan con niveles instruccionales distintos. Se aplicaron las MPG y MPA; de esta última se aplicaron las series I y II únicamente a 724 de los 988. De acuerdo a los resultados derivados de los análisis estadísticos, constatan la aparición del efecto Flynn, observándose un incremento de puntuaciones significativas para las muestras 1997-1998 respecto de la de 1964. Los incrementos más notables se pueden apreciar, especialmente, para los sujetos ubicados en los percentiles más bajos (5, 10, 25, 50). Al comparar las medias de los puntajes según la modalidad educativa (sujetos provenientes de escuelas técnicas, medias, pertenecientes a la universidad y privadas), Casé, Neer & Lopetegui (2002) refieren que la muestra del bachillerato de Bellas Artes obtiene el promedio más alto, lo cual indicaría que las características de las MPR están relacionadas con el nivel de desarrollo de habilidades particulares de los sujetos. Cabe destacar que el estudio no comprende un análisis de dificultad y discriminación de los ítems, un punto importante que pudiera afectar el cálculo de las normas percentiles.

Con una muestra de 591 escolares con edades de entre 12 y 21 años, pertenecientes a una institución de educación media, de 1º a 5º curso, Fernández, Ongarato, Saavedra & Casullo (2004) analizaron las propiedades psicométricas del test MPR para determinar su eficacia. Al realizar un análisis de dificultad de los ítems, observaron que los primeros ítems de la mayoría de las series no cuentan con un nivel de complejidad adecuado, debido a que el porcentaje de sujetos que responde de manera exitosa es muy elevado, sin mencionar que algunos de los últimos ítems incrementan su nivel de dificultad abruptamente, los ejemplos más extremos se presentan en la serie A, donde los ítems 11 y 12 descienden hasta un 7% y 5% de sujetos que respondieron con éxito, cuando en los ítems anteriores los porcentajes no descienden de 90%. En la serie D, el ítem 12 desciende a un 17% y en la serie E los ítems 11 y 12 descienden a 7% y 15% respectivamente. Otro aspecto que se puede observar es que los ítems no se encuentran ordenados de manera ascendente según su dificultad. Encontraron, también, que los índices de discriminación presentan coeficientes negativos o cerca-

nos a cero, tales como A2, A11, A12, B1, E7, E8, E9, E10 y E11. Por otro lado, se obtuvo un coeficiente $\alpha = .85$, indicando una adecuada confiabilidad. A partir de los resultados encontrados, se propone el uso de la versión abreviada I que consta de 13 reactivos (C3, D2, D3, D4, D5, D6, D7, D8, E1, E2, E3, E4, E5) que presentan una mejor homogeneidad, con un coeficiente $\alpha = .81$.

Por otro lado, Rushton, Skuy & Bons (2004) llevaron a cabo un estudio que tenía como finalidad determinar si la validez de constructo era la misma tanto para estudiantes africanos como para no africanos. La investigación se llevó a cabo en una muestra de 306 alumnos universitarios de entre 17 y 23 años de la facultad de ingeniería de la University of The Witwatersrand. Al calcular el coeficiente Alfa de Cronbach, obtuvieron $\alpha = .86$ para toda la muestra ($n=306$); lo mismo se observa para el cálculo en africanos ($n=177$) donde se obtuvo $\alpha = .86$; mientras que para los sujetos del este de India ($n=57$) se observa $\alpha = .79$; y por último, para los sujetos de raza blanca ($n=72$) se obtuvo $\alpha = .75$. De acuerdo a los datos anteriores, se confirma una adecuada confiabilidad del test. En un análisis de varianza (ANOVA), donde se contemplaron raza y sexo como factores, no encontraron diferencias significativas en lo que concierne a sexo, sino únicamente en cuanto a raza $F(2,300) = 8.34 p < 0.01$.

Mackintosh & Bennett (2005), utilizaron 97 estudiantes (48 de sexo masculino y 49 de sexo femenino) de dos colegios distintos (equivalentes al American high school senior) de Cambridge, con edades de entre 17 y 18 años, para determinar diferencias significativas entre hombres y mujeres a través de los puntajes medios obtenidos en las MPA. Los datos obtenidos en el estudio sugieren que sí existen diferencias significativas entre los puntajes obtenidos por hombres y mujeres, confirmado por $t = 2.15$ y $p < 0.05$. Además, con base en otro estudio anteriormente realizado, se revisaron los puntajes obtenidos tanto por hombres como por mujeres, en cuatro formas distintas de solución a los ítems: Progresión por parejas, distribución de tres, adición o sustracción y distribución de tres¹. De esto se desprende que los ítems para los que se requieren soluciones de tipo distribución de dos y adición/sustracción, presentan diferencias significativas en contraste con los tipos progresión de pares y distribución de tres, lo cual sugiere que en ese tipo de ítems se podría estar manifestando la presencia, al menos parcialmente, de distintos procesos u operaciones mentales; esto se ha confirmado a través de un análisis factorial donde las cargas factoriales se aprecian en más de un factor.

En otra investigación, que se llevó a cabo con 287 estudiantes de la carrera de Orientación, con edad promedio de 21 años y provenientes de la Universidad de Zulia, se tenía como objetivo revisar la confiabilidad y validez del test MPR. Para determinar la confiabilidad se utilizaron las técnicas de

¹ En un estudio, Carpenter, Just & Shell (1990) determinaron que existían ciertas reglas, comunes a ciertos ítems, para solucionar los problemas planteados por las MPA en la sección II, estas soluciones fueron: 1) Constante en una hilera, 2) distribución de tres, 3) progresión de pares, 4) adición/sustracción y 5) distribución de dos. Si se desea consultar la fuente, véase Carpenter, P. A., Just, M. A., & Schell, P. (1990). «What one intelligence test measures: A theoretical account of the processing in the Raven's progressive matrices test». *Psychological Review*, 97, 404-431.

división por mitades y (se intuye, debido a la no especificación) Alfa de Cronbach. En cuanto a la validez, se utilizaron los tests de Otis, Purdue y Dominó, así como el rendimiento académico, como criterios de contraste. A través de esto, se encontró que las MPR tienen una confiabilidad de 0.79 utilizando la técnica de división por mitades, mientras que con el coeficiente Alfa de Cronbach se obtuvo una correlación de 0.90. Por otro lado, los coeficientes de validez van de moderado a bajo, mostrando correlaciones de 0.56 contrastado con los tests de Otis y Purdue, mientras que con el test de Dominó se obtuvo una correlación de 0.61. El coeficiente de correlación más bajo fue el obtenido con el rendimiento escolar, de 0.28 (SÁNCHEZ & PIRELA, 2009).

Con una muestra de 249 niños clasificados por sus escuelas como niños con dificultades de aprendizaje, con edades de 6 a 11 años y provenientes de distintas escuelas ubicadas en las provincias de Matanzas y Cienfuegos, Cuba; Ramírez, Lorenzo & Díaz (2013) encontraron un coeficiente Alfa de Cronbach de 0.96, lo cual sugiere una confiabilidad de las MPC bastante adecuada. Además, la forma abreviada construida por los autores a partir de las series A, Ab y B se comparó con las puntuaciones de la aplicación completa, encontrando que las puntuaciones se relacionaban: $t = 6.205$ $p < 0.001$. A través de la aplicación de una regresión lineal, se determinó un 93% de probabilidad para predecir resultados similares en la aplicación de las MPR y su forma abreviada.

Con la finalidad de comprobar la viabilidad del uso de las normas obtenidas en las estandarizaciones en Colchester, Inglaterra, en una población de adultos mayores en Chile, Alarcón, Díaz, Hernández & Estrada (2012) llevaron a cabo un estudio con 102 participantes con edades de entre 60 y 86 años, con una media de 71 años. A partir de sus análisis, encontraron que las medias obtenidas por los adultos examinados contrastadas con las medias disponibles comprendían diferencias importantes en la mayoría de los rangos percentiles. En otra comparación de los puntajes brutos de acuerdo a grupos de edad (60 a 70 años y mayores de 70 años), se observó que la mayoría de las medias de puntajes para las normas percentiles disponibles diferían en comparación con las medias de los participantes, por lo que los autores concluyeron que las normas de las MPR no eran adecuadas para ser utilizadas en adultos mayores de la ciudad de Punta Arenas, Chile.

Como se puede observar, los datos acerca de las propiedades psicométricas del test MPR varían de acuerdo a la población en la que se aplica. La importancia de comprender la pertinencia del uso de este test en determinadas poblaciones, radica en la extensión de su uso. Las MPR, además de aplicarse como una herramienta de apoyo, se utiliza como instrumento de medición en otras investigaciones (en un principio se diseñó con esa finalidad). A continuación se mencionan algunas investigaciones en las cuales se han utilizados las MPR para realizar mediciones.

3.6 APLICACIÓN DEL TEST EN DIVERSOS GRUPOS

Como ya se ha mencionado, en un principio las MPR fueron diseñadas con la finalidad de servir como instrumento de medición en investigación. Sin embargo, su uso se diversificó y actualmente se utiliza tanto para investigación como en varios contextos profesionales, donde su función es de herramienta de apoyo para la toma de decisiones.

Un estudio muy reciente, realizado por González & Hernández (2013), se llevó a cabo con la finalidad de crear una batería de pruebas para la selección de alumnos de nuevo ingreso a la universidad. Entre las pruebas que utilizaron se encuentra el test MPR y se determinó que el test era adecuado para ser utilizado en la selección de aspirantes a ingenierías y licenciaturas.

Otro estudio en el cual se utilizó el test MPR, o más bien una adaptación del test MPR, fue el realizado por Elisondo & Donolo (2010). Este estudio tenía como finalidad determinar si existía una relación entre inteligencia y creatividad. Para lo cual se contrastaron los puntajes obtenidos en las MPR junto con los de otras dos pruebas llamadas: CREA (para medir creatividad) y TEI (también para medir inteligencia).

Garaigordobil (1999) realizó una investigación con el propósito de averiguar si existía relación entre una amplia gama de funciones motrices y la inteligencia verbal y no verbal. Para la evaluación del desarrollo psicomotriz, la autora fabricó una escala que incluía 39 juego a través de los cuales se pueden recoger indicadores de conducta observable. Para medir la inteligencia utilizó las pruebas WISC-R, Toni-2 y el test MPR. Los puntajes de las pruebas de inteligencia se correlacionaron con los de la escala psicomotriz y con base en ello se determinó el grado de relación entre las funciones psicomotrices y la inteligencia en sus componentes verbal/no verbal.

En un estudio, con una muestra de 190 estudiantes de secundaria, se probó determinar la relación existente entre la inteligencia, los procesos de pensamiento y el rendimiento escolar. Para realizar mediciones del rendimiento escolar se emplearon pruebas de conocimientos generales, de matemáticas y de lenguaje, mientras que para los procesos de pensamiento se utilizaron cinco problemas planteados por Rimoldi². Para medir inteligencia se utilizó el test de MPR. Para averiguar si había o no relación entre inteligencia, procesos de pensamiento y rendimiento escolar, las puntuaciones obtenidas por los alumnos en las tres mediciones se correlacionaron (Cortada, 1998).

Como se puede ver, actualmente las MPR siguen siendo utilizadas en diversas investigaciones para realizar mediciones. Los estudios anteriores son sólo algunos ejemplos de ello. No obstante,

² La fuente original no se especifica en el estudio de Cortada (1998), pero si se desean conocer los estudios sobre resolución de problemas del psicólogo Horacio Rimoldi, se puede consultar: Rimoldi, H. (1984). «Sobre solución de problemas: teoría, metodología y experimentación». *Revista de Psicología General y Aplicada*, 39(1), págs. 75-96 y Rimoldi, H. (1997). «Estructuras e imágenes en la solución de problemas». *Revista de Psicología General y Aplicada*, 50(3), págs. 285-296.

existen más investigaciones donde se ha utilizado el test de Raven para medir inteligencia general y en ocasiones se utiliza como un test de rendimiento escolar (Por ejemplo: GONZÁLEZ, CASTRO, & GONZÁLEZ, 2008; ESPINAL, MARTÍNEZ, MELO, CORDERO, & DE JESÚS, 2007; DE BORTOLI, BARRIOS, & AZPIROZ, 2002; PADIERNA, OSEGUERA, & GUDIÑO, 2009).





• CAPÍTULO 4

Método

En este capítulo se incluyen todas las consideraciones para la realización del presente estudio, como las hipótesis planteadas y los objetivos, así como las características generales de la muestra y las razones por las cuales se ha centrado el interés en las Matrices Progresivas Escala General.

4.1 JUSTIFICACIÓN

Los instrumentos psicológicos para la medición de conductas, como ya se ha mencionado con anterioridad, son herramientas de apoyo sumamente importantes para la práctica del psicólogo en diversos escenarios profesionales, ya sean clínicos, empresariales o educativos. Asimismo, los tests psicológicos son necesarios para las investigaciones con enfoques cuantitativos, donde se pretende realizar análisis estadísticos. Estos datos extraídos de los análisis deben ser fiables y válidos, porque de ello depende la veracidad de las conclusiones a las que se llega a través de una determinada investigación; sin mencionar que una de las ventajas de utilizar instrumentos de medición psicológica válidos y confiables supone economizar en tiempo y dinero.

Existen diversas organizaciones que publican normas sobre el comportamiento ético para el desarrollo y uso responsable de las pruebas: American Psychological Association, American Educational and Research Association y, en México, la Sociedad Mexicana de Psicología, encargada del Código Ético del Psicólogo. A continuación se citan 5 de sus artículos que hacen especial énfasis en la valoración psicológica.

Art. 17. El psicólogo que administra, califica, interpreta o usa técnicas de valoración, se cerciora de que éstas se basan en datos sólidos que garanticen la confiabilidad, validez y normas, así como la aplicación apropiada y usos de las técnicas y/o instrumentos que emplea. Asimismo, toma decisiones reconociendo los límites de la certidumbre con la que es posible diagnosticar, emitir juicios o hacer predicciones acerca de individuos o grupos (Sociedad Mexicana de Psicología, 2009, pág. 51).

Art. 18. El psicólogo identifica situaciones en las que ciertas técnicas de valoración o normas no son aplicables o requieren de ajustes para su administración o interpretación, debido a factores tales como el género, la edad, la raza, el origen étnico, el origen nacional, la orientación sexual, la discapacidad, el idioma o el nivel socioeconómico de los individuos

o grupos. No emplea técnicas o instrumentos solamente traducidos de otro idioma y/o cuyos datos acerca de su construcción, adaptación, validez, confiabilidad, estandarización y/o investigación de resultados que se hayan determinado con base en estudios realizados en poblaciones diferentes a la que pertenezcan los individuos o grupos que pretenda valorar (Sociedad Mexicana de Psicología, 2009, pág. 51).

Art. 19. El psicólogo que ofrece servicios de valoración a individuos pertenecientes a los pueblos indígenas de México u otros grupos, emplea técnicas de valoración y normas elaboradas ex profeso o adaptadas debidamente para esa población, y que tomen en cuenta su contexto cultural (Sociedad Mexicana de Psicología, 2009, pág. 51).

Art. 22. El psicólogo no basa sus decisiones de valoración o intervención, o sus recomendaciones, en datos, normas o resultados de pruebas psicológicas obsoletos para propósitos actuales. De manera similar, no basa tales decisiones o recomendaciones en pruebas psicológicas o medidas inútiles, o construidas con otros propósitos (Sociedad Mexicana de Psicología, 2009, pág. 52).

Art. 24. El psicólogo que ofrece valoración o procedimientos de calificación a otros profesionales describe con precisión el propósito, las normas, la validez, la confiabilidad, y las aplicaciones de los procedimientos, así como cualquier característica especial aplicable a su uso. El psicólogo selecciona servicios de calificación e interpretación (incluyendo servicios automatizados) apegados estrictamente a las normas de calidad científica que garanticen la validez y confiabilidad del programa y procedimientos. Se cerciora que tales programas y procedimientos estén basados en estudios que se hubiesen realizado en las poblaciones a las que pertenezcan los individuos bajo valoración, así mismo basa su opinión en otras consideraciones pertinentes (Sociedad Mexicana de Psicología, 2009, pág. 52).

El interés del presente estudio se centra en las MPG debido a que es un instrumento utilizado en diversos escenarios, dirigido a diversas personas de distintos países y de distintas edades y que, a pesar de que su construcción tiene ya más de 70 años (RAVEN, 2004), actualmente sigue vigente. Se siguen publicando estudios realizados con las MPR (CORTADA, 1998; GARAIGORDOBIL, 1999; DE BORTOLI, BARRIOS, & AZPIROZ, 2002; ESPINAL, MARTÍNEZ, MELO, CORDERO, & DE JESÚS, 2007; GONZÁLEZ, CASTRO, & GONZÁLEZ, 2008; PADIARNA, OSEGUERA, & GUDIÑO, 2009; ELISONDO &

DONOLO, 2010; GONZÁLEZ & HERNÁNDEZ, 2013) y se siguen realizando estudios sobre las propiedades psicométricas del test (TORRES & CUESTA, 1992; CHICO, 1997; CAIRO, CAIRO, BOUZA, & PONCE, 2000; IVANOVIC, FORNO, DURÁN, HAZBÚN, CASTRO, & IVANOVIC, 2000; DELGADO, ESCURRA, BULNES, & QUESADA, 2001; CASÉ, NEER, & LOPETEGUI, 2002; FERNÁNDEZ, ONGARATO, SAAVEDRA, & CASULLO, 2004; RUSHTON, SKUY, & BONS, 2004; MACKINTOSH & BENNETT, 2005; SÁNCHEZ & PIRELA, 2009; entre otros).

El propósito de este estudio es revisar si las MPR son confiables y válidas para su aplicación en estudiantes mexicanos, por lo tanto se responde a la pregunta ¿las MPG, a partir de su traducción al español, son aplicables y válidas en población mexicana?

La edición del manual utilizada para aquel propósito es la primera de 1993, en su tercera reimpresión del 2004 en Buenos Aires, Argentina. La traducción del manual de las MPR fue realizada por Jorge Piatigorsky (RAVEN, 2004).

A diferencia de las MPC, compuestas por tres series (A, Ab y B) y de las MPA, compuestas por las secciones I y II, las MPG están compuestas por 5 series con 12 problemas cada una. Fueron las primeras en desarrollarse y además abarca una amplia gama de edades; Mientras que las MPC y MPA fueron diseñadas para utilizarse específicamente en sujetos con ciertas edades o niveles de rasgo. De acuerdo con esto, se consideró a las MPG como la escala más adecuada para ser aplicada en la muestra de esta investigación, porque abarca los rangos de edades contemplados para esta investigación.

Es relevante averiguar si el test realmente sigue siendo adecuado para medir aquello para lo que fue diseñado. De ello dependen los resultados de futuros estudios y las decisiones tomadas por los profesionales de la salud mental que optan por utilizarlo como herramienta de apoyo.

De acuerdo con la bibliografía revisada, los estudios realizados en Latinoamérica parecen variar en relación a los coeficientes de confiabilidad y validez encontrados en la prueba. No obstante, en aquellos estudios donde se han revisado los índices de dificultad de los ítems, parece que existe una confluencia de datos (por ejemplo, CHICO, 1997; CAIRO, CAIRO, BOUZA, & PONCE, 2000; TORRES & CUESTA, 1992; IVANOVIC, FORNO, DURÁN, HAZBÚN, CASTRO, & IVANOVIC, 2000), los cuales han encontrado ítems invariables y un orden de dificultad distinto al esperado (en orden ascendente).

4.2 DISEÑO DEL ESTUDIO

La presente investigación parte de un diseño no experimental transeccional exploratorio.

4.3 MUESTRA

La muestra del presente estudio es de tipo no probabilístico y está constituida por participantes voluntarios (HERNÁNDEZ, FERNÁNDEZ, & BAPTISTA, 2006). La razón por la cual no se haya llevado a cabo un muestreo probabilístico se debió a las circunstancias. El mismo mes en el que se planearon las aplicaciones, los alumnos tenían evaluaciones y por lo tanto los directivos no pudieron ofrecer disponibilidad total de los alumnos, sino que los mismos directivos ofrecieron los grupos más disponibles, es decir, aquellos que no tenían evaluación, entrega de trabajos finales o entrega de calificaciones.

4.4 OBJETIVOS

Objetivo: revisar el nivel de confiabilidad y validez de las MPG en una población mexicana.

4.5 HIPÓTESIS

Hipótesis: el test MPG tendrá validez y confiabilidad al aplicarse en una población mexicana de alumnos de la preparatoria No. 1 de Cuernavaca, Morelos.

4.6 DEFINICIÓN DE VARIABLES

Variable independiente: población mexicana

Definición conceptual de población mexicana: «Está conformada por personas con características diversas, que comparten entre otras cosas el territorio, la historia, el idioma y algunas costumbres» (INEGI, 2015).

Definición operacional de población mexicana: alumnos de la preparatoria No. 1 de Cuernavaca, Morelos.

Variables dependientes: validez y confiabilidad

Definición conceptual de validez: «Se refiere al grado en que un instrumento realmente mide la variable que pretende medir» (HERNÁNDEZ, FERNÁNDEZ, & BAPTISTA, 2006, pág. 277).

Definición operacional de validez: a través del análisis factorial que determina el número de agrupaciones y el total de la varianza explicada.

Definición conceptual de confiabilidad: «el grado de estabilidad, previsibilidad, exactitud y consistencia con la que se realiza una evaluación» (BARRON, BROWN, EGAN, GESUALDI, & MARCHUK, 2008, pág. 29).

Definición operacional de confiabilidad: se calcula a través del coeficiente Alpha de Cronbach.

4.7 INSTRUMENTO

MPG en su 1ª edición de 1993, 3ª reimpresión de 2004 y traducidas por Jorge Piatigorsky. El test consta de 5 series (A, B, C, D, E), cada una con 12 problemas distintos, que en conjunto conforman un total de 60 problemas a resolver. Dentro de las series A y B, cada matriz cuenta con 6 posibles opciones de respuesta; dentro de las series restantes, cada matriz cuenta con 8 posibles respuestas. Al examinado se le entrega un protocolo de respuesta donde deberá registrar cada opción elegida. La consigna para iniciar la aplicación es: «Abran los cuadernos en la primera página. Van a ver algo como esto» (RAVEN, 2004) y se les muestra a los voluntarios el cuadernillo abierto con el primer problema (A1) o, en su defecto, una ampliación de la matriz.

Arriba dice serie A y en sus hojas de respuestas tienen una columna, aquí, destinada a esa serie A. Ésta es la prueba A1. Ya ven de qué se trata. Arriba hay un dibujo en el que se ha suprimido una parte. Cada una de estas figuras de abajo (RAVEN, 2004, MPG, pág. 24).

Se señalan las opciones una por una, «tienen una forma que se ajusta en el espacio en blanco, pero no todos completan bien el dibujo. La N° 1 [se muestra una opción y después el problema planteado] es totalmente errónea. Los números 2 y 3 tampoco sirven –ajustan en el espacio pero no completan bien el dibujo-. ¿Y la N° 6? Ésta tiene el dibujo correcto [se señala que el tipo de dibujo es idéntico al de la matriz] pero con una parte en blanco. [Se señala entonces la opción correcta y se observa si todos han escrito la respuesta]». En caso de que se tengan dudas sobre la comprensión de las instrucciones, se deben agregar otras explicaciones. Después, se les dice: «Sí, la N° 4 es la figura correcta. De modo que la respuesta a A1 es 4. Escriban el 4 aquí, junto al número 1 en la columna A de la hoja de respuestas» (RAVEN, 2004). Ya que se ha mostrado la forma de trabajar, se recalca que es la forma de resolver todos los casos. Además, se les indica que no deben escribir nada sobre el cuadernillo donde están los problemas. Se explica que la dificultad de los problemas irá en orden ascendente, que deben tratar de resolver los problemas sin saltar y sin regresar a problemas anteriores. Se debe comunicar que no tienen tiempo límite para resolver todos los problemas y, finalmente se observa si el problema A2 todos lo han resuelto correctamente, con la finalidad de asegurar que las instrucciones se han comprendido correctamente. Las puntuaciones obtenidas en cada serie se suman verticalmente y después se suman los totales, con lo que se obtiene el puntaje bruto. En las tablas de los percentiles se busca la edad del sujeto, debajo de la edad se busca el puntaje bruto obtenido y posteriormente se ubica el percentil y con ello se puede obtener el rango en el que se posiciona el sujeto (I, II, III, IV, V), el cual es el diagnóstico.

4.8 PROCEDIMIENTO

El presente estudio se llevó a cabo en la preparatoria N° 1 Lic. Bernabé L. de Elías de Cuernavaca, Morelos. Inicialmente se consumó una reunión en la que estuvo presente la Directora del plantel y el Secretario Académico. En esa reunión, la Directora del plantel y el Secretario Académico solicitaron la presentación de un oficio formal dirigido a la Directora, con la finalidad de mantenerla al tanto de lo que se pretendía hacer. Al cabo de una semana la Directora informaría su respuesta vía telefónica a través del Secretario Académico.

Posteriormente, se presentó un oficio donde se solicitaba el permiso de los directivos y se les explicaba a grandes rasgos en qué consistía la investigación, el manejo ulterior de los datos que se obtendrían y la finalidad con la que se realizaba el estudio. Después de la fecha pactada, se realizó una llamada dirigida al Secretario Académico para conocer la respuesta de la Directora, quien amablemente accedió y solicitó una nueva reunión para concretar la organización de la aplicación.

Se acordó que los directivos determinarían los grupos de alumnos a los cuales se tendría acceso y a partir de ello, se solicitó la participación voluntaria de los alumnos de manera verbal, debido a las circunstancias. También se acordaron los días y horas en las que se llevarían a cabo las aplicaciones de la prueba.

Las aplicaciones de la prueba se llevaron a cabo en tres diferentes salones y en un auditorio. Uno de los profesores del plantel era quien indicaba los salones en los que se realizarían las aplicaciones. Dentro de los salones, se explicaba a los alumnos el propósito de la aplicación (velando cierta información, con la finalidad de no predisponer a los examinados). Se preguntó a los alumnos sobre su participación y se invitaba a salir del aula a aquellos que no deseaban ser evaluados, aunque en ninguno de los grupos se negaron a participar.

Durante las aplicaciones, se encontraban presentes el investigador y una colaboradora. El investigador dirigió la aplicación, informando la naturaleza de los problemas del test y dando instrucciones, mientras que la colaboradora ayudaba a repartir el material de la prueba y a vigilar que los alumnos no copiaran respuestas de sus compañeros o para asistirlos en caso de que los alumnos tuvieran dudas.

Para determinar la confiabilidad de la prueba se aplicó una prueba Alfa de Cronbach, tanto de manera general como por serie, por otro lado, para averiguar la validez de constructo se realizó un análisis factorial. Este análisis se le aplicó al total de los ítems que componen las MPG y a cada una de las series por separado. Los análisis estadísticos se realizaron a través del programa informático Statistical Package for the Social Sciences en su versión 18.

4.9 CONSIDERACIONES ÉTICAS

Para garantizar la confidencialidad de los participantes y de la información proporcionada por éstos durante las aplicaciones, se les informó a los participantes que evitaran proporcionar su nombre y apellidos dentro de los protocolos de respuesta. Naturalmente, los datos demográficos son importantes para el presente estudio y para las investigaciones similares que posteriormente se pretendan realizar. Así, los datos demográficos recabados fueron: Sexo, edad, lugar de procedencia, estado civil, grado y turno al que pertenecían.

A los participantes se les explicó, previo a la aplicación, el propósito de la investigación. Sin embargo, con la finalidad de evitar que los participantes tomaran una actitud de competencia al saber que la prueba presentada fue diseñada con el propósito de obtener un CI, el uso que se haría de la información recogida se reveló al final de cada aplicación. Ahí se les comunicó la naturaleza del test aplicado y, también, se les mencionó que la información obtenida no se utilizaría con otros fines que no fueran para la presente investigación, asimismo, se les informó que se encontraban en libertad de solicitar la dimisión de su participación (a través de la eliminación de sus datos) en el estudio. Finalmente, se dijo a los participantes que no se comunicarían los resultados obtenidos en la prueba debido a la naturaleza de la investigación en curso. Mas se agradeció atentamente por su participación incondicional.

Con respecto a la colaboradora que ayudó durante las aplicaciones, se le instruyó previamente en relación a las tareas que debería desempeñar, las cuales consistieron en repartir y recolectar los materiales para la aplicación, supervisión de las aplicaciones y posibles maneras de actuar frente a situaciones no contempladas para la aplicación, pero sí previstas. La colaboradora es egresada de la carrera de psicología, lo cual la hacía una candidata más viable para ayudar en las aplicaciones.





• CAPÍTULO 5

Resultados

En este capítulo se han añadido los análisis estadísticos pertinentes para conocer las características de la muestra utilizada en esta investigación, además de las pruebas pertinentes para comprobar tanto la confiabilidad como la validez de las Matrices Progresivas de Raven Escala General.

De acuerdo con los resultados obtenidos, el promedio de edad de la muestra de la presente investigación es de 16.32 años, con una desviación estándar de 1.051. De los 123 participantes, el 53.7% de los sujetos son de sexo femenino, mientras que el 46.3% son de sexo masculino. La mayoría de los sujetos pertenecían al primer grado de preparatoria (39.8%), el resto cursaba el segundo (24.4%) y tercer grado (35.8%). El 92.7% de los examinados radica en Cuernavaca, el 4.1%, 1.6%, 0.8% y 0.8% pertenecen a Jiutepec, Temixco, Tepoztlán y Yautepec, respectivamente. Estos datos se pueden examinar en las representaciones graficas que se ofrecen a continuación.

Figura N° 7

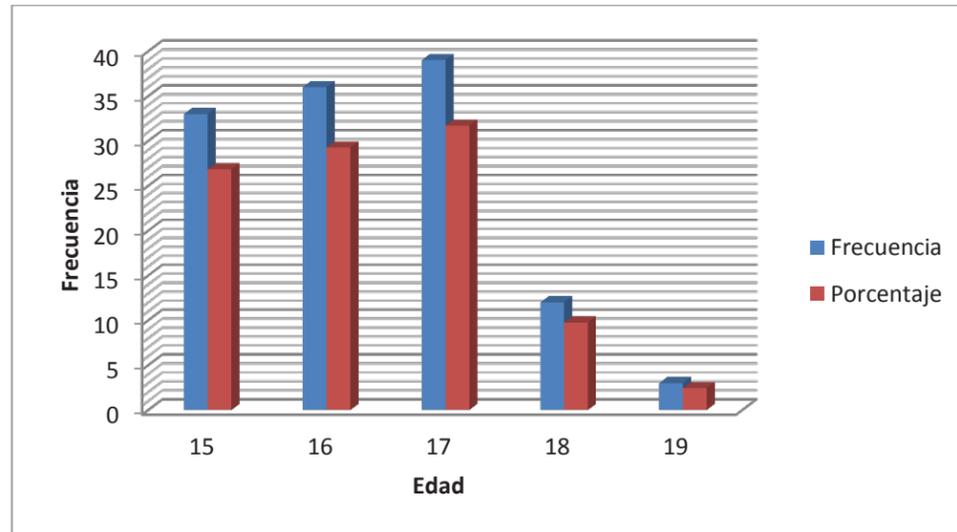


Figura N° 8

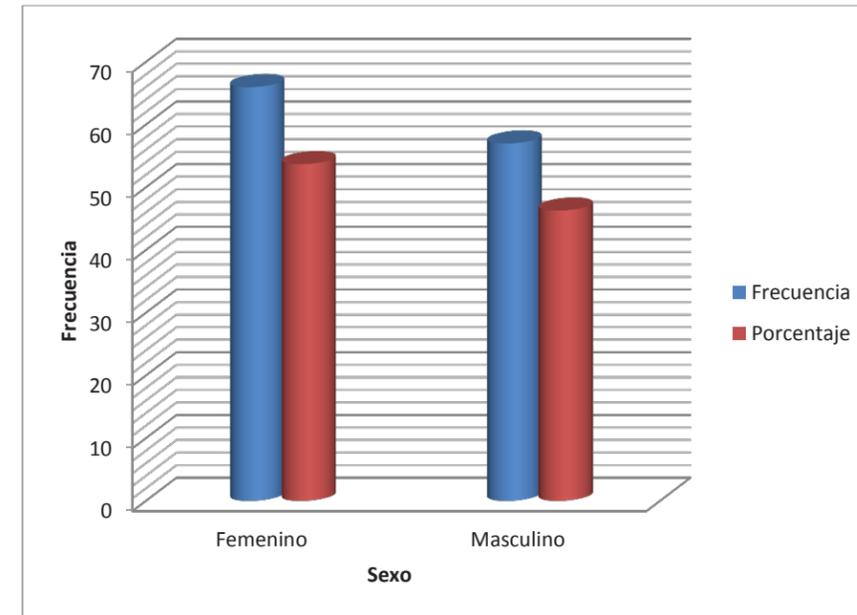
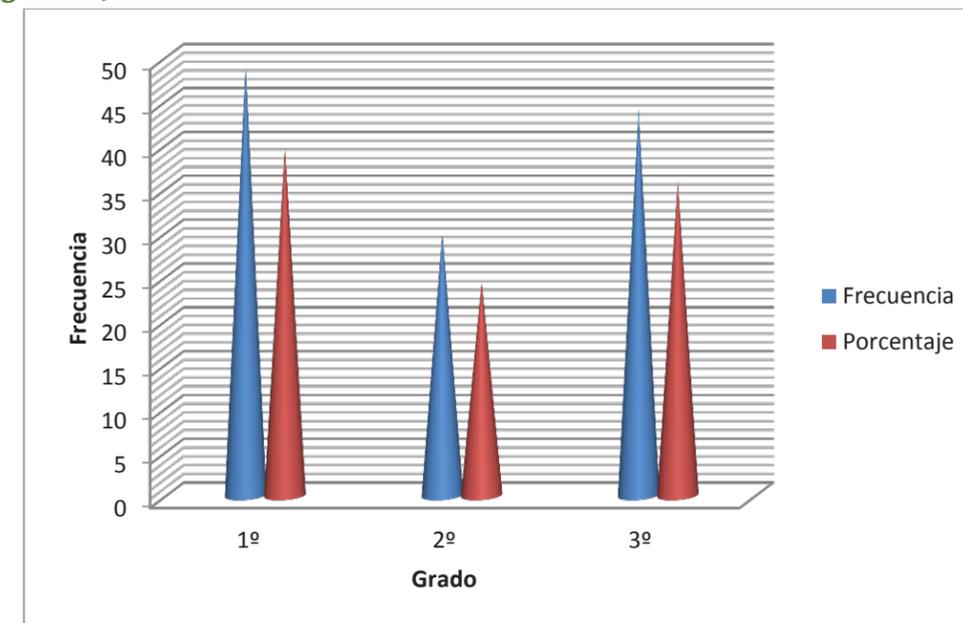


Figura N° 9



En primer lugar, para determinar la confiabilidad del test MPG, se aplicó una prueba Alfa de Cronbach. El coeficiente Alfa de Cronbach arrojado por la prueba MPG, al incluir la totalidad de sus ítems, fue de .799, un coeficiente de confiabilidad que se puede considerar bastante adecuado.

Tabla N° 9. Estadísticos de confiabilidad

ALFA DE CRONBACH	NÚMERO DE ÍTEMS
,799	60

Nota: Coeficiente Alfa de Cronbach al incluir el total de los ítems.

Sin embargo, también se aplicó la prueba Alfa de Cronbach a cada una de las escalas de las que se componen las MMPGPG, de lo cual se extrajeron los siguientes resultados: En la serie A se obtuvo un Coeficiente de confiabilidad de .549, el cual se puede considerar como un coeficiente moderado, aunque apenas se acerca a un mejor coeficiente.

Tabla N° 10. Estadísticos de confiabilidad serie A

ALFA DE CRONBACH	NÚMERO DE ÍTEMS
,549	6

Nota: Estadísticos de confiabilidad al incluir 6 ítems de la serie A.

Como se podrá observar, no se ha incluido el número total de los ítems de los que se compone la serie A, sin embargo, más adelante se explicará la razón de esta exclusión intencional. En la siguiente tabla se pueden observar los valores Alfa si se eliminara un elemento de la serie A.

Tabla N° 11. Valores Alfa si se elimina un elemento de la serie A

	MEDIA DE LA ESCALA SI SE ELIMINA UN ELEMENTO	VARIANZA DE LA ESCALA SI SE ELIMINA UN ELEMENTO	CORRELACIÓN ELEMENTO-TOTAL CORREGIDA	ALFA DE CRONBACH SI SE ELIMINA UN ELEMENTO
A7	4,3902	,797	,306	,517
A8	4,4065	,719	,440	,463
A9	4,4065	,702	,495	,445
A10	4,4472	,659	,385	,460

A11	4,4553	,709	,238	,528
A12	4,7236	,530	,207	,636

Nota: Valores Alfa si se elimina un elemento con 6 ítems de la serie A

Como se puede observar en la tabla 11, si se eliminara el ítem A12 el coeficiente Alfa sufriría un cambio positivo. Por otro lado, el coeficiente Alfa que se obtuvo de la serie B fue de .495, a pesar de que el número de ítems que se incluyó fue mayor que en el de la serie A.

Tabla N° 12. Estadísticos de confiabilidad serie B

ALFA DE CRONBACH	NÚMERO DE ÍTEMS
,495	8

Nota: Coeficiente Alfa de Cronbach con 8 ítems de la serie B.

El coeficiente Alfa obtenido en esta serie es aún más bajo que el de la serie A, aunque podría mejorar un poco si se elimina el ítem B12, como se puede observar en la tabla N° 13.

Tabla N° 13. Valores Alfa si se elimina un elemento de la serie B

	MEDIA DE LA ESCALA SI SE ELIMINA UN ELEMENTO	VARIANZA DE LA ESCALA SI SE ELIMINA UN ELEMENTO	CORRELACIÓN ELEMENTO-TOTAL CORREGIDA	ALFA DE CRONBACH SI SE ELIMINA UN ELEMENTO
B3	6,1220	1,255	,091	,499
B6	6,1626	1,186	,104	,498
B7	6,1382	1,186	,215	,475
B8	6,2195	,960	,376	,399
B9	6,2195	1,009	,287	,436
B10	6,2114	1,070	,200	,471
B11	6,3008	,851	,384	,381
B12	6,4228	,902	,187	,505

Nota: Valores Alfa si se elimina un elemento con 8 ítems de la serie B.

Con respecto a la serie C, el coeficiente de confiabilidad Alfa de Cronbach obtenido fue de .611. En esta serie se han incluido la totalidad de los ítems que componen la serie. Únicamente en esta serie y la serie E se han incluido todos los ítems de sus respectivas series para calcular el coeficiente Alfa.

Tabla N° 14. Estadísticos de confiabilidad serie C

ALFA DE CRONBACH	NÚMERO DE ÍTEMS
,611	12

Nota: Coeficiente Alfa de Cronbach al incluir 12 ítems de la serie C.

En este caso (de la serie C), el coeficiente Alfa podría incrementarse si se eliminara el elemento C12, lo que se puede corroborar en la tabla 15.

Tabla N° 15. Valores Alfa si se elimina un elemento de la serie C

	MEDIA DE LA ESCALA SI SE ELIMINA UN ELEMENTO	VARIANZA DE LA ESCALA SI SE ELIMINA UN ELEMENTO	CORRELACIÓN ELEMENTO-TOTAL CORREGIDA	ALFA DE CRONBACH SI SE ELIMINA UN ELEMENTO
C1	7,9431	3,087	,279	,589
C2	7,8862	3,315	,164	,607
C3	7,9350	3,176	,201	,601
C4	8,0163	2,688	,514	,536
C5	7,9106	3,246	,178	,604
C6	8,0000	2,951	,303	,582
C7	7,9675	3,032	,283	,587
C8	8,3740	2,629	,338	,574
C9	8,0000	3,033	,231	,596
C10	8,3577	2,658	,319	,579
C11	8,3089	2,707	,289	,588
C12	8,7805	3,304	,053	,622

Nota: Valores Alfa si se elimina un elemento de la serie C con 12 ítems.

En la serie D, para calcular el coeficiente Alfa, únicamente se omitió el ítem D1. El resultado obtenido en esta serie fue un Alfa de ,484, un coeficiente bastante bajo.

Tabla N° 16. Estadísticos de confiabilidad serie D

ALFA DE CRONBACH	NÚMERO DE ÍTEMS
,484	11

Nota: Coeficiente Alfa de Cronbach al incluir 11 ítems de la serie D.

En la serie D, nuevamente se puede apreciar un incremento de la confiabilidad, aunque ligeramente, si se elimina el ítem D12.

Tabla N° 17. Valores Alfa si se elimina un elemento de la serie D

	MEDIA DE LA ESCALA SI SE ELIMINA UN ELEMENTO	VARIANZA DE LA ESCALA SI SE ELIMINA UN ELEMENTO	CORRELACIÓN ELEMENTO-TOTAL CORREGIDA	ALFA DE CRONBACH SI SE ELIMINA UN ELEMENTO
D2	7,1220	2,206	,156	,474
D3	7,1545	2,033	,323	,436
D4	7,1707	2,110	,163	,468
D5	7,1626	2,154	,120	,477
D6	7,2114	2,070	,142	,473
D7	7,2520	1,862	,311	,420
D8	7,3171	1,710	,385	,385
D9	7,4065	1,702	,318	,410
D10	7,3333	1,765	,314	,414
D11	7,8455	1,984	,104	,495
D12	8,0000	2,410	-,213	,557

Nota: Valores Alfa si se elimina un elemento de la serie D con 12 ítems.

Para el cálculo del coeficiente Alfa de Cronbach, en la serie E se incluyó la totalidad de los ítems que componen la serie. La confiabilidad obtenida en este caso fue de ,627.

Tabla N° 18. Estadísticos de confiabilidad serie E

ALFA DE CRONBACH	NÚMERO DE ÍTEMS
,627	12

Nota: Coeficiente Alfa de Cronbach al incluir 12 ítems de la serie E.

Los resultados obtenidos con la serie E, indican que de eliminarse los ítems E1 y E12, el coeficiente de confiabilidad podría mejorar ligeramente.

Tabla N° 19. Valores Alfa si se elimina un elemento de la serie E

	MEDIA DE LA ESCALA SI SE ELIMINA UN ELEMENTO	VARIANZA DE LA ESCALA SI SE ELIMINA UN ELEMENTO	CORRELACIÓN ELEMENTO-TOTAL CORREGIDA	ALFA DE CRONBACH SI SE ELIMINA UN ELEMENTO
E1	5,3415	4,899	,094	,634
E2	5,5041	4,400	,286	,604
E3	5,3740	4,269	,481	,571
E4	5,6016	4,061	,434	,570
E5	5,5447	4,316	,314	,598
E6	5,5285	4,317	,320	,596
E7	5,8455	4,328	,301	,600
E8	5,5772	4,410	,253	,611
E9	5,6829	4,087	,406	,576
E10	5,9837	4,721	,149	,628
E11	6,1626	5,023	,114	,627
E12	6,0894	5,000	,043	,639

Nota: Valores Alfa si se elimina un elemento de la serie E con 12 ítems.

En segundo lugar, para determinar la validez de las MPG se aplicó un análisis factorial. De éste se desprende que algunos ítems pertenecientes a 3 de las 5 series contaban con varianza 0, lo que significa que el 100% de los participantes contestaron correctamente esos ítems. Específicamente son los ítems A1, A2, A3, A4, A5, A6, B1, B2, B4, B5 y D1. Es por esta razón que esos ítems fueron excluidos al momento de realizarse el cálculo del coeficiente de confiabilidad por serie. Entre los primeros resultados del análisis factorial se encuentra la tabla de comunalidades:

Tabla N° 20. Comunalidades

	INICIAL	EXTRACCIÓN
A7	1,000	,766
A8	1,000	,712
A9	1,000	,843
A10	1,000	,788
A11	1,000	,732
A12	1,000	,728
B3	1,000	,734
B6	1,000	,635
B7	1,000	,831

B8	1,000	,669
B9	1,000	,718
B10	1,000	,675
B11	1,000	,744
B12	1,000	,699
C1	1,000	,811
C2	1,000	,696
C3	1,000	,742
C4	1,000	,735
C5	1,000	,788
C6	1,000	,671
C7	1,000	,664
C8	1,000	,606
C9	1,000	,765
C10	1,000	,735
C11	1,000	,660
C12	1,000	,618
D2	1,000	,619
D3	1,000	,763
D4	1,000	,741
D5	1,000	,770
D6	1,000	,798
D7	1,000	,688
D8	1,000	,697
D9	1,000	,741
D10	1,000	,737
D11	1,000	,654
D12	1,000	,697
E1	1,000	,715
E2	1,000	,802
E3	1,000	,760
E4	1,000	,638
E5	1,000	,669
E6	1,000	,557
E7	1,000	,735
E8	1,000	,677
E9	1,000	,664
E10	1,000	,473
E11	1,000	,622
E12	1,000	,712

Nota: Tabla de comunalidades extraída del análisis factorial al incluir el total de ítems (excepto aquellos que no varían).

Como se puede observar, la mayoría de los ítems que componen las MPG explican adecuadamente los factores arrojados, excepto en el caso del ítem E10, el cual presenta un porcentaje muy bajo de explicación.

A través del análisis factorial se extrajeron 18 factores, los cuales explican un 70.799% de la varianza total. No obstante, es importante notar que el número de factores arrojados por el análisis factorial es mucho mayor que el esperado (debido a la cantidad de factores, la tabla se dividió en 3 partes con 6 factores cada una).

Tabla Nº 21. Matriz de componentes rotados

	COMPONENTES					
	1	2	3	4	5	6
A7	,701	-,116	,052	,034	,012	,118
A8	,766	-,061	,076	-,026	,006	,031
A9	,850	-,050	-,159	-,007	-,019	-,020
A10	,727	,102	,047	-,131	-,041	-,079
B3	-,014	,109	-,049	-,092	,813	-,055
B7	-,023	,085	,011	,108	,863	-,010
B8	-,037	,428	,208	,237	,000	,011
B9	,017	,043	,147	,073	,026	,518
B10	-,065	-,068	-,020	,512	,024	-,034
B12	,064	,510	,013	,121	,103	-,119
C2	-,078	-,110	-,064	-,066	,015	,747
C9	,206	,104	,102	-,091	-,071	,709
C11	,034	,212	,530	,049	-,124	-,119
D2	-,008	-,075	,083	,702	-,038	-,084
D3	-,053	,163	,025	,778	,042	,019
D5	,000	-,060	-,037	-,020	,477	,204
D6	,377	-,015	,217	,192	-,053	-,016
D7	-,096	,174	,072	,370	,024	,491
D8	-,013	-,035	,470	,360	,381	,109
D9	-,044	,122	,644	,175	-,049	,201
D10	,029	,088	,801	-,049	-,007	-,041
E3	-,008	,479	-,086	-,054	,139	,164
E4	,036	,645	,131	,008	,039	-,107
E5	-,023	,595	-,077	-,055	,075	,136
E6	-,170	,614	,225	,015	,111	,065

	COMPONENTES					
	7	8	9	10	11	12
B6	-,008	,130	,016	-,164	,397	,009
C1	-,059	,758	,062	-,032	,116	,075
C4	,346	,691	,138	,022	-,040	,100
C6	,688	,173	,070	-,024	,078	,120
C7	,430	,248	-,171	-,171	,004	-,092
C8	,669	-,025	-,061	-,055	,087	,029
C10	,077	,206	-,039	,129	-,046	,677
C12	-,088	-,055	-,119	,729	-,042	,054
D11	,058	,091	-,122	-,005	,706	-,027
E2	,048	,005	-,004	-,036	,064	,802
E7	,211	-,232	,239	,058	,348	,036
E9	,343	-,022	-,050	-,082	,549	,230
E11	,197	,023	,634	-,021	,112	-,148
E12	-,126	,111	,743	-,032	-,179	,080

	COMPONENTES					
	13	14	15	16	17	18
A11	,208	,054	,752	,035	,003	,092
A12	,785	-,004	,191	-,035	,001	-,039
B11	-,037	,766	,052	,017	,152	-,054
C3	-,116	-,105	,049	-,068	,028	,762
C5	,023	-,008	,091	,785	-,033	-,062
D4	,020	,112	-,008	-,054	,825	,035
D12	,181	,253	,043	,045	,048	,513
E1	-,282	,182	-,039	,473	-,155	,056
E8	,395	,325	-,261	,131	,158	-,001
E10	,201	,322	,193	,082	-,267	,077

Nota: Matriz de componentes rotados extraída del análisis factorial al incluir todos los ítems (excepto aquellos que no varían).

En este primer análisis se pueden observar aquellos factores en los que las cargas se han concentrado. En el factor 1 se encuentran los ítems A7, A8, A9, A10 y D6. Dentro del factor 2 se encuentran los ítems B8, B12, E3, E4, E5 y E6. En el factor 3 se agrupan los ítems C11, D8, D9 y D10. Las cargas más elevadas para el factor 4 son las de los ítems B10, D2 y D3. En el caso del factor 5, los ítems B3, B7 y D5 son los que presentan las cargas más altas. Los ítems B9, C2, C9 y D7 cargan en el factor 6. En el factor 7, están cargando los ítems C6, C7 y C8. Los ítems con las cargas más elevadas para el factor 8 son C1 y C4. En el factor 9, los únicos ítems que se están cargando son E11 y E12. En

el factor 10, sólo está cargando el ítem C12. Las cargas de los ítems B6, D11, E7 y E9 se agrupan en el factor 11. Los ítems C10 y E2 cargan en el factor 12. En el factor 13, los ítems que están cargando son A12 y E8. Las cargas de los ítems B11 y E10 se agrupan en el factor 14. En el factor 15 la única carga es la del ítem A11. En el factor 16, se agrupan las cargas de los ítems C5 y E1. En el factor 17 sólo carga el ítem D4. Finalmente, las cargas de los ítems C3 y D12 se agrupan en el factor 18.

En comparación con lo que se esperaba, es decir, 5 factores en los cuales deberían cargar los ítems de las diferentes series, el número de factores ha rebasado en gran medida la expectativa. Además, sólo los factores 1, 2, 3, 6 y 11 cuentan con un número de ítems que varía entre 4 y 6, el resto está compuesto por 3, 2 o 1 sólo ítem.

Con la finalidad de profundizar más, se realizaron análisis factoriales por cada una de las series que componen las MPG. De la serie A se extrajeron los siguientes datos:

Tabla N° 22. Comunalidades serie A

	INICIAL	EXTRACCIÓN
A7	1,000	,607
A8	1,000	,642
A9	1,000	,729
A10	1,000	,517
A11	1,000	,652
A12	1,000	,622

Nota: Tabla de comunalidades extraída del análisis factorial de la serie A.

Como se observa en la tabla 22, todos los ítems de la serie A explican adecuadamente los componentes extraídos, aunque el ítem A10 apenas alcanza un porcentaje aceptable (.517) de explicación.

El número de factores que componen la serie A son dos. La mayoría de los ítems (A7, A8, A9 y A10) se agrupan en el 1^{er} factor, mientras que los ítems A11 y A12 componen el factor 2, como se puede ver en la tabla 23. Los componentes de la serie A explican el 62.845% del total de la varianza.

Tabla N° 23. Matriz de componentes rotados serie A

	COMPONENTES	
	1	2
A7	,762	-,164
A8	,798	,076

A9	,849	,091
A10	,702	,155
A11	,027	,807
A12	,059	,787

Nota: Matriz de componentes rotados extraída del análisis factorial de la serie A.

Con respecto a la serie B, la mayoría de los ítems explican adecuadamente el total de los componentes extraídos; sin embargo, en la tabla 24 se pueden observar 3 ítems con porcentajes bajos, específicamente los ítems B6, B8 y B10.

Tabla N° 24. Comunalidades serie B

	INICIAL	EXTRACCIÓN
B3	1,000	,790
B6	1,000	,194
B7	1,000	,796
B8	1,000	,533
B9	1,000	,637
B10	1,000	,385
B11	1,000	,560
B12	1,000	,681

Nota: Tabla de comunalidades extraída del análisis factorial de la serie B

Los factores resultantes del análisis de la serie B fueron 3. Estos factores están compuestos por 2, 3 y 3 ítems respectivamente. Asimismo, los 3 factores de la serie B cuentan con un 57.206% del total de la varianza explicada.

Tabla N° 25. Matriz de componentes rotados serie B

	COMPONENTES		
	1	2	3
B3	,883	-,026	-,100
B6	-,076	-,001	,434
B7	,878	,090	,131
B8	,006	,695	,225
B9	,037	,180	,777
B10	,107	,083	,606

B11	-,065	,676	,314
B12	,169	,736	-,334

Nota: Matriz de componentes rotados extraída del análisis factorial de la serie B.

En el caso de la serie C, la tabla de comunalidades presenta diversos porcentajes bastante bajos, lo cual advierte sobre la negligencia de estos ítems para representar los factores extraídos de la serie C. Los ítems que menos son explicados por los factores son C1, C3, C6, C7, C8, C9, C10 y C11. El ítem C12 también es bajo, aunque se acerca más a un nivel aceptable.

Tabla N° 26. Comunalidades serie C

	INICIAL	EXTRACCIÓN
C1	1,000	,568
C2	1,000	,656
C3	1,000	,414
C4	1,000	,715
C5	1,000	,641
C6	1,000	,490
C7	1,000	,402
C8	1,000	,467
C9	1,000	,556
C10	1,000	,393
C11	1,000	,500
C12	1,000	,581

Nota: Tabla de comunalidades extraída del análisis factorial de la serie C.

Los factores extraídos del análisis factorial son 4. En cada uno de los factores cargan 3 ítems, C1, C3 y C4 se encuentran en el 1^{er} factor. C2, C7 y C9 pertenecen al 2^o factor. C5, C6 y C8 se encuentran en el factor 3. En el 4^o factor cargan los ítems C10, C11 y C12. Los factores arrojados por el análisis explican un 53.205% del total de la varianza.

Tabla N° 27. Matriz de componentes rotados serie C

	COMPONENTES			
	1	2	3	4
C1	,750	-,025	,051	-,041
C2	-,068	,801	-,083	,058

C3	,524	-,064	-,180	,321
C4	,785	,162	,262	,062
C5	-,154	-,188	,720	,252
C6	,284	,215	,557	-,232
C7	,194	,532	,275	-,078
C8	,117	,255	,622	,042
C9	-,020	,731	,107	,096
C10	,277	,155	,153	,519
C11	,340	-,155	,245	,549
C12	-,175	,091	-,121	,727

Nota: Matriz de componentes rotados extraída del análisis factorial de la serie C.

Como se puede apreciar, en la tabla 28 se presentan las comunalidades de la serie D. Dentro de esta serie, los ítems con menor porcentaje de explicación son el D11 y D12. En el caso del ítem D2, el porcentaje de explicación es bajo, pero también se acerca a un nivel aceptable.

Tabla N° 28. Comunalidades serie D

	INICIAL	EXTRACCIÓN
D2	1,000	,595
D3	1,000	,745
D4	1,000	,808
D5	1,000	,784
D6	1,000	,795
D7	1,000	,627
D8	1,000	,701
D9	1,000	,631
D10	1,000	,721
D11	1,000	,292
D12	1,000	,475

Nota: Tabla de comunalidades extraída del análisis factorial de la serie D.

Como se habrá podido notar, los factores extraídos de cada serie incrementan conforme se avanza. De la serie D se extrajeron 5 factores, lo cuales tienen un 65.223% del total de la varianza explicada. La mayoría de sus factores están compuestos por 2 ítems; algunos de estos ítems cuentan con un porcentaje bajo de representación de su respectivo factor, específicamente D7, D11 y D12.

Tabla N° 29. Matriz de componentes rotados serie D

	COMPONENTES				
	1	2	3	4	5
D2	,058	,724	-,045	-,186	,177
D3	,102	,841	,048	,158	,030
D4	-,042	,010	,042	,897	,011
D5	-,225	-,134	,825	,101	,156
D6	-,043	,185	,081	-,015	,867
D7	,194	,435	,474	,298	-,294
D8	,527	,205	,583	-,205	,011
D9	,733	,220	,070	,138	-,147
D10	,688	-,014	-,080	,264	,414
D11	,232	-,265	,193	,248	,264
D12	-,653	,022	,074	,205	,029

Nota: Matriz de componentes rotados extraída del análisis factorial de la serie D.

Dentro de la tabla de comunalidades de la serie E, se observan diversos porcentajes de explicación muy bajos, de lo cual se puede inferir que aquellos con porcentajes bajos no están explicando adecuadamente los factores extraídos.

Tabla N° 30. Comunalidades serie E

	INICIAL	EXTRACCIÓN
E1	1,000	,475
E2	1,000	,498
E3	1,000	,582
E4	1,000	,536
E5	1,000	,479
E6	1,000	,442
E7	1,000	,611
E8	1,000	,313
E9	1,000	,550
E10	1,000	,556
E11	1,000	,580
E12	1,000	,599

Nota: Tabla de comunalidades extraída del análisis factorial de la serie E.

Los factores obtenidos del análisis son 4. En el factor 1 se encuentran agrupadas las cargas de los ítems E4, E5, E6 y E8. Los ítems E1, E2, E3 y E4 se agrupan en el factor 2. En el factor 3 los factores que están cargando son E7, E11 y E12. En el factor 4 sólo se encuentra cargando el ítem E10. Los factores extraídos en esta serie explican el 51.845% del total de la varianza explicada.

Tabla N° 31. Matriz de componentes rotados serie E

	COMPONENTES			
	1	2	3	4
E1	-,382	,528	,222	-,027
E2	,198	,619	-,270	,060
E3	,350	,664	-,022	-,133
E4	,608	,353	,047	-,198
E5	,687	,043	,053	-,051
E6	,639	,113	,000	,143
E7	,029	,354	,681	,144
E8	,522	,035	,137	,144
E9	,118	,666	,212	,216
E10	,099	,111	,185	,707
E11	,185	-,199	,700	-,131
E12	,002	,056	,390	-,666

Nota: Matriz de componentes rotados extraída del análisis factorial de la serie E.





• CAPÍTULO 6

Conclusiones

En este capítulo se han incluido la discusión y conclusiones que se han derivado de los análisis estadísticos del presente estudio. Además, se incluyen las limitaciones del estudio que servirán para posteriores investigaciones relacionadas con ésta.

Los primeros análisis estadísticos se realizaron a través del análisis factorial. Al llevarlo a cabo, se notó que varios de los ítems de las MPG contaban con varianza cero. Esto se puede interpretar como una tasa del 100% de respuestas erróneas o 100% de respuestas correctas. En este caso, el 100% de los participantes contestaron correctamente a los ítems A1, A2, A3, A4, A5, A6, B1, B2, B4, B5 y D1.

A la luz de estos primeros resultados, y tomando en cuenta que las MPG fueron diseñadas con un orden de dificultad ascendente, se infiere que aquellos ítems son de nula utilidad para el instrumento. Esto se debe, probablemente, a un índice de dificultad bajo por parte de esos ítems.

Lo anterior, concuerda con otros estudios realizados (TORRES & CUESTA, 1992; CHICO, 1997; CAIRO, CAIRO, BOUZA & PONCE, 2000; FERNÁNDEZ, ONGARATO, SAAVEDRA & CASULLO, 2004), los cuales encontraron índices de dificultad muy bajos en la mayoría de los ítems, en algunos otros casos los índices fueron extremos, porque se ha confirmado que diversos ítems no cuentan con un adecuado índice de discriminación. Por ejemplo, Cairo, Cairo, Bouza & Ponce (2000), encontraron que un 76.7% del total de los ítems se podían considerar de dificultad inferior. También se ha confirmado que el orden de dificultad ascendente no se cumple y que, en algunos casos, las series C y D no presentan una diferencia significativa en cuanto a dificultad.

Al retirar los ítems que no variaban fue posible aplicar el análisis factorial. En este primer análisis, donde se incluyó la totalidad de los ítems (con las excepciones ya mencionadas), se observó un número de factores notablemente alto. En total se extrajeron 18 factores, cuando la expectativa se centraba en una máximo de cinco (por sus series A, B, C, D y E).

Si se observa detenidamente, las cargas esperadas en cada factor apenas se cumplen. Por ejemplo, en el caso de la serie A, la mayoría de las cargas factoriales se concentran en el componente uno (A7, A8, A9, A10); no obstante, el resto de las cargas se encuentran en los componentes 13 y 15 respectivamente. En el caso del factor dos, los ítems que presentan las cargas más altas son el B8, B12, E3, E4, E5 y E6, aunque lo esperado era una mayoría de ítems de la serie B. Además, se pueden observar factores en los cuales se agrupan cargas de muy pocos ítems. Concretamente los factores

4, 5, 7, 8, 9, 10, 12, 13, 14, 15, 16, 17 y 18 son aquellos en los que se concentran las cargas de 3, 2 y 1 ítem, respectivamente. Con esto se podría cuestionar la adecuada representación del constructo por parte de los ítems que cargan en esos factores.

Debido a la gran cantidad de factores y cargas que apenas los representan, se tomó la decisión de realizar un análisis factorial por serie, con lo cual se desprendió que la serie A se agrupaba en dos factores y sólo cuenta con un ítem con un porcentaje bajo de explicación (A10). En uno de los factores sólo están cargando los ítems A11 y A12. Por lo tanto, se podría plantear nuevamente la pregunta sobre la adecuada representación del constructo.

La serie B, por otro lado, arrojó tres factores. En el factor uno cargan dos ítems (B3 y B7), en el factor dos cargan tres ítems (B8, B11 y B12), al igual que en el factor tres (B6, B9 y B10). En la tabla de comunalidades, los porcentajes más bajos de explicación son dados por los ítems B6 y B10. Con esto se puede pensar en la eliminación de ciertos ítems, no obstante, los factores quedarían representados por muy pocos ítems.

Para la serie C, el número de factores incrementó a 4. Sin embargo, la serie C fue una de las pocas que conservó la totalidad de sus ítems, al presentar variaciones en todos ellos. Los porcentajes de explicación en la serie C caen de manera sustancial. Particularmente se puede observar en los ítems C3, C6, C7, C8, C10 y C11. Cada uno de los factores comprende cargas de tres ítems; si se retiran algunos con poco porcentaje de explicación, también se reduce bastante el número de ítems que representan cada factor y, por ende, se enfrenta una decisión difícil.

Para el análisis de la serie D únicamente se excluyó el ítem D1. Aunque ha sorprendido la cantidad de factores extraídos, los cuales suman cinco en total. Cada factor reúne las cargas de dos ítems únicamente, lo cual ya sugiere una pobre representación de cada factor. Además, si se observa la tabla de comunalidades, los ítems D11 y D12 son aquellos con porcentajes de explicación más bajos. De considerar una reducción de ítems, los factores cuatro y cinco quedarían representados por un número de ítems escueto, sin contar el resto que contaría únicamente con dos ítems.

A la serie E no se le restaron ítems para su análisis y, al igual que la serie C, arrojó 4 factores. Entre estos, los que conservan mayor número de cargas son los factores uno y dos, representados por cuatro ítems cada uno. El factor tres está representado por las cargas de tres ítems y, finalmente, el factor cuatro está representando únicamente por el ítem E10. No obstante, las comunalidades

muestran a 6 de los 12 ítems con porcentajes de explicación bajos, que se vuelve relevante al considerar la eliminación de ítems.

Todo esto, sobre el excedente número de factores, se ha observado en otras investigaciones (DELGADO, ESCURRA, BULNES, & QUESADA, 2001; MACKINTOSH & BENNETT, 2005), en las cuales se han detectado hasta 12 factores al someter al análisis factorial a las MPA, las cuales se componen de la serie I y II.

Aunque no sólo se ha aplicado el análisis factorial para comprobar la validez de las MPR. Sánchez & Pirela (2009), por ejemplo, revisaron la validez de las MPR al correlacionarlas con los tests de Otis, Purdue y Dominos, además de indicadores de rendimiento escolar. Con lo cual observaron que las correlaciones entre las MPR y los tests no eran adecuadas (con excepción del test de Dominos). Con respecto a la correlación de indicadores de rendimiento escolar y las MPR, se determinó un coeficiente poco relevante, con lo cual se concluyó que el test MPR no cuenta con validez.

Con respecto a los análisis para determinar la confiabilidad de la prueba, se llevó a cabo un Alfa de Cronbach. Con el cual se determinó un coeficiente de .799, que se considera adecuado y con lo cual se confirma que el instrumento es confiable. Sin embargo, también se decidió extraer los coeficientes Alfa por cada serie y comprobar si es posible mejorar la confiabilidad si se elimina algún elemento de la serie.

En la serie A se observó un coeficiente Alfa de .549, que se considera un coeficiente bajo. Con la serie B se obtuvo un coeficiente de .495, el cual es aún más bajo que el de la serie A. En la serie C se obtuvo un coeficiente Alfa .611, considerado adecuado. Con la serie D, el coeficiente observado fue de .484, el cual es considerado un coeficiente de confiabilidad bajo. Finalmente, con la serie E se obtuvo un coeficiente Alfa de .627, muy similar al obtenido en la serie C.

Sin embargo, al calcular el coeficiente Alfa si se elimina un elemento, se observa que en todas las series el coeficiente de confiabilidad incrementa notablemente si se elimina el ítem 12, lo cual concuerda con otras investigaciones que abordan la pertinencia de utilizar ciertos ítems (Cairo, Cairo, Bouza & Ponce, 2000; FERNÁNDEZ, ONGARATO, SAAVEDRA & CASULLO, 2004).

Al parecer, en la mayoría de las investigaciones en las que se ha comprobado la validez de las MPR, se han observado coeficientes Alfa de confiabilidad adecuados, al igual que en la presente investigación (SÁNCHEZ & PIRELA, 2009; FERNÁNDEZ, ONGARATO, SAAVEDRA & CASULLO, 2004; DELGADO, ESCURRA, BULNES & QUESADA, 2001; RUSHTON, SKUY & BONS, 2004; RAMÍREZ, LORENZO & DÍAZ, 2013).

La confiabilidad y validez de un instrumento, son fundamentales para la correcta medición de constructos o conductas, lo que lleva a tomar mejores decisiones a los profesionales. De acuerdo con la Teoría Clásica de los Tests (ABAD, GARRIDO, OLEA, & PONSODA, 2006), no es adecuado continuar con el proceso de construcción de una prueba si no es confiable ni válida. En algunas

investigaciones (CASÉ, NEER & LOPETEGUI, 2002; Ortíz, 2009) se han determinado percentiles para población argentina y mexicana, respectivamente. Sin embargo, en aquellas investigaciones se sugiere haber dado por hecho que la prueba contaba con confiabilidad y validez, un punto que se debe someter a consideración si se conoce el procedimiento adecuado para la construcción de pruebas psicométricas.

De acuerdo con los resultados obtenidos en el presente estudio y con los de otras investigaciones citadas, se sugiere, para posteriores investigaciones, llevar a cabo un análisis estadístico de los ítems, a través del cual se pueda determinar un orden de dificultad más cercano a lo que se pretendió en su diseño original. Además, ya que en diversas investigaciones se han encontrado ítems con varianza cero, es probable que el test requiera de un ajuste en el cual se eliminen algunos de los ítems.

Como se ha podido constatar en ésta y otras investigaciones, la confiabilidad de la prueba es adecuada; aunque, los análisis que se realizaron por serie demuestran que es posible mejorar el coeficiente de confiabilidad, siempre y cuando se eliminen ciertos elementos, los cuales probablemente no cuentan con una tasa de varianza adecuada.

Por otro lado, los análisis factoriales a los que se sometió la prueba demuestran un 70.799% de explicación del total de la varianza, lo cual se puede considerar adecuado para la validez. No obstante, es preciso recordar que el número de factores extraído tanto por el análisis de la totalidad de los ítems como por el de cada una de las series, es bastante alto. Tomando en cuenta que la prueba sólo pretende medir un constructo y no varios, el número de factores que se esperaba era de máximo cinco e idealmente de uno. Esto se debe a que el diseñador de la prueba pretendió medir la capacidad mental de los individuos, para lo cual diseñó las MPG con ítems fabricados con tal propósito. En tal caso, las cargas de estos ítems idealmente deberían cargar en un solo factor. No obstante, ya que la prueba consta de cinco series se asumió que el análisis podría arrojar hasta 5 factores.

Además, cuando se realizaron los análisis por serie se pudo observar una gran cantidad de ítems con un porcentaje de explicación bastante bajo. Asimismo, específicamente en las series B, C y E se obtuvieron porcentajes bajos del total de la varianza explicada (57.206%, 53.205% y 51.804% respectivamente). Sin mencionar que las series C y E permanecieron intactas (no se restaron ítems). Para otras investigaciones posteriores, esto puede representar un marco de referencia si se pretende recortar algunos ítems, con lo cual probablemente se pueda reajustar la prueba, aunque también existe la posibilidad de que se requiera la fabricación de nuevos ítems.

Uno de los estudios sobre las propiedades psicométricas de las MPR aquí citado (MACKINTOSH & BENNETT, 2005), ha planteado la posibilidad de que el número de factores tan alto se deba quizá a otra explicación. En una investigación realizada por Carpenter, Just & Schell (1990), se determinó

que existían 5 posibles formas de dar solución a los ítems de las MPR:

1. constante en una hilera,
2. distribución de tres,
3. progresión de pares,
4. adición/sustracción y
5. distribución de dos.

Desde los análisis factoriales realizados por serie, se puede cuestionar la posibilidad de que los factores expliquen un tipo de solución. No obstante, serán los estudios posteriores los que den respuesta a estas incógnitas, a través de un análisis más profundo.

Finalmente, se considera necesario llevar a cabo nuevas investigaciones en las cuales se pueda comprobar si los factores extraídos pueden ser explicados por los tipos de solución de ítems sugeridos por Carpenter, Just & Schell (1990).

También es necesario comprobar si con un reajuste de la prueba, mediante la eliminación o adición de nuevos ítems y reordenamiento de éstos, puede incrementar la confiabilidad de la prueba y quizá, en caso de no obtener una respuesta satisfactoria sobre los factores a través de los diferentes tipos de solución, averiguar si con una versión abreviada del test, el número de factores reduce y los porcentajes de explicación del total de la varianza incrementan positivamente.

6.1 LIMITACIONES DEL ESTUDIO

La muestra utilizada para el presente estudio fue de tipo no probabilística y se debió a dificultades relacionadas con el acceso a los alumnos. En una de las reuniones se explicó que se necesitaba tener acceso a las listas del total de alumnos del turno diurno, para calcular una muestra representativa. No obstante, el Secretario Académico explicó que era imposible tener acceso a los alumnos de esa manera, debido a que éstos se encontraban en periodo de evaluación y si no se encontraban haciendo examen, podrían estar entregando trabajos finales o recibiendo calificaciones; por lo tanto, se acordó que los directivos determinarían los grupos de alumnos a los cuales se tendría acceso y a partir de ello, se solicitaría su participación voluntaria a los alumnos. También se acordaron los días y horas en las que se llevarían a cabo las aplicaciones de la prueba. No obstante, para posteriores investigaciones lo más idóneo será utilizar una muestra de tipo probabilístico.

Además, las aplicaciones se realizaron en los salones de clases, algunos de los cuales eran bastante pequeños y permitían que los participantes estuvieran demasiado cerca uno del otro. Esto, aunque se puso especial atención para evitarlo, pudo facilitar la copia entre compañeros.

Especialmente dos de las aulas contaban con poca iluminación, lo cual pudo representar un factor relevante para la aplicación. Asimismo, los salones contaban con una cantidad numerosa de

alumnos, lo cual en salones muy pequeños dificultaba el control de los participantes; por ello se sugiere, para estudios posteriores, utilizar entornos más adecuados para las aplicaciones, donde exista ventilación e iluminación adecuadas, así como utilizar un número de participantes reducido por aplicación, con la finalidad de tener un mejor control.



BIBLIOGRAFÍA

- ABAD, F. J., GARRIDO, J., OLEA, J., & PONSODA, V. (2006). *Introducción a la psicometría: teoría clásica de los tests y teoría de la respuesta al ítem*. Madrid: Universidad Autónoma de Madrid.
- AMADOR, J. A., FORNS, M., & KIRCHNER, T. (2005). *Tests de factor g y factoriales*. Barcelona: Departamento de personalidad, evaluación y tratamiento psicológico facultad de psicología.
- ANASTASI, A., & URBINA, S. (1998). *Tests Psicológicos*. México, D.F.: Prentice Hall.
- ARAGÓN, L. E. (2004). «Fundamentos psicométricos en la evaluación psicológica»; en *Revista electrónica de psicología Iztacala*, 7(4), págs. 23-43.
- BARRON, K. E., BROWN, A. R., EGAN, T. E., GESUALDI, C. R., & MARCHUK, K. A. (2008). «Sobre la confiabilidad»; en M. García, A. del Castillo, R. M. Guzmán, & J. P. Martínez, *Medición en psicología: del individuo a la interacción* (págs. 29-30). Pachuca: Universidad Autónoma del Estado de Hidalgo.
- BORING, E. (1923). «Definiciones de inteligencia»; en R. J. Gregory, *Evaluación psicológica: historia, principios y aplicaciones* (pág. 170). México, D.F.: Manual Moderno.
- BRODY, E. B., & BRODY, N. (1976). «Thurstone y las habilidades mentales primarias»; en R. J. Gregory, *Evaluación psicológica: historia, principios y aplicaciones* (págs. 176-178). México, D.F.: Manual Moderno.
- EN R. J. GREGORY, *Evaluación psicológica: historia, principios y aplicaciones* (págs. 176-178). México, D.F.: Manual Moderno.
- CARROLL, J. (1993). «Resumen de Carroll»; en T. P. Hogan, *Pruebas psicológicas. Una introducción práctica* (pág. 202). México, D.F.: Manual Moderno.
- CATTELL, J. M., & HORN, J. L. (1966). «Teoría de Cattell sobre la inteligencia fluida y cristalizada»; en T. P. Hogan, *Pruebas psicológicas. Una introducción práctica* (pág. 201). México, D.F.: Manual Moderno.
- CHICO, E. (1997). «La invarianza en la estructura factorial del Raven en grupos de delincuentes y no delincuentes»; en *Psicothema*, 9(1), págs. 47-55.
- COLOM, R., & PUEYO, A. A. (1999). «El estudio de la inteligencia humana: recapitulación ante el cambio de milenio»; en *Psicothema*, 11(3), págs. 453-476.
- CORTADA, N. (1998). «Logros en educación secundaria y su relación con inteligencia y con resolución de problemas nuevos»; en *Revista latinoamericana de psicología*, 30(2), págs. 293-310.
- DAS, J. P., NAGLIERI, J. A., & KIRBY, J. R. (1994). «Procesamiento simultáneo y secuencial: la teoría PAS»; en T. P. Hogan, *Pruebas psicológicas. Una introducción práctica* (págs. 208-209). México, D.F.: Manual Moderno.

FITZPATRICK, R., DAVEY, C., BUXTON, M. J., & JONES, D. R. (2001). «La confiabilidad interna a través del coeficiente alfa de Cronbach»; en M. García, A. del Castillo, R. M. Guzmán, & J. P. Martínez, *Medición en psicología: del individuo a la interacción* (pág. 33). Pachuca: Universidad Autónoma del Estado de Hidalgo.

GARCÍA, M., DEL CASTILLO, A., GUZMÁN, R. M., & MARTÍNEZ, J. P. (2010). *Medición en psicología: del individuo a la interacción*. Pachuca: Universidad Autónoma del Estado de Hidalgo.

GARDNER, H. (1983, 1993). «Gardner y la teoría de las inteligencias múltiples»; en R. J. Gregory, *Evaluación psicológica: historia, principios y aplicaciones* (págs. 185-186). México, D.F.: Manual Moderno.

GARDNER, H. (1998). «Gardner y la teoría de las inteligencias múltiples»; en R. J. Gregory, *Evaluación psicológica: historia, principios y aplicaciones* (págs. 185-186). México, D.F.: Manual Moderno.

GONZÁLEZ, F. M. (2007). *Instrumentos de evaluación psicológica*. La Habana: Editorial Ciencias Médicas.

GREGORY, R. J. (2001). *Evaluación psicológica. Historia, principios y aplicaciones*. México, D.F.: Manual Moderno.

GRIMALDO, M. P. (2007). «La teoría de Kohlberg, una explicación del juicio moral desde el constructivismo»; en *Cultura*, 21, págs. 325-340.

HERNÁNDEZ, R., FERNÁNDEZ, C., & BAPTISTA, P. (2006). *Metodología de la investigación*. México, D.F.: McGraw Hill.

HOGAN, T. P. (2004). *Pruebas psicológicas. Una introducción práctica*. México, D.F.: Manual Moderno.

INEGI. (18 de 05 de 2015). *Población de México*. Recuperado el 18 de 05 de 2015, de Cuéntame de México: <http://cuentame.inegi.org.mx/poblacion/default.aspx?tema=P>

IVANOVIC, R., FORNO, H., DURÁN, M. C., HAZBÚN, J., CASTRO, C., & IVANOVIC, D. (2000). «Estudio de la capacidad intelectual (matrices progresivas de Raven) en escolares chilenos de 5 a 18 años I. Antecedentes generales, normas y recomendaciones»; en *Revista de psicología general y aplicada*, 53(1), págs. 5-30.

JENSEN, A. (1998). «Teoría de Jensen»; en T. P. Hogan, *Pruebas psicológicas. Una introducción práctica* (págs. 206-207). México, D.F.: Manual Moderno.

JOHNSTON, M., FRENCH, D. P., BONETTI, D., & JOHNSTON, D. W. (2004). «Medición en psicología: del individuo a la interacción»; en M. García, A. del Castillo, R. M. Guzmán, & J. P. Martínez, *Medición en psicología: del individuo a la interacción* (págs. 9-10). Pachuca: Universidad Autónoma del Estado de Hidalgo.

KAIL, R. V., & CAVANAUGH, J. C. (2006). *Desarrollo humano. Una perspectiva del ciclo vital*. México, D.F.: Thomson.

KANE, H., & BRAND, C. (1905). «The importance of Spearman's g as a psychometric, social and educational construct»; en *The Occidental Quarterly*, 3(1), págs. 7-30.

KOHLBERG, L. (1984). «Teoría de Piaget sobre el desarrollo cognitivo»; en T. P. Hogan, *Pruebas psicológicas. Una introducción práctica* (págs. 204-205). México, D.F.: Manual Moderno.

LEJARRAGA, H. (2010). «Genética del desarrollo y la conducta»; en *Arch Argent Pediatr*, 108(4), págs. 331-336.

LÓPEZ, L. M., CABRERA, G. M., PRETER, M., LUACES, Y., MIRANDA, A., REGUEIRA, B., y otros. (2010). *Compendio de instrumentos de evaluación psicológica*. La Habana: Editorial Ciencias Médicas.

MACKINTOSH, N. J., & BENNETT, E. S. (2005). «What do Raven's Matrices measure? An analysis in terms of sex differences»; en *Elsevier*, 33, págs. 663-674.

MILLER, F. M., & RAVEN, J. C. (1939). «Desarrollo de los ítems»; en J. C. Raven, *Test de matrices progresivas: escalas coloreada, general y avanzada* (págs. I 52- I 56). Buenos Aires: Paidós.

MOLERO, C., SAIZ, E., & ESTEBAN, C. (1998). «Revisión histórica del concepto inteligencia: una aproximación a la inteligencia emocional»; en *Revista Latinoamericana de Psicología*, 30(1), págs. 11-30.

NUNNALLY, J. C. (2009). *Teoría Psicométrica*. México, D. F.: Trillas.

OLIVA, A. (1997). «La controversia entre herencia y ambiente. Aportaciones de la genética de la conducta»; en *Apuntes de Psicología*, 51, págs. 21-37.

PIAGET, J. (1926,1952,1972). «Piaget y la adaptación»; en R. J. Gregory, *Evaluación psicológica: historia, principios y aplicaciones* (págs. 178-180). México, D.F.: Manual Moderno.

PICHOT, P. (1986). *Los tests mentales*. México, D. F.: Paidós.

RAVEN, J. C. (2004). *Test de matrices progresivas: escalas coloreada, general y avanzada*. Buenos Aires: Paidós.

SÁNCHEZ, M., & PIRELA, L. (2009). «Propiedades psicométricas de la prueba: Matrices Progresivas de Raven, en estudiantes de Orientación»; en *Laurus*, 15(29), págs. 76-97.

SIERRA, M. S., & CASTRO, C. (17 de 12 de 2013). *Validación y estandarización de pruebas psicológicas*. Recuperado el 2014 de 11 de 11, de Universidad de La Sabana: <http://intellectum.unisabana.edu.co:8080/jspui/bitstream/10818/9552/1/MAR%C3%8DA%20SOLEDAD%20SIERRA%20%20tesis.pdf>

SOCIEDAD MEXICANA DE PSICOLOGÍA. (2009). *Código Ético del Psicólogo*. México D.F.: Trillas.

SPEARMAN, C. (1923). «Spearman y el factor g»; en R. J. Gregory, *Evaluación psicológica: historia, principios y aplicaciones* (págs. 175-176). México, D.F.: Manual Moderno.

STERNBERG, R. J. (1986). «Definiciones de Inteligencia»; en R. J. Gregory, *Evaluación psicológica: historia, principios y aplicaciones* (pág. 170). México, D.F.: Manual Moderno.

- STERNBERG, R. J. (1986). «Sternberg y la teoría triárquica de la inteligencia»; en R. J. Gregory, *Evaluación psicológica: historia, principios y aplicaciones* (págs. 186-189). México, D.F.: Manual Moderno.
- THURSTONE, L. L. (1938). «Teoría de las capacidades mentales primarias propuesta por Thurstone»; en T. P. Hogan, *Pruebas psicológicas. Una introducción práctica* (págs. 198-200). México, D.F.: Manual Moderno.
- TORRES, E., & CUESTA, M. (1992). «Una revisión de las propiedades psicométricas del test de matrices progresivas de Raven (escala superior)»; en *Psicothema*, 4(1), págs. 261-267.
- VERNON, P. (1960). «Teorías jerárquicas»; en A. Anastasi, & S. Urbina, *Tests psicológicos* (págs. 316-318). México, D.F.: Prentice Hall.
- WILSON, R. S. (1983). «Factores hereditarios y ambientales»; en R. V. Kail, & J. C. Cavanaugh, *Desarrollo humano. Una perspectiva del ciclo vital* (págs. 226-228). México, D.F.: Thomson.
- YELA, M. (1996). «Ambiente, herencia y conducta»; en *Psicothema*, págs. 187-228.



ÍNDICE DE TABLAS

Tabla N° 1. Clasificaciones de las pruebas psicológicas		18
Tabla N° 2. Clasificaciones de las pruebas psicológicas		19
Tabla N° 3. 9 factores que componen las CMP		33
Tabla N° 4. Etapas del desarrollo cognitivo de Piaget		37
Tabla N° 5. Teoría del desarrollo moral de Kohlberg		38
Tabla N° 6. Inteligencias múltiples de Gardner		41
Tabla N° 7. Características de las pruebas de inteligencia de aplicación individual y grupal		46
Tabla N° 8. Lista de pruebas de inteligencia		47
Tabla N° 9. Estadísticos de confiabilidad		90
Tabla N° 10. Estadísticos de confiabilidad serie A		90
Tabla N° 11. Valores Alfa si se elimina un elemento de la serie A		90
Tabla N° 12. Estadísticos de confiabilidad serie B		91
Tabla N° 13. Valores Alfa si se elimina un elemento de la serie B		91
Tabla N° 14. Estadísticos de confiabilidad serie C		92
Tabla N° 15. Valores Alfa si se elimina un elemento de la serie C		92
Tabla N° 16. Estadísticos de confiabilidad serie D		92
Tabla N° 17. Valores Alfa si se elimina un elemento de la serie D		93
Tabla N° 18. Estadísticos de confiabilidad serie E		93
Tabla N° 19. Valores Alfa si se elimina un elemento de la serie E		94
Tabla N° 20. Comunalidades		94
Tabla N° 21. Matriz de componentes rotados		96
Tabla N° 22. Comunalidades serie A		98
Tabla N° 23. Matriz de componentes rotados serie A		98
Tabla N° 24. Comunalidades serie B		99
Tabla N° 25. Matriz de componentes rotados serie B		99
Tabla N° 26. Comunalidades serie C		100
Tabla N° 27. Matriz de componentes rotados serie C		100
Tabla N° 28. Comunalidades serie D		101
Tabla N° 29. Matriz de componentes rotados serie D		102
Tabla N° 30. Comunalidades serie E		102
Tabla N° 31. Matriz de componentes rotados serie E		103

ÍNDICE DE FIGURAS

Figura N° 1		35
Figura N° 2		36
Figura N° 3		39
Figura N° 4		40
Figura N° 5		42
Figura N° 6		45
Figura N° 7		88
Figura N° 8		89
Figura N° 9		89



Esta TESIS titulada,
*Revisión de la Validez y Confiabilidad de las Matrices Progresivas de Raven
Escala General en Alumnos de 15 a 19 Años,*
fue escrita por Gerardo Nájera Figueroa
para obtener el grado de Licenciado en Psicología,
por parte de la Universidad Latina (UNILA).
Este libro fue impreso en México DF
en algún momento del año 2015.