

Universidad Nacional Autónoma de México

FACULTAD DE FILOSOFIA Y LETRAS

Colegio de Psicología



EL TEST DE DOMINOS
UN ESTUDIO PSICOMETRICO

596
PSI

T E S I S

Que para obtener el título de :

P S I C O L O G A

p r e s e n t a :

CARMEN MONCADA IVAR



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.



25053.08
UNDM.12
1969

M-161080

Typo 123

A mi hermano Carlos.

A mi padre y hermanos.

A la memoria de mi madre y de mi hermano
Luis.

T. Ps. 481

Hago patente mi agradecimiento por el apoyo
y colaboración que me brindaron para la
realización de este trabajo.

Dra. Sara Margarita Zendejas Merino.

Sra. Profa. Elena Salázar Mayén.

Dr. Rogelio Díaz Guerrero.

Agradezco a la Comisión Federal
de Electricidad todas las faci-
lidades que me prestó y muy es-
pecialmente al Centro de Cálculo
por su valiosa ayuda en el trata-
miento estadístico de los datos.

I N D I C E

	Pág.
INTRODUCCION.....	1
<u>C A P I T U L O I</u>	
ANTECEDENTES Y PRESENTACION DE LA PRUEBA.....	3
<u>C A P I T U L O II</u>	
OBJETIVOS Y PROGRAMA DE TRABAJO.....	11
<u>C A P I T U L O III</u>	
ANALISIS ESTADISTICO DE LA PRUEBA.....	17
<u>C A P I T U L O IV</u>	
NORMAS DE CALIFICACION.....	38
CONCLUSIONES.....	51
BIBLIOGRAFIA.....	52

I N T R O D U C C I O N

En un país como México, considerado como subdesarrollado o en proceso de desarrollo, existen una serie de deficiencias en muchos aspectos, sin embargo, se puede afirmar que algunas de estas se han empezado a superar y mucho se ha logrado en algunas ramas, sobre todo en lo que se refiere a la técnica. Por otro lado, las disciplinas llamadas socio-humanísticas, no han tenido la evolución que debieran, tomando en cuenta que son un factor determinante para un completo y armonioso desarrollo de nuestro país.

Dentro de estas disciplinas, tenemos la Psicología, con sus diferentes especialidades, la cual se ha enfrentado a una serie de problemas de toda índole; una raquítica preparación en lo que se refiere a la investigación y experimentación, restricciones de tipo económico, falta de material humano y otros factores que no le han permitido alcanzar el nivel que corresponde a nuestro país.

Uno de los problemas a los que se enfrentan los psicólogos de diferentes especialidades; industriales, clínicos, orientadores y experimentales, es la falta de instrumentos adecuados para una medición válida y confiable de las capacidades, aptitudes, intereses y características generales de personalidad de nuestra población, tomando en cuenta que la mayoría de las pruebas que se u

utilizan en nuestro país, no se encuentran estandarizadas y adecuadas para ser aplicadas a una población como la nuestra.

Sin embargo, a la fecha existen ya varios grupos de estudios de la materia que de unos años a la fecha se han dedicado a la Psicología Experimental y Aplicada, utilizando sistemas y técnicas de investigación bastante serios, pero aún así, el problema subsiste.

Es por esto que cuando yo ingresé a trabajar al Departamento de Encuestas y Servicios Sociales de la Comisión Federal de Electricidad y al conocer la batería de pruebas que ahí se utilizaba, con fines de selección profesional, me dí cuenta que sólo una de ellas estaba estandarizada en México y todas las demás se aplicaban y calificaban con las normas de su país de origen.

Formando parte de esta batería se encontraba la prueba de "Dominós" de E. Anstey que mide factor "G" de inteligencia.

Como la prueba me pareció muy interesante y adecuada para el nivel de personas a las que se seleccionaba (ingenieros en su inmensa mayoría), pensé que valía la pena hacer el estudio completo de esa prueba, el que incluiría la validez, la confiabilidad, la consistencia interna o análisis de reactivos y los baremos adecuados a la población mexicana de nivel profesional.

Partiendo de esta base y aprovechando las facilidades y apoyo que me brindaba la jefe del Departamento, y la propia institución, inicié la investigación de dicha prueba, cuyos resultados ahora expongo.

C a p í t u l o I
ANTECEDENTES Y PRE-
SENTACION DE LA -
PRUEBA.

La prueba de "Dominós" forma parte de una larga lista de - pruebas que teniendo como base el "punto" se han elaborado desde mucho tiempo atrás, casi desde principios de la investigación experimental.

De acuerdo a las necesidades de la época y también de acuerdo a la formación profesional o al tipo de escuela a la que han pertenecido los diferentes investigadores, han sido las diversas pruebas elaboradas tomando como base el "punto". Así encontramos instrumentos para medir diferentes funciones psíquicas.

Como resultado de toda esta gama de pruebas de "puntos", - tenemos una clasificación que obedece a diversos criterios como son:

- 1.- Según su material; grupos de "puntos" informes (en desorden) y configurados (organizados en figuras geométricas o numéricas).
- 2.- Por su finalidad; pruebas experimentales (investigación básica) y pruebas de diagnóstico (examen aplicado).
- 3.- Por su objetivo específico; de aptitudes, de inteligencia y de personalidad.
- 4.- Por los problemas que plantea; de reconocimiento, retención y cálculo (aptitudes); analogías, clasificación y series (inte-

ligencia); de copia, continuación o completamiento gráficos o --- construcción de modelos (gestálticos); y de dibujo inducido (personalidad).

5.- Por la respuesta del examinado; verbales, lápiz papel y ejecución.

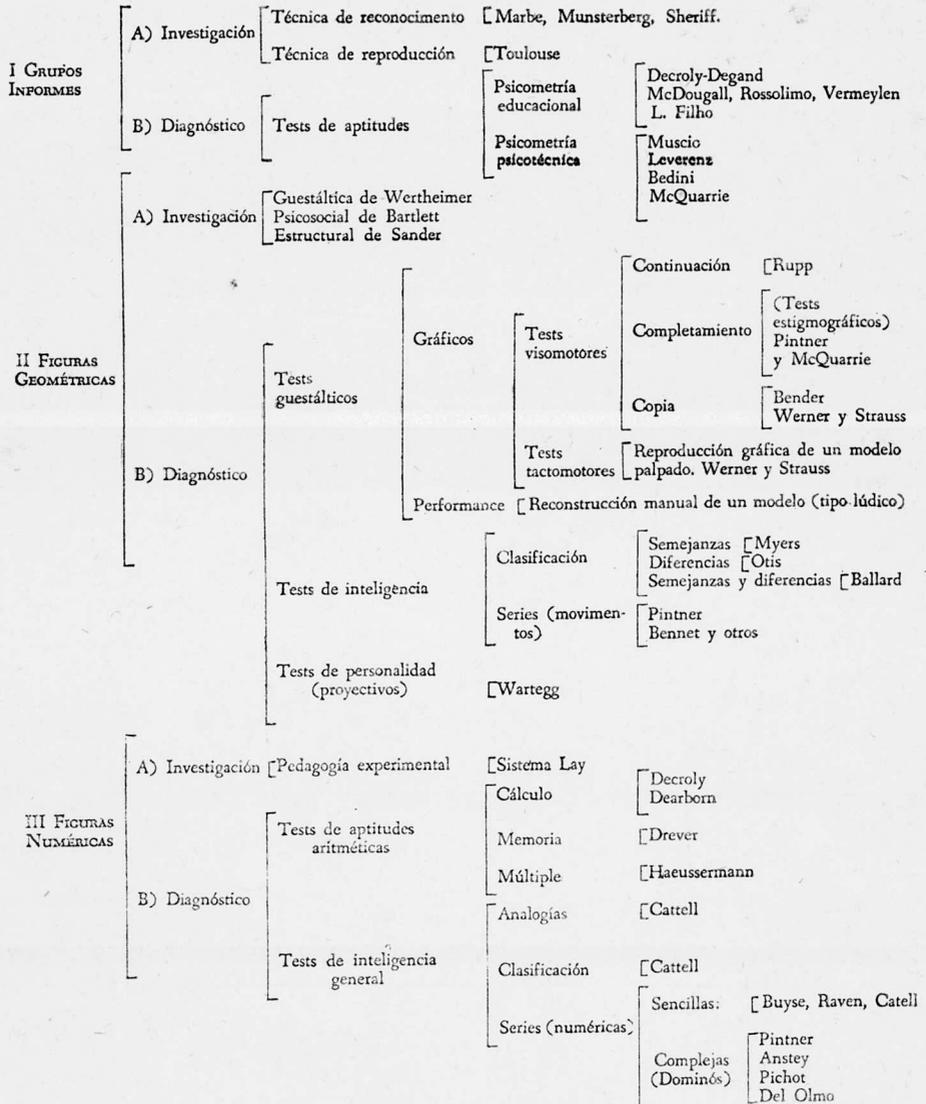
6.- Por su aplicación; educacionales, psicotécnicos y clínicos.

7.- Por su administración; individuales y colectivos.

8.- Por su fundamentación; empíricos y escolásticos (gestálticos, factoriales, proyectivos).

La clasificación de esta gran familia de pruebas a manera de cuadro sinóptico se debe a Jaime Bernstein, Director de la Biblioteca de Psicometría y en él se ve claramente el lugar que ocupa la prueba de Dominós de E. Anstey que es la que ahora nos ocupa.

ESQUEMA REFERENCIAL DEL DESARROLLO DE LOS TESTS DE PUNTOS



Como puede verse claramente en el cuadro anterior, la prueba de Anstey corresponde al grupo de pruebas más elaborado y así mismo más complejo.

Ahora bien, ya ubicados en el contexto de pruebas elaboradas tomando como base fundamental el "punto", voy a hacer una presentación de la prueba.

El test de Dominós del Psicólogo inglés Edgar Anstey, fue elaborado en 1944 para aplicarlo a los miembros de la Armada Británica como test paralelo de las Matrices Progresivas de Raven y con la intención de superar algunos de los inconvenientes de dicha prueba.

En nuestro país, concretamente en la Universidad Nacional-Autónoma de México, las matrices progresivas formaba parte de la batería utilizada para la selección de alumnos de primer ingreso, fue ahí donde se evidenció en los resultados obtenidos en su aplicación, que, debido probablemente a la constante aplicación del test, los jóvenes habían llegado a memorizar las respuestas, lo que la inutilizaba totalmente.

Este vendría a ser uno de los inconvenientes a los que hacía referencia anteriormente. Más adelante, haré mención de algunos otros, en lo que se refiere a las matrices progresivas en relación a Anstey.

Ahora bien, analíticamente y en razón a los problemas que comprende la prueba, corresponde a los siguientes grupos:

- 1.- Test de series; Es un test de educación de leyes o principios de relaciones.
- 2.- Test de continuación o completamiento de series numéricas; --

desde este punto de vista se le puede considerar como un test de completamiento de lagunas que 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:

Simetría

Alternancia y progresión simple

Asimetría

Progresión circular

Progresión compleja (series)

Combinación de principios previos

Adición y sustracción

Con relación al material que la forma, corresponde al grupo de:

Test de figuras numéricas de "puntos".- las series de cifras se representan por grupos de puntos dispuestos según un patrón (Dominós).

Debido a la utilización de las fichas de dominós en forma exclusiva, le corresponden estas otras especificaciones:

Test homogéneo perceptual y test de figuras abstractas; ya que por su apariencia familiar, las fichas parecen ofrecer un material concreto, pero por el tipo de problemas que plantea, el Dominós es un test de figuras no representacionales.

En razón de su administración es:

Autoadministrable; pues puede tomarse en forma individual o en forma colectiva con sujetos mayores de 12 años.

En razón de sus fundamentos y de su objetivo es: Test factorial.- El Anstey se fundamenta en la teoría factorial de Spearman (1904) y en su metodología psicométrica. Pues en efecto, Anstey aplicó los principios psicométricos fundamentales sustentados por los factorialistas ingleses desde sus iniciadores (1900) -- Spearman, Brown y Stephenson y que son: procurar la medida de la "inteligencia general" mediante pruebas de educación de relaciones sobre material homogéneo no verbal o perceptivo; específicamente, a través de problemas formados en términos de descubrimientos de relaciones de figuras abstractas, preferidas por su carácter no familiar, lo cual permite una evaluación más pura del factor "G".

Por otra parte, se considera un test de capacidad potencial.- El Dominós brinda al examinado el tiempo necesario para -- realizar el máximo de tareas de que sea capaz.

Finalmente se presenta la comparación entre la prueba de Anstey y las Matrices Progresivas de Raven, ya que las dos se formaron con idéntico fundamento y que sólo viene a demostrar que el balance final es favorable al Anstey, ya que hasta antes de la creación del Dominós, Raven era el que había logrado la máxima saturación del "G" y la mínima de factores específicos. Sin embargo a causa de sus 30 primeros problemas en los que la Gestalt tiene evidente peso, el Raven acusa cierta contaminación de un factor específico del grupo visual perceptivo, un factor espacial (0.15) y un residuo (0.06) de otros factores específicos no bien identificados, razón por la cual el Dominós superó en pureza de medición del factor "G" como podemos comprobarlo a través del cuadro-

siguiente, que es un resultado de los estudios hechos por Vernon- (1947) y apoyados por Baines.

C U A D R O II
COTEJO FACTORIAL ENTRE LAS MATRICES
PROGRESIVAS Y EL DOMINOS.

	G		K		n	
	M	D	M	D	M	D
Raven	0.82					
Anstey	0.79	0.82				0.12*
Vernon	0.79	0.90	0.15	0.00	0.00	0.05
Baines	0.75	0.82				(1)

G: Factor general; K: Factor espacial;

n: Factor numérico M: Matrices progresivas

D: Dominós (1944) *: En rigor factor espacio-numérico.

Hecha la presentación de la prueba, sólo me resta aclarar que la versión original de ella, elaborada por el autor en 1944 para uso de la Armada Británica, ha quedado inédita como una "Prueba cerrada", llamada así por los ingleses a las pruebas de uso exclusivo oficial. Pero posteriormente, Anstey elaboró una nueva versión (1955) con ligeras diferencias en su material (48 ítems agrupados en 8 páginas según sus principios de dificultad creciente) y a la que dió un carácter de "Prueba abierta".

(1) Bernstein Jaime.- "Manual del Test de Dominós" p.41

Esta versión de la prueba la confié para efectos de estudio, standarización y publicación al distinguido investigador uruguayo Doctor W. L. Risso, Director del Servicio Médico - Pedagógico de la Universidad del Trabajo de Uruguay. Sin embargo quisiera aclarar que la literatura que existe con relación a esta prueba es solamente la que he mencionado, por lo que prácticamente y hasta donde yo sé, este es el primer estudio completo que se hace de este test en México.

Es esta versión y los resultados obtenidos por el Doctor Risso en lo que me he basado para llevar a cabo el estudio de la prueba aplicada a población mexicana, en este caso particular, a población de nivel profesional.

Capítulo II

OBJETIVOS Y PROGRAMA DE TRABAJO.

Los objetivos principales en este estudio fueron dos:

- 1.- Análisis estadístico de la prueba en sus aspectos: Confiabilidad, Validez y Análisis del Grado de Dificultad de los Reactivos.
- 2.- Estructuración de Baremos con una población mexicana de nivel profesional.

Definidos así los objetivos a seguir, llevé a cabo su obtención mediante el siguiente programa de trabajo, dividido en dos partes: 1.- Población estudiada y 2.- Procedimiento de trabajo.

POBLACION ESTUDIADA

El primer grupo de población estudiada la integró el grupo de profesionistas de la Comisión Federal de Electricidad, a los que se les había aplicado la prueba de inteligencia de Anstey, la que formaba parte de la batería de pruebas para selección profesional destinada a trabajadores de confianza, durante el periodo comprendido de noviembre de 1960 al 31 de Diciembre de 1963 (750-casos).

Esta población se integró con sujetos todos de nivel profesional, en su mayoría ingenieros (85%), provenientes de las diversas instituciones de estudios superiores del país: Universidad Na

cional Autónoma de México, Instituto Politécnico Nacional, Instituto Tecnológico de Monterrey, etc. Las edades de los sujetos -- fluctuaron entre 19 y 42 años y todos eran de sexo masculino.

El segundo grupo de población estudiada lo integró una -- muestra de alumnos del 5o. año de la Facultad de Ingeniería de la U.N.A.M. (260 casos).

La utilización de esta población de estudiantes universita -- rios, tuvo como principal objetivo el lograr obtener la validez -- de la prueba.

PROCEDIMIENTO DE TRABAJO

Este se llevó a cabo mediante los siguientes pasos:

- 1.- Material Utilizado
- 2.- Forma y lugar de aplicación de la prueba
- 3.- Variables controladas y estipulaciones
- 4.- Datos obtenidos

1.- Material Utilizado

a) Promedio escolar.- Esta parte del material de trabajo;-- promedio escolar de los estudiantes del 5o. año de la Facultad de Ingeniería de la U.N.A.M., obtenido en el curso de toda la carre -- ra y proporcionado por la Secretaría de la Facultad, es el punto -- fundamental de la validación en estudio y tiene por finalidad de -- terminar en que medida correlacionan los puntajes alcanzados en -- Anstey con los promedios de calificación que les corresponden. -- Considerando que no existen criterios válidos para determinar la -- calificación universitaria y que en número considerable, son medi -- das aplicadas " azar", la calificación escolar resulta ser un --

factor de juicio subjetivo, que sólo eventualmente puede resultar conforme con la realidad.

Tomando en cuenta este hecho, creí que la manera de objetivar estos juicios subjetivos sería tomando el promedio escolar de los estudiantes de Ingeniería a través de toda la carrera, -- pues sería casi imposible suponer que el juicio de 30 o más personas no lograrán un equilibrio que diera como resultado un juicio más objetivo.

b) Cuestionario.- Se elaboró un cuestionario específicamente para este experimento de validación, que estaba destinado a obtener alguna información socio-económica y familiar de los estudiantes y en el que se incluía una pregunta cuya respuesta, analizada y clasificada, constituiría el instrumento para explorar la dirección de las motivaciones de los integrantes de una muestra de la población estudiantil; decimos una muestra, porque desafortunadamente no se pudo aplicar a toda esta población. De esta manera, también podríamos haber conocido su nivel socio-económico, ya que se pretendía saber la forma en que influyen estos factores en el rendimiento escolar a través de cinco años.

2.- Forma y lugar de Aplicación de la Prueba.

En la población de trabajadores profesionistas de la Comisión Federal de Electricidad (C.F.E.).

a) Forma de aplicación.- La aplicación se realizó en forma individual.

b) Lugar de aplicación.- La prueba se aplicó en las oficinas centrales de la Institución, en cubículos instalados con todas -- las características requeridas para obtener óptimos resultados

En la población de estudiantes del 5o. año de la Facultad de Ingeniería de la U.N.A.M.

a) Forma de aplicación.- Colectiva, en grupos integrados por la -- inscripción que registra cada catedrático.

b) Lugar de aplicación.- Esta se realizó en la propia Facultad de Ingeniería utilizando sus propios salones de clase y en presen -- cia del catedrático que impartía la materia.

3.- Variables Controladas y Estipulaciones.

a) Sexo, masculino.

b) Una hora de tiempo máximo para el desarrollo de la prueba.

En relación con la variable "tiempo", tanto el psicólogo -- Anstey como los psicólogos Bernstein y Risso, conceden tiempo menor de una hora para el desarrollo de la prueba, pero nosotros, -- tomando en cuenta que se trata de una prueba de poder y no de velocidad, dimos más tiempo, ya que consideramos que las pruebas de poder son más justas que las de velocidad y en este caso particular, los hechos lo han demostrado.

Sobre este punto conviene hacer la aclaración de que cuando se inició la aplicación de la prueba de Anstey en el año de -- 1960, desde los primeros casos se advirtió que durante un lapso -- no comprobadamente determinado, pero que llegaba a prolongarse -- hasta 15 minutos, la población mexicana casi no respondía al reac

tivo, es decir, no entraba a la situación de prueba; tal vez suponiendo que se trataba de un problema demasiado fácil, o tratando de hacerlo creer así al examinador, o que era un artificio cuya solución no demandaba esfuerzo mental específico, razanado y sistemático, o bien por argumentos de índole diversa que, por otro lado, expresaban características de personalidad del sujeto examinado.

El hecho es que al permitir "tiempo libre", durante un período razonable de observación para determinar el " tiempo mexicano" sobre una base real, quedó demostrado que desde el momento en que el examinado toma interés sincero en la solución de los reactivos y está conciente de que le reclaman un esfuerzo mental serio y disciplinado, se alcanza un tiempo de desarrollo de 40 a 55 minutos como promedio aproximado para resolver los 48 reactivos, pero es en extremo excepcional que el tiempo utilizado sobrepase a una hora.

En consecuencia, se estableció para el desarrollo de la prueba, una hora de tiempo máximo, ya que repetidamente se comprobaba un hecho que potencialmente se atribuye no precisamente a manifestación de lentitud, sino más bien, a un rasgo característico de nuestra población.

Estipulaciones

- a) Utilización del protocolo de prueba formulado en la Comisión Federal de Electricidad para la prueba de Anstey y cuyo cambio

corresponde sólo al hecho de no utilizar cuadernillo donde vienen los reactivos y hoja de respuestas por separado, sino que se ideó imprimir sobre la misma hoja de respuestas los reactivos y en esta forma se eliminó el uso del cuadernillo.

a) Unificación en la administración de instrucciones para el desarrollo de la prueba.

4.- Datos Obtenidos

Los resultados obtenidos en el estudio fueron:

1.- El análisis estadístico de la prueba en sus aspectos:

a) Confiabilidad

b) Validez

c) Análisis del grado de dificultad de los reactivos.

2.- Baremos Mexicanos de Nivel Profesional.

En los siguientes capítulos procederé a la descripción de los datos obtenidos.

C a p í t u l o I I I

ANÁLISIS ESTADÍSTICO- DE LA PRUEBA

- 1.- Confiabilidad
- 2.- Validez
- 3.- Análisis del Gra-
do de Dificultad-
de los Reactivos.

CONFIABILIDAD

Como ya había anunciado anteriormente, uno de los objetivos principales de este trabajo es el estudiar la prueba en todos sus aspectos y de esta manera obtener un instrumento que midiera la inteligencia general a nuestra población profesional, aunque - por otra parte, la meta es lograr estos mismos objetivos en una población general mexicana.

Uno de los aspectos del primer objetivo de nuestro trabajo es la Confiabilidad, definida por David Wechsler como: "la consistencia con la cual una prueba mide las habilidades que se propone medir".

Existen varios métodos para obtener la Confiabilidad de una prueba. Todos ellos están interesados por el grado de consistencia o acuerdo entre dos conjuntos de puntuaciones derivadas -- independientemente y pueden expresarse en función de un coeficiente de correlación, cuyo símbolo estadístico es r .

Ahora bien, los coeficientes de correlación pueden calcularse de varias maneras, según la naturaleza de los datos.

El más común es el "coeficiente de correlación momento-producto" de Pearson.

Los métodos más usuales para obtener la confiabilidad de una prueba son:

a) Confiabilidad de repetición del "test" o Confiabilidad por "re test". Este es el método que se considera más adecuado para obtener la Confiabilidad, pues si se aplica la prueba a una muestra de sujetos, se espera que pasado un tiempo no menor de seis meses con una cierta variabilidad de acuerdo al tipo de población estudiada, el rendimiento debe de ser el mismo o casi el mismo, siempre y cuando la prueba mida consistentemente la habilidad, interés, capacidad, etc. que se quiera medir.

En este caso particular, la utilización del mencionado método no fue posible, ya que los sujetos de estudio, por diversas causas, no se encontraban en posibilidad de resolver nuevamente la prueba. Dado que este es el requisito fundamental en el método de "retest", dicho método no se pudo poner en práctica.

b) El siguiente método a seguir es el de Confiabilidad de la Forma equivalente; este consiste en aplicar en la primera vez a los sujetos de estudio la prueba y en una segunda ocasión aplicar una forma equivalente de ésta, es decir, una escala paralela, obteniendo con esto idénticos resultados o muy similares. Este segundo método tampoco se pudo llevar a la práctica, ya que no existe una escala paralela de la prueba. Por otra parte, no poseo los conocimientos técnicos que me permitieran elaborar una escala que fuera equivalente y equiparable a un test tan bien estructurado como lo es el Anstey.

Así que descartados estos dos métodos, sólo me quedaba uno que sí podría llevar a cabo de acuerdo a mis posibilidades de trabajo, y que por otra parte es tan eficaz como los dos anteriores.

De esta manera, utilicé el método de "Split-half" o de división en dos mitades, método que consiste en dividir la prueba

en dos partes, lo que nos proporciona una medida de equivalencia o adecuación de las muestras de elementos.

Existen varias formas de dividir la prueba en dos partes según las características que tenga; en este caso, tomando en cuenta que la prueba es una escala de dificultad creciente y tiene un tiempo límite, lo más conveniente fue dividirla tomando en una mitad los reactivos nones y en la otra los reactivos pares, de esta manera la escala de dificultad no se alteraría.

Seleccionado y planeado el método a seguir, lo llevé a efecto mediante la aplicación de la fórmula momento-producto de Pearson que es la siguiente:

$$r = \frac{N \sum XY - (\sum X)(\sum Y)}{\sqrt{[N \sum X^2 - (\sum X)^2][N \sum Y^2 - (\sum Y)^2]}}$$

Realizados todos los cálculos, obtuve la alentadora cifra de correlación de 0.83 advirtiéndome que esta cifra corresponde a la Confiabilidad de media prueba, por lo tanto, para obtener la Confiabilidad de la prueba completa utilicé la fórmula profética de Spearman-Brown, que es la siguiente:

$$R = \frac{2r}{1+r}$$

Con esta fórmula se obtuvo la Confiabilidad completa de la prueba y la cifra resultó ser 0.91 con lo cual queda ampliamente demostrada la Confiabilidad de la prueba de Anstey en población mexicana de nivel profesional.

Antes de hablar sobre la validez de la prueba, es menester hacer una aclaración. En el transcurso de la aplicación se observó que en nuestra población algunos de los reactivos que integran la prueba, resultaron más fáciles que otros considerados con mayor índice de dificultad por el autor. Creo pertinente hablar sobre esto, puesto que al decir que se trata de una escala de dificultad ascendente, no lo es para nuestra población y me propongo demostrarlo estadísticamente en la sección de Análisis del Grado de Dificultad de los Reactivos, aspecto que trataré posteriormente.

VALIDEZ

Uno de los problemas más difíciles de salvar en el estudio completo de una prueba es el criterio de validez.

La validez de una prueba es tener la seguridad de que esta mida exactamente lo que se pretende medir. En muchas ocasiones se da por válido el nombre de la prueba como índice de lo que está midiendo, criterio que resulta por demás inaceptable ya que la validez no se puede medir a partir de términos tan generales y sin ninguna base objetiva.

Existen diversos procedimientos para determinar la validez y sus técnicas se han designado con diferentes nombres. La Asociación Psicológica Americana, en sus Recomendaciones Técnicas, ha clasificado estos procedimientos bajo cuatro categorías que ha designado como valideces: de contenido, predictiva, concurrente y de elaboración.

Validez de Contenido.- Como su nombre lo indica, consiste en el examen sistemático del contenido del test para determinar si comprende una muestra representativa de la forma de conducta que ha de medirse. Como este procedimiento es más común y aplicable para valorar pruebas de rendimiento escolar, ya sea de una habilidad específica o un curso de estudio, consideré que no era este el criterio adecuado para medir la validez del "Dominós".

Validez Predictiva.- También en este caso el nombre está indicando en que consiste este procedimiento y es la efectividad del test con una medida directa de la ejecución posterior de los sujetos.

Este tipo de información sobre la validez es el más importante para los tests utilizados en la selección y clasificación de personal y en la selección de estudiantes para ingreso a la Universidad o a Escuelas Profesionales.

De acuerdo al tipo de prueba que se estaba estudiando y también a la aplicación que tenía "selección profesional", este hubiera sido el procedimiento más adecuado para medir la validez. Pero tomando en cuenta que en un elevado porcentaje de casos, los sujetos a los que se aplicó la prueba estaban adscritos en diversos lugares de la República, no resultaba tarea fácil obtener los datos de rendimiento en el desempeño de sus labores, dato estrictamente necesario para establecer la validez de la prueba al correlacionarlo con los resultados obtenidos en Anstey, motivo por-

el que se tuvo que prescindir de la utilización de este procedimiento para la obtención de la validez. Por lo tanto recurrí al procedimiento sustitutivo de la validez predictiva que es el procedimiento llamado "validez concurrente", en el que se aplica la prueba a un grupo sobre el cual ya se tienen disponibles los datos de criterio. De este modo, las puntuaciones de tests de los estudiantes universitarios, se pueden comparar con su promedio acumulativo durante el curso de la carrera. De esta manera y de acuerdo con la muestra que mencioné en el capítulo referente a la población estudiada, procedí a la obtención de la validez de la prueba mediante este procedimiento.

De la población total, 263 estudiantes a quienes se aplicó la prueba, fueron 157 los que reunieron las estipulaciones de validación, señaladas en el procedimiento para encontrar la validez concurrente, es decir, dado que el 85% de la población profesional estudiada corresponde a ingenieros y ya que todos los ingenieros de la Comisión Federal de Electricidad, han cursado sus estudios en diferentes escuelas superiores, se entendió autorizado efectuar la validación de la prueba con estudiantes del último año (5o.) de la Facultad de Ingeniería de la Universidad Nacional Autónoma de México, siempre que el número de casos de la población general de estudiantes de dicho nivel, fuera representativo.

En segundo lugar y siempre dentro del mismo procedimiento, de validez concurrente, comparar las medias aritméticas obtenidas

de los promedios de calificación de los grupos extremos de la -- muestra estudiantil, es decir, el grupo integrado por sujetos con los más altos puntajes en Anstey y el grupo integrado por sujetos con los más bajos puntajes en Anstey. Los resultados obtenidos en ambos métodos son los siguientes:

Después de efectuar la correlación en los 157 casos de estudiantes que reunían las condiciones estipuladas, esto es, tener el promedio escolar de toda la carrera y haber resuelto la prueba el resultado fue una correlación de .31 . Esta es una correlación que a primera vista se antoja baja, pero tomando en cuenta varios factores, el criterio para juzgar esta correlación cambia totalmente.

En primer lugar se sabe, según estudios experimentales de este tipo de casos, realizados por Burt, (al que cita S. Pacoud - en el 3er. libro de Psicología Aplicada), en donde observó que: - "El test pierde sensibilidad de diferenciación a medida que se eleva, gracias a su empleo, el nivel del grupo seleccionado. De -- esto resulta una disminución de su validez".

También estos autores han encontrado que el coeficiente de correlación para obtener la validez de una prueba es significativamente más elevado cuando se obtiene en poblaciones globales o - generales y disminuye sensiblemente conforme la población es más homogénea y seleccionada, casos en los que en la práctica casi -- siempre se obtienen coeficientes que varían entre 0.30 y 0.50

Este fenómeno se comprobó en el grupo seleccionado, objeto de este estudio, esto es, una población de 750 profesionistas, en la cual el 85% eran ingenieros. Por lo tanto, el coeficiente de correlación de 0.31 se considera significativo y puede ser suficiente para afirmar que la prueba es válida para esta población.

Sin embargo, pensé que valdría la pena utilizar el segundo método de validación anteriormente citado, solamente para confirmar estos resultados.

En este procedimiento sólo se tomaron en cuenta los 39 casos de sujetos con más elevado rendimiento en la prueba y los 38 casos de sujetos con más bajo rendimiento en dicha prueba. Con estos grupos se procedió a obtener las diferencias, utilizando el método de diferencias de la media a través del índice de diferenciación t , para saber si eran estadísticamente significativas.

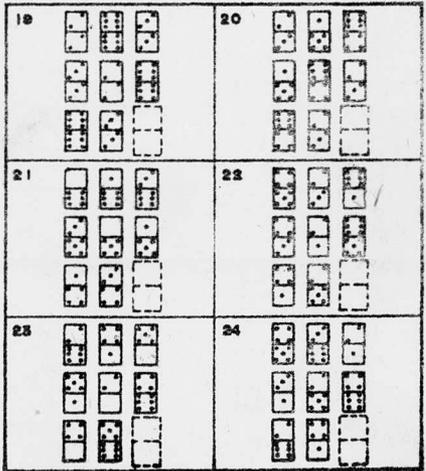
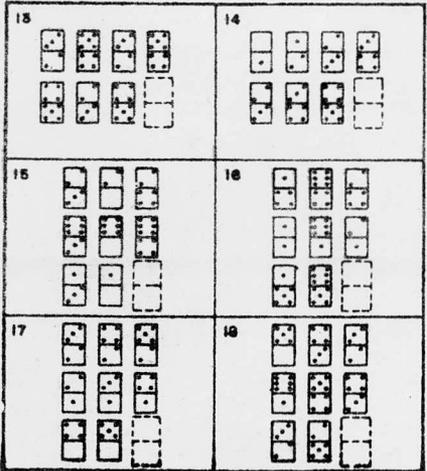
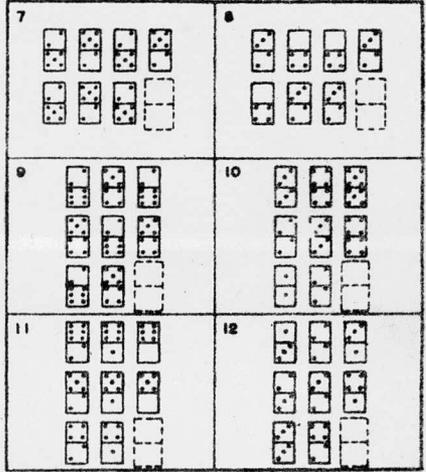
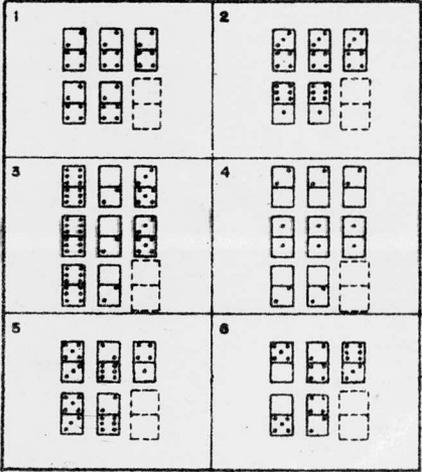
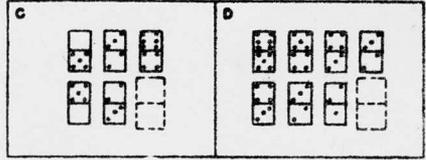
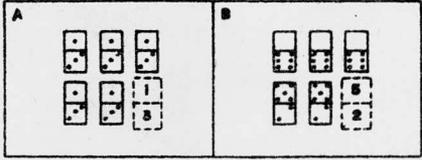
En el cuadro siguiente, presento los datos que se utilizaron.

C U A D R O III

PROMEDIO DE RENDIMIENTO ESCOLAR CORRESPONDIENTE A LOS CASOS MAS ELEVADOS Y A LOS MAS BAJOS

Puntaje Alto			Puntaje Bajo		
Calif.	Esc.	Núm.	Calif.	Esc.	Núm.
6		7	6		12
7		9	7		20
8		19	8		6
9		4	-		-

EJEMPLOS



NOMBRE _____				PUNTAJE _____	
FECHA DE NACIMIENTO		FECHA		EDAD _____	
DIA		MES		AÑO	
HORA AL INICIAR _____			DURACION _____		
HORA AL TERMINAR _____			RANHO _____		

25	26
27	28
29	30

31	32
33	34
35	36

37	38
39	40
41	42

43	44
45	46
47	48

Utilizando la fórmula: $t = \frac{M_1 - M_2}{\sigma_{dif.}}$ la t obtenida --

fue de 3.04 y tomando en cuenta dos grados de libertad, la diferencia se considera significativa estadísticamente a un nivel del 0.001 por lo que podemos concluir que la prueba de Anstey es capaz de discriminar perfectamente los sujetos capaces de los incapaces y por ende, es una prueba válida para medir el factor "G" - de inteligencia.

ANÁLISIS DEL GRADO DE DIFICULTAD DE LOS REACTIVOS

El tercer aspecto del análisis estadístico de la prueba, - es el correspondiente al análisis del grado de dificultad de los reactivos.

Durante todo el periodo de aplicación de la prueba, las observaciones que pude hacer me llevaron a la formulación de una -- hipótesis a comprobar; que el orden de los reactivos, tal y como los da el autor, no corresponde en nuestra población a una escala de dificultad creciente, por lo que me propuse hacer el estudio - estadístico de la consistencia interna de la prueba y verificar - con élllo esta hipótesis.

A continuación presento un modelo de la prueba con el orden original.

Ahora procederé a presentar los resultados obtenidos en el análisis de la prueba y con base a ello, proponer un cambio de orden de los reactivos, en la que se obtendrá como resultado una escala de dificultad creciente útil a nuestra población, proposi-

ción que es el objetivo primordial de este estudio.

El análisis de elementos o de reactivos, frecuentemente se realiza por comparación con la puntuación total del test mismo, - es decir, se calcula la correlación biserial entre la ejecución - de cada elemento y la puntuación total y de acuerdo a estos resultados, seleccionar los mejores elementos para formar el test.

Pero este tipo de análisis se utiliza generalmente cuando se trata de un test recién estructurado y en el que no se sabe si los reactivos que lo forman son los adecuados.

En el caso de la prueba de "Dominós" no sucede esto. La -- prueba está perfectamente estructurada, sólo que la idea del au-- tor fue la de proporcionar una escala de dificultad creciente y - seguramente en la población en la que él la utilizó funcionó como tal, pero en nuestra población no es así; por tal motivo, decidí hacer un análisis del grado de dificultad de cada uno de los reactivos, no para eliminar algunos, sino simplemente para proponer - un orden de dificultad creciente de acuerdo a nuestra población.- De esta manera, lo que se llevó a cabo no fue precisamente un análisis de elementos, sino más bien, un análisis del grado de dificultad de los reactivos.

Por lo tanto, inicié el trabajo, obteniendo el porcentaje de respuestas acertadas a cada uno de los cuarenta y ocho reactivos. El resultado fue muy claro, como puede verse en el cuadro número cuatro.

C U A D R O I V

PROPORCION DE LOS SUJETOS QUE DIERON UNA
RESPUESTA CORRECTA

Número de Sujetos que Contestaron Correctamente

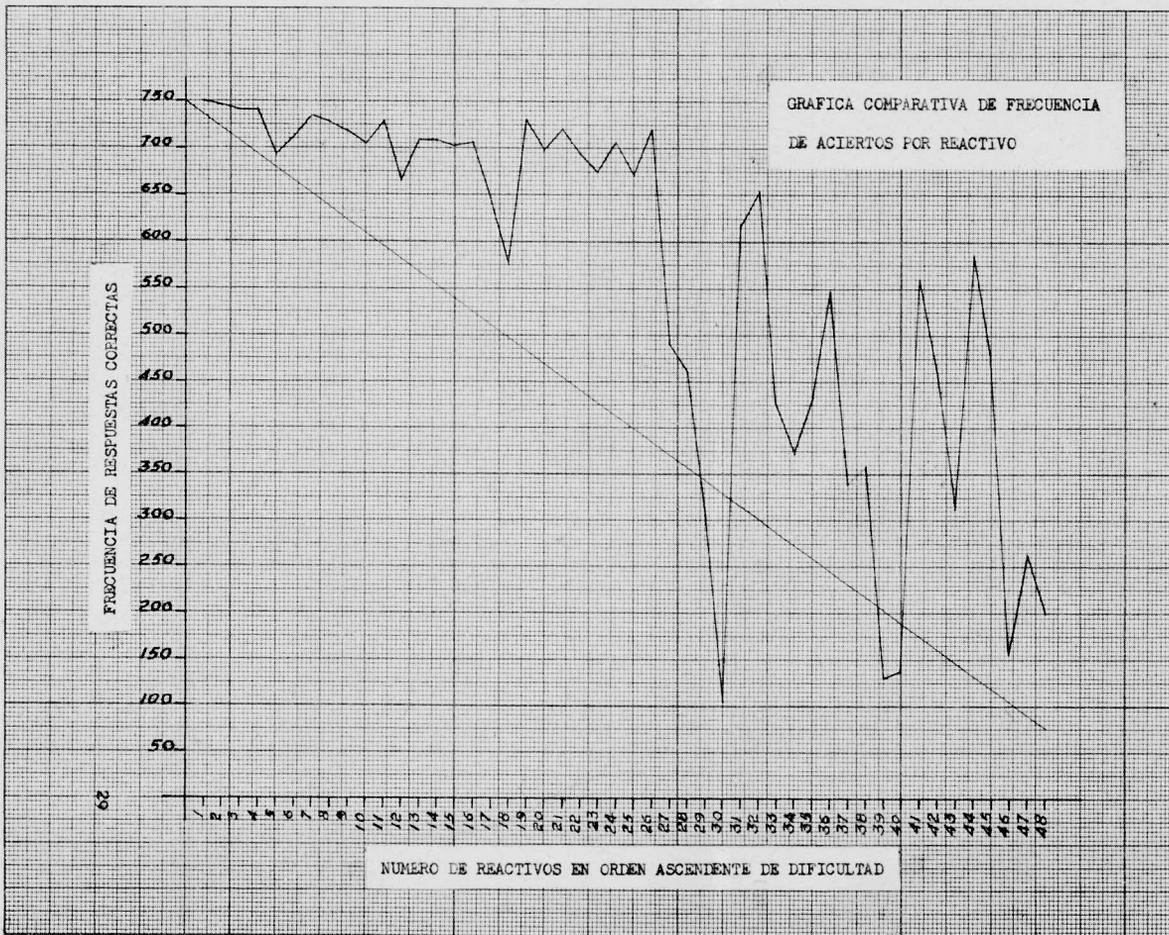
Reactivo	Núm.	%	Reactivo	Núm.	%
1	748	99.7	25	673	89.7
2	745	99.0	26	718	95.7
3	741	98.8	27	491	65.4
4	740	98.6	28	461	61.4
5	692	92.3	29	306	40.8
6	711	94.8	30	105	14.0
7	735	98.0	31	617	82.2
8	727	96.9	32	653	87.1
9	718	95.7	33	426	56.8
10	704	93.8	34	373	51.2
11	728	97.0	35	433	59.5
12	665	88.6	36	548	75.3
13	707	94.2	37	340	46.7
14	708	94.4	38	358	49.2
15	702	93.6	39	130	17.7
16	706	94.1	40	137	18.8
17	647	86.2	41	561	77.6
18	576	76.8	42	461	63.3

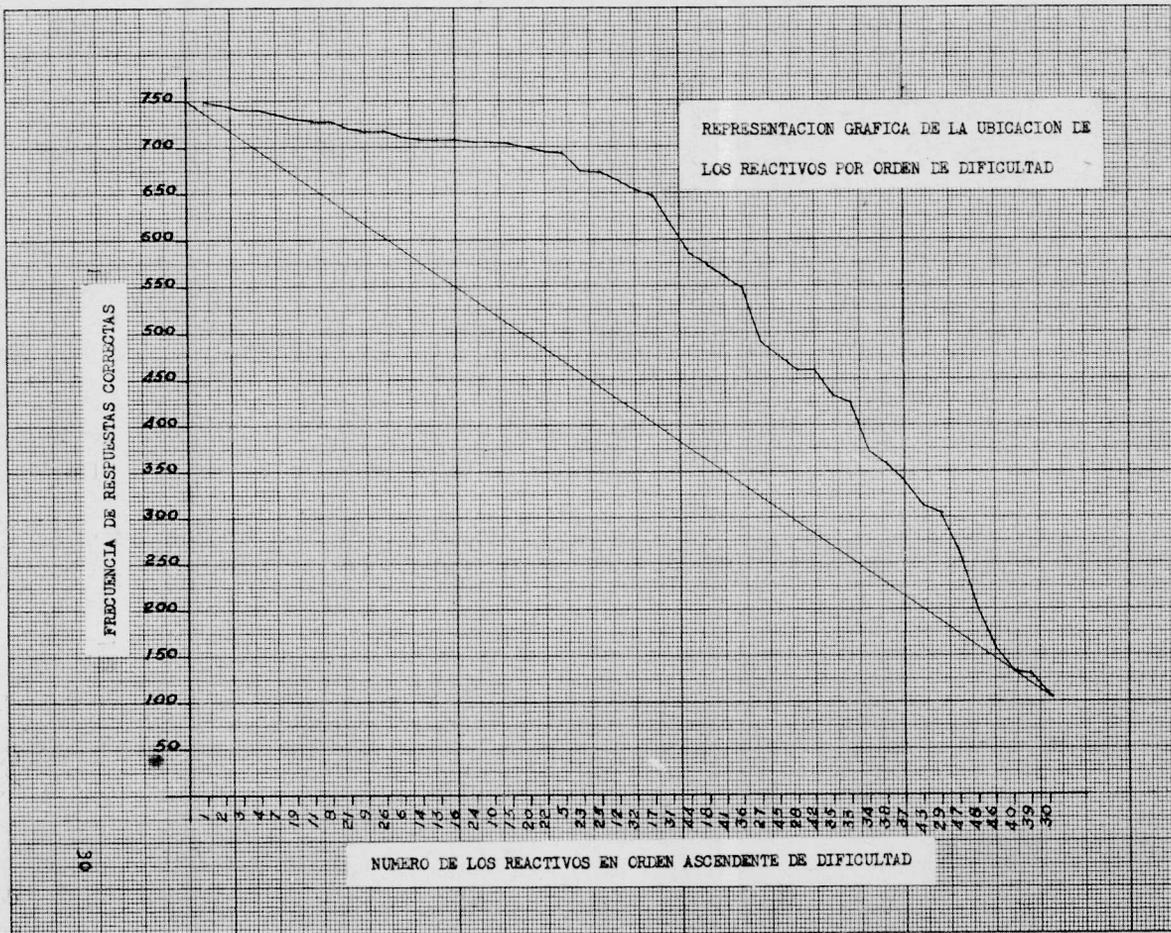
Reactivo	Núm.	%	Reactivo	Núm.	%
19	730	97.3	43	312	42.9
20	699	93.2	44	584	80.2
21	721	96.1	45	476	65.4
22	695	92.6	46	158	21.7
23	674	89.8	47	263	36.1
24	705	94.0	48	200	26.6

En el cuadro se observa como los porcentajes de respuestas correctas no disminuyen a medida que se avanza en el orden de los reactivos, esto es, el número de personas de la muestra que resuelven cada reactivo, debería de ser menor conforme se avanza en la prueba ya que se trata de una escala de dificultad creciente, pero no sucede así, los porcentajes aumentan y disminuyen de manera irregular. Este mismo fenómeno es más claro representado gráficamente. (Gráfica num. 1)

En la gráfica, la línea diagonal representa hipotéticamente el orden descendente de respuestas esperado. La línea quebrada es la que se obtuvo de acuerdo a las respuestas de la población que se estudió.

En una segunda gráfica se observa como la línea trazada de acuerdo a las respuestas correctas a la prueba y ya ordenadas en la forma que se propone, es decir, de acuerdo a un orden de dificultad creciente real para nuestra población; esta línea sigue u-





na trayectoria más afín a la línea hipotética trazada de acuerdo al orden descendente de respuestas esperado.

Estos resultados podrían haber sido razón suficiente para haber propuesto un orden diferente de reactivos que fuera verdaderamente de dificultad creciente para nuestra población, pero para presentar un trabajo más válido estadísticamente, decidí obtener por medio del método de la χ^2 y utilizando la siguiente fórmula:-

$$\chi^2 = \frac{N (AD - BC)^2}{(A + B) (C + D) (A + C) (B + D)}$$

Las diferencias significativas de cada uno de los reactivos con respecto a los 47 restantes, es decir, encontrar si existía diferencia significativa entre el reactivo número 1 y el número 2, entre el reactivo 1 y el 3, entre el 1 y el 4 y así sucesivamente, reactivo por reactivo; el número 2 con el 3, el 2 con el número 4, etc.

Los resultados obtenidos fueron muy elocuentes y marcaron definitivamente los reactivos que resultan de mayor dificultad para nuestra población. De acuerdo a estos datos, procedí a organizar el orden de dificultad de cada reactivo.

Tomando como base dos reactivos, por una parte el de más fácil solución que resultó ser el número 1 y por otra, el de mayor dificultad que resultó ser el número 30, hice una jerarquización de la χ^2 ; ascendente con respecto al reactivo número 1 y descendente con respecto al número 30 y elaboré el siguiente cuadro, en el que pueden observarse resultados idénticos y por ende-

colocan a cada reactivo en el lugar que le corresponde de acuerdo a su grado de dificultad.

C U A D R O V

CUADRO COMPARATIVO DE REACTIVOS SEGUN EL ORDEN ORIGINAL CON EL ORDEN PROPUESTO.

Reactivo Núm. 1		Orden Propuesto		Reactivo Núm. 30
X^2	Orden Original		Orden Original	X^2
0.0	1	1	1	561.9
0.6	2	2	2	556.1
2.2	3	3	3	548.3
2.7	4	4	4	546.4
5.0	7	5	7	536.9
7.5	19	6	19	527.6
8.5	11	7	11	523.9
9.0	8	8	8	522.1
12.0	21	9	21	511.2
13.5	9	10	9	505.9
13.5	26	11	26	505.8
17.2	6	12	6	493.5
18.3	14	13	14	488.2
19.2	13	14	13	486.5
19.8	16	15	16	484.8
20.3	24	16	24	483.1

Reactivo Núm. 1		Orden Propuesto	Reactivo Núm. 30	
χ^2	Orden Original		Orden Original	χ^2
20.8	10	17	10	481.3
21.9	15	18	15	478.0
23.5	20	19	20	472.9
25.6	22	20	22	466.2
27.2	5	21	5	461.2
37.2	23	22	23	432.3
37.6	25	23	25	430.8
42.0	12	24	12	418.4
48.8	32	25	32	400.4
52.2	17	26	17	391.7
69.8	31	27	31	350.0
90.1	44	28	44	307.9
95.2	18	29	18	298.3
104.9	41	30	41	280.8
113.5	36	31	36	266.1
153.2	27	32	27	207.4
164.2	45	33	45	193.3
175.6	42	34	42	179.8
175.6	28	35	28	179.8
197.5	35	36	35	155.9
203.2	33	37	33	150.2
248.2	34	38	34	110.3

Reactivo Núm. 1		Orden Propuesto	Reactivo Núm. 30	
X^2	Orden Original		Orden Original	X^2
261.8	38	39	38	100.0
278.5	37	40	37	88.2
305.7	43	41	43	71.2
311.7	29	42	29	68.0
356.8	47	43	47	44.9
397.6	48	44	48	18.6
485.1	46	45	46	6.5
514.4	40	46	40	2.5
524.5	39	47	39	1.6
561.9	30	48	30	0.0

Como puede verse en el cuadro anterior, al colocar la X^2 en orden ascendente con respecto al reactivo número 1 que resultó ser el más sencillo y al colocar la X^2 en orden descendente de acuerdo al reactivo número 30 que resultó ser el más difícil, los reactivos automáticamente quedan en el orden correspondiente al grado de dificultad que representan.

Tomando en cuenta estos resultados, el orden de los reactivos quedaría de la siguiente manera.

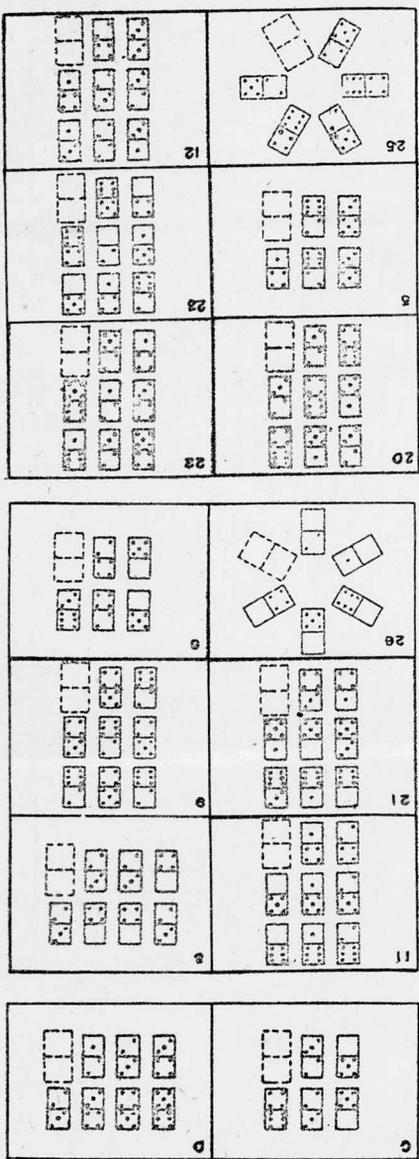
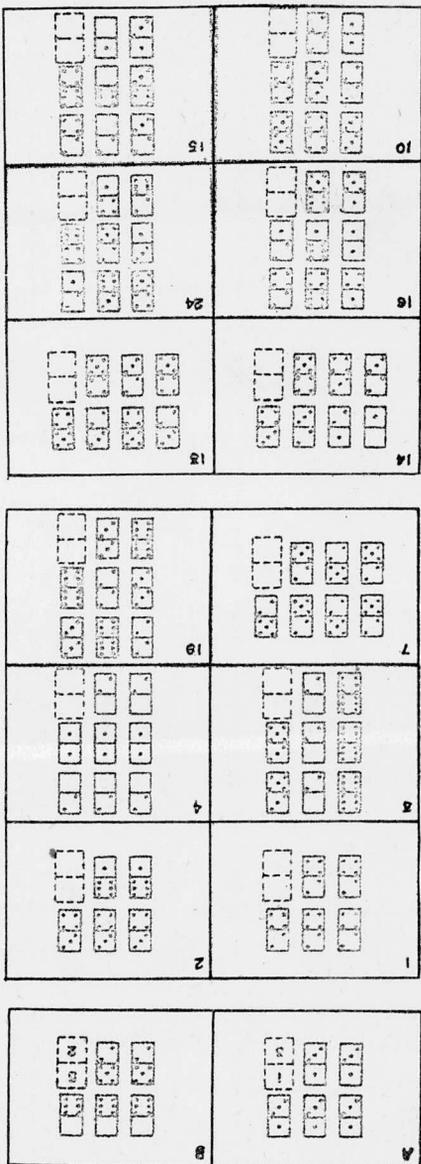
ORDEN DE REACTIVOS QUE SE PROPONE COMPARATIVAMENTE
AL ORDEN ORIGINAL

Orden Propuesto	Orden Original	Orden Propuesto	Orden Original
1	1	25	32
2	2	26	17
3	3	27	31
4	4	28	44
5	7	29	18
6	19	30	41
7	11	31	36
8	8	32	27
9	21	33	45
10	9	34	42
11	26	35	28
12	6	36	35
13	14	37	33
14	13	38	34
15	16	39	38
16	24	40	37
17	10	41	43
18	15	42	29
19	20	43	47
20	22	44	48

Orden Propuesto	Orden Original	Orden Propuesto	Orden Original
21	5	45	46
22	23	46	40
23	25	47	39
24	12	48	30

Al cambiar el orden de los reactivos de acuerdo a los resultados obtenidos, la prueba quedaría de la siguiente manera. -- Anexo aquí un modelo de la prueba tal como la propongo, como una escala ya de dificultad creciente para nuestra población.

Finalmente, quisiera agregar que el objeto principal del análisis de los reactivos es fundamentalmente desde un punto de vista económico, es decir, que a pesar de que el "Dominós" es una prueba de poder y este tipo de test, no tiene límite de tiempo, podríamos fijarle de esta manera un tiempo límite de ejecución y aún así obtener datos que sí discriminaran perfectamente, ya que como expuse en el capítulo en el que hablé acerca del tiempo de ejecución, nuestros sujetos utilizan demasiado en su resolución (1 hora), debido probablemente a que cada vez que tropiezan con un reactivo más difícil que los anteriores, no son capaces de saltarlo y seguir adelante, sino que se estacionan en él hasta que no lo resuelven; de esta forma, no sólo utilizan más tiempo, sino que la fatiga no les permite resolver después los reactivos siguientes, aún tratándose de problemas de más fácil solución.



32		17	
31		44	
18		41	

36		27	
45		28	
42		35	

33		34	
38		37	
43		29	

47		48	
46		40	
39		30	

NOMBRE _____		PUNTAJE _____	
FECHA DE NACIMIENTO _____		FECHA _____	
_____ DIA	_____ MES	_____ AÑO	EDAD _____
HORA AL INICIAR _____		DURACION _____	
HORA AL TERMINAR _____		RANGO _____	

De esta suerte, al cambiar el orden de los reactivos en la prueba, la fatiga no podría ser un factor que interfiera, pues el sujeto va entrando poco a poco dentro del test, hasta que al final se encuentra con los problemas más difíciles, pero ya para entonces está funcionando en él un proceso de aprendizaje que le va indicando que cada problema va a ser más difícil que el anterior.

Este es pues el objetivo principal del análisis. Ahora -- bien, como segundo objetivo, tenemos el hecho de que cuando una prueba es estudiada para ser aplicada a un país o población diferentes para quienes fue creada, este estudio contiene no sólo los aspectos analíticos estadísticos de ella, sino que tiene que sufrir las adaptaciones necesarias para que sea un instrumento aplicable válidamente en dicha población.

Finalmente, sólo me resta decir que espero muy pronto tener resultados de la aplicación de la prueba de "Dominós" con las modificaciones de orden en los reactivos que ahora estoy proponiendo.

Capítulo IV

NORMAS DE CALIFICACION

Después de haber efectuado el estudio analítico de la prueba de Anstey y haber obtenido una confiabilidad tan alta, comprobar su validez, de haber podido proponer, basándome en los resultados obtenidos, un cambio en el orden de los reactivos, y el hecho de comprobar que las curvas de puntaje obtenidas en este estudio, comprueban o confirman la teoría de inteligencia factor "G"-Curvas que a continuación presento.

Gráfica 1.- (Curva general de puntaje por intervalos de edad en la población estudiada), 750 casos.

La curva producida por nuestra población en intervalos de edad, muestra la tendencia clásica que establece la teoría del factor "G" de inteligencia.

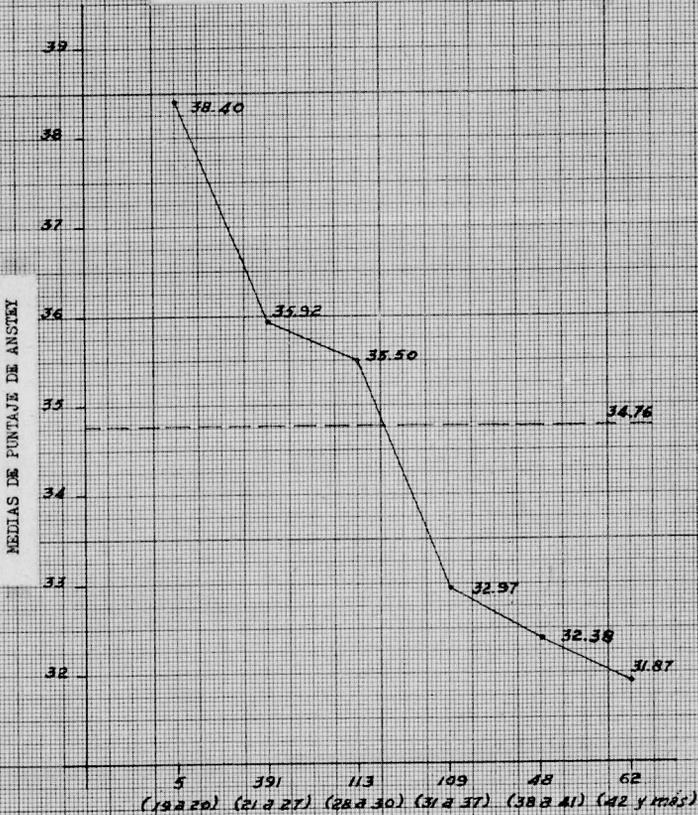
Wechsler (1) y Spearman (2), expresan que es válido concluir que el factor "G" de inteligencia se mantiene más o menos uniforme hasta los 30 años, lo cual obedece a que, generalmente, hasta esa edad, aún no se han desarrollado funciones intelectivas específicas.

En la gráfica número 1 se observa que después del intervalo de edad de 28 a 30 años, la curva cae súbitamente hasta quedar debajo de la media (34,76), lo cual teóricamente significa que, en términos generales llegando a los 30 años de edad, la influencia de la experiencia, las tendencias vocacionales y la especia--

(1) Wechsler.- "La medición de la Inteligencia del Adulto" cap. 6 p.83

(2) Spearman.- "Manual del Test de Dominós".- E. Anstey.

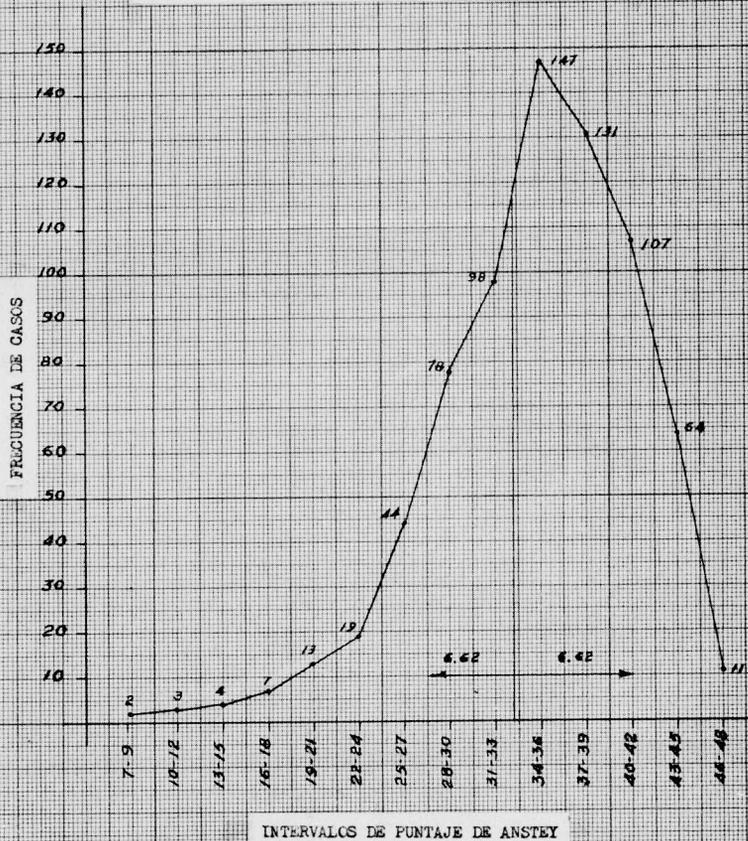
CURVA GENERAL DE PUNTAJE POR INTERVALOS DE
 EDAD DE LA POBLACION ESTUDIADA (750 CASOS)



NUMERO DE CASOS

INTERVALOS DE EDAD

CURVA DE DISTRIBUCION DE RENDIMIENTO INTELECTUAL
 DE LA POBLACION ESTUDIADA (750 CASOS DE PROFESIO
 NISTAS DE LA COMISION FEDERAL DE ELECTRICIDAD)



lización ocupacional, han desarrollado y perfeccionado las funciones específicas que el individuo ha venido utilizando en el curso de los años, pero ello simultáneamente abate el factor "G" de inteligencia. La elevación de la capacidad en esas funciones mentales específicas, redundante en detrimento del factor "G", independientemente de que la teoría también establece que la dotación natural llega a un nivel óptimo a la edad de 15 o 16 años (algunos autores la señalan hasta los 20) y a partir de esa edad va decreciendo gradualmente.

Gráfica 2.- Curva de distribución de rendimiento intelectual de la población estudiada (750 casos de trabajadores de nivel profesional).

Realizado el estudio de la curva de distribución de frecuencia obtenida con los 750 casos de la población de profesionistas de la Comisión Federal de Electricidad por medio del método estadístico, de la χ^2 , se encontró que se trata de una distribución de frecuencia normal en un nivel de confianza estadístico del 0.01 cifra que permite concluir que la curva obtenida en el estudio realizado es una curva casi normal comparada con la curva normal ideal de Gaus.

De la curva obtenida se desprende que se trata de una población muy homogénea y con rendimiento intelectual muy elevado, lo cual es lógico, considerando que se trata de población de nivel profesional especializado en una de las ramas más exigentes-escolarmente como lo es Ingeniería (85% de la población estudiada).

Con todo esto, se reafirmó en mí el deseo de obtener normas de calificación aplicables a nuestra población, pues ahora más que nunca tengo la certeza de que la prueba es un instrumento útil dentro de la Psicología Aplicada.

Por este motivo me dí a la tarea de obtener las normas de calificación válidas para una población profesional mexicana.

El ideal hubiera sido obtener normas para una población general mexicana, pero desafortunadamente no cuento con los medios y el tiempo suficientes para llevar a cabo la tarea que significaría obtener dichas normas. Sin embargo, de la fecha en que se inició este estudio, hasta el momento, la prueba se ha venido aplicando a diferentes niveles. Por ejemplo, se ha aplicado a jóvenes de nivel medio superior, es decir del ciclo escolar vocacional y pronto obtendremos los resultados finales.

Por otra parte, pienso aplicar esta prueba en un nivel escolar de enseñanza media, esto es, nivel secundario, pues creo firmemente que el conocer las normas de calificación para este nivel es de gran importancia pues con esta prueba podríamos conocer la capacidad de inteligencia general factor "G" de nuestros jóvenes estudiantes y de acuerdo a ella, encauzarlos hacia el estudio de la carrera más idónea.

Cabe hacer la aclaración de que de cuatro años a la fecha me he especializado en la rama de Orientación Profesional y por tal razón pienso que para efectuar una verdadera orientación se-

requiere de varios factores. Uno de ellos es precisamente el contar con instrumentos psicométricos válidos que nos permitan conocer las capacidades, aptitudes, intereses y rasgos de personalidad de nuestros sujetos de estudio. De tal suerte, que esta prueba puede pasar a formar parte de una batería de pruebas para un estudio de orientación profesional, por lo que es indispensable obtener normas de calificación en diversos niveles.

Haré por tanto una descripción de lo que son las normas de calificación de una prueba y, en este caso, cuales fueron los resultados obtenidos en la población estudiada.

Cuando se nos dice que un sujeto ha resuelto 15 problemas de un test de aritmética, o que ha resuelto 20 cuestiones de un test de aptitud mecánica, estos datos no nos dicen nada, porque no tenemos una idea del lugar que ocupa dentro de un grupo determinado, es decir, no existe una relación que nos indique cual es el lugar que le corresponde con respecto a un grupo de personas representativas de una población.

De manera que la calificación bruta debe interpretarse -- siempre en relación a una norma.

Las normas de los tests psicológicos representan los resultados de la ejecución del test de la muestra de tipificación de este modo, las normas siempre se establecen empíricamente, de terminado lo que un grupo representativo de personas hace realmente en el test.

Así, la puntuación en bruto o directa de un individuo es-
constrastada con la distribución de las puntuaciones obtenidas -
en la muestra de tipificación, para ver en que lugar de esta dis-
tribución queda colocado, de manera que nosotros podemos saber:-
si su puntuación coincide con el promedio de ejecución del grupo
normativo, si está ligera o significativamente por debajo de ese
promedio, o si está ligeramente por encima de él, es decir, está
en una situación, de acuerdo a su ejecución, en un nivel de supe-
rioridad.

Ahora bien, para determinar de una manera más precisa la-
posición exacta del individuo con respecto a la muestra de tipi-
ficación, la puntuación directa se reduce a alguna medida rela-
tiva.

Estas medidas relativas o puntuaciones ponderadas, tienen
una doble finalidad. En primer lugar, indican la situación rela-
tiva del individuo en el grupo normativo que permite una valora-
ción de su ejecución con referencia a otras personas. En segundo
nos dan medidas comparables que hacen posible la comparación di-
recta de la actuación del individuo en las distintas pruebas.

Por otra parte, las puntuaciones transformadas pueden ex-
presarse en las mismas unidades y referirse a la misma muestra -
normativa o a otras similares para las distintas pruebas. De es-
te modo pueden compararse la actuación relativa de un individuo-
en funciones muy diversas.

Existen varias formas de convertir las puntuaciones directas de manera que cumplan con los dos requisitos anteriores. Pero fundamentalmente, las puntuaciones de tests pertenecen a tres tipos principales que son:

- 1.- De edad o Cronológicas
- 2.- Percentiles o Deciles
- 3.- Puntuaciones Típicas

PUNTUACIONES DE EDAD O CRONOLOGICAS

Son aquellas en las que los elementos individuales están agrupados en niveles de edad, es decir, aquellos elementos resueltos por la mayoría de los sujetos de una determinada edad se colocarán en el nivel de dicha edad y la calificación obtenida se dará en función de "edad mental". Este tipo de puntuación o de escala tiene varias limitaciones que no permiten sea aplicable a todo tipo de pruebas, una de las limitaciones consiste en que el concepto mismo de la norma de edad, nos indica que su utilidad se limita en gran parte a etapas de desarrollo, pues decir que un niño de edad cronológica de 10 años tiene una edad mental de 12, nos da una idea clara y objetiva de la situación de ese niño con respecto a otros; pero cuando se trata de un adulto que no experimenta cambios notables a partir de los 15 o 16 años, el concepto de edad mental no permite una interpretación tan clara y objetiva. Otra limitación de las puntuaciones de edad mental, es que no permiten una medida para aquellas funciones que mues-

tren un cambio claro y consistente con la edad, como son los rangos de personalidad que presentan poca relación con la edad cronológica.

Debido a estas limitaciones este tipo de puntuación no se utilizó para obtener las normas de calificación o Baremos de la prueba de Anstey.

PUNTUACIONES PERCENTILES

Las puntuaciones percentiles son aquellas que se expresan en función del porcentaje de personas en el grupo normativo, que quedan por debajo de una puntuación directa determinada, esto es pueden considerarse a los percentiles como rangos en un grupo de 100. Teniendo en cuenta que, al situar por rangos, se empieza a contar por arriba, recibiendo el rango 1 la persona mejor del grupo. Con los percentiles se empieza a contar por abajo, de manera que cuanto más bajo sea el percentil, peor será la posición del individuo. Así, las ejecuciones por encima del percentil 50, representan ejecuciones superiores al promedio, y los que están por debajo de 50, ejecuciones inferiores.

Este tipo de puntuación percentil es una de las más frecuentemente utilizadas dentro de la Psicología y precisamente por éllo decidí utilizarla como norma de calificación de la prueba de Anstey.

PUNTUACION TIPICA

La puntuación típica es aquella que expresa la distancia-

del individuo en relación a la media, en función de la desviación típica de la distribución.

Existen dos tipos de puntuaciones típicas: las lineales y las normalizadas. Dentro de estas últimas, existe un tipo de transformación muy conocida que es la escala "estanona", empleada por las Fuerzas Aéreas de E.E.U.U. durante la Segunda Guerra Mundial. Esta escala proporciona un sistema de puntuaciones de un solo dígito, con una media de 5 y una DT aproximadamente de 2^2 . El nombre de "estanona" se basa en el hecho de que las puntuaciones van del 1 al 9. Así, las puntuaciones directas pueden convertirse fácilmente en estanonos, disponiendo las puntuaciones originales por orden de tamaño y asignando después estanonos de acuerdo a los porcentajes de la curva normal, como se ilustra en el siguiente cuadro.

C U A D R O V I

PORCENTAJES DE LA CURVA NORMAL PARA USARLOS
EN LA CONVERSION A ESTANONOS

Porcentaje	4	7	12	17	20	17	12	7	4
Estanono	1	2	3	4	5	6	7	8	9

Este tipo de puntuación también es ampliamente aceptado ^{dentro} de la Psicología, por lo que creí conveniente utilizarla como norma de calificación para la prueba de Anstey.

Después de estudiados los diferentes tipos de puntuación-- utilizados para obtener normas de calificación de una prueba, elegí, por ser las más comunmente usadas, las escalas de puntuación percentilar y de estanonos.

Obtuve así en la escala de percentiles dos Baremos. El -- primero de 19 a 30 años y el segundo de 31 en adelante. Estos -- dos Baremos se establecieron partiendo de la base de que no se -- encontraron diferencias estadísticamente significativas entre -- las edades de 19 a 30 años, esto es, utilizando la prueba de diferenciación t, se compararon las ejecuciones de sujetos de 19 -- años con los de 20, sujetos de 20 con los de 21 años y así sucesivamente, no encontrándose diferencias significativas estadísticamente en el intervalo de edades entre 19 a 30 años. Por lo tanto el primer Baremo corresponde a este intervalo.

A partir de los 31 años, se efectuó esta misma comparación -- y no se encontró diferencia significativa dentro de los sujetos de prueba.

Tomando en cuenta que este tipo de pruebas normalmente no se aplica a personas mayores de esta edad, consideré adecuado de -- jar abierto el intervalo del segundo Baremo. Así, este agrupa a -- sujetos de 31 años en adelante y los resultados obtenidos se pre -- sentan en el cuadro siguiente.

C U A D R O V I I
 EVALUACION DE ANSTEY
 BAREMOS MEXICANOS PARA POBLACION
 DE NIVEL PROFESIONAL

Percentiles	E D A D	
	19 a 30	31 en adelante
95	45	44
90	44	41
75	41	38
50	37	34
25	32	29
10	28	23
5	26	19

Estanonos	E D A D	
	19 a 30	
1	24	
2	29	
3	32	
4	35	
5	38	
6	41	
7	43	
8	46	
9		

De esta manera quedan establecidos Baremos de califica- -
ción para la prueba Edgard Anstey en población Mexicana de Nivel
Profesional.

Conclusiones

En el desarrollo del presente trabajo se logró un coeficiente de correlación de 0.91 en la Confiabilidad y uno de 0.31 de correlación entre los promedios escolares de estudiantes del 5o. año de la Facultad de Ingeniería de la U.N.A.M. y la ejecución de ellos mismos en la prueba de Anstey, dando un dato que justifica la Validez de la misma, el cual se confirma al encontrar a través del método de diferencias en la media que la prueba es capaz de discriminar a los sujetos capaces de los no capaces; proponer un cambio en el orden de los reactivos para obtener con esto una mejor medida de la inteligencia factor "G" en nuestra población de nivel profesional, concluyendo en la realización de los objetivos propuestos a través de:

- I El análisis estadístico de la prueba en sus tres aspectos: - Confiabilidad, Validez y Análisis del Grado de Dificultad de los Reactivos de la Prueba, se obtuvo totalmente.
- II La obtención de las normas de calificación para una población mexicana de nivel profesional.
- III La prueba, con los cambios propuestos, puede pasar ahora a formar parte de cualquier batería de estudio, ya sea de Selección o de Orientación Profesional y ser así un instrumento más adecuado para nuestra población.

B i b l i o g r a f í a

- Anastassi, Anne.- Psychological Testing.- New York.- Ed. The Macmillan Company.- 1955
- Bernstein y Risso.- Test de Dominós.- Buenos Aires.- Ed. Paidós.- 1959
- Binois Rene.- La Psychologie Appliquée.- París.- Ed. Presses Universitaires de France.- 1963
- Garret, Henry.- Las Grandes Realizaciones de la Psicología Experimental.- México.- Ed. Fondo de Cultura Económica.- 1958
- Nasse, Rocco Amelia.- "Experiencia Recogida con la Aplicación de la Prueba de Dominós, de 1956 a 1961".- 2a Jornada Psicológica.- Uruguay, Octubre -- 1961
- Palmade, Guy.- La Psychotechnique.- París.- Ed. Presses Universitaires de France.- 1963
- Picho, Pierre.- Les Tests Mentales.- Buenos Aires.- Ed. Paidós.- 1963
- Piéron, Henri. et al.- Tratado de Psicología Aplicada.- V. III.- Buenos Aires.- Ed. Kapelusz.- 1960
- Spearman, Carlos.- Las Habilidades del Hombre.- Buenos Aires.- Ed. Paidós.- 1965
- Székely, Bela.- Diccionario Encicloédico de la Psique.- Buenos Aires.- Ed. Claridad S.A.- 1958
- Wechsler, David.- La Medición de la Inteligencia del Adulto.- Ed. Universidad de New York.- 1960

