

4419 8185
TESIS
PS998
A73

**UNIVERSIDAD CATÓLICA ANDRÉS BELLO
FACULTAD DE HUMANIDADES Y EDUCACIÓN
ESCUELA DE PSICOLOGÍA**



**VALIDACIÓN DE LA VERSIÓN ABREVIADA
DEL TEST AVANZADO DE MATRICES PROGRESIVAS DE RAVEN**

*Trabajo presentado a la Escuela de Psicología como requisito
parcial para la obtención de la Licenciatura en Psicología*

**Arteaga G., María de Lourdes
Dunn M., Héctor O.**

Tutor: Lic. José Gregorio De Llano

**Caracas
Julio, 1998**

*A mi hermano Carlos, a quien
seguimos amando en sueños,
cuyo hermosísimo legado nos
permite amarlo en actos y
construir el futuro en medio de su
ausencia: Carla, Carlos Alberto,
Ricardo y su empresa.*

Héctor O. Dunn M.

*A mis padres, por recordarme
todos los días con sus palabras y
hechos que siempre están ahí
para mí.*

María de L. Arteaga.

AGRADECIMIENTOS

A nuestro padrino Gustavo Peña, por todo su apoyo, tanto metodológico como personal, incluso en aquello que no era necesario que hiciera.

A nuestro tutor José Gregorio De Llano, por su confianza e insistencia en la perfección.

A Jörg Prieler, quien desde Viena nos permitió desinteresadamente contar con información actualizada muy importante para este trabajo.

A María Alejandra Souquet y Ángel Veraza, por estar siempre ahí dispuestos a ayudar con un toque de humor.

A todos los profesores que cedieron parte de su tiempo para que nos fuera posible acceder a las personas que amablemente participaron en esta investigación; particularmente a María Alejandra Martínez, Javier Seoane (Cs. Pedagógicas), Marinel Cumare (Economía), Oswaldo Villalobos, Jaime Robles (Psicología), Kattyuska Di Eugenio (Derecho), César Sánchez (Relaciones Industriales).

A todos las personas que accedieron de forma voluntaria a formar parte de la muestra de sujetos, en especial a nuestros compañeros de clase.

A mi querido esposo Pedro Bellorín, por permitirnos robar parte de su tiempo y equipo de trabajo.

ÍNDICE DE TABLAS	v
ÍNDICE DE GRÁFICOS	vi
ÍNDICE DE ANEXOS	vii
RESUMEN	ix
INTRODUCCIÓN	10
BASES TEÓRICAS	13
El Análisis Factorial en la medición de la Inteligencia	14
Test de Matrices Progresivas de Raven	23
Desarrollos recientes de las Matrices Progresivas	32
Desarrollo de las versiones cortas	42
MÉTODO	47
Problema	47
Hipótesis	47
Variables	48
Tipo de investigación	49
Procedimiento	49
Sujetos y recolección de datos	49
Análisis de datos	53
RESULTADOS	57
DISCUSIÓN	75
CONCLUSIÓN Y RECOMENDACIONES	88
REFERENCIAS	91
ANEXOS	95

INDICE DE TABLAS

TABLA 1: Posición (alternativa incorrecta) y naturaleza de los errores más comunes	30
TABLA 2: Tamaño muestral para ambas versiones por carrera	51
TABLA 3: Propiedades psicométricas de los ítems de la versión original	60
TABLA 4: Matriz factorial del TAMPR original	63
TABLA 5: Propiedades psicométricas de los ítems de la versión abreviada	67
TABLA 6: Matriz factorial de los ítems de la versión abreviada	68
TABLA 7: Coeficientes de comunalidad de los ítems de la versión corta del TAMPR	68
TABLA 8: ANOVA de las alternativas del ítem 4	72
TABLA 9: ANOVA de las alternativas del ítem 8	73
TABLA 10: ANOVA de las alternativas del ítem 10	73
TABLA 11: ANOVA de las alternativas del ítem 12	74
TABLA 12: Media y desviación típica de las notas previas por carrera	75
TABLA 13: Correlación de ambas versiones con el rendimiento académico	76

ÍNDICE DE GRÁFICOS

GRÁFICO 1: Distribución de frecuencia de los puntajes del TAMPR original	58
GRÁFICO 2: Nivel de dificultad de los ítems del TAMPR original	61
GRÁFICO 3: Distribución de frecuencia de los puntajes del TAMPR abreviado.	64
GRÁFICO 4: Nivel de dificultad de los ítems de la versión abreviada del TAMPR.	65
GRÁFICO 5: Discrepancia del nivel de dificultad de los ítems de la forma corta en contraste con sus correspondientes en la forma larga	66

RESUMEN

La versión original del TAMPR y la forma corta de éste propuesta por Arthur y Day (1994) son comparados en cuanto a propiedades psicométricas tales como nivel de dificultad de los ítems, estructura factorial y correlación con rendimiento académico en cursos que implican pensamiento eductivo, sobre la base de una muestra de 399 estudiantes de pre-grado de la UCAB, de los cuales 277 realizaron la forma larga y los restantes 122 la abreviada. De los datos recogidos, se obtuvieron tres resultados fundamentales: (a) el coeficiente de consistencia interna Alpha de Cronbach similar, siendo ligeramente menor en la forma corta (0,73) que en la forma larga (0,83), pero mayor que la estimación realizada a partir de la fórmula Spearman - Brown (0,64), (b) estructura factorial predominantemente unitaria y similar, (c) correlación de igual magnitud y dirección de ambas pruebas respecto al rendimiento académico (0,30). En cuanto al nivel de dificultad de los ítems, la versión corta desplegó una progresión creciente, estable y monotónica que el test largo solo llegó a mostrar en los ítems de la porción central de su muestra. Se discuten estos resultados sobre la base del análisis de varianza entre las alternativas de respuesta de cada ítem y a la luz de las dos discusiones centrales encontradas en torno al Raven: la vinculación entre complejidad y medición del factor g y la independencia de carga cultural. Se proponen nuevos criterios para la elaboración de formas abreviadas del TAMPR.

INTRODUCCIÓN

Uno de los conceptos de mayor interés para la Psicología desde sus mismos inicios ha sido el de la inteligencia. El mismo fue ideado como un constructo hipotético para describir "las aptitudes requeridas para la supervivencia y el progreso..." (Anastasi, 1968/1974, pág. 220) de las cuales dispone una persona. Más específicamente, la búsqueda de indicadores empíricos para determinar la mayor o menor posibilidad de que alguien obtuviera altos estándares de desempeño en el futuro, se inició en el ámbito escolar francés con el test de Binet y Simon entre los siglos XIX y XX, alcanzándose posteriormente un auge en el desarrollo de otros instrumentos, ordenados a la clasificación de personal militar durante la II Guerra Mundial.

De esa manera se fue estableciendo en la Psicología un área de investigación en la cual se integran la formulación teórica, el desarrollo metodológico y la elaboración de herramientas de medición que dieran cuenta de aquello que se teorizaba y que, al mismo tiempo, servía como objeto de estudio per se. En primer lugar, porque se suponía la existencia de alguna estructura psíquica que explica la adaptación, el éxito o el ingenio de las personas; y ello, en consonancia con la filosofía funcionalista y empiricista que tomaba fuerza alrededor de 1900, dio lugar a la búsqueda de indicadores objetivos que dieran cuenta de la existencia de tal fenómeno.

Es así como - aunado al creciente interés por el estudio de las diferencias individuales - tuvo lugar el nacimiento de los test de inteligencia y con ello la pronta emergencia de diversas formulaciones que hipotetizaban no sólo su existencia, sino su naturaleza o constitución. Es dentro de esta diatriba que se enlaza con las técnicas psicométricas que, más allá de las decisiones y pasos necesarios para la elaboración de los tests, eran también destinadas a la validación de las pruebas para comprobar los supuestos teóricos establecidos sobre la inteligencia. Un ejemplo claro de ello lo

constituye el Análisis Factorial, producto eminentemente surgido de la práctica psicológica, que está muy ligado al estudio de la inteligencia.

En consideración de lo expuesto, es necesario advertir que éste es un trabajo de investigación cuyo objetivo es el estudio de una prueba de inteligencia por sí misma, específicamente, el Test Avanzado de Matrices Progresivas de Raven, y no se pretende el uso del mismo como herramienta para encontrar argumentos empíricos sobre alguna posición teórica respecto al constructo inteligencia.

Las Matrices Progresivas de Raven constituyen una prueba destinada a la evaluación de la capacidad intelectual general, con una carga mínima de factor cultural, ya que su resolución requiere poco del lenguaje verbal. Ha sido ampliamente utilizada en diversas áreas de la exploración psicológica, obteniéndose un alto grado de validez predictiva o diagnóstica, particularmente en los ámbitos organizacional y académico.

De esta prueba, han sido elaboradas varias versiones, como la coloreada para niños, la estándar o general para adultos y la avanzada para personas altamente capacitadas en distintas áreas de conocimiento y/o habilidades profesionales. Esta última versión -utilizada en numerosas investigaciones de validación y estudios aplicados en las áreas organizacional y académica- consta de 36 ítems y puede ser administrada en un tiempo promedio de 40 a 60 minutos. Dado que este tiempo de aplicación es mucho mas largo que el de otras medidas de habilidades cognoscitivas usadas en ámbitos aplicados, en países como Inglaterra y Estados Unidos se han desarrollado versiones cortas de la forma estándar en población psicopatológica y de la forma avanzada en contextos académicos. Tales versiones suelen tener una longitud de aproximadamente un tercio de la versión larga, y presentan la ventaja de ser aplicadas en sólo 15 minutos, conservándose similares la confiabilidad, la estructura factorial y el patrón de relaciones convergentes y divergentes del test original.

El propósito de esta investigación es pues, desarrollar una forma corta del Test Avanzado de Matrices Progresivas de Raven (TAMPR), que posea propiedades psicométricas similares a la versión original en cuanto a consistencia interna, estructura unifactorial y patrón de correlaciones con otra medida de logro intelectual, pero con un tiempo de administración sustancialmente más corto. El desarrollo de una forma corta del TAMPR permitiría -por tanto- evaluar inteligencia general en un marco temporal más corto, y proveería a investigadores, psicólogos clínicos, seleccionadores de personal y evaluadores de centros de asesoramiento, de una versión del test que satisfaga sus necesidades de optimizar el tiempo de examinación, sin que ello vaya en detrimento de la calidad de la misma.

Es importante considerar, en relación a la presente investigación, su factibilidad de actualización en el ámbito universitario dentro del cual la misma será realizada; pues, si bien el ejercicio metodológico no aparece como una limitación sustancial del trabajo - específicamente en lo referente a las técnicas y recursos empleados para el procesamiento de información -, vale la pena destacar lo trabajoso de los procedimientos de recogida de datos a llevar a cabo para construir y validar la versión abreviada del test que en este trabajo nos ocupa; lo cual limita, por ejemplo, la escogencia de una segunda o tercera medida de logro intelectual para comparar la validez convergente de las formas original y abreviada del TAMPR.

Con la finalidad de comentar el marco de referencia en el que se inserta el Test Avanzado de Matrices Progresivas de Raven, se hace una breve reseña histórica sobre las posturas conceptuales que hipotetizan acerca de la inteligencia, haciendo énfasis en el paradigma del análisis factorial y en los avances heurísticos que se han logrado empleando esta prueba como objeto de estudio. De igual manera, se expone lo referente a su estructura factorial unitaria, su independencia del factor cultural y las nuevas versiones desarrolladas a partir del test original. Todo esto tiene la finalidad de construir un basamento tanto conceptual como metodológico que permita el logro del objetivo de la presente investigación.

BASES TEÓRICAS

Las pruebas de Raven, o Matrices Progresivas, así como las Escalas de Vocabulario de Mill Hill, fueron diseñadas para evaluar lo más directa y precisamente posible los dos principales componentes del Factor g de Spearman. Se considera que estas dos pruebas al ser usadas juntas proveen información más útil y rápidamente que el resto de las pruebas de inteligencia; y que el Test Avanzado de Matrices Progresivas, una de las formas del test de Raven, por sí mismo permite acceder a evaluaciones acerca de la habilidad de las personas para forjar nuevos discernimientos de gran fundamento no verbal, es decir, su habilidad eductiva.

Muy probablemente ha sido el fundamento teórico y metodológico que ha rodeado la elaboración y evolución de las Matrices Progresivas, en general, lo que les ha conferido su carácter de herramienta fuerte, tanto para la obtención confiable y válida de indicadores relativos al constructo inteligencia como de, consecuentemente, poderosas predicciones relativas al desempeño futuro de aquellas personas cuya capacidad intelectual general ha sido explorada a través de estas pruebas. De allí que sea altamente estimada por la mayoría de los evaluadores y su amplio uso en diversas áreas en la investigación básica y en el ejercicio profesional de la Psicología.

Los párrafos anteriores resumen la forma en que es presentada una de las más recientes ediciones del Test de Matrices Progresivas de Raven, publicada en su país de origen - Inglaterra - por Raven, Court y Raven (1994). Aunque no es fácil decir lo contrario, tampoco lo es escapar de las controversias a las que se debe estar acostumbrado y al tanto en toda disciplina, científica o tecnológica. Ello es debido a que existen multitud de posiciones en cuanto filosofía, modelos, teorías, sistemas, etc., cuyos simpatizantes o detractores suelen buscar o interpretar los datos obtenidos a

la luz de lo que defienden o atacan, y ya no se trata sólo de conocer tales controversias, sino también de criticarlas.

Hasta este punto, se han mencionado - tal vez prematuramente - elementos como inteligencia, Factor g de Spearman, capacidad eductiva, rango de aplicabilidad y versiones de las matrices, poder discriminativo de las mismas, etc., en un contexto bastante desnudo aún.

No obstante, ellos son parte del esquema a partir del cual se desprende la discusión subsiguiente, mediante la cual son, además de desarrollados, hilvanados con la finalidad de exponer esas controversias asomadas anteriormente y, por supuesto, darles un sentido con el cual se puedan apreciar los marcos de referencia conceptual y empírico inherentes a las Matrices y sus desarrollos más recientes, ordenadas a enfocarse en el propósito de este trabajo: la validación de una versión abreviada del *Raven Avanzado*, como se le llama comúnmente.

El Análisis Factorial en la medición de la Inteligencia

A diferencia de los primeros tests mentales elaborados a inicios del siglo XX, como el de Binet y Simon en Francia, los tests psicológicos contemporáneos se han distinguido por su interés en utilizar métodos diferenciales para la medida de la aptitud, de manera tal que permitan un análisis de la ejecución en cuanto a los diferentes aspectos de la inteligencia. Uno de los elementos que ha contribuido al creciente interés por el desarrollo de este tipo de medidas ha sido la aplicación del Análisis Factorial al estudio de la organización de rasgos.

El objeto principal del método radica en simplificar la descripción de los datos, reduciendo el número de las variables necesarias, o dimensiones (Magnusson, 1969/1990). Por medio de dicho análisis, se comparan los diferentes tipos de actuaciones que demandan los tests, se establecen las semejanzas y relaciones funcionales entre esas actuaciones, se efectúa un

análisis interno de la correlación entre las variables acusadas por los tests y así se procura descubrir los factores subyacentes con la intención última de reducir aquellas múltiples actuaciones dadas a un mínimo de factores psicológicos significativos.

Al sustentar sus teorías e instrumentos en este enfoque metodológico, los estudiosos de la inteligencia suponen que ésta se integra por un conjunto de habilidades, las cuales - a su vez - se ven reflejadas en los factores. Así pues, los investigadores se han abocado a explicar la misma como una entidad mensurable por el número, extensión y organización de las habilidades intelectuales.

La primera teoría de organización de aptitudes basada en esta técnica de análisis estadístico fue la Teoría de los Dos Factores o Teoría Bifactorial, desarrollada por el psicólogo británico Charles Spearman (1904, 1927, c.p. Anastasi, 1968/1974 y Sattler, 1988/1996). En su formulación original, esta teoría sostenía que si dos habilidades están correlacionadas entre sí, en cierta medida cada una de ellas ha de incluir dos porciones que expliquen su varianza: una común a ambas - que determina la correlación entre esas habilidades - explicada por el factor común o Factor *g*, y otra específica de cada una - que determina la diferenciación entre ellas - o Factor *e*. Así, tal factor general más uno o más factores específicos por prueba pueden explicar la ejecución en tests de ejecución máxima, donde normalmente está implicada la inteligencia (Sattler, 1988/1996).

Tanto el Factor *g* como el *e*, son pues componentes cuantitativos de la inteligencia. Se diferencian uno del otro en su magnitud; mientras la del primero es intraindividualmente constante e interindividualmente variable, el segundo es variable intra e interindividualmente. La correlación positiva entre dos funciones cualesquiera, se atribuye pues al factor general. Cuanto más saturadas estén dos funciones con el factor *g*, más alta será la correlación entre ellas. Por otra parte, la presencia de los factores específicos tendía a hacer descender la correlación entre las funciones. Aunque en esta teoría, se

afirman dos tipos de factores, el general y el específico, sólo el factor *g* explica la correlación. Por tanto, en contraposición a otras teorías de relación de aptitudes, ésta se podría describir más característicamente como la teoría del factor único, aunque haya persistido la designación original.

Para llegar a estas apreciaciones, Spearman (1923, 1927, c.p. Norris, 1995) había notado que los resultados encontrados en varios tests de habilidad académica tendían a correlacionar bastante alto, en un rango entre 0,70 y 0,80; y sobre la base de estas correlaciones entre las pruebas que medían habilidades tales como aritmética y lectura, Spearman identificó y decidió proponer dos componentes subyacentes que podrían contar para la correlación entre las diferentes ejecuciones en pruebas académicas. Esos dos componentes de *g*, el factor general común que subyace a la inteligencia, fueron identificados como habilidades eductiva y reproductiva.

Dentro de la proposición de Spearman, el factor *g* era considerado como una energía mental general contenida en mayor cantidad en aquellas actividades mentales más complicadas, ya que el mismo está implicado en "operaciones de naturaleza deductiva, ligadas con la habilidad, velocidad, intensidad y extensión de la salida de información intelectual de una persona" (c.p. Sattler, 1988/1996, pág. 57).

Spearman (1904, 1927, c.p. Anastasi, 1968/1974) consideró que si el factor *g* aparecía a través de todas las aptitudes, proporcionaría la única base para la predicción de la ejecución del sujeto al pasar de una actividad a otra. Ya que los factores específicos - por definición - operan solamente en una actividad, su medición no aportaría mayor información acerca del nivel intelectual del individuo. Es por ello, que Spearman sugirió la construcción y utilización de tests que se ocuparan de examinar cómo el individuo resuelve relaciones abstractas, como medio para obtener las mejores medidas de *g*.

Recientemente, esta proposición de Spearman se ha visto apoyada por los resultados de un estudio en el cual se compararon los puntajes de la escala

infantil de inteligencia de Weschler (WISC-R) - una medida de factores *e* o de grupo - y la versión estándar del Test de Matrices Progresivas de Raven (TMPR) - una estimación de factor *g* -, en un grupo multiétnico de niños norteamericanos, para determinar el pronóstico de funcionamiento intelectual de los mismos tras un plazo de nueve años entre la primera y la segunda medida (Saccuzzo y Johnson, 1995).

Los resultados del estudio indicaron que tanto el WISC-R como el TMPR estándar mostraron aproximadamente igual validez predictiva. Ello, fuera del paradigma del análisis factorial, constituye una base para sustentar la idea de la estructura factorial unitaria de la cual están compuestas las matrices progresivas, dado que puede aportar información predictiva tan precisa como lo hacen las medidas de factores específicos, y así puede decirse que los puntajes obtenidos por los sujetos en el TMPR sirven como indicadores válidos que resumen el funcionamiento intelectual, en tanto que aportan información sobre un elemento común subyacente - en mayor o menor proporción - a los demás componentes intelectivos.

Otro argumento generado por disertaciones contemporáneas que apoya la proposición de Spearman acerca de la medición del factor *g* a través de la resolución de problemas que demanden razonamiento abstracto es la discusión que realizan Marshallek, Lohman y Snow (1983, c.p. Carpenter, Just y Shell, 1990) en torno a las matrices progresivas como una herramienta para la estimación del mismo factor.

A este respecto, Marshallek y cols. (1983, c.p. Carpenter y cols., 1990) señalan que una de las fuentes de la centralidad del Raven es que dado el abordaje cognoscitivo necesario para la resolución de las matrices, y el consecuente nivel de complejidad de esta tarea, son mayores los requerimientos de (a) concatenación y control de procesos para estructurar y analizar los problemas, (b) estructuración de una estrategia de abordaje sobre ello, (c) monitoreo del proceso de ejecución y (d) adaptación de dichas estrategias como una conducta de resolución.

Dicha complejidad es lo que asegura que los procesos de necesarios para resolver las matrices constituyan una vía para acceder a datos acerca del factor g, y no que sean expresión de una habilidad intelectual más pequeña que constituya un factor de grupo o específico (Snow, Kyllonen y Marshallek, 1984, c.p. Carpenter y Cols, 1990).

Retomando el curso histórico de la teoría bifactorial de Spearman, el reconocimiento e identificación que el mismo llegara a realizar acerca de la existencia de otra clase intermedia de factores que no fuera tan universal como el "g" ni tan específica como los "e" (que serían los llamados "factores de grupo") no lo llevó a replantear su postura teórica para incluir los mismos por considerarlos "pequeños y despreciables". La Teoría Bifactorial de Spearman y las leyes neogenéticas postuladas por este mismo autor, sirvieron de base a John Raven, en 1956, (Raven, s.f.) para el desarrollo de su Test de Matrices Progresivas, las cuales fueron diseñadas con la finalidad de medir el componente de habilidad eductiva, es decir, la habilidad de una persona para percibir significado a partir de un aparente desorden, desarrollar nuevos discernimientos, y percibir aquello que no siempre es obvio a simple vista (Raven y cols., 1994).

Otros autores - representantes de la corriente americana de los modelos factoriales de la inteligencia - reconocieron la existencia de los factores de grupo en la organización de los rasgos y en el estudio de las habilidades intelectivas. Las Teorías de Factores Múltiples recogen pues todas aquellas posturas teóricas que defienden la existencia de algunos factores de grupo moderadamente amplios, cada uno de los cuales puede entrar con diferentes pesos en los distintos tests.

Uno de los principales exponentes de la Teoría de Factores Múltiples fue Thurstone. Basándose en una extensa investigación realizada por él mismo y sus seguidores, Thurstone llegó a proponer alrededor de una docena de factores de grupo, que denominó aptitudes mentales primarias (French, 1951, Thurstone, 1938, Thurstone y Thurstone, 1941, c.p. Anastasi, 1968/1974 y

Sattler 1988/1992). A partir de ellas, y valiéndose de los factores más frecuentemente corroborados por otros autores, construyó entre otros el *Chicago Test of Primary Mental Abilities (PMA)*, de aplicación colectiva y que suministra un "perfil factorial" que informa sobre los factores V (Verbal), W (Fluidez verbal), M (Memoria), S (Espacial), N (Numérico), R (Razonamiento General o Inducción) y P (Rapidez Perceptiva) (Raven, s.f.).

Algunos de estos factores fueron aceptados posteriormente por Spearman, cuando la investigación ulterior sobre la "Teoría Bifactorial" llevada a cabo por sus seguidores, mostró su existencia e importancia para la comprensión de la capacidad intelectual de un individuo. Entre otros, fueron identificados los siguientes factores: Verbal (sinónimos, antónimos, vocabulario, analogías, etc.), Mecánico (relaciones espaciales), Numérico (aptitud para operaciones numéricas), Memoria (que incluye otros), Lógico (deducción e inducción), Voluntad, Habilidad (capacidad para cambiar rápidamente de una tarea mental a otra distinta), Perseverancia (opuesto al anterior) (Raven, s/f).

En la actualidad, el número de factores cognoscitivos recogidos por los diferentes investigadores sobrepasa los 100. Este hecho responde a la ausencia de distinción básica entre el factor general, los factores de grupo y los factores específicos. El que un factor hallado al realizar el análisis factorial de una batería de test sea calificado como factor específico (por ejemplo), no significa que el mismo se comportará como tal en muestras distintas de baterías de evaluación. Al incorporar los mismos tests dentro de una batería más amplia y heterogénea, el que fuese antes un factor específico puede ahora comportarse como factor de grupo. Por lo tanto, es plausible afirmar que, dependiendo de la amplitud y heterogeneidad / homogeneidad de la batería de pruebas, un mismo factor puede ser interpretado en cualquiera de los niveles (general, de grupo, específico) (Anastasi, 1968/1974).

A pesar de lo inconveniente que podría resultar la proliferación sin límites de diferentes factores que compongan la capacidad intelectual total del

individuo, vale la pena destacar que esta postura teórica presenta la ventaja de ofrecer al examinador un conjunto de puntuaciones en las diferentes aptitudes que proporcionan un perfil intelectual que muestra los puntos fuertes y débiles que caracterizan los componentes intelectivos de la conducta del individuo.

Algunos de los analistas de factores han tratado de simplificar el cuadro de las relaciones entre los rasgos, organizando a éstos en un esquema sistemático. Guilford (1967, c.p. Sattler, 1988/1992), ha propuesto un modelo semejante a una caja, al cual denomina modelo de la estructura del intelecto. El mismo clasifica los rasgos intelectuales en tres dimensiones:

1. *Operaciones*: Lo que hace el examinado (cognición, memoria, producción divergente, producción convergente y valoración).
2. *Contenidos*: La naturaleza de la información sobre la que las operaciones se ejecutan. Estos pueden ser contenidos figurativos, simbólicos, semánticos y de conducta.
3. *Productos*: La forma en que la información es concebida por el examinado. Se clasifican en unidades, clases, relaciones, sistemas, transformaciones y consecuencias.

Dado que esta clasificación incluye $5 \times 4 \times 6$ categorías, existen 120 celdillas en el modelo. En cada una de ellas se espera que haya al menos un factor, pudiendo algunas contener más de uno. Cada factor es finalmente descrito en función de las tres dimensiones. Este modelo, ha sido criticado en tres aspectos fundamentalmente:

1. Es incapaz de reproducir la naturaleza jerárquica de los datos obtenidos de tests de inteligencia, con factores menores, factores mayores y un factor general. Lo que se señala más directamente en cuanto a esto es que Guilford no menciona en su modelo la comunalidad o factor general básico subyacente a toda actividad

intelectual, hallazgo tal que se encuentra presente en la mayoría de los estudios hechos sobre pruebas de inteligencia (Eysenck, 1967).

2. Un gran número de los factores señalados por el modelo de Guilford no muestran una validez externa que no esté explicada por factores generales o grupales (Vernon, 1965).
3. Según Horn y Knapp (1973, c.p. Sattler, 1988/1996), los resultados de análisis factoriales realizados no apoyan la teoría de Guilford.

Un esquema alternativo propuesto para la organización de los factores es el de las llamadas Teorías Jerárquicas (Vernon, 1950, Cattell y Horn, 1967, c.p. Sattler, 1988/1996; Cattell, 1963; Gustafsson, 1984). De modo general, Vernon (1965) resume estas teorías describiendo la estructura de la inteligencia como un árbol genealógico y jerárquico en cuya cúspide queda ubicado el factor g de Spearman. En el nivel siguiente, se encuentran dos amplios factores de grupo, que corresponden a las aptitudes verbal-educativa y práctica-mecánica respectivamente. Estos factores principales pueden, a su vez, subdividirse y ser más reducidos si se prosigue en su análisis. En el nivel inferior de la jerarquía, se hallan los factores específicos.

Luego, Cattell (1963) y Cattell y Horn (1967, c.p. Sattler, 1988/1996) postularon dos tipos de inteligencia: (a) fluída (eficiencia mental esencialmente no verbal y con relativa independencia de la cultura), y (b) cristalizada (habilidades y conocimientos adquiridos que dependen en gran medida de la exposición de la individuo a la cultura). La versión más reciente de este modelo, realizada por Horn (1985), propone una jerarquía de cuatro niveles que en forma ascendente van desde las funciones de detección sensorial visual y auditiva a los procesos de asociación tanto a corto como a largo plazo, los procesos de organización perceptual y los procesos de deducción de relaciones que constituyen las capacidades fluída y cristalizada. Además, sostiene que las capacidades en las jerarquías inferiores correlacionan bajo con las ubicadas en la parte superior del modelo.

Uno de los enfoques integradores dentro de las teorías jerárquicas de la inteligencia, en su rama analítico - factorial, es el de Jan Eric Gustafsson (1984), quien realiza un resumen semejante al elaborado por Vernon en 1965. Así, del mismo modo, ubica en el nivel más alto el factor g o inteligencia general concebido de la misma manera que Spearman. En el nivel medio, coloca tres factores amplios, el de inteligencia cristalizada, inteligencia fluida y visualización general.

Dentro del modelo de Gustafsson, la inteligencia fluida es en esencia la misma que la inteligencia general y se encuentra asociada a las capacidades adaptativas no verbales. Por su parte, la inteligencia cristalizada es concebida por Gustafsson como la representación de una dimensión relativamente pequeña del conocimiento que se generaliza menos a las subsecuentes situaciones de solución de problemas y aprendizaje, en comparación con la inteligencia fluida. Finalmente, en el nivel inferior de la jerarquía, se encuentran los factores primarios, los cuales básicamente se corresponderían a los señalados por Thurstone (1938, c.p. Sattler, 1988/1996) y Guilford (1967, c.p. Sattler, 1988/1996).

Lo que se advierte hasta ahora, es que todas estas posturas acerca de la medición y constitución del constructo inteligencia se ubican en torno a dos elementos principales, a saber: (a) la base metodológica de la cual parten, (b) la base conceptual que proveen para el desarrollo de los tests.

El punto (a) es el relativo a la Teoría Factorial, que provee un conjunto de operaciones cuantitativas para esclarecer, mediante la lógica de la correlación y de la varianza, cuáles son las dimensiones subyacentes a los elementos que constituyen la muestra de conducta de un test y, así, concluir sobre la naturaleza de la inteligencia que es propuesta por cada marco conceptual. En torno al segundo aspecto, cada una de estas teorías provee una concepción de lo que es la inteligencia y, así mismo, sobre su base se han generado los diversos procedimientos de medición de tal constructo.

Se establece un proceso cíclico entre las teorías substanciales de la inteligencia y la metodología de estudio desarrollada en relación a las mismas, en el cual (a) se parte de la proposición de un concepto y descripción de lo que es capacidad intelectual (factor g ó e , factores de grupo, jerarquía de factores, etc.), (b) se elaboran instrumentos que intentan dar cuenta de la existencia de una forma particular de inteligencia y, en conjunción con todo ello, (c) se utiliza el análisis factorial como paradigma de análisis de los datos extraídos de (b), dando lugar a la verificación o no de lo supuesto desde (a), donde se dio origen a ese ciclo.

Test de Matrices Progresivas de Raven

Una de las pruebas de inteligencia más conocidas y utilizadas (Arthur y Day, 1994; Raven, s.f.) - sin escapar al proceso cíclico antes descrito - fue la elaborada por Raven, quien construyó su Test de Matrices Progresivas con el objetivo de medir el Factor g , es decir, el factor común y fundamental en todas las funciones cognitivas humanas (Raven, s.f.). Como fuera mencionado anteriormente, la magnitud de este factor es individualmente constante e interindividualmente variable: constante en todas las habilidades de un individuo y ampliamente variable de un individuo a otro.

La intención de medir g a través del test de Raven, se basa en la concepción del razonamiento como un componente de las capacidades de procesamiento mental que pueden ser independientes del contenido y del contexto en el cual se enmarcan los problemas que resuelve una persona, pues las matrices progresivas constituyen muestras de conducta de razonamiento abstracto, por lo que se consideran medidas ideales de tales procesos (Richardson, 1991).

En este punto, es oportuno señalar un apareamiento entre dos nociones, una relativa al factor g y la otra a las matrices de Raven como herramienta de medida de éste. Raven (s/f) describe su prueba como una tarea de ejecución

máxima que debe ser resuelta con claridad de pensamiento. Por otra parte, Sattler (1988/1996) comenta que las pruebas con altas cargas de *g* normalmente solicitan un esfuerzo mental consciente por parte del individuo que las resuelve, como el involucrado en el razonamiento, la comprensión y las tareas de prueba de hipótesis; actividades éstas que se desempeñan durante la ejecución en las matrices.

Igualmente, se ha reportado que la estructura del instrumento creado por John Raven ha permitido obtener correlaciones entre los puntajes del TMPR y otras medidas de logro intelectual que sugieren que el proceso fundamental es de naturaleza general en lugar de específica (Court y Raven, 1982, c.p. Arthur, Barrett y Doverspike, 1990).

Para medir la inteligencia como un constructo unitario, Raven diseñó 60 láminas de figuras geométricas abstractas lacunarias (ordenadas en cinco series denominadas A, B, C, D y E, de 12 ítems cada una), en complejidad creciente, y cuya resolución requiere del examinado la educación de relaciones y precisión en la discriminación en las primeras 24 matrices, y educación de correlaciones, comprensión de analogías, permutación, alternación del modelo y otras relaciones lógicas en las últimas 36. Ésta constituye la versión estándar o general para adultos, aplicable a individuos cuyas edades oscilan entre los 12 y 65 años de edad. En su primera estandarización, realizada en 1936, se dan normas percentiles para cada intervalo de medio año entre los 8 y catorce años, y para intervalos de 5 años, entre los 20 y 65. Estas normas por edad se basaron en muestras británicas que incluyeron 1.407 niños, 3.665 adultos en el servicio militar a quienes se aplicó el test durante la segunda guerra mundial y 2.192 adultos civiles (Anastasi, 1968/1974).

A partir de la versión estándar, se desarrolló una medida de la habilidad mental general de orden superior que constituye la versión avanzada (TAMPR), destinada a personas con puntajes superiores al promedio, es decir, una progresión natural para personas que podrían considerar la forma general como muy fácil y obtienen puntajes por encima de 50/60; y, para su

resolución, sus 36 ítems deben ser abordados por el sujeto educiendo únicamente correlaciones; es decir, ante un reactivo y una relación, el individuo debe hallar el ítem correlativo (Raven, s.f.).

Esta última versión de las Matrices Progresivas, como su predecesora, consiste en una serie de ítems homogéneos en dificultad creciente, y requiere del sujeto la escogencia de la pieza - opción (entre otras ocho) que completa mejor una serie de patrones presentados a través de las hileras de diseños (matriz progresiva) (Arthur y Day, 1994).

El APM fue diseñado originalmente en 1943 para ser usado por la British War Office Selection Boards, la cual necesitaba una versión de las Matrices Progresivas que fuera más difícil, y proveyera mejor discriminación en los niveles superiores que la versión estándar del test. Las Matrices Progresivas Avanzadas fueron entonces preparadas para diferenciar entre las personas de habilidad intelectual superior, y pueden ser usadas para selección de técnicos de alto nivel, posiciones gerenciales y estudiantes de cursos avanzados tanto técnicos como científicos. Así, el APM es capaz de proveer información útil sobre la tasa probable de progreso y eventual éxito en la educación y el trabajo (Matarazzo, 1990).

El APM tiene dos componentes o sets, publicados en cuadernillos separados. El Set I comprende un pequeño conjunto de 12 ítems que cubre todos los procesos intelectuales mostrados en el test completo, y que puede servir de varios propósitos:

1. Como conjunto de ítems de práctica, para que los sujetos se familiaricen con el test y disminuyan su ansiedad. Si se considera deseable, a las personas a ser evaluadas con el Set II se les puede dar el Set I para que lo lleven a casa y practiquen, y minimizar los efectos de la práctica previa, adiestramiento y familiaridad.
2. Como un test muy corto, para clasificar a los adultos en el 10% superior, el 80% promedio y el 10% más brillante.

3. En situaciones de incertidumbre, para decidir si aplicar subsiguientemente la versión de matrices estándar o la avanzada. Sobre todo si se requiere discriminar bien sea entre niveles de habilidad general o alta.

El Set II, de 36 ítems, es similar en presentación y procesos de razonamiento requerido a los elementos del Set I, pero en el Set II los problemas se incrementan marcadamente en dificultad y llegan a ser mucho más complejos. Al contrario de las versiones estándar y las computarizadas del Raven desarrolladas últimamente, en las que se contabiliza el tiempo, las matrices Avanzadas pueden ser administradas considerando el tiempo o no. Si se requiere probar la capacidad total de razonamiento de una persona, el Set II se lleva a cabo sin límites de tiempo; pero si se requiere evaluar la eficiencia en el razonamiento - lo cual acentúa su carácter como test de ejecución máxima -, se aplica el Set II con límites de tiempo (Raven y cols., 1994).

A través de los problemas del Raven, Carpenter y cols. (1990) encontraron cinco tipos diferentes de reglas que gobiernan cada uno de los elementos que componen las matrices; a saber:

1. Constante en una fila: el mismo valor ocurre a través de toda una fila, pero ese valor cambia en la columna siguiente (ver Anexo A, figura B, donde la localización del componente oscuro es constante dentro de cada fila; en la primera fila, la localización está en la mitad superior del diamante; en la fila del medio está en la mitad inferior del diamante; y en la última fila está en ambas mitades).
2. Progresión cuantitativa: ocurre un incremento o decremento cuantitativo en atributos tales como tamaño, posición o número, entre las figuras adyacentes en las matrices (ver Anexo A, figura A, donde el número de cuadrados negros en cada elemento se incrementa a lo largo de la fila de uno a dos a tres).

3. Adición o sustracción de elementos de la figura: el primer elemento de una columna es añadido a (yuxtapuesto o superpuesto) o sustraído del segundo elemento para producir el tercero (ver Anexo A, figura B, donde el elemento geométrico en la columna uno yuxtapuesto al elemento en la columna dos, produce el elemento en la columna 3).
4. Distribución de tres valores: tres valores de un atributo categórico o nominal (tal como un tipo de figura) se distribuyen a través de una fila (ver Anexo A, figura D, donde las tres formas geométricas – diamante, cuadrado y triángulo – siguen una regla de distribución, y las tres texturas de líneas – negra, rayada y vacía – también siguen una regla de distribución).
5. Distribución de dos valores: dos valores de un atributo categórico se distribuyen a través de una fila y el tercer valor es nulo. (ver Anexo A, figura C, donde los elementos geométricos tales como la línea vertical, la línea horizontal y la “V” en la primera fila, siguen una distribución de dos valores).

La resolución de cada matriz puede estar gobernada por varias reglas al mismo tiempo, sobre todo a medida que se avanza en el test y aumenta correlativamente el nivel de dificultad de los problemas. Tale reglas deben ser consideradas todas por el sujeto para converger en la alternativa de respuesta correcta.

Algunas de las alternativas de respuesta incorrectas, están diseñadas para satisfacer un conjunto de reglas incompletas. Por ejemplo, si un sujeto indujo la regla A pero no la B o la C (ver Anexo A, figura D) podría escoger la alternativa dos u ocho. De igual modo, induciendo la regla B pero omitiendo la A y la C llevaría al sujeto a elegir la alternativa tres, lo cual resultaría de nuevo en una respuesta incorrecta.

Con el fin de obtener más datos sobre la ejecución en el test de Raven y revelar más sobre los procesos y el contenido del pensamiento durante la solución de cada uno de los problemas del mismo, Carpenter y cols. (1990) diseñaron un experimento en el cual le pidieron a 22 estudiantes del Carnegie Mellon y de la Universidad de Pittsburgh (quienes participaron por un pago de \$10) que describieran las reglas que ellos inducían en la escogencia de las respuestas de cada problema, es decir, se tomó un protocolo verbal de su ejecución en el Raven Avanzado.

Se obtuvo que la forma en que los sujetos resolvían un problema era descomponiéndolo en diferentes sub-problemas sucesivamente más pequeños procediendo entonces a resolver cada uno de éstos. La inducción de las reglas fue incremental en dos aspectos, primero en los problemas que contienen más de una regla, éstas eran descritas una a la vez, con largos intervalos entre las descripciones de la misma, lo cual confirma que eran inducidas una por vez. Segundo, la inducción de cada regla consistía en muchos pasos pequeños, reflejados en la comparación de figuras adyacentes o vecinas. Estos aspectos de procesamiento incremental fueron características ubicuas en todos los sujetos cuando éstos intentaban solucionar los problemas, y no pareció ser una fuente de diferencias individuales.

Por otra parte, también es importante conocer los errores cometidos en los intentos por resolver cada ítem y cómo éstos se distribuyen entre las ocho figuras dadas como posibles soluciones (ver Tabla 1). Como el APM se diseñó para ser esencialmente un test de claridad, precisión y pensamiento, si es necesario bajo presión de tiempo, el objetivo principal de comparar los errores es asegurar (a) que desde el momento en que cualquier ítem llega a ser "psicológicamente problemático" no haya figuras erróneas que engañen o guíen por mal camino, y (b) que, antes de que este punto sea alcanzado, la gente no escoja las respuestas correctas a pesar de un razonamiento equivocado.

Se han identificado cuatro tipos de errores (Tabla 1):

1. Soluciones incompletas (correlación incompleta): estos son errores que provienen del fracaso de entender todas las variables que determinan la naturaleza de la figura correcta requerida para completar el patrón del problema. En lugar de ello, se escoge una figura que es sólo parcialmente correcta.
2. Líneas de razonamiento arbitrario (principio equivocado): la figura escogida sugiere que la persona evaluada está usando principios cuantitativos de razón cualitativamente diferentes de las demandadas por el ítem.
3. Elecciones sobredeterminadas (confluencia de ideas): estos son errores que involucran fracaso en determinar la cualidades irrelevantes en la figura escogida, y entonces se selecciona una que combine tantas características individuales como sea posible de las mostradas en la matriz a ser completada. Esto puede ser visto como una forma de pensamiento sobreinclusivo.
4. Repeticiones: involucra selección de una figura idéntica a una de las tres figuras inmediatamente adyacentes al espacio a ser completado en la matriz.

Por otro lado, cabe esperar que este nuevo test que constituye la forma avanzada de las pruebas de Raven, al estar elaborado sobre los mismos supuestos de homogeneidad de los elementos y libertad de factor verbal o carga cultural, conforme un instrumento de estructura factorial unitaria y similar a su predecesor estándar.

Tabla 1. Posición (alternativa incorrecta) y naturaleza de los errores más comunes.¹

A: correlaciones incompletas
B: principio equivocado

C: confluencia de ideas
D: repetición

Problema número	Porcentaje de aciertos	Figuras correctas	Error más común		2° error más común	
			Posición	Naturaleza	Posición	Naturaleza
1	98	5	2	A	1	A
2	96	1	2	D	6 - 7	C - A
3	93	7	6	C	4	B
4	91	4	2	B	8	B
5	90	3	7	A	1	A
6	89,5	1	7	D	2	D
7	89,5	6	2	C	4	D
8	87	1	4	D	5	C
9	88	8	3	B	4	B
10	88	4	8	B	7	C
11	84	5	4	B	1	A
12	84	6	5	B	4 - 7	B - B
13	79	2	7	A	6 - 8	A - D
14	80	1	5	B	4	B
15	79	2	4	C	6	A
16	73	4	1	B	5	C
17	69	6	3	A	4	A
18	65	7	5	A	1	A
19	64	3	5	A	8	A
20	59	8	2	A	4	C
21	58	8	1	A	4	B
22	50	7	8	C	2	A
23	49	6	8	A	5	B
24	49	3	1	B	8	B
25	45	7	2	B	8	C
26	37,5	2	6	A	5	A
27	38	7	4	B	-	-
28	32	5	4	A	3	A
29	29	6	2	A	7	A
30	26	5	4	A	8	A
31	24	4	8	A	2	A
32	16,5	8	4	A	5	A
33	24	5	7	A	4	B
34	13	1	3	B	2	A
35	14	3	1	B	4 - 7	B - C
36	5	2	1	B	8	B

Para corroborar tal hipótesis, Arthur y Woehr (1993) llevaron a cabo un análisis factorial confirmatorio para examinar la dimensionalidad del TAMPR; así, se trabajó partiendo de dos postulados posibles para explicar la estructura de factores de este instrumento, a saber: (a) un modelo de dos componentes factoriales y (b) un modelo de factor único, los cuales comportan dos puntos de vista discrepantes acerca de la estructura subyacente al test en cuestión.

¹ Fuente: Raven, Court y Raven, 1994, página APM12.

De esta manera, luego de haber sido aplicada la prueba a 363 estudiantes pregraduados de Texas A&M University, se obtuvo de los datos que la separación de los reactivos en dos subtests no resultaba en niveles más altos de homogeneidad de los ítems, caso tal que tendría lugar si los mismos constituyeran factores distintos. En cambio, se mantuvo la hipótesis del modelo (b) como la mejor forma de representar la estructura factorial del TAMPR, sobre todo por el alto grado de homogeneidad existente entre los ítems del mismo (Arthur y Woehr, 1993).

Otras investigaciones han igualmente convergido en la conclusión de que el test de Raven mide procesos que son centrales en la inteligencia analítica. Ésta se entiende como la habilidad para razonar y resolver problemas que envuelven información nueva no relacionada con una base explícita de conocimiento derivado de la escolaridad o de la experiencia previa, lo cual fue etiquetado por Cattell (1963) como inteligencia fluida; en contraste con la inteligencia cristalizada, la cual refleja más directamente el conocimiento adquirido previamente y las habilidades que han sido cristalizadas.

Las diferencias individuales en el test de Raven correlacionan altamente con aquellas encontradas en otros tests cognoscitivos complejos. Por ejemplo, según Snow y cols. (1984, c.p. Carpenter y Cols, 1990) la centralidad del test de Raven entre los tests psicométricos se puede ilustrar gráficamente en el centro de una diana junto a otros tests de razonamiento complejo, mientras que otros tests más simples pueden localizarse hacia la periferia y pueden ser dispuestos espacialmente de acuerdo a su contenido (Anexo B). Así, la complejidad es máxima cerca del centro - donde se ubica el Raven - y disminuye hacia la periferia. Los tests en el anillo que circunda al Raven involucran razonamiento abstracto, inducción y/o deducción de relaciones. Para los tests de complejidad media, hay una ubicación espacial como una función del contenido del test, con celdas separadas para tests espaciales, numéricos y verbales.

Desarrollos recientes de las Matrices Progresivas

Recientemente otros autores han construido formas más elaboradas del TMPR original. Una de ellas se refiere al uso de las matrices progresivas como un instrumento de evaluación por computadora (Kubinger, Formann y Farkas, 1991). Para el desarrollo de esta modalidad se escogieron sujetos de grupos tan distintos como niños en edad escolar, pacientes psiquiátricos y empleados federales, a quienes les fue aplicada la versión alemana, tanto la de lápiz y papel como la computarizada. El objetivo final era establecer la equivalencia entre ambas formas de explorar la inteligencia y, a pesar de que los sujetos de todos los grupos obtuvieron en promedio en la versión computarizada 13 puntos de CI más que en la forma original del test, se conservaron todas las propiedades psicométricas a través de las dos versiones (confiabilidad, estructura factorial unitaria, orden de dificultad creciente de los ítems, etc.).

Una línea de investigación que intentó proveer una perspectiva general en cuanto al uso de procedimientos de evaluación por computadora fue el Proyecto Leicester/DHSS (Beauman y French, 1987). En este estudio de campo clínico de ocho procedimientos psicométricos automatizados, Beaumont y French examinaron cómo los diversos tests afrontarían los rigores de las condiciones clínicas de evaluación en la vida real. La administración de los ocho tests se condujo a través de 367 sujetos en cinco ambientes clínicos, usando las presentaciones computarizadas de los tests existentes, entre ellas las Matrices Progresivas de Raven en su versión estándar.

En este estudio se demostró que era posible producir versiones por computadora psicométricamente paralelas de tests existentes, pero se encontró que algunos de ellos eran menos dóciles a la computarización, en el sentido de que se produjeron diferencias significativamente diferentes entre las versiones de papel y lápiz respecto a las basadas en computadora, aunque lograran buenos niveles de confiabilidad. En el caso de las Matrices Progresivas, éste fue uno de los tests en el que se obtuvieron puntajes más altos en la versión convencional. No obstante, los autores (Beaumont y French, 1987) destacaron

que del conjunto de pruebas usadas, el test de Raven fue uno de los dos en que se contabilizó el tiempo con un reloj presentado en la pantalla, lo cual posiblemente incrementaba el sentido de urgencia en los sujetos durante la resolución de las matrices.

A pesar del buen desempeño de las Matrices de Raven en este estudio, en términos de puntajes y confiabilidad, Beaumont y French consideraron que las diferencias absolutas entre las versiones indicaron que el formato automatizado podría no ser usado rutinariamente en lugar del tradicional basado en papel y lápiz. Los investigadores propusieron que las dificultades encontradas en las diferencias entre formatos se podrían resolver si se hicieran mejoras en la calidad gráfica de los ítems del test, y así proporcionar una verdadera versión del test psicométricamente paralela.

Como se puede observar, en los dos estudios anteriormente descritos (Kubinger y cols., 1991; y, Beaumont y French, 1987) se encontró que las evaluaciones con las nuevas versiones por computadora arrojaban puntajes significativamente diferentes a los obtenidos con la versión convencional, bien por exceso o por defecto; aunque el funcionamiento psicométrico de las Matrices se sigue considerando paralelo en todos los casos. Sin embargo, sería prudente agregar que la percepción del formato utilizado, en cualquiera de sus bases, debe ser también paralelo, que a fin de cuentas fue lo que sugirieron Beaumont y French a partir de la discusión de sus resultados.

Estos nuevos formatos computarizados de las pruebas psicométricas parecen tener una importante influencia en el tipo de resultados que se obtienen de las evaluaciones y, tal como lo sentencia Norris (1995), se comprende que para obtener una concepción más precisa del valor que tiene la evaluación por computadora, es necesario encontrar un procedimiento validado y bien probado que haya sido transferido de su formato existente de papel y lápiz a un formato computarizado. Entre ellos, comenta Norris, "uno de los instrumentos que ha ganado mayor aceptación a nivel mundial en términos de validez y confiabilidad son las diferentes formas de las pruebas de Raven";

por lo que cabe destacar que lo mismo puede considerarse a la hora de proponer o validar cualquier otro tipo de versión, v.g. una forma abreviada, ya que ello hace más viable el paralelismo psicométrico y la aplicabilidad a una nueva población.

Otro trabajo realizado a partir del TMPR fue llevado a cabo en la Universidad de San Diego (California, EEUU), consistente en el desarrollo y validación de una forma paralela del test de Raven (Johnson, Sacuzzo y Guertin, 1994). Con tal objeto, se elaboró la nueva versión tras construirse 60 nuevos reactivos, cada uno de los cuales correspondía a los otros tantos que contiene la versión estándar original, siendo comparables los unos y los otros en términos de la estrategia subyacente requerida para solucionar los ítems - como fuera explicado anteriormente - y el nivel de dificultad de los mismos. Una vez que ambas formas fueran aplicadas a un grupo de 449 niños entre 6 y 13 años de edad, procedentes de comunidades diversas en el área de San Diego, la nueva versión paralela mostró un coeficiente de confiabilidad de 0,90, al mismo tiempo que el coeficiente de consistencia interna de Kuder-Richardson alcanzó niveles idénticos en ambas formas del test - 0,94 -. Ello es indicador del éxito alcanzado por el estudio, en tanto que se obtuvo, además de una prueba paralela de tan buena calidad psicométrica como la original, una nueva herramienta de gran utilidad para la investigación.

Además de formas o versiones distintas del instrumento en sí, como la computarizada y la paralela de lápiz y papel mencionadas, también se probaron las propiedades psicométricas del TAMPR con un procedimiento de aplicación alternativo, mejor conocido como la administración dinámica (Larson, Alderton y Kaupp, 1991). Luego de examinar 808 reclutas de la marina norteamericana utilizando este procedimiento, los resultados indicaron que a pesar de que los puntajes de las matrices progresivas se incrementaron significativamente, las propiedades psicométricas del instrumento permanecían estables - confiabilidad y validez de constructo -. En consideración del tiempo que consume la aplicación del TAMPR a través del procedimiento dinámico, y

dada la ausencia de beneficios psicométricos que superen al estilo original de administración, no parece haber razón alguna por la cual sustituir esta última por la primera.

El uso de este test - el TMPR original - en diversos países europeos indicó también la aplicabilidad de las normas dadas a los mismos. Si bien Raven (s.f.) construyó las matrices progresivas con la finalidad de medir g con la menor carga posible de factor cultural, dado que su comprensión y resolución no requieren actividad verbal alguna por parte del individuo, Anastasi (1968/1974) señala que estudios realizados sobre algunas culturas no europeas han hecho surgir dudas sobre la adecuación de este test para grupos de medio ambiente muy distinto, llegando incluso a considerar que el test refleja (a) la cuantía de la educación y (b) que era susceptible a una considerable práctica.

No obstante, un ejemplo empírico que contrasta con esta afirmación, se observa en una revisión de las aplicaciones más frecuentes del TMPR en Asia, particularmente en Taiwan, China, Hong Kong, Singapur, Filipinas y el subcontinente indio (Court, 1991). Según éste, tanto la versión estándar, como la coloreada y la avanzada del Raven, arrojaron resultados más satisfactorios que otros tests comúnmente utilizados en países occidentales desarrollados, que sí necesitan de la mediación del lenguaje para ser comprendidos y resueltos, de manera tal que el éxito de los sujetos sometidos a dichas evaluaciones comportaba mayor dependencia de las oportunidades culturales a las que habían estado expuestos. Así, la utilidad de las pruebas de Raven como medida de habilidades generales se vio reforzada en el contexto asiático.

En torno a la mediación del lenguaje, se probaron los efectos de la verbalización concurrente en la resolución de las matrices progresivas de Raven en la versión avanzada por computadora (DeShon, Chan y Weissbein, 1995), para determinar si la ejecución a través de todos los ítems es independiente de los procesos cognoscitivos involucrados en el habla,

considerada como expresión del pensamiento. Así, una muestra de estudiantes universitarios fueron asignados aleatoriamente a dos grupos, uno de *verbalización* (42 Ss) y otro de *no verbalización* (35 Ss) durante la realización de las tareas de correlación en el test, mientras que a los 90 Ss de control les fue aplicada la versión original de lápiz y papel. A los estudiantes de la muestra de verbalización se les instruyó que describieran las matrices y que dijeran cualquier cosa que les viniera a la mente, aún cuando no fuera relevante, mientras ellos iban resolviendo las matrices.

Los resultados de esta investigación, mostraron que la ejecución de los sujetos en cualquiera de los grupos de no verbalización no diferían significativamente, mientras que el desempeño de los sujetos con verbalización fue significativamente más bajo. De esta manera, se puede decir que las matrices progresivas de Raven pueden ser resueltas mediante procesos visoespaciales, y no dependen tanto de los procesos de análisis verbal como mediación cognoscitiva o apoyo compensatorio para su resolución, ya que mas bien pueden constituir un elemento distractor durante la ejecución de esta tarea.

En otro contexto, Templer (1992) realizó un estudio con la versión estándar para el Centro Correccional del Desierto del Sur en el estado de Nevada (EEUU) y estableció normas para un grupo de prisioneros compuesto por ciudadanos estadounidenses de razas blanca y negra en su mayoría, más otros de origen mexicano, cubano, asiático e indoamericano. Con tal finalidad, se obtuvo información normativa para (a) el grupo de prisioneros blancos, (b) negros y (c) todos los grupos étnicos en conjunto, dividiendo cada conjunto de normas en tres categorías de edad. De esta manera, se encontró que aún cuando los sujetos con edad menor a 35 años puntuaron más alto que los mayores, independientemente del grupo racial, existe en general un solapamiento substancial en la distribución de los puntajes de sujetos blancos y negros. Esto señala que, a pesar de las diferencias que puedan tener los sujetos en cuanto a las oportunidades culturales a las que son expuestos

durante su experiencia de vida (según el grupo étnico del cual procedan), las diferencias entre ellos en la resolución de las matrices progresivas no alcanzaron el nivel de significancia necesario para apoyar la aseveración de Anastasi sobre la no independencia de factor cultural de las Matrices.

Un punto de equilibrio encontrado entre las posiciones expuestas por Anastasi y Raven en relación a la independencia de factor cultural del TMPR, lo establece un estudio realizado con la finalidad de determinar la adecuación de la prueba de Raven para varios grupos étnicos en Suráfrica (Owen, 1992). Fueron seleccionados estudiantes de raza blanca, no blancos, aborígenes y negros, de la escuela secundaria, a quienes les fue aplicada la versión estándar. Una vez obtenidos los datos, se encontró que las propiedades psicométricas del instrumento se conservaban similares a través de las muestras grupales, dado que los mencionados grupos étnicos mostraron diferencias muy pequeñas en confiabilidad, el orden de dificultad creciente de los ítems y sus poderes discriminativos correspondientes, así como la carga del factor principal para cada uno de los elementos. Estos resultados apoyan la posición sostenida por Raven acerca de la independencia de factor cultural en el TMPR.

En cambio, a pesar de la similitud observada en las propiedades psicométricas del test aplicado a los cuatro grupos muestrales, fue advertido un comportamiento particular de los datos dentro de cada conjunto muestral, específicamente entre los grupos de estudiantes negros y los blancos. En este sentido, los ítems que arrojaron el mayor poder discriminativo dentro de los subgrupos de sujetos, resultaron ser los mismos que establecían mayores diferencias entre los subconjuntos raciales de la muestra total; por lo tanto, Owen (1992) concluyó que, a pesar de las propiedades similares del test a través de los grupos, no resulta adecuado hacer uso de normas comunes para blancos y negros debido a las grandes diferencias de medias obtenidas a favor de los primeros.

De esta manera, mientras no haya resultados consistentes que permitan establecer alguna conclusión acerca de la mayor o menor carga de factor cultural que puedan poseer las matrices progresivas, por el momento pareciera que lo más sensato es que los investigadores se dediquen a elaborar las versiones de la prueba que se ajusten a las necesidades de las distintas poblaciones de estudio, y establecer la información normativa que corresponda en cada caso. Pues si bien se ha encontrado que las propiedades psicométricas ciertamente se mantienen, no queda claro aún si los puntajes medios logrados por los distintos grupos de referencia son del todo diferentes o no. Es allí, cuando los aplicadores de pruebas, sobre todo aquellos que buscan datos con fines de diagnóstico diferencial y metas de intervención - como en las áreas clínica, educativa y organizacional - deben estar conscientes de si la evaluación que realizan a través de la información normativa extraída de otros contextos, es justa o no para los sujetos con los cuales se está tratando.

Un apoyo empírico de las aseveraciones hechas anteriormente lo constituye la estandarización polaca del TMPR en sus tres formas, realizada por Jaworska y Szustrowa (1993), a partir de la cual fue necesario establecer normas para los siguientes grupos de referencia:

1. Forma estándar (versión de 1956): (a) niños de 5 años y 11 meses a 15 años y 11 meses de edad; (b) jóvenes y adultos de 16 años en adelante.
2. Forma avanzada (versión de 1962): (a) estudiantes de primaria y secundaria entre 12 años y 6 meses a 19 años y 5 meses de edad; (b) estudiantes de todas las universidades polacas.
3. Forma coloreada: niños entre 3 años 11 meses a 9 años 11 meses de edad.

Nuevamente se encontró que las propiedades psicométricas del instrumento se mantienen, no ocurriendo lo mismo con los puntajes obtenidos por las distintas poblaciones a través del mismo, al igual que en el estudio

normativo llevado a cabo por Measso, Zappalà, Cavarzeran, Crook y cols. (1993) de la versión coloreada del TMPR para adultos sanos en los países escandinavos. Estos investigadores aplicaron el test a 894 adultos seleccionados aleatoriamente, y observaron que el género, la edad y el nivel de educación influenciaron significativamente la ejecución tanto en el test completo como en los diferentes subsets que constituyen al mismo, razón por la cual se requirió el establecimiento de normas distintas que consideraran las variables mencionadas.

La versión avanzada de las Matrices Progresivas (TAMPR), es una de las que ha sido más ampliamente usada con adultos de habilidad mental general superior (Raven, Court y Raven, 1985, c.p. Arthur y Day, 1994). Esta versión también ha sido utilizada para identificar potencial académico en estudiantes pertenecientes a sectores no representativos de la población norteamericana (Mills y Tissot, 1995). En tal investigación, se comparó la ejecución de 347 estudiantes de 9° grado en el TAMPR con el School and College Ability Test, con el objeto de identificar aquellos estudiantes de poblaciones minoritarias que podrían ser promovidos a cursos de nivel avanzado con alta probabilidad de éxito. Los resultados arrojaron una alta capacidad del TAMPR para identificar potenciales académicos en minorías étnicas.

Mills, Ablard y Brody (1993), así mismo, afirman que existe evidencia acerca de que la inclusión de cualquiera de las tres formas de las Matrices Progresivas (estándar, coloreada o avanzada) en una batería de evaluación ciertamente puede ampliar el rango de habilidades examinadas, particularmente en el caso de estudiantes talentosos. No obstante, también comentan que, al constituir las MPR un estimador de factor g, su utilidad se limita cuando hay necesidad de establecer diagnósticos diferenciales y/o pronósticos de estos niños, cuando presentan - por ejemplo - condiciones paralelas como dificultades de aprendizaje, específicas o evolutivas.

Esta misma idea se refuerza en el contexto de los servicios de rehabilitación, cuando el TMPR es utilizado dentro de una batería de

evaluación destinada a la orientación vocacional de usuarios de este servicio, tales como personas con amputaciones, impedimentos ortopédicos, o que padezcan alguna enfermedad mental o trastorno de personalidad. En tal sentido, Kern, Bordieri y Taylor (1993), aplicaron en forma aleatoria tanto la versión estándar de las matrices progresivas como el WAIS-R a un grupo de 64 sujetos, con edades comprendidas entre 17 y 64 años, con las características anteriormente mencionadas, a fin de determinar la utilidad de sustituir el WAIS-R por el TMPR, dadas las ventajas que posee este último con respecto al tiempo de aplicación. Sin embargo, se concluyó que cada test por separado aporta información valiosa, por lo que se considera importante la inclusión de ambos en una batería de evaluación para asesoramiento vocacional.

Ello se debe a que la validez predictiva del TMPR y el WAIS-R, si bien están relacionadas (Kern y cols., 1993), en el sentido de que a partir de ambos instrumentos se puede establecer el nivel de funcionamiento futuro de una persona, como en el caso de niños con dotación superior (Mills y cols., 1993 y Mills y Tissot, 1995), usuarios de servicios vocacionales (Kern y cols., 1993), etc., la utilidad del pronóstico que cada uno aporta es aprovechable a escalas distintas. Tal situación se origina precisamente en lo que intenta medir cada uno de estos tests, pues como se sabe, el TMPR está diseñado para estimar nivel intelectual general, mientras que el WAIS-R lo está para la medición de habilidades intelectivas específicas o de grupo. En este sentido, se prevé que la validez de criterio de las pruebas de Raven se ubica a nivel general, es decir, a partir de sus puntajes se puede predecir el estándar global de funcionamiento futuro de alguien que haya sido evaluado con esta herramienta; en cambio, no aporta información sobre cuáles habilidades puede utilizar la persona con mayor probabilidad de éxito para su adaptación o sobre qué áreas de funcionamiento puede abocar su capacidad intelectual, tal como lo puede expresar un perfil de un protocolo de WAIS-R.

Se puede observar que la intención original por la cual fue desarrollado el TAMPR - medir habilidad mental superior en adultos - no ha sido limitación

para llevarlo a distintos contextos. Por ejemplo, ha sido objeto de investigaciones en ámbitos organizacionales, como la validación para conductores de transporte de productos petroleros (Arthur y cols., 1990) y para controladores aéreos (Ackerman, 1992). La versión avanzada también es considerada útil para poblaciones cuyo procesamiento de lenguaje debe ser minimizado en la evaluación (Van Den Broek y Bradshaw, 1994).

El TAMPR también ha encontrado utilidad predictiva en otros ámbitos, como el psicopatológico. En tal sentido, el formato de las matrices progresivas, y el hecho de que es un test usado para evaluar habilidad de razonamiento en modalidad visual, las hace ideales en situaciones de diagnóstico clínico, pues la principal atracción de usar el TMPR para evaluar habilidad intelectual es que las respuestas no requieren necesariamente verbalización ni habilidades de manipulación. El modo de presentación del test y el contenido, y los métodos de los sujetos para responder claramente lo convierte en una técnica de evaluación de amplio rango, que puede incluso ser usada en un espectro de edades que va desde los más jóvenes a los mayores, y puede incluso abarcar varias discapacidades físicas tales como sordera o afasia, así como también capacidad intelectual sub-normal (Norris, 1995).

Respecto al ámbito clínico, se encontró un estudio británico que relacionó el Test Nacional de Lectura para Adultos con el TAMPR con fines de detectar déficit adquiridos en inteligencia general de adultos. Como resultado se obtuvo que la prueba de lectura tiene mayor probabilidad de diagnosticar falsos - negativos, mientras que las Matrices Progresivas fueron más sensibles para detectar tales déficit adquiridos por neuropatología (Van Den Broeck y Bradshaw, 1994). Esto nuevamente habla del rango de aplicabilidad y poder discriminativo de las matrices de Raven respecto a otras pruebas, cuya centralidad en la estimación de Factor g es de carácter periférico, siguiendo el modelo representativo presentado por Snow y cols. (1984, c.p. Carpenter y Cols, 1990) (Anexo B).

Desarrollo de las versiones cortas

Es importante considerar no sólo el valor heurístico de las medidas de la inteligencia, como el TAMPR, sino otros aspectos más básicos respecto a su aplicación. Por ejemplo, las pruebas Raven presentan las ventajas de fácil manejo, puntuación e interpretación de resultados; no obstante, se considera que para ámbitos aplicados requiere un tiempo mucho más largo para ser administrado que otras medidas de rendimiento intelectual (Arthur y Day, 1994). Por ejemplo, pruebas como el Wesman Personnel Classification Test requiere sólo de 28 minutos para su aplicación, al igual que el Wonderlic Personnel Test, para el cual sólo necesitan 12 minutos. Es por ello que han habido varios intentos de construir formas cortas, no sólo del Raven, sino de otras pruebas aún mas complejas como el Wais-R (Crawford, Allan y Jack, 1992).

Un ejemplo de esto lo constituye la validación de una forma corta del WAIS-R para pacientes clínicos con cuadros depresivos (Grossman, Mednitski, Dennis, Scharff y cols., 1993). Tal estudio se llevó a cabo en 31 sujetos con edades comprendidas entre 16 y 57 años, a quienes se les aplicó la versión abreviada, constituida por cuatro subtests, cuya administración requirió una inversión de tiempo de 13 a 24 minutos, lo cual - aparte de la ventaja que comporta reducir considerablemente el tiempo de aplicación y corrección - representa una forma más asequible de estimar el funcionamiento intelectual de pacientes depresivos, quienes clínicamente disponen de poca proactividad. Así mismo, tanto la versión estándar como la avanzada del TAMPR, han sido acortadas en diversas oportunidades, tanto en población psicopatológica (Wytek, Opgenoorth y Presslich, 1984) como normal, obteniéndose un valioso instrumento que permite sondear la inteligencia general, en menor tiempo, y sin perturbar sensiblemente las propiedades psicométricas de las versiones originales.

La última versión abreviada del TAMPR de la cual se tiene conocimiento, es la llevada a cabo por Arthur y Day (1994) en población universitaria del

Suroeste de Estados Unidos. El instrumento fue acortado a un tercio de su longitud original, es decir, de 36 ítems se llegó a 12, reduciendo igualmente el tiempo de administración, que pasó de 40-60 minutos a 15 minutos. Las propiedades psicométricas del instrumento se conservaron intactas; a saber: la confiabilidad, estructura factorial unitaria y validez convergente, es decir, el patrón de correlaciones con otras medidas de logro intelectual, tales como el Wesman Personnel Classification Test y el Wonderlic Personnel Test.

En lo tocante a este trabajo de Arthur y Day (1994), uno de sus mayores aportes es la descripción de cómo elaborar una versión corta del test de Raven. Los criterios tomados en cuenta por ellos son extraídos del análisis de ítems realizado luego de haber aplicado la forma larga a una muestra de sujetos inicial (300 Ss), así como también haber administrado las dos pruebas antes mencionadas, con las cuales obtuvieron el perfil de correlaciones convergentes.

De tal análisis de datos se consideran los siguientes criterios, cada uno correspondiente a un paso en la toma de decisión sobre la escogencia del reactivo en la nueva versión:

1. nivel de dificultad: la primera característica que desea conservar al elaborar resumir el instrumento original es que se conserve el perfil de dificultad creciente de los ítems que éste posee. Para ello, se dividió el continuo de reactivos en grupos consecutivos de tres cada uno, por ejemplo: 1-2-3, 4-5-6... 34,35,36. Luego, dentro de cada segmento sería extraído uno de los reactivos, 12 en total, de acuerdo a los criterios que siguen en este proceso.
2. correlación corregida ítem-test: es decir, la correlación existente entre las respuestas registrada para cada ítem y el puntaje total obtenido por los sujetos en el test. La corrección significa que para el cálculo de cada correlación se elimina el ítem que se está correlacionando, de modo que éste no funcione como factor de inflación en el valor resultante al incluir correlaciones consigo mismo. Con estos datos, se

escoge de cada grupo de tres elementos, aquel cuya magnitud de la relación sea la mayor. El proceso de decisión puede terminar aquí; sin embargo, puede ocurrir que dos o los tres ítems del subconjunto sostengan correlaciones iguales con el test; para esta situación, se cuenta con un último criterio.

3. *alpha* si el ítem es eliminado: esto es, la estimación de la confiabilidad del test total mediante el coeficiente de consistencia interna *alpha* de Cronbach tantas veces como elementos hay, eliminando uno de ellos por vez. En este sentido, los mejores reactivos son aquellos en los cuales el nivel de confiabilidad disminuye cuando son eliminados del cálculo.

Una vez realizado este proceso de toma de decisión, los ítems que resultaron fueron por Arthur y Day (1994) fueron: 1, 4, 8, 11, 15, 18, 21, 23, 25, 30, 31 y 35; cuyo funcionamiento como prueba independiente se probó mediante su aplicación a una nueva muestra (200 Ss), calcular sus propiedades psicométricas y compararlas con las del test original.

El contraste de los parámetros que caracterizan a cada versión del test permitió concluir que el acortamiento del test fue exitoso en cuanto a niveles de confiabilidad (larga 0,84; corta 0,73), estructura factorial similar y unitaria, y patrón de correlación con otras medidas de ejecución intelectual semejantes, encontrándose éstas entre 0,20 y 0,50 en ambas formas del test.

En resumen, de acuerdo a toda la exposición precedente, se pueden ubicar tres temas bastante importantes en la actualidad, en torno a las pruebas de Raven:

1. la discusión acerca de la hipótesis de ausencia de carga cultural o factor verbal en la resolución de las Matrices Progresivas; hasta el momento, lo que se ha encontrado es que el rango de puntajes varía de una población a otra, mientras que las propiedades psicométricas se mantienen prácticamente inalteradas, especialmente la

confiabilidad y la estructura factorial predominantemente unitaria. Ello da cuenta de que en efecto el factor g subyace a las medidas logradas con este test; no obstante, no queda clara la independencia cultural del mismo, y no se ha encontrado una explicación satisfactoria a esta diferencia en la tendencia central a través de distintas muestras de sujetos.

2. la alta validez de criterio que posee, al servir de predictor eficaz del rendimiento global futuro de los individuos, en el sentido de que funciona como un resumen de la ejecución intelectual global del individuo. También se reconoce su debilidad para elaborar diagnósticos diferenciales a partir de sus resultados u obtener perfiles de fortalezas y debilidades cuando se trata de diseñar una intervención sobre la base del aprovechamiento de factores intelectivos de grupo o específicos.
3. la corriente actual de desarrollos relativos a las matrices de Raven, bien por el lado de la creación de formatos computarizados, o bien por la elaboración de versiones cortas, tanto en lápiz y papel como en computadora. Estas últimas, en general, traen como beneficio el acortamiento del tiempo de evaluación con una calidad de información muy similar a la del test original. La importancia de esto deviene en su aplicación a diversas áreas dentro del ejercicio de la psicología; por ejemplo en procesos de reclutamiento y selección de personal, evaluaciones clínicas a pacientes hospitalizados o con impedimentos motores, dificultades atencionales, etc. Así, con el uso de las versiones cortas, incluyendo las de pruebas distintas al *Raven*, se pueden obtener estimaciones bastante aproximadas y válidas, siguiendo procedimientos menos intrusivos para los sujetos a evaluar.

Es a partir de la consideración de esos tres argumentos que surgió el interés por validar una versión abreviada del TAMPR en contexto académico universitario, con la finalidad de examinar la adecuabilidad de la versión

existente y, si se presenta la necesidad, proponer a partir de los datos encontrados lo que podría constituir un instrumento abreviado mejor adaptado a las propiedades de la población estudiada; el cual despliegue parámetros similares a su predecesor largo, con un tiempo de administración substancialmente más corto, y que satisfaga las necesidades de evaluación y calidad de la información de los psicólogos en dicha población, obteniendo también el beneficio de realizar una valoración más justa del nivel intelectual de la misma al proveer a los evaluadores de una medida adaptada al comportamiento de esta variable en el mencionado grupo de referencia.

MÉTODO

PROBLEMA

El propósito de este estudio es la validación de la forma corta del Test Avanzado de Matrices Progresivas de Raven (TAMPR), desarrollada originalmente en Inglaterra por Arthur y Day (1994), en la población compuesta por los estudiantes de la Universidad Católica Andrés Bello. Ello con la finalidad de explorar la adecuación del uso de una forma abreviada de esta prueba en dicho contexto, mediante la comparación de propiedades psicométricas tales como confiabilidad, estructura factorial y patrón de correlación con otras medidas de ejecución, sostenidas por ambas versiones, es decir, la original o larga y la abreviada o corta.

HIPÓTESIS

En esta investigación se espera:

1. Obtener un coeficiente de confiabilidad (α) de aproximadamente 0,84 para la versión original del TAMPR (Arthur y Day, 1994).
2. Una reducción en el índice de confiabilidad de la forma corta respecto del original, mas el mismo debe ser mayor al predicho según la fórmula Spearman-Brown para la estimación de la confiabilidad cuando se altera la longitud del primer test.
3. Que tanto la versión original como la corta sostengan un patrón similar de correlaciones con otras medidas de ejecución intelectual - como el rendimiento académico en cursos académicos cuyo contenido sea predominantemente formal -, en el sentido de que sean positivas, directas, significativas, y que sus magnitudes no difieran sensiblemente.

4. Obtener, tanto en la forma larga como en la abreviada, una estructura factorial similar y preferentemente unitaria.

VARIABLES

La escogencia de las variables utilizadas en este trabajo responde a criterios relativos a: (a) el constructo que se pretende medir a través del test sometido a estudio, (b) la variable de ejecución considerada para comparar la validez convergente entre las dos versiones del test, y (c) criterios de muestreo de sujetos, aspecto éste que será descrito con mayor detalle más adelante. Se exponen las definiciones conceptuales y operacionales, en los casos que aplican, bajo las cuales fueron concebidas las variables utilizadas, y en cuya aceptación fueron tratadas para establecer la aproximación teórica al momento de interpretar los datos. Las variables son:

<i>Capacidad intelectual general</i>	
Definición conceptual	Es la capacidad presente de una persona de desempeñarse cognoscitivamente para adaptarse a su medio ambiente (Sattler, 1988/1996), haciendo uso de su más alta claridad de pensamiento y con independencia de los conocimientos adquiridos (Raven, s.f.).
Definición operacional	Puntaje obtenido por un sujeto, bien en la versión larga o bien en la corta del TAMPR, tras comparar formas y razonar por analogías, educiendo correlaciones entre elementos que componen una matriz progresiva lacunaria, en condiciones de disponer de tiempo ilimitado (Raven, s.f.).
<i>Rendimiento académico</i>	
Definición operacional	Promedio de la evaluación continua previa al examen final, que obtiene cada uno de los sujetos la primera vez que hayan cursado Estadística I en la carrera que actualmente estudian en la UCAB; en el caso de los estudiantes de Derecho, se consideran las calificaciones de Lógica Jurídica.

<i>Carrera</i>	
Definición operacional	Escuela de la UCAB en la cual el sujeto reporta estar cursando estudios.

TIPO DE INVESTIGACIÓN

Se trata de una investigación psicométrica, en la cual el objeto de estudio es un instrumento de medida (el TAMPR) y no el constructo que el mismo intenta indicar (Factor g de Inteligencia). El trabajo comprende dos conjuntos de unidades de observación, obtenidos ambos de la misma población:

1. Muestra A: Aplicación del TAMPR original, cuyas respuestas y puntajes totales serán los datos que permitirán realizar el análisis psicométrico para obtener los parámetros de la prueba: confiabilidad, validez, análisis de ítems y correlación con la medida de ejecución académica.
2. Muestra B: Aplicación de la forma abreviada a una nueva muestra, para realizar el mismo análisis psicométrico, y así comparar sus parámetros con los de la prueba larga, para probar su poder o calidad de información aportada al evaluar inteligencia general.

La aplicación de ambos tests se llevó a cabo simultáneamente, a sujetos presentes dentro del mismo grupo, siguiéndose una proporción de distribución de 3:1, es decir, por cada tres pruebas largas se administraba una corta.

PROCEDIMIENTO

Sujetos y recolección de datos

La población a partir de la cual se muestrearon los sujetos de investigación, está constituida por aproximadamente 10.000 estudiantes de pre-grado de la UCAB. El procedimiento de selección empleado fue uno no

probabilístico e incidental (Kerlinger, 1988/1990). La aproximación a la muestra se hizo a través de profesores de las distintas Escuelas de la universidad, quienes facilitaban los grupos de sujetos cediendo de una a dos horas de sus clases; así como también, otros alumnos que en forma eventual y voluntaria se ofrecieron para ser evaluados. Dado que la resolución de las Matrices consume un tiempo que va de unos 15 minutos - para la versión corta - a unos 40 minutos o más - para la versión larga -, no fueron utilizadas otras medidas psicométricamente formales; pues éstas implicarían tanto mayor fatiga de los sujetos como un mayor consumo de tiempo, no sólo académico sino personal.

En tal sentido, la representatividad de la muestra tal vez se no se vea respaldada por el tipo de muestreo, pues no están presentes dos procedimientos deseables como la aleatorización y la estratificación. No obstante, una representatividad aproximada se podría apreciar de acuerdo a dos elementos: (a) el tamaño muestral, que en este caso se ajusta al utilizado en la mayoría de los estudios que sobre las Matrices Progresivas fueron revisados (ver Bases Teóricas), y (b) la adecuación estadística de las distribuciones de puntajes del rendimiento académico, como propiedad de la población, aspecto éste que se revisará en la exposición de Resultados.

Como se mencionara anteriormente, la recolección de datos constó de dos muestras según la versión del test que se aplicara. La cantidad total de unidades de observación alcanzó los 399 sujetos, correspondiendo 277 de ellos a quienes les fuera administrada la versión original, y los restantes 122 a los que se les aplicó la forma abreviada. Los sujetos examinados provinieron de las Escuelas de Psicología, Administración y Contaduría, Relaciones Industriales, Ciencias Pedagógicas, Derecho, Economía y Comunicación Social (ver Tabla 2).

Tabla 2: Tamaño muestral para ambas versiones por carrera.

	<i>Versión Original</i>	<i>Versión Abreviada</i>
Carrera	n	n
Psicología	109	40
Relaciones Industriales	86	2
Administración y Contaduría	2	0
Ciencias Pedagógicas	16	16
Derecho	14	5
Economía	50	57
Comunicación Social	0	2

La recolección de datos comprendía la obtención de tres informaciones: (a) la carrera cursada por el sujeto, (b) su calificación previa en el curso seleccionado - como se describe en Variables -, y (c) las respuestas a cada ítem del test y el puntaje bruto obtenido en la versión del TAMPR que haya recibido la persona para su resolución.

Una vez que se disponía de los grupos de sujetos a ser evaluados, se procuraba hacer la distribución de las versiones del TAMPR en una proporción de una forma corta por cada tres largas. Los cuadernillos de la forma consistían en copias del formato original pertenecientes a la Escuela de Psicología de la UCAB, mientras que los de la versión abreviada se editaron especialmente para este estudio, obteniendo copias de los doce ítems necesarios para conformarla (Arthur y Day, 1994), los cuales fueron encuadernados en el orden original y les fue cambiado su número de identificación dentro del test largo por el correspondiente a su posición dentro de la forma corta. Así mismo, fueron diseñados los protocolos de respuesta para cada versión (Anexos C y D).

Una vez que se pedía a los sujetos escribir sus datos: Carrera y Calificación previa; se procedía dar las instrucciones del test, de las cuales hay un resumen incluido en el protocolo de respuesta. Las instrucciones utilizadas para este estudio constituyen una integración de las que se exponen para administrar colectivamente tanto el Set I como el Set II del TAMPR (Raven y

cols., 1994). Ello motivado por dos razones; la primera es que el Set I no sería administrado y, por lo tanto, los sujetos no disponían de una práctica previa a la resolución del Set II, así que se hacía necesario explicar claramente en qué consistía la tarea.

Esta licencia en la exposición de las instrucciones es avalada por el propio manual de administración de las Matrices, dado que "no es necesario apegarse rígidamente a ningún discurso particular" (Raven y cols., 1994, pág. APM50), siempre y cuando se aseguren ciertos requerimientos, como (a) que los sujetos comprendan lo que van a hacer y el método de pensamiento necesario para resolver los problemas; (b) que el test sea administrado a todos en la misma forma y (c) que el procedimiento adoptado corresponda al tipo de datos de datos que se desean obtener.

Antes de toda instrucción, se exponía brevemente el uso que tendrían los datos a recolectar, es decir, el objetivo de aquella evaluación y, por lo tanto, el objetivo del estudio del cual formaban parte aquellos sujetos. Así, en aplicación colectiva, se decía al grupo que recibirían una cuadernillo con los problemas y la correspondiente hoja de respuesta, unos de la versión original de 36 ítems y otros la corta de 12 ítems. Para cumplir con los puntos (a) y (b), el discurso contenía los siguientes puntos:

1. lo que el test intenta medir: capacidad intelectual general, o habilidad para resolver problemas nuevos independientemente del conocimiento adquirido y libre de factores lingüísticos.
2. el contenido del test: matrices que cumplen una progresión, cuya regla debe ser educida por el sujeto para encontrar aquel paso -o figura- de la progresión que se adecuara al espacio vacío al final de la misma, respuesta tal que debía corresponderse a una de las alternativas presentadas en la parte inferior de la hoja problema, debajo de la matriz. Por otra parte, también se decía que el método de trabajo es el mismo en todos los ítems y que podían explorar los

pasos de las progresiones en los sentidos izquierda-derecha y arriba-abajo, y que se podía re-explorar la matriz para verificar la decisión tomada.

3. la respuesta al problema: explicación sobre cómo utilizar la hoja de respuesta, de manera tal que los sujetos se aseguraran de escribir el número de su escogencia en el renglón correspondiente al ítem que estaban resolviendo. Además, señalar que la respuesta correcta es sólo una en cada problema.

Las claves del discurso para llenar el requisito (c) eran:

1. resolución continua de las matrices problema, sin saltar alguna ni volver atrás;
2. tiempo ilimitado para completar el test;

ello era necesario para poder construir una base de datos adecuada, los más completa posible, para el análisis de datos, que debía contar con la mayor cantidad de datos válidos posible, es decir, sin vacíos de respuesta o ausencia de algún dato correspondiente a algún sujeto; de modo que se pudiera establecer una estimación de las propiedades psicométricas del test sobre la base una resolución total.

Análisis de datos

Los parámetros psicométricos de ambas formas del test fueron obtenidos mediante dos pasos: (a) el vaciado de los datos brutos en la hoja de cálculo Microsoft Excel 7.0 ®, para luego (b) ser procesados a través del paquete estadístico SPSS ®.

Las respuestas de los sujetos a los tests se realizó en dos hojas de cálculo, cada una correspondiente a una versión del TAMPR. En ellas se colocaban los siguientes datos: (a) estatus de la respuesta para cada ítem -

correcta: 1; incorrecta: 0 -, cuya sumatoria resultaba en (b) el puntaje total obtenido por el sujeto en el test. Así mismo, se registró la alternativa escogida por el evaluado, de acuerdo a números del 1 al 8, siendo éstos los mismos que fueron escritos en la hoja de respuesta.

Posteriormente, se inscribían los datos personales: (d) carrera, indicada por un número asignado arbitrariamente a cada Escuela; y (e) el puntaje relativo a la variable rendimiento académico. Un último dato, añadido a posteriori, fue (f) el puntaje estándar Z en la distribución estadística de su carrera de la variable rendimiento académico. El siguiente es un ejemplo del vaciado de datos, de acuerdo a lo descrito anteriormente:

ítem 1	alter 1	...	ítem 36	alter 36	total	carrera	previa	Z
1	7	...	0	2	25	1	10	+0,85

donde, hipotéticamente, el sujeto respondió correctamente el ítem 1 con la opción 7... respondió incorrectamente el ítem 36 con la opción 2, obtuvo un puntaje total de 25 puntos, estudia Psicología en la UCAB, y calificó con 10 puntos la primera vez que cursó Matemáticas y Estadística I, que lo ubica 0,85 desviaciones típicas de la media de calificaciones previas al examen final de ese curso en su carrera.

Una vez completada esta base de datos, las hojas de cálculo fueron transferidas a SPSS®, y se prosiguió con la estimación de:

1. Confiabilidad: coeficiente de consistencia interna *alpha de Cronbach (a)* para cada versión del TAMPR.

2. Validez de constructo:

Análisis factorial. De acuerdo al instrumento de cálculo estadístico utilizado, el análisis realizado corresponde al de Componentes Principales, con autovalores iguales o mayores a

1 (≥ 1) y rotación varimax. Se reportan como reactivos componentes de factor aquellos cuya carga factorial es igual o mayor 0,35. La comparación entre las formas del test se establece a partir de la inspección ocular de sus estructuras factoriales en tanto en cuanto al número de componentes principales obtenidos, proporción de elementos que participan en los mismos, autovalor de los factores y proporción en que éstos explican la varianza de los ítems. El objetivo es verificar la existencia tanto de una estructura factorial unitaria como similar entre las dos muestras de reactivos.

Validez convergente: correlación Producto - Momento de Pearson entre los puntajes brutos obtenidos por los sujetos en el TAMPR - sea su forma original o abreviada - y el puntaje estandarizado Z de su rendimiento académico (tal como se define en el apartado Variables de este capítulo). Esto constituye, más allá de la estimación del poder predictivo de las Matrices, un parámetro que permite comparar la magnitud y sentido de las relaciones entre las medidas aportadas por el test y otras medidas de ejecución que comportan en algún grado factores intelectivos.

3. Análisis de ítems: Los parámetros psicométricos a considerar abarcan los señalados por Arthir y Day (1994), a saber: (a) media del ítem o proporción de sujetos que lo contestaron correctamente, cuyo valor complementario constituye el nivel de dificultad; (b) correlación entre el estatus de las respuestas dadas al ítem (correcto - incorrecto) y el puntaje total obtenido en el test, lo cual también se denomina correlación ítem-test; (c) *alpha* si el ítem es eliminado, es decir, el coeficiente de confiabilidad por consistencia interna que tendría el test si algún ítem es eliminado de la muestra total. Además, se presenta un Análisis de Varianza simple (ANOVA), ítem por ítem,

entre los puntajes totales obtenidos por los sujetos en el test, agrupados éstos como muestras independientes según la opción de respuesta escogida durante la resolución del test, lo cual constituye el análisis de distractores de los ítems versus su respuesta correcta.

RESULTADOS

Los datos a continuación presentados parten de una muestra de estudiantes de pre - grado de la UCAB inscritos en las escuelas de Psicología, Relaciones Industriales, Administración y Contaduría, Ciencias Pedagógicas, Derecho, Economía y Comunicación Social, cuyo tamaño - sobre la base de un aproximado de 10 sujetos por ítem que contenga la versión del test en cuestión (Nunnally, 1973) - alcanzó un número de 399 individuos, de los cuales 277 fueron examinados con la versión original o larga del TAMPR, y los restantes 122 lo fueron mediante la forma abreviada del mismo (Ver tabla 2).

Tabla 2: Tamaño muestral para ambas versiones por carrera.

	<i>Versión Original</i>	<i>Versión Abreviada</i>
Carrera	n	n
Psicología	109	40
Relaciones Industriales	86	2
Administración y Contaduría	2	0
Ciencias Pedagógicas	16	16
Derecho	14	5
Economía	50	57
Comunicación Social	0	2

El recorrido teórico de puntajes del TAMPR original puede ir de 0 a 36 puntos. Luego de la administración y puntuación de los protocolos correspondientes a esta prueba, la distribución encontrada ocupa un rango entre 1 y 35 puntos, donde se encontraron los siguientes indicadores de tendencia central (Anexo E):

1. Las puntuaciones totales obtenidas con mayor frecuencia fueron 25 y 26, ubicándose en cada una de ellas 25 de los 277 sujetos pertenecientes a esta parte de la muestra. Al realizar un examen visual de la curva (Gráfico 1), se puede observar cierto carácter multimodal en la misma, dado que existen otros puntajes que si, bien

no representan el modo de la distribución, sobrepasan en frecuencia otras puntuaciones. No obstante, dichos puntos del recorrido no se alejan de la tendencia central.

2. El punto de la distribución que divide las observaciones en dos porciones equivalentes en frecuencia (50%) lo representaría un puntaje de 24,5, valor muy cercano al de la media aritmética de 24,3; lo cual, para efectos prácticos, constituiría el mismo puntaje.



Gráfico 1: Distribución de frecuencia de los puntajes del TAMPR original.

La dispersión de los puntajes alrededor de estos valores centrales posee una desviación típica de 5,48 puntos, por lo que cerca del 70% de la población estudiada obtiene puntajes entre 18 y 29 puntos. A partir de aquí, es posible hacer ciertos señalamientos en torno a la normalidad de la distribución de puntajes del TAMPR. Nótese la relación existente entre las medidas de tendencia central, donde la media es menor que la mediana, y ésta a su vez es menor que la moda, que es propio de una distribución coleada hacia adentro.

Así mismo, se observa que los puntos entre los cuales se encuentra la mayor porción de puntajes, da como resultado una larga cola proyectada hacia los valores más pequeños del recorrido de puntajes posibles. Estadísticamente, estas propiedades de la distribución quedan reflejadas en sus indicadores de forma, a saber: (a) asimetría $-0,88$, y (b) Kurtosis $2,09$, que habla de una gran concentración de puntajes alrededor de la media, lo cual le aporta un carácter leptocúrtico (Anexo E).

La estimación del nivel de confiabilidad de la versión original del TAMPR - calculada a través del coeficiente de consistencia interna Alpha de Cronbach - se ubicó en $0,838$, valor tal que responde al alto grado de homogeneidad existente entre los reactivos que componen la prueba (Anexo F). Las correlaciones entre las respuestas a los ítems y el puntaje total obtenido en la prueba, se ubican entre $0,1941$ (ítem 32) y $0,45$ (ítems 9, 16 y 23), mientras que la confiabilidad Alpha del test si algún ítem es eliminado se mantiene en todos los casos cercana a su valor original ($0,838$), encontrándose las diferencias a partir del tercer decimal (ver Tabla 3). Las correlaciones ítem - test señaladas corresponden a elementos cuyo acierto o desacierto permitirían predecir con un considerable grado de probabilidad la ejecución de los sujetos en el test total, es decir, su ubicación en el continuo latente del constructo inteligencia (entendido éste como factor g).

La distribución del nivel de dificultad de los ítems del test original para esta población se puede caracterizar en tres grupos consecutivos de dificultad creciente, pero con características distintivas entre sí (Anexo G). Se puede apreciar un primer segmento que va del ítem 1 al 12, considerados como fáciles para la muestra de sujetos examinada, que ostenta una distribución más bien rectangular, y despliega un recorrido de $0,06$ (ítem 3) a $0,16$ (ítem 11), valores estos que representan la proporción de sujetos que los contestan incorrectamente.

Tabla 3: Propiedades psicométricas de los ítems de la versión original.

<i>Ítem</i>	<i>Nivel de Facilidad</i>	<i>Correlación Ítem-test</i>	<i>Alpha (si el ítem es eliminado)</i>				
1	0,900	0,2782	0,8334	18	0,670	0,358	0,831
2	0,912	0,398	0,831	19	0,797	0,330	0,831
3	0,940	0,370	0,832	20	0,666	0,331	0,831
4	0,892	0,397	0,830	21	0,607	0,375	0,830
5	0,849	0,381	0,830	22	0,559	0,403	0,829
6	0,884	0,411	0,830	23	0,607	0,455	0,827
7	0,912	0,232	0,836	24	0,531	0,348	0,831
8	0,908	0,382	0,831	25	0,547	0,371	0,830
9	0,837	0,4561	0,828	26	0,476	0,217	0,836
10	0,904	0,382	0,831	27	0,357	0,246	0,834
11	0,932	0,393	0,831	28	0,381	0,321	0,832
12	0,892	0,371	0,831	29	0,265	0,260	0,834
13	0,6310	0,291	0,833	30	0,472	0,247	0,834
14	0,853	0,380	0,830	31	0,333	0,316	0,832
15	0,813	0,324	0,832	32	0,361	0,194	0,836
16	0,813	0,453	0,828	33	0,472	0,286	0,833
17	0,7857	0,320	0,832	34	0,365	0,346	0,831
				35	0,309	0,345	0,831
				36	0,091	0,016	0,838

El siguiente segmento de elementos, del ítem 13 al 24, conllevan un aumento progresivo y básicamente monotonico en su nivel de dificultad. Sin embargo, se pueden observar dos elementos que comportan cambios bruscos. El ítem 13 – primero de esta serie – tiene un nivel de dificultad de 0,369, y muestra gráficamente un aumento en su proporción de desaciertos fácilmente evidente, en contraste con los ítems predecesores y sucesores (ítem 12: 0,1071, ítem 14: 0,1468).

Otro cambio brusco observado en esta serie lo constituye el ítem 19 (dificultad: 0,2024), el cual recupera el índice de aciertos obtenidos por los sujetos de la muestra; es decir, disminuye su nivel de dificultad en comparación con su predecesor (ítem 18, dificultad: 0,3294) y su sucesor (ítem 20, dificultad: 0,3333). Salvo estos dos elementos discrepantes dentro de la serie, el resto de los reactivos comportan un perfil de aumento constante de la dificultad del test, que se ubica en niveles que van de fácil (ítem 14, dificultad: 0,1468) a promedio (ítem 24, dificultad: 0,4682).

El último conjunto de ítems es el que comprende los reactivos más difíciles de la prueba. Sin embargo, su comportamiento es irregular. Hay un subconjunto de ítems (24, 25, 26, 30 y 33) que muestra un nivel de dificultad alrededor del promedio, con una proporción de aciertos entre 0,4524 (ítem 25) y 0,5278 (ítem 30). Así mismo, hay otro subconjunto de elementos (27, 28, 31, 32, 34 y 35) con dificultad superior al promedio entre 0,619 (ítem 28) y 0,6905 (ítem 35). Entre las irregularidades descritas en párrafos anteriores, caracterizadas por picos y valles, y que pueden ser observadas en el Gráfico 2, destacan por su dificultad evidentemente superior los ítems 29 (0,7341) y 36 (0,9082).

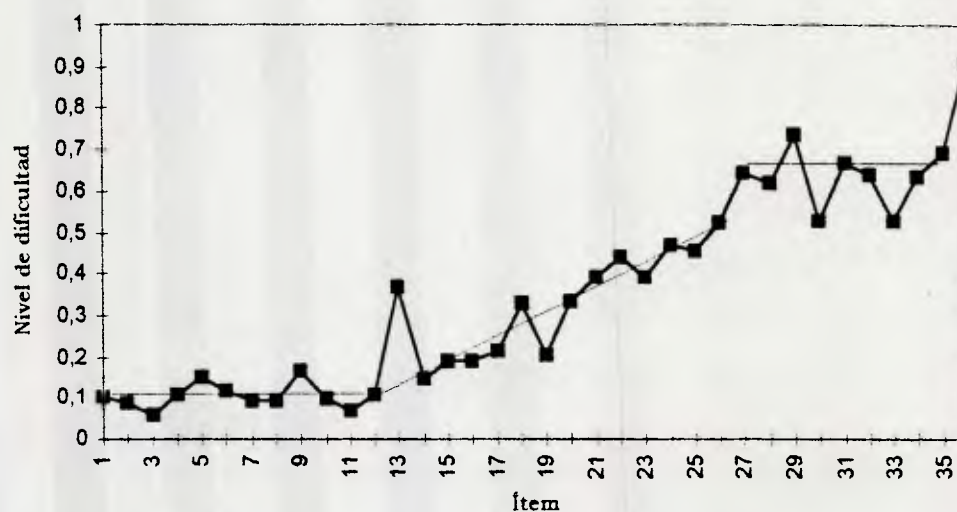


Gráfico 2: Nivel de dificultad de los ítems en el TAMPR original. Obsérvense las líneas de regresión a la mediana que muestra el aumento progresivo del nivel de dificultad por grupo de reactivos.

De esta manera, el ordenamiento de los ítems en el test original responde más a un incremento de dificultad por subconjuntos de reactivos que a una obligada progresión continua de la misma de un ítem a otro. Así, el crecimiento estable y monotónico que se observa en la dificultad de los ítems de la porción central del test señala el conjunto de elementos que ejercen el mayor poder discriminativo dentro de la población evaluada, puesto que la

distribución plana de la primera y la última sección de reactivos conforman - respectivamente - el piso y el techo de ejecución de dicha población.

La evaluación de la validez de constructo del TAMPR en su versión original se sustenta en dos procedimientos: (a) el análisis factorial o, más específicamente, el análisis de componentes principales; y (b) las correlaciones de los puntajes totales obtenidos en el test y otra medida de ejecución correspondiente a los puntajes estandarizados de las calificaciones obtenidas por los mismos sujetos en cursos académicos (Anexo H).

El análisis de componentes principales del TAMPR en su versión original arrojó un componente de autovalor igual a 6,037 y cuyo porcentaje de varianza explicada 16,8%. Estos valores resultan mayores que los obtenidos por otros componentes - que podrían considerarse residuales -, cuyos autovalores van desde 1,036 hasta 1,919 y el porcentaje de varianza que explican recorre un rango de 2,9 a 5,3. Además, al observar la Tabla 4, se puede apreciar cómo la mayoría de los ítems se encuentran exclusivamente cargados del componente 1. Otros, aún cuando cargan fundamentalmente de este factor también están compuestos de otros más pequeños que, en comparación con el primero, podrían considerarse residuales.

Los ítems muestran una proporción de varianza común - comunalidad - que se ubica en un rango de 0,36 a 0,72, lo cual implica que existe una serie de factores comunes a los ítems que componen el test; al considerar que casi la totalidad de ellos poseen una carga significativa de un mismo factor, se presume que estos coeficientes de comunalidad son relativos a una estructura factorial fundamentalmente unitaria.

Cabe señalar que la mayoría de los elementos tienen una carga en el factor 1 entre 0,38 y 0,56, habiendo solamente ocho de ellos cuya carga es menor a 0,35 que, según la regla de extracción de cargas factoriales realizada durante el cálculo estadístico, son aquellos cuyos valores no se muestran en la tabla correspondiente a la matriz factorial. En cuanto a los once factores

restantes encontrados, se observa que a los componentes dos y tres se les puede atribuir alguna consideración dado que sus cargas superan un coeficiente de 0,35 en cinco ítems cada uno, mientras que el resto de los componentes sólo llegan a tener una participación aislada en uno o dos ítems.

Tabla 4: Matriz factorial del TAMPR original.

Incluye sólo los componentes más grandes, y muestra sólo las cargas de factor iguales o superiores a 0,35.

Ítem	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
9	0,558				
16	0,520				
2	0,516				
23	0,508				
6	0,508				
11	0,496				
4	0,493				
3	0,486	-0,367			
10	0,477				
8	0,474				
14	0,466				
12	0,464				
5	0,463				
21	0,448				
22	0,445				
18	0,444				
25	0,421			0,399	
19	0,400		0,371		
20	0,394				
15	0,390		0,359		
17	0,386				
1	0,385	-0,361			
24	0,382				
35	0,377				
28					
36		0,515			
31		0,416			
30		0,360			
29					
32			0,482		0,359
34			0,376		
26				0,520	
7					0,374
13	0,355				
27			0,370		
33					

En cuanto a la versión abreviada administrada, el recorrido posible de puntajes brutos va de 0 a 12, encontrándose en la muestra un rango entre 1 y 12, cuya tendencia central mostró un comportamiento semejante al observado en la prueba larga; es decir, la media aritmética (7,93) es ligeramente menor que la mediana (8,00), que a su vez es visiblemente menor que la moda (10,00) (Gráfico 3). La dispersión alrededor de la tendencia central refleja que el 70% medio de la muestra alcanza puntajes entre 5 y 11, donde la desviación estándar es de 2,70 puntos (Anexo I).

La diferencia entre las distribuciones de puntajes de la prueba larga y la corta se observa en su forma. A pesar de ser ambas coleadas hacia adentro, dicho fenómeno es más notable para la forma original que en la abreviada, donde el valor de la asimetría es de $-0,38$. Así mismo, se ve en esta última distribución un aplanamiento de la curva que la hace más cercana a una normal unimodal, semisimétrica y campanoide, con una kurtosis de $-0,62$ (Anexo I).

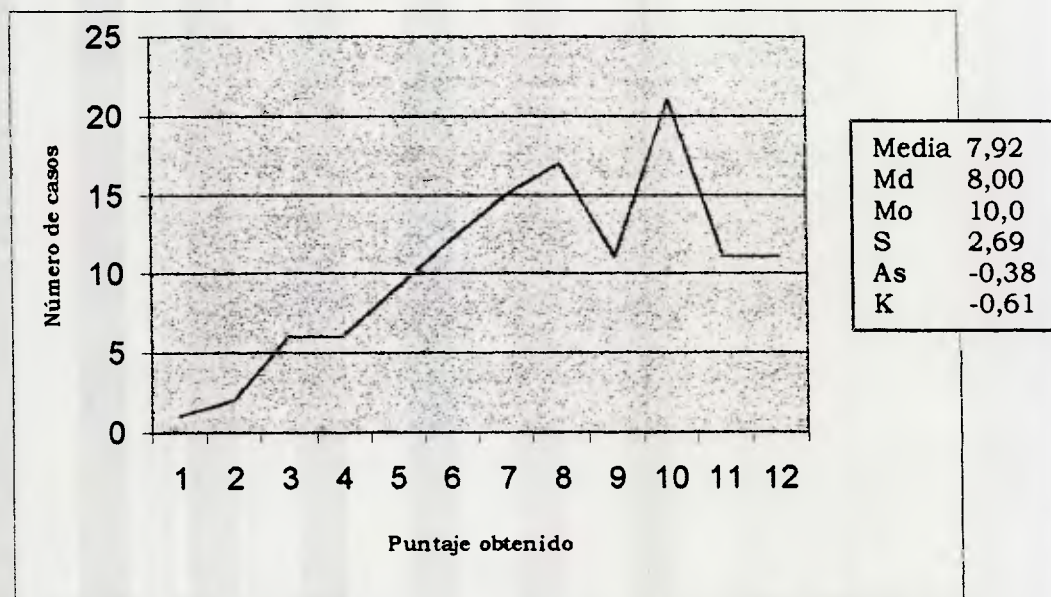


Gráfico 3: Distribución de frecuencia de los puntajes del TAMPR abreviado.

El coeficiente de consistencia interna obtenido – equivalente a su nivel de confiabilidad – es diez puntos menor que el de la versión original, lo cual corresponde a un valor Alpha de 0,73 (Anexo J). Al mismo tiempo, este valor supera en nueve puntos al valor esperado según la fórmula Spearman – Brown, a través de la cual se estimó que el nivel de confiabilidad que alcanzaría el TAMPR si se acorta su longitud a un tercio de la original, se obtendría una coeficiente de confiabilidad de 0,64.

El nivel de dificultad de los doce ítems funcionando como una prueba independiente corre progresivamente de 0,11 hasta 0,62, lo cual implica que son resueltos correctamente entre un 62% de los sujetos en el primer ítem (que corresponde al ítem 1 de la prueba original) y un 11% de ellos en el último elemento (ítem 35 en el test largo) (Anexo K). Así mismo, el crecimiento progresivo del nivel de dificultad, como se puede advertir en el Gráfico 4, muestra un ajuste muy notorio a la línea de regresión mediana, lo cual respalda una función creciente, monotónica y estable de la dificultad.

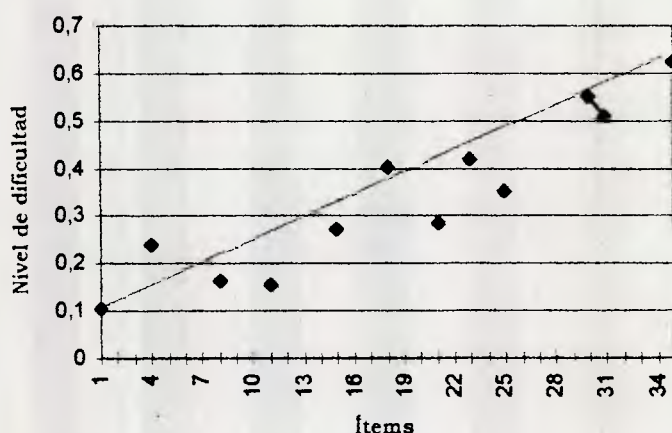


Gráfico 4: Nivel de dificultad de los ítems de la versión abreviada del TAMPR.

Además, obsérvese que la dificultad de los primeros ítems en la versión corta es ligeramente más alta que la de sus correspondientes en la versión larga; mientras que lo contrario ocurre hacia el final de la versión abreviada, donde la dificultad de los elementos es ligeramente inferior a sus correlativos en la versión larga (ver Gráfico 5).

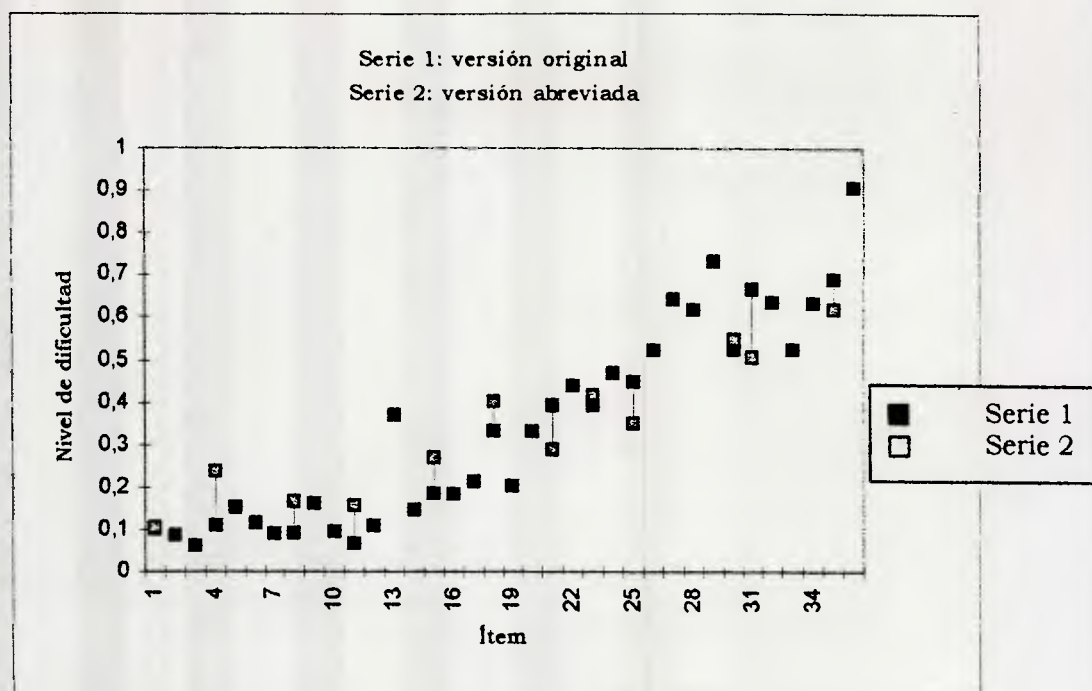


Gráfico 5: Discrepancia del nivel de dificultad de los ítems de la forma corta en contraste con sus correspondientes en la forma larga.

Este fenómeno, puede explicarse mediante la regresión a la línea mediana que sufre el nivel de dificultad de los ítems al evaluar su comportamiento dentro de la muestra en contraste a su funcionamiento como una prueba independiente, también puede encontrar su explicación en otros aspectos distintivos de la resolución de las matrices avanzadas que se presentan cuando los sujetos se enfrentan a una forma u otra del test, implicaciones éstas que posee un carácter más psicológico que estadístico y que se discuten posteriormente en este trabajo.

En la versión corta también se observa un comportamiento psicométrico bastante homogéneo en los ítems, lo cual se evidencia en correlaciones ítem - test total cuya dispersión es semejante a la encontrada en la prueba larga, es decir, entre 0,20 y 0,46. Al ser eliminado un ítem del cálculo de la confiabilidad, su valor disminuye muy ligeramente, aunque se torna un poco más amplio el rango entre los valores mínimo y máximo (Anexo J). Como se puede ver, el índice de confiabilidad del test es 0,73 y, al ser eliminado uno de

los ítems, los valores de este coeficiente oscilan entre 0,6983 y 0,7287 (Tabla 5).

Tabla 5: Propiedades psicométricas de los ítems de la versión abreviada.

<i>Ítem</i>	<i>Correlación ítem – test (corregida)</i>	<i>Alpha (si el ítem es eliminado)</i>
1	0,1979	0,7287
2	0,4090	0,7060
3	0,3753	0,7112
4	0,3431	0,7148
5	0,4223	0,7039
6	0,3138	0,7182
7	0,4326	0,7018
8	0,4561	0,6983
9	0,2726	0,7243
10	0,4249	0,7029
11	0,3515	0,7138
12	0,3388	0,7153
Número de casos = 122 Número de Ítems = 12 Alpha = 0,7294		

El análisis de los componentes principales (Anexo L) arrojó un total de tres factores, entre los cuales hay uno predominante, presente a través de todos los ítems, con cargas ubicadas entre 0,44 y 0,59, cuyo autovalor es 3,08, y el porcentaje de varianza que explicada es de 25,6%. Nótese que este dato es mayor al mostrado a partir del análisis de componentes principales de la versión larga, donde el porcentaje de varianza explicada alcanza un 16,8%. De esto cabe señalar que la forma abreviada del test pareciera quedar limpia de componentes residuales al no ser considerados en su muestra dos tercios del número de ítems existentes originalmente, lo cual hace que esta versión sea más acentuadamente unifactorial que su predecesor largo (Tabla 6).

Como consecuencia de ello, también se obtuvo un rango de cargas factoriales en los ítems de la versión corta más homogéneo que el de la original. Así, mientras los elementos de esta última poseen cargas de factor que pueden ser desde menores a 0,35 hasta 0,56; en la primera, dichas cargas se encuentran entre 0,38 y 0,59.

Tabla 6: Matriz Factorial para los ítems de la versión abreviada.

<i>Ítem</i>	<i>Factor 1</i>	<i>Factor 2</i>	<i>Factor 3</i>
8	0,59089		
7	0,58298		
5	0,57271		
2	0,57058	0,43083	
10	0,56887		
3	0,54521	0,46235	
4	0,51102	0,48084	- 0,39485
11	0,47689	- 0,46457	
6	0,44804		
12	0,43817	- 0,52460	0,36321
1		0,45031	0,43073
19	0,37766		0,61341

La predominante estructura unifactorial de la versión corta se apoya, además, en unos coeficientes de comunalidad desplegados entre 0,23 y 0,61, los cuales reflejan la existencia de una porción aceptable de varianza compartida y relativa en su mayor porción al primer componente factorial, a lo cual subyace la medición de un mismo constructo (Tabla 7).

Tabla 7: Coeficientes de comunalidad de los ítems de la versión corta del TAMPR.

<i>Ítem</i>	1	2	3	4	5	6	7	8	9	10	11	12
Comunalidad	0,48	0,61	0,58	0,65	0,36	0,23	0,49	0,47	0,52	0,42	0,46	0,60

Por su parte, los restantes componentes dos y tres, con autovalores de 1,55 y 1,23, y varianza explicada de 12,9 y 10,2 respectivamente, logran agrupar seis reactivos para el segundo factor y cuatro para el tercero. El comportamiento de estos componentes es semejante el mostrado por sus homólogos en tamaño mostrados en la matriz factorial del test original; esto, en el sentido de que las cargas de factor de los elementos que los componen, aunque significativas en magnitud, son inversas en signo si se comparan los

valores observados en los primeros ítems con los de los últimos reactivos (Tabla 6).

La significación de ello puede estar en el patrón de respuesta ejecutado por los sujetos, entendiéndose a través de estos valores que quienes aciertan los ítems ubicados hacia el inicio de la prueba, desaciertan los que están hacia el final, lo que constituiría un componente factorial relativo al nivel de dificultad de los ítems en cuestión y no a la medición de un constructo distinto del factor g (Franco y Prengler, 1980).

En este punto, se hace valedera una revisión de las propiedades psicométricas de los ítems de la forma corta dentro de la versión original (Anexo M). Así, se encontró que al aislar las respuestas de los sujetos a estos reactivos dentro de su ejecución en la versión larga del TAMPR, el nivel de confiabilidad estimado alcanzó un coeficiente Alpha de 0,65, muy similar al obtenido mediante la corrección Spearman - Brown, según la cual acortar el TAMPR a un tercio de su longitud original conllevaría obtener una confiabilidad de 0,64. Al mismo tiempo, estas magnitudes del coeficiente son - respectivamente - ocho y nueve puntos de correlación menores que el obtenido a partir de los datos de los doce ítems trabajando como una prueba independiente, cuyo coeficiente de consistencia interna Alpha es de 0,73. Por su parte, las correlaciones ítem - escala corren entre 0,18 y 0,37, recorrido de valores éste que es muy similar al mostrado por los datos del formato abreviado del TAMPR.

El coeficiente Alpha si se elimina algún elemento de la muestra se ubica muy cercano al valor inicial, que para este conjunto de matrices problema es de 0,65, tal como se ha observado en las otras estimaciones llevadas a cabo sobre cada versión de la prueba. En este caso, donde se evalúan los doce reactivos del formato abreviado dentro del original, los coeficientes corregidos de Alpha ocupan un rango entre 0,61 y 0,66. Lo que se advierte es que mientras menor es la magnitud del coeficiente de confiabilidad, los valores corregidos que resultan cuando se elimina uno de los elementos amplían su

rango. Por ejemplo, dentro de la versión larga, donde Alpha es igual a 0,83, las diferencias en los coeficientes sólo se dan a partir del tercer decimal; mientras que siendo el valor de Alpha de 0,73 en la versión corta, sus coeficientes corregidos resultan entre 0,69 y 0,73.

La importancia de estos resultados deviene de la alteración de la confiabilidad que comporta la eliminación de un determinado ítem de la muestra. De este modo, se encuentran implicaciones tanto para la conservación de un elemento dentro de la muestra original como para su factible escogencia en la elaboración de una forma abreviada. Si la extracción de uno de ellos resulta en un aumento significativo del nivel de confiabilidad del test total, entonces constituye un reactivo que - para beneficio de éste - puede ser eliminado, así como también quedaría en duda su elegibilidad para conformar una versión corta.

Mientras tanto, si al eliminar un ítem en el cálculo de la confiabilidad, este coeficiente disminuye significativamente su valor, esto no sólo resultaría en su conservación dentro del test original, sino que aumenta la posibilidad de pertenecer a la selección de reactivos de la forma corta.

De los resultados anteriormente señalados se extrae un comportamiento fundamentalmente rectangular, en el que ninguno de los valores corregidos del coeficiente de confiabilidad se aleja apreciablemente de aquel que fuera estimado para la muestra de reactivos en cuestión, es decir, los 36 ítems del TAMPR original, los doce ítems dentro del mismo, y estos doce elementos como prueba independiente.

Los niveles de dificultad de estos reactivos, es decir, los correspondientes a la forma corta dentro de la original, tienen un recorrido que va de 0,11 a 0,68, cuya progresión creciente es estable (Anexo N).

La evaluación de las alternativas de respuesta de los ítems se llevó a cabo a través de un análisis de varianza entre éstas, tomando en cuenta el puntaje total obtenido por los sujetos que escogieron cada una de ellas al

responder a determinado reactivo (Anexo O). De este modo, se espera en primer lugar que haya un mayor número de sujetos que escojan la alternativa correcta mientras que la frecuencia con que son elegidos los distractores muestre una distribución rectangular y de menor frecuencia. Además, es deseable obtener una diferencia significativa entre la media de puntajes de los sujetos que eligen la alternativa correcta – siendo a su vez dicha media superior – y la media de los sujetos que optan por los distractores. Por último, también se espera que no exista diferencia significativa entre la media de los puntajes de los distractores.

Es importante conocer la distribución de las escogencias realizadas a través de las ocho figuras dadas como posibles soluciones en los intentos por resolver cada ítem. Tal como se comenta en el manual del TAMPR (Raven y cols., 1994), las matrices fueron diseñadas para constituir esencialmente un test de ejecución máxima que debe ser afrontado con claridad y precisión de pensamiento, si es necesario bajo presión de tiempo.

El objetivo principal de comparar la distribución de las respuestas a los ítems, es ubicar a partir de qué punto la prueba se torna “psicológicamente problemática” (Raven y cols., 1994), es decir, cuáles son los ítems en los cuales los sujetos comienzan a fallar y desviar su atención hacia las alternativas distractoras en lugar de la opción correcta que completa la progresión de la matriz. Mediante este análisis se pueden detectar figuras que engañan a los sujetos, y asegurar que antes de que se alcance tal punto, la gente no escoja las respuestas correctas a pesar de un razonamiento equivocado.

Del análisis de varianza se obtuvo que las diferencias de medias entre la alternativa correcta y las incorrectas a través de todos los ítems fueron significativas, obteniéndose una probabilidad de ocurrencia al azar menor a 0,01 en todos los casos. Aún así, se encontraron tres conjuntos de matrices problemas con características distintivas, que se relacionan con su posición en el test, su nivel de dificultad, e implicaciones en torno a la resolución de los ítems a medida que se avanza en la prueba.

Hay un primer conjunto de reactivos que cumplen a cabalidad con las hipótesis señaladas anteriormente, y cuya estructura permite que la posibilidad de que ocurran respuestas correctas tras un proceso eductivo erróneo sea mínima. Ellos son los ítems 1, 2, 3 y 4, en los cuales se observa una clara preferencia de los examinados por escoger la alternativa correcta, lo cual ciertamente contribuye a su ubicación dentro del test por ser los reactivos menos difíciles de la prueba.

El resto de las alternativas son marcadas por un número mucho menor de sujetos siguiendo una distribución prácticamente aleatoria, es decir, plana o rectangular. Además de ello, la escogencia de las alternativas correctas va relacionada con medias de puntajes obtenidos significativamente superiores a la de aquellos sujetos que escogen el resto de las alternativas, no habiendo diferencias entre quienes optan por los distractores.

Como se observa en el ejemplo presentado en la Tabla 8, correspondiente al ítem 4 del TAMPR abreviado, quienes fallan en acertar los primeros ítems de la prueba son quienes obtienen los puntajes de la porción inferior del continuo latente que indica este test. Ello implica un fracaso en educir el tipo de relaciones involucradas en las matrices avanzadas, por lo que este grupo de sujetos deberían ser preferentemente evaluados mediante un test de menor dificultad, como el de las matrices estándar.

Tabla 8: ANOVA de las alternativas del ítem 4.

<i>Alternativa</i>	<i>n</i>	<i>Media</i>	<i>S</i>
1	2	6.0	5.66
2	1	2.0	
3	3	7.0	2.65
4	6	4.2	2.14
5 *	103	8.5	2.33
6	3	5.3	2.52
7	1	4.0	
8	3	3.3	4.04

* Respuesta correcta.

Subsiguientemente, se encuentra un conjunto de reactivos de mediana dificultad, donde un número considerable de sujetos se concentra en optar por determinada alternativa incorrecta, siendo su promedio de puntajes obtenidos de bajo a medio. En los ítems 5, 6, 7, 8 y 9 (ver ejemplo en Tabla 9) - los cuales ocupan la porción central de la prueba -, dentro de su función discriminativa se incluyen tanto nuevas reglas de progresión como elementos de figuras que concurren en las matrices y las tornan más complicadas. De allí que el fracaso en estos problemas vendría constituido por la no advertencia del sujeto de una de las reglas, y optar por aquella alternativa que cumple con requerimientos parciales de la progresión de la matriz.

Tabla 9: ANOVA de las alternativas del ítem 8.

<i>Alternativa</i>	<i>n</i>	<i>Media</i>	<i>Desviación</i>
1	3	5.33	1.53
3	18	6.33	2.22
4	2	3.50	0.70
5	13	5.85	2.23
6 *	70	9.30	2.20
7	1	7.0	
8	15	6.40	2.20

Los distractores más poderosos se muestran en cursiva. Obsérvese el número de sujetos que las escoge y el puntaje medio que alcanzan respecto a quienes responden correctamente (* Respuesta correcta).

En el último segmento, ubicado hacia el final del test, están aquellos ítems en los cuales hay una alternativa que por alguna razón atrae la atención de unos pocos sujetos con puntajes similares a los del grupo que respondió correctamente, y lo constituyen los ítems 10 y 11 (ver ejemplo en la Tabla 10).

Tabla 10: ANOVA de las alternativas del ítem 10.

<i>Alternativa</i>	<i>n</i>	<i>Media</i>	<i>S</i>
1	9	5.11	2.93
2	13	6.62	1.98
3	4	9.0	1.41
4	6	6.33	2.50
5 *	56	9.61	2.05
6	13	6.15	2.51
7	5	6.2	1.10
8	16	7.0	2.22

En estos ítems es necesario mencionar que la posición que ocupan en el test, lógicamente, deviene de su nivel de dificultad, y éste le es conferido por el número y tipo de reglas implicadas en sus progresiones. Así, un subconjunto de sujetos que podrían eventualmente tener éxito en este punto de la prueba, ven atraída su atención fuera de la alternativa correcta al generar una solución sólo parcialmente correcta.

El ítem 12 tiene un carácter distintivo, cuyo comportamiento en relación a los demás se torna un poco desorganizado puesto que combina características del segundo y tercer tipo de reactivos descritos anteriormente (Tabla 11), lo cual se puede atribuir a la dificultad que presenta dentro de la muestra de sujetos evaluada. En este caso particular, el ítem 12 constituye una proyección hacia arriba en la medición del continuo latente de g ; así, responder correctamente este ítem, además de poco frecuente, tiene una probabilidad de ocurrencia al azar inferior a 0,0001, y ello implicaría que el evaluado ha sido exitoso en la resolución de todas o casi todas las matrices anteriores.

Tabla 11: ANOVA de las alternativas del ítem 12.

<i>Alternativa</i>	<i>n</i>	<i>media</i>	<i>S</i>
1	44	7.02	2.37
2	3	6.67	3.51
3 *	46	9.65	2.18
4	8	6.38	3.11
5	5	8.4	1.82
6	3	7.67	1.15
7	11	5.64	2.46
8	2	8.0	0.00

* Respuesta correcta.

Es por todo ello que se puede encontrar que quienes fallan en acertar este elemento, obtienen puntajes ubicados en posiciones superiores del recorrido. Además, se puede observar cómo aumenta la media de puntajes obtenidos, lo cual puede estar debido a que en este punto de la prueba hay

sujetos que no intentan darle solución a la matriz, lo cual infla el valor ya que sólo están considerados los sujetos que afrontaron este problema, siendo estos - al mismo tiempo - los que pertenecen al grupo de los más aventajados intelectualmente.

Finalmente, el indicador estadístico que resta para comparar el comportamiento psicométrico del TAMPR es la correlación que los puntajes obtenidos a través de éstos sostienen con otras medidas de ejecución, lo cual constituye otro indicio de la validez de constructo, a través de la evaluación de la convergencia o divergencia de las medidas correlacionadas (Anexo H). En este caso, la variable de criterio es la calificación lograda por los sujetos evaluados en Estadística I (Lógica Jurídica en el caso de los estudiantes de la Escuela de Derecho), la cual se expresa como puntaje estandarizado Z por Escuela, a fin de eliminar las diferencias entre las distribuciones (Anexo P).

A continuación, en la Tabla 12, se presenta un análisis descriptivo de esta variable de comparación, en el cual se muestran las medias y desviaciones típicas de la misma, especificándose el curso considerado por carrera:

Tabla 12: Media y desviación típica de las notas previas por carrera.

<i>Carrera</i>	<i>Materia</i>	<i>Media</i>	<i>Desviación</i>
Psicología	Estadística I	7,765	2,67
Relaciones Industriales	Estadística I	13,162	2,758
Administración y Contaduría	Estadística I	n.d.	n.d.
Ciencias Pedagógicas	Estadística I	10,857	2,193
Derecho	Lógica	15,846	2,035
Economía	Estadística I	11,633	2,395
Comunicación Social		n.d.	n.d.

n.d.: no disponible. No se estimaron los estadísticos de Administración y Contaduría, ya que dicha muestra sólo la compone un conjunto poco representativo de dos sujetos, y por lo tanto no están incluidos en el cálculo de la correlación. Igualmente quedaron excluidos los sujetos de Comunicación Social, quienes no reportaron sus calificaciones.

A pesar de que las medias observadas son distintas, la dispersión y la forma de todas las distribuciones son semejantes, manteniéndose el grueso de

la población a un poco más de 2 puntos alrededor de los promedios, con una ligera asimetría coleada hacia adentro y forma similar a la de una distribución semi - simétrica campanoide.

Las diferencias entre el aplanamiento o abultamiento de las curvas hacia el centro, resultaron en que las distribuciones de calificaciones de Derecho y Ciencias Pedagógicas son leptocúrticas, es decir, altas y delgadas; mientras que las de Psicología, Economía y Relaciones Industriales son mesocúrticas, con una leve tendencia al aplanamiento (Ver tabla 13).

Tabla 13: Estadísticos de forma de las distribuciones de las calificaciones por carrera.

<i>Carrera</i>	<i>Asimetría</i>	<i>Kurtosis</i>
Psicología	- 0,394	- 0,251
Relaciones Industriales	- 0,084	- 0,076
Administración y Contaduría	n.d.	n.d.
Ciencias Pedagógicas	- 0,404	0,946
Derecho	- 0,541	- 1,488
Economía	- 0,019	0,109
Comunicación Social	n.d.	n.d.

Sobre la base de estos puntajes, se observó que tanto la versión original como la abreviada sostienen una correlación de 0,30 con la medida de rendimiento académico, cuya magnitud es significativa al poseer una probabilidad de ocurrencia al azar inferior a 0,001 (Anexo L). Aún así, lo que cabe destacar de este resultado no es la proporción en que dicha medida contribuye a la validez convergente de lo que miden las matrices, sino la similitud de la correlación existente entre las dos formas del test y dicho criterio (Tabla 14).

Tabla 14: Correlación de ambas versiones con el rendimiento académico.

<i>Versión del test</i>	<i>Original</i>	<i>Abreviada</i>
Correlación	0,3081	0,3049

DISCUSIÓN

Los resultados en esta sección discutidos se fundamentan en la comparación del comportamiento psicométrico de las versiones larga y abreviada del TAMPR, así como el de esta última con el obtenido por la investigación predecesora realizada por Arthur y Day (1994), con miras a establecer una conclusión sobre los alcances en utilidad y calidad de la información que se logran en la población estudiada.

En primer lugar, como resultado global, todas las suposiciones formuladas para este estudio fueron confirmadas; como se expone a continuación:

Confiabilidad

Una de las claves que define lo que es un test es que éstos constituyen una muestra de conducta (Anastassi, 1968/1974), la cual a su vez conforma una serie de indicadores observables de un constructo hipotético. De este modo, para obtener una medición precisa de dicho constructo, la muestra de indicadores debe ser representativa del mismo y en la medida que lo sea el nivel de confiabilidad del instrumento será mayor. Por lo tanto, mientras más grande es una muestra de conducta, teóricamente su posibilidad de replicar los mismos resultados también es más grande. Así, al alterar la longitud de esa muestra, también teóricamente se altera su nivel de confiabilidad, alteración esta no relacionada con fuentes de error ocurridos durante la examinación, sino al cambio ocurrido en la cantidad de indicadores que se poseen del constructo (Magnusson, 1969/1990).

Dado que a través de la versión corta del TAMPR se reduce la muestra original de 36 matrices a 12 de ellas, es decir, se acorta la longitud del test a un tercio del original, cabe esperar teóricamente una reducción en la confiabilidad. De acuerdo a la fórmula Spearman - Brown, un test con las características del TAMPR, cuya confiabilidad está en torno a 0,84 - tanto en el

trabajo de Arthur y Day (1994) como en el presente -, debería alcanzar un coeficiente de 0,65 al reducirse su longitud a la proporción mencionada. Sin embargo, el descenso observado en el nivel de confiabilidad de la versión corta de 12 ítems es menor al esperado (0,73) - equivalente al de la investigación predecesora - manteniéndose un buen estatus a pesar de la drástica reducción en su longitud, lo cual es consecuencia de la adecuada selección de los ítems que mejor representan la muestra de la versión original.

Validez

La validez como propiedad psicométrica de un test significa que éste, en su rol de instrumento de medida, aporte lo más cercanamente posible indicadores observables de aquel constructo para el cual fue concebido. Retomando el concepto de muestreo de conductas, se tiene que así como la representatividad de ésta contribuye a que las estimaciones obtenidas sean confiables, también en la medida en que se cuenta con una selección de elementos que genuinamente se correspondan a expresiones del constructo a medir, también se obtendrá lo que se conoce como una estimación válida del mismo.

Como se describiera en las bases teóricas, las matrices constituyen patrones gráficos que siguen una regla de progresión y que están libres de componente lingüístico, diseñados por Raven (s/f) con la finalidad de medir el factor g de inteligencia, bajo la suposición de que en efecto lo hacen, tal como recientemente lo han sustentado nuevas disertaciones metodológicas en torno al *Raven* (Carpenter, Just y Shell, 1990).

En este orden de ideas, la validez de contenido, es decir, la validez de la muestra de reactivos per se que conforman el test, se limita en este caso a la selección de ítems a partir de la forma original del TAMPR realizada por Arthur y Day (1994) para conformar la nueva prueba, que se sustenta en el proceso de toma de decisiones fundamentado en el nivel de confiabilidad creciente de los

elementos, su correlación corregida con los puntajes obtenidos en el test total y el coeficiente de confiabilidad calculado si se elimina un ítem.

De la aplicación de los dos ítems seleccionados a una muestra de sujetos como prueba independiente, se encontró que tanto la versión larga como la abreviada sostienen una estructura factorial similar compuesta por: (a) un factor predominante a través de la mayoría de los reactivos, que podría representar la proporción de factor *g* que miden los ítems de ambas versiones y (b) como lo señala Tavella (1978) y Franco y Prengler (1980) otros dos factores más pequeños cuya existencia podría deberse al nivel de dificultad creciente de los ítems, aspecto éste que contribuye a la centralidad de las pruebas de Raven como medida de *g* (Carpenter y cols. 1990; ver Anexo B).

A pesar de que la elaboración de las matrices problema está concebida bajo fundamentos teóricos que definen lo que es una estructura factorial unitaria, como lo es el pensamiento abstracto, claro y preciso para afrontar situaciones nuevas y con libertad de componentes culturales (Arthur y Day, 1994; Carpenter y cols., 1990; DeShon, Chan y Weissbein, 1995; Raven y cols, 1994), no es nueva la presencia de algunos factores residuales adicionales diferentes de uno grande y presente a través de la mayor parte de los ítems (Franco y Prengler, 1980). La existencia de dichos componentes parece tener una explicación viable basada en el muestreo de sujetos a quienes se le aplica el test.

Si se retoma la exposición de los Resultados, y se observan las cargas de factor presentes en los componentes residuales más grandes (el 2 y el 3, ver Tablas 4 y 6), se advierte que éstas denotan un patrón de ejecución tal vez asociado a un grupo de sujetos para quienes el test representa una dificultad superior a la usual que podrían percibir aquellos que obtienen las mayores puntuaciones. Dicho patrón se corresponde con una proporción inversa de los aciertos y desaciertos en la solución de los problemas que se encuentran en la porción media inferior del test en contraste con los que están en la porción media superior del mismo, correspondiendo cada uno de estos sectores del test

a los componentes antes mencionados. Es allí donde se vislumbra la posibilidad de que tales componentes residuales correspondan a un factor de dificultad implicado en las matrices problema.

El análisis de varianza de las alternativas de respuesta hecho para cada ítem de la forma corta permite evidenciar la existencia de tales grupos de sujetos para quienes, a partir de cierto punto, según su capacidad intelectual, la prueba comienza a ser "psicológicamente problemática", y es que a partir de la porción media del test los distractores empiezan a desplegar un comportamiento cada vez más semejante al de las respuestas correctas (ver Tablas 8, 9, 10 y 11). Ello influye, por lo tanto, en el aumento del nivel de dificultad creciente de los ítems, asociado a la percepción de tales sujetos de que la opción que completa correctamente la matriz es otra que se ha descubierto, desde luego, bajo un proceso eductivo erróneo.

Para discutir este resultado, que a simple vista pareciera ser contraproducente, cabe considerar a Spearman (1955, c.p. Tavella, 1978), quien señala que los tests más confiables son aquellos cuyos ítems tienen un nivel de dificultad de 0,5; es decir, que son realizados correctamente en su totalidad sólo por el 50% de la población. La razón de ello es que separan a ésta en dos gruesas porciones correspondientes de la cual resulta una categorización que dicotomiza artificialmente el continuo latente que se intenta medir, entonces es mayor la probabilidad de que se repliquen los resultados que aporta un test, y por ende es más confiable.

Sin embargo, en la misma discusión hecha por Spearman, también se señala que la dificultad de los ítems es una propiedad relativamente artificial, pues depende de la población sobre cuya base se recogen los datos de evaluación, es en esencia una adecuación accidental a los sujetos examinados.

El hecho de que todos los ítems de un test sean resueltos correctamente sólo por el 50% de la población reduce su poder discriminativo ya que divide a los sujetos en dos grandes grupos donde no se conocen los niveles medios

entre ellos o, mejor dicho, su ubicación más precisa dentro del continuo latente estimado. De allí que, todo test de ejecución máxima cuenta o debe contar con una muestra de elementos cuya posibilidad de ser acertados se distribuya progresivamente desde los más fáciles hasta los más difíciles, es decir, pudiendo establecer estimaciones que abarquen desde los sujetos en los niveles inferiores de magnitud del constructo hasta los superiores.

Este factor de dificultad denota su importancia por dos razones:

1. Rango de aplicabilidad del test: Esto quiere decir que dependiendo de la dificultad de los ítems que componen un test se deriva el tipo de población a la cual puede ser aplicado. Entonces, dado que el TAMPR original fue diseñado como progresión hacia arriba para poblaciones de alto rendimiento intelectual, quienes consideran los problemas de la versión estándar como fáciles (Raven, Court y Raven, 1994), es congruente, por lo tanto, encontrar que un muestreo realizado a partir de éste – constituido por la versión corta – no sólo presente rangos de dificultad semejantes sino que también posea este componente dentro de su estructura factorial, conservándose como en la prueba original en un estatus de componente francamente secundario al factor g, asociado sólo a una parte de la población, aquella para la cual el test es “psicológicamente problemático” (Raven y cols., 1994).
2. Validez de criterio: el aporte del nivel de dificultad de los ítems contribuye al poder discriminativo del test, y juega un papel importante en las estimaciones que a partir de los puntajes obtenidos en el test se puedan hacer respecto a una medida de ejecución futura.

La interpretación del significado de los factores en los tests de Raven en general se dificulta un poco dado el carácter no lingüístico de sus elementos, por lo que la existencia en ambas versiones de un factor grande presente en magnitudes significativas a través de la mayoría de los ítems y que permite un

considerable grado de comunalidad entre ellos permite señalar, además de la existencia de una estructura factorial predominantemente unitaria, que ese componente principal corresponde al factor g de inteligencia, dado que teóricamente para su medición fueron diseñadas las matrices progresivas. Estas conclusiones son congruentes con el parámetro que como requisito precede a la validez, es decir la confiabilidad, que en este caso corresponde a la consistencia interna, homogeneidad o correlación existente entre las respuestas a todos los ítems, la cual en ambas formas del test se conserva considerablemente alta, que indica que todos los elementos comportan un mismo constructo.

Otro parámetro mediante el cual se evalúa la validez de constructo, relativo a dos tipos de estimaciones estadísticas mediante las cuales se prueba la correlación de la medida del test en cuestión con otros tipos de medidas, de lo cual se espera obtener: (a) una validez convergente, aquella cuya correlación es de una magnitud considerable o significativa y que es deseable cuando se supone que las medidas de comparación intentan medir lo mismo que el test que se está sometiendo a prueba, (b) la validez divergente, donde la expectativa es que la correlación entre el test y otra medida que se hipotetiza no relacionada sea cercana a cero.

En el caso de esta investigación, se correlacionaron los puntajes obtenidos por los sujetos en las versiones del test que les fueron aplicadas con una medida de ejecución académica, cuya finalidad no era establecer el tipo de validez de criterio ni lograr que representara una estimación de validez convergente, sino observar la magnitud y dirección de la relación que sostiene cada forma del TAMPR con el rendimiento académico. Así, se obtuvo que tanto la versión larga como la corta poseen el mismo nivel de relación respecto al criterio considerado, siendo además de magnitud significativa y en proporción directa; es decir, que los puntajes altos obtenidos en cualquier forma del TAMPR se relacionan con los puntajes superiores del rendimiento académico, e igualmente se espera respecto a los puntajes medios y bajos.

Cabe destacar que el valor encontrado se ubica en el recorrido medio del rango de correlaciones obtenido por Arthur y Day (1994). Ellos correlacionaron las medidas de ambas versiones del TAMPR con otros tests de ejecución intelectual utilizados en el reclutamiento y selección de personal. Allí las correlaciones variaron entre 0,20 y 0,50; encontrándose, además, que la magnitud de las relaciones de cada forma de las matrices con las otras medidas era similar.

Por otra parte, aunque existe una correlación significativa, no se espera que sea muy alta, dadas las características del criterio, el cual se podría hipotetizar que posee una estructura factorial más heterogénea que la de las matrices progresivas. Si bien las pruebas de Raven y los cursos de estadística y lógica pueden compartir una buena proporción de varianza en torno al componente eductivo del factor g, la ejecución dentro de tales cursos también depende del componente reproductivo del factor g, así como posiblemente de otros factores específicos o de grupo.

Análisis de ítems

Los parámetros para comparar el comportamiento de los ítems en la versiones original y abreviada del TAMPR son aquellos según los cuales Arthur y Day (1994) realizaron la selección para formar la prueba corta.

En primer lugar, las dos versiones del test mantienen un crecimiento progresivo del nivel de dificultad desde el primer ítem hasta el último, con una pequeña diferencia en el perfil mostrado por cada prueba. Del análisis de datos se desprende que los primeros ítems de la versión larga adquieren un índice de dificultad ligeramente superior al de sus correspondientes en la prueba corta, luego hacia el centro de las pruebas, el índice de aciertos se torna equivalente para finalmente invertirse la relación hacia los últimos elementos, donde la dificultad de la prueba corta es ligeramente inferior a la observada en la prueba larga.

Este fenómeno se hace previsible o puede ser explicado por dos factores:

1. La discrepancia entre la dificultad estadística del ítem y la dificultad psicológica que le atribuye el sujeto al mismo durante la resolución del test. Raven y Cols. (1994) establecen una diferencia en el incremento del nivel de dificultad entre el Set I (12 ítems de prueba) y el Set II (36 ítems) advirtiéndolo que los saltos en el incremento de la dificultad en el set de prueba son mayores que los existentes en el set de evaluación. De allí que para un sujeto que esté resolviendo la versión corta, tal como si tratara de resolver el Set I, va a percibir una mayor dificultad en los ítems subsiguientes que aquellos sujetos que resuelven la prueba larga, dado que los saltos en esta última son paulatinos, en contraste con los de aquella que son más abruptos. Entonces, aunque los procesos de pensamiento para encontrar la respuesta correcta sean los mismos, el salto en la dificultad juega un papel importante en el éxito del sujeto.
2. Efectos de fatiga: la explicación anterior puede funcionar para comprender por qué los primeros ítems adquieren mayor dificultad en la prueba corta. No obstante, la otra pregunta es por qué disminuye la dificultad hacia el final de la prueba. La respuesta a este fenómeno puede encontrarse dentro de una de las denominadas fuentes de error: la fatiga (Magnusson, 1969/1990). Lo que se supone es que durante la ejecución en la prueba larga, el constante trabajo de explorar las matrices, encontrar su regla de progresión y encontrar una respuesta, tarea tal que requiere de claridad de pensamiento y un importante consumo de energía cognoscitiva, hace que los sujetos desacierten en mayor proporción cuando tratan de resolver los últimos ítems de la prueba larga, bien sea por cansancio, falta de motivación en invertir energía para lograr la solución, o una percepción de dificultad incrementada que incide sobre sus expectativas de ejecución. En cambio, para la prueba corta, aunque

se perciban mayores saltos de dificultad de un ítem a otro, los sujetos pueden completar la prueba o el test con una mayor claridad de pensamiento debido a un menor gasto intelectual.

Es oportuno retomar la discrepancia existente en el crecimiento progresivo de ambas pruebas. Del análisis gráfico de los Resultados se desprende que la línea de regresión a la mediana que posee cada perfil, además de diferente, señala características de la población; aspecto éste que vuelve a introducirse aquí, luego de haber sido considerado anteriormente en la discusión de la estructura factorial.

La línea de regresión de la dificultad de los ítems en la versión abreviada crece estable y monótonicamente, lo cual da una señal sobre su adecuabilidad en la población estudiada. En contraste, la línea de regresión presente en la prueba original despliega un crecimiento segmentado en tres cambios de tendencia que marcan tres grupos consecutivos de reactivos. La primera y la última porción poseen una tendencia plana, cada una correspondiente a los niveles más fácil y más difícil de la prueba para esta población. Mientras que es la porción central la que muestra un perfil semejante al encontrado en la versión abreviada.

Nuevamente, retomando nociones relativas al nivel de dificultad y el poder discriminativo de los ítems, se entiende que este fenómeno pareciera pertenecer a características de la población evaluada y no a propiedades esenciales de los ítems; ya que la versión larga muestra grados de dificultad creciente de mayor estabilidad en otras poblaciones (Arthur y Day, 1994; Raven y cols., 1994); pero tampoco se pueden negar las diferencias encontradas de un país a otro en cuanto a puntajes y recorrido de los mismos (Court, 1991; Franco y Prengler, 1980; Jaworska y Szustrowa, 1993; Owen, 1992).

En tal sentido, lo que se plantea es que el grupo de ítems de la versión larga que contribuye más válidamente a la estimación de g en la UCAB es el de

la porción central, alrededor de los ítems 12 y 30. Por lo tanto, a pesar de lo adecuado que ha resultado ser la versión corta propuesta por Arthur y Day (1994), una nueva sugerencia podría estar enmarcada en esa sección del TAMPR original.

El siguiente parámetro de comparación lo constituyen (a) la correlación ítem - test y (b) el coeficiente alpha si el ítem es eliminado. Para ambas pruebas, los valores de estos parámetros se despliegan en distribuciones prácticamente rectangulares, donde la calidad de los ítems es bastante homogénea tal como lo confirma su nivel de consistencia interna y su estructura factorial. De este planteamiento, se puede especular que tales indicadores no constituyen criterios suficientemente sólidos para suponer que una versión corta pueda ser casi tan buena como la original. Por una parte, se podría decir que la selección de ítems realizada es ciertamente eficaz, debido a que su desempeño psicométrico es bastante similar al de la muestra de la cual partió. Sin embargo, sobre la base de lo expuesto al principio de este párrafo, también se podría especular que si se seleccionan otros ítems de semejante nivel de dificultad con menor correlación ítem - test, y por ende, mayor alpha si se borra, los resultados obtenidos no diferirían mucho de los actuales.

Una particularidad del comportamiento de los elementos lo constituyen los ítems 1 y 36 de la prueba larga, que tienen implicaciones a la hora de una posible selección para conformar una prueba corta. Estos reactivos se colocan muy por fuera del perfil de los parámetros psicométricos considerados, mostrando respectivamente un nivel de dificultad sumamente bajo y sumamente alto, al igual que sus correlaciones ítem - test y un incremento del alpha si son eliminados que los distingue de la distribución plana que despliega el resto de los ítems. Estas propiedades reflejan un bajo poder discriminativo, puesto que:

1. En el caso del ítem 1, casi la totalidad de los sujetos - por lo menos dentro de la población examinada - logra acertar la respuesta, por lo

que constituye un problema de muy bajo piso para el grupo de sujetos en cuestión.

2. En el caso del ítem 36, casi la totalidad de los sujetos responde incorrectamente a la matriz, siendo éste un problema de muy alto techo para la misma población.

Sin embargo, estos ítems adquieren importancia ya que su utilidad dentro de la prueba podría ser la detección de aquellos sujetos cuyo desempeño intelectual se caracterice bien por ubicarse en el extremo inferior de la distribución, o en el superior, aspecto tal que implica que el peso de estos elementos deviene de aquello que lo diferencia de los demás; es decir, el ítem 1 es potente en la medida en que un grupo de sujetos desacierta su solución y el ítem 36 toma relevancia en la medida que otro grupo de sujetos es capaz de resolverlo correctamente.

Lo que se ha discutido en estas líneas, a partir del análisis de ítems, podría ser la base de un criterio adicional para el muestreo de los elementos que van a constituir una versión abreviada, que estaría relacionado a la finalidad por la cual la nueva forma es aplicada en cuanto a población blanco, necesidades de discriminación, validez de criterio, etc.

CONCLUSION Y RECOMENDACIONES

De la comparación psicométrica entre las versiones original y abreviada del TAMPR que resulta del presente estudio, se concluye que éste último constituye una medida confiable y válida de la capacidad intelectual general dentro de la población estudiada, puesto que mantiene un comportamiento similar entre ésta y otras reseñadas en estudios previos de versiones cortas del TAMPR.

Los resultados encontrados en este trabajo coinciden con la apreciación que se obtiene al revisar las disertaciones teóricas y metodológicas existentes en torno a las matrices progresivas de Raven, las cuales se concentran en discutir acerca de dos aspectos fundamentales: (a) la centralidad de estas pruebas como medida válida del factor g de inteligencia y (b) la independencia de componentes culturales.

El primer punto, como en todas las investigaciones que constituyen el marco conceptual de esta investigación, se verifica una vez mas; es decir, en efecto existe una estructura factorial predominantemente unitaria que, de acuerdo al basamento teórico sobre el cual fueron construidas las matrices, supone medir el factor g de inteligencia.

No obstante, dados los niveles de dificultad de los ítems y la consecuente variación del rango de puntajes que obtienen los sujetos de una población a otra, dar respuesta a una franca libertad de factores culturales queda fuera del alcance de este estudio, así como tampoco se han encontrado argumentos contundentes o definitivos acerca de este punto de la discusión.

Precisamente, es entorno a estas diferencias entre las poblaciones que surgen los principales motivos para generar un estudio de este tipo, ya que en ello están implicados el poder discriminativo de la prueba, su validez de criterio o capacidad de predecir ejecución futura en tareas determinadas y su adecuación dentro de distintos grupos de referencia, así como también razones de tipo social relativas a la justicia y/o equidad de las medidas obtenidas.

En este caso particular, tal como se desprende de los resultados y la discusión, se puede decir que uno de los hallazgos mas importantes aquí lo constituye haber encontrado el sector de ítems de la versión larga que mejor discriminan las magnitudes de factor g entre los estudiantes de la UCAB; es decir, aquellos que rodean la porción media de la muestra de reactivos entre los cuales se puede encontrar el piso y el techo que circunscriben la utilidad del TAMPR original para esta población.

Aún cuando la versión corta propuesta por Arthur y Day en 1994 logró tener un comportamiento psicométrico adecuado, no se podría negar que el conjunto de elementos mencionados anteriormente puedan constituir una submuestra mas adecuada para la elaboración de una versión corta destinada al grupo de referencia que compone los estudiantes de esta universidad.

De allí que además de los criterios de selección de ítems propuestos por dichos investigadores que han demostrado su utilidad como parámetros generales, los resultados actuales permiten sugerir otros criterios de carácter mas bien particular cuando se trata de elaborar una versión abreviada del TAMPR de acuerdo a la finalidad con que ésta se va a construir.

Es por ello que, siendo la forma del test puesta a prueba en esta investigación un submuestreo de elementos que pretenden estimar el factor g de inteligencia, que mantiene su calidad psicométrica casi tan buena como la del test original, además de la considerable reducción en el tiempo de evaluación y decremento de los efectos de fatiga durante su resolución - , su utilidad primordial deviene del alcance de la evaluación que realiza el aplicador que la escoge dentro de su batería de tests.

La versión corta actual es una forma general que resume o funge de promedio para todas las versiones cortas posibles que se puedan crear, lo cual tiene dos implicaciones básicas:

1. Si lo que se busca son estimaciones para llevar a cabo pre-clasificaciones, triajes, despistajes, obtener datos con fines de investigación, y

donde el gasto de tiempo constituye un aspecto logístico a considerar, es recomendable el uso de la versión corta en lugar de la original.

2. Dado que la investigación psicométrica es el fundamento de toda investigación en psicología, puesto que de ella se obtienen los instrumentos para obtener los indicadores de los constructos que ésta abarca, siempre resulta oportuno señalar formas de mejorar y afinar este tipo de herramientas. Es por ello que se sugiere la prueba y reformulación de nuevas versiones cortas destinadas principalmente a necesidades relacionadas a poblaciones particulares. Así, al permanecer inalteradas las propiedades psicométricas de la prueba relativas a la confiabilidad, estructura factorial y análisis de ítem en distintas muestras de población, lo que cabe seguir afinando es su poder discriminativo y predictivo respecto a otras poblaciones o ambientes, considerando variables tales como género, nivel socioeconómico, determinadas tareas técnicas u ocupacionales, profesiones, patologías, estudios de postgrado, etc.

REFERENCIAS

- Ackerman, P. (1992). Predicting individual differences in complex skill acquisition: Dynamics of ability determinants. *Journal of Applied Psychology*, 77, 598-614.
- Anastasi, A. (1968/1974). *Tests Psicológicos*. 3ª edición. España: Aguilar.
- Arthur, W., Barret, G. y Doverspike, D. (1990). Validation of an information processing test battery for the prediction of handling accidents among petroleum product transport drivers. *Journal of Applied Psychology*, 75, 621-628.
- Arthur, W. y Day, D. (1994). Development of a short form for the Raven Advanced Progressive Matrices Test. *Educational and Psychological Measurement*, 77, 394-403.
- Arthur, W. y Woehr, D. (1993). A confirmatory factor analytic study examining dimensionality of the Raven's Advanced Progressive Matrices. *Educational and Psychological Measurement*, 53, 471-478.
- Beaumont, J. y French, C. (1987). A clinical field study of eight automated psychometric procedures: the Leicester/DHSS Project. *International Journal of Man - Machine Studies*, 26, 661-682.
- Carpenter, P., Just, M. y Shell, P. (1990). What one Intelligence Test Measures: A Theoretical Account of the Processing in the Raven Progressive Matrices Test. *Psychological Review*, 97, 404-431.
- Cattell, R. (1963). Theory of Fluid and Crystallized Intelligence: A Critical Experiment. *Journal of Educational Psychology*, 54, 1-22.
- Court, J. (1991). Asian Applications of Raven's Progressive Matrices. *Psychologia: An International Journal of Psychology in the Orient*, 34, 75-85. (De *Psychological Abstracts*, 1992, 79, resumen nº 11844).
- Crawford, J., Allan, K. y Jack, A. (1992). Short-forms of the UK WAIS-R: Regression equations and their predictive validity in a general population sample. *British Journal of Clinical Psychology*, 31, 191-202.
- DeShon, R., Chan, D. y Weissbein, D. (1995). Verbal overshadowing effects on Raven's Advanced Progressive Matrices: Evidence for multidimensional performance determinants. *Intelligence*, 21, 135-155.
- Eysenk, H. (1967). Intelligence assessment: A theoretical and experimental approach. *British Journal of Educational Psychology*, 37, 81-98.

- Franco, L. y Prengler, A. (1980). *Análisis psicométrico del Test de Matrices Progresivas de Raven en su forma avanzada en una muestra de estudiantes de la UCAB* (Trabajo de grado). Venezuela: Universidad Católica Andrés Bello, Escuela de Psicología.
- Grossman, I., Mednitsky, S., Dennis, B., Scharff, L., y cols. (1993). Validation of an "amazingly" short form of the Wais-R for a clinically depressed sample. *Journal of Psychoeducational Assessment*, 11, 173-181.
- Gustafsson, J. (1984). A unifying model for the structure of intellectual abilities. *Intelligence*, 8, 179-208.
- Horn, J. (1985). Remodeling old models of intelligence. En Wolman, B. (Ed.), *Handbook of intelligence*. Nueva York: Willey.
- Jaworska, A. y Szustrowa, T. (1993). Polish standardization of RPM. Polish Psychological Bulletin, 24, 303-307. (De *Psychological Abstracts*, 1993, 81, resumen n° 43545).
- Johnson, N., Sacuzzo, D. y Guertin, T. (1994). The development and validation of a reliable alternate form for Raven's Standard Progressive Matrices. *Assessment*, 1, 315-319.
- Kerlinger, F. (1988/1990). *Investigación del comportamiento*. 2° edición. México: Trillas.
- Kern, R., Bordieri, J. y Taylor, D. (1993). A coparison of Raven's Standard Progressive Matrices with the Wechsler Adult Intelligence Scale - Revised for individuals receiving rehabilitation services. *Vocational Evaluation & Work Adjustment Bulletin*, 26, 53-56.
- Kubinger, K., Formann, A. y Farkas, M. (1991). Psychometric shortcomings of Raven's Progressive Matrices, in particular for computerized testing. *European Review of Applied Psychology / Revue Européenne de Psychologie Appliquée*, 41, 295-300.
- Larson, G., Alderton, D. y Kaupp, M. (1991). Dynamic administration of a general intelligence test. *Learning & Individual Differences*, 3, 123-134.
- Matarazzo, J. (1990). Psychological assessment versus psychological testing. *American Psychologist*, 45, 999-1017.
- Magnusson, D. (1969/1990). *Teoría de los Tests: Psicometría diferencial, psicología aplicada, orientación vocacional*. 2° edición. México: Trillas.
- Measso, G., Zappalà, G., Cavarzeran, F., Crook, T. y cols. (1993). Raven's colores progressive matrices: A normative study of a random sample of healthy adults. *Acta Neurologica Scandinavica*, 88, 70-74. (De *Psychological Abstract*, 1994, 81, resumen n° 4004).

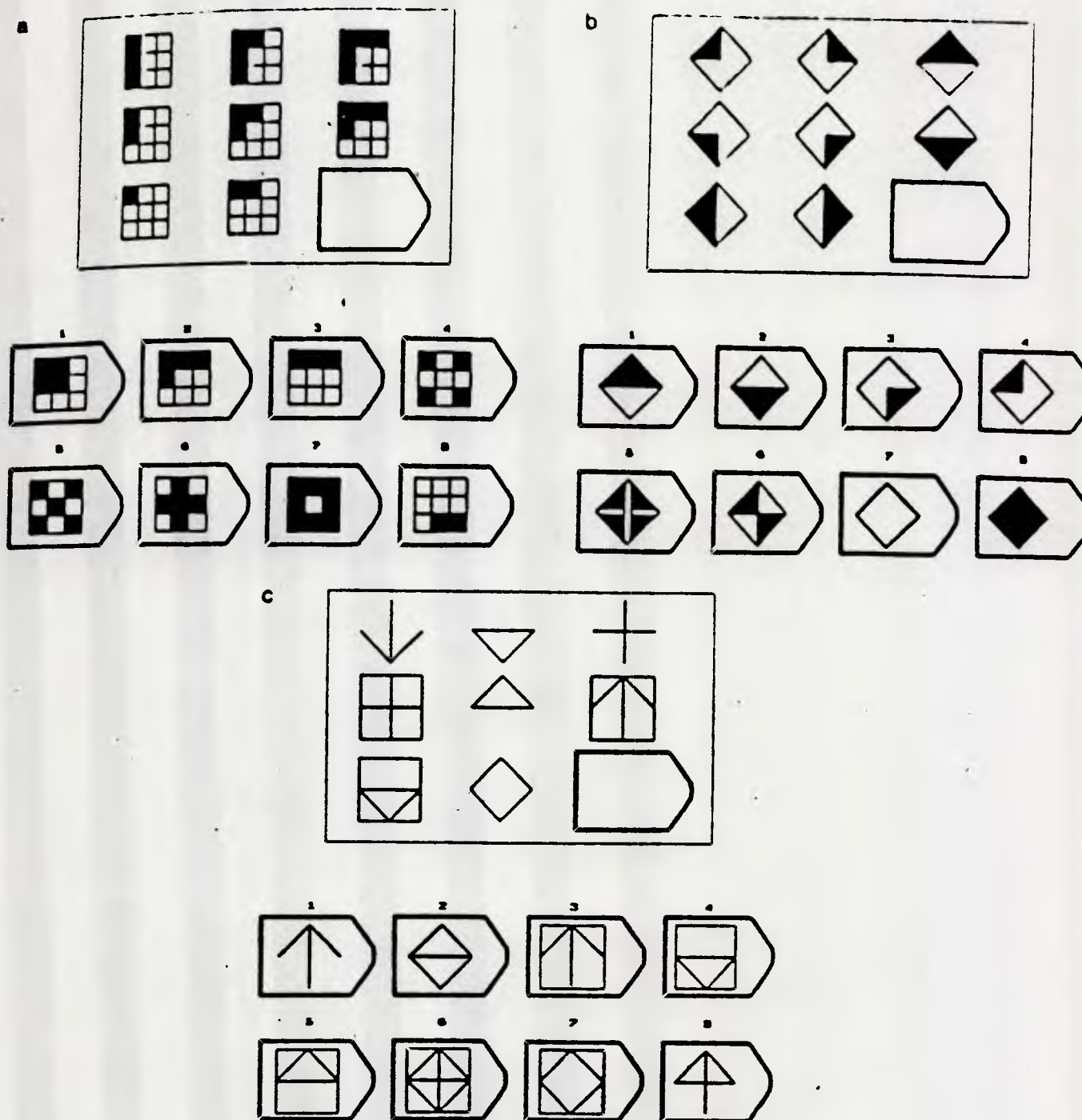
- Mills, C.; Ablard, K. y Brody, L. (1993). The Raven's Progressive Matrices: Its usefulness for identifying gifted/talented students. Special issue: Longitudinal studies in gifted education. *Roeper Review*, 15, 183-186. (De *Psychological Abstract*, 1993, 80, resumen n° 35967).
- Mills, C. y Tissot, S. (1995). Identifying academic potential in students from under-represented populations: Is using the Raven's Progressive Matrices a good idea. *Gifted Child Quarterly*, 39, 209-217.
- Norris, M. (1995). *Assessing intellectual ability: A comparison between computer based and paper-pencil format using Raven's Standard Progressive Matrices Test* (Reporte de Proyecto Final). Inglaterra: University of Westminster, School of Computer Science and Information System Engineering.
- Nunnally, J. (1973). *Introducción a la medición psicológica*. Buenos Aires: Paidós.
- Owen, K. (1992). The suitability of Raven's Standard Progressive Matrices for various groups in South Africa. *Personality and Individual Differences*, 13, 149-159.
- Raven, J. (Sin fecha). *Test de matrices progresivas para la medida de la capacidad intelectual: manual*. Buenos Aires: Paidós.
- Raven, J.C., Court, J. y Raven, J. (1994) *Advance progressive matrices. Raven manual: Section 4*. Gran Bretaña: Oxford Psychologists Press.
- Richardson, K. (1991). Reasoning with Raven: In and out of context. *British Journal of Educational Psychology*, 61, 129-138.
- Saccuzzo, D. y Johnson, N. (1995). Traditional psychometric tests and proportionate representation: An intervention and program evaluation study. *Psychological Assessment*, 7, 183-194.
- Sattler, J. (1988/1992). *Evaluación infantil*. 3ª edición. México: El Manual Moderno.
- Schuhfried, A. (sin fecha). Raven's Advanced Progressive Matrices (APM). (De: www.schuhfried.co.at/e/wts/apm.htm).
- Schuhfried, A. (sin fecha). Raven's Coloured Progressive Matrices (CPM). (De: www.schuhfried.co.at/e/wts/cpm.htm).
- Schuhfried, A. (sin fecha). Raven's Standard Progressive Matrices (SPM). (De: www.schuhfried.co.at/e/wts/spm.htm).
- Tavella, N. (1978). *Análisis de los ítems en la construcción de instrumentos psicométricos*. México: Trillas.

- Templer, D. (1992). Prison norms for Raven's Standard Progressive Matrices. *Perceptual & Motor Skills*, 74, 1193-1194.
- Van Den Broek, M. y Bradshaw, C. (1994). Detection of acquired deficits in general intelligence using the National Adult Reading Test and Raven's Standard Progressive Matrices. *British Journal of Clinical Psychology*, 33, 509-515.
- Vernon, P. (1965). Ability factors and environmental influences. *American psychologist*, 20, 723-733.
- Wytek, R., Opgenoorth, E. y Presslich, O. (1984). Development of a new shortened version of Raven's Matrices Test for an application and rough assessment of present intellectual capacity within psychopathological investigation. *Psychopatology*, 17, 49-58.

ANEXOS

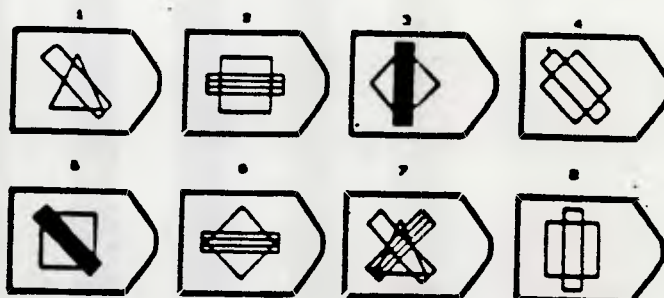
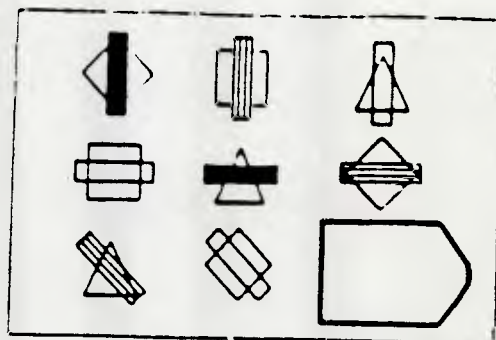
ANEXO A

Reglas de resolución de las matrices avanzadas



Problemas que ilustran la resolución de las Matrices Progresivas de Raven : Matriz A (Regla de la progresión cuantitativa), el número de cuadrados negros en el tope de cada fila se incrementa de uno en uno desde la primera columna a la segunda y a la tercera. El número de cuadrados negros se mantiene constante dentro de una fila, pero cambia entre las filas de tres, a dos, a uno. La respuesta correcta es 3. Matriz B (La regla de adición de figura), el elemento geométrico en la primera columna se superpone en el elemento geométrico de la segunda columna para componer el de la tercera. La posición del elemento oscurecido queda constante en una fila pero pasa de estar en el tope al fondo y finalmente arribes. La respuesta correcta es 8. Matriz C (Regla de distribución de dos valores), cada elemento geométrico, tal como la línea horizontal, la línea vertical y la V aparecen dos veces en una fila y el tercer valor es nulo. La respuesta correcta es 5.

(Tomado de Carpenter, Just y Shell, 1990, pág. 409)



...Continuación Anexo A

Matriz D (Regla de Distribución de tres valores), la variación entre las tres formas geométricas - diamante, cuadrado y triángulo - y las tres texturas de líneas - oscura, rayada y vacía - es gobernada por una regla de distribución de tres valores. La orientación de la línea - vertical, horizontal u oblicua - lo está bajo la regla de constante en una fila. La respuesta correcta es 5.

(Tomado de Carpenter, Just y Shell, 1990, pág. 407)

ANEXO B
Niveles de intercorrelación entre varios tests de habilidades
que ilustra la centralidad del test de Raven²



² Fuente: Carpenter, Just y Shell (1990), página 406.

ANEXO C

Hoja de respuesta de la versión original del TAMPR

UNIVERSIDAD CATÓLICA ANDRÉS BELLO
FACULTAD DE HUMANIDADES Y EDUCACIÓN
ESCUELA DE PSICOLOGÍA
TRABAJO DE GRADO

Test Avanzado de Matrices Progresivas de Raven

HOJA DE RESPUESTAS

Carrera _____ Calificación previa _____

INSTRUCCIONES: Abra su cuaderno en la primera página; observe que en la parte superior hay un dibujo en el que se ha omitido un trozo. Cada uno de estos trozos abajo tiene el tamaño adecuado para ajustarse al espacio, pero no todos completan el dibujo. Si, por ejemplo, el trozo número 4 es el correcto, escriba 4 en la casilla "Opción" correspondiente al número del "Ítem" que está resolviendo. Los problemas son simples al comienzo y se vuelven más difíciles a medida que avancen. Si presta atención a la manera como se resuelven los fáciles, los últimos le resultarán menos difíciles. Trabaje solo, resuelva todos los problemas, sin saltar ninguno ni volver atrás. Dispone de todo el tiempo que desee. Puede comenzar.

Ítem	Opción	Puntaje
1		
2		
3		
4		
5		
6		
7		
8		
9		
10		
11		
12		

Ítem	Opción	Puntaje
13		
14		
15		
16		
17		
18		
19		
20		
21		
22		
23		
24		

Ítem	Opción	Puntaje
25		
26		
27		
28		
29		
30		
31		
32		
33		
34		
35		
36		

PUNTAJE BRUTO _____

ANEXO D
Hoja de respuesta de la versión abreviada del TAMPR

UNIVERSIDAD CATÓLICA ANDRÉS BELLO
FACULTAD DE HUMANIDADES Y EDUCACIÓN
ESCUELA DE PSICOLOGÍA
TRABAJO DE GRADO

Test Avanzado de Matrices Progresivas de Raven
- versión abreviada -

HOJA DE RESPUESTAS

Carrera _____ Calificación previa _____

INSTRUCCIONES: Abra su cuaderno en la primera página; observe que en la parte superior hay un dibujo en el que se ha omitido un trozo. Cada uno de estos trozos abajo tiene el tamaño adecuado para ajustarse al espacio, pero no todos completan el dibujo. Si, por ejemplo, el trozo número 4 es el correcto, escriba 4 en la casilla "Opción" correspondiente al número del "Ítem" que está resolviendo. Los problemas son simples al comienzo y se vuelven más difíciles a medida que avancen. Si presta atención a la manera como se resuelven los fáciles, los últimos le resultarán menos difíciles. Trabaje solo, resuelva todos los problemas, sin saltar ninguno ni volver atrás. Dispone de todo el tiempo que desee. Puede comenzar.

Ítem	Opción	Puntaje
1		
2		
3		
4		
5		
6		
7		
8		
9		
10		
11		
12		

PUNTAJE BRUTO _____

ANEXO E:
Estadísticos descriptivos de la versión original del TAMPR

Mean	23,581	Median	24,000	Mode	25,000
Std dev	5,631	Kurtosis	1,684	S E Kurt	,292
Skewness	-,881	S E Skew	,146	Minimum	1,000
Maximum	35,000				

* Multiple modes exist. The smallest value is shown.

Valid cases 277

Value	Freq	Pct	Cum Pct	Value	Freq	Pct	Cum Pct	Value	Freq	Pct	Cum Pct
1,00	1	0	0	17,00	6	2	12	27,00	13	5	75
3,00	1	0	1	18,00	10	4	16	28,00	14	5	80
4,00	1	0	1	19,00	12	4	20	29,00	14	5	85
5,00	1	0	1	20,00	12	4	25	30,00	18	6	92
6,00	1	0	2	21,00	20	7	32	31,00	13	5	96
8,00	2	1	3	22,00	18	6	38	32,00	3	1	97
12,00	2	1	3	23,00	18	6	45	33,00	1	0	98
14,00	4	1	5	24,00	21	8	52	34,00	4	1	99
15,00	9	3	8	25,00	25	9	61	35,00	2	1	100
16,00	6	2	10	26,00	25	9	70				

ANEXO F

Propiedades psicométricas de la versión larga del TAMPR

RELIABILITY ANALYSIS - SCALE (ALPHA)

Reliability Coefficients

N of Cases = 252.0

N of Items = 36

Alpha = .8377

RELIABILITY ANALYSIS - SCALE (ALPHA)

Item-total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item- Total Correlation	Alpha if Item Deleted
I1A	22.6389	32.1599	.2837	.8350
I2A	22.6270	31.8444	.4039	.8328
I3A	22.5992	32.1535	.3734	.8339
I4A	22.6468	31.7194	.4005	.8325
I5A	22.6905	31.5293	.3863	.8324
I6A	22.6548	31.6214	.4143	.8321
I7A	22.6270	32.3862	.2322	.8361
I8A	22.6310	31.8991	.3777	.8332
I9A	22.7024	31.1661	.4627	.8304
I10A	22.6349	31.8582	.3821	.8331
I11A	22.6071	32.0323	.3930	.8334
I12A	22.6468	31.8150	.3726	.8331
I13A	22.9087	31.4060	.2904	.8351
I14A	22.6865	31.5467	.3868	.8324
I15A	22.7262	31.6578	.3195	.8340
I16A	22.7262	31.0841	.4543	.8304
I17A	22.7540	31.5647	.3203	.8339
I18A	22.8690	31.1103	.3582	.8328
I19A	22.7421	31.5547	.3309	.8336
I20A	22.8730	31.2348	.3325	.8336
I21A	22.9325	30.8918	.3828	.8320
I22A	22.9802	30.7526	.4012	.8314
I23A	22.9325	30.5253	.4529	.8296
I24A	23.0079	31.0358	.3461	.8333
I25A	22.9921	30.9242	.3679	.8325
I26A	23.0635	31.7569	.2139	.8377
I27A	23.1825	31.6638	.2442	.8365
I28A	23.1587	31.2337	.3205	.8341
I29A	23.2738	31.7056	.2632	.8357
I30A	23.0675	31.5771	.2466	.8366
I31A	23.2063	31.3278	.3145	.8342
I32A	23.1786	31.9162	.1960	.8381
I33A	23.0675	31.3301	.2917	.8351
I34A	23.1746	31.1168	.3463	.8332
I35A	23.2302	31.2377	.3402	.8334
I36A	23.4484	33.0850	.0136	.8402

ANEXO G

Nivel de dificultad de los ítems de la versión larga del TAMPR

		Mean	Std Dev	Cases
1.	I10A	.9048	.2941	252.0
2.	I11A	.9325	.2513	252.0
3.	I12A	.8929	.3099	252.0
4.	I13A	.6310	.4835	252.0
5.	I14A	.8532	.3546	252.0
6.	I15A	.8135	.3903	252.0
7.	I16A	.8135	.3903	252.0
8.	I17A	.7857	.4111	252.0
9.	I18A	.6706	.4709	252.0
10.	I19A	.7976	.4026	252.0
11.	I1A	.9008	.2995	252.0
13.	I21A	.6071	.4894	252.0
14.	I22A	.5595	.4974	252.0
15.	I23A	.6071	.4894	252.0
16.	I24A	.5317	.5000	252.0
17.	I25A	.5476	.4987	252.0
18.	I26A	.4762	.5004	252.0
19.	I27A	.3571	.4801	252.0
20.	I28A	.3810	.4866	252.0
21.	I29A	.2659	.4427	252.0
22.	I2A	.9127	.2828	252.0
23.	I30A	.4722	.5002	252.0
24.	I31A	.3333	.4723	252.0
25.	I32A	.3611	.4813	252.0
26.	I33A	.4722	.5002	252.0
27.	I34A	.3651	.4824	252.0
28.	I35A	.3095	.4632	252.0
29.	I36A	.0913	.2886	252.0
30.	I3A	.9405	.2371	252.0
31.	I4A	.8929	.3099	252.0
32.	I5A	.8492	.3586	252.0
33.	I6A	.8849	.3198	252.0
34.	I7A	.9127	.2828	252.0
35.	I8A	.9087	.2886	252.0
36.	I9A	.8373	.3698	252.0

ANEXO H

Validez de constructo del TAMPR original

ANEXO H1

Análisis factorial de la versión original del TAMPR

----- FACTOR ANALYSIS -----

Analysis number 1 Listwise deletion of cases with missing values

Extraction 1 for analysis 1, Principal Components Analysis (PC)

Initial Statistics:

Variable	Communality *	Factor	Eigenvalue	Pct of Var	Cum Pct
I10A	1.00000 *	1	6.03700	16.8	16.8
I11A	1.00000 *	2	1.91909	5.3	22.1
I12A	1.00000 *	3	1.62634	4.5	26.6
I13A	1.00000 *	4	1.40269	3.9	30.5
I14A	1.00000 *	5	1.36581	3.8	34.3
I15A	1.00000 *	6	1.26125	3.5	37.8
I16A	1.00000 *	7	1.20062	3.3	41.1
I17A	1.00000 *	8	1.16728	3.2	44.4
I18A	1.00000 *	9	1.14465	3.2	47.6
I19A	1.00000 *	10	1.11604	3.1	50.7
I1A	1.00000 *	11	1.09993	3.1	53.7
I20A	1.00000 *	12	1.03626	2.9	56.6
I21A	1.00000 *	13	.99245	2.8	59.4
I22A	1.00000 *	14	.91164	2.5	61.9
I23A	1.00000 *	15	.90640	2.5	64.4
I24A	1.00000 *	16	.85066	2.4	66.8
I25A	1.00000 *	17	.84127	2.3	69.1
I26A	1.00000 *	18	.81882	2.3	71.4
I27A	1.00000 *	19	.78833	2.2	73.6
I28A	1.00000 *	20	.78144	2.2	75.7
I29A	1.00000 *	21	.75631	2.1	77.8
I2A	1.00000 *	22	.73282	2.0	79.9
I30A	1.00000 *	23	.71831	2.0	81.9
I31A	1.00000 *	24	.69984	1.9	83.8
I32A	1.00000 *	25	.67630	1.9	85.7
I33A	1.00000 *	26	.63147	1.8	87.5
I34A	1.00000 *	27	.57328	1.6	89.0
I35A	1.00000 *	28	.54075	1.5	90.5
I36A	1.00000 *	29	.52861	1.5	92.0
I3A	1.00000 *	30	.48059	1.3	93.4
I4A	1.00000 *	31	.47073	1.3	94.7
I5A	1.00000 *	32	.45792	1.3	95.9
I6A	1.00000 *	33	.42037	1.2	97.1
I7A	1.00000 *	34	.40748	1.1	98.2

----- FACTOR ANALYSIS -----

Variable	Communality	* Factor	Eigenvalue	Pct of Var	Cum Pct
I8A	1.00000	*	35	.34554	1.0
I9A	1.00000	*	36	.29170	.8
					100.0

PC extracted 12 factors.

Factor Matrix:

	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
I9A	.55879				
I16A	.52077				
I2A	.51692				
I23A	.50874				
I6A	.50864				
I11A	.49656				
I4A	.49301				
I3A	.48652	-.36733			
I10A	.47759				
I8A	.47428				
I14A	.46697				
I12A	.46423				
I5A	.46350				
I21A	.44813				
I22A	.44576				
I18A	.44448				
I25A	.42115			.39968	
I19A	.40054		.37155		
I20A	.39431				
I15A	.39047		.35904		
I17A	.38649				
I1A	.38523	-.36148			
I24A	.38224				
I35A	.37718				
I28A					
I36A		.51537			
I31A		.41667			
I30A		.36007			
I29A					
I32A			.48241	.35975	
I34A			.37675		
I26A				.52083	

----- FACTOR ANALYSIS -----

	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
I7A				.37455	

I13A	.35549				
I27A			.37041		
I33A					

	Factor 6	Factor 7	Factor 8	Factor 9	Factor 10
--	----------	----------	----------	----------	-----------

I9A					
I16A					
I2A					
I23A					
I6A					
I11A					
I4A					
I3A					
I10A					
I8A					
I14A					
I12A					
I5A					
I21A					
I22A					
I18A					
I25A					
I19A					
I20A					
I15A					
I17A					
I1A					
I24A					
I35A					
I28A					
I36A					
I31A		.39158			
I30A					
I29A					
I32A					
I34A					
I26A					
I7A					

----- FACTOR ANALYSIS -----

	Factor 6	Factor 7	Factor 8	Factor 9	Factor 10
I13A			.45911	.36794	
I27A				.40523	
I33A					

	Factor 11	Factor 12
I9A		
I16A		
I2A		
I23A		
I6A		
I11A		
I4A		
I3A		
I10A		
I8A	.42752	
I14A		
I12A		
I5A		
I21A		
I22A		
I18A		
I25A		
I19A		
I20A		
I15A		
I17A		
I1A		
I24A		
I35A		
I28A		
I36A		
I31A		
I30A		
I29A		
I32A		
I34A		
I26A		
I7A		
I13A		
I27A		

----- FACTOR ANALYSIS -----

Factor 11 Factor 12
I33A .35355

Final Statistics:

Variable	Communality *	Factor	Eigenvalue	Pct of Var	Cum Pct
I10A	.64407 *	1	6.03700	16.8	16.8
I11A	.54952 *	2	1.91909	5.3	22.1
I12A	.46719 *	3	1.62634	4.5	26.6
I13A	.70773 *	4	1.40269	3.9	30.5
I14A	.62738 *	5	1.36581	3.8	34.3
I15A	.48973 *	6	1.26125	3.5	37.8
I16A	.60665 *	7	1.20062	3.3	41.1
I17A	.51686 *	8	1.16728	3.2	44.4
I18A	.49985 *	9	1.14465	3.2	47.6
I19A	.36352 *	10	1.11604	3.1	50.7
I1A	.51708 *	11	1.09993	3.1	53.7
I20A	.42893 *	12	1.03626	2.9	56.6
I21A	.65258 *				
I22A	.59114 *				
I23A	.53044 *				
I24A	.57789 *				
I25A	.49612 *				
I26A	.53808 *				
I27A	.71949 *				
I28A	.45220 *				
I29A	.49351 *				
I2A	.53751 *				
I30A	.68175 *				
I31A	.61891 *				
I32A	.67569 *				
I33A	.54121 *				
I34A	.60497 *				
I35A	.52802 *				
I36A	.62129 *				
I3A	.58763 *				
I4A	.58779 *				
I5A	.50238 *				
I6A	.55578 *				
I7A	.61315 *				
I8A	.65303 *				
I9A	.59788 *				

ANEXO H2
Correlación entre la versión original del TAMPR
y el rendimiento académico

Variable	Cases	Mean	Std Dev
PTJEBRUT	277	23,5812	5,6310
ZETA	247	-,0740	,9714

Variables	Cases	Cross-Prod Dev	Variance-Covar
PTJEBRUT ZETA	247	398,8936	1,6215

- - Correlation Coefficients - -

	PTJEBRUT	ZETA
PTJEBRUT	1,0000 (277) P= ,	,3081 (247) P= ,000
ZETA	,3081 (247) P= ,000	1,0000 (247) P= ,

(Coefficient / (Cases) / 2-tailed Significance)

" . " is printed if a coefficient cannot be computed

ANEXO I

Estadísticos descriptivos de la versión corta del TAMPR

Mean	7,926	Median	8,000	Mode	10,000
Std dev	2,697	Kurtosis	-,618	S E Kurt	,435
Skewness	-,380	S E Skew	,219	Minimum	1,000
Maximum	12,000				

Valid cases 122

Value	Freq	Pct	Cum Pct	Value	Freq	Pct	Cum Pct	Value	Freq	Pct	Cum Pct
1,00	1	1	1	5,00	9	7	20	9,00	11	9	65
2,00	2	2	2	6,00	12	10	30	10,00	21	17	82
3,00	6	5	7	7,00	15	12	42	11,00	11	9	91
4,00	6	5	12	8,00	17	14	56	12,00	11	9	100

ANEXO J **Propiedades psicométricas de la versión abreviada del TAMPR**

RELIABILITY ANALYSIS - SCALE (ALPHA)

Reliability Coefficients

N of Cases = 122.0

N of Items = 12

Alpha = .7294

Item-total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item- Total Correlation	Alpha if Item Deleted
I1A	7.0328	6.8584	.1979	.7287
I2A	7.1639	6.2208	.4090	.7060
I3A	7.0902	6.4298	.3753	.7112
I4A	7.0820	6.5056	.3431	.7148
I5A	7.1967	6.1428	.4223	.7039
I7A	7.3279	5.9908	.4326	.7018
I6A	7.2131	6.3509	.3138	.7182
I8A	7.3443	5.9301	.4561	.6983
I9A	7.2787	6.3845	.2726	.7243
I10A	7.4754	5.9870	.4249	.7029
I11A	7.4344	6.1486	.3515	.7138
I12A	7.5492	6.2166	.3388	.7153

ANEXO K
Nivel de dificultad de los ítems de la versión
abreviada del TAMPR

		Mean	Std Dev	Cases
1.	I10A	.4508	.4996	122.0
2.	I11A	.4918	.5020	122.0
3.	I12A	.3770	.4866	122.0
4.	I1A	.8934	.3098	122.0
5.	I2A	.7623	.4274	122.0
6.	I3A	.8361	.3717	122.0
7.	I4A	.8443	.3641	122.0
8.	I5A	.7295	.4460	122.0
9.	I7A	.5984	.4923	122.0
10.	I6A	.7131	.4542	122.0
11.	I8A	.5820	.4953	122.0
12.	I9A	.6475	.4797	122.0

ANEXO L
Validez de constructo del TAMPR abreviado

ANEXO L1
Análisis factorial de la versión corta del TAMPR

PC extracted 3 factors.

Factor Matrix:

	Factor 1	Factor 2	Factor 3
I8A	.59089		
I7A	.58298		
I5A	.57271		
I2A	.57058	.43083	
I10A	.56887		
I3A	.54521	.46235	
I4A	.51102	.48084	-.39485
I11A	.47689	-.46457	
I6A	.44804		
I12A	.43817	-.52460	.36321
I1A		.45031	.43073
I9A	.37766		.61341

Final Statistics:

Variable	Communality *	Factor	Eigenvalue	Pct of Var	Cum Pct
I10A	.42224 *	1	3.07608	25.6	25.6
I11A	.45776 *	2	1.55280	12.9	38.6
I12A	.59912 *	3	1.22975	10.2	48.8
I1A	.47701 *				
I2A	.60631 *				
I3A	.57782 *				
I4A	.64825 *				
I5A	.36081 *				
I6A	.22675 *				
I7A	.49341 *				
I8A	.46941 *				
I9A	.51972 *				

Skipping rotation 1 for extraction 1 in analysis 1
Analysis number 1 Listwise deletion of cases with missing values

Extraction 1 for analysis 1, Principal Components Analysis (PC)

Initial Statistics:

Variable	Communality *	Factor	Eigenvalue	Pct of Var	Cum Pct
I10A	1.00000 *	1	3.07608	25.6	25.6
I11A	1.00000 *	2	1.55280	12.9	38.6
I12A	1.00000 *	3	1.22975	10.2	48.8
I1A	1.00000 *	4	.96973	8.1	56.9
I2A	1.00000 *	5	.90463	7.5	64.4
I3A	1.00000 *	6	.84225	7.0	71.5
I4A	1.00000 *	7	.73317	6.1	77.6
I5A	1.00000 *	8	.66710	5.6	83.1
I6A	1.00000 *	9	.60419	5.0	88.2
I7A	1.00000 *	10	.54031	4.5	92.7
I8A	1.00000 *	11	.44890	3.7	96.4
I9A	1.00000 *	12	.43110	3.6	100.0

ANEXO L2
Correlación entre la versión corta del TAMPR
y el rendimiento académico

Variable	Cases	Mean	Std Dev
PTJEBRUT	122	7,9262	2,6973
ZETA	116	,1896	1,2051

Variables	Cases	Cross-Prod Dev	Variance-Covar
PTJEBRUT ZETA	116	113,2941	,9852

- - Correlation Coefficients - -

	PTJEBRUT	ZETA
PTJEBRUT	1,0000 (122) P= ,	,3049 (116) P= ,001
ZETA	,3049 (116) P= ,001	1,0000 (116) P= ,

(Coefficient / (Cases) / 2-tailed Significance)

". " is printed if a coefficient cannot be computed

ANEXO M
Propiedades psicométricas de los 12 ítems de la versión corta
dentro de la larga

RELIABILITY ANALYSIS - SCALE (ALPHA)

Reliability Coefficients

N of Cases = 262.0

N of Items = 12

Alpha = .6505

Item-total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item- Total Correlation	Alpha if Item Deleted
I1A	7.1412	4.8344	.2418	.6387
I4A	7.1412	4.7807	.2828	.6335
I8A	7.1260	4.7465	.3409	.6273
I11A	7.0992	4.8177	.3514	.6291
I15A	7.2214	4.5638	.3252	.6251
I18A	7.3588	4.3919	.3293	.6233
I21A	7.4160	4.2822	.3675	.6155
I23A	7.4198	4.3058	.3537	.6184
I25A	7.4809	4.2659	.3623	.6165
I30A	7.5573	4.6308	.1756	.6552
I31A	7.6985	4.5792	.2253	.6438
I35A	7.7176	4.5176	.2647	.6359

ANEXO N
Nivel de dificultad de los 12 ítems de la versión abreviada
dentro de la original

		Mean	Std Dev	Cases
1.	I1A	.8931	.3095	262.0
2.	I4A	.8931	.3095	262.0
3.	I8A	.9084	.2890	262.0
4.	I11A	.9351	.2468	262.0
5.	I15A	.8130	.3907	262.0
6.	I18A	.6756	.4691	262.0
7.	I21A	.6183	.4867	262.0
8.	I23A	.6145	.4876	262.0
9.	I25A	.5534	.4981	262.0
10.	I30A	.4771	.5004	262.0
11.	I31A	.3359	.4732	262.0
12.	I35A	.3168	.4661	262.0

ANEXO O

Análisis de varianza de las alternativas de los ítems de la versión corta del TAMPR

----- O N E W A Y -----

Variable PTBRUT

By Variable I10B

Analysis of Variance

Source	D.F.	Sum of Squares	Mean Squares	F Ratio	F Prob.
Between Groups	7	341.1875	48.7411	10.3060	.0000
Within Groups	114	539.1486	4.7294		
Total	121	880.3361			

Group	Count	Standard Mean	Standard Deviation	Error	95 Pct Conf Int for Mean
Grp 1	9	5.1111	2.9345	.9782	2.8555 TO 7.3667
Grp 2	13	6.6154	1.9807	.5493	5.4185 TO 7.8123
Grp 3	4	9.0000	1.4142	.7071	6.7497 TO 11.2503
Grp 4	6	6.3333	2.5033	1.0220	3.7063 TO 8.9604
Grp 5	56	9.6071	2.0510	.2741	9.0579 TO 10.1564
Grp 6	13	6.1538	2.5115	.6936	4.6362 TO 7.6715
Grp 7	5	6.2000	1.0954	.4899	4.8398 TO 7.5602
Grp 8	16	7.0000	2.2211	.5553	5.8165 TO 8.1835
Total	122	7.9262	2.6973	.2442	7.4428 TO 8.4097

Levene Test for Homogeneity of Variances

Statistic	df1	df2	2-tail Sig.
1.4377	7	114	.197

Multiple Range Tests: Tukey-HSD test with significance level .050

The difference between two means is significant if
 $MEAN(J) - MEAN(I) \geq 1.5378 * RANGE * \sqrt{1/N(I) + 1/N(J)}$
 with the following value(s) for RANGE: 4.37

----- O N E W A Y -----

Variable PTBRUT

By Variable I11B

Analysis of Variance

Source	D.F.	Sum of Squares	Mean Squares	F Ratio	F Prob.
Between Groups	7	257.0196	36.7171	6.7153	.0000
Within Groups	114	623.3164	5.4677		
Total	121	880.3361			

Group	Count	Standard Mean	Standard Deviation	Error	95 Pct Conf Int for Mean
Grp 1	1	2.0000			
Grp 2	4	8.0000	1.8257	.9129	5.0949 TO 10.9051

Grp 3	9	6.1111	2.4721	.8240	4.2109	TO	8.0113
Grp 4	60	9.3000	2.4167	.3120	8.6757	TO	9.9243
Grp 5	13	6.7692	1.5359	.4260	5.8411	TO	7.6974
Grp 6	12	6.2500	2.4168	.6977	4.7144	TO	7.7856
Grp 7	9	6.7778	2.8186	.9395	4.6112	TO	8.9443
Grp 8	14	6.8571	2.2138	.5917	5.5789	TO	8.1354
Total	122	7.9262	2.6973	.2442	7.4428	TO	8.4097

Levene Test for Homogeneity of Variances

Statistic	df1	df2	2-tail Sig.
.8262	7	114	.568

Variable PTBRUT
By Variable I11B

Multiple Range Tests: Tukey-HSD test with significance level .050

The difference between two means is significant if
 $MEAN(J) - MEAN(I) \geq 1.6534 * RANGE * \sqrt{1/N(I) + 1/N(J)}$
 with the following value(s) for RANGE: 4.37

----- ONEWAY -----

Variable PTBRUT
By Variable I12B

Analysis of Variance

Source	D.F.	Sum of Squares	Mean Squares	F Ratio	F Prob.
Between Groups	7	255.9702	36.5672	6.6766	.0000
Within Groups	114	624.3658	5.4769		
Total	121	880.3361			

Group	Count	Mean	Standard Deviation	Standard Error	95 Pct Conf Int for Mean
Grp 1	44	7.0227	2.3673	.3569	6.3030 TO 7.7425
Grp 2	3	6.6667	3.5119	2.0276	-2.0574 TO 15.3908
Grp 3	46	9.6522	2.1829	.3219	9.0039 TO 10.3004
Grp 4	8	6.3750	3.1139	1.1009	3.7717 TO 8.9783
Grp 5	5	8.4000	1.8166	.8124	6.1444 TO 10.6556
Grp 6	3	7.6667	1.1547	.6667	4.7982 TO 10.5351
Grp 7	11	5.6364	2.4606	.7419	3.9833 TO 7.2894
Grp 8	2	8.0000	.0000	.0000	8.0000 TO 8.0000
Total	122	7.9262	2.6973	.2442	7.4428 TO 8.4097

Levene Test for Homogeneity of Variances

Statistic	df1	df2	2-tail Sig.
1.3068	7	114	.253

Variable PTBRUT
By Variable I12B

Multiple Range Tests: Tukey-HSD test with significance level .050

The difference between two means is significant if
 $MEAN(J) - MEAN(I) \geq 1.6548 * RANGE * \sqrt{1/N(I) + 1/N(J)}$
 with the following value(s) for RANGE: 4.37

----- ONEWAY -----

Variable PTBRUT
 By Variable I1B

Analysis of Variance

Source	D.F.	Sum of Squares	Mean Squares	F Ratio	F Prob.
Between Groups	6	132.2726	22.0454	3.3891	.0041
Within Groups	115	748.0635	6.5049		
Total	121	880.3361			

Group	Count	Standard Mean	Standard Deviation	Error	95 Pct Conf Int for Mean
Grp 1	2	8.0000	2.8284	2.0000	-17.4124 TO 33.4124
Grp 2	4	3.2500	2.0616	1.0308	-.0303 TO 6.5303
Grp 3	3	5.3333	3.5119	2.0276	-3.3908 TO 14.0574
Grp 4	2	7.5000	.7071	.5000	1.1469 TO 13.8531
Grp 5	109	8.2110	2.5498	.2442	7.7269 TO 8.6951
Grp 6	1	4.0000			
Grp 8	1	8.0000			
Total	122	7.9262	2.6973	.2442	7.4428 TO 8.4097

Levene Test for Homogeneity of Variances

Statistic	df1	df2	2-tail Sig.
1.4096	6	115	.217

----- ONEWAY -----

Variable PTBRUT
 By Variable I1B

Multiple Range Tests: Tukey-HSD test with significance level .050

The difference between two means is significant if
 $MEAN(J) - MEAN(I) \geq 1.8035 * RANGE * \sqrt{1/N(I) + 1/N(J)}$
 with the following value(s) for RANGE: 4.24

----- ONEWAY -----

Variable PTBRUT
 By Variable I2B

Analysis of Variance

Source	D.F.	Sum of Squares	Mean Squares	F Ratio	F Prob.
Between Groups	7	295.3876	42.1982	8.2240	.0000
Within Groups	114	584.9485	5.1311		
Total	121	880.3361			

Group	Count	Standard Mean	Standard Deviation	Error	95 Pct Conf Int for Mean	
Grp 1	1	6.0000				
Grp 2	9	3.8889	1.9003	.6334	2.4282 TO	5.3496
Grp 3	2	4.0000	2.8284	2.0000	-21.4124 TO	29.4124
Grp 4	94	8.7021	2.1344	.2201	8.2650 TO	9.1393
Grp 5	1	4.0000				
Grp 6	1	4.0000				
Grp 7	4	7.0000	3.7417	1.8708	1.0463 TO	12.9537
Grp 8	10	6.4000	3.0258	.9568	4.2355 TO	8.5645
Total	122	7.9262	2.6973	.2442	7.4428 TO	8.4097

Levene Test for Homogeneity of Variances

Statistic	df1	df2	2-tail Sig.
1.8450	7	114	.085

Variable PTBRUT
By Variable I2B

Multiple Range Tests: Tukey-HSD test with significance level .050

The difference between two means is significant if
 $MEAN(J) - MEAN(I) \geq 1.6017 * RANGE * \sqrt{1/N(I) + 1/N(J)}$
 with the following value(s) for RANGE: 4.37

----- O N E W A Y -----

Variable PTBRUT
By Variable I3B

Analysis of Variance

Source	D.F.	Sum of Squares	Mean Squares	F Ratio	F Prob.
Between Groups	4	217.0125	54.2531	9.5694	.0000
Within Groups	117	663.3235	5.6694		
Total	121	880.3361			

Group	Count	Standard Mean	Standard Deviation	Error	95 Pct Conf Int for Mean	
Grp 1	102	8.5098	2.3620	.2339	8.0459 TO	8.9737
Grp 2	4	5.0000	2.8284	1.4142	.4994 TO	9.5006
Grp 3	8	4.5000	2.6186	.9258	2.3108 TO	6.6892
Grp 7	6	5.6667	2.3381	.9545	3.2130 TO	8.1203
Grp 8	2	4.5000	.7071	.5000	-1.8531 TO	10.8531
Total	122	7.9262	2.6973	.2442	7.4428 TO	8.4097

TOTAL 1.0000 12.0000

Levene Test for Homogeneity of Variances

Statistic	df1	df2	2-tail Sig.
.8826	4	117	.477

Variable PTBRUT
By Variable I3B

Multiple Range Tests: Tukey-HSD test with significance level .050

The difference between two means is significant if
 $MEAN(J) - MEAN(I) \geq 1.6837 * RANGE * \sqrt{1/N(I) + 1/N(J)}$
 with the following value(s) for RANGE: 3.92

----- O N E W A Y -----

Variable PTBRUT
By Variable I4B

Analysis of Variance

Source	D.F.	Sum of Squares	Mean Squares	F Ratio	F Prob.
Between Groups	7	214.6160	30.6594	5.2502	.0000
Within Groups	114	665.7201	5.8396		
Total	121	880.3361			

Group	Count	Standard Mean	Standard Deviation	Error	95 Pct Conf Int for Mean
Grp 1	2	6.0000	5.6569	4.0000	-44.8248 TO 56.8248
Grp 2	1	2.0000			
Grp 3	3	7.0000	2.6458	1.5275	.4275 TO 13.5725
Grp 4	6	4.1667	2.1370	.8724	1.9241 TO 6.4093
Grp 5	103	8.4563	2.3254	.2291	8.0018 TO 8.9108
Grp 6	3	5.3333	2.5166	1.4530	-.9183 TO 11.5850
Grp 7	1	4.0000			
Grp 8	3	5.3333	4.0415	2.3333	-4.7063 TO 15.3730

Total 122 7.9262 2.6973 .2442 7.4428 TO 8.4097

Levene Test for Homogeneity of Variances

Statistic	df1	df2	2-tail Sig.
1.9645	7	114	.066

Variable PTBRUT
By Variable I4B

Multiple Range Tests: Tukey-HSD test with significance level .050

The difference between two means is significant if
 $MEAN(J) - MEAN(I) \geq 1.7087 * RANGE * \sqrt{1/N(I) + 1/N(J)}$
 with the following value(s) for RANGE: 4.37

----- O N E W A Y -----

Variable PTBRUT
By Variable I5B

Analysis of Variance

Source	D.F.	Sum of Squares	Mean Squares	F Ratio	F Prob.
Between Groups	4	249.6138	62.4035	11.5759	.0000
Within Groups	117	630.7222	5.3908		
Total	121	880.3361			

Group	Count	Standard Mean	Standard Deviation	Error	95 Pct Conf Int for Mean
Grp 2	90	8.7222	2.4450	.2577	8.2101 TO 9.2343
Grp 4	6	5.0000	1.6733	.6831	3.2440 TO 6.7560
Grp 6	3	3.3333	.5774	.3333	1.8991 TO 4.7676
Grp 7	2	8.0000	1.4142	1.0000	-4.7062 TO 20.7062
Grp 8	21	6.0000	2.0248	.4419	5.0783 TO 6.9217
Total	122	7.9262	2.6973	.2442	7.4428 TO 8.4097

Levene Test for Homogeneity of Variances

Statistic	df1	df2	2-tail Sig.
2.1981	4	117	.073

Variable PTBRUT
By Variable I5B

Multiple Range Tests: Tukey-HSD test with significance level .050

The difference between two means is significant if
 $MEAN(J) - MEAN(I) \geq 1.6418 * RANGE * \sqrt{1/N(I) + 1/N(J)}$
 with the following value(s) for RANGE: 3.92

----- O N E W A Y -----

Variable PTBRUT
By Variable I6B

Analysis of Variance

Source	D.F.	Sum of Squares	Mean Squares	F Ratio	F Prob.
Between Groups	6	236.8268	39.4711	7.0538	.0000
Within Groups	115	643.5093	5.5957		
Total	121	880.3361			

Group	Count	Standard Mean	Standard Deviation	Error	95 Pct Conf Int for Mean	
Grp 1	22	6.5909	1.9678	.4195	5.7184 TO	7.4634
Grp 3	1	8.0000				
Grp 4	1	5.0000				
Grp 5	8	5.3750	2.7742	.9808	3.0557 TO	7.6943
Grp 6	2	2.5000	.7071	.5000	-3.8531 TO	8.8531
Grp 7	87	8.7126	2.4300	.2605	8.1947 TO	9.2305
Grp 8	1	3.0000				
Total	122	7.9262	2.6973	.2442	7.4428 TO	8.4097

Levene Test for Homogeneity of Variances

Statistic	df1	df2	2-tail Sig.
1.6402	6	115	.142

Variable PTBRUT
By Variable 16B

Multiple Range Tests: Tukey-HSD test with significance level .050

The difference between two means is significant if
 $MEAN(J) - MEAN(I) \geq 1.6727 * RANGE * \sqrt{1/N(I) + 1/N(J)}$
 with the following value(s) for RANGE: 4.24

----- O N E W A Y -----

Variable PTBRUT
By Variable 17B

Analysis of Variance

Source	D.F.	Sum of Squares	Mean Squares	F Ratio	F Prob.
Between Groups	7	355.0911	50.7273	11.0099	.0000
Within Groups	114	525.2449	4.6074		
Total	121	880.3361			
Group	Count	Standard Mean	Standard Deviation	Error	95 Pct Conf Int for Mean
Grp 1	15	4.6667	2.1931	.5662	3.4522 TO 5.8811
Grp 2	1	5.0000			
Grp 3	3	7.0000	1.0000	.5774	4.5158 TO 9.4842
Grp 4	25	6.7600	1.7388	.3478	6.0423 TO 7.4777
Grp 5	1	2.0000			
Grp 6	1	8.0000			
Grp 7	3	7.3333	2.3094	1.3333	1.5964 TO 13.0703
Grp 8	73	9.1781	2.2751	.2663	8.6473 TO 9.7089
Total	122	7.9262	2.6973	.2442	7.4428 TO 8.4097

Levene Test for Homogeneity of Variances

Statistic	df1	df2	2-tail Sig.
1.4831	7	114	.180

Variable PTBRUT
By Variable 17B

Multiple Range Tests: Tukey-HSD test with significance level .050

The difference between two means is significant if
 $MEAN(J) - MEAN(I) \geq 1.5178 * RANGE * \sqrt{1/N(I) + 1/N(J)}$
 with the following value(s) for RANGE: 4.37

----- O N E W A Y -----

Variable PTBRUT
 By Variable I8B

Analysis of Variance

Source	D.F.	Sum of Squares	Mean Squares	F Ratio	F Prob.
Between Groups	6	329.1771	54.8628	11.4472	.0000
Within Groups	115	551.1590	4.7927		
Total	121	880.3361			

Group	Count	Standard Mean	Standard Deviation	Error	95 Pct Conf Int for Mean	
Grp 1	3	5.3333	1.5275	.8819	1.5387 TO	9.1280
Grp 3	18	6.3333	2.2229	.5239	5.2279 TO	7.4387
Grp 4	2	3.5000	.7071	.5000	-2.8531 TO	9.8531
Grp 5	13	5.8462	2.2303	.6186	4.4984 TO	7.1939
Grp 6	70	9.3000	2.2024	.2632	8.7748 TO	9.8252
Grp 7	1	7.0000				
Grp 8	15	6.4000	2.1974	.5674	5.1831 TO	7.6169

Total 122 7.9262 2.6973 .2442 7.4428 TO 8.4097

Levene Test for Homogeneity of Variances

Statistic	df1	df2	2-tail Sig.
.6947	6	115	.654

Variable PTBRUT
 By Variable I8B

Multiple Range Tests: Tukey-HSD test with significance level .050

The difference between two means is significant if
 $MEAN(J) - MEAN(I) \geq 1.5480 * RANGE * \sqrt{1/N(I) + 1/N(J)}$
 with the following value(s) for RANGE: 4.24

----- O N E W A Y -----

Variable PTBRUT
By Variable I9B

Analysis of Variance

Source	D.F.	Sum of Squares	Mean Squares	F Ratio	F Prob.
Between Groups	7	281.7643	40.2520	7.6661	.0000
Within Groups	114	,598.5718	5.2506		
Total	121	880.3361			

Group	Count	Standard Mean	Standard Deviation	Error	95 Pct Conf Int for Mean
Grp 1	4	5.5000	1.9149	.9574	2.4531 TO 8.5469
Grp 2	15	6.3333	2.4689	.6375	4.9661 TO 7.7005
Grp 3	1	3.0000			
Grp 4	9	9.3333	1.6583	.5528	8.0586 TO 10.6080
Grp 5	5	6.6000	3.7815	1.6912	1.9047 TO 11.2953
Grp 6	3	6.3333	1.5275	.8819	2.5387 TO 10.1280
Grp 7	78	8.7564	2.2689	.2569	8.2449 TO 9.2680
Grp 8	7	4.0000	1.9149	.7237	2.2291 TO 5.7709
Total	122	7.9262	2.6973	.2442	7.4428 TO 8.4097

Levene Test for Homogeneity of Variances

Statistic	df1	df2	2-tail Sig.
1.7076	7	114	.114

Variable PTBRUT
By Variable I9B

Multiple Range Tests: Tukey-HSD test with significance level .050
The difference between two means is significant if
 $MEAN(J) - MEAN(I) \geq 1.6203 * RANGE * \sqrt{1/N(I) + 1/N(J)}$
with the following value(s) for RANGE: 4.37

ANEXO P

Estadísticos descriptivos de las calificaciones por carrera

Variable PSICOLOGÍA

Mean	7,765	S.E. Mean	,270
Std Dev	2,670	Kurtosis	-,055
S.E. Kurt	,483	Skewness	-,011
S.E. Skew	,244	Minimum	2,00
Maximum	15,00		
Valid observations -	98	Missing observations -	179

Variable CIENCIAS PEDAGÓGICAS

Mean	10,857	S.E. Mean	,829
Std Dev	2,193	Kurtosis	-1,366
S.E. Kurt	1,587	Skewness	,252
S.E. Skew	,794	Minimum	8,00
Maximum	14,00		
Valid observations -	7	Missing observations -	270

Variable ECONOMÍA

Mean	11,633	S.E. Mean	,342
Std Dev	2,395	Kurtosis	-,173
S.E. Kurt	,668	Skewness	,123
S.E. Skew	,340	Minimum	7,00
Maximum	17,00		
Valid observations -	49	Missing observations -	228

Variable RELACIONES INDUSTRIALES

Mean	13,162	S.E. Mean	,308
Std Dev	2,758	Kurtosis	-,076
S.E. Kurt	,532	Skewness	-,084
S.E. Skew	,269	Minimum	6,00
Maximum	20,00		
Valid observations -	80	Missing observations -	197

Variable DERECHO

Mean	15,846	S.E. Mean	,564
Std Dev	2,035	Kurtosis	1,769
S.E. Kurt	1,191	Skewness	-,807
S.E. Skew	,616	Minimum	11,00
Maximum	19,00		
Valid observations -	13	Missing observations -	264

ANEXO Q

Estadísticos descriptivos de las respuestas a los ítems en la versión larga del TAMPR

Number of valid observations (listwise) = 252.00

Variable I1

Mean	.899	Std Dev	.302
Variance	.091	Kurtosis	5.119
S.E. Kurt	.292	Skewness	-2.661
S.E. Skew	.146	Minimum	.00
Maximum	1.00		

Valid observations - 277 Missing observations - 0

Variable I2

Mean	.917	Std Dev	.276
Variance	.076	Kurtosis	7.286
S.E. Kurt	.292	Skewness	-3.039
S.E. Skew	.146	Minimum	.00
Maximum	1.00		

Valid observations - 277 Missing observations - 0

Variable I3

Mean	.942	Std Dev	.234
Variance	.055	Kurtosis	12.622
S.E. Kurt	.292	Skewness	-3.812
S.E. Skew	.146	Minimum	.00
Maximum	1.00		

Valid observations - 277 Missing observations - 0

Variable I4

Mean	.895	Std Dev	.307
Variance	.094	Kurtosis	4.742
S.E. Kurt	.292	Skewness	-2.590
S.E. Skew	.147	Minimum	.00
Maximum	1.00		

Valid observations - 276 Missing observations - 1

Variable I5

Mean	.848	Std Dev	.359
Variance	.129	Kurtosis	1.828
S.E. Kurt	.292	Skewness	-1.953
S.E. Skew	.146	Minimum	.00
Maximum	1.00		

Valid observations - 277 Missing observations - 0

Variable I6

Mean	.888	Std Dev	.316
Variance	.100	Kurtosis	4.158
S.E. Kurt	.292	Skewness	-2.475
S.E. Skew	.146	Minimum	.00
Maximum	1.00		

Valid observations -	277	Missing observations -	0
----------------------	-----	------------------------	---

Variable I7

Mean	.910	Std Dev	.287
Variance	.082	Kurtosis	6.314
S.E. Kurt	.292	Skewness	-2.876
S.E. Skew	.146	Minimum	.00
Maximum	1.00		

Valid observations -	277	Missing observations -	0
----------------------	-----	------------------------	---

Variable I8

Mean	.913	Std Dev	.282
Variance	.080	Kurtosis	6.738
S.E. Kurt	.292	Skewness	-2.948
S.E. Skew	.147	Minimum	.00
Maximum	1.00		

Valid observations -	276	Missing observations -	1
----------------------	-----	------------------------	---

Variable I9

Mean	.841	Std Dev	.366
Variance	.134	Kurtosis	1.533
S.E. Kurt	.292	Skewness	-1.877
S.E. Skew	.146	Minimum	.00
Maximum	1.00		

Valid observations -	277	Missing observations -	0
----------------------	-----	------------------------	---

Variable I10

Mean	.910	Std Dev	.287
Variance	.082	Kurtosis	6.314
S.E. Kurt	.292	Skewness	-2.876
S.E. Skew	.146	Minimum	.00
Maximum	1.00		

Valid observations -	277	Missing observations -	0
----------------------	-----	------------------------	---

Variable I11

Mean	.939	Std Dev	.240
Variance	.058	Kurtosis	11.589
S.E. Kurt	.292	Skewness	-3.675
S.E. Skew	.146	Minimum	.00
Maximum	1.00		

Valid observations -	277	Missing observations -	0
----------------------	-----	------------------------	---

Variable I12

Mean	.903	Std Dev	.297
Variance	.088	Kurtosis	5.487
S.E. Kurt	.292	Skewness	-2.729
S.E. Skew	.146	Minimum	.00
Maximum	1.00		

Valid observations -	277	Missing observations -	0
----------------------	-----	------------------------	---

Variable I13

Mean	.635	Std Dev	.482
Variance	.233	Kurtosis	-1.692
S.E. Kurt	.292	Skewness	-.566
S.E. Skew	.146	Minimum	.00
Maximum	1.00		

Valid observations -	277	Missing observations -	0
----------------------	-----	------------------------	---

Variable I14

Mean	.863	Std Dev	.345
Variance	.119	Kurtosis	2.515
S.E. Kurt	.292	Skewness	-2.121
S.E. Skew	.146	Minimum	.00
Maximum	1.00		

Valid observations -	277	Missing observations -	0
----------------------	-----	------------------------	---

Variable I15

Mean	.812	Std Dev	.391
Variance	.153	Kurtosis	.590
S.E. Kurt	.292	Skewness	-1.608
S.E. Skew	.146	Minimum	.00
Maximum	1.00		

Valid observations -	277	Missing observations -	0
----------------------	-----	------------------------	---

Variable I16

Mean	.827	Std Dev	.379
Variance	.144	Kurtosis	1.020
S.E. Kurt	.292	Skewness	-1.736
S.E. Skew	.146	Minimum	.00
Maximum	1.00		

Valid observations -	277	Missing observations -	0
----------------------	-----	------------------------	---

Variable I17

Mean	.783	Std Dev	.413
Variance	.170	Kurtosis	-.087
S.E. Kurt	.292	Skewness	-1.383
S.E. Skew	.146	Minimum	.00
Maximum	1.00		

Valid observations -	277	Missing observations -	0
----------------------	-----	------------------------	---

Variable I18

Mean	.678	Std Dev	.468
Variance	.219	Kurtosis	-1.427
S.E. Kurt	.292	Skewness	-.764
S.E. Skew	.147	Minimum	.00
Maximum	1.00		

Valid observations -	276	Missing observations -	1
----------------------	-----	------------------------	---

Variable I19

Mean	.808	Std Dev	.395
Variance	.156	Kurtosis	.475
S.E. Kurt	.292	Skewness	-1.572
S.E. Skew	.147	Minimum	.00
Maximum	1.00		

Valid observations -	276	Missing observations -	1
----------------------	-----	------------------------	---

Variable I20

Mean	.675	Std Dev	.469
Variance	.220	Kurtosis	-1.445
S.E. Kurt	.292	Skewness	-.752
S.E. Skew	.146	Minimum	.00
Maximum	1.00		

Valid observations -	277	Missing observations -	0
----------------------	-----	------------------------	---

Variable I21

Mean	.629	Std Dev	.484
Variance	.234	Kurtosis	-1.724
S.E. Kurt	.293	Skewness	-.537
S.E. Skew	.147	Minimum	.00
Maximum	1.00		

Valid observations - 275 Missing observations - 2

Variable I22

Mean	.570	Std Dev	.496
Variance	.246	Kurtosis	-1.932
S.E. Kurt	.292	Skewness	-.286
S.E. Skew	.146	Minimum	.00
Maximum	1.00		

Valid observations - 277 Missing observations - 0

Variable I23

Mean	.621	Std Dev	.486
Variance	.236	Kurtosis	-1.761
S.E. Kurt	.292	Skewness	-.501
S.E. Skew	.146	Minimum	.00
Maximum	1.00		

Valid observations - 277 Missing observations - 0

Variable I24

Mean	.529	Std Dev	.500
Variance	.250	Kurtosis	-2.001
S.E. Kurt	.293	Skewness	-.118
S.E. Skew	.147	Minimum	.00
Maximum	1.00		

Valid observations - 274 Missing observations - 3

Variable I25

Mean	.554	Std Dev	.498
Variance	.248	Kurtosis	-1.966
S.E. Kurt	.292	Skewness	-.220
S.E. Skew	.147	Minimum	.00
Maximum	1.00		

Valid observations - 276 Missing observations - 1

Variable I26

Mean	.473	Std Dev	.500
Variance	.250	Kurtosis	-2.003
S.E. Kurt	.293	Skewness	.110
S.E. Skew	.147	Minimum	.00
Maximum	1.00		

Valid observations - 275 Missing observations - 2

Variable I27

Mean	.364	Std Dev	.482
Variance	.232	Kurtosis	-1.689
S.E. Kurt	.294	Skewness	.569
S.E. Skew	.148	Minimum	.00
Maximum	1.00		

Valid observations - 272 Missing observations - 5

Variable I28

Mean	.379	Std Dev	.486
Variance	.236	Kurtosis	-1.760
S.E. Kurt	.294	Skewness	.503
S.E. Skew	.148	Minimum	.00
Maximum	1.00		

Valid observations - 272 Missing observations - 5

Variable I29

Mean	.270	Std Dev	.445
Variance	.198	Kurtosis	-.922
S.E. Kurt	.293	Skewness	1.041
S.E. Skew	.147	Minimum	.00
Maximum	1.00		

Valid observations - 274 Missing observations - 3

Variable I30

Mean	.471	Std Dev	.500
Variance	.250	Kurtosis	-2.001
S.E. Kurt	.293	Skewness	.118
S.E. Skew	.147	Minimum	.00
Maximum	1.00		

Valid observations - 274 Missing observations - 3

Variable I31

Mean	.333	Std Dev	.472
Variance	.223	Kurtosis	-1.506
S.E. Kurt	.294	Skewness	.711
S.E. Skew	.147	Minimum	.00
Maximum	1.00		

Valid observations - 273 Missing observations - 4

Variable I32

Mean	.372	Std Dev	.484
Variance	.235	Kurtosis	-1.730
S.E. Kurt	.293	Skewness	.531
S.E. Skew	.147	Minimum	.00
Maximum	1.00		

Valid observations - 274 Missing observations - 3

Variable I33

Mean	.478	Std Dev	.500
Variance	.250	Kurtosis	-2.007
S.E. Kurt	.297	Skewness	.090
S.E. Skew	.149	Minimum	.00
Maximum	1.00		

Valid observations - 268 Missing observations - 9

Variable I34

Mean	.351	Std Dev	.478
Variance	.229	Kurtosis	-1.615
S.E. Kurt	.295	Skewness	.630
S.E. Skew	.148	Minimum	.00
Maximum	1.00		

Valid observations - 271 Missing observations - 6

Variable I35

Mean	.311	Std Dev	.464
Variance	.215	Kurtosis	-1.337
S.E. Kurt	.295	Skewness	.821
S.E. Skew	.148	Minimum	.00
Maximum	1.00		

Valid observations - 270 Missing observations - 7

Variable I36

Mean	.094	Std Dev	.293
Variance	.086	Kurtosis	5.836
S.E. Kurt	.298	Skewness	2.791
S.E. Skew	.150	Minimum	.00
Maximum	1.00		