

315011
11

UNIVERSIDAD

SALESIANA

INCORPORADA A LA
UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO

ESCUELA DE PSICOLOGIA



ESTANDARIZACION DE LA PRUEBA DE DOMINOS PARA UNA
MUESTRA DE LA POBLACION MEXICANA

297861

TESIS PROFESIONAL

QUE PARA OBTENER EL TITULO DE:

LICENCIADO EN PSICOLOGIA

P R E S E N T A :

ARTURO ROA QUINTANAR

MEXICO, D.F.

2001



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

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

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

AGRADECIMIENTOS:

A MI PADRE, QUIEN ME ENSEÑÓ EL VALOR DE LA PERSEVERANCIA Y ME BRINDÓ SU APOYO INCONDICIONAL, SU CARIÑO Y SU CONFIANZA.

A MI MADRE, POR ENSEÑARME LA IMPORTANCIA DE HACER BIEN LAS COSAS Y NO RENDIRME EN EL INTENTO.

A BÁRBARA, EL AMOR DE MI VIDA, POR TODO EL CARIÑO Y APOYO SIN LÍMITES QUE FUE COMO UN ALICIENTE PARA LOGRAR ESTA META TAN IMPORTANTE EN MI VIDA.

A MIS AMIGOS, POR SU APOYO, SUS CONSEJOS Y CRÍTICAS SIEMPRE OPORTUNAS.

A LA DRA. OLGA SANTA MARÍA, POR TRANSMITIRME SU ESPÍRITU DE LUCHA Y DESEOS DE SUPERACIÓN CONTÍNUA.

AL PROFESOR AGUSTÍN SOLÍS SÁNCHEZ, POR SU ASESORÍA INCONDICIONAL, POR TRANSMITIRME SUS CONOCIMIENTOS Y POR DISFRUTAR ESTA TESIS TANTO COMO YO.

QUIERO DEDICAR LA PRESENTE TESIS PRIMERO QUE NADA A MIS PADRES, YA QUE SE QUE ESTE LOGRO LO VAN A DISFRUTAR TANTO COMO YO.

TAMBIÉN SE LA DEDICO A LA PROFESORA MA. ELIZABETH RAMÍREZ LÓPEZ, PORQUE SIN DUDA ESTE TRABAJO LLEVA PARTE DE ELLA, YA QUE VIVIÓ, GOZÓ Y SUFRIÓ CON MIGO TODO EL PROCESO DE SU REALIZACIÓN.

“EL PRESENTE TRABAJO FUE REALIZADO CON LA
SUPERVISIÓN Y ASESORÍA ACADÉMICA DE LA
LIC. MARÍA ELIZABETH RAMÍREZ LÓPEZ”

INDICE

AGRADECIMIENTOS	i
DEDICATORIA	ii
REFERENCIA DE SUPERVISIÓN	iii
ÍNDICE GENERAL	iv
LISTA DE TABLAS	vi
LISTA DE GRÁFICAS	vii
INTRODUCCIÓN	1
CAPITULO I	
RESEÑA DE LA MEDICIÓN EN PSICOLOGÍA	5
1.1 Introducción	5
1.2 Perspectiva histórica	8
1.2.1 Los test y la evaluación psicológica a principios del siglo XX. La importancia de la guerra.	13
1.3 Diferencias individuales: medición y significado.	17
1.3.1 La Medición en Psicología: Diferencias Individuales.	21
1.4 Clasificación de los Tests.	25
1.4.1 Diferencia entre test y prueba.	26
1.4.2 Tipos de test.	28
1.5 El Test de Dominós de Anstey.	33
1.5.1 Lo que mide el test de Dominós. Características	34
1.5.2 Puntuación y Calificación	37
1.5.3 Estandarización	39
1.5.4 Valores estadísticos	40
1.6 Factor general	42
1.6.1 La inteligencia y el factor G	44
1.7 Conceptos y Principios de La Medición y Estadística	49
1.7.1 Las variables en estadística	53
1.7.2 Centiles o Porcentiles	56
1.7.3 Medidas de tendencia central o promedios	57
1.7.4 Variabilidad	59
1.7.5 Puntuaciones estándar y la curva normal	63
1.7.6 Comprobación de hipótesis	66

CAPÍTULO II

METODOLOGÍA	70
2.1 Planteamiento del Problema	70
2.2 Hipótesis	71
2.2.1 Hipótesis de Trabajo	71
2.2.2 Hipótesis Nula	72
2.3 Sujetos	72
2.4 Escenario	73
2.5 Material	74
2.6 Instrumento	74
2.7 Diseño	75
2.8 Procedimiento	75

CAPÍTULO III

RESULTADOS Y SU ANÁLISIS	81
3.1 Resultados	81
3.2 Análisis de resultados	97

CAPÍTULO IV

CONCLUSIONES Y RECOMENDACIONES	110
4.1 Conclusiones	110
4.2 Recomendaciones	114
REFERENCIAS BIBLIOGRÁFICAS	118
ANEXO A	121
ANEXO B	122
ANEXO C	123

LISTA DE TABLAS

Número	Descripción	Página
1	Recorrido histórico de la psicometría desde 1910	15
2	Valores estadísticos del Dominós para la población de Uruguay.	40
3	Normas percentilares para el Dominós, versión 1955.	41
3.1	Distribución de los sujetos de la muestra por edad y sexo	82
3.2	Distribución de frecuencias de las puntuaciones	83
3.3	Distribución de los sujetos según el percentil y por rangos de edad	84
3.4	Distribución de los sujetos según un rango percentil más amplio y por rangos de edad	85
3.5	Distribución de los sujetos según el percentil y por edades	86
3.6	Distribución de los sujetos según un rango percentil más amplio y por edades	87
3.7	Distribución de los sujetos según el percentil, edad y sexo	88
3.8	Distribución de los sujetos según un rango percentil más amplio, edad y sexo	89
3.9	Puntuaciones de distribución y variabilidad	90-95
3.10	Comprobación de hipótesis. Prueba z por rangos entre Uruguay y México	96
3.10.1	Prueba z para la población mexicana, hombres vs. mujeres	97

LISTA DE GRÁFICAS

Número	Descripción	Página
1	Curva de distribución de los diferentes grupos de edad.	103
2	Distribución de frecuencias de la población general.	104

INTRODUCCION

La historia de la medición o evaluación de las características psicológicas de los individuos se remonta hacia el año 1000 a.C. en China. Analizando la historia, desde los Chinos hasta la época actual, se puede vislumbrar un interés constante de la humanidad de conocer a fondo las diferencias individuales entre los sujetos y lograr una clasificación de ellos. Esta constante búsqueda ha llevado a los estudiosos e investigadores a crear instrumentos de medición que faciliten dicha tarea, los cuales en psicología se llaman "Test psicológicos o psicométricos".

En su afán por medir diferencias, se han creado múltiples y muy variados tipos de tests. Los tests de habilidades, de aptitudes, de personalidad, los de inteligencia, etc. Todos ellos con características independientes y sobre todo con una base teórica que los sustenta.

Pero además de la teoría, existe una parte fundamental en todas las pruebas que permite su uso cotidiano sin duda de que lo que se pretende medir es correcto, esto es, su contextualización dentro de una población determinada.

Lo anterior hace referencia a que cada rincón del planeta, hablando de su población, tiene sus características típicas; en cada país y en cada región, cambian considerablemente los parámetros de lo que es normal y lo que no es; hablando con términos estadísticos, la media de las puntuaciones puede variar de una muestra a otra.

La consecuencia lógica al medir a una población con los parámetros de otra sería simplemente que las mediciones o su clasificación dentro de una escala sería errónea.

A pesar de que esto es y ha sido de conocimiento común a lo largo del desarrollo de la psicometría en México y en todo el mundo, se sigue cometiendo el error de aplicar pruebas y clasificar a los sujetos con tests que no han sido diseñados específicamente para esa población.

Eso mismo ha pasado en México desde hace más de 40 años con la prueba de Dominós, cuando el profesor Edgard Anstey diseñó la prueba para medir la inteligencia de los sujetos británicos y tiempo después se comenzó a aplicar a la población Mexicana sin hacer las precisiones estadísticas necesarias, es decir, sin adecuar las tablas de clasificación o los baremos para la población que estaba siendo evaluada; ya que desde el inicio de su aplicación en México, se han

utilizado para la clasificación de los sujetos evaluados, los baremos hechos para la población de Montevideo, Uruguay.

Es importante señalar que, estadísticamente es un error tratar de hacer comparaciones y establecer aseveraciones entre dos grupos cuando los promedios de las puntuaciones que se comparan son diferentes.

Lo anterior significa que ha sido un error clasificar a los sujetos Mexicanos según su inteligencia utilizando los baremos que fueron diseñados para la población de Uruguay por el profesor W. L. Risso en el año de 1955.

Debido a eso, la presente investigación pretende demostrar, a través del análisis estadístico de la media y la prueba z para la comprobación de hipótesis, que la medida promedio de inteligencia de los sujetos Mexicanos es diferente que la de los sujetos Uruguayos, y por lo tanto, ha sido mal clasificado durante todos estos años. Por eso, se considera también de importancia que, a través de los resultados obtenidos, se presenten las tablas o baremos apropiados para la población Mexicana.

Para tal efecto, a lo largo de la presente investigación se hará un recorrido en el tiempo para profundizar en la historia de la medición, pasando por uno de los objetivos de ésta, que es precisamente el establecimiento de las diferencias

individuales. Una vez conocido lo anterior, se pasará a hacer una revisión de la clasificación de los tests y sus características, esto con el fin de poder ubicar en forma general el test de Dominós de Anstey para después pasar a conocer algunas de las características más importantes de la prueba, lo que mide, así como sus valores estadísticos y de estandarización . Debido a que la prueba se basa en la teoría del factor general de inteligencia, se hará un acercamiento a dicha teoría y su relación con la inteligencia. Después, y siendo tal vez una de las partes fundamentales en la presente investigación, se abordarán los conceptos y principios de la medición y estadística, a fin de tener una referencia clara y precisa con la cual enmarcar el presente trabajo; sin esta parte del trabajo las aseveraciones y sustentaciones tal vez no tendrían ninguna validez, en especial para un profesionalista alejado de la estadística como lo es un psicólogo.

Si bien la forma de abordar el problema parece un tanto alejado de la práctica psicológica, definitivamente no es así, ya que gran parte del trabajo que se hace con los individuos esta basado en el lugar que éste ocupa dentro de la clasificación que se le de en diversos ámbitos, incluyendo el intelectual, y siempre en el rango que va de la "normalidad" a la "anormalidad". Es por eso que se debe contar con instrumentos que estén "calibrados" para la población en la que pretenden ser utilizados, siendo ese precisamente el objeto de la presente investigación, adecuar la prueba de Dominós para ser usada en la población mexicana.

CAPITULO I

RESEÑA DE LA MEDICIÓN EN PSICOLOGÍA

1.1 Introducción.

Desde hace muchos años, la humanidad ha tenido la necesidad de establecer parámetros específicos para la medición de las características del ser humano: el más alto, el más fuerte, el más rápido, el más inteligente, etc. parte de esta necesidad se ha visto reflejada en los esfuerzos para la creación de instrumentos de medición, los cuales en psicología se denominan “ test psicométricos ”.

Es evidente que las personas son diferentes unas de otras, pero ¿cómo son esas diferencias?, ¿de qué clase?, ¿cuáles son sus implicaciones?. Algunas diferencias entre las personas pueden ser muy patentes, incluso para los

observadores más casuales y en muchas ocasiones se requiere de algún tipo de instrumento de medición para poder establecer comparaciones.

Algunas personas son obesas y otras delgadas, unas corren los 100 metros en menos de 10 segundos mientras a otras les tomaría 15 segundos o más. Pero las personas no solo difieren en aspecto y habilidades físicas, sino que varían también en cuanto a personalidad y habilidades psicológicas. Para muchas personas las matemáticas les pueden resultar sencillas, mientras que otros tienen grandes dificultades para efectuar los cálculos más sencillos.

La medición en psicología a través de los test, se ha utilizado con múltiples y muy variados objetivos, siendo la medición de la inteligencia uno de los temas de mayor interés desde hace mucho tiempo y la cual ha sido estudiada en diversas partes del mundo.

Gran parte de los países industrializados de la época actual, evalúan de forma cotidiana la inteligencia de sus habitantes, sin embargo, muy pocos de estos países cuentan con instrumentos adecuados o con los parámetros justos para medir y ubicar en una clasificación a los sujetos que examinan.

Debido a esto, el estudio de las diferencias individuales ha sido un punto de continuo interés para la psicología. Pero no solo eso, el interés no se centra en las diferencias que existen entre un grupo de individuos de la misma región

geográfica o la misma raza, sino que, las investigaciones se han interesado en las diferencias existentes entre poblaciones que se encuentran a miles de kilómetros de distancia.

Cada país y cada región tiene sus características propias, en cada parte del planeta cambian considerablemente los parámetros de normalidad y anormalidad, y esto mismo ocurre con la inteligencia. Es evidente que pueden existir diferencias entre la inteligencia de los sujetos de una misma región, pero las diferencias se podrían hacer más evidentes entre regiones o entre países.

Para la investigación en psicología, uno de los mayores problemas ha sido el descubrir la naturaleza y amplitud de esas diferencias entre los individuos en forma precisa. Aún cuando en una plática cotidiana se puede decir que una persona es "lista o tonta", esas distinciones tan "a groso modo" no han sido suficientes. Es por eso que a lo largo de la historia ha sido necesario desarrollar instrumentos que ayuden a proporcionar datos más precisos. A pesar de que existen muchos métodos de medición y evaluación tales como las observaciones, las escalas de estimación, las muestras de trabajo y las medidas de ejecución, se considera que los Test Psicológicos, representan el mejor método de evaluación desarrollado hasta el momento.

De igual forma, cada test ofrece información muy valiosa sobre las características de los sujetos en una región o en un país determinado, sin embargo,

es muy posible que la información obtenida se distorsione considerablemente si la prueba es aplicada en otro país a sujetos diferentes.

En el presente capítulo, se hará un recorrido por la historia de la medición y su desarrollo, hasta la creación de los test psicológicos; la inteligencia y los test que la miden; el factor general de la inteligencia, el test de Dominós y también los métodos estadísticos involucrados en la presente investigación.

1.2 Perspectiva Histórica.

“Solo la historia de la ciencia puede clarificar en forma adecuada el significado de la ciencia contemporánea. Solo ella puede integrar el efímero presente a la marcha perenne hacia el enronquecimiento del género humano.”

(Gusdorf, G., en Sattler, 1993)

La historia muestra que las concepciones de la psicología, la psicometría, así como, en particular de la inteligencia han ido evolucionando, dejando ver un progreso general desde una primera etapa en la que no existía una definición, ni un método para probarlas, hasta el momento en que se da un auge a los instrumentos de medición estandarizados.

Así conociendo los orígenes de la medición en psicología, es posible vislumbrar ese desarrollo y comprender más a fondo las inquietudes de esos primeros investigadores, cuyo legado es la base de todo el conocimiento actual.

Según Aiken (1995), los precursores de la medición de las diferencias individuales fueron Platón y Aristóteles, los cuales escribieron acerca de éstas precediendo a los chinos.

Tratando de encontrar los primeros indicios de estas disciplinas, es necesario remontarse hasta los años "1000 a. C. , a China ... cuando el emperador decide evaluar el grado de competencia de sus oficiales..., los puestos eran cubiertos por personas con buenas puntuaciones en exámenes que cubrían en campos como la música, equitación, leyes civiles, escritura ... y conocimiento de ceremonias" (García, 1993). Más tarde, durante el siglo XIX, en Inglaterra y Francia "se guiaron por el antiguo sistema chino a fin de diseñar los exámenes para el servicio militar" (Aiken,1995).

Sin embargo en la Edad Media, en Europa, el interés por las diferencias individuales era casi inexistente, y la estructura social al igual que las actividades a las que una persona estaba sujeta, determinaba la clase social a la cual pertenecía.

Ya en el siglo XVI, Europa es convulsionada por los grandes cambios, en donde se ve transformada y constituida como capitalista y menos doctriniana, creyendo en la idea de que las personas eran únicas y estaban bien habilitadas para hacer valorar sus capacidades y mejorar su destino. Esta concepción se consolida con el renacimiento, restableciendo el individualismo.

Iriarte (1948), dice que en 1529, nace en España Juan Huarte de San Juan, quien escribiera un libro llamado: *Examen de ingenios para las ciencias (Examen de los Ingenios)*, donde se muestra las diferencias de habilidades entre los hombres y el género de letras que a cada uno responde en particular. La base de su teoría es que la diferente capacidad de los hombres proviene de su naturaleza.

Este es un ejemplo auténtico de los antecedentes de la medida de las diferencias individuales y grupales, sobre todo de la Psicología científica, ya que es posible deducir que de este libro parte la moderna Psicología Diferencial.

Sin embargo, el hombre siguió cuestionándose sobre los fenómenos que tenía al alcance, en la búsqueda constante de leyes y de una comprobación científica de las mismas, siendo difícil ésta tarea cuando las medidas de los fenómenos naturales no son confiables.

En el siglo XVI “los jesuitas comienzan a utilizar test para la evaluación y clasificación de sus estudiantes” (Iriarte, 1948), sin embargo aún no eran

considerados como test en el sentido de que no llenaban los requerimientos actuales para estas pruebas, como son la validez y confiabilidad.

Es aquí cuando surge el interés de otras disciplinas como la Física, elaborando instrumentos más eficaces en la medición, sobre todo por la inmensa contribución de Charles Darwin (Ville, 1990) y por el surgimiento de la psicología científica, que el estudio de las Diferencias Individuales aumenta considerablemente durante la segunda mitad del siglo XIX.

En Inglaterra, a finales del siglo XIX, F. Galton (en Anastasi, 1970), instala su laboratorio: Antropometric Laboratory; lo cual constituye el inicio del estudio científico de las diferencias individuales. Es aquí donde nacen los test sensoriales y motores.

Sin embargo, no es sino hasta 1890 que "aparece por primera vez la palabra test, aplicada al contexto de la medida en Psicología, en un artículo de Catell, 'Mental test and Measurement' " (García, 1993).

Según Pichot (1991), Catell crea este testimonio para designar una serie de pruebas psicológicas utilizadas en el examen de las diferencias individuales de estudiantes universitarios. Dichas pruebas estaban encaminadas a medir funciones sensoriomotrices elementales.

Poco tiempo después, analizando un poco la historia, García (1993), establece cuatro puntos importantes en el desarrollo histórico de la utilización de las matemáticas en Psicología:

1.- De 1870 a 1930, los trabajos matemáticos son desarrollados por experimentalistas con alguna formación en física o ingeniería.

2.- Desde 1930 hasta 1940.

a) Gran auge de la psicometría, con trabajos más extensos y elaborados.

b) Enfrentamiento entre la Psicología Matemática y Psicología Experimental.

3.- De 1940 a 1960, sobre todo durante y después de la Segunda Guerra Mundial, se restablecen los trabajos de la psicología matemática y la experimental. Se desarrolla la Teoría de la información y Teoría de Juegos.

4.- A partir de 1960 se anexan dos nuevas características:

a) El uso masivo de los ordenadores y de la informática por psicólogos.

b) La utilización de modelos matemáticos cada vez más sofisticados.

Como se puede observar, las matemáticas y sobre todo la estadística han contribuido enormemente en el desarrollo de la evaluación en psicología, ya que es ésta la que se encarga de darle objetividad a los datos que se recaban con los ítems de cada test.

La medición y evaluación a lo largo de la historia ha estado rodeada de significativas modificaciones tanto en su objetivo como en los métodos utilizados; dirigiéndose en un principio al estudio de las características físicas y psicológicas como tales, pasando posteriormente a evaluar ciertas interacciones de los individuos con su medio ambiente.

1.2.1 Los test y la evaluación psicológica a principios del siglo XX. La importancia de la guerra.

“... el profano inteligente tenía probablemente poca confianza en el empleo o valor del test ... Después de la guerra ese mismo profano creyó que los psicólogos habían establecido un método simple y relativamente perfecto de la medida de la inteligencia” (Pichot, 1991).

Para 1904, el psicólogo Alfred Binet con el Dr. Simon, desarrollan un procedimiento para identificar a los niños que se creía que no podrían obtener suficientes beneficios de las clases escolares normales, (Aiken, 1995; Pichot, 1991; Sattler, 1993). Así elaboraron una prueba que se administraba en forma

individual, la cual consistía de 30 pruebas con orden creciente de dificultad. En 1905 publicaron en *L' Année Psychologique*, un artículo titulado: "Nuevos métodos para el diagnóstico del nivel intelectual de los anormales" (Pichot, 1991). Siendo este el primer test mental práctico.

No fue sino hasta 1908, que se realizó una revisión de la prueba, donde se introduce el concepto de "*edad mental*" como una forma de cuantificar el desempeño general de la persona que presenta la prueba" (Corbella, 1994).

La mayoría de los autores concuerdan en que la primera y segunda guerra mundial fueron los más importantes propulsores de la psicometría, de la investigación y del perfeccionamiento de los instrumentos de medición. Hasta la primera guerra mundial, estos nuevos instrumentos de medición quedaron limitados a los test de inteligencia y de aptitudes, aplicados a la educación y en un segundo término a la orientación profesional.

Ya que en 1917, los E.U. tuvieron que reclutar a un ejército rápidamente y elegir su puesto, "...esta aplicación masiva tuvo profunda resonancia; Yakes, Brisges y Harwik, establecieron una batería de test ... apoyado en la escala Binet-Simon ... consolidando el Army test " (Pichot, 1991). Este test de aplicación colectiva fue elaborado en 2 versiones: los test Alfa y los Beta. El primero de ellos de aplicación general y el segundo para los analfabetas.

Es así como se gesta el principio de toda una generación de test de lápiz y papel de aplicación colectiva y de test manipulativos.

Entre los desarrollos alrededor de los test "... se encuentran los mejoramientos en la metodología estadística y los avances tecnológicos en la preparación y calificación de los test, así como el análisis de resultados" (Aiken,1995).

Es imposible concluir un recuento de tal magnitud en el desarrollo de la Psicometría, sin mencionar los más recientes logros alcanzados, para lo que Fernandez-Ballesteros (1993) y Martínez (1995) * , hacen un largo recorrido desde 1910, proporcionando la fecha y el suceso:

Tabla 1

1910	Kent y Rosanoff establecen un incipiente valoración del "Test de asociación de palabras" de Jung.
1912	Rossolino publica en método de perfiles a través del cual cuantificar los aspectos diferenciales entre componentes innatos y adquiridos.
1914	Woodworth construye el primer test colectivo de medida del ajuste personal, el <i>Personal Data Sheet</i> , que se aplicará a miles de soldados con fines de selección durante la Primera Guerra Mundial.
1915	Aparece el "Test de Laberintos" de Porteus, con el que se pretende la evaluación de las capacidades mentales libre de influencias culturales.
1916	Primera adaptación norteamericana de las escalas Binet que adopta el nombre de "Escala Stanford-Binet para la medición de la inteligencia"
1918	Se emprende la evaluación de la inteligencia a miles de soldados aliados en la Primera Guerra Mundial mediante los test <i>Army Alpha</i> y <i>Army Beta</i> .
1919	Gesell edita la primera "Escala de Desarrollo" para la evaluación de las primeras edades de la vida.
1921	Rorschach publica su <i>Psychodiagnostik</i> , presentando su técnica de evaluación de la personalidad mediante la interpretación de manchas de tinta.
1926	Hartshorne y May realizan la primera crítica de los rasgos y test de personalidad a través de sus estudios sobre la honestidad. - Goodenough edita el "Test de dibujo de la figura humana", el cual supone el primer antecedente de las técnicas de dibujo en la Evaluación Psicológica.

* Los datos de la tabla de Martínez, 1995 se muestran con un asterisco al inicio del párrafo

- 1934 Jacob L. Moreno publica su obra *Who Shall Survive*, en la que sienta las bases de los tests sociométricos y situacionales.
- 1935 Von Lennep publica la primera técnica temática proyectiva, su "Test de cuatro cuadros".
- 1935 Vygotsky formula el concepto de "zona de desarrollo próximo", antecedente de la importancia de la evaluación de estrategias cognitivas frente a productos intelectuales.
- *1935 Primera máquina para la calificación de pruebas
- 1936 Se publica la primera escala de medida del desarrollo social, la *Vineland Social Maturity Scale* de Doll.
- *1936 Primer volumen de *Psichometrika*.
- *1937 Revisión de la escala Stanford-Binet.
- 1938 Murray presenta, junto a su teoría de la personalidad (elaborada con Morgan), la técnica proyectiva "Test de Apercepción Temática" (TAT).
- 1939 David Wechsler publica la primera edición de su "Medida de Inteligencia del Adulto"
- 1942 Hathaway y McKinley presentan el primer inventario clínico de evaluación de entidades nosológicas psiquiátricas, el Minnesota Multiphasic Personality Inventory.
- **1944 Edgard Anstey publica en Inglaterra la primera versión de su "Test de Dominó", a partir de los mismos principios que las matrices construidas por Raven (Fernandez-Ballesteros, 1993).¹
- 1947 Halstead ... y Luria... sientan las bases de la evaluación neurológica.
- *1949 Se publicó el Wechsler para niños
- 1950 Shapiro presenta... el trabajo *Un enfoque experimental al diagnóstico mediante test psicológicos...*
- 1957 Cronbach y Gleser publican *Psychological Test and Personnel Decision*, donde se propone una teoría formalizada a aplicar en el proceso de evaluación.
- *1960 Se edita la forma L-M de la prueba Stanford-Binet
- 1963 Staats publica su *Complex Human Behavior*, donde se sientan las bases para su evaluación conductual integradora.
- 1965 Kanfer y Saslow, con la publicación de su trabajo *Behavioral Analysis*, desarrollan un nuevo modelo de evaluación conductual.
- 1967 Anastasi, en su trabajo *Psychology, Psychologists and Psychological Testing*, denuncia la enorme separación entre la Evaluación psicológica y la Psicología.
- *1970 Aumenta substancialmente el uso de las computadoras.
- 1972 Feuerstein edita su *The Dynamic Assessment of Retarded Performers*, que supone el inicio de un modelo cognitivo en la evaluación del funcionamiento intelectual.
- 1974 Moos publica *The social Climate Scales: An Overview*, que representa un consolidado esfuerzo en la evaluación del ambiente.
- *1975 Incremento en las técnicas de evaluación conductual.
- *1981 Revisión de la escala Wechsler para adultos.
- *1987 Revisión del inventario Psicológico de California
- *1987 Se publica el DSM-III-R
- *1989 Se publicó el MMPI-II
- *1990 Escala para niños III de Wechsler
- *1992 Se publicó la 11ª edición del *The Mental Measurements Yearbook*.
- ** 2001 El reto continua...

¹ **Inscrito a la tabla original por el autor para la ubicación histórica de la aparición del test de Dominó

Parece evidente que la evaluación de las características del ser humano y por supuesto los test psicológicos, sufrieron un impulso enorme a partir de la Primera Guerra Mundial, claro está debido a la marcada urgencia de la situación; sin embargo, parece aún más patente que la evaluación psicológica era y es, un paso inminente para el desarrollo no sólo de la psicología como materia, sino de la sociedad en general, ya que con el paso de los años, se ha extendido el propio objeto de la evaluación, " ... pasando del énfasis en el análisis de las características individuales a la evaluación de interacciones persona-ambiente y de ahí, a la evaluación del propio ambiente y a la valoración de intervenciones (véanse Fernández-Ballesteros, 1980, 1983; Pelechano, 1988)." (Fernández-Ballesteros, 1993).

1.3 Diferencias individuales: Medición y significado.

"Cualquier cosa que existe, existe en cantidad (Thorndike, 1918)

Cualquier cosa que existe en cantidad, puede medirse (Mc Call, 1939)

(Aiken, 1995)

El objetivo primordial del descubrimiento científico y del desarrollo tecnológico consiste en producir nuevo conocimiento, el cual puede ser difundido a través de informes en revistas científicas, libros, etc.

El estudio de las diferencias individuales ha tenido lugar desde que el hombre existe sobre la tierra, sin embargo, su estudio objetivo tuvo lugar desde el momento en que se incluyó la medición en las observaciones. “Toda medición está relacionada con el ordenamiento de una clase definida de objetos con respecto a algún atributo o atributos” (Sternberg, 1987).

En la mayoría de las investigaciones existe un divorcio entre el quehacer científico y la realidad social, dando como resultado la simple descripción de los problemas sin llegar, en la gran mayoría de los casos, a ofrecer sugerencias aplicables, ni permitir análisis cuyos resultados orienten a la actuación de los individuos.

Es aquí donde radica la importancia de la medición, ya que, más que un simple método que nos pone al alcance ciertos datos, es necesario orientarla y enfocarla dentro de los procesos que se estudian, no como fenómenos aislados, sino que hay que contextualizarlos y aplicarlos a la población en cuestión, para así hacerla positiva y, sobre todo, funcional.

“La medición en psicología, como en cualquier otra ciencia, se inicia con la identificación de elementos del mundo real con los elementos o construcciones de un sistema lógico abstracto, al cual llamamos modelo” (Deledalle, 1995).

Las construcciones de la teoría psicológica están expresadas por medio de palabras, no obstante, tienden a traducirse en significados connotativos. Los modelos matemáticos difieren de los modelos puramente verbales en que los primeros emplean sistemas matemáticos exactos, generalmente de alto orden, mediante los cuales las construcciones elementales pueden manipularse a fin de facilitar las deducciones del modelo.

“La profundización en el uso de estas técnicas estadísticas desembocó en el método de correlación, mediante el cual se hace posible la representación gráfica y la formulación cuantitativa del grado de relación entre dos variables” (Rojas Soriano,1995).

Así mismo, la ciencia psicológica tiene como base el método científico y la experimentación, y parte de éste le permite llegar a conclusiones y, por tanto, establecer leyes acerca del comportamiento, lo cual también le permite abordar las investigaciones de la religión y las creencias del ser humano, sus motivaciones, y demás rasgos que lo conforman como un todo.

Cuando se habla de los instrumentos de medición en psicología, se hace referencia a aquellos procedimientos u operaciones que permiten llegar a obtener, objetiva y certeramente, información acerca de la expresión de los fenómenos que se suceden en esa unidad bio-psico-social, que es la conducta humana.

“Estos instrumentos, contruidos para medir aspectos específicos de la conducta humana se conocen con el nombre de PRUEBAS PSICOLÓGICAS” (Arnau, 1992).

La necesidad de medir cuantitativamente los atributos o características de un segmento de conducta, al ser observada por quienes se interesan por comprender y predecir el comportamiento de los seres humanos, surgió realmente durante la primera mitad de este siglo.

A Galton (1869, en Floria 1990), se le puede considerar el primer representante de la psicología realizada mediante los procedimientos científico-naturales en Gran Bretaña, aunque desde luego, de una psicología que apenas tenía puntos comunes con la representada por el alemán Wundt. Si la de éste obedecía a las exigencias de un modelo físico-fisiológico, la de Galton estaba inspirada en el modelo biológico implantado por Darwin.

La medición y evaluación psicológica es una subdisciplina de la Psicología Científica. Esto quiere decir que está basada en los hallazgos de las distintas áreas de la psicología. Especialmente de la psicología de las diferencias individuales, de la personalidad, cognitiva, del aprendizaje, de la psicofisiología y psiconeurología, de la psicopatología, de la Psicología social, así también de algunos desarrollos de la psicología aplicada y sus técnicas, que han sido posibles gracias a los avances metodológicos de la psicología experimental y matemática.

1.3.1 La medición en psicología: Diferencias individuales.

“Debe recordarse que muchas medidas menos perfectas han resultado muy útiles en psicología ...” (Zigler,1973 en Mc Conell,1990)

De un modo muy amplio se podría definir la ciencia experimental como el tipo de conocimiento que se adquiere mediante la utilización del método experimental. Siguiendo este criterio, la psicología experimental puede caracterizarse por la aplicación del método experimental al estudio de los procesos psicológicos.

“La psicología experimental, por tanto, no se define en función de lo que estudia ..., sino en función de la técnica de investigación que emplea...” (Arnau, 1992). Esto implica que cualquier área del saber psicológico puede ser explorada de manera experimental.

En el desarrollo de la psicología como ciencia, ha sido necesario realizar un sin número de estudios a fin de poder determinar si los seres humanos difieren entre si y en qué grado se presentan estas diferencias, lo que ha originado, desde hace aproximadamente un siglo “... el desarrollo de la instrumentación adecuada para evaluar cuantitativamente las diferencias y semejanzas entre los individuos” (Morales, 1993).

Según Floria (1990), se ha discutido extensamente, a quién puede atribuirse el nombre de “padre de la medición mental”; entre los que se hacen acreedores a tal designación se puede nombrar a Sir Francis Galton (1822-1911), James Mckeen Catell (1860-1944) y Alfred Binet (1857-1911). Pero lo verdaderamente importante es que los tres personajes han contribuido considerablemente al desarrollo de este campo.

La teoría de los rasgos de la época moderna es pariente muy cercana de la teoría Darwiniana de la Evolución. Una de las primeras personas en interesarse en la medición de los rasgos fue Sir Francis Galton (1869; Castellan,1987; Floria,1990; Arnau, 1992; Morales, 1993). Darwin había dicho que la evolución era un asunto de supervivencia del más apto, lo que Galton interpretó como la supervivencia del mejor dotado intelectualmente. Asumía que las diferencias intelectuales se debían principalmente a la herencia y no a la capacitación.

Entre los principales problemas que encontró estaban el identificar los rasgos psicológicos relacionados a la superioridad, así como la manera de medir estos rasgos.

Para cuando ambas dificultades fueros solucionadas a su satisfacción, había (Corbella, 1994; Mc Connell, 1990):

- Ideado la primera prueba de inteligencia.
- Efectuado el primer estudio científico de las diferencias individuales.
- Probado que las experiencias de la niñez tienen un efecto en el pensamiento adulto.
- Formulado el primer cuestionario psicológico.
- Hecho el primer estudio de gemelos.
- Establecido el uso de huellas dactilares como identificación personal.
- Utilizado por primera vez las correlaciones estadísticas en un estudio psicológico.
- Medido las capacidades sensoriales y motoras de cerca de 10,000 personas.

La contribución principal de Galton a la teoría de los rasgos fue, con seguridad, la de sus estudios detallados sobre la teoría de las diferencias individuales. Galton publicó su primer libro sobre la teoría de los rasgos en 1869 (Mc Connell, 1990); *Hereditary Genius*, fue un intento por probar que la inteligencia y la creatividad tendían a ser “de familia”. Lo que hizo Galton en realidad fue “... demostrar que los hijos son más similares a sus padres que a una muestra al azar de individuos no emparentados” (Floria, 1990).

Así, la teoría de los rasgos - según lo define Galton y la mayoría de los otros psicólogos - es el "... estudio parcial de las similitudes humanas" (Deladalle, 1995).

Bachelard (1989), dice al respecto que la teoría de los rasgos también se interesa en el estudio de las diferencias individuales, lo que son y de dónde provienen. Galton definió los rasgos como patrones medibles y consistentes del desempeño humano y del carácter.

Se inicia de esta manera, el estudio científico de las diferencias individuales y la evaluación de las pruebas mentales, apoyado en la instrumentación matemática.

Como instrumento metodológico de su psicología diferencial-individual , Galton (1969, en Paulus, 1990) estableció el TEST MENTAL en la medida en que este posibilita la constatación de las diferencias entre los componentes de un grupo.

El objetivo del test "... consiste en mostrar y medir las diferencias; estas se hacen patentes en los resultados, sin que sea necesario entrar en los procesos ..." (Paulus, 1990).

En el caso de la capacidad intelectual, lo que se busca es proporcionar ciertas formas de dar un orden a los individuos con respecto a ciertos atributos, ya sean capacidad o conducta.

El psicólogo debe esforzarse en precisar los métodos que le permitan desprender la significación integral de cada conducta, sin apelar a lo que el sujeto diga de sí mismo ni a una interpretación del observador, que traería consigo el riesgo de proyectar sus propios problemas.

Si bien es posible establecer una psicología científica que tome en cuenta la singularidad y la subjetividad en toda conducta humana, es imposible tener un conocimiento científico de un individuo determinado, no en teoría, sino como una imposibilidad en sí misma. Jamás podrán tenerse en cuenta todos los datos de la vida biológica, social y/o psicológica de un hombre, aún con la prueba más completa y complicada.

1.4 Clasificación de los test.

Como ha podido observarse durante el presente capítulo, la historia de los test psicológicos no es exactamente la más reciente, sino que data de muchos años atrás. A través de los años, y con el constante interés de medir las capacidades del ser humano, se han desarrollado un sin número de tests psicológicos o

instrumentos de medición, que claro está, se pueden dividir en muchas clases o tipos. A continuación se hará una revisión de la clasificación de los test, no sin antes hacer algunas precisiones de las diferencias -o similitudes- entre los términos que tan comúnmente se utilizan.

1.4.1 Diferencia entre Test y Prueba.

“Los términos test, prueba o instrumento de medición se usan indistintamente, cualquiera de los tres es correcto. Ya que estos son considerados de uso universal en la lengua hispana” (Aiken,1995). En este mismo caso se encuentran los términos test mental y test psicológico.

Etimológicamente la palabra “test” proviene del latín testis, que significa testigo. A la psicología llega a través del inglés, donde, como verbo significa: examinar, probar, ensayar; como sustantivo: prueba, ensayo, examen, reactivo.

Dentro de la psicología son muchos los autores que han intentado una definición al término de test psicológico (García, 1993):

“Reactivo que aplicado a un sujeto, revela y da testimonio de la índole o grado de su instrucción, aptitud o manera de ser” (Yela,1987).

“Artefacto para obtener una muestra de la conducta de un individuo” (Allen y Yen, 1979).

“ Instrumento de evaluación cuantitativa de los atributos psicológicos de un individuo” (Anstey, 1976).

“ Medida objetiva y normalizada de una muestra de la conducta “ (Anastasy, 1970).

Cualquiera de ellas puede ser válida, y todas juntas dan una buena idea de lo que es un test en psicología, “... un instrumento normalizado para medir algún aspecto práctico o cognoscitivo de la conducta de un individuo y/o un grupo de individuos” (García, 1993).

Así el término Test se convirtió en un conjunto de ítems; por lo general, cada ítem o tarea daba un resultado binario (éxito o fracaso). “El resultado del test pasó a ser expresado mediante el número de ítems realizados con éxito y la escala resultante fue denominada “escala de puntos” (Sternberg, 1987).

El concepto de prueba psicológica varía, aún cuando se puede encontrar que hay factores en común que permiten sintetizar y establecer una definición general que podría enunciarse como : “un instrumento de medición psicológica es aquella

técnica metodológica producida artificialmente, que obedece a reglas explícitas y coloca al o a los individuos en condiciones experimentales, con el fin de extraer el segmento del comportamiento a estudiar y que le permita la comparación estadística conductora a la estadística cualitativa, tipológica o cuantitativa de la característica o características que se están evaluando (Morales, 1993).

Pierón (1952, en Pichot, 1991), dice que “el test mental es una prueba definida, que implica una tarea a realizar, idéntica para todos los sujetos examinados con una técnica precisa para la apreciación del éxito o fracaso, o para la puntuación numérica del éxito. Esta tarea puede poner en juego ya sea conocimientos adquiridos, ya sea funciones sensoriomotrices o mentales”, tales como la coordinación motriz fina o gruesa o la misma inteligencia.

En la presente investigación, se podría utilizar para fines prácticos y de simplificación, la definición establecida por Pichot (1991): “ una prueba psicológica es una situación experimental estandarizada que sirve de estímulo a un comportamiento y que, manipulada estadísticamente, permite clasificar a los sujetos tipológica o cuantitativamente”.

1.4.2 Tipos de Test.

Del mismo modo en que existen muchos tipos de comportamientos, también existen muchos tipos de test, por lo tanto, hacer una descripción detallada de

cada uno de ellos sería un trabajo exhaustivo que no arrojaría ningún fruto. Es por eso que se hará una revisión de la clasificación de los test en sus aspectos más gruesos, con el fin de poder ubicar lo más preciso posible el test de Dominós de Edgar Anstey, objeto de la presente investigación.

Primero, los test que pueden ser administrados a una sola persona a la vez, son conocidos como **tests individuales**. La situación involucrada alrededor de un test individual es similar a la de una terapia personal, en donde el examinador (la persona que administra el test), aplica el test sólo a una persona a la vez, de la misma manera que los psicoterapeutas individuales atienden sólo a una persona. Por el contrario, los **test grupales o colectivos** pueden ser administrados a más de una persona a la vez por el mismo examinador. Exactamente como lo hace un psicoterapeuta en una sesión de terapia familiar o de grupo.

Para poder hacer una clasificación de los test, sólo se puede realizar de acuerdo al tipo de comportamiento que cada una de ellas pretendan medir.

Los **test de habilidades** contienen ítems que pueden ser cuantificados en términos de velocidad, precisión o ambos. En un test de habilidad, la mayor velocidad o la mejor precisión en las respuestas, obtiene las puntuaciones más altas en una característica particular.

Cuando la prueba está diseñada para medir el número de respuestas correctas que un sujeto puede dar en un tiempo previamente determinado, se está hablando de un **test de rendimiento o velocidad**. Por otro lado, existen algunos casos en que la prueba no tiene límite de tiempo; aquí se trata de medir hasta dónde puede llegar el sujeto o cuántas respuestas correctas puede hacer, independientemente del tiempo que se tarde, este tipo de tests son llamados **tests de potencia, de capacidad potencial o de poder ("Power tests")**. En algunos casos, en este tipo de pruebas se incluye un tiempo límite, solo con el objeto de facilitar su aplicación y no perder tiempo innecesariamente.

Citando al profesor Anstey (1959), "El test de Dominós es fundamentalmente un test de poder, pues los sujetos menos inteligentes nunca resolverán los problemas más difíciles, aún cuando se les de un tiempo ilimitado para hacerlo. No obstante, la imposición de un límite razonable de tiempo (30 min.) facilita la administración del test".

Otra característica que establece diferencias entre los tests es que éstos pueden ser Verbales o No Verbales. Los test Verbales son aquellos que sus ítems están formados por material que involucra capacidades verbales - valga la redundancia -, como sería la lectura o el vocabulario, y por el contrario, los tests No Verbales, involucran aspectos más bien perceptuales visuales.

Históricamente, se han hecho distinciones entre potencial, aptitud e inteligencia, como los diferentes tipos de habilidades. La *capacidad* o potencial se refiere al aprendizaje previo. Por ejemplo, un test que mide o evalúa cuántas palabras tu puedes definir correctamente es llamado test de *capacidad* de vocabulario, aunque muchas veces es omitido el término “capacidad”, por parecer obvia la especificación.

Por otro lado, la *aptitud* se refiere al potencial para el aprendizaje o la adquisición de una cierta destreza o habilidad. Claro está que dicho aprendizaje se debe llevar a cabo administrando alguna cantidad de entrenamiento, educación o a través de la experiencia. La aptitud musical se refiere en parte a que tan bien tu puedes ser capaz de aprender a tocar algún instrumento musical después de tomar cierto número de lecciones.

Tradicionalmente diferenciada de la aptitud y la capacidad, la *inteligencia* se refiere a la capacidad general de las personas para resolver problemas, adaptarse a situaciones cambiantes, pensar en forma abstracta y aprovechar la experiencia. Cuando se dice que una persona es “lista”, probablemente se está haciendo referencia a la inteligencia.

Las distinciones entre capacidad, aptitud e inteligencia, no han sido tan precisas como se quisiera; como sea, las tres están altamente relacionadas. Por ejemplo, los intentos de separar el *aprendizaje previo* de la *capacidad de aprendizaje* no

han tenido mucho éxito. Existe un considerable traslape entre los test de capacidad, aptitud e inteligencia. Como resultado de este traslape, los tres conceptos pueden involucrarse por el término **habilidad humana**.

Hay una clara distinción entre los test de habilidades y los test de personalidad. Mientras que los test de habilidades son relacionados con capacidad y potencial, los **test de personalidad** son relacionados con los contenidos latentes y manifiestos del individuo. Por ejemplo, la tendencia de una persona a mostrar un comportamiento o respuesta en particular ante una situación determinada. El aislarse de los demás, por ejemplo, no requiere de alguna habilidad particular, sin embargo, algunas personas lo prefieren típicamente o tienen disposición a aislarse ante ciertas situaciones. Los test de personalidad miden el comportamiento típico de las personas.

Existen muchos test de personalidad, los cuales pueden ser reunidos en dos grandes grupos. Los llamados test de **personalidad estructurados** proveen un informe del perfil de personalidad del individuo, dependiendo del tipo de respuestas que dé el sujeto a las preguntas previamente establecidas en la prueba, las cuales en su mayoría, requieren que el individuo escoja entre dos o más alternativas de respuesta como "cierto o falso".

En contraste a este tipo de test, los **test proyectivos** son más bien inestructurados o no estructurados. En un test proyectivo de personalidad, los

estímulos (los materiales del test) y las respuestas dadas por el sujeto o ambas, son ambiguas. Es decir, estas pruebas están diseñadas en forma “parcial”, para que el individuo las “complete” con los contenidos de su historia personal.

Hasta aquí se ha hecho una clasificación “a groso modo” de los tipos de test, con lo cual se puede ubicar precisamente al test de Dominós de Anstey como un test no verbal de habilidades humanas, específicamente, como un test de inteligencia, de potencia (“power test”) por no requerir de un tiempo específico de realización y cuya aplicación puede ser individual o colectiva.

1.5 El test de Dominós de Anstey.²

Como se ha visto, el desarrollo de los tests de inteligencia es muy extenso, pero la historia de los tests de puntos en general y específicamente la de los Dominós, termina en el test de Anstey. Es en este test donde la historia alcanza su más definida y alta realización.

La primera versión del test de Dominós fue elaborada por el psicólogo Inglés Edgard Anstey en 1944, para el uso del Ejército Británico, primero como test paralelo de las matrices Progresivas de Raven y luego como prueba substitutiva de ésta.

² Tomado del Manual del test de Dominós (Ver referencia bibliográfica).

La versión original del test de Dominós (1944) que preparó Anstey para la armada Británica, integrada por 44 ítems, ha quedado inédita, como un "closed test" - según denominan los ingleses a las pruebas de uso oficial o exclusivo -. Pero, posteriormente, Anstey preparó una nueva versión (1955), ligeramente diferente en su material (48 ítems agrupados, en 8 páginas según sus principios y complejidad creciente), a la que le dio, en cambio, el carácter de "open test".

Anstey confió su versión de 1955 al distinguido investigador uruguayo Dr. W.L. Risso, con el fin de obtener su estandarización y publicación, quien obtuvo un baremo nacional para el Uruguay, y el cual se ha venido utilizando indistintamente para la población Mexicana.

1.5.1 Lo que mide el test de Dominós. Características

El test de Dominós es un test no verbal de inteligencia. Ha sido construido con el propósito de establecer la capacidad de una persona para aprehender correctamente el número de puntos de grupos de dominós, descubrir el principio ordenador de esos grupos y mediante la aplicación de ese principio, averiguar cuántos puntos le corresponde a cada mitad en blanco de un dominó de ese grupo para completar aquel patrón.

El rendimiento en este test no depende en un grado considerable del ambiente, la educación o la experiencia del examinado. No depende directamente del conocimiento previo del juego de dominós.

El test de Dominós mide la inteligencia de los sujetos basándose en la teoría del factor General, esta prueba posee una saturación G de 0.82 (en comparación con una saturación G de 0.79 para el Test de matrices progresivas), y una pequeña saturación espacio numérica de 0.12.

Para hacer un acercamiento más preciso a las características del Dominós, se deberá entonces dividirlo según las notas siguientes:

En razón de sus problemas

Test de series: Es un test de educación de leyes o principios de relaciones.

Test de series numéricas o más precisamente *Test de continuación o completamiento de series numéricas* . Puede considerárselo como un *test de completamiento de lagunas*. Ese completamiento entraña la tarea de terminar series de números dispuestos según diversos planes, patrones o principios. Los principios que regulan las secuencias en las series de Anstey son los siguientes:

	<i>Página</i>
Simetría	1
Alternancia y progresión simple	2 y 3
Asimetría	4
Progresión circular	5
Progresión compleja (series)	6
Combinación de principios previos	7
Adición y sustracción	8

En razón de su material

Test de figuras numéricas de puntos: Las series de cifras se representan por grupos de puntos dispuestos según un patrón.

Test de series de dominós: por utilizar las fichas de dominós en forma exclusiva, le caven estas otras especificaciones

Test homogéneo perceptual. Test de figuras abstractas: en su apariencia visual las fichas ofrecen un material concreto y familiar; pero por la índole no real de los problemas que plantea, que es lo que importa, El Dominós es un test de figuras no representacionales.

En razón de su administración

Autoadministrable. Puede tomarse en forma individual o colectiva.

Sin límite de tiempo: ("Power Test").

En razón de sus fundamentos y su objetivo.

Test factorial. El test de Dominós de Anstey se fundamenta en la teoría factorial de Spearman (1904) y en su metodología psicométrica. Anstey aplicó, en efecto, los principios psicométricos fundamentales sustentados por los factorialistas ingleses, desde sus iniciadores (1900) Spearman, Brown y Stephenson: procurar la medición de la "inteligencia general" mediante pruebas de educación de relaciones sobre material homogéneo no verbal o perceptivo; específicamente, a través de problemas formulados en términos de descubrimiento de relaciones entre series de figuras abstractas, preferidas por su carácter no familiar, lo cual permite una evaluación más pura de G.

Test de factor G. Su propósito de medir la capacidad intelectual con la máxima pureza, vincula al Dominós con los tests del factor G de Spearman-Stephenson (1934); Test perceptual, de Penrose y Raven (1936); Matrices Progresivas, de Raven (1938); test de inteligencia no cultural, de Cattell (1940); Percepción de patrones, de Penrose(1943).

Test de capacidad potencial ("Power Test"). El dominós brinda al examinado el tiempo que necesite para realizar el máximo de tareas que sea capaz. A diferencia de los tests de velocidad, que hacen la evaluación por la rapidez de respuesta.

1.5.2 Puntuación y calificación.

En esta prueba los ítems están, en general, dispuestos en orden de dificultad creciente, pero no absolutamente, dado que el primer ítem correspondiente a un nuevo principio suele ser algo más fácil que el último ítem correspondiente al principio anterior.

La calificación de la prueba se hace notar por su sencillez, se acredita un punto por cada respuesta totalmente correcta. La posibilidad de dar una respuesta correcta por azar es de 1/49, y los protocolos se pueden calificar a razón de uno por minuto.

El puntaje bruto al que se llega al final del proceso no constituye un dato suficiente para la calificación, ya que un mismo puntaje tiene un significado diferente para cada grupo de edad. Se hace necesario entonces, usar una escala que permita una rápida ubicación del sujeto dentro del grupo de edad al que pertenezca.

En el caso de los adultos, Anstey sugiere que la distribución de los puntajes es aproximadamente la siguiente:

Puntajes	Población
48-41	10 % superior
40-33	20% siguiente
32-27	20% siguiente
26-20	20% siguiente
19-11	10% siguiente
10-0	10% inferior

Puntaje medio en adultos: 26/27.

1.5.3 Estandarización

Población

La población sobre la cual se realizó la estandarización en Uruguay se integró por 1736 sujetos de 12 a 30 años de edad, alumnos de escuelas montevidéanas, los cuales fueron elegidos al azar, y en cuanto a edades, sexo y nivel de escolaridad, representaban la totalidad de la población estudiantil.

Según el profesor Risso (1957), las curvas de distribución de frecuencia de los datos encontrados, resultan moderadamente asimétricas en sentido negativo.

Confiabilidad

En su investigación, el Dr. Risso encontró que en el test de Dominós, el coeficiente de confiabilidad tiende a crecer con la edad, yendo de 0,850 a 0,910. Para el total de la población examinada, el coeficiente es 0,854.

Validez

En el caso de la investigación en Montevideo, Uruguay, se tomó como criterio para la validación del Dominós, el test de Matrices Progresivas de Raven; dado que el análisis factorial de ambos los caracteriza como test con fuerte

saturación en G y relativamente incontaminados por otros factores, como el numérico o el espacial. El coeficiente de correlación obtenido por Riso en dicha investigación tiene un valor de 0,550.

Teniendo el Raven y el Dominós coeficientes de confiabilidad próximos a la unidad, el coeficiente de validez resulta aproximadamente igual al coeficiente de correlación entre ambos (0,550 en este caso).

1.5.4 Valores estadísticos

El cuadro de valores estadísticos encontrado por el Dr. Riso para la población de Montevideo, Uruguay fue la siguiente:

Tabla 2.

Edad	12-13	14-15	16-17	18-30	población general
N	= 336	855	340	205	1736
Media	= 25,79	26,83	27,68	29,20	27,32
Mediana	= 26,68	27,65	28,50	30,48	28,16
Sigma	= 8,10	8,01	8,46	7,56	8,16
Coeficiente de Conf.	= 0,858	0,855	0,875	0,853	0,863

De igual forma, el Dr. Risso utilizó una escala para una rápida ubicación de los sujetos dentro del grupo de edad al que pertenecían. Para ello, elaboró una escala de percentiles (o centillas) de quince grados, para los diversos niveles de edad examinados, a saber: 12-13; 14-15; 16-17 y 18 y + años. Las centillas o percentiles tomados como base fueron: 1; 5; 10; 20; 25; 30; 40; 50; 60; 70; 75; 80; 90; 95; 99.

Con tales datos, obtuvo una tabla de normas percentilares para el test de Dominós (versión 1955), calculadas sobre 1736 alumnos adolescentes y adultos de Uruguay. La tabla es la siguiente:

Tabla 3.

Edad	12-13	14-15	16-17	18-30	población general
Percentiles					
1	4	5	6	8	5
5	9	11	12	16	12
10	14	15	16	20	17
20	21	22	23	24	22
25	22	23	24	25	23
30	23	24	25	26	25
40	25	26	27	29	27
50	27	28	29	31	29
60	29	30	31	33	31
70	31	32	33	35	33
75	32	33	34	36	34
80	33	34	35	37	35
90	35	37	39	40	37
95	38	39	41	41	40
99	42	43	44	45	44

Por último, para poder hacer una ubicación precisa de cada sujeto calificado con esa tabla, fue necesario hacer una tabla de conversión de percentiles a rangos; la tabla sugerida por el Dr. Risso es la siguiente:

Percentiles		Rangos
95	SUPERIOR
90 Y 75	SUPERIOR AL TERMINO MEDIO
50	TERMINO MEDIO
25 Y 10	INFERIOR AL TERMINO MEDIO
5	DEFICIENTE

Hasta aquí se han expuesto los valores obtenidos en la estandarización del Dominós para la población de Montevideo, Uruguay, los cuales han sido utilizados por igual para la calificación y clasificación de los sujetos mexicanos. Sin embargo, como se pretende demostrar en la presente investigación, su uso para la población mexicana ha sido un error, ya que los valores estadísticos son diferentes y , por lo tanto las tablas y baremos deben también ser diferentes para la correcta clasificación de los sujetos mexicanos según su inteligencia.

1.6 El factor General.

Junto con el desarrollo de los tests construidos dentro de la tradición de Binet, Terman y Wechsler, han surgido numerosos instrumentos de evaluación de

la inteligencia en relación con diferentes teorías sobre la organización y naturaleza de la inteligencia, los cuales se han puesto a prueba mediante el empleo del análisis factorial. Gran parte de estas pruebas han sido utilizadas a lo largo de los años con fines predictivos en el contexto de la selección y clasificación de personal y en el académico.

El test de Dominós de Edgard Anstey es uno de los que tienen su sustento teórico en el análisis factorial, siendo su base más específica el Factor General de la inteligencia o factor G.

En realidad la expresión "análisis factorial" engloba un conjunto de técnicas que, al ser aplicadas a las subpruebas de un test, tratan de encontrar qué o cuánto tienen en común las subpruebas entre sí. Con esto se pretende encontrar los factores comunes que existen en una misma prueba. El análisis factorial combina las calificaciones para reducir el número de variables o resumir los datos. La meta del análisis factorial es el de obtener una variable o un factor que reemplace a varias medidas distintas. Por eso, la utilidad primordial es la de instrumento teórico o de explicación más que la de ayuda práctica para la toma de decisiones.

El análisis factorial es una técnica estadística para determinar el número de factores necesarios para justificar las intercorrelaciones entre un conjunto de variables. Su meta es encontrar grupos de variables que midan el mismo factor. Este proceso de agrupamiento sirve para reducir el número de conceptos de

explicación que se necesitan en el sistema. Un factor no es más que un compuesto lineal dado, formado a partir de algunas de las variables originales.

Lo anterior ha llevado a los estudiosos de la inteligencia a preguntarse si tiene sentido hablar de la inteligencia como una capacidad de carácter general, que se halla presente siempre que se enfrenta con algún problema o si, por el contrario, lo que existen son diferentes aptitudes o capacidades específicas en las que las personas se diferencian cuando han de afrontar la solución de determinados problemas o tareas.

Como respuesta a las interrogantes anteriores, Jesús Alonso Tapia (1986), menciona que el uso del análisis factorial ha permitido poner de manifiesto que la inteligencia no es un rasgo unitario pero que no es un conjunto de rasgos inconexos. Las correlaciones existentes entre los distintos tests que definen las habilidades primarias sugiere la existencia de un factor general de inteligencia conocido como factor G. La interpretación de los datos del test de Dominós exige que se conozca a qué se hace referencia con tales constructos, punto que se expondrá a continuación.

1.6.1 La inteligencia y el factor G.

Cuando se habla de Capacidades Intelectuales o de Inteligencia en general, se hace referencia a un conjunto de habilidades o capacidades por separado, que al

trabajar en conjunto dan al ser humano la posibilidad de resolver problemas o de encontrar la relación entre dos o más situaciones, entre otras cosas. Entonces, ¿qué es lo que se evalúa cuando se utiliza pruebas que miden el factor G? ¿qué implica que un sujeto puntúe alto o bajo en estas pruebas?

“Desde un punto de vista matemático, el factor G es lo que un conjunto amplio y diverso de tests tienen en común. Es un constructo hipotético mediante el que se hace referencia a una dimensión lineal en la que cabe ordenar las puntuaciones que representan el promedio ponderado de los resultados obtenidos por distintas personas en un número determinado de tests.” (Alonso Tapia, Jesús, 1986, en Fernández-Ballesteros, 1993).

En principio, no cabe considerar al factor G como un factor único, ya que en la medida en que el factor G es una función de la composición de la matriz de correlaciones a partir de la que se obtiene y dado que los resultados de la matriz pueden ser diferentes debido al tamaño de la muestra examinada y al número y tipo de tests utilizados. Sin embargo, en realidad esto no constituye un gran problema en la medida en que se defina con anterioridad cuál es el dominio o área en relación con el cual se trata de determinar qué variables pertenecen al mismo y cómo saturan estas en factor G de ese dominio. Como ejemplo se puede mencionar el área o dominio de las habilidades como el de la conducta voluntariamente orientada a metas claramente diferenciadas y clasificables de acuerdo con la escala de calidad, frente al área de la personalidad donde no ocurre

lo mismo. Por eso, dado que todas las medidas que caen dentro del dominio de las habilidades mentales correlacionan positivamente, con que los tests sean suficientemente amplios y su contenido sea heterogéneo, cabe suponer que tal correlación se debe a que dependen de un factor general. Por otro lado, no todos los tests dependen de dicho factor de igual forma, lo cual dificulta la determinación de la naturaleza psicológica de G.

El primero en intentar delucidar la naturaleza de G fue Spearman (1927, en Alonso Tapia, 1986). Spearman, tras analizar factorialmente más de cien tests, todos con contenido homogéneo, y después de observar en que se parecían los tests cuya ejecución dependía en mayor grado del factor G y cuáles eran las diferencias de tales tests de aquellos en los cuales dicho factor no ejercía influjo alguno, "llegó a la conclusión de que el proceso psicológico que subyace a los resultados que reflejan la existencia del factor G es un proceso *noegenético* y de abstracción. El término *noegenesis* hace referencia a tres aspectos de la vida mental del sujeto: a la capacidad de introspección que permite al sujeto observar lo que ocurre dentro de su mente, a la capacidad de educir las relaciones esenciales existentes entre dos o más ideas que se tienen en mente y a la capacidad de educir correlatos, esto es, a la capacidad que una persona tiene, supuesto que tenga en su mente una idea y una relación, de concebir la idea original en esa relación". (Alonso Tapia, Jesús, 1986, en Fernández-Ballesteros, 1993).

De hecho, entre las tareas que más saturan en G están las de:

- * Matrices progresivas
- * Analogías verbales
- * Continuación de series
- * Razonamiento sobre problemas aritméticos
- * Comprensión de párrafos
- * Analogías figurales y clasificaciones de figuras.

Parece pues, que la complejidad de la tarea y la cantidad de actividad o manipulación mental consciente que requiere su solución son los principales determinantes del grado en que una tarea satura en G, y dado que, en realidad la complejidad de la tarea no es más que la oportunidad para que un sujeto desarrolle su actividad mental, *“puede decirse que el elemento crucial responsable del grado en que un test satura en G es la cantidad de manipulación mental consciente desencadenada por la información presentada al sujeto”*. (Alonso Tapia, Jesús, 1986, en Fernández-Ballesteros, 1993). Se recalca el aspecto *consciente* de dicha actividad porque puede ser que en algunas ejecuciones complejas el sujeto conteste en forma automática e inconsciente como resultado del aprendizaje, situación en que las diferencias individuales correlacionan menos con G que en el caso de tareas en que todos los sujetos se hallan en fase de aprendizaje. *“De hecho, cierto grado de novedad parece necesario para que un test sature en el factor G, ya que es la novedad, fundamentalmente, la que elicit la actividad mental consciente* (Jensen, 1980, 1982, en Alonso Tapia, 1986).

Dentro de los principales tests de factor G están los Tests de Dominós, de los cuales existen numerosas variantes. El primero de ellos fue creado en 1944 por Anstey en Inglaterra, a partir de los mismos principios que las matrices construidas por Raven. Dentro de las variantes de estos tests se encuentran el TIG-1 (Test de Inteligencia General), el D-48, D70 y el TIG-2. Estos tests son medidas casi puras del factor G, encontrándose saturaciones de hasta 0,86 en dicho factor, y los índices de fiabilidad obtenidos con estas pruebas están en general por encima de 0,85.

El hecho de que algunas pruebas de inteligencia como las escalas Wechsler estén diseñadas para medir en forma aislada algunos de los factores que las integran, no quiere decir que sean más exactas o precisas, simplemente pone de manifiesto una puntuación para cada una de las áreas en las que fue dividida, puntuaciones que al final, arrojan una sola a partir de su manipulación matemática.

Como se puede observar, la utilización de pruebas de factor G, como el test de Dominós de Edgard Anstey, representa una forma confiable de medir la inteligencia de los sujetos en forma por demás rápida y sencilla, sin tener que recurrir a tablas u operaciones matemáticas complicadas.

1.7 Conceptos y principios de la medición y estadística.

Para muchos estudiantes y más para aquellos que dicen que su carrera “no tiene nada que ver con matemáticas”, la estadística no es más que un conjunto de nombres y fórmulas extrañas que tienden a complicar las cosas cuando se trata de manipular información basada en números. Sin embargo, contrariamente a lo que opinan los estudiantes universitarios de carreras humanísticas, la estadística es una herramienta imprescindible de cualquier disciplina académica e incluso de la vida cotidiana.

Durante siglos, el hombre ha estado recopilando datos estadísticos, por ejemplo, se necesitaba conocer la cantidad de alimento necesario para dar de comer a una tribu o población; la cantidad de guerreros, armas y caballos y cuántos necesitaba de cada uno para vencer al enemigo. Actualmente, se dispone de una gran cantidad de datos relativos a la industria, negocios, deportes, aspectos sociales, médicos, psicológicos y a cientos de otras actividades humanas. Desde un punto de vista muy estricto, cada ser humano y lo que hace, es entonces, un dato estadístico que puede ser analizado e interpretado simplemente como un número.

La ciencia estadística puede considerarse también como un método que sirve para analizar datos, esto es, para organizar y dar significado a una gran cantidad de información. La metodología estadística puede considerarse según tres planteamientos, como descriptiva, correlacional, y deductiva o inferencial.

La estadística descriptiva se encarga, como su nombre lo indica, de describir o proporcionar información de la muestra que estudia. La correlacional, trata de establecer si dos muestras o alguna característica de ellas están relacionadas entre sí y por último, la inferencial o deductiva, se encarga de formular deducciones de la población a partir de la que se obtuvo la muestra.

Pero la estadística tuvo que pasar por un amplio recorrido antes de llegar a lo que es hoy en día. En los siglos XVII y XVIII, apostadores profesionales pidieron a algunos matemáticos que desarrollaran principios que pudieran mejorar las oportunidades de ganar con los naipes y los dados. Los dos matemáticos más notables que intervinieron en este primer y más importante estudio de la probabilidad, fueron Bernoulli y DeMoivre. En la década de 1730, el segundo, DeMoivre, desarrolló la ecuación de la curva de distribución normal. Durante las dos primeras décadas del siglo XIX, otros dos matemáticos, Laplace y Gauss, realizaron importantes trabajos sobre el cálculo de probabilidades. Su labor consistió en la aplicación de los principios de la probabilidad a la astronomía.

“Durante el siglo XVIII, la ciencia estadística tuvo aplicaciones de tipo matemático, político y gubernamental. A principios del siglo XIX, Quetelet, un famoso investigador belga, aplicó la estadística en la investigación de problemas sociales y educativos. Walker (1929) atribuye a Quetelet el desarrollo de la teoría estadística como método general de investigación aplicable a todas las ciencias de

la observación. Sin duda alguna, la persona que ejerció con mayor influencia en la introducción y en el empleo de la estadística en las ciencias sociales, fue Francis Galton. En el transcurso de su larga vida contribuyó notablemente en los estudios de la herencia y de la eugenesia, de la Psicología, de la Antropometría y de la Estadística. Se le atribuyen los conocimientos actuales que se tienen acerca de la correlación, es decir, la medida de la concordancia entre dos variables. El matemático Pearson colaboró con Galton en años posteriores, y participó en la creación de muchas de las fórmulas de correlación y regresión que se utilizan hoy en día. Entre las atribuciones importantes de Galton hay que citar el desarrollo de los centiles (o porcentiles),”(Downie-Heat, 1986)

El famoso psicólogo estadounidense James McKeen Catell estudió en Europa en la década de 1880 y estuvo en comunicación con Galton y otros estadísticos europeos. A su regreso a Estados Unidos, él y sus discípulos, incluyeron a E. L. Thorndike, empezaron a aplicar los métodos estadísticos en problemas psicológicos y educativos. La influencia de estos hombres fue importante; al cabo de unos años ya se impartían cursos de estadística teórica y aplicada en las universidades de ese país.

En este siglo XX se han aplicado nuevas técnicas y métodos en el estudio de muestras pequeñas. Las principales aportaciones en la teoría de las muestras pequeñas fueron las del estadístico inglés R. A. Fisher. Aún cuando la mayoría de sus métodos fueron desarrollados en el campo agrícola o biológico, no transcurrió

mucho tiempo antes de que los sociólogos reconocieran su utilidad y aplicaran sus ideas. Actualmente, la estadística es la principal herramienta metodológica del investigador en las ciencias sociales.

En psicología se pueden utilizar e incluso se utilizan muy frecuentemente los tres tipos de estadística: descriptiva, correlacional y deductiva. Por ejemplo, cada vez que se hace una prueba psicológica, se está describiendo a detalle la muestra seleccionada en función de lo que se esté midiendo, su diferencia por sexos, edades, escolaridad, estratos sociales o razas. Los resultados de análisis correlacional sirven para estudiar la validez y confiabilidad, e incluso forman parte importante en la misma elaboración de los test al identificar si dos o más subtests están correlacionados o no. Y por último, se pueden establecer predicciones de rendimiento de una población a partir de los resultados obtenidos por la muestra.

Como se puede observar, el empleo de la estadística en actividades humanísticas no se encuentra totalmente desligado, por el contrario, su uso ofrece la posibilidad de conocer mejor al objeto de estudio de la psicología: el hombre y su mente. A continuación se hará una revisión general de los conceptos estadísticos que fueron utilizados en la metodología de la presente investigación.

1.7.1 Las variables en estadística.

Tipos de Medidas

Se pueden clasificar los datos en dos tipos: *continuos y discontinuos*. Los continuos pueden considerarse como puntos de una recta. La magnitud y precisión de las medidas que pueden realizarse a lo largo de esta recta dependerán de la forma como se haga la medición y del instrumento que se utilice. Cantidades en metros, minutos, pies y libras, son ejemplos de datos continuos. Con ellos se pueden expresar medidas de diversos grados de precisión. Por ejemplo, es posible descomponer horas en minutos, minutos en segundos, y con instrumentos muy elaborados se pueden llevar a cabo mediciones cada vez más y más precisas.

Sin embargo, cuando no se cuenta con la instrumentación más precisa e incluso con ella, es necesario considerar que todas las mediciones tienen una tolerancia o margen. Es decir, en una medición del peso de un niño que marque 35 Kg., no significa que pese exactamente eso, sino que es la marca redondeada más cercana a su peso. Estos 35 Kg. se considera que incluyen una tolerancia o un margen, esto es, el segmento de 1 Kg. que se extiende desde 34.5 Kg. hasta 35.5 Kg. Cada medición en una escala de datos continuos tiene un límite inferior y uno superior.

Los datos discontinuos (o discretos, como también se les llaman) provienen de mediciones que sólo pueden expresarse como cantidades enteras. Contar personas, por ejemplo, únicamente puede hacerse en enteros, a diferencia de los metros, que pueden expresarse en unidades decrecientemente pequeñas o crecientemente mayores.

Aun cuando las calificaciones de una prueba sean discretas, suelen tratarse como si fueran continuas, bajo el supuesto de que si fuera necesario, se podrían hacer distinciones más precisas.

Escalas de Medición

Stevens (1946, 1958, 1968, en Downie, 1986) reconoce cuatro tipos de escalas: *nominal, ordinal, de intervalos y de razones*. Las escalas nominales son usadas como medidas de identidad. Los números pueden ser usados para identificar objetos, personas o clases en forma de indicativo o de "etiqueta", es decir, el número sirve para dar un nombre. Los números que los atletas llevan en la espalda constituyen una escala nominal en su forma más simple.

En una escala ordinal, los números indican el orden o la jerarquía de individuos u objetos. Las medidas ordinales se disponen, en orden, valga la redundancia, desde la más alta hasta la más baja o viceversa. Las personas de un grupo pueden ordenarse según sus rasgos físicos o sus características mentales.

Pero tales medidas no pueden expresar cuánto más alto o cuánto más inteligente es un individuo respecto a otro. Sin embargo, no pueden obtenerse grandes resultados estadísticos con las medidas ordinales, excepto determinar la mediana y los centiles, o bien, calcular coeficientes de correlación de rangos.

En una escala de intervalos, sí se proporcionan números que reflejan diferencias entre los individuos u objetos. En el caso de esta escala, las unidades de medición son exactamente iguales. Como ejemplos de las escalas de intervalos se tiene la del tiempo, considerada según el calendario común, y las puntuaciones obtenidas en pruebas o tests de inteligencia. En el último caso se suponen unidades iguales de medición. Muchos valores estadísticos se utilizan con las escalas de intervalos: la media (aritmética), la desviación estándar y el coeficiente de correlación (producto-momento). Así mismo, las pruebas de significación estadística más utilizadas, la prueba *t* y la prueba *F*, pueden utilizarse con tales datos. Las escalas de intervalos indican que un individuo u objeto es tantas unidades más grande o más pequeño, más pesado o más ligero, más inteligente o menos inteligente que otro, etcétera.

El tipo de escala más refinado es la escala de razones. La principal diferencia entre esta escala y la anterior, es que en este tipo sí existe un cero absoluto. Las escalas de razones más comunes corresponden a medidas de longitud, capacidad, peso, sonido, entre otras. Cuando se emplea una escala de razones, los números indican cocientes entre ciertas magnitudes de los objetos, y

los datos obtenidos con tales escalas pueden ser sometidos a tratamientos estadísticos más elaborados.

1.7.2 Centiles o Percentiles.

Al hacer referencia a un centil (o punto de centil) se puede definir como un punto específico en una distribución que tiene un porcentaje dado de los casos por debajo de él. Por ejemplo el centil 75 (C 75) es el punto o puntuación de una distribución tal que el 75% de los casos están por debajo de él.

Los centiles se utilizan mucho para informar los resultados de las pruebas estandarizadas. En su favor puede decirse que son muy fáciles de entender. Dado que los centiles se parecen a los porcentajes, no hay una gran dificultad para comprender su significado. Así mismo, proporcionan una indicación adecuada del rango o jerarquía de un individuo en un grupo.

Sin embargo, los centiles tienen serias limitaciones, y muchos de los que elaboran y hacen uso de pruebas ya no se molestan por ellas. Por ejemplo, en el centro de una distribución, el uso de puntuaciones centiles tiende a exagerar diferencias realmente inexistentes. Los centiles son unidades desiguales de medición y no pueden ser tratados aritméticamente. Esto es, no hay justificación para promediarlos, combinarlos o tratarlos de modo matemático. “ En lo que se refiere a la estadística, los centiles apenas tienen utilidad. No se puede hacer nada más con ellos. Si se desea manejar datos que se han reducido a centiles, deberán convertirse de nuevo en puntuaciones no elaboradas y operar luego con éstas.

Dado que los centiles son unidades de medida desiguales, algunos estadísticos consideran que es mejor descartarlos. En el caso de pruebas estandarizadas los centiles se utilizan cada vez menos como método de información ", (Downie - Heath, 1986).

1.7.3 Medidas de tendencia central o Promedios.

Una vez que se han recopilado una serie de datos, éstos tienen que ponerse en una forma que haga posible resumirlos e interpretarlos fácilmente. Generalmente, uno de los primeros valores obtenidos para una serie de datos es una medida de tendencia central o promedio. El promedio es una cifra única que representa a un grupo de cifras.

De manera general existen tres promedios: la media , \bar{X} , la mediana, M_d y la moda, M_o .

La *media aritmética* es el promedio de uso más frecuente, y es también la medida más útil de la tendencia central. Este valor es el promedio que se encuentra usualmente. La media no es más que el promedio aritmético y para calcularla, se suman todas las calificaciones y se dividen entre el número de ellas.

La *moda* (o modo) se define como el valor que se observa más frecuentemente en una distribución de puntuaciones o de medidas. En el caso de

que los datos estén agrupados, la moda se puede definir como el punto medio del intervalo que contiene mayor número de casos.

La *mediana* se define como el punto en una distribución con 50% de las puntuaciones a cada lado de él. Esto es, la mediana es el punto medio de una distribución, es la calificación que divide la distribución en mitades; las calificaciones caen 50% hacia arriba y 50% hacia abajo.

De los tres promedios presentados, la moda es el valor menos confiable. Se utiliza propiamente con datos nominales, pero puede emplearse con datos ordinales, de intervalo o de razón. Sin embargo, por lo general se usa poco, dado que no se presta para procesos de calcular. Sin embargo, es fácilmente determinada por quién desea una rápida medida de tendencia central.

La mediana es el promedio adecuado para usarse con datos ordinales, o cuando una distribución se aparta de lo normal. En realidad muchas distribuciones, al ser examinadas demuestran no ser normales. La mediana es la medida de tendencia central apropiada cuando los datos están clasificados o cuando la presencia de pocas calificaciones hacia los extremos tiene probabilidades de distorsionar un promedio aritmético.

La media se utiliza como el promedio apropiado para datos de intervalo o de razón cuando las distribuciones son simétricas o aproximadas a la forma

normal. En la mayoría de los cálculos estadísticos, es esencial que la media se emplee como el promedio.

1.7.4 Variabilidad.

Los promedios o medidas de tendencia central son muy útiles cuando se trata de analizar y representar la información obtenida a través de los datos de una muestra, sin embargo, la información no es la suficiente como para llegar a conclusiones contundentes. Una de las cosas que es necesario saber es cómo difieren las puntuaciones dentro de la distribución, es decir, cómo están distribuidas independientemente de las medidas de tendencia central. Para ser más específicos, es necesario saber la variabilidad o dispersión de las puntuaciones. Por ejemplo, no es suficiente saber que tres grupos tienen una media de 90 en una prueba, a primera vista parecería que los tres grupos son muy similares, pero si se ve que las puntuaciones mínimas y máximas para cada uno de ellos, es decir, que la amplitud de variación cambia para cada uno, se observa que los tres grupos realmente difieren en su desempeño en la prueba. Por tanto, junto con una calificación promedio para cada grupo, debe determinarse una medida de variabilidad o dispersión respecto a la media. A continuación se presentarán tres de ellas.

Amplitud de Variación.

La amplitud de variación se define como “la diferencia entre la puntuación más alta y la más baja en una distribución, más uno”. (Downie-Heath,1986). Como valor estadístico, la amplitud de variación deja mucho que desear, ya que, en algunos casos, un sólo valor en los extremos de la distribución, da lugar a diferencias apreciables. Es en estos casos donde se deduce que esta medida estadística no es confiable. Dada esta característica, la amplitud de variación no es un valor estadístico muy útil.

En ocasiones, la confiabilidad de la amplitud de variación aumenta al eliminar el 5% superior e inferior de las calificaciones, y luego determinar ésta para el restante 90% central de los valores. Puesto que a esta distribución se le han quitado las puntuaciones más baja y más alta, debe ser más confiable que la amplitud de variación real.

Esta es la medida menos confiable de todas, y se utiliza sólo cuando hay urgencia de obtener una medida de variabilidad. Puede usarse con datos ordinales, de intervalo o de razón.

La desviación estándar y variancia.

Las medidas de variabilidad muestran la distribución o dispersión de las puntuaciones a partir de un punto dado, generalmente la media. El índice que se usa comúnmente, es la desviación estándar. “La desviación estándar (s) es la raíz

cuadrada de la desviación promedio al cuadrado de las calificaciones a partir de la media". (Brown, 1980).

La desviación estándar da un índice de la amplitud de la dispersión de calificaciones en torno a la media; cuanto mayor sea la desviación estándar, tanto más ampliamente estarán dispersas las calificaciones.

Para obtener un conocimiento práctico del significado de la desviación estándar (s), se puede hacer referencia a la curva normal. En una distribución normal existe una relación especificable entre la proporción de casos que caen dentro de ciertos límites de calificaciones y la desviación estándar. En una distribución normal, 68% de las calificaciones caen dentro de más o menos $1s$ a cada lado de la media. Si se miden dos unidades estándares a cada lado de la media, se obtienen dos puntos que delimitan aproximadamente el 95% de área bajo la curva. Cuando se marcan tres unidades de desviación estándar, se incluye más de 99% del área. De hecho, casi 13 casos de cada 10 000 quedarán a cada lado de la media, más allá de los puntos de $+3s$ y $-3s$.

Variación.

La variancia es el cuadrado de la desviación estándar. Es una medida de la cantidad total de variabilidad en un conjunto de calificaciones. La variancia tiene la propiedad importante de la aditividad, que no posee la desviación estándar.

Esta propiedad hace que la variancia sea un procedimiento estadístico útil para conceptualizar ciertas propiedades de las pruebas (tales como la confiabilidad y la validez) y para determinar los efectos de las diversas condiciones de la ejecución en las pruebas.

La desviación estándar (y variancia) se utiliza siempre que una distribución se aproxime a la distribución normal. Es la base de gran parte de la estadística. Como la medida más confiable de variabilidad, se emplea con datos de intervalo y de razón.

Desviación cuartílica.

La *desviación cuartílica* (Q), es otro nombre de la semiamplitud intercuartílica. Por definición, Q1 es el punto que tiene 25% de las puntuaciones por debajo de él en la distribución. En una distribución normal, si se toma la mediana, y se suma y resta una desviación cuartílica a cada lado de ella, se abarca aproximadamente 50% de los casos. Por tanto, 4 desviaciones cuartílicas abarcan toda la amplitud de variación como lo hacen también 6 desviaciones estándares, cuando N es grande.

Esta medida estadística se utiliza cuando la mediana se usa como promedio; esto es, cuando los datos se apartan notablemente de lo normal. Es también la medida que ha de usarse con datos ordinales.

1.7.5 Puntuaciones estándares y la curva normal.

Es muy común que las calificaciones brutas sean transformadas a otras escalas para facilitar el análisis y la interpretación. Estas calificaciones son derivadas o transformadas. Un tipo de calificación derivada particularmente útil es la calificación estándar. Una calificación o puntuación estándar (z) expresa la ejecución de una persona en función de su desviación de la media en unidades de desviación estándar. Dicho en otras palabras, los valores z , como se les conoce comúnmente, indican cuántas unidades de desviación estándar se aparta una puntuación no elaborada a partir de la media.

Debido a que las tres desviaciones estándares a uno y otro lado de la media incluyen prácticamente todos los casos, se deduce que la puntuación z más alta que se observa es 3, y la más baja -3. Es posible describir la distribución de valores z diciendo que tienen una media igual a cero y una desviación estándar de 1. Siempre que se considera una puntuación estándar, se podrá saber exactamente donde está situado un valor en una distribución. Por otro lado, las puntuaciones estándar pueden ser manejadas en forma matemática. Se puede observar también que al cambiar una distribución de puntuaciones comunes a puntuaciones z , no cambia la forma de la distribución original de las puntuaciones.

Curva Normal.

En el siglo XVIII los jugadores profesionales se interesaron en las posibilidades que tenían de ganar en diversos juegos de azar, y pidieron ayuda a algunos matemáticos. DeMoivre (1733, en Downie-Heath, 1986) fue el primero en obtener la ecuación de la curva normal. A principios del siglo XIX Gauss y Laplace ampliaron el desarrollo del concepto de tal curva y el de probabilidad. Casi en la misma época los errores de observación cometidos por los astrónomos quedaron representados por una curva de tal tipo. Hoy la curva normal se conoce también como “curva de error”, “curva de campana”, “curva de Gaus” o “curva de DeMoivre”.

Actualmente ya es bien conocida la forma de esa curva. Su altura máxima está en la media. Se dice también que la ordenada máxima está en la media. Las demás ordenadas son más cortas que ésta. También se dice que la curva normal es asintótica, esto quiere decir que, teóricamente las colas o extremidades de la curva nunca tocan la línea de base, sino que se prolongan indefinidamente en uno y otro sentidos. En la práctica, sin embargo, tres desviaciones estándar a uno y otro lado de la media incluirán prácticamente todos los casos. Como se mencionó anteriormente, la asimetría de la curva normal es cero y por su “agudeza” se describe como mesocúrtica.

En todos los estudios de psicología se supone que ciertas características están distribuidas normalmente. Probablemente ninguna distribución tenga exactamente la forma de la distribución normal. Muchas de las distribuciones de frecuencias se aproximan a la normal, y se supone que tienen la citada distribución normal. En el grado en que las distribuciones difieren de la normal, se presentan los errores. La curva normal es importante no sólo porque las puntuaciones se supongan “normalmente distribuidas”, sino porque las distribuciones de muestreo de diversos valores estadísticos se sabe o se supone que son normales. De aquí que la importancia de la curva normal radica, sobre todo, en los valores estadísticos de muestreo.

Áreas bajo la curva normal

Las áreas bajo la curva normal están establecidas por unidades de desviación estándar, en donde 1 desviación estándar tomada a cada lado de la media incluye un área de 68.26% del área total, o aproximadamente dos tercios de los casos. En términos de probabilidad se puede establecer que las oportunidades de que una puntuación, en una muestra distribuida normalmente, caiga dentro del área de 1 desviación estándar a cada lado de la media es de dos de cada tres. Una segunda desviación estándar después de la primera a cada lado de la media, toma 13.59% de área. Toda el área incluida por dos unidades de desviación estándar a ambos lados de la media, comprenden más del 95% de los casos. Si se continúa, se considera una tercera desviación a cada lado de la media, tomando

otra parte igual a 2.15% del área total. La suma de todas las áreas abarcadas por estas 6 unidades de desviación estándar es igual a 99.74 % del total. De esto se deduce que 0.26% de los casos están más allá de tres desviaciones estándar a partir de la media, lo cual significa 26 casos en 10 000.

Con estas medidas se puede saber con exactitud el número de casos que caen dentro de cierta área debajo de la curva normal, o incluso, la probabilidad de que una puntuación caiga en un área específica. Como se puede ver, la curva normal y sus valores estadísticos, son de suma importancia en la elaboración de pruebas y en la clasificación de los sujetos a partir de ellas.

1.7.6 Comprobación de hipótesis.

Hasta aquí se ha descrito la forma en que los datos de una muestra pueden ser manejados para obtener puntuaciones muy valiosas que a su vez, van a ser de gran utilidad para comprender la naturaleza de la muestra que se está analizando, es decir, se puede conocer a un grupo de personas en función de la distribución de sus puntuaciones, de las medidas de tendencia central pero sobre todo en función de su variabilidad.

Como se explicó anteriormente, se puede tener a dos muestras con medias estadísticas iguales, pero no necesariamente se podría decir que esos grupos son iguales, o que provienen de la misma población, depende mucho de la forma en

que estén distribuidas las puntuaciones. Pero, por otro lado es importante señalar que también cabe la posibilidad de que se tenga a dos grupos con medias diferentes y que provengan de la misma población.

En este punto vale la pena hacer un abordaje más profundo del problema de la diferencia entre dos medias. Estadísticamente se puede decir que es un error considerar que dos grupos fueron tomados de la misma población si tienen medias diferentes, es decir, que en una distribución de sus puntuaciones, la media de alguna de las muestras estuviera desplazada hacia uno u otro lado de la curva normal y ambas medias no coincidieran en el mismo lugar. Sin embargo, hay que tomar en cuenta que existe un rango de tolerancia hacia uno y otro lado de la media en donde las puntuaciones siguen siendo consideradas como promedio. De aquí que es necesario hacer una prueba de hipótesis para determinar si la diferencia entre dos medias es significativa o no, antes de cometer el error de asegurar que provienen de la misma población o que en realidad son diferentes.

“Por otra parte, si los datos de muestra ocasionaron que el investigador rechazara H_0 , esta decisión fue un error. Tal error, rechazar H_0 cuando es verdadera en la población se denomina *error de tipo I*, o *error α* . La probabilidad de cometer un error de tipo I se define como nivel de significancia alfa (α) . . . Los niveles que se encuentran con más frecuencia son 0.05, 0.01 y 0.001. Un nivel α de 0.05 implica que la probabilidad de cometer un error de tipo I sólo por casualidad,

es 5 en 100. . . Cuando α se establece en 0.001, las posibilidades de cometer un error de tipo I se reducen a 1 en 1 000 “. (Downie-Heat, 1986).

Qué pasa si se tienen dos muestras diferentes y se encuentra que ambas tienen medias distintas. La respuesta parece muy obvia, en este caso, se puede casi asegurar que los valores estadísticos encontrados corresponden a poblaciones diferentes. Estadísticamente sería un error considerar a dos grupos como iguales si su media no cae en el mismo punto o si se alejan significativamente.

Por ejemplo, la estatura promedio (media) del hombre mexicano es 1.70 mts y alguien toma una muestra grande de sujetos en algún lugar, los mide y dice que su estatura media es de 1.85 mts. Sería una equivocación considerar que estos otros sujetos forman parte de la misma población, es decir, muy probablemente estos sujetos no sean mexicanos. Continuando, qué pasaría si se clasificara a “los otros sujetos” con los parámetros mexicanos, probablemente estarían en un rango por arriba o muy por arriba del promedio. Sin embargo, si estos sujetos fueran clasificados con una escala propia o que sea para la población a la que pertenecen, probablemente la gran mayoría de ellos sean considerados con una estatura promedio. Lo mismo sucede con la medición de la inteligencia; no se puede considerar que dos grupos de sujetos pertenecen a la misma población si sus medias no corresponden una con la otra o se alejan significativamente entre ellas y por tanto, sería un error clasificarlos con los mismos parámetros.

Para conocer si la diferencia entre dos medias es significativa o no, se debe realizar una prueba z o prueba de hipótesis y su resultado es una buena herramienta para determinar si los grupos pertenecen a la misma población o a una diferente.

CAPITULO II

METODOLOGIA

2.1 Planteamiento del Problema.

Con respecto a la utilización de test psicométricos en México, es muy común encontrar que éstos son aplicados y calificados según los parámetros establecidos para la población con la cual fueron creados.

En muy pocas ocasiones se cuenta con pruebas estandarizadas específicamente con muestras de población Mexicana, lo cual constituye un gran problema, ya que la información ofrecida por los test después de su aplicación y calificación, es decir, el resultado final no es del todo real, ya que los sujetos están siendo comparados con los parámetros de otra población.

A partir de este planteamiento, surgió la necesidad de conocer la veracidad de los resultados arrojados por el test de Dominós de Edgard Anstey, al ser

clasificados los sujetos Mexicanos según los baremos Uruguayos, obtenidos por el Dr. Washington Risso en el año 1957.

La presente investigación pretende determinar si los baremos Uruguayos del test de Dominós utilizados en México, clasifican correctamente a los sujetos Mexicanos según su inteligencia.

De ahí que es necesario saber si existen diferencias significativas entre las puntuaciones medias del nivel de inteligencia de los sujetos Uruguayos y las de los sujetos Mexicanos, a fin de determinar si es necesario elaborar tablas de clasificación (Baremos) específicas para la población Mexicana.

2.2 Hipótesis.

Las hipótesis de la presente investigación son las siguientes:

2.2.1 Hipótesis de trabajo.

H₁: Existen diferencias significativas entre la media de las puntuaciones de los sujetos Uruguayos y la media de las puntuaciones de los sujetos Mexicanos a un nivel de significancia de 0.05

$$H_1: \bar{X}_1 \neq \bar{X}_2$$

2.2.2 Hipótesis Nula.

Ho: No existen diferencias significativas entre la media de las puntuaciones de los sujetos Uruguayos y la media de las puntuaciones de los sujetos Mexicanos a un nivel de significancia de 0.05

$$H_0: \bar{X}_1 = \bar{X}_2$$

2.3 Sujetos

El presente trabajo de investigación se llevó a cabo apegándose en lo posible, a las características de la muestra utilizada para la estandarización del Dominós en Montevideo, Uruguay. Lo anterior con el objeto de evitar la intervención de algunas variables que pudieran afectar los resultados de la prueba al aplicarse en México, y por lo tanto, que se afectaran los resultados de la investigación.

Todos y cada uno de los sujetos de la muestra utilizada para la realización de la investigación aquí presentada debían cubrir, sin excepción alguna, todas las características previamente señaladas por el profesor Washington L. Risso. A continuación se describen las características de la muestra seleccionada para formar parte de la presente investigación:

La muestra fue constituida en su totalidad por sujetos de: (a) nacionalidad Mexicana, (b) cuyas edades fueran de entre 12 años y 30 años 11 meses, (c) de sexo indistinto, (d) con escolaridad mínima de 5º año de primaria y (e) de nivel socioeconómico cultural medio o alto.

El diseño de la muestra es de tipo No probabilístico. Fue tomada de diversas escuelas del Distrito Federal y área metropolitana, desde 6º año de primaria hasta nivel universitario. Se hicieron aplicaciones hasta completar un total de 843 sujetos, según fuera necesario para equilibrar la muestra en cuanto a edad y sexo. El número de grupos y el de sujetos por grupo fue variable, los cuales debían contar con un mínimo de 12 sujetos y un máximo de 30.

Fueron eliminados de la muestra en total 5 sujetos (3 hombres y 2 mujeres), ya que se consideró que habían manipulado la prueba para dar un resultado muy bajo en la misma.

2.4 Escenario

El escenario fue variable, estuvo formado por diversas aulas y salones de clases de escuelas primarias, secundarias, preparatorias y universidades, las cuales tenían unas medidas que fluctuaban entre los 4.5 X 4.5 mts., hasta 8 x 8 mts. Todas ellas debían contar con bancas o pupitres para cada uno de los sujetos. Debían tener adecuada ventilación e iluminación.

En cada aplicación se procedió a dar las indicaciones a los sujetos, tan extensas como fuera necesario, hasta asegurarse de que se habían comprendido las instrucciones. Se concedió a cada sujeto un tiempo de 30 minutos para la realización de la prueba, sin ser totalmente estrictos en el tiempo total de ejecución debido a la naturaleza de la prueba.

2.5 Material

La prueba que se utilizó fue el test de Dominós de Edgard Anstey: cuaderno de problemas, hoja con instrucciones y ejemplos, hoja de respuestas y lápiz con goma.

Hoja de calificación

Tabla de baremos

Calculadora y Computadora (para el manejo estadístico de la información)

2.6 Instrumento

El test de inteligencia de Dominós de Edgard Anstey es un test gráfico no verbal, que mide la inteligencia basándose en el Factor General o factor "G". Consta de 48 problemas, impresos en cuaderno a razón de 6 diseños por página. Los cuadernos contienen una página preliminar donde se encuentran impresas las instrucciones a seguir por el sujeto, así como 4 ejemplos para el mejor entendimiento de la tarea. Cada diseño consta de un grupo de fichas de dominó

que guardan una cierta relación entre sí, y una ficha en blanco para que el examinado señale la cantidad de puntos que corresponde a la solución del problema.

Los 48 Items que forman la prueba están dispuestos en orden de dificultad creciente, contando con 30 minutos para su realización. La prueba se califica otorgando un punto por cada Item contestado correctamente, la sumatoria de las puntuaciones se busca en una tabla de calificación según la edad del sujeto, la cual proporciona el rango al que pertenece el sujeto según su inteligencia. El test de Dominós tiene una validez de 0,550 y un confiabilidad de 0,854.

2.7 Diseño

El diseño de la presente investigación es de tipo *comparativo*, ya que la muestra mexicana fue trabajada como independiente y los resultados fueron comparados con la muestra uruguaya a fin de encontrar diferencias entre ambas.

2.8 Procedimiento

La aplicación de las pruebas comenzó como parte del procedimiento de admisión para escuelas secundarias o preparatorias, y posteriormente, al ver que la distribución de las puntuaciones de los sujetos se agrupaba mayormente en el rango Superior al Término Medio, se decidió iniciar la aplicación de pruebas para

finés de investigación, procurando que la muestra fuera equitativa en edades y sexo de los sujetos, apegándose a la muestra original. La aplicación fue colectiva, otorgándose un tiempo de 30 minutos para contestar el máximo de reactivos posibles, el tiempo comenzaba a contar hasta después de dar las instrucciones y estar seguros de que se habían entendido perfectamente. Al final del tiempo establecido se recogían las hojas de respuesta, otorgando algunos minutos extra a los sujetos que lo solicitaban, considerando que es un test de potencia y no de rendimiento.

Se procedió a contabilizar las respuestas de los sujetos tomando en cuenta el número de ítem contestado acertada o erróneamente, sin embargo, se consideró que dicha información no era de utilidad para la presente investigación, por lo cual no fue tomada en cuenta, aunque podría ser utilizada para otro estudio en particular.

Una vez hecha la aplicación de pruebas al grupo total de sujetos, se realizó la calificación de las hojas de respuestas con las plantillas de corrección contenidas en el manual; de ahí se fue agrupando a los sujetos por rangos de edad y de acuerdo a la puntuación obtenida por cada uno de ellos, dando como resultado la tabla 3.2. Esta primera tabla de distribución se considera la piedra angular de la presente investigación.

También se procedió a la clasificación de cada uno de los 843 sujetos en 5 rangos según el Baremo Montevideo, contenido en el manual. Los rangos que ahí se manejan son: Percentil 5 - Deficiente, Percentil 10 y 25 - Inferior al Término medio, Percentil 50 - Término Medio, Percentil 75 y 90 - Superior al Término Medio, Percentil 95 - Superior. De lo anterior resultó una tabla de distribución de frecuencias de las puntuaciones (Tabla 3.3), la cual proporcionó información de la forma en que se agrupaban los sujetos de acuerdo al rango que les correspondía, dando la pauta para continuar con la investigación al ofrecer datos que indicaban la posibilidad de que las hipótesis eran correctas.

Al ir clasificando a los sujetos, se observó que en algunos casos, un número considerable de ellos se ubicaba dentro del rango "Superior al Término Medio" y no dentro del rango llamado "Término Medio", que sería lo normal, por lo que se procedió a hacer unas tablas de distribución de frecuencias según el rango percentilar de inteligencia al que pertenecían los sujetos, a fin de poder hacer comparaciones visuales claras de la forma en que se distribuían las puntuaciones en la muestra. En un principio, se hicieron dichas tablas tomando en cuenta únicamente los rangos de edad y los rangos percentilares que aparecen en la hoja de corrección de la prueba de Dominós, es decir, 12-13, 14-15, 16-17, 18-30 años y población general por un lado, y por otro se agruparon según el percentil 5, 10, 25, 50, 75, 90 y 95; sin embargo, los datos obtenidos en la tabla 3.3, dieron la pauta para que se hicieran nuevas distribuciones. En estas otras se tomaron en cuenta

nuevos rangos percentilares y agrupaciones diferentes por edad, incluso divididos por edad y sexo. (Tablas de la 3.4 a la 3.8)

Al hacerse evidente que las distribuciones cambiaban, es decir, que no se agrupaban del centro hacia los extremos, según las medias obtenidas para la población uruguaya, surgió la inquietud de conocer cómo estaban las puntuaciones medias de la muestra (\bar{X}) o muestras mexicanas, y qué tanto variaban con respecto a lo observado con los uruguayos.

Para lograrlo fue necesario obtener algunos datos estadísticos de la muestra mexicana. Cabe señalar que ésta tiene un total de 843 sujetos y fue dividida en subgrupos o muestras más pequeñas para fines de estudio y para hacer una buena comparación de los datos que permitiera llegar a conclusiones más contundentes y claras. Con las puntuaciones de los sujetos, se hicieron diferentes grupos: (a) divididos por edad, (b) por edad y sexo al mismo tiempo y (c) por rangos de edad, obteniendo para cada uno de ellos la media, desviación estándar, varianza y el error estándar de la media; estas mismas medidas se obtuvieron también para la población general (843 sujetos). Los resultados están presentados en la tabla 3.9.

Después de realizar todas estas comparaciones, era evidente que las muestras mexicana y la uruguaya no eran iguales, tenían medias diferentes y su distribución también era distinta. Sin embargo, se debía establecer si la diferencia era lo suficientemente significativa para asegurar que ambas muestras provenían

de poblaciones distintas y poder así elaborar y proponer baremos específicos para la población mexicana.

A fin de poder establecer una diferencia precisa entre las puntuaciones observadas entre los grupos, se procedió a realizar una prueba de hipótesis o prueba z . Se utilizó ésta debido a que es la que se usa como prueba de significancia de la diferencia entre dos medias, usando datos independientes y muestras grandes. No se uso la t de Student o razón t debido a que ésta se usa para muestras pequeñas. Como se puede observar en el apéndice A, el uso de esta fórmula requiere de conocer previamente el error estándar de la media y el error estándar de la diferencia entre las medias, ambas fórmulas se muestran en el mismo apéndice que las anteriores.

Las pruebas z se hicieron básicamente entre los diferentes grupos de las muestras mexicana y uruguaya (tabla 3.10), aunque se hicieron también varias comparaciones entre los diversos grupos de la misma muestra mexicana, a fin de determinar si existían diferencias entre las medias de dichos grupos y poder establecer así la forma en que se debían hacer los baremos, si es que los grupos eran iguales o diferentes entre si. Cabe recordar que para poder considerar que ambos grupos eran diferentes, era necesario observar una puntuación z mayor a 1.96 o menor que -1.96, a un nivel de significancia de 0.05. Los resultados de las pruebas de hipótesis se describirán más adelante.

Por último, se procedió a la realización de los baremos mexicanos para la clasificación de los sujetos mexicanos según su inteligencia. Lo anterior se realizó mediante la obtención de las puntuaciones correspondientes para cada centil (con los datos de la tabla 3.2), según los rangos de edad descritos anteriormente y para la población general. Los baremos propuestos se muestran en el Apéndice A.

Una vez obtenidos los baremos se procedió a obtener la amplitud de los rangos para la clasificación de los sujetos de acuerdo a su edad y a la puntuación obtenida. La tabla con dicha amplitud de los rangos se presenta más adelante. (Apéndice B)

A continuación se presentan los resultados obtenidos a través del manejo estadístico de la información recabada, para posteriormente pasar al análisis de datos y presentación de las conclusiones de la presente investigación.

CAPITULO III

RESULTADOS Y SU ANÁLISIS

3.1 RESULTADOS

En esta sección se procederá a presentar los resultados obtenidos de la aplicación de la prueba de Dominós a una muestra mexicana de 843 sujetos, así como la agrupación de dichos resultados para fines de estudio.

En primer lugar, en la tabla 3.1 se observa la forma en que la muestra fue distribuida de acuerdo a la edad y sexo de los sujetos, incluso por rangos de edad, observándose que de la muestra total que corresponde a 843 sujetos, 401 eran hombres y 442 mujeres. Además se puede apreciar que el número de sujetos para cada rango de edad es muy parecido, a excepción del rango I que tiene un número menor de sujetos. Por otra parte, lo que si llama la atención es que la distribución de los sujetos por sexo es ligeramente irregular, aunque esta diferencia entre el

número de hombres y mujeres en las muestras individuales no representa ninguna diferencia significativa que pudiera alterar los resultados de la investigación.

TABLA 3.1
DISTRIBUCIÓN DE LOS SUJETOS POR EDAD Y SEXO

EDAD	HOMBRES	MUJERES	TOTAL
12 AÑOS	22	50	72
13 AÑOS	27	37	64
12-13 AÑOS	49	87	136
14 AÑOS	75	61	136
15 AÑOS	74	75	149
14-15 AÑOS	149	136	285
16 AÑOS	71	51	122
17 AÑOS	39	51	90
16-17 AÑOS	110	102	212
18 AÑOS	17	19	36
19 AÑOS	30	40	70
20+ AÑOS	46	58	104
18-20+ AÑOS	93	117	210
POB. GENERAL	401	442	843

Posteriormente, en la tabla 3.2, se observa la forma en que se distribuyeron las puntuaciones de los sujetos, según el rango de edad al que pertenecían y según la población general. Aquí se hace notar como las puntuaciones de cada grupo se concentran más en los rangos de puntuaciones que van de 25 a 30 aciertos y de 31 a 36. Incluso en la población general, un mayor número de sujetos (236) se agrupa entre las puntuaciones que van de 31 a 36 aciertos. Cabe señalar que aquí lo esperado era que en todos los rangos de edad (excepto el de 18 a 30 años), la

mayor concentración de sujetos estuviera entre las puntuaciones de 25 a 30, ya que son estas las que corresponden a la media en los diferentes rangos de edad según el Baremo Montevideo propuesto por el profesor Washington Risso. Dicha observación es de gran importancia ya que es a partir de ahí que surge el interés en investigar qué tanta diferencia existía entre la muestra mexicana y la uruguaya; por qué no se distribuía uniformemente la muestra mexicana y lo más importante, saber si la diferencia encontrada entre las medias de ambas muestras era lo suficientemente significativa como para hacer tablas de clasificación propias para los sujetos mexicanos.

TABLA 3.2
DISTRIBUCION DE FRECUENCIAS DE LAS PUNTUACIONES

PUNTUACION	12-13 AÑOS	14-15 AÑOS	16-17 AÑOS	18-30 AÑOS	POBL. GRAL
43-48	0	1	6	15	22
37-42	10	36	49	56	151
31-36	43	69	80	71	263
25-30	44	96	42	43	225
19-24	22	49	20	19	110
13-18	11	24	8	5	48
7-12	5	9	6	0	20
1-6	1	1	1	1	4

En la tabla 3.3 encontramos la distribución de los sujetos según el percentil y por rangos de edad, observándose una distribución bastante uniforme.

TABLA 3.3

**DISTRIBUCIÓN DE LOS SUJETOS SEGÚN EL PERCENTIL
Y POR RANGOS DE EDAD**

Rango	Percentil	12 – 13	14 - 15	16 - 17	18 - 30	P.G.
SUP.	95	8	22	14	27	71
S.T.M	90	12	15	18	18	63
S.T.M	75	23	42	65	36	166
T.M	50	66	134	80	95	375
I.T.M	25	16	55	25	25	121
I.T.M	10	7	8	3	6	24
DEF.	5	4	9	7	3	23
		136	285	212	210	843

Aquí se puede observar cómo los sujetos se van agrupando uniformemente en el centro de la tabla, es decir, en el percentil 50 que corresponde a la media; sin embargo, los resultados contrastan con los de la tabla anterior (3.2), lo cual hizo que surgiera la duda de los resultados llevando a hacer nuevas tablas de distribución cada vez más detalladas, las cuales se muestran a continuación.

TABLA 3.4

**DISTRIBUCIÓN DE LOS SUJETOS SEGÚN EL PERCENTIL
Y POR RANGOS DE EDAD**

Rango	Percentil	12 -13	14 - 15	16 - 17	18 - 30	P.G.
	99	2	1	2	5	10
SUP.	95	6	21	12	22	61
S.T.M	90	12	15	18	18	63
	80	13	31	50	26	120
S.T.M	75	10	11	15	10	46
	70	10	14	17	10	51
	60	14	41	21	38	114
T.M	50	21	45	22	24	112
	40	16	23	16	17	72
	30	5	11	4	6	26
I.T.M	25	4	9	11	9	33
	20	12	46	14	16	88
I.T.M	10	7	7	3	6	23
DEF.	5	4	9	6	2	21
	1	0	1	1	1	3
		136	285	212	210	843

En la tabla 3.4 se puede observar claramente que la distribución cambió drásticamente al tomar en cuenta un mayor número de percentiles para hacer la tabla. Aquí es evidente que los sujetos tienen una distribución más irregular y además no se agrupan en el percentil 50, que es lo esperado. Por ejemplo, se puede mencionar que para la población general un número mayor de sujetos (120) se agrupa mucho más arriba del término medio, es decir, se concentró más en el percentil 80, lo cual dista mucho de ser lo esperado.

A continuación se puede observar la forma en que las distribuciones se van haciendo más irregulares a medida que se van haciendo más específicas, hasta llegar al punto donde la tabla comprende 14 percentiles y su distribución es de acuerdo a la edad de los sujetos e incluso por sexo.

En la tabla 3.5 se observa nuevamente una distribución más uniforme, aunque en la edad de 17 años hubo un mayor número de sujetos agrupados en el rango Superior al Término Medio (S.T.M.) que en el de Término Medio.

En la siguiente tabla (3.6) se hace más evidente esta discrepancia, observándose incluso en la edad de 20 años y más, que un mayor número de sujetos se agruparon en el percentil 95 (Superior) en lugar del percentil 50 (Término Medio).

TABLA 3.5
DISTRIBUCIÓN DE LOS SUJETOS SEGÚN EL PERCENTIL
Y POR EDADES

Rango	Percentil	12	13	14	15	16	17	18	19	20+	P.G.
SUP.	95	3	5	11	11	7	7	3	5	19	71
S.T.M	90	4	8	5	10	6	12	1	6	11	63
S.T.M	75	12	11	21	21	34	31	8	6	22	166
T.M	50	35	31	66	68	50	30	19	36	40	375
I.T.M	25	10	6	23	32	20	5	4	13	8	121
I.T.M	10	5	2	6	2	2	1	0	4	2	24
DEF.	5	3	1	4	5	3	4	1	0	2	23
		72	64	136	149	122	90	36	70	104	843

TABLA 3.6

**DISTRIBUCIÓN DE LOS SUJETOS SEGÚN EL PERCENTIL
Y POR EDADES**

Rango	Percentil	12	13	14	15	16	17	18	19	20+	P.G.
SUP.	99	2	0	1	0	1	1	0	3	2	10
SUP.	95	1	5	10	11	6	6	3	2	17	61
S.T.M	90	4	8	5	10	6	12	1	6	11	63
S.T.M.	80	7	6	14	17	24	26	6	4	16	120
S.T.M	75	5	5	7	4	10	5	2	2	6	46
	70	4	6	7	7	8	9	3	5	2	51
	60	5	9	20	21	14	7	8	22	8	114
T.M	50	13	8	21	24	18	4	3	5	16	112
	40	10	6	10	13	8	8	5	4	8	72
	30	3	2	8	3	2	2	0	0	6	26
I.T.M	25	3	1	3	6	10	1	2	1	6	33
I.T.M.	20	7	5	20	26	10	4	2	12	2	88
I.T.M	10	5	2	5	2	2	1	0	4	2	23
DEF.	5	3	1	4	5	3	3	0	0	2	21
DEF.	1	0	0	1	0	0	1	1	0	0	3
		72	64	136	149	122	90	36	70	104	843

En la tabla 3.7 se observa nuevamente una distribución uniforme de los sujetos a pesar de que se dividió el grupo total en diferentes grupos de edad, incluyendo una división por sexo. Aquí se hace notar claramente que en este caso la distribución no depende de su especificación en edad o sexo, sino más bien depende de la amplitud de los rangos percentilares que se toman en cuenta.

TABLA 3.7

DISTRIBUCIÓN DE LOS SUJETOS SEGÚN EL PERCENTIL,
EDAD Y SEXO

Rango	Percentil	12 H	12 M	13H	13 M	14 H	14 M	15 H	15 M	16 H	16 M	17 H	17 M	18 H	18 M	19 H	19 M	20+ H	20+ M	P G
SUP.	95	1	2	1	4	8	3	5	6	5	2	2	5	1	2	3	2	9	10	71
S.T. M	90	1	3	4	4	2	3	4	6	4	2	3	9	0	1	2	4	3	8	63
S.T. M	75	1	11	7	4	10	11	12	9	20	14	15	16	2	6	3	3	8	14	166
T.M	50	10	25	12	19	36	30	34	34	27	23	15	15	11	8	14	22	20	20	375
I.T.M	25	5	5	3	3	13	10	16	16	12	8	2	3	3	1	6	7	4	4	121
I.T.M	10	3	2	0	2	4	2	0	2	1	1	1	0	0	0	2	2	0	2	24
DEF.	5	1	2	0	1	2	2	3	2	2	1	1	3	0	1	0	0	2	0	23
		22	50	27	37	75	61	74	75	71	51	39	51	17	19	30	40	46	58	843

En la tabla 3.8 se corrobora la observación hecha anteriormente, en donde se hace mención de que la diferencia en la distribución depende de la amplitud de los rangos percentilares que se toman en cuenta y no de la edad o sexo de los sujetos. En esta tabla se vuelve a ver que un mayor número de sujetos se agrupa en el percentil 80 en las edades de 16 años hombres, 17 años de ambos sexos, 18 años mujeres, 20+ años y en la población general. Llama la atención que en el grupo de 15 años mujeres los sujetos se agruparon más en el percentil 20.

TABLA 3.8

DISTRIBUCIÓN DE LOS SUJETOS SEGÚN EL PERCENTIL,
EDAD Y SEXO

Rango	Percentil	12 H	12 M	13H M	13 M	14 H	14 M	15 H	15 M	16 H	16 M	17 H	17 M	18 H	18 M	19 H	19 M	20+ H	20+ M	P G
SUP.	99	1	1	0	0	0	1	0	0	1	0	1	0	0	0	1	2	0	2	10
SUP.	95	0	1	1	4	8	2	5	6	4	2	1	5	1	2	2	0	9	8	61
S.T.M	90	1	3	4	4	2	3	4	6	4	2	3	9	0	1	2	4	3	8	63
S.T.M.	80	1	6	3	3	6	8	10	7	15	9	13	13	1	5	3	1	4	12	120
S.T.M	75	0	5	4	1	4	3	2	2	5	5	2	3	1	1	0	2	4	2	46
	70	0	4	3	3	4	3	4	3	4	4	5	4	1	2	2	3	0	2	51
	60	0	5	4	5	11	9	9	12	5	9	4	3	6	2	11	11	2	6	114
T.M	50	4	9	2	6	13	8	13	11	12	6	2	2	1	2	0	5	12	4	112
	40	4	6	1	5	6	4	7	6	5	3	2	6	3	2	1	3	2	6	72
	30	2	1	2	0	2	6	1	2	1	1	2	0	0	0	0	0	4	2	26
I.T.M	25	2	1	1	0	2	1	4	2	9	1	1	0	1	1	0	1	2	4	33
I.T.M.	20	3	4	2	3	11	9	12	14	3	7	1	3	2	0	6	6	2	0	88
I.T.M	10	3	2	0	2	3	2	0	2	1	1	1	0	0	0	2	2	0	2	23
DEF.	5	1	2	0	1	3	1	3	2	2	1	1	2	0	0	0	0	2	0	21
DEF.	1	0	0	0	0	0	1	0	0	0	0	0	1	0	1	0	0	0	0	3
		22	50	27	37	75	61	74	75	71	51	39	51	17	19	30	40	46	58	843

En la siguiente tabla (3.9) se pueden observar las puntuaciones de distribución y variabilidad para los diferentes grupos de edad en los que se dividió la muestra, incluidos los grupos por edad y sexo, así como para los distintos rangos y para la muestra total, a la cual se le llamará Población General.

En la primera parte de la tabla, se observan las puntuaciones obtenidas para el grupo de edad de 12 años (hombres y mujeres), 13 años (hombres y mujeres), así como el Rango I, que incluye las dos edades sin distinción de sexo.

Las medidas y puntuaciones fueron las siguientes:

N = Número de sujetos

\bar{X} = Media aritmética de las puntuaciones de la muestra

S = Desviación estándar

S^2 = Varianza

$S\bar{x}$ = Error estándar de la media

TABLA 3.9

PUNTUACIONES DE DISTRIBUCIÓN Y VARIABILIDAD

12 AÑOS HOMBRES	12 AÑOS MUJERES	12 AÑOS TOTAL
$N = 22$	$N = 50$	$N = 72$
$\bar{X} = 23.04$	$\bar{X} = 27.52$	$\bar{X} = 26.15$
$S = 7.8$	$S = 6.84$	$S = 7.15$
$S^2 = 60.86$	$S^2 = 46.81$	$S^2 = 51.11$
$S\bar{x} = 1.66$	$S\bar{x} = 0.9673$	$S\bar{x} = 0.8426$

13 AÑOS HOMBRES	13 AÑOS MUJERES	13 AÑOS TOTAL
$N = 27$	$N = 37$	$N = 64$
$\bar{X} = 29.77$	$\bar{X} = 28.21$	$\bar{X} = 28.88$
$S = 5.27$	$S = 5.25$	$S = 6.5$
$S^2 = 5.27$	$S^2 = 7.25$	$S^2 = 6.5$
$S\bar{x} = 1.014$	$S\bar{x} = 1.2$	$S\bar{x} = 0.8125$

RANGO I (12-13 AÑOS)

$N = 136$
 $\bar{X} = 27.43$
 $S = 6.85$
 $S^2 = 46.9$
 $S\bar{x} = 0.5874$

Como se puede apreciar en la tabla, lo más importante que hay que resaltar es la forma en que cambian las puntuaciones medias de los diferentes grupos, las cuales van de $\bar{X}= 26.15$ para el grupo de 12 años total, hasta $\bar{X}= 28.88$ para el grupo de 13 años total, y con una media de $\bar{X}= 27.43$ para el Rango I, que incluye ambas edades. En el siguiente capítulo se explicarán más a fondo estas puntuaciones y se verá si estas diferencias son significativas o no.

En la siguiente parte de la tabla 3.9 se pueden observar las mismas puntuaciones pero ahora para el Rango II, que incluye las edades de 14 y 15 años. De igual forma, se obtuvieron dichas medidas para los diferentes grupos de edad divididos también por sexo.

Aquí se puede observar también la forma en que cambia la media para cada grupo de edad, aunque sin lugar a duda, este es un grupo con características más uniformes, desde el número de sujetos por grupo, hasta las medias aritméticas parciales y totales.

TABLA 3.9 Cont.

PUNTUACIONES DE DISTRIBUCIÓN Y VARIABILIDAD

14 AÑOS HOMBRES	14 AÑOS MUJERES	14 AÑOS TOTAL
$N = 75$	$N = 61$	$N = 136$
$X = 28.04$	$X = 27.52$	$X = 27.8$
$S = 7.72$	$S = 7.5$	$S = 7.62$
$S^2 = 59.58$	$S^2 = 56.22$	$S^2 = 58.07$
$S_{\bar{x}} = 0.8914$	$S_{\bar{x}} = 0.9603$	$S_{\bar{x}} = 0.6534$

15 AÑOS HOMBRES	15 AÑOS MUJERES	15 AÑOS TOTAL
$N = 74$	$N = 75$	$N = 149$
$\bar{X} = 28.12$	$\bar{X} = 28.01$	$\bar{X} = 28.06$
$S = 7.31$	$S = 7.71$	$S = 7.517$
$S^2 = 53.51$	$S^2 = 59.45$	$S^2 = 56.5$
$S_{\bar{x}} = 0.85$	$S_{\bar{x}} = 0.89$	$S_{\bar{x}} = 0.616$

RANGO II 14-15 AÑOS
$N = 285$
$\bar{X} = 27.94$
$S = 7.566$
$S^2 = 57.25$
$S_{\bar{x}} = 0.448$

La siguiente parte de la misma tabla está conformada por las puntuaciones obtenidas por los sujetos que integran el Rango III, es decir, aquellos sujetos con edades de 16 y 17 años. De igual forma, se hicieron las divisiones por sexo para cada edad y se obtuvieron las mismas puntuaciones.

Aquí se puede observar nuevamente que las medias de los grupos son más irregulares, además de que han cambiado radicalmente con respecto al Rango I. En este Rango III se obtuvo una puntuación $\bar{X} = 31.45$, contra una $\bar{X} = 27.94$ para el Rango I. Inclusive para el mismo Rango III existen diferencias entre las medias de cada edad que lo conforma, observándose una $\bar{X} = 30.6$ para 16 años total y una $\bar{X} = 32.62$ para 17 años total. En el siguiente capítulo se expondrá si estas diferencias son significativas o no.

TABLA 3.9 (Cont.)

PUNTUACIONES DE DISTRIBUCIÓN Y VARIABILIDAD (Cont.)

16 AÑOS HOMBRES	16 AÑOS MUJERES	16 AÑOS TOTAL
$N = 71$	$N = 51$	$N = 122$
$\bar{X} = 30.83$	$\bar{X} = 30.27$	$\bar{X} = 30.6$
$S = 7.5$	$S = 6.87$	$S = 7.25$
$S^2 = 56.28$	$S^2 = 47.25$	$S^2 = 52.51$
$Sx = 0.89$	$Sx = 0.962$	$Sx = 0.6564$

17 AÑOS HOMBRES	17 AÑOS MUJERES	17 AÑOS TOTAL
$N = 39$	$N = 51$	$N = 90$
$\bar{X} = 32.43$	$\bar{X} = 32.76$	$\bar{X} = 32.62$
$S = 6.92$	$S = 8.269$	$S = 7.71$
$S^2 = 47.89$	$S^2 = 68.376$	$S^2 = 59.5$
$Sx = 1.11$	$Sx = 1.16$	$Sx = 0.813$

RANGO III (16-17 AÑOS)
$N = 212$
$\bar{X} = 31.45$
$S = 7.45$
$S^2 = 55.48$
$Sx = 0.5117$

A continuación se muestran los resultados obtenidos para el Rango IV, que agrupa a los sujetos cuyas edades iban de 18 a 30 años. Anteriormente se había hecho otra agrupación, en la cual el rango IV concentraba a los sujetos de 18 y 19 años y un último rango V, agrupaba a los sujetos de 20 y más años. Sin embargo, al ir haciendo el análisis estadístico, se observó que la diferencia entre ambos rangos no era significativa y además se tomó en cuenta que era más sencillo

conservar las agrupaciones propuestas por el profesor Washington Risso para las tablas uruguayas.

En este segmento de la tabla 3.9 se puede observar mayor variabilidad en las diferentes medidas, desde el número de sujetos totales para 18 y 19 años, la media de los sujetos de 18 años (hombres y mujeres), hasta la desviación estándar de éstos últimos. Más adelante se hará una revisión de estas puntuaciones para analizar si estas diferencias son significativas o no.

18 AÑOS HOMBRES	18 AÑOS MUJERES	18 AÑOS TOTAL
$N = 17$	$N = 19$	$N = 36$
$\bar{X} = 31.53$	$\bar{X} = 33.63$	$\bar{X} = 32.64$
$S = 5.97$	$S = 8.15$	$S = 7.2$
$S^2 = 35.66$	$S^2 = 66.44$	$S^2 = 51.9$
$S_{\bar{x}} = 1.45$	$S_{\bar{x}} = 1.87$	$S_{\bar{x}} = 1.2$

19 AÑOS HOMBRES	19 AÑOS MUJERES	19 AÑOS TOTAL
$N = 30$	$N = 40$	$N = 70$
$\bar{X} = 32.17$	$\bar{X} = 31.975$	$\bar{X} = 32.057$
$S = 7.19$	$S = 6.62$	$S = 6.87$
$S^2 = 51.67$	$S^2 = 43.82$	$S^2 = 47.188$
$S_{\bar{x}} = 1.313$	$S_{\bar{x}} = 1.047$	$S_{\bar{x}} = 0.821$

En esta parte de la tabla, se puede recalcar la diferencia entre las medias del grupo de 20 años, las cuales son de $\bar{X} = 32.85$ para los hombres y $\bar{X} = 34.97$ para las mujeres. Como se ha mencionado, el análisis de los datos se hará más adelante.

20+	AÑOS HOMBRES	20+	AÑOS MUJERES	20+	AÑOS TOTAL
$N = 46$		$N = 58$		$N = 104$	
$X = 32.85$		$X = 34.97$		$X = 34.03$	
$S = 6.94$		$S = 6.61$		$S = 6.76$	
$S^2 = 48.22$		$S^2 = 43.688$		$S^2 = 45.69$	
$Sx = 1.02$		$Sx = 0.868$		$Sx = 0.663$	

RANGO IV = 18-20+ AÑOS	
$N = 210$	
$X = 33.133$	
$S = 6.87$	
$S^2 = 47.2$	
$Sx = 0.474$	

Por último, en la parte final de la tabla 3.9, se muestran las puntuaciones para la población general, la cual integra a la muestra total de 843 sujetos, divididos por sexo. Aquí se puede observar que el grupo total es bastante uniforme tanto en el número de sujetos, como en la media, desviación estándar y error estándar de la media.

TABLA 3.9 (Cont.)

PUNTUACIONES DE DISTRIBUCIÓN Y VARIABILIDAD (Cont.)

POBLACIÓN GENERAL		
HOMBRES	MUJERES	TOTAL
$N = 401$	$N = 442$	$N = 843$
$X = 29.83$	$X = 30.23$	$X = 30.04$
$S = 7.2$	$S = 7.31$	$S = 7.26$
$S^2 = 51.76$	$S^2 = 53.45$	$S^2 = 52.29$
$Sx = 0.36$	$Sx = 0.347$	$Sx = 0.025$

La siguiente tabla que se mostrarán los resultados es la 3.10, en la cual están plasmados los puntajes obtenidos en la prueba z para diversos grupos de comparación. En la primera parte se pueden observar los puntajes obtenidos al hacer las diversas comparaciones de los grupos por rangos de edad entre Uruguay y México. Cabe recalcar que todos los puntajes son negativos y que van desde $Z = -2.11$ hasta $Z = -8.58$, mostrando claramente que hay diferencias. El significado de dichas diferencias se explica en el capítulo siguiente.

TABLA 3.10

COMPROBACIÓN DE HIPÓTESIS

PRUEBA Z POR RANGOS ENTRE URUGUAY Y MÉXICO

RANGO I (12-13 AÑOS) URUGUAY Vs MÉXICO	Z= -2.23
RANGO II (14-15 AÑOS) URUGUAY Vs MÉXICO	Z= -2.11
RANGO III (16-17 AÑOS) URUGUAY Vs MÉXICO	Z= -5.5
RANGO IV (18-30 AÑOS) URUGUAY Vs MÉXICO	Z= -4.84
RANGO V (POBLACIÓN GENERAL) URUGUAY Vs MÉXICO	Z= -8.58

La siguiente parte de la tabla muestra los resultados obtenidos para la misma prueba z, sólo que ahora las comparaciones fueron hechas entre sexos, por edades, pero con la misma población mexicana. Aquí se puede observar que las puntuaciones son muy chicas, algunas positivas y otras negativas, las cuales van

de $Z= 0.09$ a $Z= -2.33$. Las puntuaciones y su significado se explicarán en el siguiente capítulo.

TABLA 3.10.1

PRUEBA Z PARA LA POBLACIÓN MEXICANA POR EDADES

HOMBRES Vs. MUJERES

GRUPO DE 12 AÑOS HOMBRES Vs. 12 AÑOS MUJERES	$Z= -2.33$
GRUPO DE 13 AÑOS HOMBRES Vs. 13 AÑOS MUJERES	$Z= 0.997$
GRUPO DE 14 AÑOS HOMBRES Vs. 14 AÑOS MUJERES	$Z= 0.397$
GRUPO DE 15 AÑOS HOMBRES Vs. 15 AÑOS MUJERES	$Z= 0.09$
GRUPO DE 16 AÑOS HOMBRES Vs. 16 AÑOS MUJERES	$Z= 0.427$
GRUPO DE 17 AÑOS HOMBRES Vs. 17 AÑOS MUJERES	$Z= -0.206$
GRUPO DE 18 AÑOS HOMBRES Vs. 18 AÑOS MUJERES	$Z= -0.89$
GRUPO DE 19 AÑOS HOMBRES Vs. 19 AÑOS MUJERES	$Z= 0.101$
GRUPO DE 20+ AÑOS HOMBRES Vs. 20+ AÑOS MUJERES	$Z= -1.58$
POB. GENERAL HOMBRES Vs. POB. GENERAL MUJERES	$Z= -0.8$

3.2 ANÁLISIS DE RESULTADOS

Después de haber realizado una presentación de los datos y resultados obtenidos a lo largo de la investigación, en el presente apartado se procederá a hacer el debido análisis de los datos, abordando el *cómo* y el *por qué* de algunas

comparaciones, hasta llegar a mostrar *qué* fue lo que se encontró en forma de conclusión.

A manera de referencia, cabe señalar que el presente trabajo representa un esfuerzo por demostrar las diferencias individuales que existen entre los diversos grupos étnicos, en este caso en particular, entre los sujetos uruguayos y los mexicanos, utilizando como punto de comparación el nivel de inteligencia mostrado por unos y otros a través de la prueba de Dominós del Dr. Edgard Anstey (1955) (Baremo uruguayo por el profesor Washington L. Riso, 1955). Por tanto, esta investigación tiene el afán de enfatizar la importancia del uso correcto de las pruebas psicométricas y la relevancia que tiene su estandarización y/o tipificación para la muestra en que pretende ser (y ha sido) utilizada.

De hecho, la psicometría no tendría razón de ser si no siguiera estrictamente las reglas de estandarización y validación para sacar una nueva prueba al mercado, las cuales es sabido que se hacen en función de la población con la que se obtuvo la muestra. Siguiendo con esta línea y siendo tal vez muy afanosos al respecto, el hecho es que ninguna comparación de puntuaciones hecha en función de una medición, tendría razón de ser si no estuviera basada en parámetros previamente establecidos, los cuales permitan ubicar perfectamente dicha puntuación por arriba o por abajo del promedio esperado. La medición ha existido casi desde tiempos inmemorables y su éxito reside precisamente en que está basada en normas paramétricas que permiten especificar el *más que* y el *menos*

que en función de la población en cuestión. En este caso en particular la población a la que se hace referencia es la mexicana, y lo que se pone en tela de juicio son precisamente las normas paramétricas utilizadas para establecer las diferencias entre los sujetos.

Pasando directamente al análisis de los datos recabados, es importante hacer mención de que el grupo o muestra total utilizado para la realización de este trabajo, fue en ocasiones subdividido en grupos más pequeños para fines de investigación, tomando en cuenta la edad y/o sexo de los sujetos, a fin de poder hacer comparaciones más detalladas que permitieran llegar a conclusiones más precisas y contundentes. Por lo anterior en ocasiones los datos, su presentación o uso no corresponden con la forma típica de hacerlo desde un punto de vista meramente estadístico.

Se comienza el análisis con la referencia de la tabla 3.1, donde es bastante notorio que la cantidad de hombres y mujeres en la población general (843 sujetos) es sumamente homogéneo, existiendo únicamente una diferencia de 41 individuos más a favor de las mujeres, lo cual indica que los resultados que arroje el presente trabajo, pueden ser generalizados para ambos sexos debido a que ninguno de ellos tiene un peso estadísticamente mayor con respecto al otro. En este caso, la diferencia significa el 4.8 % de los casos totales. En la estandarización del Dr. Risso (1955) se desconoce la diferencia entre el número de hombres y mujeres. El análisis de la diferencia entre sexos es un tema que genera mucha

polémica, y en esta investigación por supuesto no se ha dejado de lado, aunque por fines prácticos, de orden y continuidad, el análisis de los resultados se muestra más adelante.

La siguiente parte del análisis corresponde a los datos de la tabla 3.2, en donde se puede ver claramente que las puntuaciones de los sujetos se agrupan más en el rango siguiente superior al esperado, esto es, por arriba del rango donde se supone se encuentran las medias de los diferentes grupos de edad. Esta observación fue contundente en la investigación, ya que es a partir de esta diferencia en las distribuciones que se inició un proceso científico de investigación como tal. Lo anterior basado en el supuesto de que no es estadísticamente correcto que los sujetos se agrupen en un rango que no les corresponde, a menos que hubiera algún error de muestreo o que los sujetos pertenecieran a poblaciones diferentes. A manera de ejemplo y buscando una forma de justificar la realización de la investigación, se puede mencionar que sería un error que la embajada de Alemania hiciera una medición de la altura de algunos mexicanos y dijera que todos ellos se encuentran por debajo del promedio, incluso los que miden 1.70 mts; lo que les haría falta decir es que se están basando en una altura promedio de 1.75 mts de los alemanes y no en el 1.70 de los mexicanos. Lo mismo sucedería si viene un ciudadano de Mongolia que mide 1.65 y decimos que su estatura está por debajo del promedio; igualmente haría falta expresar que es con respecto a la media mexicana y no a la media de Mongolia. Pues lo mismo sucede con la inteligencia, existen parámetros y normas para cada población y sería erróneo

hacer comparaciones de los sujetos de un lugar con los parámetros de otro. De ahí fue que surgió la inquietud de verificar si las tablas uruguayas eran aplicables a los sujetos mexicanos.

Para tratar de indagar más sobre la naturaleza de las diferencias, se realizó una tabla de frecuencias similar a la de la tabla 3.2, sólo que ahora se basaba en el rango percentilar al que pertenecían las puntuaciones. Los datos obtenidos en la tabla 3.3 fueron ciertamente desconcertantes al ver que las puntuaciones se agruparon en el percentil 50 (como es lo esperado), sin embargo, esto tiene su explicación en el hecho de que el rango en el que se mueven las puntuaciones para considerarse como término medio es muy amplio, por lo tanto, hubiera sido realmente extraño que las puntuaciones de los sujetos se agruparan en un percentil diferente al 50.

Este hallazgo hubiese sido suficiente para abandonar la investigación y dedicarse a encontrar cualquier otra cosa interesante en el ramo de la psicología, sin embargo, la forma de agruparse las puntuaciones en la tabla 3.2 distaba mucho de ser "normal", y por lo tanto, era un buen argumento para seguir buscando una explicación. Debido a lo anterior, se hicieron nuevas tablas, cada vez más específicas, y el resultado fue muy favorecedor para la investigación, ya que en las tablas 3.4, 3.6 y 3.8, se logra apreciar que las puntuaciones de los sujetos ya no se agrupan en el centro de la tabla, sino que tienden más bien a separarse del centro.

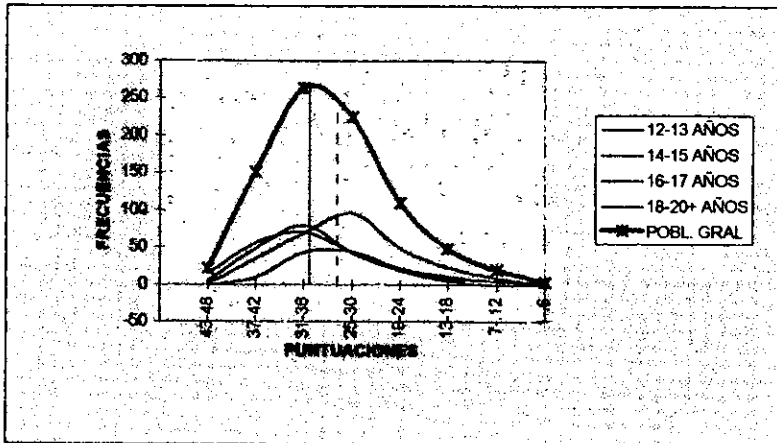
Se puede apreciar también que la forma en que se desplazan es más drástica en las edades mayores a 16 años, siendo mayormente homogénea entre las edades de 12 a 15 años, independientemente del sexo.

De igual forma, a lo largo de la tabla 3.9 se puede apreciar la manera en que las medidas de tendencia central, así como las de distribución van cambiando dependiendo de la edad de los sujetos o incluso del sexo; dichas variaciones son evidentes dentro de la misma tabla 3.9, aunque al compararlas con las medidas y puntuaciones de la población uruguaya también se pueden apreciar diferencias. Conforme iba avanzando la investigación era más evidente y estadísticamente más probable que las diferencias entre ambas poblaciones fueran significativas.

A continuación se presentan algunas gráficas que muestran la forma en que las medias de los diferentes grupos se desplazan hacia uno u otro lado de la curva de distribución, lo cual hará más evidente la diferencia existente entre ambos grupos de comparación.

GRÁFICA 1

CURVA DE DISTRIBUCIÓN DE LOS DIFERENTES GRUPOS DE EDAD.
Comparación con la media uruguaya

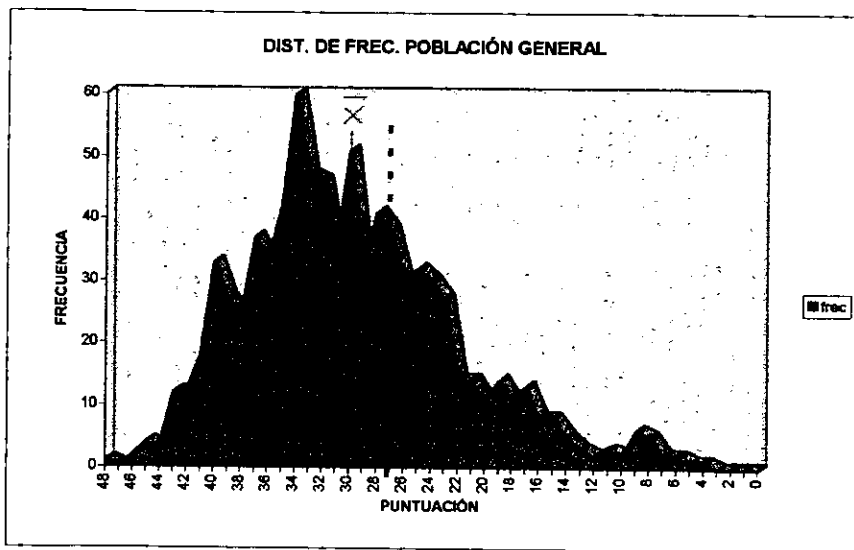


— \bar{X} de la muestra mexicana

- - - \bar{X} de la muestra uruguaya

GRÁFICA 2

DISTRIBUCION DE FRECUENCIAS DE LA POBLACION GENERAL.
Comparación con la media uruguaya



- \bar{X} de la muestra mexicana
- - - \bar{X} de la muestra uruguaya

Véase en la gráfica 1, la forma en que las diferentes curvas de distribución de los diferentes grupos de edad se inclinan hacia la izquierda con respecto al punto donde se supone tendría que estar la media, el cual está representado por la línea vertical, especialmente la gráfica que corresponde a la curva de distribución de la población general, la cual, además se acerca mucho más a la forma de la curva de distribución normal. En la gráfica 2 es más evidente la forma en que se desplaza la media de la población mexicana con respecto a la media de la población uruguaya.

Lo mostrado en estas dos gráficas sucede también con el resto de los grupos de edad, observándose claramente que la media para cada uno de ellos, está desplazada con respecto al punto donde debería de estar, lo cual es estadísticamente imposible, ya que en teoría se supone que la media de la población es igual a la media muestral. Pasando lo anterior al actual punto de interés, se quiere decir que en teoría la media de la muestra de los sujetos mexicanos es igual a la media de la población uruguaya, o por lo menos eso es lo que se había dado por hecho desde hace más de 45 años, ya que, como se ha venido mostrando, en realidad existen diferencias entre los diversos grupos estudiados en la presente investigación. Ahora lo que resta es tratar de demostrar si las diferencias encontradas entre los grupos son significativamente diferentes desde un punto de vista estadístico, lo cual daría como resultado la elaboración de baremos adecuados o más bien tipificados para la población mexicana.

Con el objeto de establecer la dimensión de las diferencias se realizó la prueba de hipótesis o prueba Z, también conocida como prueba de la significancia de la diferencia entre dos medias. Es muy importante señalar que en la presente investigación, cada una de las muestras que se obtuvieron a partir de la subdivisión del total de los sujetos fueron tratadas estadísticamente como muestras independientes para fines de estudio.

De igual forma, cabe recordar que la puntuación z para un nivel de significancia $\alpha = 0.05$ es de 1.96 y que la puntuación z para una nivel $\alpha = 0.01$ es de 2.58.

Después de haber hecho la especificación anterior, se procedió a analizar los resultados de la prueba de hipótesis, donde se estudió la significancia de la diferencia entre las puntuaciones medias de los sujetos uruguayos contra las puntuaciones medias de los mexicanos, primero por rangos de edad y después para el total de las poblaciones.

Para el Rango I (12-13 años), se observa una $z = -2.23$, lo que indica que si se considera un nivel $\alpha = 0.05$ las medias de los grupos son diferentes, es decir, que pertenecen a distintas poblaciones. Sin embargo, al disminuir la probabilidad de error a un nivel $\alpha = 0.01$, tal afirmación no podría ser válida.

Para el Rango II (14-15 años) sucede prácticamente lo mismo, ya que se obtuvo una $z = -2.11$, lo que indica que con una probabilidad de cometer un error del 5% ($\alpha = 0.05$), se puede asegurar que las medias de los grupos son diferentes, aunque con una probabilidad de error del 1% ($\alpha = 0.01$), dicha aseveración no es correcta.

Para el Rango III (16-17 años) se obtuvo una $z = -5.5$, lo cual no deja lugar a dudas, a cualquier nivel de significancia, de que ambas medias son totalmente diferentes, es decir, que pertenecen a poblaciones distintas. Cabe señalar que este fue el rango donde se hizo más evidente la diferencia entre las muestras.

Para el Rango IV (18-30 años), una vez más la prueba Z demostró que ambas muestras pertenecen a poblaciones diferentes al tener una $z = -4.84$, tomando en cuenta un nivel de significancia $\alpha = 0.01$

Por último, para el llamado Rango V (Población General) se obtuvo una $z = -8.58$, lo cual indica claramente que la media de los sujetos mexicanos y la de los uruguayos son diferentes entre sí a un nivel de significancia $\alpha = 0.01$ como mínimo.

Cabe señalar que en la prueba de hipótesis entre los sujetos mexicanos y uruguayos, todas las puntuaciones obtenidas fueron negativas, lo que quiere decir

que en todos los casos la media de los uruguayos se situaba por debajo de la de los mexicanos, con sus diferencias en cuanto a inteligencia se refiere. En otras palabras se puede decir que los mexicanos son más inteligentes que los uruguayos, y que por más de 45 años se había aceptado , o más bien se había dado por hecho, que la inteligencia era igual en unos y otros, casi siempre argumentando la similitud de las razas por ser ambas latinas. Hasta ahora no era válido pensar en que existan diferencias significativas entre dos poblaciones física o culturalmente similares y que es válido dudarlo hasta el grado de elaborar tablas con los parámetros propios para cada una de los aspectos que se quieran medir, especialmente si se pretenden usar parámetros que pertenecen a otra población.

A fin de corroborar el proceso estadístico y para tratar de establecer si existían diferencias entre hombres y mujeres, se llevó a cabo una prueba Z entre las puntuaciones medias de cada grupo de edad tomando en cuenta el sexo, observándose lo siguiente:

Como se puede ver en la tabla 3.10.1 casi todas las puntuaciones, exceptuando la primera, se mantienen por debajo del valor de z necesario para considerar a ambos grupos como iguales con un nivel $\alpha = 0.05$, es decir, $z = 1.96$; aunque disminuyendo el nivel de error al 1% , se puede asegurar que no existe diferencia significativa entre los sujetos mexicanos en función del sexo. Dicho en otras palabras, y contrariamente a lo que pensarían algunos grupos machistas o feministas, la inteligencia de los mexicanos y las mexicanas es estadísticamente

igual, al menos basándose en el factor general de inteligencia, ya que es conocido que si existen algunas diferencias entre hombres y mujeres en algunos factores específicos de la inteligencia.

CAPITULO IV

CONCLUSIONES Y RECOMENDACIONES

4.1 CONCLUSIONES

Como se ha podido observar en el presente trabajo, desde el imperio chino en el año 1000 a. C. hasta nuestros días, la historia de la medición se ha enmarcado por una gran necesidad de conocer las diferencias entre los sujetos para poder dar una calificación de “más que” o “mejor que”, o en su defecto todo lo contrario. Esta inquietud fue cambiando y evolucionando con los años llegando a establecer con cierto rigor la forma en que alguna característica humana o capacidad fuera medida. Pero no fue sino hasta que las necesidades de algunos países obligaron a encontrar la forma de hacer mediciones masivas que fueran suficientemente válidas para hacer un reclutamiento óptimo de personas con ciertas características; de ahí la importancia de la primera guerra mundial para la psicometría.

Sin duda, parte fundamental en la elaboración de las pruebas, incluyendo las psicométricas, es el fundamento estadístico en el que se basan para poder llegar a hacer generalizaciones, sin lo cual una prueba, cualquiera que esta sea, carece de validez y por supuesto de confiabilidad.

De igual forma, a pesar de que para la realización del presente estudio se utilizó una prueba psicométrica que mide la inteligencia de los sujetos a partir del factor general, no es la inteligencia la que esta en tela de juicio, a pesar de que si se encontraron diferencias, sino la inconsciencia de algunos investigadores, incluyendo psicólogos, para aplicar las pruebas psicológicas como el Dominós sin un respaldo estadístico que asegure que los parámetros para calificar y clasificar a los sujetos corresponde a la población en que se está realizando la medición. En otras palabras, no se puede asegurar que un sujeto mexicano tiene un percentil 50 en una prueba de inteligencia, si la media de los sujetos mexicanos es diferente a la media de los sujetos con la que se hizo y estandarizó dicha prueba.

En la presente investigación se ha puesto especial interés en demostrar que ha sido un error clasificar a los sujetos mexicanos según su inteligencia utilizando los baremos uruguayos de la prueba de Dominós de Edgard Anstey. ¿Que acaso los usuarios de la prueba en México no llegaron a pensar que tal vez las poblaciones eran diferentes? ¿O tal vez pensaron que por estar en América y hablar español deberían ser igual de inteligentes? Ahora es sabido que no hay que dar nada por hecho, que muy probablemente haya diferencias, que se debe

realizar todo el proceso de estandarización para una prueba antes de querer aplicarla en una población diferente a la que se utilizó para su estandarización.

Actualmente se puede observar cotidianamente en empresas, escuelas primarias, secundarias, universidades o consultorios psicológicos públicos y privados, usar indistintamente todo tipo de pruebas psicométricas sin antes estar estadísticamente seguros que no hay diferencias significativas entre ambos grupos, con lo que se está muy cerca de realizar un grave error, como con el Dominós, ya que de eso depende en gran medida el ingreso de algunas personas a las universidades o empresas, pero lo que es más grave aún, es que se pueden cometer errores a la hora de elaborar un diagnóstico psicológico.

A lo largo del presente trabajo se han podido dar referencias claras de que la muestra utilizada para la estandarización de la prueba de Dominós hecha en Uruguay y la muestra mexicana utilizada para esta investigación, tienen medidas de tendencia central y de dispersión diferentes, y se ha hecho evidente de que dichas diferencias son significativas como para elaborar tablas o baremos de clasificación propias para la población mexicana. Si hay que medir la inteligencia de los sujetos, hay que medirla bien.

En conclusión como resultado de la presente investigación se puede afirmar que sí existen diferencias significativas entre la media de las puntuaciones de los sujetos uruguayos y la media de las puntuaciones de los sujetos mexicanos a un nivel de significancia $\alpha = 0.01$, por lo tanto, se acepta la hipótesis de trabajo y se rechaza la hipótesis nula.

Como resultado del trabajo de investigación, se elaboraron y ahora se presentan los Baremos Mexicanos del Test de Dominós para la clasificación de la inteligencia, basados en 843 sujetos de 12 a 30 años del Distrito Federal y área metropolitana, los cuales se muestran en el Anexo A .

Otra conclusión importante a la que se ha podido llegar es que las diferencias en la inteligencia encontradas entre los hombres y mujeres de la muestra no son significativas, recalcando que la falta de similitud entre las puntuaciones no es suficientemente significativa para hacer baremos específicos para hombres y mujeres por separado.

De igual forma, haciendo pruebas de hipótesis entre los diferentes grupos comparativos, se encontró que entre algunos de ellos no había diferencias, por lo que algunos rangos de edad pudieron haber sido agrupados.

La inteligencia, siendo una parte importante de la psique de un individuo, debe ser tratada como una parte integral de la mente de un ser humano, y el solo hecho de utilizar lápiz y papel para medirla, no significa que tiene menos importancia o que debe ser tratada a la ligera. A pesar de que el Dominós parece ser una forma un tanto simplista de medir la inteligencia, hay que recordar que su fundamento radica en el factor general, lo cual la convierte en una forma rápida y confiable de explorar la inteligencia en un individuo o en un grupo de ellos.

4.2 RECOMENDACIONES

Después de haber presentado el trabajo y sus conclusiones, hay varias cosas que pudieron haberse hecho para mejorar los resultados o para poder hacerlos más contundentes. A continuación se ofrecen algunas sugerencias para aquellos que estén interesados en seguir los pasos de esta investigación:

Sin duda, uno de los aspectos más fundamentales es el tamaño de la muestra, ya que a pesar de que es significativa para la población del Distrito Federal, podría hacerse con un número mayor de sujetos. Con esto, el rango de puntuaciones sería mucho más extenso y se podría estar completamente seguro de que se tomaron puntuaciones a ambos extremos de la curva de distribución; por lo tanto, las conclusiones serían aún más precisas y contundentes. El número de sujetos recomendado es de 2000 o más.

Sería recomendable también tener un control más estricto para la contabilización de los sujetos por edad y sexo al momento de ir aplicando las pruebas, a fin de que cada grupo comparativo fuera del mismo tamaño y con el mismo número de hombres y mujeres en cada uno.

Si han de realizar esta investigación, se recomienda seguir el procedimiento descrito en la metodología, la base de todo es la simplicidad.

Definitivamente hay que sugerir el uso de un paquete estadístico apropiado a sus necesidades; no hay que dejar el trabajo de la parte estadística en manos de nadie más, ya que hay que conocer la estadística y manejarla adecuadamente.

Si se desea continuar con este trabajo a partir del punto en que se encuentra, sería recomendable comprobar que los baremos sugeridos son adecuados para la población mexicana del Distrito Federal, y más aún, sería muy importante demostrar que dichos baremos pueden ser utilizados para toda la población mexicana, sin restricciones. De no ser así, no hay que dudar en proponer baremos diferentes para cada entidad que así lo requiera.

Esta investigación puede hacerse con cualquier prueba psicológica utilizada en nuestro país, de hecho, debería hacerse algo parecido con cada una de ellas antes de sacar su versión en español para México. Hay que recordar que las

pruebas, aunque sean españolas, podrían no ser adecuadas estadísticamente para la población mexicana. La simple traducción de ellas no es suficiente.

Si se quiere llegar más lejos, se podría hacer un análisis detallado de la prueba en sí, tratando de encontrar si en realidad mide la inteligencia o no. De igual forma se podría establecer si el grado de dificultad de cada ítem es adecuado y se encuentra en el orden apropiado a su dificultad. Se hace esta sugerencia ya que se encontró que el ítem N° 30 es mucho más difícil que muchos con numeración posterior.

Se espera que la realización de esta investigación sirva de aliciente para los investigadores jóvenes, en particular psicólogos, para seguir haciendo este tipo de trabajos, ya que son precisamente las pruebas psicológicas y más particularmente las diferencias individuales las que competen a la psicología, siendo ellos los que precisamente tendrían que dedicarse a la tarea de hacer pruebas psicológicas, o en su defecto, estandarizarlas . Los temas psicológicos competen a los psicólogos, pero definitivamente hay que usar la estadística adecuadamente para poder hacer bien el trabajo. Si hay que medir la inteligencia, o lo que sea, hay que medirla bien.

“ Cualquier cosa que existe, existe en cantidad (Thorndike, 1918)

Cualquier cosa que existe en cantidad puede medirse (Mc Call, 1939) “

(en Aiken, 1995)

“ Cualquier cosa que ha de medirse... mídase bien “ (Arturo Roa, 2001)

REFERENCIAS BIBLIOGRÁFICAS

- Aiken, L. (1996). Test psicológicos y evaluación. (8ª ed.). México: Prentice-Hall. (Tr. Psychological testing and assessment. Trad: Verónica de Parres impreso en México).
- Amón, J. (1980). Estadística para psicólogos I. Madrid: Pirámide.
- Anastasi, A. (1970). Test psicológicos. México: Planeta.
- Anstey, Edgard. El test de Dominós. Manual. México: Planeta.
- Brown, Frederick G.(1980). Principios de la medición en psicología y educación. México, El manual moderno.
- Corbella, J. (1994). Descubrir la psicología. México: Planeta.
- Deledalle, G. (1965). Psicología Contemporánea. Tr. Española por Luis Cuellar. Barcelona; México: Labor.
- Downie, N. M., Heath, R. W. (1986). Métodos estadísticos aplicados. (5ª ed.). México: Harla.
- Fernández-Ballesteros, R. (1993). Introducción a la psicología I. (2ª ed.). Madrid: Pirámide.

- Floria, G.B., Fenollosa, M. R. (1984). Psicología General. Tr. J. Ferrer, Aleu. Barcelona: Plaza & Janes.
- García, E. (1993). Introducción a la psicometría. México: Siglo XXI.
- Ibáñez Brambilia, B. (1990). Manual para la elaboración de tesis. México: Trillas.
- Iriarte, M. (1948). El doctor Juan Huarte de San Juan y su "examen de ingenios". Contribución a la psicología diferencial. Madrid: CSIC.
- Martínez, A. (1995). Psicometría. Teoría de los tests psicológicos y educativos. España: Síntesis Psicológica.
- Morales, J. F., Fernández Dols, J. M. (1990). Patrones para el diseño de la psicología social. Madrid: Morata.
- Muñiz, J. (1994) Teoría clásica de los test.. España: Pirámide.
- Paul, K. (1993). The handbook of psychological testing. Gran Bretaña: Routledge.
- Paulus, J. (1970). Los fundamentos de la psicología. Madrid: Guadarrama
- Pichot, P. (1991). Test Mentales. México: Paidós.

Rojas Soriano, R. (1982). El proceso de la investigación científica. (2ª ed.).

México: Trillas.

Sattler, J. (1993). Evaluación de la inteligencia infantil y habilidades especiales. México: Manual Moderno.

Sellier, J. L. (1980). Conocer a los demás por los test: su comprensión y respuesta. España: Mensajero. (Paris 1973, España 1976. Trad: Juan José Ferrero)

baremos mexicanos

ANEXO A

PROPUESTA DE BAREMO MÉXICO D.F.

EDAD \ CENTIL	12-13	14-15	16-17	18-20+	POBL. GRAL
99	42	42	46	47	46
95	38	40	42	44	42
90	36	38	41	42	40
80	34	35	38	40	37
75	33	34	37	39	36
70	32	32	36	37	35
60	30	30	34	35	33
50	27	28	32	33	30
40	26	27	31	32	29
30	25	25	29	30	26
25	23	23	27	28	25
20	21	21	26	27	24
10	17	17	20	23	19
5	13	14	15	20	15
1	7	8	8	13	8

EDAD \ CENTIL	12-13	14-15	16-17	18-20+	POBL. GRAL
95	38	40	42	44	42
90	36	38	41	42	40
75	33	34	37	39	36
50	27	28	32	33	30
25	23	23	27	28	25
10	17	17	20	23	19
5	13	14	15	20	15

APÉNDICE B

TABLA DE CONVERSIÓN DE PERCENTILES A RANGOS

<i>PERCENTIL</i>	<i>RANGO</i>
95	SUPERIOR
90 Y 75	SUPERIOR AL TÉRMINO MEDIO
50	TÉRMINO MEDIO
25 Y 10	INFERIOR AL TÉRMINO MEDIO
5	INFERIOR

APÉNDICE C

BAREMO MONTEVIDEO

EDAD \ CENTIL	12-13	14-15	16-17	18-20+	POBL. GRAL
99	42	43	44	45	44
95	38	39	41	41	40
90	35	37	39	40	37
80	33	34	35	37	35
75	32	33	34	36	34
70	31	32	33	35	33
60	29	30	31	33	31
50	27	28	29	31	29
40	25	26	27	29	27
30	23	24	25	26	25
25	22	23	24	25	23
20	21	22	23	24	22
10	14	15	16	20	17
5	9	11	12	16	12
1	4	5	6	8	5

EDAD \ CENTIL	12-13	14-15	16-17	18-20+	POBL. GRAL
95	38	39	41	41	40
90	35	37	39	40	37
75	32	33	34	36	34
50	27	28	29	31	29
25	22	23	24	25	23
10	14	15	16	20	17
5	9	11	12	16	12