

8292

Claudia [Hernández Zavala

← "Las pruebas psicológicas en el campo educativo" : uso en la detección y diagnóstico de los problemas de aprendizaje en México

**INDICE**



	Pag.
INTRODUCCION .....	1
CAPITULO 1 EVALUACION .....	5
1.1 Conceptos y Definiciones .....	6
1.2 Evaluación como Medición .....	9
1.3 Evaluación con Referencia a Normas .....	11
1.4 Evaluación con Referencia a Criterios .....	12
CAPITULO 2 INSTRUMENTOS DE MEDICION EN PSICOLOGIA Y EDUCACION .....	13
2.1 Historia de las Pruebas Psicológicas .....	14
2.2 Definición y Clasificación .....	20
CAPITULO 3 ASPECTOS PSICOMETRICOS DE LOS INSTRUMENTOS DE MEDICION .....	30
3.1 Principios fundamentales de la Medición .....	31
3.1.1 Error en la Medición .....	35
3.1.2 Confiabilidad .....	35
3.1.3 Validez .....	37
3.1.4 Normatividad .....	39
CAPITULO 4 OBJETIVOS Y METODOLOGIA .....	41
CAPITULO 5 PRUEBAS PSICOLOGICAS .....	45
5.1 Escala de Desarrollo de Arnold Gesell .....	47
5.2 Prueba Guestáltica Visomotora de Lauretta Bender .....	53
5.3 Prueba Monterrey .....	57
5.4 Escala de Inteligencia de Stanford-Binet (Terman-Merrill 1960)	62
5.5 Prueba de la Figura Humana de Florence Goodenough .....	67
5.6 Escalas de Inteligencia Wechsler (Niveles Preescolar y Primaria) .....	76
5.7 Batería de Evaluación Intelectual Secuencial y Simultánea Kaufman .....	84
5.8 Sistema de Evaluación Multicultural y Pluralístico S.O.M.P.A. .	88

	Pag.
<b>CAPITULO 6 REVISION Y UTILIZACION DE LAS PRUEBAS PSICOLOGICAS EN EL CAMPO EDUCATIVO EN MEXICO .....</b>	<b>94</b>
<b>Como Instrumento de Investigación</b>	
<b>Como Instrumento de Psicodiagnóstico</b>	
<b>Cuadros de Análisis y Comentarios .....</b>	<b>148</b>
<b>CONCLUSIONES .....</b>	<b>159</b>
<b>SUGERENCIAS .....</b>	<b>163</b>
<b>ANEXO .....</b>	<b>165</b>
<b>BIBLIOGRAFIA .....</b>	<b>167</b>

## ***INTRODUCCION***

En los últimos años el sistema educativo de nuestro país ha tenido diversos cambios. Los programas de estudio han sido revisados en forma continua en un intento de actualizar y optimizar los sistemas de enseñanza para el logro de sus objetivos.

La educación básica tiene como finalidad lograr el desarrollo armónico e integral de los educandos, por ello considera necesario orientar e impulsar la creación de proyectos que beneficien la acción educativa. Sin embargo hay alumnos que van quedando rezagados o presentan serias dificultades durante su escolaridad.

A pesar de que éstas dificultades se detectan en su mayoría, al inicio de la escolaridad, no son consecuencia exclusiva de los métodos, planes, programas o maestros, ya que existen diferentes factores inherentes al propio estudiante que inciden directamente en el aprendizaje como pueden ser los trastornos orgánicos, psicológicos, ó los factores del ambiente familiar y socio-cultural en los que el niño está inmerso.

Otros factores que pueden repercutir en las dificultades escolares son los relacionados con la motivación del niño para aprender o los que provienen de la aplicación equivocada y rígida de los programas y métodos de enseñanza.

Por otro lado, también hay escolares que aún contando con una inteligencia adecuada, un buen nivel de madurez y los recursos culturales suficientes, llegan a presentar dificultades para aprender.

Para dar atención a éstos problemas educativos que a diario se presentan en el aula, se han creado, por parte de la Secretaría de Educación Pública (S.E.P.) Centros de Atención Especializada como los Centros de Atención Psicopedagógica de Educación Preescolar (C.A.P.E.P) y los centros de la Dirección General de Educación Especial (D.G.E.E.) encargados de proporcionar atención a los alumnos, que por algún motivo presentan dificultades en el proceso de la enseñanza-aprendizaje.

Los educandos remitidos a éstas Instituciones generalmente presentan problemas en alguna o varias áreas como: en el lenguaje, percepción y coordinación, conducta, atención, concentración, dificultades en la adquisición de la lecto-escritura y el cálculo matemático, lo que limita o entorpece su aprovechamiento.

El objetivo de éstas Instituciones es proporcionar al niño la atención integral que requiera, acorde a su problemática, para incorporarse así al sistema regular de enseñanza sin ninguna dificultad.

Para la detección de la ó las problemáticas que obtaculizan el aprendizaje, se utilizan Baterías de Psicodiagnóstico con las cuales se pretende medir las capacidades y habilidades del alumno y, establecer un Diagnóstico Diferencial para canalizarlo al tratamiento que se requiera.

Los diversos programas de atención que proporcionan los Centros, cuentan a su vez con pruebas específicas de área, como son las exploraciones del lenguaje, de psicomotricidad y perceptual, proyectivas, de escritura y matemáticas, entre otras. Una

ez obtenida ésta información, el especialista podrá planear e implementar el programa e atención para cada caso.

En diversas investigaciones se han tratado temas relacionados con las pruebas psicológicas, sin embargo, no se ha llevado a cabo un estudio sistematizado que permita valorar lo adecuado o no de éstos instrumentos de medición, el tipo de población, en la que actualmente son utilizados para la detección y diagnóstico de los problemas que intervienen en el aprendizaje.

Para fundamentar la presente investigación en los capítulos 1 ,2 y 3 se abordará el Marco Teórico, en el que se tratan los temas relacionados con la evaluación y sus diferentes aspectos como son medición, normas y criterios.

Se revisarán los Instrumentos de medición en psicología, partiendo de la historia de las pruebas, su diferenciación y clasificación, exponiendo los aspectos psicométricos fundamentales para la estandarización de las mismas.

En el capítulo 4 se contemplan los objetivos a alcanzar con la investigación, así como a metodología a seguir.

En los capítulos 5 y 6 se presenta la fundamentación teórica de las pruebas a revisar, así como las investigaciones, revisiones y modificaciones que en el ámbito de la Psicometría en México, se han llevado a cabo con cada una de ellas, incluyendo el uso que se les da en el sector educativo.

Para finalizar se realiza el Análisis del material sujeto a investigación con sus conclusiones, comentarios y sugerencias sobre el uso de las pruebas en los Centros mencionados.

*"Todos los niños del mundo siguen el mismo proceso de desarrollo, todos han de nadurar pasando por cada una de las etapas correspondientes. Lo que varía en los primeros años de evolución es la velocidad y ésta depende de las características de cada cultura".*

Diario de Jornadas  
Gabinete Técnico  
de Educación Preescolar.  
Barcelona, España.

# ***CAPITULO 1 : EVALUACION***

## 1.1 CONCEPTOS Y DEFINICIONES

La evaluación es un proceso que se utiliza en distintos ámbitos y bajo diversos motivos. El presente trabajo se circunscribe principalmente al campo educativo en el que se contempla a la evaluación como un proceso dinámico y sistemático, ubicándose como parte integral y fundamental de la acción educativa.

Con la planeación de la enseñanza a través de objetivos, la evaluación se hace cada vez más compleja y necesaria, siendo ésta un proceso que implica mucho más que la realización de una prueba.

Implica un juicio de valor, no puede quedarse en un nivel puramente descriptivo, sino que poco se limita a verificar o refutar una hipótesis. Debe establecerse si el factor que se está evaluando favorece o entorpece el aprendizaje y en que medida lo hace.

La evaluación puede describir un problema, estimar su gravedad, identificar los factores relacionados con él, sugerir áreas de mejoramiento y obtener indicadores para mejorar el plan de tratamiento. (Sattler, 1988:292)

Si se parte de que la evaluación "es el proceso de delineamiento, obtención y organización de información útil para juzgar posibilidades de acción" (Chadwick en P.N., 1987:159) se obtendrán elementos importantes que ponen no solo énfasis en el concepto, sino en lo que se está juzgando, como lo es en la toma de decisiones, particularmente significativas en la actualidad.

En el ámbito educativo, la evaluación permite tomar decisiones concernientes al proceso de enseñanza-aprendizaje, entre otras acciones.

Una de las funciones más importantes de la evaluación educativa es proporcionar un sistema de control de calidad, es decir, vigilar el logro de los objetivos de la instrucción, así como determinar otras acciones a seguir e identificar defectos y/o limitaciones en los planes y programas educativos. (Anderson, et. al., 1977:148)

La función primordial de la escuela y el maestro es dirigir, orientar y facilitar el aprendizaje de los alumnos; el grado de eficiencia con que se efectúe la evaluación, podrá determinar el éxito o fracaso de la enseñanza.

Para Mehrens (1982:21), la evaluación que los maestros llevan a cabo debe ir acompañada a:

- a) Obtener información acerca del comportamiento inicial de sus alumnos.

- ) Fijar y afinar metas realistas para cada uno de ellos.
- ) Evaluar el grado en que se alcanzaron los objetivos.
- ) Determinar, evaluar y perfeccionar las técnicas didácticas.

Existen, en general, dos tipos de evaluación que, de ser utilizadas en la forma adecuada le permitirán al maestro obtener un mejor conocimiento de sus alumnos y por lo tanto un mayor aprovechamiento escolar, estas son: **la Evaluación Informal y la Evaluación Formal.**

**La Evaluación Informal** se refiere a la observación no sistemática de cualquier hecho que sucede en el salón de clases, permitiéndole al maestro tomar decisiones en cualquier momento o a cualquier situación que se le presente, por ej.: cambiar un tema cuando los alumnos pierden interés, observar cuando un estudiante no responde a la enseñanza, proporcionar un nuevo material de apoyo para que se de el aprendizaje, etc...

Este tipo de evaluación es variable en cantidad, ya que no supone ningún registro en especial que permita sistematizarla, sin embargo, en cuanto a calidad, es una valiosa información para ser tomada como punto de partida en la toma de decisiones.

**La Evaluación Formal** cuenta con todos los elementos indispensables para ser sistemática, como son listas de comprobación, registros, pruebas escritas, etc...

Este tipo de evaluación permite obtener información objetiva, supone campos de observación bien definidos y estructurados, sentando las bases para elegir que alternativas son las más adecuadas para el propósito que se persigue.

Estos dos tipos de evaluación son de suma importancia pues le permiten al maestro mejor y oportuna intervención en el proceso de la enseñanza.

La evaluación es "el proceso para definir, obtener y proporcionar la información indispensable para juzgar las alternativas en una decisión." (Mehrens, 1982:25)

Una buena decisión es aquella que se basa en una información precisa y relevante, considerando todas las alternativas, los posibles resultados y la probabilidad y conveniencia de esos resultados.

Para Georgia, (1975:21-22) el proceso de la evaluación, debe considerar cinco aspectos básicos:

1. Determinar lo que se pretende evaluar.
2. Definir lo que se desea evaluar en términos de comportamiento.
3. Seleccionar las situaciones adecuadas para la observación.
4. Registrar el comportamiento del alumno.
5. Resumir los datos recabados.

La evaluación supone dos operaciones fundamentales: la de obtener datos objetivos (medición), y la interpretación de los datos obtenidos. "Medir es asignar números a propiedades o fenómenos a través de la comparación de estos con una unidad establecida." (Livas, 1988:11)

La evaluación, en cambio, excede a la medición ya que el elemento que las distingue es la interpretación de los resultados.

Más allá de la medición, "la evaluación supone la existencia de juicios de valor." (Georgia, 1975:20) Se mide la capacidad del sujeto y se interpretan los resultados a través del proceso de la evaluación.

La evaluación implica el desarrollo de distintas estrategias o procedimientos cuantitativos y cualitativos, dependiendo de la naturaleza de los fenómenos que se pretenden evaluar.

Se ha mencionado que la evaluación es parte fundamental en la toma de decisiones, en la educación las decisiones institucionales son las que se toman con referencia a los alumnos principalmente, sea en grupo o como individuo.

Estas decisiones tienen que ver con la necesidad de que se les oriente tanto en el proceso de enseñanza-aprendizaje como con los contenidos programáticos y, lo más importante, con la detección de alguna posible alteración que limite o interfiera en el aprendizaje escolar.

La evaluación es parte del proceso de reunir los datos necesarios para la toma de decisiones educacionales, y la administración de pruebas es parte fundamental de la misma.

## **.2 LA EVALUACION COMO MEDICION**

Tanto educadores como psicólogos están interesados en aquellas mediciones que van a servir de base a la evaluación educativa, y más allá, las que permitan una interpretación clara y confiable de las posibles dificultades que presenta un alumno en el aprendizaje.

Para evaluar un proceso o logro de un objetivo, es necesario medirlo de alguna manera, de modo que sea posible manejarlo cuantitativamente, lo que implica el desarrollo de instrumentos adecuados para medir aquello que se pretende evaluar. Si el instrumento arroja datos erróneos o se mide algo distinto del objeto de evaluación, el proceso resultará ineficiente o equivocado.

El resultado de una medición no tiene en sí mismo significado, es preciso ubicarlo correctamente a un patrón, norma o parámetro.

La medición es una parte necesaria de la evaluación. Cuando se habla de evaluación y medición de un sujeto, se suele pensar principalmente en pruebas, y para que éstos instrumentos sean confiables, se requiere que sean estandarizados, es decir que estén bien definidos y de acuerdo a la población que se está evaluando. (Nunnally, 1987:15)

Los instrumentos de medición psicológica proporcionan la comprensión de las habilidades del niño y dan sugerencias para su mejor desarrollo, ayudan a valorar la motivación y las capacidades para establecer programas educativos y terapéuticos apropiados.

Sin embargo, no deben concebirse como una solución final a sus dificultades, sino ser un punto de partida para su atención.

La evaluación psicológica según Sattler (1988:299) se debe basar en la información obtenida de :

- Pruebas Objetivas
- Material de entrevista
- Observaciones específicas
- Observaciones controladas
- Rendimiento académico
- Información de otras personas
- Evaluación médica

En particular las pruebas psicométricas se pueden considerar como "un conjunto de ítems o preguntas destinadas a producir ciertos tipos de conductas que al ser presentadas en condiciones estandarizadas, producen puntuaciones que tienen propiedades psicométricas deseables." (Salvia, 1987:3)-

La evaluación psicológica a través de pruebas, permite estudiar no solo los patrones de comportamiento, sino también los índices cualitativos que revelan dificultades en la conciencia y control del desarrollo cognoscitivo.

Es necesario insistir una vez más que las pruebas evalúan las capacidades de un sujeto pero no son el único recurso de que se debe disponer para emitir un juicio, y ningún procedimiento por separado puede ser considerado como un criterio definitivo para determinar lo que un estudiante debe hacer.

El puntaje obtenido en una prueba, no debe ser base fundamental para la evaluación de un alumno, pero si posibilita el explorar tanto las habilidades como las deficiencias en aspectos cualitativos y cuantitativos. (Sattler, 1988:3)

Otro aspecto importante a considerar en la evaluación es la **formación y habilidad de quien aplica los instrumentos de medición**, pues de ello depende la interpretación de los resultados y por consiguiente el diagnóstico, independientemente de que el tipo de prueba que se seleccione, deberá utilizarse tal para lo cual fué diseñada.

### **.3 EVALUACION CON REFERENCIA A NORMA**

La evaluación educativa de los últimos tiempos ha sufrido una importante transformación en lo que se refiere a la forma para llevar a cabo la interpretación de los resultados de la medición psicológica.

La posición más conocida es la llamada Evaluación por Normas, que supone la clasificación de lo que un estudiante puede hacer respecto a otros estudiantes.

El término Norma, se refiere a una medida de comparación establecida al someter a un estudiante a otros estudiantes. (Anderson, et. al., 1977:157) Es decir, los resultados de la medición psicológica se interpretan al comparar el rendimiento de un alumno con el rendimiento de los demás.

Este tipo de evaluación proporciona poca información, pues no dice lo que un estudiante puede hacer por sí mismo, sino qué puede hacer en relación con otros. Se supone que las "aptitudes" son determinantes de un aprendizaje; las considera como un pronóstico del rendimiento escolar, sin tomar en cuenta otros elementos que afectan el proceso del aprendizaje.

En la Evaluación por Normas, la fuente de significado es la ejecución del grupo, lo que implica que los resultados no se refieren al conocimiento que tiene el alumno, sino que son generalizables a otras situaciones de evaluación con el mismo grupo. (Livas, 1973:21)

La mayoría de las pruebas psicológicas estandarizadas, están diseñadas por referencia a una Norma; ésta referencia tiene la particularidad de "clasificar" a los alumnos y determinar quien es más apto que otro.

## **1.4 EVALUACION CON REFERENCIA A CRITERIOS**

En la evaluación con referencia a Criterios, se pretende la eliminación de los inconvenientes de las pruebas normativas y con esto dar respuesta a la necesidad de mejorar el proceso educativo del alumno, sin necesidad de compararlo con los demás.

En educación, éste tipo de evaluación es reciente, y a consecuencia de **la propuesta de la tecnología educativa, se recomienda principalmente para la planeación de programas adecuados, así como para la evaluación de conductas específicas y situaciones dominadas por los estudiantes.**

Las preguntas relacionadas con las evaluaciones por Criterios están vinculadas con los objetivos instruccionales. Cuando se comenzó a generalizar la enseñanza a través de objetivos de aprendizaje, se propició la evaluación por Criterios, pues los objetivos abarcan tanto el contenido, como la conducta esperada por un alumno.

Las medidas psicológicas relacionadas con el Criterio, ponen de manifiesto la destreza con los contenidos y definen claramente las habilidades de un estudiante en relación con lo que él sabe y es capaz de hacer.

Implica el interpretar los resultados de las pruebas aplicadas, indicando que las habilidades son las mejor desarrolladas y que el conocimiento es el que ha llegado a dominar el individuo.

Éste tipo de evaluación no es muy común; sin embargo, en la actualidad los criterios de evaluación e interpretación se comienzan a inclinar por ésta alternativa.

***CAPITULO 2: INSTRUMENTOS DE  
MEDICION EN PSICOLOGIA Y  
EDUCACION***

En diferentes campos como: las Industrias, los Hospitales, el Ejército y especialmente el sector Educativo, la utilización de pruebas psicológicas desempeñan un papel importante como apoyo para la toma de decisiones.

Específicamente en el ámbito educativo son parte esencial en el proceso de la evaluación. En el presente capítulo se considera importante mencionar, en primera instancia, lo más relevante en el desarrollo histórico de las pruebas psicológicas, tanto en los países como en México, para después abordar la definición y clasificación que dan origen de los investigadores.

## **2.1 HISTORIA DE LAS PRUEBAS PSICOLOGICAS**

El uso de las pruebas psicológicas se remonta a la antigüedad en Grecia, en donde los griegos constituían un complemento al sistema educativo, estimando en el individuo el dominio de sus habilidades tanto físicas como intelectuales.

Desde entonces, el uso de las pruebas ha recorrido una larga trayectoria. Los métodos tradicionales de evaluación eran únicamente la historia del caso y la técnica de observación.

En la actualidad, la inclusión de las pruebas en la evaluación de un estudiante ha permitido aportar elementos importantes de diagnóstico.

Hacia 1830, se comenzaron a realizar diversos estudios con fundamentos psicológicos sistemáticos en un intento por medir la capacidad mental.

A partir del nacimiento de la psicología diferencial, se comenzaron a desarrollar las primeras pruebas mentales. Uno de los problemas que estimularon el desarrollo de las pruebas fué el identificar a los sujetos deficientes mentales para lo cual se elaboraron pruebas en el campo de la psicología clínica.

De los primeros médicos en trabajar la rehabilitación de los débiles mentales fué Binet, fundando en 1837 la primera escuela dedicada a la educación de los niños mentalmente deficientes. (Serrano, 1984:16)

Sin duda su mayor contribución fué la de diseñar, en 1848, técnicas para el entrenamiento sensorial y muscular; elaborando, entre otras, una Prueba de Ajustes de manos la cual lleva su nombre.

En 1838 el psiquiatra J. D. Esquirol, marcó ciertas pautas para clasificar a los débiles mentales, distinguiendo claramente entre incapacidad y enfermedad mental. En un intento por estructurar un sistema de clasificación de los diversos grados de retraso, llegó a la conclusión de que "el criterio más fiable del nivel intelectual del individuo lo constituye el uso que hace del lenguaje." (Anastasi, 1971:6)

Posiblemente esto fue el inicio de la serie de pruebas de inteligencia que en la actualidad se conocen como de contenido verbal.

En general la mayoría de los psicólogos experimentales del siglo XIX se abocaron a formular descripciones generales acerca de la conducta humana, restándole importancia a las diferencias individuales, o al menos no eran tan significativas.

Fue hacia 1859-69, con la contribución de la biología Darwinista, que se desarrollaron numerosas investigaciones. Sir Francis Galton, desempeñó una gran influencia en el desarrollo de las pruebas, mostrando interés sobre todo por los factores hereditarios y las diferencias individuales en el desempeño de un sujeto.

Galton desarrolló pruebas de discriminación sensorial y coordinación motora. Creó el primer laboratorio antropométrico con el fin de medir las características físicas y sensoriomotoras de los seres humanos. (Morales, 1990:14)

Su contribución al desarrollo de métodos estadísticos para el análisis de los datos acerca de las diferencias individuales, sentó las bases para que K. Pearson formulara en 1894 el coeficiente de correlación.

El psicólogo norteamericano J. Catell, fundó en 1890 su propio laboratorio, en donde hizo exhaustivamente el desarrollo de los instrumentos de medición psicológica, usando por primera vez el término "prueba mental". (Morales, 1990:15)

En dichas pruebas se incluían medidas para determinar la discriminación sensorial y motora utilizándose principalmente en sujetos con deficiencia mental.

A partir de entonces surgieron un gran número de investigaciones relacionadas con el funcionamiento intelectual, en un intento por aclarar con mayor precisión el funcionamiento a través de mediciones psicológicas.

Entre 1895 y 1905, los investigadores franceses A. Binet y V. Henry criticaron la mayoría de las series de pruebas hasta entonces existentes, considerándolas en gran medida sensoriales únicamente, por lo que su trabajo se orientó hacia el estudio de los procesos mentales superiores. (Sattler, 1988:30)

Binet y colaboradores dedicaron muchos años de investigación a la medición de la ligencia. Constantemente se llevaron a cabo revisiones de las escalas, tanto por sus res como por otros investigadores como Terman, Merrill, Yerkes, entre otros.

En 1904 el psicólogo inglés Spearman, desarrolló una teoría de dos factores de la ligencia, el factor G (general) y el factor S (específico), dando así las bases para lo se conoce como el Análisis Factorial. (Cerda, 1978:79)

Hacia 1905, Thorndike estudia la inteligencia animal y formula los Principios del endizaje, así como de la construcción de pruebas psicológicas.

Con la introducción de las Escalas Binet-Simon, florece el movimiento de las ebas, y se da a partir de la Psicometría y de la Psicología Diferencial, el desarrollo del odiagnóstico, siendo paralelo el nacimiento de la Psicología Clínica.

A principios del siglo XX, en varios países, el desarrollo de las pruebas de ligencia proliferó. Mientras que en Inglaterra los investigadores estaban interesados nálisis estadísticos, en Norteamérica se concentraron en las ideas de Bion.

En Alemania se estudiaba la psicopatología y las funciones más complejas, en Francia xperimentación clínica, surgiendo así la línea psiquiátrica con Charcot y Freud. tler, 1988:30)

Ya en 1912, los avances en el campo de la psicometría eran amplios. Las estigaciones continuaban, surgiendo así nuevos conceptos y criterios para la ización y evaluación de las pruebas psicológicas.

En ésta fecha W. Stern introdujo el concepto de Cociente Mental o Coeficiente lectual (C.I.)

Hacia 1915, comenzaron a aparecer las llamadas pruebas estandarizadas de ovechamiento para diversas materias como: aritmética, ortografía, lenguaje, caligrafía, más específicas del campo educativo.

Con éste avance se simplificó el trabajo en forma individual, ahorrándose tiempo en aplicación y evaluación. Sin embargo la aplicación de éste tipo de pruebas no estuvo relación con su revisión técnica olvidándose con frecuencia de su estado, aun, ipiente. (Anastasi, 1971:12)

En los años 30 continuaron desarrollándose numerosas pruebas con distintos etivos. En 1936 Piaget público Los Orígenes de la Inteligencia, lo cual sirvió como damento teórico para la interpretación de pruebas psicológicas.

. Bender, en 1938, diseñó la Prueba Guestáltica Visomotora basada en los ítems fundamentales de la Teoría de la Guestalt desarrollada por Wertheimer, Kohler y Koffka (1923)

Después A. Gesell ideó la Escala de Madurez, que lleva su nombre, apoyada en el esquema del desarrollo evolutivo normal.

En 1939 apareció la escala Wechsler-Belleuve, llevándose a cabo, por primera vez, la estandarización de una prueba para adultos en función de su edad cronológica. (Cerda, 1978:10)

Después se llevarían a cabo prolongaciones de ésta prueba, para adaptar la escala de adultos a niños, incluyendo nuevos reactivos.

Con la difusión del Psicoanálisis en Estados Unidos, se ejerció una gran influencia en las pruebas, por lo que se favorecieron algunas de ellas en mayor o menor grado al estarse en los postulados de la teoría Psicoanalítica. (Cerda, 1978:80)

Algunas de ellas fueron el Rorschach en 1921, el T.A.T. en 1942, la prueba de asociación de Rosezweg en 1944; las cuales son consideradas como pruebas de tipo proyectivo.

Desde los años 40 a la fecha, numerosas son las pruebas que se han desarrollado, algunas de las cuales han sido desechadas y otras continúan vigentes.

### **Desarrollo en México.**

El desarrollo de la psicometría en México no ha sido paralelo al desarrollo de la psicología y sus demás áreas. Son muy pocos los trabajos de investigación en relación con los realizados en otros países, y los que aquí existen no se les da difusión.

Con la fundación del primer laboratorio experimental, en 1916, se cimentaron las bases que permitieron el desarrollo de ésta disciplina en nuestro país.

Correspondió al Dr. Rafael Santamarina ser uno de los iniciadores de la psicología aplicada al realizar y coordinar proyectos de adaptación y estandarización de pruebas, en un momento en que entonces se conocía como Sección de Psicopedagogía e Higiene de la S.E.P. (Santamarina, et.al., 1994:49)

En 1921, Santamarina inició la traducción y adaptación de la recién diseñada Escala de Binet-Simon (1916) para aplicarla a los escolares mexicanos, presentando su primera validación provisional ante el Primer Congreso Mexicano del Niño.

Al concluir éste trabajo, el autor propuso el uso de la Escala "tanto para la formación de grupos homogéneos en las escuelas, como para la selección de los sospechosos de enfermedad mental" (Valderrama, et.al., 1994:42)

Los trabajos iniciados por Santamarina, se continuaron a partir de 1925. De ésta época se realizaron estandarizaciones de pruebas como la de Descudres (lenguaje); Binet-Simon (de complemento); Vermeulen (para medir el desarrollo mental); escala de Rey (lectura); prueba de Haggerty (comprensión de lectura), entre otras. Estas pruebas, se aplicaron a estudiantes de escuelas primarias y secundarias.

Al igual que en otros países, en México surgieron nuevas revisiones y adaptaciones de la Escala de Inteligencia Binet-Simon-Terman, como la que realizaron Bernal y Chertu en 1926 y la de Boder y Zuñiga, en el mismo año.

Algunos de estos trabajos, se aplicaron a los escolares, otros, se utilizaron en las escuelas de Educación Superior, como la de Agricultura, Normal de Maestros, de Policía y la Penitenciaria. En ésta última, por ejemplo, el objetivo de utilizar la Escala era complementar los estudios de los reclusos para otorgar su libertad.

Por otro lado, Montana Hastings, en 1929, realizó un trabajo de investigación con niños asistentes a escuelas secundarias. Las pruebas que se utilizaron fueron el Test de Binet-Simon y la Prueba Otis; el propósito de esto era "clasificar en grupos a todos los niños según su habilidad" (Valderrama, et. al., 1994:45)

Hacia 1936, se estableció el Departamento de Psicopedagogía e Higiene de la S.E.P., donde se llevaron a cabo diversas mediciones antropométricas y psicométricas.

En 1938, en el Estado de México se inició la aplicación de pruebas psicológicas en la escuela Normal de Profesores y, hacia 1940, se fundó el Instituto Nacional de Pedagogía, donde se implementaron numerosos trabajos de adaptación de pruebas aplicables a la realidad primaria, principalmente.

Entre ellas se mencionan: la Prueba Detroit Engels; Pintner Cuningham; Otis "A" y "B"; Prueba Nacional de Inteligencia y Prueba Terman Colectivo. (Serrano,1984:20)

En 1940 se revisó nuevamente la Escala Binet-Simon, ésta vez por los Dres. Ortega y Gasset. Su uso se limitó a los niños 5o. y 6o. de primaria y de 1o. de secundaria, esto con el propósito de conocer la habilidad mental de éstos alumnos.

En la década de 1950, continuaron los trabajos acerca de la psicometría en México, se aplicaron pruebas como la de Richard Meili y la estandarización de la Prueba Individual de Inteligencia de Kohs. (Valderrama, 1994:41)

En 1964, el Dr. Rogelio Díaz Guerrero fundó el Centro de Investigaciones en Ciencias del Comportamiento del Centro Electrónico de Cálculo de la UNAM. Anteriormente el Centro se estableció en la Facultad de Filosofía y Letras y, finalmente en 1973, se fundó el Instituto de Ciencias del Comportamiento y de la Actitud Pública, y Asociación Civil, (INCCAPAC). (Valderrama, et, al.,1994:111)

En éste Instituto los trabajos que principalmente se han implementado se relacionan con el desarrollo cognoscitivo y de la personalidad del mexicano, en comparación con el norteamericano, utilizando una serie de pruebas psicométricas.

Al crear éste tipo de investigaciones, se da inicio a numerosos trabajos que sirvieron de base para Tesis Profesionales; revisión de pruebas mentales y un acercamiento a las realidades socioculturales de nuestro país.

Diversas investigaciones se han llevado a cabo en éste Instituto, como revisiones y estandarizaciones de Instrumentos de Medición Psicológica, principalmente con un enfoque transcultural.

Hacia 1970, el mismo Dr. Díaz Guerrero funda el Centro de Investigaciones Psicológicas (CIPAC), en donde se llevan a cabo estudios relacionados con pruebas de desarrollo intelectual para niños 3 a 5 años de edad de la Cd. de México y regiones cercanas "...si bien estos materiales se han enseñado en las prácticas y clases de psicometría en la UNAM, su difusión entre los investigadores y profesionistas de la psicología en México, ha sido relativamente pequeña" (Valderrama, et, al., 1994:120)

Ante la necesidad de contar con un Instrumento que permitiera evaluar el rendimiento académico de los alumnos asistentes a escuelas primarias, se crea en 1974, en la Cd. de Monterrey una prueba que cubriera con estas expectativas.

La Prueba Monterrey, como se le llamó, fué dirigida por la Dra. Gómez Palacio y un grupo de colaboradores.

Hacia 1975 se desarrolla un proyecto titulado "Investigación sobre el Desarrollo de la Personalidad del Escolar Mexicano" (IDPEM), el cual pretendía obtener estandarizaciones y normas de calificación de diversos Instrumentos de Medición.

La muestra poblacional para ésta investigación, sirvió para diversos trabajos de Tesis, entre ellos la estandarización de la Escala de Inteligencia WISC-RM, en 1982, dirigido por la Dra. Gómez Palacio en la D.G.E.E.

En éste mismo año, en la Dirección mencionada, se llevaron a cabo diversos trabajos de revisión, adaptación y estandarización de pruebas psicométricas como El Sistema de

uación Multicultural y Pluralístico, SOMPA y, la Batería de Evaluación Intelectual  
ultánea y Secuencial, Kaufman, dirigidas también por la Dra. Gómez Palacio.

## 1.2 DEFINICION Y CLASIFICACION

Existen varios procedimientos para evaluar la conducta de una persona, entre ellos se encuentran las pruebas psicológicas que suelen ser útiles para éste fin.

Uno de los campos en donde las pruebas han encontrado su primera y más provechosa aplicación ha sido en el campo educativo, proporcionando a los maestros y psicólogos escolares la detección e identificación de los problemas de aprendizaje.

En ambientes educativos en donde se utilizan pruebas psicológicas, se espera poder adecuar a los alumnos de acuerdo a sus aptitudes, para implementar, así, diferentes métodos de instrucción.

La aplicación de pruebas en educación, puede aportar numerosas ventajas, Serrano (1984:15) considera que:

- El trabajo escolar se organiza científicamente.
- Se puede seleccionar a los alumnos para cursar carreras de educación superior.
- Seleccionar instrumentos para la orientación educativa y vocacional.
- Para conocer la personalidad de los alumnos.
- Para identificar quienes requieren de educación especial.

En educación especial comúnmente se especifica que el conjunto de datos acerca del funcionamiento intelectual de un sujeto, debe estar incluido en el proceso de toma de decisiones y provenir de una "evaluación intelectual".

### **Definición:**

La función básica de una prueba psicológica es medir las diferencias entre los individuos. La palabra PRUEBA es derivada del latín "TESTIS", que significa en inglés "TEST", siendo empleada ésta sin traducción en diversos países del mundo.

Para el desarrollo de éste trabajo, se utilizará el término "PRUEBA" que corresponde a la traducción del término en inglés, y por ser el apropiado a nuestra lengua.

Las pruebas se utilizan para seleccionar, diagnosticar, elaborar hipótesis, predecir y evaluar. "Una prueba es un instrumento de experimentación que permite observar cierto tipo de conductas y medirlas en forma objetiva." (Forns, 1980: 35)

Existen varios tipos de pruebas de acuerdo a diversas teorías o concepciones que se han desarrollado en la psicología y en función de los aspectos que se pretenden medir.

Al hablar de pruebas psicológicas, es común identificarlas con el término de "inteligencia". Numerosas definiciones se han propuesto en función de las teorías, describiendo así su origen y desarrollo.

Durante décadas se han venido utilizando varias denominaciones para describir lo que el sujeto es capaz de hacer, desde describir conceptos como capacidad, inteligencia, personalidad, aptitud, tratando de entender la conducta humana.

Es por esto que se han creado diversas pruebas, las cuales intentan medir diferentes aspectos en un sujeto dándonos una idea de su eficiencia intelectual.

Las pruebas son consideradas como métodos para la observación sistemática de la conducta de una persona "no se mide directamente a la persona sino que se infiere en sus características a partir de las respuestas dadas." (Brown, 1980:1)

Una prueba tiene por objeto "determinar ya sea el desarrollo mental de un individuo, o sea el nivel de su saber, la existencia y el grado en el de una aptitud dada." (Collin, 1987:2)

Hay pruebas que se diseñan para medir el desarrollo de las capacidades de un individuo en un área específica, otras miden el desarrollo de capacidades en diversos campos permitiendo identificar áreas fuertes y áreas débiles en un sujeto, así como determinar la naturaleza de éstas.

Una prueba puede ser considerada "como un conjunto de tareas o preguntas diseñadas a producir cierto tipo de conductas que al ser presentadas bajo condiciones estandarizadas producen puntuaciones que tienen propiedades psicométricas deseables." (Collin, 1987:3)

En el siguiente apartado de éste capítulo se especificaran los diversos tipos de pruebas existentes, seleccionando las que para éste trabajo son esenciales.

### **Clasificación:**

Existen diversos tipos de pruebas que pueden medir rasgos o características de un sujeto y, dependiendo de la información que se necesite, será el tipo de prueba que se utilizará.

En educación con frecuencia se necesita obtener información acerca del rendimiento de los niños, así como del desarrollo educativo, de los métodos de enseñanza y, en otros casos, esta información permite la ubicación de los alumnos que requieren de un apoyo especial.

Para obtener esta información, no es suficiente la aplicación de pruebas elaboradas por los maestros, que si bien es cierto son medidas de evaluación significativas, hay situaciones en que se hace necesaria la información que se obtiene a través de otro tipo de pruebas como las pruebas psicológicas.

Cuando a menudo una prueba psicológica se asemeja a una prueba escolar, la diferencia existe, es que la primera debe presentar un carácter psicométrico, es decir estar regida por reglas estrictas y uniformes que permitan la medida exacta del fenómeno estudiado. "... la tarea puede consistir en poner de manifiesto conocimientos adquiridos (prueba pedagógica) o bien, funciones sensorio-motoras o mentales (prueba psicológica)." (Cerde, 1978:75)

La selección de una prueba se determina por una parte, de acuerdo al propósito del examen y otra, por la eficiencia con la que ese propósito sea alcanzado.

Mehrens, (1982:40) considera para la elección de una prueba las siguientes características:

1. Disponibilidad de la información necesaria.
2. Facilidad de interpretación.
3. Administración y calificación.
4. Normas y escalas.
5. Validez.
6. Confiabilidad.
7. Calidad de los usuarios.
8. Costos.
9. Formatos.

Quien evalúa a un alumno, debe tener en cuenta tanto las limitaciones de la prueba, como estas difieren en términos de sus características técnicas, como en el propósito para el cual se va a valorar.

Para evaluar a un niño que presenta algún problema en el aprendizaje, generalmente se utilizan las pruebas estandarizadas.

**Una prueba estandarizada es aquella que es elaborada con métodos para seguir muestras de conducta uniformes, es decir, la serie de preguntas e ucciones de la prueba son fijos; cuenta con un tiempo limitado (algunas de ellas) la icación está definida, es uniforme y suele ser objetiva. Éste tipo de pruebas, eralmente están construidas por expertos en medición. (Mehrens, 1982:393)**

**Esto es lo que da a las pruebas estandarizadas el carácter de normativas. Para desheere (1978:25) una prueba de éste tipo debe responder a las siguientes encias:**

- 1. En la construcción de la prueba, la materia y la dificultad de las cuestiones deben ser controladas sistemáticamente.**
- 2. La estandarización debe cumplir con la aplicación y correlación lo más uniforme posible.**
- 3. La clasificación se realiza en función de las normas resultantes a la previa aplicación de un grupo de sujetos, con lo que se sitúa a cada una de las respuestas en una distribución estadística, a los que define como Contraste.**
- 4. La validez de la prueba se obtiene cuando las respuestas a las cuestiones planteadas dan una medida correcta del comportamiento solicitado.**
- 5. Si las consideraciones no cambian, la repetición del examen debe conducir siempre al mismo resultado o a otro muy próximo, a lo que se llama Confiabilidad o Constancia.**

**Las características para considerar a una prueba como estandarizada, se basan su estandarización, objetividad, confiabilidad y validez.**

**Estos aspectos se analizarán en el capítulo 3 del presente trabajo, sin embargo sideramos conveniente hacer mención en éstos, ya que nos estamos refiriendo a las ebas estandarizadas.**

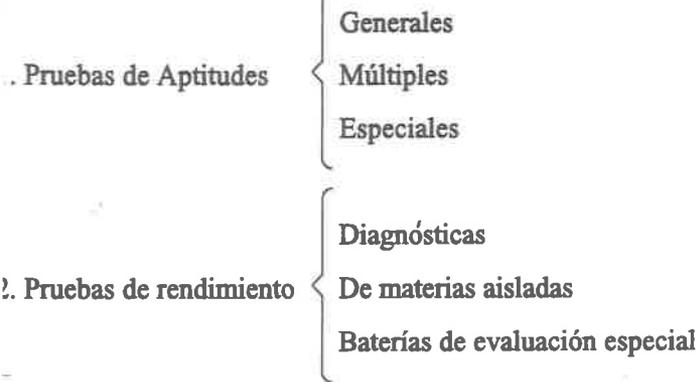
**Las pruebas pueden clasificarse de diversas maneras, cada autor lleva a cabo una ificación de acuerdo a lo que miden, por ejemplo, se pueden agrupar de acuerdo a sus edimientos de aplicación, como las pruebas individuales o las colectivas; o en base a instrucciones, como las pruebas verbales o las pruebas escritas.**

**Otras, pueden basarse en los propósitos de la prueba y el uso de los resultados, como pruebas basadas en normas o las basadas en criterios.**

**Con base a la complejidad de las conductas, personalidad y aptitudes estudiadas en un eto a través de las pruebas, así como la multiplicidad de éstas, es difícil establecer hoy**

lía, clasificaciones específicas. Es por esto que las divisiones generalmente se fican a causa de los enfoques con que cuentan los autores de éstas. Lo que no hay perder de vista es el seleccionar la prueba, aplicarla y evaluarla en los términos para al fué creada.

La clasificación que hace Mehrens(1982:394), la considera como una clasificación general, tomando en cuenta lo que las pruebas se proponen medir:



3. Inventarios de Intereses Personalidad y Actitudes

El autor de ésta clasificación considera a las dos primeras categorías como pruebas de medición máxima, es decir, que se espera que el resultado sea superior.

Continuando con Mehrens (1982:418), hay una diferencia entre las pruebas de aptitud y las de Rendimiento que tiene que ver con la cobertura y el rango de experiencia, estableciéndola de la siguiente manera: .

1. Los contenidos de las pruebas de aptitud general son más amplios.
2. Las pruebas de rendimiento tienen que ver con materias académicas.
3. Las pruebas de rendimiento miden un aprendizaje reciente, mientras que las aptitudes, un aprendizaje más remoto.
4. Al parecer, las pruebas de aptitud tienen índices de transmisión genética superiores.
5. Los propósitos difieren en cuanto a predecir futuras ejecuciones (aptitud), y medir el nivel presente de conocimientos (rendimiento).

La tercera categoría que se señala no está considerada como una medida de tipo cuantitativa, como es el caso de las dos anteriores. No se esperan respuestas falsas o correctas, lo que significa que su diseño es totalmente diferente. Sus aspectos son esencialmente cualitativos y no poseen cualidades psicométricas, denominándolas pruebas de ejecución característica o típica.

En este tipo de pruebas, el término "medida" es aún discutible ya que no hay una clara cuantificación de la inteligencia de las aptitudes, en donde lo que se trata de valorar es una cualidad.

Según Arias (1981:75) las pruebas se clasifican con base a: la forma de realizarlas, la forma de aplicarlas, el tiempo empleado y las características que miden:



Para elegir que tipo de prueba se debe utilizar, es necesario considerar el objetivo que se pretende alcanzar, así como las características técnicas de la prueba y la medida de las conductas muestreadas que representan.

Al revisar las clasificaciones que hacen algunos autores, coincidimos que para ésta investigación, las pruebas que se revisarán están incluidas en la selección que hace W. Wrensen en el rubro de: Pruebas de Aptitudes Psicológicas.

Las Pruebas de Aptitudes Psicológicas, son aquellas "que pretenden medir algunos aspectos que no son producto directo del aprendizaje..." pero tampoco se pueden considerar como aspectos innatos de esas aptitudes. (Forns, 1980:37)

Dentro de las pruebas de Aptitud se distinguen dos tipos:

a) Pruebas de Inteligencia.

b) Pruebas de Aptitudes específicas.

a) Las pruebas de Inteligencia o también llamadas como de Aptitudes cognoscitivas, son pruebas que "intentan poner de manifiesto la capacidad de una persona para realizar con éxito actividades que se desarrollan en el nivel simbólico." (Forns, 1980:37)

Esto significa, explorar en el sujeto la capacidad para representar mental, gráfica o simbólicamente una actividad dada.

Así, se cuenta con pruebas que exploran: madurez mental, habilidades, aptitudes mentales primarias, de factor "G", de aptitudes diferenciales, de razonamiento lógico,

b) Las pruebas de Aptitudes específicas, como las sensoriomotoras, aptitud espacial, mecánica, de memoria, atención perceptiva, etc... tienden a comprobar la presencia o no de estas aptitudes con objeto de determinar la realización de actividades.

En éste tipo de pruebas, generalmente se halla implícita la posibilidad de predecir con éxito o fracaso determinadas tareas que lleva a cabo el examinado.

Con objeto de entender a lo que se le llama Aptitud, se presentan una serie de definiciones realizadas por diversos investigadores:

**APTITUD:**

- "Capacidad innata y progresiva de adaptación adecuada y consciente a las situaciones de la vida." (Planchard, 1978)

"Un estado o conjunto de características consideradas como sintomáticas de una capacidad individual para adquirir por medio de la ejercitación, algún conocimiento..." (Szekely, 1966)

"...el potencial de un estudiante para aprovechar la enseñanza en áreas específicas." (Salvia, 1981)

"La capacidad potencial de un individuo en un tipo específico de actividad." (Mehrens, 1982)

Para Mehrens (1982:420), las Pruebas de Aptitudes se clasifican en cuatro categorías:

1. Pruebas individuales, que proporcionan una medida general de la inteligencia y permiten hacer una predicción general acerca de futuros éxitos académicos.
2. Pruebas de grupo, que proporcionan una medida general de la aptitud.
3. Pruebas que proporcionan medidas de Aptitudes múltiples, utilizadas principalmente cuando se desea hacer predicciones diferenciales.
4. Pruebas que representan medidas de algún tipo específico de Aptitud, y que son empleadas cuando se trata de predecir el éxito en una ocupación o curso especial.

En algunas instituciones educativas de otros países, se utilizan generalmente las pruebas colectivas, con el objeto de realizar una selección entre un grupo de alumnos, como por ejemplo, ubicarlos por grados o niveles, en México, casi no se utilizan pruebas grupales.

Algunas veces resulta más adecuado administrar pruebas individuales para contar con una descripción más detallada del alumno, así como en casos especiales, observar de cerca su comportamiento y reacciones ante los estímulos presentados, lo cual establecerá un diagnóstico mejor. De ésta manera se podrá definir con claridad la toma de decisiones, ayudando, así, tanto al alumno como al maestro.

La mayoría de las pruebas de inteligencia miden habilidades similares, algunas de ellas hacen una diferenciación entre los reactivos verbales y los no verbales.

Una prueba de ejecución es aquella que exige una manipulación de objetos más que una respuesta oral. Éste estilo de pruebas, son particularmente útiles para niños que

tantos problemas en el área de lenguaje, por lo tanto son recomendables en la aplicación de otras pruebas.

Sea cual fuere el tipo de prueba que se elija, intelectual, específica, individual o grupal, verbal o no verbal deben éstas cumplir con una función primordial, que es la de **medir y predecir las capacidades y habilidades** con las cuales cuenta un sujeto para el desempeño, y a partir de ahí proporcionar, en caso de ser necesario, la **instrucción adecuada**.

***CAPITULO 3: ASPECTOS  
PSICOMETRICOS DE LOS  
INSTRUMENTOS DE MEDICION***

En el capítulo 1, se señaló la importancia de la evaluación en el sistema escolar, mencionándose que, tanto la medición como la evaluación, son indispensables en la toma de decisiones educativas.

Este capítulo tiene por objeto, exponer algunos conceptos y procedimientos, utilizados en la construcción de pruebas psicológicas.

Lo que interesa es lo relacionado con la elaboración de pruebas estandarizadas. Partiendo de los principios fundamentales de la medición, indispensables para llevar a cabo el proceso de la evaluación.

Se ha mencionado que con frecuencia se necesita información relacionada con el aprendizaje de los alumnos y que dicha información se puede obtener a través de diferentes técnicas de evaluación.

Los maestros, generalmente recurren a las pruebas elaboradas por ellos mismos y a técnicas de observación, con objeto de medir o evaluar el aprovechamiento de sus alumnos, entre otras cosas.

En ocasiones, ésta información no es suficiente, sobre todo si se trata de detectar en el alumno una posible alteración que repercuta en su aprendizaje, por tanto, se hace necesario utilizar pruebas psicológicas estandarizadas.

La Estandarización es uno de los aspectos del proceso de construcción de una prueba y tiene su fundamento en el contenido, administración y calificación de ésta.

Si la construcción de una prueba cumple con estas propiedades y además dispone de Criterios de Ejecución, Criterios de Validez y Confiabilidad, se considera que cumple con los criterios de Estandarización.

La mayoría de las pruebas estandarizadas, son elaboradas por editores profesionales, quienes cuentan con expertos en medición y en planes de estudio.

**"El proceso mediante el cual se establecen las normas de una prueba psicológica conocido con el nombre de Estandarización" (Szekely,1966:18)**

### **3.1 PRINCIPIOS FUNDAMENTALES DE LA MEDICION**

Desde el punto de vista de la psicología, el estudio de las diferencias individuales ha sido un tema de continuo análisis.

Así como hay personas que difieren en el aspecto físico, también difieren en sus características de personalidad, en sus habilidades y en sus aptitudes.

Las preguntas que con mayor frecuencia se realizan, se refieren a la naturaleza y magnitud de esas diferencias, y es por eso que en el campo de la psicología diferencial, se buscado permanentemente el establecimiento de rasgos para describir a las personas, valiéndose así a sus características, las cuales varían mucho unas a otras.

El campo educativo también se enfrenta a situaciones tales como la búsqueda de un instrumento de medición que colabore en el proceso de la evaluación y que permita evaluar tanto al alumno, en su proceso de aprendizaje, como a los programas y contenidos de estudio.

Para lograr éste conocimiento se podrían utilizar diversos métodos o técnicas que permitan una mejor toma de decisiones. Lo que interesa para éste trabajo es conocer la medición que se obtiene a través de las pruebas psicológicas, consideradas por los psicólogos como un método adecuado para detectar oportunamente los problemas que afectan a los alumnos en el salón de clases, además de ser el motivo principal de ésta medición.

En términos generales, la medición supone la adopción del número para describir determinados aspectos de un fenómeno.

En el contexto de las pruebas, la medición es "asignar números a la conducta, de acuerdo a reglas preestablecidas." (Brown,1980:33) Por tanto, la medición de cualquier característica, implica la utilización de determinados procedimientos, de acuerdo a reglas específicas para obtener como resultado la asignación de valores numéricos a la ejecución de una persona.

Así, la estadística aplicada a la psicología, determina sus propiedades cuantitativas, y proporciona elementos indispensables para la interpretación de los resultados de la medición.

Al método de las pruebas, se han incorporado los principios básicos de la estadística aplicada, desarrollándose así nuevos recursos que se han adaptado a sus propias necesidades, parte de estos recursos constituye la llamada Psicometría.

Para Szekely ( 1966:49), la Psicometría comprende concepciones teóricas y técnicas relacionadas con:

- La construcción de pruebas.
- Los sistemas de unidades aplicables a las funciones psicológicas y a la conducta.

La evaluación de las pruebas psicológicas.

La validez.

La confiabilidad del instrumento.

Los fundamentos psicológicos de una prueba se ven reflejados en su contenido y lo que se propone medir, lo cual da a las pruebas el carácter de instrumento psicológico y, para que estas se conviertan en escalas de medida, se recurre a los métodos estadísticos.

Con objeto de entender la precisión de las mediciones psicológicas y educativas, es necesario conocer la naturaleza de la escala a utilizar.

Una escala es definida como "un sistema para asignar valores numéricos a ciertas características o rasgos mensurables." (Sattler, 1988:12 )

Existen varios métodos para ordenar datos, estos métodos o técnicas de medición son los que varios autores contemplan como escalas y comúnmente se utilizan cuatro. Haciendo la distinción que realiza Sattler (1988:13) se tiene:

**1. Escalas Nominales:** En ésta escala, a cada cosa que se está midiendo se le asigna un número a nombre distinto. Por lo general, las escalas de éste tipo representan categorías que se excluyen mutuamente, no hay orden en su organización. Ésta escala únicamente sirve para hacer clasificaciones.

**2. Escalas Ordinales:** En éstas, la variable bajo medición se ordena y jerarquiza, sea cual fuere la diferencia de magnitud entre los puntajes.

**3. Escalas de Intervalos:** En donde se obtiene una unidad específica de medición, de tal manera que la distancia entre cualesquier dos números adyacentes, es idéntica a la de cualesquier otros dos números.

La técnica tiene que especificar un procedimiento para determinar los intervalos idénticos, pero no ofrece ninguna operación a partir de "cero".

1. **Escalas de Razón:** En ésta técnica, existen intervalos idénticos, y además un punto "cero absoluto". Como la mayoría de las características psicológicas impiden medir un cero absoluto, estas escalas no son utilizadas.

Las escalas que con mayor frecuencia se emplean en psicología y educación son las de Razón y las de Intervalos, sobre todo ésta última, pues no solo interesa la calificación, sino la magnitud de las diferencias entre los individuos.

La Escala de Intervalos ofrece un método para transformar estadísticamente las calificaciones a una escala de unidades iguales: Calificaciones Estándar. Estas expresan la ejecución de una persona en términos de su desviación de la media en unidades de desviación estándar." (Brown, 1980:213)

Generalmente estas escalas son frecuentemente utilizadas en la medición y evaluación con referencia a normas.

## **.1.1 ERROR EN LA MEDICION**

Como todo tipo de medición se encuentra afectada por errores, la medición psicológica no está exenta de éstos.

Una de las principales metas de las pruebas psicológicas es que "las calificaciones obtenidas reflejen las calificaciones reales con tan poco error como sea posible." (Brown, 1980:66)

El sustento que apoya la introducción de los conceptos de Confiabilidad y Validez, se centra en los errores de medición. Para Brown (1980:56) existen dos tipos de errores:

**Errores al Azar:** Los cuales se dan cuando una variable afecta a la consistencia de la ejecución de la prueba. Por tanto, es indispensable la identificación y control de esos errores para considerar así la confiabilidad del instrumento.

**Errores Constantes:** Son los que producen efectos sistemáticos sobre la ejecución; pero son irrelevantes para los propósitos de la medición. Básicamente se deben a los propios instrumentos de la medición y son cruciales en la determinación de la validez.

Estos errores tienen tres causas principales:

- Los inherentes a la prueba. Como son el muestreo de los reactivos.
- Los asociados a las condiciones de la aplicación. Como la situación física, instrucciones, registros de tiempo, factores distractores.
- Los errores debido a fluctuaciones en las características del individuo. Como educación, maduración, cambios de ambiente, salud y atención.

Tomando en cuenta lo expuesto, se puede decir que una prueba psicológica "es un instrumento para apreciar objetiva y cuantitativamente aspectos del fenómeno humano en un determinado nivel de medición y con cierto margen de error, en las condiciones de una situación controlada." (Tavella, 1972:9)

Así, tenemos que todas las escalas de medición se ven afectadas por errores, sin embargo esto no impide o afecta su empleo siempre y cuando no excedan de cierta magnitud y sean identificados y controlados.

## **.1.2 CONFIABILIDAD**

Se mencionó que debido al error vinculado con la falta de confiabilidad de una prueba, crea expectativas en el resultado de la aplicación y puede ser no del todo objetivo, sobre todo cuando los errores no son identificados ni controlados.

La confiabilidad indica hasta que punto una prueba mide aquello para lo cual fue diseñada, para lo que se pretende medir, partiendo del supuesto de que al medir una característica hay que contar con los errores de medida.

Psicométricamente, "una escala es confiable cuando aplicada dos o más veces a un mismo sujeto, se obtiene el mismo resultado o diferencias aceptables." (Tavella, 1972:11)

La confiabilidad en una prueba depende de que el instrumento de medidas sea confiable, de la manera que se obtengan los mismos resultados al medir el rasgo nuevamente bajo condiciones similares del examinado. (Magnusson, 1969:77)

Esto es lo que le da a una prueba consistencia, sin ésta, los resultados de la aplicación varían diferentes en cada ocasión.

Para estimar el grado o coeficiente de confiabilidad de una prueba, se utilizan generalmente tres métodos: (Salvia, 1987:76-78 )

**1. Confiabilidad del Test-Retest:** Es un índice de estabilidad cuyo procedimiento es el aplicar al sujeto la prueba dos veces y, entre cada aplicación, considerar un intervalo de tiempo no muy largo, puesto que mientras más próximas se apliquen, mayor será la confiabilidad al existir mínimas probabilidades de cambio en períodos cortos.

Las puntuaciones obtenidas se correlacionan (coeficiente de correlación) obteniéndose así el Coeficiente de Confiabilidad.

**2. Confiabilidad de Formas Alternas o Equivalentes:** Se define como dos pruebas que miden el mismo rasgo en la misma amplitud y han sido estandarizadas con la misma población. Dichas formas alternas no contienen los mismos reactivos, se supone que las medias y las varianzas de las dos formas de prueba deben ser iguales.

Para calcular el Coeficiente de Confiabilidad se aplican las dos formas a la misma población y se correlacionan ambas puntuaciones.

**3. Confiabilidad de Consistencia Interna:** Éste procedimiento consiste en dividir la prueba en dos pruebas. Después de la aplicación, se correlacionan ambos grupos de puntuación y se obtiene una estimación de cada uno de ellos, a lo que se denomina

**Cálculo de Confiabilidad de mitades separadas.** La forma más común, es dividirla por pares, impares o por grado de dificultad.

Así como hay factores que determinan la consistencia de una prueba, también los hay que afectan su confiabilidad, como la longitud de prueba, el intervalo del Test-Retest, la amplitud del rango, las respuestas aleatorias y la variación dentro de la situación de prueba.

Una de las principales razones por las que se obtiene un coeficiente de confiabilidad para poder calcular la cantidad de error que generalmente se adhiere a la puntuación de un sujeto. (Salvia, 1987: 82)

### **3.1.3 VALIDEZ**

La característica más importante de una prueba sin lugar a dudas es su Validez. Si una prueba no cuenta con pruebas de validez, no se podrá conocer en realidad que es lo que mide, por tanto no será posible interpretar sus calificaciones.

**Quien construye pruebas debe proporcionar suficientes datos para saber que características mide la prueba y que criterios predice.** Si una prueba no es válida para propios fines, no podrá usarse con confianza y, como se utilizan para diferentes fines, ningún tipo de validez es único para todas las pruebas.

Tanto la Validez como la Consistencia y la Confiabilidad son términos genéricos que estrechamente relacionados, no podría hacerse referencia a uno, dejando de lado a los otros.

Así como se señaló que la confiabilidad se sustenta en las proporciones de las calificaciones libres de error, la validez se define como "la proporción de varianzas reales que es relevante para los fines del examen." (Sattler, 1988:23-24)

Es decir, la validez de una prueba se define por la extensión con que mide un rasgo, o la relación existente entre las calificaciones y medidas de criterio externo.

Desde el punto de vista psicológico, los resultados de una prueba tienen significado si se relacionan con otras variables significativas. Los puntos principales a observar en la validez son el establecer lo que mide la prueba y hasta que punto la prueba mide lo que se desea medir.

Existen tres clases principales para clasificar o construir una prueba.

- . Validez de Criterio.
- . Validez de Contenido.
- . Validez de Construcción.

Sattler (1988: 22-24) define estos criterios de la siguiente manera:

**La Validez de Criterio "se refiere a la relación entre los resultados de una prueba y algún tipo de criterio o resultado."** El criterio debe tener propiedades métricas de utilidad, se debe medir con facilidad y caracer de vicios y, lo más importante, servir para el fin de la prueba.

sin ésta relación complementaria no podría determinarse que la prueba mida el rasgo característica para lo cual fué estructurada.

Existen a su vez dos tipos de validez relacionada con el criterio: La Concurrente y la Predictiva. La primera ésta relacionada con el diagnóstico y se refiere cuando los resultados de la prueba están relacionados o no con alguna medida de criterio actual.

La segunda o de pronóstico se refiere a la correlación entre los puntajes de la prueba y el desempeño, conforme a un criterio importante.

**La Validez de Contenido, se refiere a la evaluación sistemática del contenido de la prueba con objeto de determinar si los reactivos seleccionados son o no representativos de lo que se pretende medir.**

**La Validez de Construcción** pretende medir el grado del rasgo psicológico y se logra a través de varios procedimientos diseñados para relacionar los reactivos con los constructos teóricos.

Sattler (1988:24) considera tres factores primordiales que pueden afectar la validez de prueba:

- . Factores relacionados con la prueba misma.
- . Factores relacionados con el criterio.
- . Intervención de sucesos inesperados.

### 3.1.4 NORMATIVIDAD

Las normas de las pruebas psicológicas constituyen marcos de referencia para la interpretación de las puntuaciones.

Cuando se interpreta la ejecución de una persona en una prueba, se debe conocer con qué parámetro se va a comparar, es decir, cuál es la norma de comparación. "Si se compara al sujeto con un grupo de referencia correcto, la muestra de individuos a la cual ha sido considerada la norma debe ser una representación justa de la población relevante" (Magnusson, 1969:291)

En los apartados anteriores se señaló que una prueba debe ser confiable y válida y a ser significativa, pero, si las normas son inadecuadas las puntuaciones serán incorrectas y por tanto, alteran el resultado de la aplicación.

En el ámbito educativo, las mediciones psicológicas son de mayor utilidad cuando las normas han sido obtenidas de UN GRUPO NORMATIVO REPRESENTATIVO, es decir, "la muestra normativa es importante porque es el grupo de individuos con el que se compara a la persona examinada." (Salvia, 1987:121)

La muestra normativa debe contemplar los mismos tipos de personas en la misma proporción que la población de referencia, y debe ser lo suficientemente numerosa para ser estable y proporcionar así, un rango completo de puntuaciones derivadas. "Las normas deben ser relevantes en términos de los propósitos del examen y ser utilizadas correctamente." (Salvia, 1987:123)

Para que esto se cumpla, es necesario considerar tres factores primordiales: representatividad, número de sujetos y relevancia de las normas. (Salvia, 1987:110)

La representatividad se refiere a las variables demográficas con lo que la prueba pretende medir y depender del contenido y de la estructura que ésta mide.

Se apoya en que "La muestra normativa contenga los mismos tipos de personas que la población que las normas pretenden representar." (Salvia, 1987:110)

Esto se refiere a la necesidad de tomar en cuenta las variables de: edad, grado escolar, sexo, aculturación, factores geográficos, inteligencia, fecha de las normas y características especiales de la población.

El número de sujetos debe ser lo suficientemente grande como para poder garantizar la estabilidad. Si la muestra normativa es mayor, más estables serán las normas. Si la

ba abarca un número de edades o grados, la muestra debe contener por lo menos 100 personas por edad y grado.

La relevancia de las normas, "conciérne al grado en que las personas de la muestra nativa proporcionarían comparaciones que sean relevantes en términos del propósito por el que se aplicó la prueba." (Salvia, 1987:119) Para tal efecto, existen normas nacionales, locales, basadas en grupos particulares y de población especial.

Cada situación de medición es única, por tanto no es posible decir si una prueba es buena o mala. Más bien se debe combinar el conocimiento de la situación, con el de los principios de la medición psicológica y hacer una evaluación razonada de la utilidad de la prueba para cada situación en particular.

"A menos que sepamos lo que mide una prueba y con que precisión lo hace, no podemos interpretar las calificaciones que obtengamos con ningún grado de confianza o exactitud." (Brown, 1980:212)

Brown clasifica las normas de ejecución de las pruebas en tres categorías:

En comparación con la ejecución de otras personas (grupo de norma o referencia).

En términos del grado de destreza en un contenido definido o dominio de habilidades (contenido de referencia).

En términos del nivel de rendimiento sobre un criterio externo (estas tienen la ventaja de combinar la validez con los datos normativos).

De acuerdo a éste autor las calificaciones relacionadas con el contenido son apropiadas para examinar contenidos académicos y que todas estas al menos, implican una base normativa. Por tanto, considera que "...las relacionadas con el contenido y las normas son en realidad dos modos de observar la misma ejecución; por ello se pueden utilizar en la misma prueba dos tipos de calificación." (1980:213).

Generalmente se espera que en los manuales de las pruebas psicológicas se especificuen los criterios de Validez, Confiabilidad y Normatividad de las mismas.

## ***CAPITULO 4: OBJETIVOS Y METODOLOGIA***

De acuerdo al marco teórico desarrollado en los capítulos anteriores, se señala que la acción es un proceso continuo, dinámico y sistemático inherente al proceso de enseñanza-aprendizaje, formando parte integral y fundamental de la acción educativa.

En las Instituciones dedicadas a la atención de los niños que presentan alteraciones en el proceso de enseñanza-aprendizaje, la evaluación se realiza a través de la utilización de bases psicológicas estandarizadas es decir, pruebas que estén bien definidas y de acuerdo a la población que se está evaluando.

Por tanto consideramos de utilidad llevar a cabo una revisión a través de un análisis de los aspectos psicométricos acerca de los instrumentos de medición psicológica que con mayor frecuencia son utilizadas para la detección y diagnóstico de los problemas que existen en el aprendizaje escolar de los niños mexicanos.

## **OBJETIVOS DE LA INVESTIGACION.**

1. OBTENER Y SISTEMATIZAR LA INFORMACION ACERCA DE LAS INVESTIGACIONES REALIZADAS POR DIVERSAS INSTITUCIONES ESPECIALIZADAS EN EL CAMPO DE LA PSICOMETRIA, SOBRE EL USO DE LAS PRUEBAS PSICOLOGICAS EN EL CAMPO EDUCATIVO.
2. DESCRIBIR EL MARCO TEORICO CONCEPTUAL DEL QUE PARTIERON LOS AUTORES PARA LA ELABORACION DE SUS PRUEBAS.
3. INDICAR LA PERTINENCIA DE ESTOS INSTRUMENTOS, PARA NUESTRO PAIS, EN TERMINOS DE VALIDEZ, CONFIABILIDAD Y NORMATIVIDAD.
4. ANALIZAR LA EFECTIVIDAD EN LA DETECCION Y DIAGNOSTICO DE LOS PROBLEMAS DE APRENDIZAJE EN MEXICO, A TRAVES DE LAS PRUEBAS PSICOMETRICAS QUE ACTUALMENTE SE UTILIZAN.
5. PROPORCIONAR A LOS PROFESIONISTAS INVOLUCRADOS EN EL PSICODIAGNOSTICO Y EN EL CAMPO DE LA EDUCACION, UN DOCUMENTO UTIL QUE LES PERMITA INFORMARSE Y REFLEXIONAR

## **SOBRE LOS INSTRUMENTOS DE EVALUACION QUE UTILIZAN PARA LA TOMA DE DECISIONES.**

### **METODOLOGIA.**

La presente investigación documental pretende ubicarse en un contexto que permita analizar los criterios seleccionados para la utilización de las Pruebas Psicológicas en el ámbito educativo.

Para su elaboración el objeto de estudio se sitúa en el Campo de la Psicología Educativa, cuestionando el modo en que son utilizadas las Pruebas Psicométricas en los centros dependientes de la Dirección General de Educación Preescolar y la Dirección General de Educación Especial.

La investigación será utilizada en forma selectiva, sistemática y eminentemente crítica, avocándose a lo esencial y significativo de acuerdo con los objetivos planteados.

Para recopilar la información presentarla y analizarla, se emplean las técnicas que conforman el sistema de "Sustentación documental" como son las fichas bibliográficas, de campo y la entrevista.

El ámbito de estudio seleccionado se limita a los Centros de Atención Educativa de Educación Preescolar y Centros de la Dirección General de Educación Especial, que atienden a niños del nivel preescolar y primaria con alteraciones del desarrollo y/o aprendizaje.

Las fuentes de información consideradas para éste trabajo son:

- El personal directivo de los Centros.
- Investigadores en el campo de la Psicometría.
- Trabajos de Tesis de diferentes Universidades del D.F. (U.P.N., U.N.A.M., IBERO, U.V.M., entre otras)
- Libros y Documentos dedicados a la Psicología Educativa, Psicometría, Evaluación y Medición.
- Manuales de las Pruebas Psicológicas utilizadas.
- Lineamientos Normativos de los Centros.

La realización del trabajo implica tres etapas:

**Etapas**

**Etapas 1**  
Recopilación, selección y clasificación de la información.

**Etapas 2**

Codificación y Sistematización del material recopilado.

**Etapas 3**

Análisis de la información.

***CAPITULO 5: PRUEBAS  
PSICOLOGICAS***

Dada la importancia que han adquirido las pruebas psicológicas como instrumentos de apoyo en la detección y diagnóstico de los problemas que interfieren en el aprendizaje se considera necesario realizar una revisión del origen y desarrollo de las Pruebas en México.

Es importante conocer las características fundamentales que ofrece cada instrumento, refiriéndose al sustento teórico en que se basan, así como los aspectos psicométricos (Validez, Confiabilidad y Normatividad) con los cuales la prueba fue Estandarizada, a fin de analizar su conveniencia de uso en la población infantil mexicana.

Las pruebas que para éste trabajo han sido seleccionadas son las que, según los Criterios Normativos establecidos por la S.E.P., ( 1984 ), forman parte de la batería de pruebas utilizada para llevar a cabo la evaluación psicológica de la población infantil mexicana, en la detección de los problemas que interfieren en el aprendizaje.

De acuerdo a la clasificación que hace Mehrens (1982:394) las pruebas a revisar se consideran:

#### **PRUEBAS DE APTITUDES:**

##### **1) Especiales:**

- Escala de Desarrollo de Arnold Gesell.
- Prueba Gestáltica Visomotora de Loretta Bender.
- Prueba Monterrey.

##### **2) Generales:**

- Escala de Inteligencia Stanford-Binet. (Terman-Merrill)
- Prueba de la Figura Humana de Florence Goodenough.
- Escalas Wechsler de Inteligencia para los Niveles Preescolar y Primaria.
- Batería de Evaluación Intelectual Secuencial Simultánea. Kaufman.

##### **3) Múltiples:**

- Batería S.O.M.P.A. Sistema de Evaluación Multicultural y Pluralístico.

# **.1 ESCALAS EVOLUTIVAS DEL DESARROLLO DE ARNOLD GESELL**

## **UNDAMENTACION TEORICA**

En un contexto de observación, medición y búsqueda de relación entre el desarrollo y maduración surge el trabajo de Arnold Gesell.

Parte de una posición dualista: desarrollo y maduración, términos sustentadores del concepto de crecimiento. Retoma de la biología las nociones de evolución y adaptación como soporte para el análisis del fenómeno del crecimiento humano.

Gesell propone un método de observación objetiva que permite hacer una descripción de los distintos momentos de desarrollo, para de ahí generar un método de medición para individuos particulares, llamado modelo de Evaluación "ESCALAS DE DESARROLLO GESELL"; que permiten ubicar el momento de desarrollo y maduración por el cual pasa un individuo en determinada edad, y saber si se encuentra "atrasado", "adelantado" o si es "normal" en relación con lo que se observa habitualmente en otros niños de la misma edad.

Aunque Gesell enfatiza que el modelo "no ha de usarse de manera inflexible ni con reglas fijas, pues cada niño tiene un modo de crecimiento único, es por tanto necesario usar con técnicas variables, elásticas y adaptables a la observación e interpretación del individuo" (Canton y Tadeo, 1990:28)

Cuando el niño es evaluado se encuentran conductas ya establecidas, "modos permanentes", al igual que se encuentran restos de su comportamiento, "modos transitorios" y, esbozos de conductas nuevas.

El sujeto puede ser mayor o menor que la edad asignada a cada nivel siendo lo importante describir el grado de madurez que de manera real ha alcanzado, porque no se trata de normas rígidas.

En 1946 fué cuando presentó por primera vez su diagnóstico del desarrollo al público en un congreso "anoparlante" (Gesell y Amatruda, 1979:13) aunque sus trabajos en éste campo comenzaron de manera sistemática en 1938.

Las Escalas de Desarrollo son herramientas de interpretación que no están concebidas para calificar al niño, sino... para observar las distintas facetas de su personalidad, las variaciones de sus reacciones, su comportamiento general y su adaptación a las exigencias ambientales de toda la situación creada.

Gesell plantea que en la primera época de vida existe una especie de paralelismo entre desarrollo mental y neurológico, por ello la mayoría de las conductas se relacionan con adquisiciones en el dominio de la motricidad y los sentidos. De ahí que sus mejores expresiones sean conductas psico-sensorio-motrices" (Gesell y Amatruda, 1979:21)

Las Escalas de Desarrollo incluyen aspectos sencillos y conocidos de la vida infantil, para determinar varias etapas de madurez en individuos con edades comprendidas entre un amplio rango de edad: desde las cuatro semanas, hasta los 16 años, aunque ocupa principalmente del período que va del crecimiento hasta los 3 ó 5 años" (Gesell y Amatruda, 1979:21)

Para la construcción de las Escalas de Desarrollo, Gesell parte de la correlación de la maduración y el desarrollo a partir de la edad cronológica en donde el término edad aparece como punto de confluencia entre lo que un individuo puede hacer, ya que tiene el tiempo biológico que se lo permite y, lo que sabe hacer. Esto es, como se ha llevado a cabo su adaptación al medio sea éste físico o social, así el término edad no es la base de medida, sino el momento de detención que permite valorar como es que la relación entre lo biológico y lo psicológico se está llevando a cabo.

Así, "la maduración biológica y las relaciones entre el individuo y su medio producen aparición, perfeccionamiento o desaparición de conductas o funciones. Se considera que los factores que influyen en el desarrollo son tanto externos (entorno), como internos (duración)." (Cantón y Tadeo, 1990:19)

Gesell parte del fraccionamiento de la conducta humana en cuatro campos diferentes: Desarrollo Motor, Conducta Adaptativa, Desarrollo del Lenguaje y Conducta Personal-Social, que "evolucionan con relativa independencia y al mismo tiempo conservan una unidad fundamental" (Gesell y Amatruda, 1979:21), ésta separación se hace únicamente como metodología para diseñar las Escalas de Desarrollo.

- 1.- Desarrollo Motor: ejecución de movimientos corporales, desde los más globales hasta los más finos, que implican actividad muscular; coordinación de procesos sensoriales y perceptuales, abarca lo que es el control del cuerpo en el área física.
- 2.- Conducta Adaptativa: coordinación visomotora, es decir coordinación de movimientos oculares y manuales.
- 3.- Desarrollo de Lenguaje: no solo es la expresión verbal, sino incluye toda la comunicación observable: gesticulaciones, balbuceos, expresiones motoras, palabras, oraciones, imitación de personas cercanas, expresión facial, gestos, etc.
- 4.- Conducta Personal-Social: implica la conjugación de las áreas anteriores y depende del medio en que se ha desarrollado el sujeto y de la influencia de los

factores intrínsecos al crecimiento. Describe la manera individual y característica de reaccionar ante las situaciones del sujeto.

Gesell considera al niño "como un sistema de acción en crecimiento a través del cual se desarrolla, que puede ser revelado por la forma en que se comporta" (Gesell y Amatruda, 1979:37)

Así el desarrollo es un proceso continuo que empieza desde la gestación y permite el estudio científico, la valoración diagnóstica del sujeto y la conducta como el indicador de desarrollo. Es decir la conducta crece y asume formas características a medida que se desarrolla, la "organización de la conducta... su dirección general va de la cabeza a los pies, de segmentos proximales a los distales y de céfalo caudal" (Gesell y Amatruda, 1979:37)

Las seis primeras etapas corresponden a un largo período de inmadurez relativa, durante el cual tiene lugar el crecimiento físico y mental. Sin embargo los años en que ocurren los cambios más dramáticos son los que corresponden a la infancia y a la adolescencia, años de preparación y perfeccionamiento, entre los cuales se intercalan períodos de equilibrio relativo y de desequilibrio transitorio" (Cantón y Tadeo, 1990:19)

## DESCRIPCION DE LA PRUEBA

Las Escalas de Desarrollo de Arnold Gesell se aplican de manera individual. Se han utilizado en el diagnóstico de trastornos neurológicos por médicos; en la evaluación del desarrollo por psicólogos; en la mejor conducción del niño por educadores; en el campo de la acción educativa y en la orientación a padres.

Cada examen o aplicación de la prueba lleva de 10 a 20 minutos y en éste tiempo se evalúan, ordenadamente, las diversas funciones de que el niño es capaz: control corporal, fuerza, presión, percepción visual, oído, coordinación visomanual, manipulación, locomoción, vocalización y reacciones sociales. El examen, en conjunto, está ideado para detectar comportamientos significativos para la estimación del grado de madurez del sistema nervioso.

Este diagnóstico evolutivo es la observación de formas o modos de conducta y su comparación mediante la comparación con "normas tipificadas"; describe pruebas estandarizadas para cada uno de los campos de conducta que explora, (motriz, lenguaje, intelectual y personal social).

El examen del desarrollo de la conducta a través de las escalas de desarrollo, lucra tres pasos:

**) La entrevista preeliminar:**

Se realiza para obtener información general en lo que respecta al niño, es una guía para el examinador porque incluye la descripción del desarrollo pasado del niño y de sus condiciones actuales.

Esta entrevista "debe ser corta y de orientación... buscando entablar una relación con el progenitor y el niño" (Gesell y Amatruda, 1979:118)

**) La aplicación de pruebas:**

Representa un procedimiento tipificado para observar y valorar el curso del desarrollo y la conducta de la vida diaria del niño.

El examinador explora el comportamiento del niño en un clima de libertad de acción, haciendo las preguntas hacia la zona próxima superior hasta que establece el límite máximo de las habilidades que esté explorando.

Primero se explora el área motora, luego el área de lenguaje, posteriormente el área de conducta adaptativa y, por último el área de la conducta personal social.

El material que se utiliza en ésta prueba es "aros de 3 colores, aro colgante, sonajero, conjunto de cubos, bolita, taza, campanilla, aro y cinta, taza y cubos, bolita y campanilla, tablero de madera, pelota pequeña, caja de pruebas con bloque y varillas, libro de genes, papel y lápiz, lamina de dibujos, pelota grande, formas coloreadas" (Gesell y Amatruda, 1979:48)

Con la aplicación de las Escalas Evolutivas se ha llegado según Gesell a adquirir conocimiento de la identidad en el proceso de desarrollo en la medida que determina los modos de reacción de niños "sanos" y "defectuosos", de los "dotados" y de los "afectados" por lesión ó por enfermedad en el que se pueden especificar niveles y grados de desarrollo en niños de seriación de madurez ya que las Escalas están vinculadas más al estado de madurez que a las habilidades o capacidades.

Para la calificación, se concede signo más (+) cuando el niño exhibe el correspondiente modo de conducta dependiendo de su edad o del esquema evolutivo fijado para explorar su desarrollo; se concede doble signo más (++) siempre que el niño exhibe un modo temporario debido a que en su lugar presenta un modo de conducta

maduro; y se concede signo menos (-) cuando el niño fracasa con exhibir un modo anente de conducta y además no exhibe un modo temporario debido a que en su r presenta un modo de conducta menos maduro. Aunque hay que tener en cuenta que hay respuestas buenas o malas toda respuesta es apropiada para algún nivel de rrollo, aun cuando desde el punto de vista adulto, la respuesta no sea adecuada al nulo": (Gesell y Amatruda, 1979:351)

Quando se hace el inventario de presencia y ausencia de modos de conducta, no se ura el comportamiento infantil, sino se juzga la estimación final es decir, no se an y promedian los signos; porque la significancia corresponde al conjunto total del nóstico, y éste se valora al determinar en que medida el comportamiento infantil sponde mejor a un nivel de edad que a otro.

El Coeficiente de Desarrollo (C.D.), es la relación existente entre la edad de madurez edad física actual, expresada en valor perceptual:

$$C.D.= \frac{\text{edad de madurez}}{\text{edad cronológica}} * 100$$

El "diagnóstico traduce los valores de la conducta en valores de edad. Los modos o as de conducta tipificados se usan para identificar y valorar la conducta que un cierto o exhibe" (Gesell y Amatruda, 1979:36), en términos de modos de conducta alizados.

Así, tanto la entrevista como la aplicación de las pruebas permiten describir el nivel adurez del niño y catalogarlo como "inteligente", "normal" o "anormal", de ahí que valuación no sea el resultado de un promedio de habilidades calificadas, sino la ración de como el niño se enfrenta al medio dependiendo de su edad.

### 3) Resumen de los resultados y revisión diagnóstica del examen:

La labor diagnóstica consiste en establecer distinciones entre los retardos, que son génitos y los que son adquiridos; entre los que son permanentes y los que son omáticos.

Gesell en el desarrollo de sus Escalas Evolutivas, no deja de manera clara los criterios validez, confiabilidad y normatividad utilizados en la construcción de éstas. Aunque a establecer y normalizar los rasgos promedio de desarrollo, estudio observando a tenares de niños "criados en hogares americanos de clase media, que representaban un po relativamente estrecho en cuanto al fondo cultural" (Cantón y Tadeo, 1990:22)

también utilizó el método de mellizos que consiste en dar un cierto entrenamiento isivo en alguna actividad a uno de ellos, manteniendo al otro en imposibilidad de itar la función investigada.

Uno de los trabajos más sistemáticos que realizó Gesell fué en 1927 en la clínica de la versidad de Yale. Consistió en un estudio de tipo longitudinal del curso normal del rollo de la conducta en el infante. Los datos se obtuvieron a través de repetidas rvacaciones de 107 infantes que constituían una muestra normal y homogénea, solo se yeron niños sanos sin ningún defecto conocido.

Se eligieron de forma que los padres fueran de estatus socioeconómico medio, todos niños eran americanos de nacimiento. A los infantes se les examinó a las 4, 6 y 8 anas de nacidos; estudios posteriores se hicieron a los 18 meses y a los 2.0, 3.0, 4.0, y 6.0 años.

Las repeticiones de los exámenes continuaron durante un período de 10 años de edad dos los casos en que fué posible. Solo se examinaron 28 de 60 niños hasta la edad de ños; a los 6 años solo se incluyeron 18 casos. En cada edad se observó ximadamente el mismo número de niños que de niñas.

El estudio de las muestras normativas se completó con extensas observaciones de s clínicos (niños con desviaciones y defectos en el desarrollo), que fueron estudiados n consultorio externo "en conexión con el trabajo de diagnóstico terapéutico y oramiento de una nursery de orientación para niños entre los 18 meses y 4 años de l. Los protocolos e historias clínicas incluyeron detallados informes de la conducta, o condiciones naturales clínicas y normativas". (Gesell y Amatruda, 1979:27)

Ésta obra de Arnold Gesell se construyó observando regular y sistemáticamente a un mo grupo de niños a través de los años (estudio longitudinal), y haciendo ervaciones ocasionales con niños anormales (estudio transversal).

Así a partir de tales observaciones, describe los fenómenos correspondientes que elacionados entre sí y con la maduración biológica, dan origen a determinados modos onducta.

## REVISIONES Y MODIFICACIONES

No se encontró información acerca de éste aspecto.

## **.2 PRUEBA GUESTALTICA VISOMOTORA. L. BENDER.**

### **UNDAMENTACION TEORICA**

La prueba giestáltica visomotora de Laurretta Bender está fundamentada principalmente en la Teoría de la Giestalt desarrollada por Wertheimer, Kohler y Kofka 1923.

Esta teoría permitió plantear una nueva comprensión para aprender las relaciones del todo y sus partes, expresando que cualquier estímulo será percibido como un todo cuyo todo cuyas propiedades no pueden derivarse por la simple adición de sus partes y relaciones. (Bender, 1982:17)

La función giestáltica se define como "aquella función del organismo integrado por la que éste responde a una constelación de estímulos dada por un todo, siendo la respuesta a una constelación, un patrón, una giestalt". (Bender, 1982:24)

El diseño de los estímulos que componen la prueba, está apegado a las leyes establecidas por la teoría de la Giestalt. La parte motriz, está representada en la ejecución de ejercer una actividad de copiado de cada uno de los modelos que se ofrecen.

Los fundamentos teóricos de ésta técnica proponen evaluar, a través de la copia, aspectos del desarrollo perceptomotor, así como las limitaciones, alteraciones y defectos que se pudieran presentar en la percepción visual vinculada a la ejecución motriz, los cuales con frecuencia se relacionan con causas de tipo orgánico y en otras, a problemas funcionales o a deficiencias leves o severas en el desarrollo.

Por ésta razón el Bender en una batería psicológica, es una de las pruebas que con mayor frecuencia se emplea. Su autora, la Dra. Laurretta Bender, la desarrolló en 1938 como un instrumento de evaluación clínica en poblaciones adultas y como una evaluación de desarrollo visomotor en niños.

### **DESCRIPCION DE LA PRUEBA**

Por su material corresponde a una prueba de lápiz y papel, por su forma de aplicación es una prueba individual.

Las instrucciones para su aplicación son verbales, se pide al sujeto que copie una serie de figuras (9) que se encuentran en tarjetas. Estas se presentan una a una, a la vez que el

o tiene una hoja y lápiz del no. 2 con goma. No se permite utilizar material de apoyo la ejecución de éstas figuras.

La copia de las figuras "constituye una prueba que permite establecer el nivel de maduración infantil de la función giestáltica visomotora." (Bender, 1982:156)

Mauretta Bender da normas para niños de 4 a 11 años de edad, sin embargo Sattler (1988:578) considera que éstas normas de grupo son limitadas pues la confiabilidad es no satisfactoria; la validez, opina, es satisfactoria en tanto se utilice como una medida de desarrollo percepto-motor.

Bender da una puntuación de 1 a 7 para cada figura en relación a las tablas de estandarización que ella misma establece, así obtiene una edad de maduración para cada niño y saca un promedio para la edad de maduración del sujeto.

Este sistema de calificación proporciona normas percentiles y puntaje estándar, las cuales las obtuvo Bender como resultado de sus investigaciones sobre la maduración visomotora en niños, a través de diferentes grupos.

La estandarización de la prueba se dió sobre los resultados de una población de 800 niños pertenecientes a distintos grados de una escuela suburbana de la Cd. de N. York. (Sattler, 1988:578)

## REVISIONES Y MODIFICACIONES

Una de las ventajas de la prueba Bender radica en el hecho de que puede ser interpretada de maneras distintas.

A medida que se fué empleando más la prueba, psicólogos e investigadores sintieron la necesidad de contar con un sistema objetivo de puntuación que fuera al mismo tiempo preciso y confiable.

El resultado ha sido una gran variedad de Sistemas de Puntuación y Esquemas de Clasificación, la mayoría de los cuales han sido diseñados y estandarizados en Estados Unidos y son para un grupo particular de sujetos solamente, como es el caso de la técnica propuesta por Pascal y Sutell en 1951 aplicable a personas de 15 a 50 años de edad.

Otro Sistema de Evaluación es el desarrollado por Santucci y Galifret Granjon (1952-1954) para niños de 6 a 10 años de edad, construida para detectar posibles discordancias entre nivel mental y nivel de organización espacial.

Watkins, desarrolló una investigación en 1980 en donde propone un Sistema de puntuación para determinar la existencia de un problema visual y de percepción en niños adolescentes.

Como éstas, a partir de la creación de la prueba por L. Bender se han venido desarrollando gran cantidad de investigaciones, sumándose a éstas la propuesta por la Elizabeth Koppitz.

Este Sistema ha sido considerado por los estudiosos en la materia como el más claro y objetivo para calificar la prueba, y sobre todo para evaluar a la población infantil. Fué diseñado en 1964, tomando en cuenta para su normatividad a niños en edad escolar entre 5 y 10 años de edad, asistentes a Jardín de Niños y hasta cuarto de Primaria.

Los sujetos de la muestra, provenían de escuelas públicas de los Estados Unidos, los cuales fueron recolectados de los protocolos de 1200 alumnos. (Koppitz, 1963:11)

El objetivo de éste estudio, fué el identificar los puntos clave para detectar alteraciones orgánicas, y establecer niveles de maduración propios de los niños, a través de los dibujos de la prueba en una edad determinada.

La misma autora publicó en 1975 un segundo Sistema de Puntuación para medir la maduración emocional, así como un nuevo conjunto de datos normativos utilizados para convertir puntuaciones de error a equivalente de edad y a rangos de percentil. (Salvia, 1975:316)

Con éstas investigaciones, Koppitz trató de proveer distintas formas de analizar los resultados del Bender en niños, de tal manera que se pudiera evaluar tanto el desarrollo intelectual, ajuste emocional y posible deterioro neurológico en base a un sólo protocolo. (Koppitz, 1975:20)

En la revisión de 1975, Koppitz estableció normas basada en una muestra de 1104 niños, entre 5 años y 11 años 11 meses, todos ellos residentes de áreas rurales, pueblos pequeños, suburbios y grandes centros del Oeste, Sur y Noreste de los Estados Unidos. (Koppitz, 1975:55)

La muestra en su mayoría fué integrada por población de raza blanca (86%) y el resto por diferentes etnias.

En relación a ésta muestra, Sattler opina que no es representativa de ese país, ya que la distribución geográfica no fué equilibrada, además de carecer de información socioeconómica de la misma. (1988:257)

La correlación media sugiere que aunque el puntaje global tiene confiabilidad, no es tan eficiente como para tomar decisiones de diagnóstico o de ubicación. (Sattler, 1988:260)

Sin embargo el rendimiento obtenido en la prueba puede ser utilizado como criterio diagnóstico diferencial en niños con Daño Cerebral, impedimentos perceptuales o problemas emocionales. (Salvia, 1981:320)

La aplicación se realizó en forma individual, de manera independiente, por tanto la confiabilidad es aceptable pues se basó en muestras de niños normales y con retraso mental.

La validez del Sistema de Calificación del Desarrollo propuesto por Koppitz, depende de cómo se use la prueba. Si se utiliza para evaluar el desarrollo percepto-motor, el sistema parece tener validez aceptable. (Sattler, 1988:253)

En la revisión de 1975, se reportan hallazgos sobre factores que suponen pueden afectar el rendimiento en la escala, concluyendo que la prueba sirve especialmente "como medida de la maduración de los niños al integrar funciones perceptuales y motoras". (Salvia, 1981:317)

La validez comparativa de éste Sistema, se ha establecido al correlacionarlo con pruebas como Frostig, Benton, V. M. I., entre otras. En relación a las pruebas de inteligencia la correlación ha sido variable.

El Bender es un instrumento útil en cualquier batería psicológica, sin embargo varios autores opinan que no debe usarse para emitir un diagnóstico definitivo, sino servir como instrumento que posibilite hipótesis sobre la habilidad percepto-motora del niño.

Para Salvia la validez de la prueba no está claramente establecida, pues no se ha demostrado que mida percepción visomotora. "Según los expertos, el copiar nueve formas o diseños es una medida de la percepción visual exclusivamente" (1981:317) por lo que se recomienda ser precavido en la interpretación.

Las principales contribuciones del trabajo realizado por la Dra. Koppitz, son los Criterios y Objetivos de Puntuación. Una vez que se realiza el análisis de los reactivos, se muestra el puntaje en la tabla 6 (anexo 1) con lo que se obtiene un nivel de maduración en percepción visomotora al comparar el puntaje obtenido por el niño, con el de otros de la misma edad cronológica. (Koppitz, 1980:231)

### **.3 PRUEBA MONTERREY**

#### **UNDAMENTACION TEORICA**

Quizá la teoría de mayor influencia en el desarrollo cognoscitivo del niño sea la propuesta por el psicólogo suizo Jean Piaget que a la par con sus colaboradores define la Teoría Psicogenética.

Este enfoque psicogenético se apoya en diversas ciencias como: la lógica, matemática, biología, física, sociología, fisiología; para abordar a la psicología.

Sobre la base de las nociones: estructura, génesis y equilibrio, Piaget desarrolló una teoría de la Inteligencia. Explica el desarrollo de la inteligencia tomando en cuenta el efecto de 4 factores: maduración del sistema nervioso, experiencia del sujeto sobre los objetos, factores sociales y equilibrio.

Piaget llegó a la conclusión de que el niño construye su mundo a partir de acciones y reacciones al relacionarse con objetos, acontecimientos y procesos que conforman su realidad. (Programa de Educación Preescolar, libro I S.E.P., 1981:12)

El desarrollo cognoscitivo como señala Piaget, se organiza en períodos o estadios, uno de los cuales posee a su vez diversos subestadios que cuentan con características propias y se originan en un nivel anterior, siendo el sustento o punto de partida para el siguiente nivel, dando así un progreso en el desarrollo.

En cada estadio se forman una serie de esquemas característicos que incluyen la percepción, el pensamiento y la conducta. Piaget distingue 4 períodos:

1. Período Sensomotor: Hasta los 2 años aproximadamente, se caracteriza por la aparición de los primeros conocimientos; el niño comienza a hacer uso de la imitación, la memoria y el pensamiento.
2. Período Preoperacional: De los 2 a los 7 años aproximadamente. En éste período comienza el lenguaje y las primeras funciones simbólicas, sin que exista aun el razonamiento lógico.
3. Período de las Operaciones Concretas: Va de los 8 años a los 11 años de edad. Aparecen los conceptos de número, tiempo, espacio y velocidad; se desarrollan operaciones intelectuales como: ordenar, disociar y combinar objetos concretos.

5. Periodo de las Operaciones Formales: De 11 a 15 años aproximadamente. Inicia con la aparición de nuevas estructuras lógicas, operaciones con conceptos y relaciones entre ellos.

Piaget también toma en cuenta el desarrollo afectivo-social en las relaciones niño-niño, ya que considera que proporcionan la base emocional que permite el desarrollo general.

## DESCRIPCION DE LA PRUEBA

La Prueba Monterrey es "un instrumento de evaluación psicológica de enfoque piagetiano que evalúa las construcciones de los niños con distintos materiales y la explicación verbal de los niños sobre sus construcciones." (Carrasco, 1983:19)

Se aplica a niños que asisten a los Grupos Integrados, entre los 6 y 9 años de edad, cursan el 1er y 2do grado de Primaria (repetidores) en escuelas regulares; y que enfrentan problemas de aprendizaje.

La Prueba Monterrey junto con el Cuestionario Psicopedagógico<sup>(1)</sup> y la Prueba de Habilidades Escolares<sup>(2)</sup> forman parte de los filtros de selección para ingresar a un grupo integrado.

Con los resultados en la Prueba Monterrey, se puede conocer el momento por el que atraviesa el niño con respecto a la lecto-escritura; ubicándolo en el estadio correspondiente, según la teoría psicogenética de Piaget.

Las personas facultadas para emplear e interpretar la Prueba, son aquellas que conocen los fundamentos del marco teórico de la Psicogenética.

El material de aplicación lo forman: un protocolo que se utiliza para la Prueba de Lecto-Escritura y para la Lógico-Matemática; 24 bloques (de diferente forma, tamaño y color); 10 reglas (de diferentes tamaños, variando 1cm de longitud); 38 fichas (en 2 colores)

Las áreas que se pueden explorar con la Prueba Monterrey (DGEE, Prueba Psicopedagógica de Detección de Problemas de Aprendizaje) son las siguientes:

Se utilizan para obtener datos acerca de las características de la población que se reprobó y se aplica con el fin de detectar que niños presentan dificultades significativas en la lecto-escritura y el cálculo

## **. Funciones Superiores.**

- a) **Visomotricidad:** Habilidad para reproducir estímulos visuales, por medio de la actividad motora.
- b) **Discriminación perceptual figura-fondo:** Habilidad para diferenciar una figura dentro de un campo perceptual.
- c) **Orientación espacial:** Percepción de un objeto, en relación a la posición del observador.
- d) **Memoria visual inmediata:** Habilidad para reproducir de memoria secuencias de estímulos visuales.
- e) **Memoria secuencial visual:** Habilidad para recordar y reproducir, en un orden determinado, una secuencia de estímulos visuales.
- f) **Memoria secuencial auditiva (articulación):** Habilidad para retener y reproducir de memoria secuencias de estímulos recibidos auditivamente.
- g) **Actividad perceptual:** Interpretación de lo que el niño ve, lo cual depende de su grado de maduración y experiencia.
- h) **Estructuras espacio-temporales:** Habilidad para reproducir impresiones no verbales, temporalmente organizadas.

## **. Lecto-Escritura:**

- a) **Análisis:** Proceso de separación de los elementos de una unidad; por medio de la síntesis se combinan e integran una estructura original.
- b) **Análisis y síntesis fonético:** Habilidad para analizar símbolos auditivos; así como el manejo del fonema y de la sílaba.
- c) **Escritura (Dictado):** Explora la percepción auditiva-fonética y su traducción a grafías (sílabas, palabras, etc.).
- d) **Lectura:** Explora la percepción viso-gráfica y su traducción motora fonemática y verbal.
- e) **Comprensión:** Evalúa el nivel de comprensión de una lectura, a partir de la relación de las respuestas con el contenido.

a Prueba Monterrey (Carrasco, 1983), consta de dos áreas de exploración, la noción elemental del número natural y la noción elemental de la lengua escrita; así como de un tipo de evaluación.

#### 1. Noción Elemental del Número Natural:

Evalúa el razonamiento lógico-matemático. Cuenta con 3 pruebas: Clasificación, Seriación y Conservación; cada una de las cuales divide a los niños en 2 momentos (operatorios y preoperatorios).

- a) Clasificación lógica: Es una forma de sintetizar información, lo cual permite al individuo organizar mentalmente el mundo que lo rodea. La clasificación cuenta con dos propiedades fundamentales: Comprensión - aspecto cualitativo de los objetos (forma, tamaño, etc.); y extensión - aspecto cuantitativo (cantidad de elementos que pertenecen a un criterio seleccionado).
- b) Seriación: Consiste en ordenar elementos de tamaños diferentes, pero iguales en forma; se pide que se forme una serie ordenando los materiales del más chico al más grande. La relación es asimétrica ya que, si se invierte el orden de la comparación, la relación que se establece ya no es la misma.
- c) Conservación: Los niños se dejan guiar por los aspectos perceptuales de cualquier material, al modificarse su distribución; el niño explica el cambio como algo definitivo.

#### 2. Noción Elemental de la Lengua-Escrita:

Cuenta con dos pruebas: Noción gramatical de la oración escrita, y Noción de la palabra escrita; las cuales tienen 3 niveles de conceptualización (concreto, presilábico-simbólico y lingüístico) y a su vez varias categorías.

- a) Nociones elementales de la lengua escrita: Explora las diferentes concepciones que tiene el niño acerca de la oración, lo importante no es el descifrado correcto, sino que el niño deduzca. Se trata de una adquisición gradual, basada en el razonamiento y los juicios que se va formando el niño del sistema de escritura.
- b) Noción de la palabra escrita: Pretende conocer el sistema conceptual que está en la base de la estructura del niño; es decir, si el niño ya concibe que una palabra puede ser escrita. Al igual que en la oración escrita el niño pasa por diferentes etapas: Concreta - las letras no le dicen nada, necesita de un dibujo; Simbólica - la escritura todavía no tiene significado; Lingüística - el

niño acepta la relación entre significado y significante, otorgándole valor a lo escrito.

#### Perfil de evaluación:

El perfil analítico fué introducido, para lograr una mayor aproximación del nivel cognoscitivo en que se encuentra el niño; permitiendo verificar los distintos niveles de la construcción del pensamiento del niño.

Conforme a los resultados de la Prueba Monterrey, los menores pueden caer dentro de 3 perfiles:

- a) Bajo. El niño es remitido a COEC y es canalizado.
- b) Alto. Es un niño de nivel operatorio que no necesita el servicio de G.I., puede continuar el proceso de lecto-escritura y matemáticas en un grupo regular.
- c) Perfil para G.I. El niño no obtiene ninguno de los 2 perfiles anteriores y es atendido en un G.I., con el fin de alcanzar una "nivelación" y así reintegrarse al sistema de primaria regular.

#### REVISIONES Y MODIFICACIONES

La prueba Monterrey fué creada en 1974 por la Dirección General de Educación Especial. A partir de su creación ha tenido varias revisiones, la última revisión de éste documento se realizó en 1981 la cual "cuenta con un cambio cualitativo, ya que incorpora la evaluación del desarrollo cognitivo de la LENGUA ESCRITA; reorganiza la prueba de LOGICA y LOGICO-MATEMATICA; y propone un PERFIL realmente válido para la medida de tales nociones." (Prueba Monterrey para Grupos Integrados, 1981 : febrero).

Sin embargo, en la bibliografía revisada, no se encontró información relacionada con las modificaciones que se hicieron a la prueba. Las referencias son escasas, y no se encuentran datos acerca de la metodología que se siguió para estas revisiones.

## 4 ESCALA DE INTELIGENCIA STANFORD-BINET (TERMAN-MERRILL)

### FUNDAMENTACION TEORICA

A partir de los trabajos realizados por Charles Darwin acerca de los factores hereditarios y su influencia en las diferencias individuales, se desarrollaron numerosas investigaciones relacionadas con la medición de las habilidades en los seres humanos. Ésto, se creó la necesidad de elaborar instrumentos de medición psicológica.

En 1905 con la construcción de la Escala Binet-Simon, se desarrollaron nuevas opciones que dieran cuenta del nivel de funcionamiento intelectual de los individuos.

Alfred Binet aportó un conjunto variado de habilidades e intereses al estudio del aprendizaje y la inteligencia, aplicando el método científico al estudio de las diferencias individuales.

La construcción de la Escala Binet-Simon fué considerada como la primer prueba de inteligencia en donde se intentaba demostrar algún interés por los datos normativos. (Miller 1988: 85,86)

Esta Escala fué una de las primeras pruebas mentales dirigidas a medir la inteligencia a través de los procesos mentales superiores, en lugar de las funciones sensoriales que a entonces se habían venido trabajando. (Sattler, 1988:30)

La teoría que subyace a ésta prueba, es la concepción que sus autores tienen acerca de la inteligencia. En colaboración con Theodore Simon, Binet elaboró un instrumento que, de acuerdo a la problemática del momento, pudiera medir la inteligencia en general, considerando que ésta "no se expresa en segmentos de conducta, sino más bien como una actividad mental combinada en la cual todo proceso funciona como un total unificado". (Sattler, 1978:75)

El definir la inteligencia no es tarea fácil, para Terman, la inteligencia no solo es la habilidad para aprobar las pruebas de una escala determinada, lo cual significa que no hay ninguna prueba que mida en una sola forma la habilidad para manejar todos los tipos de materiales en todos los niveles. (1921:131)

Así, cuando nos referimos a la aplicación de pruebas psicológicas para medir la inteligencia, es necesario revisar las teorías relacionadas con su naturaleza.

## **DESCRIPCION DE LA PRUEBA**

La Escala Stanford-Binet, es una prueba de aplicación individual. El tiempo de aplicación es de 45 min. a 1 hr. para los niños pequeños. Agrupa a los sujetos en niveles de

Los períodos de edad están divididos en grupos o bloques por cada 6 meses. Las pruebas dentro de cada nivel, son de dificultad uniforme, las tareas a realizar varían desde la manipulación sencilla hasta el razonamiento abstracto.

Los materiales que se necesitan para su aplicación y calificación incluyen una caja de juguetes comunes para niños pequeños, dos conjuntos de tarjetas impresas, un formato de registro y un manual de aplicación y calificación en donde se especifica la ejecución de cada prueba que acredita a cada reactivo.

La edad mental del sujeto se obtiene partiendo de la edad basal o año base, restando los meses correspondientes por cada reactivo que se apruebe más allá de su

Para obtener así la edad mental y considerando la edad cronológica en años y meses, se puede obtener el Coeficiente Intelectual (C.I.) utilizando las tablas normativas que figuran en el manual. Se presentan dos revisiones, la de 1960 (Stanford-Binet:309) y la de 1973 (Stanford-Binet:417), ambas revisiones elaboradas por Pinneau.

Una de las principales ventajas del Stanford-Binet es la cantidad de datos diagnósticos así como la extensa experiencia clínica que se ha acumulado alrededor de esta prueba. Sin embargo el uso de la clasificación de niveles de C.I. lleva a ciertos sesgos, como el no aplicarse rígidamente, ni excluir datos relativos al individuo como, experiencias de aprendizaje, ambiente socio-cultural, adaptación social, etc... (Anastassi,1971:207)

## **REVISIONES Y MODIFICACIONES**

La Escala original fué creada en París, Francia en 1905 por Alfred Binet y sus colaboradores. La prueba se construyó con el fin de incluir en sus niveles máximos el nivel General de la Inteligencia y limitar la influencia de factores de grupo o aptitudes heredadas.

Originalmente la prueba consistía en 30 subpruebas en orden creciente de dificultad. La estandarización fué hecha en niños de edad escolar considerados como "normales" y en una institución para niños "deficientes mentales".

En 1908 la Escala fué revisada por sus propios autores, reconociendo las acciones de la prueba en cuanto a la validez de las normas. En ésta revisión se paron los reactivos de acuerdo a la edad. La muestra fué más amplia, idarizándose por edades y empleándose por primera vez el concepto de Edad Mental I.). (Morales, 1978:79)

En 1911 la Escala Binet-Simon fué revisada nuevamente los cambios principales istieron en ampliar los reactivos a niveles superiores, introduciendo un nivel adulto. itieron algunos puntos de la revisión anterior, lo cual propició que proliferaran iones, traducciones, adaptaciones y estandarizaciones de la prueba.

En 1916 en la Universidad de Stanford, Louis Terman revisó y amplió la Escala para so en Estados Unidos. Con ésta revisión, Terman intentó proporcionar normas de portamiento intelectual en niños nacidos en Estados Unidos entre 3.0 y 16.0 años de , llamándola Revisión y Extensión Stanford de la Escala Binet-Simon de igencia. (Salvia, 1981:238)

La revisión de la Escala Binet-Simon por Terman, incorpora los rasgos que terizaban a las escalas de tipo Binet, como el uso de las normas de edad, los tipos de iones mentales correspondientes y el concepto de medida de una "inteligencia ral". Se utilizó por primera vez el concepto de C.I. (Coeficiente Intelectual).

Las pruebas empleadas eran variadas y heterogéneas, intentándose así probar la vidad funcional combinada, que concibe a la inteligencia como la suma de los esos del pensamiento que se hallan implicados en la adaptación mental." fford-Binet, 1975:2).

La Escala se tipificó en una muestra aproximada de 1 millar de niños y 400 adultos. elección de la muestra se hizo tratando de obtener una muestra representativa de la ación americana en general.

Aunque ésta revisión proporcionó una medida valida y confiable para un amplio po, se consideró inadecuada para probar las capacidades que se hallaban por debajo a edad mental de 4.0 años y niveles adultos. Se encontró tambien que algunas de las bas de la escala tenían una validez baja: las instrucciones de aplicación y valoración cían de precisión necesaria.

En 1937, después de 21 años la Escala fué nuevamente revisada, ésta vez por nan y Merrill. Se conservaron las características que distinguían a las pruebas tipo et, utilizándose los mismos supuestos, métodos y principios.

Se introdujeron reactivos nuevos, la muestra poblacional fué de 3,184 sujetos a los que se les aplicó las formas L y M. Se incluyeron nuevas pruebas para los niveles escolar y adulto, extendiéndose así los niveles de edad. (Sattler, 1988:90)

La contribución de ésta revisión fué una mayor objetividad en la valoración, mayor variedad de aplicación así como mayor variedad de contenido y forma.

Las evidencias de Confiabilidad y Validez se obtuvieron de las comparaciones que existían las valoraciones logradas por el grupo tipificado en la forma L, con las obtenidas en la forma M. Ambas formas se aplicaron con intervalo de una semana.

Dado que la Confiabilidad es función a la vez de la edad y la magnitud del C.I., se calcularon coeficientes de confiabilidad para cada escala por separado.

Los procedimientos de selección de los elementos empleados en la elaboración de la Escala de 1937 dan evidencia de su confiabilidad. Los distintos criterios fueron: edad tal como en el Stanford-Binet de 1916, edad cronológica y consistencia interna.

Otros elementos concernientes a la validez del Stanford-Binet, se basan en el análisis factorial de la Escala de 1937.

En 1960 apareció una nueva revisión (Terman-Merrill, 1960) Al respecto, Mehrens afirma que "la edición de 1960, no es en realidad una revisión, sino una combinación de mejores ítems incluidos en las dos formas anteriores de la edición de 1937" (2:423)

No se obtuvo para su estandarización ningún grupo nuevo, únicamente se usó una muestra de 4,498 sujetos a quienes se les había aplicado la Escala entre los años de 1950 y 1954, con objeto de constatar los cambios en la dificultad de los reactivos.

El grupo incluía representantes de todas las edades, desde los 2.6 hasta los 18 años; la distribución no fué proporcional entre las edades o grupos de edad mental.

Los elementos conservados fueron seleccionados porque continuaban satisfaciendo los requisitos de Validez y Confiabilidad de la revisión anterior.

Uno de los aspectos más importantes de la nueva forma, (L M) era la desaparición de las tablas de C.I. de la Escala anterior, apareciendo las tablas de Pinneau, "que presentan una menor desviación para las edades de 2.0 a los 18 años" (Sattler, 1988:91) El C.I. conceptualizado de ésta manera, es una forma básica de puntaje estándar normalizado con una media de 100 y una desviación estándar de 16.

Algunos autores opinan que la revisión de 1960 tiene puntos débiles, como el modo en que elaboraron los C.I.; los procedimientos para substituir a la reestandarización; la

mación escasa relacionada con las características demográficas y la saturación de ruido verbal.

Desde el punto de vista técnico, según Mehrens (1982 :423), la prueba es un instrumento bien elaborado. La mayoría de los coeficientes de confiabilidad son superiores a .90.

Los coeficientes de Validez predictiva y concurrente asociados a los criterios de calificaciones escalares y puntuaciones en las pruebas de rendimiento tienden a ser entre .75 y .85, por tanto, considera que la prueba es de gran utilidad como predictora de éxitos académicos.

Durante el período de 1971-72 se revisaron, en Estados Unidos, las normas de la Escala Stanford-Binet forma L M, publicándose hasta un año después. En ésta ocasión el estudio para la estandarización consistió en una muestra de 2,100 niños, aproximadamente para cada nivel de edad.

A diferencia de las normas de 1960, las de 1972 comprendían población negra y blancas con apellidos hispanos y blancos, excluyéndose de la muestra a los sujetos cuyo idioma principal no fuera el inglés.

Las normas de 1972 producen un C.I. de desviación media de 100 y una desviación estándar de 16. Sin embargo se revelan ciertas tendencias en los cambios de rendimiento intelectual, relacionados con la edad cronológica y nivel de habilidad.

Los cambios más significativos se dan en los niveles más jóvenes de la Escala. Cambios en la media de los puntos del C.I. de -9.7 a -10.8, en el rango de edad de 2.0 a 3.0 años). (Sattler, 1988:92)

Por tanto las normas de 1972 proporcionan un C.I. por lo general más bajo que las normas de 1960. Thorndike (1973) opina que los antecedentes culturales de los niños han cambiado con los años, y la exposición a la televisión pudo ser uno de los principales factores que causaron cambios en los preescolares. Por lo que, el ajuste hacia abajo de los puntajes tal vez indique que los niños de la muestra normativa más reciente son más brillantes que los anteriores.

## **.5 LA PRUEBA DEL DIBUJO DE LA FIGURA HUMANA. GOODENOUGH**

### **FUNDAMENTACION TEORICA**

Lorence Goodenough, creó un método para medir la inteligencia general a través del análisis objetivo de los elementos y detalles del dibujo de un hombre.

Parte de la siguiente hipótesis: "cuando el niño traza la figura humana, sobre el papel dibuja lo que ve sino lo que sabe a su respecto, y, por tanto no efectúa un trabajo intelectual, sino una expresión de su capacidad artística sino un repertorio conceptual". (F. Goodenough, 1971:13)

Goodenough afirmó que dicho repertorio conceptual, se incrementa con la edad tal, lo cual se refleja en el dibujo que hace un niño de 5 a 10 años.

Por lo tanto el análisis psicométrico de ese saber, manifestado en la representación de la figura humana en cada nivel de edad, proporciona un patrón para establecer un dibujo determinado, la edad mental.

La autora sostuvo que al medir ese saber, media Inteligencia puesto que el acto de dibujar implica funciones como: "asociación, observación analítica, discriminación, memoria de detalles, sentido espacial, juicio, abstracción, coordinación visomanual y estabilidad". (F. Goodenough, 1971:13)

Lo que significa que al dibujar la figura de un hombre, el niño lleva a cabo las siguientes actividades mentales: asocia los rasgos gráficos con el objeto real, analiza los componentes del objeto y lo representa, valora y selecciona los elementos característicos, toma en cuenta las relaciones espaciales (posición), formula juicios de relaciones cuantitativas (proporción), abstrae (reduce y simplifica) las partes del objeto en rasgos gráficos, coordina su actividad psicomotriz y adapta el esquema gráfico a su concepto de lo representado.

### **DESCRIPCION DE LA PRUEBA**

La prueba consiste en dibujar un hombre. F. Goodenough escogió la figura humana considerarla la más familiar para el niño.

Esta prueba se aplica a niños entre los 5 y 12 años de edad.

La Escala quedó constituida por 51 ítems que se ajustan a varios criterios, los catalogó orden de complejidad creciente. Luego computó el número de ítems, y calculó el ajuste para medir la madurez mental.

La prueba de dibujar un hombre se evalúa mediante 8 criterios:

- a) Cantidad de detalles.
- b) Proporcionalidad.
- c) Bidimensionalidad.
- d) Transparencias.
- e) Congruencia.
- f) Plasticidad.
- g) Coordinación visomotora.
- h) Perfil.

La consigna es: "Dibuja la figura de un hombre, realiza el mejor dibujo que puedas", se proporciona al sujeto una hoja blanca y un lápiz del No. 2, se puede aplicar en la grupal o individual.

Se evalúa registrando en su respectivo protocolo la presencia de detalles acertados, dividiendo el puntaje obtenido a edad mental y éste a Cociente Intelectual (C.I.).

La forma de computar los ítems según F. Goodenough, no reside en consideraciones técnicas, o en generalizaciones establecidas por estudios anteriores, sino en un análisis empírico de miles de niños entre los 2 y los 15 años de edad de diferente nacionalidad, sexo, grado y rendimiento escolar.

De ésta manera F. Goodenough procuró estandarizar su prueba sobre una muestra que realizara las diferencias culturales, de modo que en el examen, ellas no afectasen a la evaluación.

La Confiabilidad del Goodenough (el Dibujo de la Figura de un Hombre), con la prueba de Terman-Merrill oscila entre .80 y .90 para los niños entre 4 y 10 años; y de .76 la edad mental de la prueba de Stanford Binet con grupos de edades entre los 4 y los 10 años.

Su error probable en la estimación del Coeficiente Intelectual dado es de 5.4 para los niños de 5 a 10 años de edad". (Goodenough, 1971:15)

Goodenough empleó la prueba del Dibujo de la Figura de un Hombre, en el estudio del éxito escolar, encontrando una alta Correlación.

Encontró que las niñas evidencian superioridad cualitativa en sus dibujos en relación con los niños de la misma edad.

En la construcción de su prueba F. Goodenough reconoció la posibilidad del Dibujo de la Figura Humana para el análisis de la personalidad, sin embargo, su propósito al diseñar su sistema de calificación estuvo enfocado a la obtención del Coeficiente Intelectual actual.

## REVISIONES Y MODIFICACIONES

En 1926 en Estados Unidos, F. Goodenough, dió un giro al enfoque con que se había estado investigando e interpretando el Dibujo de la Figura Humana, lo ubicó como un instrumento de Medición Mental.

Goodenough (1957) fué la primera en realizar una evaluación cuantitativa de desarrollo infantil a partir del Dibujo de la Figura de un Hombre, contribuyendo así al desarrollo de la Psicología Evolutiva y de la Psicometría.

A fines de 1920 recogió aproximadamente 4000 dibujos de la Figura Humana de niños preescolares y de los primeros grados de las escuelas primarias públicas de Perth Amboy, Nueva Jersey; de ahí seleccionó un grupo de 100 dibujos en base a cada nivel de edad desde 4 años 0 meses a 9 años 11 meses, con ésta muestra comenzó su trabajo para la construcción de su escala.

En 1963, aparece la prueba de Harris-Goodenough que es la revisión de la prueba de Dibujo de un Hombre de F. Goodenough.

Harris consideró que la prueba construida por F. Goodenough como medida de inteligencia era utilizada en múltiples ocasiones, y se interesó en hacer una revaloración de los planteamientos de dicha prueba, siendo sus objetivos:

Ampliar el rango de edad incluyendo los años de la adolescencia aproximadamente hasta los 16 años.

Abarcar la exploración de nuevos items que pudieran aumentar la validez y confiabilidad mediante el muestreo de aspectos diferentes de la "aptitud cognitiva", creación de la Escala de Calidad.

Desarrollar formas alternativas de la Escala del Dibujo de un Hombre, introduciendo el Dibujo de una Mujer, y el Dibujo de Si Mismo (como base para uso proyectivo).

Diagnosticar Madurez Conceptual (nivel de abstracción), y no una Inteligencia unitaria.

La modificación practicada por el doctor Dale Harris (1963), al Dibujo de la Figura n Hombre, se dio a conocer en 1964. (Pallares, 1971)

La prueba se puede aplicar en forma grupal o individual a sujetos entre 3 y 16 años de edad, continua siendo una prueba de lápiz y papel.

El propósito es medir la Madurez Intelectual (Madurez Conceptual), que "consiste en la habilidad para formar conceptos de carácter creciente abstracto para percibir, abstraer y generalizar" (Salvia, 1980:290), en el reconocimiento de detalles.

La cual requiere de:

1) La capacidad de percibir o de discriminar entre similitudes y diferencias; 2) La capacidad de abstraer o de clasificar los objetos de acuerdo con esas similitudes y diferencias y 3) La capacidad de generalizar, es decir de asignar un objeto recién experimentado a una clase correcta, de acuerdo con características, propiedades o atributos diferenciados. Estas tres funciones tomadas en conjunto abarcan el proceso de formación de conceptos" (Harris, 1981:2), que implican la Madurez Conceptual.

La Escala de Goodenough-Harris consta de 71 ítems para la figura femenina y 73 ítems para la figura masculina que se valoran por su presencia o ausencia, siempre y cuando cubran el requisito a valorar.

Harris en su manual incluye las tablas de:

- 1) Dibujo del Hombre hecho por niños.
- 2) Dibujo del Hombre hecho por niñas.
- 3) Dibujo de la Mujer hecho por niños.
- 4) Dibujo de la Mujer hecho por niñas.

El Dibujo de Si Mismo no se califica de manera formal, es decir no hay puntajes con o grado de validez o confiabilidad, tiene un valor meramente proyectivo para medir la personalidad o afectividad.

Los dibujos de un hombre y una mujer pueden ser calificados mediante dos métodos o escalas. Escala de Calidad (Método cualitativo), y Escala de Puntos (Método cuantitativo).

En la "calificación cuantitativa, se otorgan puntos sobre la base de la cantidad de rasgos del dibujo... la calidad del dibujo se califica en una escala de 12 puntos". (Salvia, 1981:290)

**Escala de puntos:** fué construida ítem por ítem, las normas de la escala se proporcionan por separado para los hombres y las mujeres de 3 a 15 años de edad por intervalos de un año. Los puntajes crudos se convierten en estándar, con una media de 10 y una desviación estándar de 15". (Sattler, 1988:213)

La muestra para la estandarización de ésta escala constó de 2975 niños y niñas de varias áreas geográficas de E.U.A. seleccionados por ser representativos del censo de 1970, en cada nivel de edad de 5 a 15 años se incluyeron 75 niños.

**Escala de calidad:** ésta escala reporta un índice adecuado, válido y confiable para los niños que realizan los dibujos sobre la figura humana: Discrimina el desarrollo intelectual de una manera casi tan adecuada como las escalas de punto, según Harris ésta es útil como instrumento descriptivo, tiende a amplificar las diferencias sexuales observadas en las escalas de punto principalmente después de los 8 a los 9 años de edad.

Se basa en una escala de 12 reactivos con el 1 indica la categoría más baja y, con el 12 la más alta. Es un procedimiento de "calificación rápida que consiste en comparar el dibujo del niño con 12 modelos que representan un continuo de la menor excelencia a la máxima". (Sattler, 1988:213)

Fué construida y estandarizada (con dibujos del hombre y de la mujer), sobre la base de un grupo de jueces que clasificaron 240 dibujos (realizados por 20 niños en cada nivel de edad), en doce categorías.

Para establecer los criterios de validez y confiabilidad se "usaron tres grupos de niños, provenientes de niños de edades homogéneas (38 y 37 niñas entre 7 y 12 años de edad), y de niños de edades heterogéneas.

Para establecer la "validez de la escala de calidad, Harris reporta que tiene una correlación de .76 a .96 con la escala de punto". (Salvia, 1981:294) Por lo que se

Hay dos tipos de puntuaciones crudas de la Escala de Puntos y de la Escala de Edad que pueden ser convertidas en C.I.s de puntuaciones estándar.

Además los valores de las Escalas de Dibujo de un Hombre y de una Mujer pueden combinarse para obtener una estimación más confiable del rendimiento del sujeto en la prueba. Ésta combinación de valores se realiza "tomando en cuenta la mitad de la suma de los puntajes estándar equivalentes de C.I. que pueden obtenerse de los cuadros estadísticos... El promedio así obtenido resulta ser una estimación más exacta, desde el punto de vista estadístico, de la capacidad medida por el test del dibujo que aquella obtenida de cada dibujo solo". (Harris, 1981:120)

Los datos normativos de la Prueba de Goodenough-Harris se realizaron por medio de la administración de pruebas de dibujo a miles de niños de cuatro áreas geográficas de los Estados Unidos: área central: costa de Atlántica y Nueva Inglaterra; área sur: costa Occidental y Sur; área del Medio Oeste y del Medio Este.

Los datos normativos finales (estandarización) resultaron de la selección de 75 niños en cada nivel de edad de las diferentes zonas geográficas, se eligió un mismo número de niños y de niñas tomándose también en cuenta las ocupaciones de sus padres, según el censo de 1950.

Por lo tanto las normas y la estandarización de la prueba ésta basada en "2975 niños, representativas de la distribución ocupacional de los Estados Unidos en 1950, y están distribuidas en cuatro áreas geográficas". (Harris, 1980:113)

No obstante que la muestra "estandarizada fué excelente, las normas necesitan ajustarse con intervalos de medios y cuartos años para niños pequeños". (Sattler, 1980:214)

La prueba se validó con una muestra de 100 dibujos, se distribuyeron en cada nivel de edad un igual número de dibujos, además de tomar en cuenta la distribución de ocupaciones de los padres de acuerdo al censo de 1950 (de los padres de los niños que realizaron los dibujos).

Primeramente se hizo el análisis de 50 casos, y después sobre 100 casos (dibujos), cubriendo que se obtenían resultados mejores sobre 100 dibujos.

Para la selección de los ítems de la prueba se utilizaron los siguientes criterios: "1) Los ítems debían acusar un incremento regular y bastante rápido con la edad en el porcentaje de niños que aprueban el punto. 2) Los ítems debían mostrar una relación con una medida general de la inteligencia. 3) Los ítems deberían diferenciar entre los niños que obtienen un alto puntaje en la escala en conjunto y los niños que obtenían un puntaje bajo en conjunto." (Harris, 1980:87)

así a partir del porcentaje de niños que obtuvieron un punto en el ítem en cada edad azaron las curvas para ese ítem.

también se consideraron: 1) desde diferentes ángulos los efectos que determinadas ciencias o aprendizaje ejercen sobre el rendimiento en el dibujo, y 2) Relacionando lísticamente la escala con otras medidas de inteligencia, y con otras aptitudes afines ez a las características evaluadas por las escalas del dibujo.

Existen diferentes estudios de correlación para determinar la validez:

Correlación en el rendimiento obtenido en la prueba de Goodenough-Harris con puntuaciones obtenidas en otras medidas de inteligencia; las correlaciones fluctúan "entre .05 y .92 estando la mayoría en un rango entre .50 y .80". (Salvia, 1980:244)

Correlación entre el rendimiento de la prueba Goodenough-Harris y la escala original de Florencia Goodenough fluctúan entre .91 y .98.

Correlación entre el rendimiento de Goodenough-Harris y pruebas perceptomotoras oscilan entre .39 y .83, con una correlación media de .46. (Sattler, 1988:213)

En 1967 Elizabeth Koppitz retoma la prueba del Dibujo de un Hombre de F. denough (1926) y la evaluación de Dale Harris (1963), para la construcción de su la, en la que une el enfoque evolutivo y proyectivo para evaluar e interpretar el ajo de la Figura Humana.

Koppitz da nuevamente un cambio a la investigación e interpretación del Dibujo de la ra Humana, mientras F. Goodenough y Dale Harris la utilizan para medir la madurez tal de manera prioritaria; y Karen Machover (1949) como técnica que busca signos ecesidades inconscientes, conflictos y rasgos de personalidad; Koppitz la utiliza para ar el nivel madurativo y emocional de niño.

Su sistema de evaluación "surgió de la necesidad que experimento en su trabajo ico de contar con un método integrado y sistemático de interpretar los DFH". (ppitz, 1987:18)

El trabajo de E. Koppitz sigue la línea teórica de Harry Stack Sullivan, quien sostiene el Dibujo de la Figura Humana refleja primordialmente el nivel evolutivo de los os y sus relaciones interpersonales: actitudes hacia Si mismo y hacia las personas ificativas de su vida.

Koppitz afirma que el Dibujo de la Figura Humana "muestra tanto una estructura de contenido, como un cierto estilo que le es peculiar de ese niño". (Koppitz, 1987:19)

Es decir que la estructura del dibujo de un niño está determinada por su edad y su nivel de madurez mental, mientras que su estilo del dibujo refleja sus actitudes y capacidades que le son importantes en un momento dado.

Koppitz considera que la instrucción ambigua de "dibújame una persona entera" lleva al niño a mirar dentro de sí mismo y capturar la esencia de una persona y, como la persona que mejor conoce es EL MISMO, entonces su dibujo se convierte en un retrato de su ser interior y de sus actitudes ante la vida.

La prueba del Dibujo de la Figura Humana de E. Koppitz consta de 30 ítems evolutivos que permiten evaluar el nivel de madurez mental; y de 30 indicadores funcionales.

Se puede aplicar de manera individual y grupal, a niños entre los 5 y los 12 años de edad, el material que se utiliza es una hoja blanca y un lápiz del No. 2.

Koppitz sugiere que cada vez que se aplique ésta prueba se preste atención a la estructura en la cual dibuja el niño, la actitud, sus comentarios espontáneos, la cantidad de tiempo que emplea y la cantidad de papel que utiliza.

Este método traduce los puntajes obtenidos a categorías amplias del funcionamiento actual, es decir, éste método no es específico a un C.I. o rango percentil sino un nivel tal que según Koppitz, es suficiente para la clasificación de grupos o para detectar retraso mental.

Dado que los niños y las niñas maduran a ritmo diferente, Koppitz separa los datos de niños y niñas por edad, igual que Dale Harris.

Los estudios de Confiabilidad, los realizaron E. Koppitz y otra psicóloga, a través de un protocolo DFH de 10 alumnos de segundo grado elegidos al azar y de 15 niños remitidos a un psicólogo educativo por problemas de aprendizaje o conducta. Se verificó en cada uno de los 25 protocolos la presencia de 30 ítems evolutivos". (Koppitz, 1987:24)

Ambas computaron un total de 466 ítems diferentes para todos los dibujos. De éstos 433 fueron asignados por las dos psicólogas (95%), mientras que los 23 restantes (5%) fueron asignados solo por una o por otra. Encontrando confiabilidad porque el número promedio de ítems asignados a cada protocolo fué de 19.

La muestra para determinar la frecuencia de los 30 ítems evolutivos (estudio piloto), estuvo corroborada por 1856 niños y niñas entre los 5 y 12 años de edad, de escuelas públicas de Jardín de niños y Primarias.

Cada profesor aplicó la prueba del Dibujo de la Figura Humana de manera colectiva en las escuelas primarias y en el caso de Jardín de Niños la aplicó E. Koppitz de manera individual.

Koppitz dividió los puntajes de los ítems en cuatro categorías de frecuencia de aparición.

- a) Ítems esperados presentes en 86% a 100% de los dibujos de la figura humana, la omisión de cualquiera de estos ítems es señal de inmadurez mental.
- b) Ítems comunes aparecen en el 51% a 85% de los dibujos de la figura humana.
- c) Ítems bastante comunes aparecen en el 16% a 50% de los dibujos de la figura humana, la omisión de estos ítems al igual que los comunes no es importante desde el punto de vista diagnóstico.
- d) Ítems excepcionales presentes en el 15% o menos de los dibujos de la figura humana.

A través de diferentes estudios de validez, Koppitz "estableció que solo unos pocos ítems, primordialmente el cabello y las prendas de vestir, estaban relacionados tanto con la edad como con las experiencias sexuales y socioculturales del niño". (Koppitz, 1987:31)

Koppitz encontró una correlación significativa entre su prueba y las subescalas del WISC y Stanford Binet con niños de escuela primaria que tenían problemas emocionales y de aprendizaje, las correlaciones obtenidas fueron de .70 y .80.

Debido a las diferentes épocas y culturas Koppitz no atribuye demasiada importancia psicológica a la presencia u omisión de ítems de vestimenta o de el cabello en los dibujos de la figura humana. "Ciertamente, un dibujo infantil solo puede ser evaluado si el estilo de ropa y peinado son conocidos en su ambiente". (Koppitz, 1987:34)

## **.6 ESCALAS DE INTELIGENCIA PARA LOS NIVELES PREESCOLAR Y PRIMARIA. WECHSLER (WISC-WPSI)**

### **UNDAMENTACION TEORICA**

Las Escalas Wechsler son instrumentos de medición intelectual para: Preescolares (WPPSI) de 4 años 6 meses a 6 años de edad; Escolares (WISC) de 6 años a 16 años de edad; y Adultos (WAIS) de 15 años de edad en adelante. Para el desarrollo de la presente investigación, únicamente serán consideradas las escalas WPPSI y WISC, por ser utilizadas en los Centros de Atención.

Las Escalas Wechsler son baterías formadas "por varias subescalas; cada una de las cuales, si se considera aisladamente, puede proporcionar la medida de una actitud específica y si se combinan para formar una puntuación única nos da una medida global de la capacidad intelectual." (Manual WPPSI, 1981:1)

Las Escalas Wechsler, nacieron a partir de las críticas realizadas a la escala Stanford-Binet (se basa en la Inteligencia Verbal, y mide la Inteligencia Global); Wechsler apoyó en la teoría bifactorial de Spearman; en la que se considera que la inteligencia es la capacidad de cada individuo de establecer relaciones, desde las más simples hasta las más complejas. (Morales, 1975:88)

Partiendo de lo anterior Wechsler considera la capacidad intelectual como un todo, que puede ser analizada separadamente de sus partes; esto no significa que las habilidades sean independientes sino que se trata de la "capacidad global de un individuo para entender y enfrentarse al mundo que lo rodea." (Salvia, 1981:242) Con su medición evalúa la inteligencia.

Al realizar sus escalas, Wechsler intenta lograr una prueba con la cual se puedan medir los componentes de la inteligencia y las habilidades específicas, "para a través del análisis poder obtener una idea del funcionamiento general del sujeto." (Forns, 1970:71), sin olvidar que se deben tomar en cuenta motivaciones y tensiones impulsivas que influyen en el sujeto en todo momento de su vida.

Considera que la "palabra" y la inteligencia no tienen un desarrollo paralelo, siendo las pruebas de ejecución, pruebas de inteligencia concreta y de cualidades prácticas. Es por ello que agrupa sus escalas en verbal y de ejecución, y con las cuales, según Forns (1970:73) se pueden hacer las siguientes comparaciones:

- ) Inter-individual, ofrece un valor de orientación permitiendo situar a cada individuo en relación a los demás;
- ) Personal a edades diferentes, da un valor de fidelidad;
- ) Intra-sujetos, en un mismo individuo tiene un gran valor diagnóstico.

Wechsler, suprime la edad y evalúa en términos de dispersión, homogeneizando las categorías de las pruebas utilizadas. Introduce el concepto de C.I., considerándolo como potencia mental de un sujeto. Se trata de un consciente de eficiencia, que permita pasar no solo a un sujeto en relación a su edad cronológica, sino también en relación con otros sujetos de su propia edad." (Forns, 1980:72) Además realiza su clasificación de la inteligencia de acuerdo a que "Los datos cuantitativos se basan generalmente en medidas derivadas de diversas pruebas de inteligencia, y la clasificación dependerá del método de unidad psicometría que se utilice (Percentiles, C.I., calificaciones T, etc.) a partir de las cuales la significación podrá ser diferente." (Morales, 1975:103)

Wechsler "propone que para cada nivel de inteligencia se tenga un intervalo de clase que marque un rango de C.I. que se encuentre a una distancia medida a partir de la media expresará en términos de la desviación estándar." (Morales, 1975:105) Su esquema de clasificación es simétrico, comprende muchas clases e incluye porcentajes para cada categoría dada; ya que la estadística de inteligencia, obtenida sobre una muestra aleatoria, será válida solo para los grupos de la muestra. Por otro lado utiliza al C.I. como base fundamental de la inteligencia, ya que indica la relatividad de las medidas.

Las escalas Wechsler están constituidas por una serie de subpruebas, las cuales según Morales (1975:107) son: "por una parte escalas heterogéneas intertest", ya que miden diferentes funciones y factores en cada una de ellas; y son "escalas homogéneas intratest", ya que cada subprueba mide un solo factor en toda su amplitud."

## DESCRIPCION DE LA PRUEBA

En la estructura de la subprueba, la clasificación de los reactivos se divide en tres tipos:

- 1) Toda persona los puede contestar tenga o no instrucción escolar;
- 2) Deben de contestar las personas que han adquirido instrucción escolar media;

) Serán contestadas por sujetos que tengan cierta especialización o alguna instrucción especial. (Morales, 1975:107)

Las Escalas emplean un sistema diferencial de calificación para las subpruebas, ya cuentan con respuestas que se califican con aprobado-reprobado; otras reciben calificación de 0, 1 ó 2; con límite de tiempo, en donde se otorgan créditos extra a las soluciones más rápidas.

Las puntuaciones crudas obtenidas en las tres escalas Wechsler se convierten a puntuaciones en escala con una media de 10 y una desviación estándar de 3. Las puntuaciones en escala para las subpruebas verbales, las de ejecución y todas las subpruebas combinadas, son sumadas y luego convertidas para obtener C.I.s verbales, de ejecución y totales." (Salvia, 1981:244)

Las escalas cuentan con C.I.s de desviación con una media de 100 y una desviación estándar de 15; tanto en el WPPSI como en el WISC-R, las puntuaciones crudas pueden ser convertidas a edades de prueba.

Las Escalas Wechsler se administran por subpruebas, incluyen tareas Verbales y de ejecución proporcionando así C.I. para cada una de ellas, además de un C.I. Total. Son unidades de puntos y se aplican subpruebas idénticas a todos los sujetos.

La escala para el nivel escolar WISC, se encuentra formada por 12 subpruebas, 6 verbales y 6 de ejecución; y es formada por el siguiente material: Manual; protocolo; 23 tarjetas de figuras incompletas; 13 series de tarjetas para ordenación de dibujos; lámina con hilera de árboles; 3 tarjetas con los problemas de aritmética 16, 17, 18 y 19 cubos y impresiones con diseños; 5 rompecabezas; hoja de claves y laberintos.

El WPPSI, consta de once subpruebas, 6 verbales y 5 de ejecución; las cuales se reducen, para mantener el interés y la colaboración de los niños pequeños. Las subpruebas de: información, vocabulario, aritmética, semejanzas, comprensión, figuras incompletas, laberintos y diseño con prismas; fueron tomadas del WISC (se incluyeron ítems más fáciles y se eliminaron los de mayor dificultad; tanto en contenido como en forma de aplicación); frases, casa de animales y diseños geométricos son nuevas; se añaden a dígitos, ordenamiento de figuras, composición de objetos y claves del WISC.

El material por el cual está formada es el siguiente: Manual; Protocolo; hoja con laberintos; hoja con diseños geométricos; 1 tablero y 28 cilindros de colores para casa de animales, 1 libreta con los reactivos de figuras incompletas y los diseños prismas; 1 libreta con los problemas de aritmética y los diseños geométricos; 1 caja con 14 prismas.

as Escalas Wechsler cuentan con las siguientes subpruebas:

ubpruebas Verbales.

. Información. (WISC-WPPSI) - Mide la cantidad de conocimientos adquiridos por la experiencia cultural temprana; así como la habilidad para responder a preguntas específicas.

. Semejanzas. (WISC-WPPSI) - Mide la formación de conceptos verbales y la identificación de puntos en común.

. Aritmética. (WISC-WPPSI)- Mide la habilidad para el razonamiento numérico y la resolución de problemas (en el WPPSI, no se usan explícitamente números).

. Vocabulario. (WISC-WPPSI) - Mide el desarrollo del lenguaje, la habilidad de aprendizaje y la definición de palabras; así como la reserva de conocimientos.

. Comprensión (WISC-WPPSI) - Mide el juicio social y la habilidad para comprender instrucciones verbales, costumbres y normas específicas.

. Retención de Dígitos. (WISC)(WPPSI-Frases) - Mide la memoria a corto plazo, así como la atención.

ubpruebas de ejecución:

. Figuras incompletas. (WISC-WPPSI) - Mide la habilidad para distinguir detalles esenciales e identificar partes ausentes de figuras; requiere de concentración, organización y memoria visual.

. Ordenamiento de dibujos (WISC) - Mide la habilidad para el razonamiento no verbal y la planeación, así como la comprensión de una situación total.

. Diseño con cubos. (WISC)(WPPSI-Diseño con prismas, tiene dos caras para trabajar) - Mide la coordinación motora y la organización perceptual.

. Ensamble de objetos. (WISC)(WPPSI-Casa de animales) - Mide la habilidad perceptual.

. Claves. (WISC)(WPPSI-Diseños geométricos) - Mide coordinación visomotora, velocidad de operaciones mentales y memoria a corto plazo. Se relaciona con las habilidades o deficiencias más especializadas del niño.

Laberintos. (WISC-WPPSI) - Mide la habilidad para planear, y la organización perceptual. En el WPPSI, tienen forma horizontal.

El tiempo de aplicación para el WPPSI y para el WISC, es de aproximadamente una y media; además cada subprueba tiene sus propios criterios de discontinuación.

## REVISIONES Y MODIFICACIONES

La escala original, Wechsler-Belleuve de Inteligencia (1939), fué diseñada para medir la inteligencia en adultos, en 1955 se revisó y denominó Escala Wechsler de Inteligencia para Adultos (WAIS). En 1949, fué elaborada la Escala Wechsler de Inteligencia para el nivel escolar (WISC). En 1967 se desarrolló en cooperación con la división de pruebas de The Psychological Corporation; una extensión descendente del WISC, originando así la Escala Wechsler de Inteligencia para los Niveles Preescolar y Primario (WPPSI).

El WPPSI, es una escala para evaluar la inteligencia y capacidad del niño preescolar. El propósito es el diagnosticar el modo o grado en que influye en el funcionamiento general de los sujetos sus dotes y deficiencias. Es al mismo tiempo una prolongación de WISC y una escala independiente, para afrontar los problemas psicométricos que se presentan al examinar a niños.

El período entre los 4 y 6 años de edad, es crucial para el desarrollo intelectual del niño; ya que se enfrenta por primera vez a la educación formal, tiene contactos sociales más amplios, y se encuentra preparado para realizar tareas de una dificultad razonable, siempre que aprende y cuando se mantenga su interés y atención. El WPPSI se basa en el supuesto de que todas esas capacidades "pueden ser sistemáticamente evaluadas, con una batería de instrumentos psicológicos apropiados". (Manual-WPPSI, 1981:1)

El rápido crecimiento y maduración de sus distintas cualidades, rasgos y habilidades, como los procesos y niveles en los que tienen lugar, se denominan Desarrollo Mental. Existe una dificultad para diferenciar entre lo que es intelectual, neuronal, sensorio-motriz o referente a la maduración; pues aunque las funciones continúen mejorando con la edad, no significa que las habilidades se incrementen de igual modo.

El potencial intelectual del niño no se puede evaluar en sus primeros años del mismo modo como se hace más adelante; porque no se miden las mismas funciones, por eso "es esencial disponer de instrumentos no solamente adecuados, sino capaces de medir éstas

lidades de una forma continua." (Manual-WPPSI, 1981:2); las subescalas del WPPSI permiten comparar las ejecuciones de las mismas habilidades en una edad con otra, en edades posteriores.

Para realizar la distribución de la muestra de Estandarización del WPPSI, "se seleccionaron 45 centros de examinación, dentro de las comunidades de las diferentes zonas, con un total de 116 investigadores. Las aplicaciones se llevaron a cabo entre octubre de 1963 y mayo de 1966." (Manual-WPPSI, 1981:8)

Como norma se tomaron grupos representativos de niños estadounidenses de 4 a 6 1/2 años, de acuerdo con los reportes del censo de 1960, se estableció un cierto número de casos, así como del número de niños que representarían cada clase social.

Fueron 5 las variables que se utilizaron para la estratificación de la muestra y determinación del número de casos empleados:

1. Edad y sexo: 100 niños y niñas en cada grupo de edad (6 grupos en total), de los 4 a los 6 1/2 años, con un intervalo de 6 meses entre cada grupo, siendo en total 1,200 casos.

2. Región geográfica: Se tomaron las 4 regiones de E.U.A., la cantidad de casos fue proporcional a la población.

3. Residencia Urbana-Rural: Areas Urbanas - 2,500 habitantes o más; Areas Rurales - menos de 2,500 habitantes.

4. Color: Se incluyeron Blancos y No Blancos en proporción al censo de 1960, respetando región geográfica y tipo de residencia.

5. Ocupación del padre: Se dividió en ocho niveles, si no se obtenía información del padre se tomaba el nivel educativo de la madre. (Manual-WPPSI, 1981:7)

Se establecieron 5 niveles educativos, conforme a los años de estudio terminados: 8 años o menos, 9-11 años, 12 años, 13-15 años, y 16 años o más.

El coeficiente de confiabilidad para cada subescala se determinó mediante la técnica de división por mitades Split-Half (no es apropiada para escalas que requieren correlación); y con la fórmula de Spearman-Brown.

Además, "los datos para los Estados Unidos muestran buenos coeficientes de confiabilidad aunque no óptimos, lo que puede deberse (al igual que en México o en

quier otro país) a que el error de medición es debido, en parte, al desarrollo ológico y normal del niño." (Morales, 1975:120)

Para ver la estabilidad de sus puntuaciones, con respecto al tiempo, se realizó un estudio de reaplicación (test-retest) a 50 niños de Georgia; de alrededor de 5 años de edad, aplicándose desde 48 hasta 117 días después de la primera aplicación, con un intervalo promedio de 11 semanas.

En el WPPSI se conserva el C.I. como medida más efectiva para expresar el potencial intelectual de un niño en relación con otros de su misma edad; por tanto se trata de una prueba con validez a las Normas. La consistencia de los C.I., en varias aplicaciones, no significa que un niño deba tener un C.I. idéntico al volver a ser examinado. Pero en la interpretación de resultados, al obtener diferencias muy grandes (20 puntos o más) se debe buscar una explicación.

Para ver la relación del WPPSI y las escalas de inteligencia Stanford-Binet (forma L); El Peabody Picture Vocabulary Test, forma A (Dunn, 1959); El Pictorial Test of Intelligence (French, 1964); se realizó un estudio (1965-1966) donde se aplicaron las escalas a 98 niños entre los 4 y 5 años, en una escuela de San José California. Como resultado se obtuvo que el C.I. verbal del WPPSI correlaciona más con el C.I. del Stanford-Binet, pero no así el C.I. de la escala total. (Manual, 1981:12 y 21)

Battler (1988) opina que el WPPSI, posee excelentes propiedades psicométricas; proporciona información diagnóstica útil; además de que las normas de confiabilidad y validez son excelentes. Aunque existe dificultad al calificar algunas respuestas, debido a decisiones que dependen del examinador; además se presentan dificultades con niños marginados, por las posibles cargas emocionales que pueden afectar debido a sus condiciones de vida. Para Mehrens (1982), el WPPSI es difícil para niños de 4 años con poca capacidad.

Por otro lado La Escala de Inteligencia Para el Nivel Escolar (WISC) realizada en 1974, fué revisada y reestandarizada en 1974, en Estados Unidos; como una labor de cooperación entre el Autor y la Compañía Editorial, así como los miembros de Test Division of The Psychological Corporation, entre otros. (Manual-WISC-R, 4:Prefacio)

La revisión se realizó 25 años después de su aplicación original; duró poco más de 3 años y se inició por los comentarios críticos hechos hacia el WISC. Fueron preparados nuevos reactivos y materiales, se modificaron instrucciones y se ensayaron los cambios para desechar o sustituir lo que se creyó inoperante.

La escala en su totalidad permaneció estructural y contextualmente sin alteraciones, se conservaron el 72% de los reactivos; la nueva escala se denominó Escala de Inteligencia

el Nivel Escolar Revisada (WISC-R), se aplica a niños de 6 a 16 años de edad y en estandarización se incluyeron tanto niños blancos como de otras razas.

A pesar de que Wechsler no utilizaba "La edad mental, para calcular los C.I., incluyó, el manual del WISC-R, una tabla de equivalentes de edad de prueba para los puntajes brutos." (Sattler, 1988:140) Además por ser la escala una medida científica, práctica y válida; el C.I. se conservó como un aspecto esencial de la revisión.

La escala WISC-R, presenta una "confiabilidad, validez y estandarización excelentes; los procedimientos de estandarización; un manual óptimo y criterios de calificación claros." (Sattler, 1988:141) Sin embargo se considera cierta restricción en las normas de

## **.7 BATERIA DE EVALUACION INTELECTUAL SECUENCIAL SIMULTANEA KAUFMAN. MEXICO. D.F.**

### **UNDAMENTACION TEORICA**

El K-ABC (Batería de Evaluación Intellectual Secuencial Simultánea Kaufman) es un instrumento que permite evaluar la inteligencia de los niños cuyo rango de edad oscila entre los 2 años 6 meses y los 12 años de edad.

El Kaufman es una prueba referida a la norma que evalúa la inteligencia de los niños a través de la ejecución en tareas cognitivas, administradas y calificadas bajo condiciones conocidas.

Los autores de ésta Batería, Alan S. Kaufman y Nadeen L. Kaufman, consideran que la inteligencia es el resultado del estilo de cada individuo para procesar la información y resolver problemas que se le presentan, midiéndola a través de diversas tareas que se agrupan en: Tareas de procesamiento Mental Secuencial y Simultáneo y Tareas de Aprendizaje Escolar.

Los dos tipos de procesamiento, cuentan con el mismo peso, y las personas pueden funcionar tanto con un estilo como con el otro, dependiendo de la tarea a realizar.

La Teoría que subyace a ésta Batería es producto de tres conceptualizaciones fundamentales:

1. La aproximación neurológica de Luria y de los norteamericanos Das, Kirby y Jarman.
2. Los estudios efectuados por psicólogos cognoscitivistas y experimentales.
3. Y, los experimentos efectuados por especialistas en el cerebro.

El enfoque de Luria, es de tipo Sistémico, es decir, plantea que las funciones mentales están organizadas en sistemas que trabajan concertadamente, ejerciendo cada una de ellas un importante papel dentro de esos sistemas, "cualquier actividad consciente, constituye parte de un sistema funcional complejo, y se realiza con base al trabajo concertado de los bloques del cerebro, cada uno de los cuales contribuye a su realización". (Molina Castro, 1988:Ponencia)

Por otro lado, el enfoque que proponen Das, Kirby y Jarman, sugiere que las tareas no se clasifican de acuerdo con un nivel de complejidad, sino de acuerdo a la categoría de procesamiento de la información de cada individuo.

Para los psicólogos Cognoscitivistas, la aproximación de Luria es aceptable, partiendo del planteamiento de que hay dos tipos de procesamiento de información; como también coinciden con los especialistas del cerebro al asignar funciones a cada uno de los hemisferios.

Autores como Lezak, Klove, Noback y Demerest, especialistas en las funciones del cerebro, sostienen que hay diferencias en cuanto a la manera en que la información es procesada por cada Hemisferio.

Señalan que ambos hemisferios tienen diferentes estilos así, el hemisferio izquierdo tiene mayor habilidad para el Procesamiento Secuencial de la Información, mientras que el hemisferio derecho es más apto para el Procesamiento Simultáneo. (Molina Castro, 1988: Ponencia)

Para Kaufman, la localización de las estructuras cerebrales no son relevantes, lo importante en la Batería K-ABC son los procesos para obtener la información.

Así, Kaufman considera que el Procesamiento Secuencial es la manera que un problema se resuelve a través del arreglo de los estímulos en un orden sencillo o lineal, mientras que en el Procesamiento Simultáneo, la información tiene que ser organizada y sintetizada de manera total tomando en cuenta los estímulos simultáneamente. (Molina Castro, 1988: Ponencia)

## DESCRIPCION DE LA PRUEBA

La Batería de Evaluación Kaufman, es un instrumento que permite evaluar la inteligencia de los niños.

En su versión original, la prueba fue diseñada para ser aplicada a niños cuyo rango de edad es de 2.6 a 12.6 años de edad; consta de 191 reactivos para medir el Procesamiento Simultáneo y 156 reactivos de Rendimiento.

El Kaufman es una prueba de aplicación individual. Consta de tres atriles diseñados de modo que el examinador pueda sentarse frente o a un lado del niño y observar las instrucciones de administración y calificación.

Dentro del compartimento de los atriles, se encuentra material de aplicación. Por lo tanto se incluye una caja con el material necesario para aplicar otras subpruebas.

as subpruebas están ordenadas de acuerdo al orden de administración, en algunas ocasiones es necesario saltarse una o más subpruebas para encontrar la siguiente que ha de aplicarse según la edad del niño.

La Bateria consta de dos Manuales, uno de Administración y Clasificación, y otro de Interpretación. También cuenta con un protocolo de registro individual.

La Prueba ésta formada por 4 Escalas que a su vez contienen subescalas:

- . Escala de Procesamiento Secuencial.
- . Escala de Procesamiento Simultáneo.
- . Escala de Procesamiento Mental Compuesto.
- . Escala de Rendimiento.
- . Escala no Verbal.

Los autores de la Bateria, consideran que el K-ABC puede constituirse en un valioso instrumento para llevar a cabo un diagnóstico diferencial.

## REVISIONES Y MODIFICACIONES

Esta Bateria como ya se señaló, se originó en Estados Unidos a finales de 1982. La muestra que se utilizó para su estandarización estuvo compuesta por 580 niños cuyo rango de edad fué de 3.0 a 12.6 años. Hubo distribución racial equitativa de blancos, negros e hispanos. Los propósitos que llevaron a la creación de ésta Bateria fueron:

- Desarrollar una Bateria que midiera los dos tipos de procesamiento.
- Proporcionar una validación cruzada a través del análisis factorial.
- Estudiar los cambios del desarrollo en ambos procesos, considerados desde edades tempranas (preescolar) hasta nivel primaria.

Al término de éste trabajo, se llevó a cabo un segundo estudio, el cual tuvo como finalidad:

Proporcionar una validación cruzada independiente de los dos tipos de procesamiento, (secuencial y simultáneo) mediante el análisis factorial, a través de los resultados de la muestra de estandarización de la Batería.

Evaluar la Validez de Constructor.

Estudiar los cambios en el desarrollo de los dos procesos en el rango de los niños preescolares a los de educación elemental (primaria).

La muestra en éste caso fué de 2,000 niños entre 2.6 y 12.6 años de edad, estratificada por edad y sexo, nivel socioeconómico y región geográfica. Se incluyeron 72.5% de población blanca; 15.6% de población negra; 7.8% hispana y 4.1% de indios americanos, asiáticos u otras razas.

Desde la raíz del surgimiento de la Batería Kaufman se han generado numerosas investigaciones, principalmente en Estados Unidos, el propósito, en general, de estas ha sido la validación del propio instrumento.

## **8 SISTEMA DE EVALUACION MULTICULTURAL Y PLURALISTICO. S.O.M.P.A.**

### **UNDAMENTACION TEORICA**

a base fundamental de la que parte la creación del SOMPA es que: "la sociedad norteamericana es pluralística, cultural..." (Estandarización, DGEE:1982).

Esta concepción postula que existen diferentes grupos raciales y que cada grupo tiene una relación única con la cultura angloamericana.

Fischer (1978) considera que las diferencias lingüísticas y culturales deben ser tomadas en cuenta al interpretar una prueba, ya sea ésta de medición de habilidades, aptitudes, capacidades, etc... dado que todo comportamiento o conducta es aprendido en un determinado contexto sociocultural estructurado.

La Batería de pruebas SOMPA nace para responder a esta diversidad cultural y tomar en cuenta el medio por el cual atraviesa al niño antes de llegar al desarrollo en el que se encuentra cuando se le aplica la batería.

El niño es considerado como un ser biopsicosocial que se relaciona con su cuerpo y con los demás (medio ambiente). A partir de esto la batería parte de tres modelos para realizar el diagnóstico integral del sujeto.

#### **Modelo Médico:**

En donde el concepto de "normal" significa ausencia de síntomas patológicos (Sattler: 1982:241)

Este modelo tiene como base tres postulados:

1. Los signos patológicos tienen por causa una condición biológica en un niño.

2. Las características socioculturales del niño no son tan indispensables para un diagnóstico de normalidad biológica, precisa.

3. Las medidas que constituyen éste modelo evalúan las conductas o rasgos afectados por una disfunción biológica.

El propósito de éste modelo es evaluar el estado del organismo: anomalías biológicas, signos de enfermedad, impedimentos sensoriomotores, etc... "que se encuentren interfiriendo con las funciones fisiológicas del organismo" (Estandarización, 1982:13)

### **Modelo Social:**

El objetivo de éste modelo es determinar si la conducta del niño es normal en una situación determinada. Ésta se define "como la conducta que se conforma con las expectativas de los demás miembros de grupo, en tanto que lo anormal, es aquella que no cumple con las expectativas". (Estandarización, 1982:242)

Uno de los postulados de éste modelo es que amplía el concepto de normalidad incluyendo que existen muchas definiciones de "normal" por la existencia de diferentes normas sociales y sus diferentes papeles". (Estandarización, 1982:242)

El beneficio de ésta premisa según Sattler (1988), es que a un niño se le puede juzgar normal en un contexto social, y anormal en otro. Recalcando que la normalidad no es un atributo del niño, sino la definición dentro de determinado sistema social, con respecto a un contexto social específico.

### **Modelo Pluralístico:**

Considera al comportamiento adaptativo, como la habilidad del individuo para desempeñar roles progresivamente complejos, en un círculo de sistemas sociales.

Según Sattler (1988), éste modelo postula lo siguiente:

Todas las pruebas son específicas en una cultura, y miden el aprendizaje, comportamiento, actitudes, pero de esa cultura en especial.

Los niños deben compararse con otros de idénticos antecedentes, antes de sacar cualquier conclusión definitiva sobre su inteligencia.

La Batería de pruebas SOMPA es un Sistema de Evaluación Multicultural porque se refiere a la evaluación de niños precedentes de diversas culturas; y Pluralístico porque se refiere al uso de Normas múltiples en concordancia con la cultura.

## DESCRIPCIÓN DE LA PRUEBA

Es una Batería de medidas que "tienen por objeto incorporar información médica, individual y pluralística en la evaluación de la conducta cognoscitiva, percepto-motora y conductiva de niños negros, blancos e hispanos". (Sattler, 1988:241)

Se aplica a niños entre los 5.0 y 11.0 años de edad, permitiendo realizar una evaluación equilibrada para tomar decisiones educativas y de colocación de grado escolar, libres de toda discriminación racial y cultural.

La Batería de pruebas SOMPA, contiene 10 medidas dentro de sus tres modelos de evaluación.

**Modelo Médico:** Consta de seis medidas, que pretenden dar respuesta a si es el niño un individuo biológicamente normal, lo cual proporciona de ésta forma un panorama del estado de salud del niño, tanto pasado como presente.

**Escalas de Destreza Física:** constan de 29 tareas, miden la integridad y capacidades de las vías sensoriomotrices del niño.

Estas Escalas, a su vez están agrupadas en seis escalas:

- Ambulación.
- Equilibrio.
- Colocación.
- Secuencia motora fina.
- Destreza de dedos.
- Movimiento involuntario.

Los 6 puntajes crudos se convierten en escalares, basados en una media de 50 y una desviación estándar de 15.

Según Mercer (1978) las tareas de destreza física pueden interpretarse sin hacer referencia a los antecedentes socioculturales del niño.

**Prueba Gestáltica Visomotora de L. Bender:** Tiene como propósito "evaluar la madurez perceptual del niño y su posible daño neurológico." (Investigación SOMPA, DGEE, 1980:114)

**Medición de Agudeza Auditiva:** La debe realizar una persona capacitada en el ramo. En la medición se considera que el oído del niño es subnormal si existe un umbral de audición de 35 decibeles o más para cualquier frecuencia, o de 25 a 35.

a calificación de ésta escala consiste en registrar los niveles no corregidos de audición en decibeles.

**Medición de la Agudeza Auditiva:** Para ésta evaluación se administra la prueba de Snellen. Los puntajes obtenidos se registran en el Perfil del SOMPA, considerándose como subnormal la visión de 20/40 o menores.

**Relación Peso por Altura:** La altura se registra reduciéndola al centímetro más cercano, y el peso al kilogramo más cercano.

Los puntajes escalares correspondientes se proporcionan por separado para niños y niñas, y se basan en una medida de 50 o una desviación estándar de 15.

**Inventarios de Salud:** Constan de una serie de 45 preguntas estructuradas que se hacen a la persona que ésta a cargo del niño en el hogar.

Los Inventarios de Salud están formados de la siguiente manera: Inventario Prenatal y Posnatal, de Traumatismo, de Enfermedades, de Visión, de Audición.

Las preguntas tienen varios incisos, la calificación debe ser sumada para obtener el puntaje crudo total de cada inventario.

**Modelo Social:** El objetivo de éste modelo es valorar el papel del niño en relación diversos grupos sociales. Las medidas que se utilizan son:

**Escala de Inteligencia Wechsler WISC-R.** La aplicación de ésta Escala en éste modelo, es concebida como "predicador del rendimiento académico en las escuelas públicas, y se les llama Nivel de Funcionamiento Escolar." (Sattler, 1988:243)

La segunda medida es El Inventario de la Conducta Adaptativa (ABIC) para niños, que mide seis áreas de conducta: Familia, Compañeros, Comunidad, Escuela, Consumidor, Autocuidado.

La administración del Inventario (ABIC) consta de una entrevista a los padres del niño, tiene 242 preguntas divididas en dos secciones. La primera para todos los niños; la segunda, ordenada por niveles de edad.

Cada reactivo se califica con 1 o 5 tipos de respuesta. La fuerza principal del ABIC "radica en que es un instrumento estandarizado, que mide factores como el papel de la conducta en el hogar, vecindario y comunidad." (Sattler, 1988:278)

**Modelo Pluralístico:** Por medio de éste modelo y a través de las escalas culturales y la administración del WISC-R, se estima el Potencial de Aprendizaje, es "una información basada en una comparación entre la ejecución del individuo en Escala de Inteligencia y la ejecución de otros individuos pertenecientes a si mismo contexto sociocultural." (Estandarización, DGEE, 1982:36)

En el respecto, Sattler opina que la Validez para el uso del Potencial de Aprendizaje no se ha establecido. (1988:580)

El instrumento utilizado en éste modelo para evaluar o medir las características sociales y culturales de la familia del niño son: las Escalas Socioculturales, cuyos ítems se utilizan para predecir o interpretar la ejecución del niño en el WISC-R.

Estas Escalas constan de 11 preguntas sobre 4 áreas: Tamaño de la familia, Estructura familiar, Posición socioeconómica, Aculturación urbana.

Así, los puntajes crudos a quienes se les asigna un valor para las escalas, se convierten en puntajes escalares, que corresponden a los antecedentes étnicos del niño.

La administración del SOMPA se realiza en dos partes:

a. Se realiza una entrevista a los padres del niño, la cual consta de 3 medidas: Inventario de Conducta Adaptativa para niños; Escalas Socioculturales; e Inventario de Historia de Salud.

b. Se aplica al niño la Escala de Inteligencia Wechsler WISC-R; la Prueba Gestáltico Visomotora Bender; Agudeza Visual; Auditiva; y Relación Peso-Altura.

Según Sattler (1988), la administración de la Batería de pruebas SOMPA, se lleva a cabo aproximadamente 5 horas.

La muestra de estandarización de la Batería constó de 2085 niños, de las escuelas públicas de California, entre 5 y 12 años de edad. Los niños seleccionados pertenecían a los grupos étnicos de California, 699 blancos, 696 negros, 690 hispanos; seleccionados al

La aplicación de las pruebas (WISC-R y Bender) se realizó por psicólogos pertenecientes al mismo grupo étnico de los niños en estudio.

## **MODIFICACIONES Y REVISIONES**

La idea de desarrollar la Batería de pruebas SOMPA, según su autora, la Dra. Jane Mercer, surge de la necesidad detectada a través de las investigaciones efectuadas en Riverside, California, entre 1963 y 1965.

Los resultados de la investigación consistieron en que el maestro de grupo "detectaba un bajo nivel del rendimiento escolar, a los niños que se sospechaba podían tener retraso mental. Estos niños eran remitidos al psicólogo quien les aplicaba la Escala de inteligencia. Los niños negros e hispanoparlantes obtenían un C.I. menor de 79, en una proporción mayor que los niños norteamericanos, por lo tanto eran con más frecuencia admitidos a programas para niños retardados." (Estandarización, DGEE, 1982:8)

En el estudio se observó que los niños pertenecientes a estatus socioeconómicos bajos, en la escuela eran catalogados como retardados mentales, pero de acuerdo a sus padres y comunidad eran normales.

Se hizo también un estudio epidemiológico clínico para localizar por vía directa, a los niños con síntomas de retraso mental. Así entre 1972 y 1973 se emprendió la estandarización de la Batería SOMPA en Estados Unidos, y en 1978 Lewis y Mercer presentaron una Alternativa de Diagnóstico Integral, con todo un marco teórico referencial. Este año fué publicado el SOMPA por la Asociación Americana de Psicología.

En su país de origen, la Batería ha tenido diferentes críticas, entre ellas sobresalen "la falta de normatividades nacionales adecuadas, la representatividad de la muestra y la validez de pruebas de validez de sus medidas..." (Sattler, 1988:244)

***CAPITULO 6: REVISION DE LAS  
RUEBAS PSICOLOGICAS EN EL CAMPO  
EDUCATIVO EN MEXICO***

Después de haber realizado una investigación bibliográfica acerca de las bases teóricas y técnicas de los Instrumentos de Medición Psicológica, surge la pregunta ¿Qué pruebas se han dado en materia de evaluación en el ámbito educativo mexicano?.

En este capítulo se retomarán las pruebas anteriores expuestas como;

Instrumentos de investigación, revisando los trabajos de Tesis realizadas en diferentes Universidades del País, así como Investigaciones realizadas por Instituciones dedicadas a la Psicología y Psicopedagogía interesadas en las pruebas Psicométricas;

Instrumentos de Psicodiagnóstico, revisando el uso que, actualmente se les da en los Centros de Atención Especializada, (C.A.P.E.P. y D.G.E.E.) ya que éstas pruebas forman parte de la Batería de Diagnóstico Psicológico.

Finalmente se presentará en un cuadro de Análisis y Comentarios acerca de las pruebas revisadas.

## **SCALAS EVOLUTIVAS DEL DESARROLLO DE ARNOLD GESELL**

### **COMO INSTRUMENTO DE INVESTIGACIÓN**

En 1978 en la Universidad Iberoamericana (D.F.), se llevó a cabo una investigación (de licenciatura) sobre la "Aplicación del esquema evolutivo de Arnold Gesell a la lactante".

En 1982 Gutiérrez Ruíz y López, realizaron un "Programa de entrenamiento en el diagnóstico de Gesell y Amatruda a asistentes educativas, pertenecientes a un Centro de Desarrollo Infantil", siendo uno de sus principales objetivos:

Identificar la distribución de actividades del personal responsable de la sala de lactantes.

Evaluar el desarrollo y detección de problemas de lactantes, mediante el diagnóstico normal y anormal del niño de Arnold Gesell y Amatruda.

Orientar y capacitar al personal responsable en el diagnóstico normal y anormal del niño.

Esta investigación tomo como modelo el trabajo realizado por Alcázar y Galván en

En 1984 Castellanos, Orióstegui Pérez, y Servín Ortiz realizaron una investigación "El nivel de desarrollo de los menores que trabajan de acuerdo a la Teoría de Arnold Gesell".

Esta investigación se realizó durante 5 meses en 1982 en 19 mercados de 3 delegaciones políticas del D.F., con 18 menores entre los 13 y 14 años de edad que se encontraban trabajando.

Se ocuparon dos cuestionarios, uno de 35 reactivos para conocer el nivel socioeconómico y otro, basado en la Teoría de Arnold Gesell para conocer su nivel de desarrollo. Para éste último cuestionario crearon un banco de 320 reactivos, que fueron seleccionados mediante estudios piloto hasta quedar constituido por 88 reactivos de los cuales 37, son específicos de 13 años y 51 de otras edades cronológicas de desarrollo (de 15 años de edad).

Así, concluyeron que de acuerdo a los resultados del segundo cuestionario ningún menor se encuentra en el nivel de desarrollo que le corresponde según Gesell, ya que el porcentaje se encuentra por debajo del mismo de acuerdo a las normas que marca

...ll, descartándose de éste modo la posibilidad de que el trabajo afecte el desarrollo de menores en sus diferentes áreas.

En 1987 Vázquez Becil pretende demostrar que el hecho de asistir a una guardería entre un lapso de 5 años resulta estimulante en el desarrollo perceptual de un niño. Para determinar esto Vázquez Becil diseña una prueba que toma como marco referencial las Escalas de Evolutivas de Arnold Gesell (conducta motriz, social, adaptativa y de lenguaje), el Método de Evaluación de la Percepción Visual de Marianne Frostig.

Para el desarrollo del trabajo manejó dos grupos de niños entre los 5 y 6 años de edad; cada grupo se le aplicó la prueba de manera individual, un examen pediátrico, una consulta clínica psiquiátrica, y se incluyó la opinión de la educadora para detectar y registrar aquellos niños con algún problema de aprendizaje.

Grupo experimental: 20 niños (10 niños y 10 niñas) con una permanencia de más de 5 años en la guardería con la estimulación correspondiente al 3er grado de preescolar y los últimos años de estancia.

Grupo control: 20 niños (10 niños y 10 niñas), sin haber asistido a ninguna guardería en los años anteriores al ingreso de Jardín de Niños y la estimulación correspondiente al 3er grado de preescolar.

Vázquez Becil estableció los resultados de acuerdo a los puntajes de su prueba, concluyendo que la asistencia a un centro de Desarrollo Infantil desde temprana edad, tiene una influencia positiva en el desarrollo del conocimiento perceptual, aunque no en todos los niños.

En 1989, Hernández Sánchez observa la "psicomotricidad de niño Down de 1 a 3 años" a través de un análisis comparativo y de las características marcadas por Arnold Gesell. Este estudio se basa en los parámetros marcados por A. Gesell en su libro el "Diagnóstico del Desarrollo Normal y Anormal del niño" publicado en Argentina en 1979.

En 1989 Rincón Gallardo, realiza una investigación cuyo objetivo es ver que tanto el encuentro temprano favorece la afectividad y la estimulación materna, para lograr dicho objetivo se manejaron dos grupos de madres con sus recién nacidos del Instituto de Neonatología de nivel socioeconómico medio, un grupo con encuentro y otro grupo sin encuentro.

Para el seguimiento se utilizaron las Escalas Evolutivas de A. Gesell para los niños y la prueba de Perfil de la conducta materna de Bayley para las madres.

De acuerdo a los resultados, Rincón Gallardo concluye que el encuentro temprano no se refiere a la afectividad y estimulación materna: Argumentó que la prueba de Gesell usada para su investigación fue la estandarizada por el doctor Cravioto, (al revisar la bibliografía utilizada para éste trabajo, no se hace referencia a dicha estandarización).

En 1990 Cantón Arjona y Tadeo Madrazo realizan un estudio sobre la Construcción y presentación de las Escalas Evolutivas de A. Gesell, donde lo que hacen únicamente es la adaptación de dichas Escalas, con algunos cambios de presentación en cuanto al formato de la prueba, se basan en el libro de "Diagnóstico del Desarrollo Normal y Anormal del niño" publicado en Argentina en 1979.

## **COMO INSTRUMENTO DE PSICODIAGNOSTICO**

Este instrumento es parte de la Batería de pruebas de la DGEE, específicamente de las Escalas de Educación Especial y de los Centros de Intervención Temprana.

Estos Centros atienden a niños con alguna deficiencia en el desarrollo o con algún riesgo potencial de alto riesgo establecido en algún área del desarrollo.

En la DGEE, Las Escalas se llegan a utilizar en niños menores de 6.0 años, con algunos trastornos del desarrollo y consideran que los resultados de ésta prueba son útiles para el diseño de programas de atención en los niños.

El Manual que utilizan para el diagnóstico es el libro de "Diagnóstico del Desarrollo Normal y Anormal del Niño", publicado en Argentina en 1979. No se cuenta con la adaptación para la población infantil en México.

En los CAPEP, ésta prueba no forma parte de la Batería por que se considera que se usa en niños con daños severos en las diferentes áreas del desarrollo. El objetivo de los CAPEP es atender a los niños con leves alteraciones en el desarrollo, exclusivamente.

## **PRUEBA GUESTALTICO VISOMOTORA DE LAURETTA BENDER**

### **COMO INSTRUMENTO DE INVESTIGACIÓN**

Numerosas son las investigaciones que se han realizado en torno al Bender. Revisando la bibliografía nos damos cuenta de la extensa aplicabilidad de éste instrumento, tanto para el trabajo clínico como para el psicopedagógico, llegando a usarse para detectar problemas de lectura, aritmética; predecir el desempeño escolar; identificar y evaluar alteraciones emocionales, síntomas conductuales; diagnóstico de lesión cerebral, retraso mental; grado de desarrollo de la percepción visomotora, etc...

A pesar de la gran utilidad que se le ha conferido a éste instrumento en la investigación, en México no se cuenta aun con Normas de Calificación estandarizadas a la población infantil.

La mayoría de los trabajos de investigación, a nivel de Tesis de Licenciatura y grado, se refieren principalmente a estudios comparativos entre pruebas; a la aplicación de alteraciones en determinados ámbitos (clínicos, escolares, laborales) a la comprobación de la efectividad de la prueba, así como a la conveniencia de las normas de calificación. Sin embargo no se ha dado difusión a las propuestas que han surgido de otros investigadores (Rivera, 1974; Reguera y Rentería, 1979, Fuentes y Martínez, 1980).

A continuación mencionan trabajos de Tesis más representativos:

En 1974 Rivera llevó a cabo una investigación cuyos objetivos principales fueron: determinar la ejecución normal de niños mexicanos en la prueba Bender; establecer las normas de desarrollo para la población infantil mexicana y comparar la ejecución de niños mexicanos y americanos en la prueba.

Los sujetos de estudio fueron 1248 niños urbanos de escuelas públicas y privadas, residentes en el D.F. La aplicación y evaluación fué de acuerdo a las instrucciones elaboradas por Koppitz (1963).

Los resultados de ésta investigación indicaron atraso desde 6 meses hasta dos años de retraso en el desarrollo visomotor de los niños mexicanos respecto a los americanos.

Esto reveló que las normas propuestas por Koppitz no son las adecuadas para culturas diferentes de las norteamericanas y sugiere, desde entonces, una estandarización urgente de las normas de calificación de la prueba.

Dueñas Calderón y González Barney, en 1975, hicieron una investigación para ver si los niños con Daño Cerebral efectivamente presentan los indicadores ológicos que sugiere Koppitz.

Noverola Mondragón, en 1976, llevo a cabo una investigación en la cual, entre otros tivos, trato de establecer una comparación entre niños con diagnostico de D.C. y s normales respecto a sus habilidades percepto-motrices.

Reguera Rodríguez y Rentería Escalona, en 1979, trabajaron sobre un estudio parativo del Bender en niños y niñas de determinada edad, proponiendo entre sus lusiones, la necesidad de tener normas que se adapten a la población infantil icana.

Un estudio en niños mexicanos con las tablas normativas propuestas por Koppitz y ucci-Grajon se realizó en 1980 por Fabián Porchas. El objetivo fué el comparar los tados de la aplicación de ambas normas a la misma población.

En 1981, Mass Bustamante y Nuñez Amezcua realizaron una investigación acerca La Detección y Desarrollo del Aprendizaje a través de la lecto-escritura mediante ro pruebas psicométricas", entre ellas se considero al Bender.

Ray Bazán y Alvarez Páramo, desarrollaron en 1982 en la Cd. de Guadalajara, Jal., investigación sobre la utilidad del Bender y otras pruebas en la discriminación de lemas de aprendizaje en niños de alto y bajo riesgo, en un estudio de tipificación.

En 1990, se llevó a cabo un trabajo propuesto por Fuentes Montañés y Martínez valo con el titulo "Normalización del Test Guestáltico Visomotor de Lauretta Bender Escolares Mexicanos". Éste estudio se realizó con una muestra de 1440 niños icanos, asistentes a Jardín de Niños y Primaria, tanto inscritos en el sector oficial io en el sector privado, en el Distrito Federal.

Uno de los objetivos principales de ésta investigación fué el de establecer si los gnósticos basados en normas que no corresponden a la población escolar mexicana

adecuados, y si existe diferencia alguna entre la población infantil norteamericana y mexicana.

Para la aplicación y calificación de los protocolos de los niños, se basaron en el Sistema de Puntuación establecido por Koppitz (1963). Los resultados obtenidos en su investigación muestran que los datos normativos propuestos por la autora no son aplicables a la población mexicana, ya que presentan una diferencia de tres años y medio con respecto a la población norteamericana.

Por otro lado, se tuvo una entrevista con el Dr. Rogelio Díaz Guerrero, encargado del CAPAC con el objeto de conocer si en la Institución que el coordina, se han llevado a cabo revisiones o estandarizaciones del Bender, a lo cual nos informa que no existe a el momento ningún trabajo relacionado con ésta prueba.

En la entrevista realizada con la Dra. Enedina Villegas, docente e investigadora de la U.A.M., nos señaló que no se han llevado a cabo, en esa casa de estudios, estandarizaciones de las pruebas que se mencionan, remitiéndonos a los trabajos de Tesis de la Facultad de Psicología. Únicamente mencionó que se está trabajando sobre la validación y estandarización del Bender, en cuanto a las normas de calificación, por el Psicólogo Rogelio Díaz Gutiérrez, la cual se inició en Noviembre del año de 1993 pero no fue posible proporcionarnos ningún avance.

Como se puede observar, el Bender ha sido una prueba que se ha utilizado para diversas investigaciones, sin embargo se continúa con el Sistema de Calificación empleado en otras poblaciones, y a pesar de que se ha demostrado en algunos trabajos la importancia de su revisión, no ha habido continuidad en la investigación.

## COMO INSTRUMENTO DE PSICODIAGNOSTICO

En relación a los CAPEP, así como en la DGEE nos encontramos con que para la validación y evaluación del Bender, se utiliza la traducción del Manual de E. Koppitz (1963-1973).

Ésta prueba forma parte de la Batería básica, y el objetivo de su utilización es valorar el desarrollo de la percepción visomotora y establecer en el nivel de madurez en los niños escolares y escolares, así como obtener datos significativos de una posible alteración psicológica. (Lineamientos para la Evaluación Psicológica, DGEE, 1984:26)

Este instrumento es utilizado desde hace bastante tiempo, inicialmente se empleó el forma de Evaluación propuesto por la Dra. Bender, pero con el paso de los años se optó la evaluación propuesta por Koppitz, después de analizarla y considerarla como un ma objetivo y propio para la población infantil que ahí se atiende.

En ambas instituciones, nos encontramos con una tabla que se presenta a continuación:

## DATOS NORMATIVOS DE LA CALIFICACION DEL BENDER

### Puntuaciones individuales del Bender y Edades Equivalentes.

(Manual de normas y procedimientos del área de Psicología, 1988:93-B)

Puntuación del Bender	Edad Equivalente
21	4-0
20	4-0
19	4-1
18	4-2 a 4-3
17	4-4 a 4-5
16	4-6 a 4-7
15	4-8 a 4-9
14	4-10 a 4-11
13	5-0 a 5-1
12	5-2 a 5-3
11	5-4 a 5-5
10	5-6 a 5-8
9	5-9 a 5-11
8	6-0 a 6-5
7	6-6 a 6-11
6	7-0 a 7-5
5	7-6 a 7-11
4	8-0 a 8-5
3	8-6 a 8-11
2	9-0 a 9-11
1	10-0 a 10-11
0	11-0 a 11-11

Al cotejar ésta tabla con el Sistema de Evaluación propuesto por Koppitz, tabla 6 (Anexo 1), se advierte que no hay una línea de correspondencia, lo que nos lleva a pensar que el Sistema de Calificación utilizado ha sido establecido por otros autores.

Desafortunadamente en los Manuales de Lineamientos Normativos para la Evaluación Psicológica con que cuenta cada Centro la tabla aparece sin justificar su presentación. Por lo tanto no hay evidencia de que haya sido estandarizada o adecuada para la población a la que ahí se atiende.

Por tanto, pensamos que es necesario reconsiderar el Sistema de Puntuación que se emplea para no incurrir en una interpretación equivocada en relación a la edad de maduración visomotriz de cada niño, y por ende, establecer un diagnóstico poco certero.

## PRUEBA MONTERREY

### COMO INSTRUMENTO DE INVESTIGACIÓN

A través de un convenio entre la Dirección General de Educación Especial (SEP) y la Secretaría de Servicios Sociales y Culturales del Estado de Nuevo León, en la Ciudad de Monterrey; se llevó a cabo en 1974, la construcción de una prueba psicopedagógica. Sus objetivos fueron: Detectar, diagnosticar y reeducar a niños que presentan problemas de aprendizaje en las escuelas públicas de ésta ciudad. (DGEE, Plan Nuevo León, 1974: 9)

La prueba Monterrey, como se le llamó fué elaborada bajo la dirección de la Dra. Margarita Gómez Palacio.

Hasta 1981, la Prueba Monterrey, contaba ya con 5 revisiones, de las cuales no se menciona en ninguna de ellas como se llevó a cabo el proceso de estandarización; sino solamente se da información relacionada con el contenido y forma de la prueba.

Como anterior se debe a que las fuentes de información son limitadas. Las entrevistas se intentaron realizar con las personas involucradas en la investigación no se pudieron llevar a cabo. Únicamente la Licenciada Magdalena Del Río <sup>(1)</sup>, informó<sup>(2)</sup> que la Prueba Monterrey fué elaborada y modificada en diversas ocasiones pero nunca fué estandarizada por lo que no existen documentos acerca de esos datos.

En lo que respecta a otro tipo de investigaciones o trabajos de tesis, salvo el estudio de la continuación se menciona no se encontró ningún otro.

En 1984 Rubio Traspeña, Coordinadora de Grupos Integrados, en el Estado de Nuevo León, realizó un instructivo de aplicación de la Prueba Monterrey; ya que consideraba que los aplicadores, debían ser seleccionados y capacitados previamente. En su elaboración incluyó aspectos teóricos acerca del desarrollo de la inteligencia del niño; la forma de ubicar al niño en el perfil de la Prueba; y modificó algunas consignas confusas y poco comprensibles.

Rubio conformó la muestra utilizada con el personal del Grupo Integrado donde trabajaba; maestros, médicos, trabajadoras sociales y psicólogos, formando 2 grupos: Control, 7 sujetos, capacitados tradicionalmente; y Experimental, 7 sujetos, capacitados en éste instructivo.

Los resultados obtenidos por Rubio fueron: Debido a la preparación con que cuentan los maestros, fué difícil para ellos entender la teoría piagetiana; existió una tendencia a

ar a los niños en niveles superiores al que les corresponde; fué mejor la aplicación y  
icación de la Prueba realizada por el grupo experimental.

## **COMO INSTRUMENTO DE PSICODIAGNOSTICO**

En todos los Grupos Integrados<sup>(1)</sup> existentes en la Ciudad de México, la Prueba  
terrey es un instrumento esencial para la detección y selección de los niños que  
entran problemas en el proceso de la adquisición de la lecto-escritura y el cálculo  
mático.

La Prueba es aplicada a los niños cuya edad cronológica se ubica entre los 6 y 9 años  
dad, en caso de que un niño no corresponda a Grupo Integrado según los resultados  
a Prueba, se canalizara a COEC para que se le realice un estudio psicológico  
encial.

## SCALA DE INTELIGENCIA STANFORD-BINET (TERMAN-MERRILL)

### OMO INSTRUMENTO DE INVESTIGACIÓN

pesar de que la Escala Stanford-Binet ha sido utilizada como punto de partida para rosas investigaciones y adaptaciones en otros países, en México es muy escasa la ografía, tanto del material de Tesis como para las Instituciones dedicadas al diagnóstico.

as primeras pruebas de inteligencia que se utilizaron en México, fueron, al parecer, igen europeo. Éste hecho contribuyo a la formación de investigadores mexicanos ados al estudio de los problemas psicológicos.

En 1918 el Dr. Rafael Santamarina comenzó el estudio y adaptación de la Escala al de Binet-Simon en México.

En 1921 presentó su primera adaptación provisional en el Primer Congreso americano del Niño. Una segunda adaptación fué publicada en el V Congreso, como ado de los trabajos que se habían venido realizando a través de la Oficina de tificaciones del Departamento de Salubridad Publica, y más tarde del Departamento icopedagogía y Médico Escolar.

Por la misma época, el Profesor David Pablo Boder, encargado de la Sección de técnica, del mencionado Departamento, realizó una traducción de la Revisión ord de la Escala Binet-Simon, con algunas modificaciones apriori.

sta adaptación se realizó directamente de la Escala Francesa por considerarla más que la adaptada a niños de Estados Unidos. (Ortega y López, 1940:20)

En 1940 el Instituto Nacional de Psicopedagogía, realiza un Estudio de nderización de la prueba Colectiva de Inteligencia General Forma "A" de L. M. an, dirigido por los Dres. Pablo Ortega y Matías López Jr. (Ortega y López, 1940)

ara llevar a cabo el proceso de estandarización de la prueba, se aplicó a niños de 5o. de primaria, así como de 1o. de secundaria. La primera aplicación se realizó en a un total de 587 alumnos de los grados mencionados.

a segunda aplicación se realizó en 1939 a un grupo de 2,938 alumnos de los mismos os; la tercera muestra se obtuvo en julio de 1939, de un total de 1,379 alumnos. (García y López, 1940:39)

Los autores de la investigación señalan que únicamente se tomó en cuenta el grado de madurez de los sujetos, pues la finalidad del estudio pretendía conocer la habilidad mental de los alumnos que cursaban el 5o., 6o. de primaria y 1o. de secundaria por lo que la prueba únicamente podría ser aplicada a los alumnos de estos grados.

En relación a los estudios de Tesis, las investigaciones son escasas, únicamente se encuentran algunos estudios de correlación entre la Escala Terman-Merrill con otras pruebas como: Wechsler, Rtgerts, Bender; Normalizaciones de la Escala para medir la personalidad; Estudio de correlación entre Bender y Terman-Merrill en Niños con Retardos Mentales; Efectos de diferentes contingencias de reforzamiento sobre la conducta en la Prueba Stanford-Binet y Matrices Progresivas.

No se encontró ningún trabajo de adaptación o estandarización de la Escala aplicado a la práctica, salvo el trabajo antes mencionado.

## COMO INSTRUMENTO DE PSICODIAGNOSTICO

En los Centros de Atención (CAPEP) y (DGEE) la Escala Stanford-Binet Terman-Merrill (1975) forma parte de la Batería de pruebas básica, seleccionada para el diagnóstico.

El Manual que utilizan ambas Instituciones es: Medida de la Inteligencia, (Método de aplicación y empleo de las pruebas de Stanford-Binet, tercera revisión L y M, 1960). El Manual es una traducción del Inglés y adaptación al Castellano por el Dr. José Germain y Sureda, ésta impreso en España en 1975 e incluye las tablas para el Cálculo de los cocientes de Samuel R. Pinneau (1960 y 1972).

El uso de ésta Escala en los Centros, según el personal entrevistado, obedece principalmente a la Validez y Confiabilidad del instrumento, a la facilidad y rapidez en su aplicación e interpretación, así como al material que contiene, ya que éste es atractivo para los niños.

Por lo que se pudo observar, el material de aplicación con que cuentan los psicólogos en su mayoría, una reproducción del material original.

Esto se debe principalmente al escaso presupuesto que se otorga a los Centros, así como a lo costoso de la prueba original. Aunque se ha tratado de conservar el material de

abpruebas, no todo es homogéneo. Los manuales de aplicación e interpretación son is fotostáticas del original.

n relación a los manuales, la traducción y adaptación española presenta algunos eptos propios del idioma que en nuestro país no son utilizados, así en la aplicación aquí se hace se han cambiado **arbitrariamente**, sin llevar a cabo ningún estudio al, guiándose las modificaciones por la frecuencia de respuestas de los niños y la riciencia del personal que aplica el instrumento.

continuación se presentan los reactivos originales y las modificaciones que se han o:

### Vocabulario de Imágenes (Años: II.6; III y IV)

Uno de los dibujos que se presentan es una chaqueta, el criterio de evaluación ivo es: chaqueta o americana. En los Centros se considera positivo saco o traje.

### Diferencias (Año VI, reactivo 2)

La pregunta es la diferencia entre una zapatilla y un zapato. Aquí la palabra que se tuye es zapatilla por pantufla.

n el mismo año VI, reactivo 5, se pregunta "la punta de un bastón es redonda; la de avajaja es...", se sustituye a la palabra navaja por cuchillo.

En el año VII, reactivo 2 la pregunta es: "en que se parecen una manzana y un cotón", se sustituye la palabra melocotón por durazno o naranja.

En el año VIII, reactivo 1, se exploran palabras de vocabulario:

Traducción:	Centros:
Naranja	Naranja
Sobre	Llave
Paja	Sobre
Charco	Prisa
Grifo	Charco
Bata	Bata
Rugido	Pestanas
Pestanas	Reposar
Marte	Mirada

Jugar	Tostar
Tostar	Paja
Discurso	Mosaico
Destreza	Morena
Morena	Rugido
Bozal	Enojo
Prisa	Marte
Particularidad	Tolerar
Inapreciable	Lamentar
Mirada	Discurso
Tolerar	Bozal
Desproporcionado	Astuto
Loto	Estrofa
Astuto	Percibe
Mosaico	Altanero
Estrofa	Desproporcionado
Lamentar	Superficial
Ocre	Destreza
Reposar	Particularidad
Ambar	Incrustación

En el mismo bloque de edad, año VIII, reactivo 2 se explora Memoria de Cuentos:

Traducción:

"La jaca traviesa"

Había una vez una niña que se llamaba Lola. Vivía en un cortijo con su hermano Ricardo. Un día el padre de estos niños les regalo una jaca andaluza. Se divertían mucho con ella. Una vez Ricardo la estaba montando, se asustó la jaca y echo a correr al galope. Sobre Ricardo se cayó en un charco. Como se reía Lola cuando le vio!. Estaba cubierto de pies a cabeza."

Centros:

"El caballo travieso"

Había una vez una niña que se llamaba Lola. Vivía en un rancho con su hermano Ricardo. Un día el padre de estos niños les regaló un caballo. Se divertían mucho con él. Una vez que Ricardo lo estaba montando, se asustó el caballo y echó a correr a galope. El

e Ricardo se cayó en un charco. Como se reía Lola cuando lo vió!. Estaba cubierto do de pies a cabeza."

Como se puede observar, se han modificado algunos términos y conceptos que son necesarios, no cuentan con un análisis de reactivos que los apoye, como lo han los investigadores para considerar la adaptación de una prueba.(Landesheere, :25)

En relación a la calificación de la Escala, se señala que en los CAPEP, la tabla que se usa para determinar el Coeficiente Intelectual de los niños, es la propuesta por Pinneau 1960, mientras que la tabla que se utiliza en la DGEE, es la elaborada por el mismo en 1972.

Los criterios que se tomaron en ambas Instituciones no son válidos, ya que por un lado se menciona que la población con la que fueron normadas las tablas de 1960 es muy similar a la población infantil mexicana en cuanto a nivel socioeconómico y cultural. Por otro lado, el criterio de utilizar las tablas de 1972 es porque son las "actuales".

El resultado de ésta diversificación de niveles intelectuales con que los niños son diagnosticados y canalizados es totalmente inseguro y ambigüo, a la ya de por sí crítica que se hace alrededor de las pruebas psicológicas.

Para la interpretación de la Escala, se ha establecido para ambas instituciones el Sistema de Clasificación propuesto por Sattler (1988:523-527), que identifica tipos de niveles o categorías.

## RUEBA DEL DIBUJO DE LA FIGURA HUMANA

### OMO INSTRUMENTO DE INVESTIGACIÓN

En 1965, Ahumada Rodríguez realizó una investigación con el objetivo de conocer sucede cuando dos pruebas construidas y estandarizadas fuera de una sociocultura ntroducidas en ella y se comparan entre si.

La muestra estuvo constituida por 68 niños entre los 6 y 12 años de edad, de la tra del IDPEM.

Los instrumentos utilizados para comparación de puntajes fueron el WISC y la prueba dibujo de un Hombre (Goodenough). Se rechazó para el estudio a aquellos niños que alcanzaron un C.I. de 80 y que registraron en el HIT (Prueba de Percepción y validad) problemas de personalidad.

A partir de los puntajes obtenidos en ambas pruebas, Ahumada Rodríguez concluyó

La prueba de Goodenough tiene baja correlación con los puntajes del WISC (Total, de Ejecución y Verbal), y que no es una prueba apta para medir inteligencia global, porque parece ser que en México, es un instrumento de mayor validez para niñas que para niños, además los puntajes de los niños mexicanos son significativamente bajos en relación con los puntajes de los niños norteamericanos.

Las calificaciones de los varones en el Goodenough, no se correlacionan por arriba de .80 con ningún subtest del WISC a diferencia de las mujeres.

La intervalidez entre el WISC y el Goodenough disminuye al aumentar la edad de los sujetos. La escala verbal del WISC tiene menor intervalidez con la del Goodenough que la escala de ejecución. Los subtest del WISC, comprensión y retención de dígitos, carecen de factores comunes con la prueba de Goodenough.

En 1965, Reyes Lagunes realizó una investigación en la que pretendía conocer los res que intervienen en la personalidad infantil y su desarrollo ligado a los años lares. La muestra constó de 68 niños (muestra del IDPEM).

Los instrumentos utilizados fueron el WISC y la prueba de Goodenough, en donde los adadores igualaron la forma de aplicación e interpretación para obtener acuerdo con los ólogos encargados de la muestra norteamericana e intercambiar experiencias respecto aplicación y calificación.

eyes Lagunes concluyó que:

Los factores socioculturales modifican de una o de otra forma los resultados de las subpruebas, algunas situaciones presentadas en varios reactivos fueron desconocidas para el niño. Además de que fueron escasos los cambios en el grado de dificultad de las subpruebas de la escala de Ejecución, considerándose que sufre menor grado de impacto sociocultural.

eyes Lagunes sugirió que dada la tendencia de aumentar el rendimiento intelectual, al aumento de edad, es necesario realizar una investigación con el fin de saber ¿son los factores que influyen en éste cambio.

En 1966 Gomez Aldama, realizó una investigación en donde se plantean tres objetivos:

Esclarecer el problema que se presenta a nuestro sistema escolar público, en la realización de adecuados exámenes de admisión al primer año de educación primaria.

Informar sobre el resultado estadístico de correlación entre las pruebas Goodenough y Harris-Goodenough.

Presentar por primera vez en México la prueba de Harris-Goodenough.

La muestra constó primeramente de 150 alumnos, de 4 Jardines de Niños, dos del norte de la ciudad y dos del sur, con edades de 5 años a 5 años 11 meses. Se administró la prueba de Goodenough y la de Harris-Goodenough, la muestra final estuvo constituida por 106 casos (muestra del IDPEM).

Gómez Aldama señala que la prueba de Harris-Goodenough se basa en la original de Goodenough, en cuanto a construcción y marco teórico.

No se encontraron diferencias en el rendimiento global de la muestra cuando se midió por zona escolar; sin embargo, al obtener el resultado global de la prueba de Goodenough, ésta se desvió de manera significativa con el ideal de la muestra representativa de los Estados Unidos, puesto que para la transformación de calificaciones originales se utilizaron las tablas de la estandarización original, por lo que es factible que esto haya afectado la evaluación real de niños con características socioculturales educativas diferentes a los que sirvieron para desarrollar las normas originales.

ambién concluye que es suficiente utilizar la prueba de Goodenough para determinar el grado de desarrollo conceptual general de niños preescolares, por tratarse de una prueba más sencilla y conocida que la de Harris-Goodenough.

En 1968 Vallejos Dellaluna realizó una investigación donde buscó las causas que ocasionan variaciones en los resultados de la prueba de Raven y Goodenough. La muestra fue de 100 varones entre los 6.6 y 7.6 años de edad, en una escuela primaria particular.

Concluyó que las dos pruebas miden lo mismo, la correlación fué de .73; la prueba de Goodenough es un instrumento útil para el exámen del desarrollo mental de los niños. Concluyó que la prueba de Raven es mejor porque no influye el aspecto sociocultural en la de Goodenough.

En 1968 Patrón Castillo hace un estudio de validación cruzada con el WISC, Goodenough y Harris-Goodenough, analizando sus resultados en diferentes edades, sexos y tipos de escuelas.

Sus hipótesis fueron:

Que la prueba de Goodenough y Harris-Goodenough tienen mayor correlación en estudios con niños pequeños y menor correlación en niños en edad avanzada.

Que éstas dos pruebas pueden ser sustituidas por el WISC, considerando la baja correlación de la Figura Humana por intervenir en ella factores proyectivos.

Investigar si los puntajes de Harris-Goodenough funcionan en niños mexicanos o si por razones de cultura producen calificaciones inferiores.

A partir de la muestra del IDPEM, se obtuvo una muestra representativa de 660 casos, mitad hombres y mitad mujeres, con edades entre los 6 y 12 años, de diferentes tipos de escuelas (particulares y públicas). Las pruebas fueron aplicadas de manera individual a los casos que fueron la muestra final.

A partir del análisis de los puntajes, Patrón Castillo concluye:

Que la validación cruzada entre los puntajes de las pruebas, a pesar de tratarse de pruebas americanas que no han sido estandarizadas en nuestro país, es significativa.

Que la prueba de Goodenough tiene un alto grado de correlación con Harris-Goodenough, a diferencia del dibujo de la figura femenina. Y que, aunque

las medidas de puntajes aumentan con la edad, no alcanzan el puntaje obtenido en las normas de Estados Unidos.

Que la prueba de Goodenough es suficiente para determinar el grado de desarrollo conceptual de niños preescolares, por tratarse de una técnica sencilla, rápida y conocida.

En 1969 Ahumada Rodríguez realizó una investigación cuyo objetivo fué confrontar /WISC, el Goodenough y el Harris-Goodenough, para conocer hasta que grado miden mismo y contribuir al entendimiento del individuo en constante proceso de cambio; medio de las curvas de desarrollo de puntuaciones con la edad.

Tomó la muestra del IDPEM, y a través de la aplicación de estos instrumentos concluyó que la diferencia en puntuaciones por sexo, son estadísticamente más significativas que por edad. También concluyó que por ser instrumentos no construidos en México, las puntuaciones de los escolares mexicanos, en comparación con las normas de estandarización de Estados Unidos, "se ven penalizadas y ninguno de los subgrupos alcanza el nivel de normalidad". (Ahumada, 1969:143)

Por tanto sugiere que la prueba del WISC, Goodenough y Harris-Goodenough se validen en nuestra cultura para un uso adecuado.

En 1976 Canudas González llevó a cabo una investigación para saber si las pruebas Rutgers-Drawin, Detroit-Engel, Frostig, Goodenough y la prueba del Dibujo de la cara Humana de E. Koppitz, poseen validez predictiva sobre el posterior comportamiento académico en los primeros grados de instrucción primaria.

La muestra estuvo formada por alumnos de cuatro generaciones de 1971 a 1975, de escuela particular incorporada a la SEP, donde se considera que los niños que ahí están son de clase media alta.

El número total de alumnos fué de 248 (118 niños y 130 niñas) cuyas edades variaban entre los 5.6 años y 7.6 años en la fecha de aplicación de las pruebas; la selección de estas se realizó en la escuela durante los meses de agosto de 1971 a 1974 y febrero de 1975, antes de que cada generación iniciara las clases del primer grado de instrucción primaria.

A la generación de 1971 se le administraron las pruebas de Rutgers-Drawing, Detroit-Engel y Goodenough.

Las generaciones de 1972 y 1973 se les aplicaron las pruebas de Rutgers-Drawing, Detroit-Engel y el DFH de E. Koppitz.

La generación de 1974 se le aplicaron todas las pruebas anteriores además de la de Frostig.

El total de pruebas de Goodenough aplicadas fué de 96, y 144 de la prueba de DFH.

Se correlacionó los puntajes obtenidos de todas las pruebas, con las calificaciones académicas de los niños y con los puntajes de las pruebas pedagógicas administradas por los maestros para evaluar el curso.

Se concluye que las correlaciones más altas obtenidas entre los puntajes de cada una de las pruebas, calificaciones y puntajes de las pruebas pedagógicas de cada grupo, se obtuvieron en las pruebas de Frostig, Rutgers-Drawing y Detroit-Engel, lo cual indica que estas pruebas miden conductas que un niño requiere para cursar satisfactoriamente los primeros grados de instrucción primaria, como son coordinación visomotora, discriminación perceptual, seguimiento de instrucciones.

Los puntajes más bajos fueron en la prueba de Goodenough y la prueba del DFH de E. Koppitz, por lo que considera que éstas pruebas no miden conductas relacionadas con la capacidad del niño para aprender los conocimientos de los primeros grados de instrucción primaria.

En 1978 Ramírez López realizó un estudio, siendo su objetivo primario ver si los niños diagnosticados con trastornos de aprendizaje, a través de elementos clínicos y de la prueba de Frostig, cuyo C.I. esté por debajo del promedio y que además sus edades sean los 9 años, son susceptibles de mejoría en sus áreas perceptuales bajas, merced al entrenamiento percepto-motor de Frostig en un lapso de 15 semanas.

Los instrumentos utilizados fueron la prueba de Goodenough, y la prueba de Frostig, nuestra fué seleccionada de acuerdo a la forma que sigue la admisión de pacientes la Unidad de Salud del Centro Comunitario de Salud Mental del IMSS.

Después de aplicadas e interpretadas las pruebas Ramírez concluyó que los resultados del entrenamiento percepto-motor no fueron estadísticamente significativos. Sin embargo la observación clínica si existen cambios en el comportamiento.

Se concluye que en los puntajes de la Prueba de Goodenough, los sujetos que habían obtenido en el test los más bajos puntajes, en el retest obtuvieron alguna de las más altas calificaciones y los sujetos que tenían puntajes altos en el test, bajaron en el retest."

írez, 1978:110) Concluyendo que hubo interferencia de variables extrañas no controladas ni controladas a tiempo.

En 1971 Acle Tomasini realizó un estudio para saber si existe correlación estadística entre el WISC y el Harris-Goodenough trabajó con 45 niños varones de nivel socioeconómico alto, de una escuela particular cuyas edades fluctuaron entre los 7 y 11 años de edad dividiéndolos en cinco niveles de edad 7, 8, 9, 10 y 11 años, las formas se administraron de manera individual.

Acle Tomasini concluye que no existe una correlación estadísticamente significativa entre ambas pruebas, y que ambas fueron susceptibles de evaluar el desarrollo intelectual al intentar medir distintas funciones.

En 1971 Pallares Díaz averiguó la Confiabilidad de la prueba de Harris-Goodenough, como instrumento de medición intelectual, mediante el Método de retest con un intervalo aproximado de dos meses entre ambas aplicaciones.

La muestra se hizo al azar, fué un grupo de varones entre los 7 y 10 años de edad de escuela particular de los cuatro primeros grados escolares, en un principio estuvo formada por 60 niños que se dividieron en cuatro niveles según su edad cronológica se redujeron a 3, reduciéndose la muestra a 57 sujetos.

La prueba de Harris-Goodenough se aplicó de manera individual, y con la interpretación de Pallares concluyó que la correlación de confiabilidad de la escala total en los dos grupos ésta situada en un nivel de significancia de 01 y que sirve de base para afirmar que ésta técnica se puede aplicar e interpretar con firmeza en sujetos con características similares a la muestra estudiada.

En 1984 Rodríguez Sierra realizó una investigación para ver si existe relación de las dopaminérgicas centrales con aprendizaje, memoria y retardo mental, tratando de resolver los problemas de aprendizaje que se manifiestan en los niños con retardo mental para demostrar la utilidad terapéutica de los medicamentos que fueron administrados.

La muestra constó de 10 niños de ambos sexos, entre 6 y 13 años de edad, estudiantes de diferentes niveles del Centro de estudios de psicología humana, después se descartó a los que no quisieron participar, por que sus padres se opusieron a que continuara participando.

Primeramente se les aplicó el WISC, Bender y Harris-Goodenough, después se les administró Fosfatidilcolina y Benadon 300mg, durante cuatro meses y por último se

ieron a aplicar las pruebas. Las diferencias reportadas antes y después del camiento en los puntajes de cada una de las pruebas no resultaban ser lísticamente significativas.

Uno de los últimos estudios que utiliza la prueba de Harris-Goodenough es el de hez Carlos en 1987, siendo su objetivo ver si existen diferencias de identificación y ocimiento de la figura paterna en niños de madres solteras y de familias integradas.

En 1979 Heyerdahl G. comparó los resultados obtenidos por E. Koppitz y los ntrados por ella en niños mexicanos. Sus resultados mostraron diferencias culturales ficativas y diferencias en los indicadores emocionales de ambos.

En 1980 Barocio y De la Teja hicieron una investigación para comparar los tados obtenidos por E. Koppitz en niños americanos y los obtenidos con niños canos de nivel escolar. Concluyeron que sí existen diferencias significativas en los s evolutivos conforme aumentaba la edad de los niños. Además de que los Dibujos Figura Humana de los niños mexicanos presentan determinadas características, que icuentran en niños norteamericanos con lesión cerebral, por lo que consideran que es probable que la deficiencia alimenticia de los niños mexicanos, traducida en durez y deficiencias neurológicas, se manifiesta a través de las características idas en el dibujo de la figura humana. (Barocio, 1980)

En 1981 Vargas Oseguera realiza una investigación para ver si los niños de nivel económico bajo tienen mayor dificultad en el desarrollo de la Percepción Visual; si e diferencia entre éste desarrollo en niños y niñas; y si hay relación entre la igencia y el desarrollo de la Percepción Visual. La muestra fué de 100 sujetos, 50 de sexo de una escuela primaria pública y otros 100 sujetos 50 de cada sexo de una ela primaria particular. Utilizó la prueba del dibujo de la figura humana de Elizabeth pitz, y la Prueba de Frostig.

Concluye que los niños de nivel socioeconómico bajo tienen mayor dificultad en el rrollo de la percepción visual que los niños de nivel alto; las niñas tienen un mejor rrollo de la percepción visual que los niños en ambos niveles socioeconómicos; y sí e una relación entre la inteligencia y el desarrollo de la percepción visual, lo que luye a través de la interpretación de los puntajes de las pruebas.

En 1983 Lanz Mora realiza una investigación a partir del problema de la normalización de pruebas aquí en México en donde analiza las respuestas de los niños canos ante la prueba del dibujo de la figura humana de E. Koppitz, principalmente lo respecta a los indicadores emocionales.

La muestra estuvo constituida de 150 niños (73 niñas y 77 niños) entre los 6 y 12 años de edad, de dos escuelas primarias del Sur de la ciudad, una de turno matutino y otra de vespertino, de clase baja.

La prueba del DFH se aplicó de manera colectiva en grupos de 10 y 12 niños; a 48 niños con problemas de conducta se les aplicó la prueba de Bender a 8 de los cuales, se detectó un probable daño cerebral.

Lanz Mora dió los resultados de acuerdo a sus hipótesis:

No hubo diferencia entre los puntajes aceptados por Koppitz para ser considerados como indicadores emocionales y los encontrados en la muestra. Sin embargo los puntajes obtenidos de la muestra son mayores al 16%.

Los niños con problemas de conducta, presentaron con mayor frecuencia, indicadores emocionales que los niños sin problemas. No existen diferencias significativas entre un grupo y otro.

Las niñas y los niños presentaron indicadores propios de cada sexo. Pero según los puntajes y la homogeneidad de la muestra no hubo ninguna diferencia.

Los niños que asisten al turno matutino y vespertino, tienen diferencias en los puntajes. Ésta hipótesis se rechazó.

En 1985 Venegas Padilla realiza una investigación, cuyo objetivo es analizar a niños con problemas de rendimiento escolar pertenecientes a la comunidad de Tepito.

Sus hipótesis fueron: Los niños con mayores alteraciones en su madurez presentan mayores problemas emocionales y la madurez de un niño ésta relacionada directamente a la madurez emocional o viceversa.

Los instrumentos que se utilizaron fueron la prueba de Bender, la del DFH de E. Koppitz y un cuestionario médico, psicológico y socioeconómico. La muestra después de un estudio sistemático quedó constituida por 70 niños (35 niñas y 35 niños), entre los 6 y 12 años de edad, que fueran reprobados y con bajas calificaciones, pertenecientes a las escuelas de la comunidad.

Las pruebas se aplicaron de manera individual, la prueba del DFH de E. Koppitz permitió detectar alteraciones evolutivas relacionadas a la edad mental y emocional en nivel de edad.

En éste estudio se elaboraron normas de acuerdo a la muestra siguiendo el proceso de Koppitz.

Benegas Padilla concluye que la madurez no se relaciona en forma directa con el nivel emocional, aunque constituye un factor de peso cuando presenta actitudes emocionales inadecuadas.

## COMO INSTRUMENTO DE PSICODIAGNOSTICO

En los CAPEP la prueba del dibujo de la figura humana es un instrumento que se utiliza como una prueba de inteligencia a partir de los parámetros de F. Goodenough. Normalmente los niños son remitidos a estos Centros por las educadoras de los Jardines de Infancia, son enviados porque obtienen un rendimiento muy pobre en su ejecución. De manera que la prueba se utiliza como un elemento diagnóstico para obtener información acerca de los niños.

La prueba de Harris-Goodenough, es utilizada muy pocas veces como prueba de inteligencia en algunos Centros dependientes de la DGEE, ya que se le da prioridad a la cualitativa, considerando que el dibujo de la figura humana en determinada edad se utiliza para observar la integración del esquema corporal y el concepto que el niño tiene de sí mismo.

También en algunos de estos Centros se utiliza la prueba del dibujo de la figura humana bajo el criterio de Koppitz.

Se aplican a niños cuyas edades fluctúan entre los 5 y 12 años de edad aunque para los mayores de 12 años se sugiere se utilice tan solo como rapport". (Lineamientos, DGEE, 1984:26)

Con la aplicación de ésta prueba "se obtienen datos generales sobre el niño. No es conveniente reportar el Coeficiente Intelectual" (Lineamientos, DGEE, 1984:26), normalmente se proporciona una interpretación cualitativa.

## **ESCALAS DE INTELIGENCIA PARA LOS NIVELES PREESCOLAR Y PRIMARIA. (WPPSI-WISC)**

### **COMO INSTRUMENTO DE INVESTIGACIÓN (WPPSI)**

En 1977 De la Parra Aguilera, pretende comprobar si existen diferencias en el rendimiento intelectual de los preescolares debidas al nivel socioeconómico al que pertenecen. Utilizó 78 niños de guarderías del Departamento del Distrito Federal, hijos de obreros (50%) y de pepenadores (50%). Aplicó la traducción del WPPSI sin ninguna modificación.

Los resultados obtenidos por De la Parra fueron: Existieron diferencias significativas en el rendimiento intelectual entre los dos grupos; los resultados correlacionales de las subpruebas fueron satisfactorios, por lo que De la Parra consideró que la traducción tuvo su validez de construcción.

En 1981 Coto Villa, realizó una investigación para determinar el nivel intelectual de los mexicanos con el WPPSI. Tenía como objetivos traducir, adaptar y detectar el nivel de confiabilidad y validez de la prueba.

Realizó la aplicación a 60 niños (jardín de niños particular) de nivel socioeconómico medio alto, clasificados en cuatro grupos: de 15 niños cada uno, de los 4 a los 6 años de edad.

La confiabilidad se obtuvo por medio de la fórmula de Pearson y la validez por medio de la diferenciación con la edad.

Validez total: con el aumento de la edad cronológica, se dió un avance progresivo en las puntuaciones, consideró válidas las subpruebas de información, vocabulario, aritmética, laberintos, completamiento de figuras, diseño geométrico, diseño con bloques y comprensión.

Confiabilidad total: las subpruebas de información, vocabulario, comprensión y aritméticas incrementaron su nivel de significancia al valor de 0.01; en los cuatro grupos de edad estudiados las subpruebas de información, vocabulario, aritmética, completamiento de figuras y laberintos, resultaron confiables.

En 1984 Soberanis Quintero investigó si existen diferencias significativas entre las normas normales y su C.I. proporcionado, para lo cual comparó los resultados de las escalas: Terman Merrill, WISC y WPPSI.

Integró la muestra por 30 niños (50% de cada sexo) de 5 a 7 años de edad, que iban al servicio de psicología del Instituto de Salud Mental del DIF.

Soberanis concluyó que según los resultados de las correlaciones de las tres escalas, se pueden considerar válidas y confiables para realizar una evaluación.

En 1987 Caparroso Chávez, realizó un estudio exploratorio entre los problemas de conducta y la disfunción cerebral mínima evaluado a través de las pruebas: Bender, WISC, y el Dibujo de la Figura Humana. Utilizó 21 niños de 4 a 6 años de edad pertenecientes a Centros de Desarrollo Infantil de la SEP.

Como las pruebas fueron aplicadas sin ninguna modificación, se utilizaron como estándares para la detección de la lesión cerebral.

En 1990 Castelan Martínez realizó una normalización del WPPSI (en la cual siguió los mismos principios utilizados por Wechsler para validar y confiabilizar sus escalas), a pesar de las diferencias estadísticas de los resultados obtenidos por los niños utilizados en el estudio y de los niños americanos utilizados para la estandarización de esa prueba.

Constituyó la muestra por 12 escuelas del D.F., formando tres grupos de 4 escuelas en cada uno: particular, oficial y CENDI; utilizó 270 sujetos (entre los 4 y 7 años de edad) distribuidos equitativamente en cada grupo. Se aplicó el WPPSI según la traducción del Manual Moderno.

Los resultados obtenidos por Castelan fueron los siguientes:

Existieron diferencias estadísticamente significativas (.05) entre las normas de la prueba original y las obtenidas en ésta investigación.

Se mantuvo una desviación estándar de 3 con media de 15 para cada subprueba, y una desviación estándar de 15 y una media de 100 para los puntajes totales.

Castelan consideró que la traducción del WPPSI resultó válida y confiable al aplicarse a los niños mexicanos, y que los tres subsistemas educativos presentaron características homogéneas, encontrándose puntajes más altos en las escuelas particulares.

## **COMO INSTRUMENTO DE PSICODIAGNOSTICO**

El WPPSI está contemplado como un instrumento de psicodiagnóstico que puede ser usado a los niños que asisten al servicio de CAPEP.

sin embargo no es una prueba que se utilice en forma regular, debido a que no todos los psicólogos de los Centros cuentan con el material. Algunos no saben aplicarlo o interpretar y, dado el número de niños que ahí se atienden, prefieren utilizar la Escala de Binet-Merrill (Traducción, 1975) ya que su aplicación e interpretación es más accesible y rápida.

Cuando es utilizado el WPPSI, el material de aplicación con el que cuentan los especialistas, no es el original, sino que ha sido material que el propio personal ha preparado.

Los manuales de aplicación e interpretación son en su mayoría copias de la traducción original (1981).

En relación a la DGEE se informa que la prueba WPPSI, es un instrumento que no es utilizado, debido a que la población que les es remitida corresponde en su mayoría al nivel escolar primario.

Cuando es necesario aplicar un instrumento para preescolares, utilizan la escala de Binet-Merrill (1975).

## **COMO INSTRUMENTO DE INVESTIGACIÓN. WISC.**

En 1973, el Dr. Rogelio Díaz Guerrero, dirigió el INCCAPAC<sup>(1)</sup> el cual ha enfocado el eje de sus investigaciones al estudio de la Psicometría en México, y en el que se desarrolló la IDPEM<sup>(2)</sup> (Morales, 1975:201) fundada originalmente en Austin, Texas y la cual al aplicarse en México se pretendía: Entrenar psicólogos en el campo de la investigación y obtener la estandarización y normas de calificación de diversos instrumentos de medición psicológica que cubrieran los requerimientos del medio mexicano.

La muestra utilizada para el desarrollo de la IDPEM se constituyó por 450 escolares en total, los cuales pertenecían al sistema escolar tanto privado como oficial del D.F., dividiéndose en tres grupos: Sistema escolar "Centro" (100 niños), Sistema escolar "Privado" (200 niños), sistema escolar "Gubernamental" Unidad Independencia (150 niños). Se conformaron tres niveles por edad y grado escolar: Niños de 6 años (primer

le primaria); niños de 9 años (cuarto año de primaria) y niños de 12 años (primero de secundaria).

En esta forma el diseño experimental cubriría doce años de investigación a través de 6 años de estudio. Ya que la muestra fue superpuesta "los sujetos que se encontraban en el primer año de primaria al iniciarse la investigación, estaban ya en cuarto grado al cuarto año de aplicación, obteniéndose de este modo dos muestras comparables de sujetos en el primer grado y con la misma edad, (...), el grupo II de niños fue investigado inicialmente cuando se encontraban en cuarto año de primaria y al llegar al primer grado de secundaria pertenecía a la muestra de aquellos niños que se hallaban en este grado al iniciarse la investigación, y así sucesivamente." (Morales, 1975:202)

La Bateria aplicada en México en 1965, se realizaba en dos sesiones y era la siguiente:

- . HIT- (Holtzman Inkblot Technique) Prueba de percepción y personalidad.
- . WISC- Escala de Inteligencia.
- . Prueba de Fraccionamiento Visual- Explora el aprendizaje transferencial.
- . Prueba de Clasificación de Objetos- (Sortig Test) Indaga aspectos de formación de conceptos y su evolución con el crecimiento.
- . Test de Witkin- Define la agudeza perceptiva visual.
- . Prueba del Dibujo de la Figura Humana- (Goodenough) Mide la Inteligencia.
- . Test de Estilo Conceptual- Evalúa la forma en que el sujeto categoriza dos eventos.
- . Estimaciones de Tiempo- Se tiene que hacer la estimación, de lo que dura un minuto, sin contar ni ver un reloj.
- . Estimación de Conducta- Son observaciones realizadas al término de la aplicación de la Bateria, en la cual se sintetiza objetivamente el comportamiento del sujeto.

Para la estandarización del WISC en México a través de la IDPEM, los pasos que se siguieron fueron los siguientes:

- . Obtención de una muestra extensa y representativa de la población (escolares urbanos estables; de ambos sexos; de tres niveles de sistemas escolares; y tres niveles económicos, 450 niños en total).
- Era requisito indispensable que durante los seis años de la investigación los niños

permanecieran en la Cd. de México, aunque no vivieran en el mismo domicilio ni fueran a la misma escuela.

La traducción de la prueba realizada en Puerto Rico, se corrigió y adaptó a la cultura mexicana, tratando de no alterar el contenido ni la dificultad de los reactivos.

Al administrarse, se siguieron las instrucciones del Manual, incluyendo modificaciones para México en cuanto a sintaxis.

Se analizó la dificultad de los reactivos, para establecer el orden pertinente para México, encontrando "diferencias importantes entre la escala original (...), la traducción de Puerto Rico y la muestra para México." (Morales,1975:123)

Se llevó a cabo un análisis por edades y se transformaron las puntuaciones crudas a CIs, en función del porcentaje de respuestas.

a muestra empleada fué la correspondiente al primer año de aplicación de ésta investigación, no se utilizó la misma muestra en los siguientes años debido a la influencia en la reaplicación; por lo que se obtuvieron muestras representativas de las variables que definían la muestra original de los años subsecuentes ( 2do, 3o, 5o y 6o de primaria, 1o y 3o de secundaria).

La confiabilidad se estableció por medio del test-retest, obteniéndose "un alto coeficiente de confiabilidad temporal significativa al 0.01 y al 0.05." (Morales, 1975:123). En la primera aplicación se comprobó la validez de constructo del WISC, mediante el análisis factorial; además se realizó un análisis factorial por edades, y un análisis de varianza ( ANOVAS ) para poder probar el poder discriminativo por edad de subtest, diferenciación por sexo y por nivel socioeconómico.

Al terminar el período de investigación de la IDPEM, en el INCCAPAC pretendían observar y registrar lo observado: la evolución de la personalidad de los escolares desde el primero de primaria hasta el tercero de secundaria; así como las diferencias motivadas por la clase socioeconómica y por el impacto cultural. Todo ello según la comparación estadística de la muestra con la muestra de Estados Unidos (Austin, Texas).

Con el inicio y desarrollo de la IDPEM, fueron realizadas (abordando diferentes temas y puntos específicos) diversas tesis y trabajos de investigación por medio de los cuales se conocen los resultados obtenidos por el INCCAPAC. Estas investigaciones

1 Díaz Guerrero "se han enseñando en las prácticas y clases de psicometría en la UAM, su difusión entre los investigadores y profesionistas de la psicología en México es relativamente pequeña.." (Valderrama, 1994:121)

Por otro lado entre las investigaciones y trabajos de tesis (Licenciatura y Posgrado) realizadas en distintas Universidades de México acerca del WISC; se encontró que en algunas de ellas utilizaron la prueba sin ninguna modificación observando solo los resultados obtenidos por los niños, por ejemplo:

En 1969 Fuchs Modiano realizó una investigación para conocer la relación que existe entre el Test Fuchs-Lara (el cual mide la frustración y agresividad); el Coeficiente Intelectual actual y el grado de C.I. de los niños (por medio del WISC).

En 1971 Torres Lucero, realizó una validación cruzada entre el WISC y el Raven, en un grupo de 50 menores infractores; apoyándose en los trabajos de la IDPEM.

En 1980 Fernández Banueles, realizó una correlación de la escala ejecutiva del WISC con los test de Figuras de Torrance "pensando creativamente con dibujos", en 100 sujetos de 7 a 9 años de edad, alumnos de escuelas oficiales del D.F.

En 1965, Ahumada Rodríguez realizó una investigación con el objetivo de conocer cómo sucede cuando dos pruebas mentales (WISC y Goodenough), construidos y validados fuera de una sociocultura son introducidos en ella y se comparan entre sí.

Estudio 68 casos de la muestra de la IDPEM, entre los 6 y 12 años de edad; rechazó a los niños que en el WISC, no alcanzaron un C.I. de 80; o por registrar en el HIT síntomas graves de personalidad.

Los resultados que Ahumada obtuvo fueron los siguientes:

La media de CIs que obtuvo en el WISC fué: Ejecución 101.86; Verbal 97.01; Global 100.0. Las correlaciones entre el Goodenough y el WISC fluctuaron entre .51 y .84.

En el Goodenough, obtuvo correlaciones bajas con los CIs Global, Verbal y de Ejecución del WISC; Ahumada no lo consideró como una buena estimación de la Inteligencia Global, observándose además, mayor validez del instrumento para niñas que para los niños.

Obtuvo en el WISC, mejores calificaciones en la escala de ejecución que en la verbal; media de C.I. global, aumentó hasta 10 unidades con la edad. Por tanto, consideró que México deben existir normas de calificación para cada edad.

La intervalidez entre el WISC y el Goodenough, disminuyó al aumentar la edad. La verbal del WISC, tuvo menor correlación con el Goodenough que la de ejecución. subpruebas de Comprensión y Retención de Dígitos carecieron de factores comunes el Goodenough.

En ese mismo año Reyes Lagunes, realizó una investigación en la que pretende car los factores que intervienen en la personalidad infantil y su desarrollo, ligado a niños escolares.

Utilizó la Escala de Goodenough y el WISC (administrado sin omitir ninguna prueba), los aplicadores, tuvieron que igualar hasta donde fué posible la forma de aplicación, para "obtener acuerdo con los psicólogos encargados de la muestra americana e intercambiar experiencias respecto a la aplicación y calificación." (Reyes, 1965:34)

Utilizó 68 niños de la muestra de la IDPEM, obtuvo los siguientes resultados:

Los factores socioculturales, modificaron de una u otra forma los resultados de las pruebas, algunas situaciones presentadas en varios reactivos, fueron desconocidas al escolar mexicano.

Según los resultados estadísticos, Reyes obtuvo una Media de CI aproximadamente 100 y una Desviación Estándar de 15. Fueron escasos los cambios que realizó en el índice de dificultad de las subpruebas de la escala de ejecución, debido a que considera sufren menor grado de impacto sociocultural.

Existió tendencia de aumentar el rendimiento intelectual en relación al aumento de la edad, por lo que sugiere realizar una investigación con el fin de saber cuales son los factores que influyen, y poder llevar a cabo los ajustes necesarios a la Escala.

En 1965 Jarquin Fagoaga, realizó un estudio exploratorio para conocer el funcionamiento del WISC en un grupo de niños mexicanos, con el propósito de sentar las bases para una futura estandarización.

Constituyó la muestra por 198 escolares de la Cd. de México, entre los 5 y 15 años de edad; 50% de cada sexo, no repetidores escolares, de Jardín de Niños, primaria y

idaria con diferente rendimiento escolar y diferente nivel sociocultural. Utilizó 2 años particulares y 7 escuelas mixtas oficiales.

Jarquín tradujo el WISC en forma experimental y con autorización, para su aplicación y estandarización, conservando su contenido y modificando solo preguntas conforme a la cultura mexicana.

Aplicó individualmente las 10 subpruebas básicas en su totalidad y en el orden normal. Realizó la calificación conforme al Manual del WISC.

Los resultados obtenidos por Jarquín fueron los siguientes: En las Escalas, la distribución fue normal, el grupo obtuvo un rango entre 70 y 142 puntos en la Escala con una Media de 75 y una desviación de 91. Encontró correlación entre el orden normal de los reactivos del WISC y el obtenido en su investigación, lo que apoyó la habilidad de la prueba para su aplicación en México; influyeron en la variación del rendimiento intelectual, el nivel sociocultural, sexo, grado escolar y edad.

En 1966 Hurtado de Mendoza, consideró que en México la mayoría de los sujetos examinados con el WISC, obtenían puntajes bajos, clasificados como "deficientes" según las normas del autor. Por tanto, "creyó necesario examinar detenidamente la dificultad de los reactivos del subtest de Comprensión presentan para los escolares de la Cd. de Veracruz, analizando las principales variables que puedan intervenir en la misma." (Hurtado, 1966:19)

Seleccionó la muestra según la relación que debe existir entre edad cronológica y grado escolar, la dividió en tres grupos: Normales Pedagógicos, Retrasados y Retardados.

Utilizó 1,481 casos (833 hombres y 648 mujeres), de 20 escuelas (primarias y secundarias, tanto públicas como privadas). Clasificó la población en 3 niveles, según su condición económica: 25% bajo, 45% medio y 30% alto.

Efectuó las aplicaciones en un año, a los niños de 1o de primaria, de manera individual, y a los grupos restantes, colectivamente. Utilizó reactivos y valoraciones de la prueba, que en Puerto Rico se había hecho de la prueba.

Hurtado obtuvo los siguientes resultados:

Edad Cronológica: encontró diferencias en los puntajes obtenidos desde los 8 años, aumentando gradualmente hasta los 15 años; los puntajes fueron inferiores a los de Wechsler, Hurtado lo atribuyó a la administración colectiva.

**Grado escolar:** se detectaron incrementos debido a la escolaridad. El grado de dificultad de los reactivos lo obtuvo a través de los porcentajes de las respuestas correctas.

**Ocupación de los padres:** los factores socioeconómicos, y el nivel ocupacional participaron significativamente en la obtención de puntajes.

**Nivel socioeconómico:** la influencia del medio socioeconómico fué decisiva, ya que los niveles más altos obtuvieron mayores puntajes.

**Dificultad de los reactivos:** a partir del reactivo 8, obtuvo variaciones considerables. Propone que el orden de los reactivos para niños mexicanos sea diferente:

Wechsler - 1 2 3 4 5 6 7 8 9 10 11 12 13 14

Jalapa - 1 2 3 4 5 6 7 11 14 8 12 10 9 13

En 1969 Ahumada Rodríguez, realizó un segundo estudio donde analizó el plejo Intelectual de estudiantes mexicanos; utilizó 3 pruebas de inteligencia infantil: VISC, la Prueba de la Figura Humana de Goodenough y la de Harris. Los objetivos planteó fueron: Confrontación de las pruebas entre si, para conocer hasta que grado en lo mismo y contribuir al entendimiento del individuo en constante proceso de bio. La muestra que utilizó fué la de la IDPEM.

legó a las siguientes conclusiones:

**Análisis factorial:** en todas las edades obtuvo una buena consistencia interna y un buen grado de validez de construcción empírica. En las escala verbal del WISC, obtuvo una media de CI de 94.65 a 104.10, y en la escala de ejecución, de 99.10 a 112.55.

**Curvas de desarrollo de las subpruebas y las escalas:** observó un consistente ascenso en las subpruebas de información, comprensión y semejanzas.

**Curvas de desarrollo de los cocientes intelectuales :** los 3 cocientes intelectuales del WISC , tendieron a incrementarse con el aumento de edad .

**Curvas por edad y sexo:** los resultados favorecieron generalmente a los varones, aunque las mujeres rindieron mejor en la prueba Símbolos en Dígitos, por tanto se consideró que se debió a la obediencia de las instrucciones.

En 1969, Ehrlich Quintero Nelva, realizó una investigación acerca de la influencia del sexo y el grado de escolaridad en niños de un estrato socioeconómico bajo. Ehrlich comparó los resultados obtenidos, con 147 casos estudiados por la IDPEM; utilizó el manual del WISC, revisado por la IDPEM y, el cuestionario del Estudio Demográfico Escolar Mexicano.

Formó la muestra con 64 niños (ambos sexos) de 6 años de edad, que asistían a los cursos de Orientación Nutricional (CON) del Instituto Nacional de Pediatría (INP), de diferentes colonias proletarias del D.F.

Los resultados obtenidos fueron los siguientes: No encontró diferencias significativas de que la escolaridad nula o el jardín de Niños influyan en el rendimiento actual; ni tampoco con respecto al sexo, lo cual contrasta con los resultados de la IDPEM. Al igual que ésta, obtuvo un mejor rendimiento en los varones que en las niñas.

En 1986, Flores Luna realizó una correlación entre el WISC y el ITPA (Prueba de Integridad Intraindividual de las funciones psicológicas y psicolingüísticas involucradas en la comunicación), con el fin de conocer la relación entre la edad mental y la edad lingüística; contar con un instrumento válido y confiable para el diagnóstico de las funciones psicolingüísticas en población mexicana con problemas de aprendizaje.

Formó la muestra con 60 niños (ambos sexos) de 6 a 13 años (10 de cada edad), que iban a la primaria. Los resultados fueron: existieron diferencias significativas entre la edad mental y la edad psicolingüística, aunque se encuentran relacionadas y, obtuvo puntajes más elevados a mayor edad, escolaridad y desarrollo biológico.

En 1983, la Dra. Margarita Gómez Palacio, en colaboración con Eligio R. Padilla y Daniel Roll, llevaron a cabo la Adaptación y Estandarización del WISC-R para México (WISC-RM) a través del Departamento de Investigación, dependiente de la Dirección Técnica de la Dirección General de Educación Especial, a cargo de la Profa. Marta Arraiza. Esto en colaboración con la Universidad de Nuevo México, para formar parte de la Bateria de Pruebas SOMPA (1982) y con el fin de aportar instrumentos y criterios de calificación adaptados a la cultura mexicana.

El WISC-R y el WISC-RM están contruidos por las mismas 12 subescalas del WISC-R. De las 10 subescalas se consideran básicas. Las subescalas Retención de Dígitos y Símbolos, no se utilizan para calcular el CI y se consideran complementarias. (WISC-RM, Manual, 1984:5).

En el WISC-RM, se procuró conservar en todo lo posible los elementos del WISC-R 4) tanto en su uso como en aceptación.

Entre los cambios principales que se realizaron se encuentran los siguientes:

Valor de edad cubierta: WISC 5 a 15 años

WISC-R y WISC-RM 6 a 17 años.

Se cambió el contenido de algunos reactivos.

Se modificó la clasificación y el orden de algunos reactivos en las 2 subescalas.

Se alternó la aplicación de las subescalas verbales y de ejecución.

Se revisaron las instrucciones de aplicación para eliminar posibles ambigüedades.

Las instrucciones de calificación de las subpruebas de Semejanzas, Vocabulario y Comprensión, se ampliaron, incluyéndose mayor variabilidad de respuestas.

Si el niño fracasa en el primer reactivo de cualquier subprueba, el examinador debe demostrar o dar la respuesta correcta.

Para la estandarización del WISC-RM, la muestra utilizada fué de 1,100 niños de primarias y secundarias del D.F., fué la misma que se empleó para la estandarización del IPA en México, en 1982, debido a que formaba parte de esa Batería.

Debe señalarse que el WISC-RM como prueba que mide la inteligencia de un sujeto, se puede utilizar tanto en forma independiente, utilizando las Normas del Manual (1984), como en forma conjunta, como lo establecen las Normas empleadas en las Escalas Culturales de la Batería de pruebas SOMPA.

Los cambios en las subpruebas del WISC-R, tanto en contenido como en aplicación y calificación para la estandarización, fueron los siguientes:

Subescalas Verbales:

Información. Con respecto al contenido, fueron eliminados 9 reactivos, sustituyéndose por 7 nuevos; se cambió el orden de los reactivos. Para la aplicación, los niños entre los 11 y 12 años de edad, empiezan con el reactivo 5 y los de 13 años con el reactivo 11.

- . Semejanzas. En el contenido se cambió el orden de los reactivos. En la aplicación y calificación no hubo ningún cambio.
- . Aritmética. En el contenido, se cambió el orden de los reactivos 13 y 14. En la aplicación y calificación no hubo cambios.
- . Vocabulario. En cuanto a contenido, se eliminaron 7 reactivos y se sustituyeron por 6; además se cambió el orden de éstos. Para la calificación, todos los reactivos reciben una puntuación de 2, 1 o 0.
- . Comprensión. En contenido, se eliminaron 5 reactivos sustituyéndose por otros 5; también se cambió el orden. No hubo cambios en aplicación o calificación.
- . Retención de Dígitos (complementaria). No hubo cambios en cuanto a contenido. En la aplicación, se incluyeron 7 reactivos, tanto en las series de orden progresivo como en las de orden inverso. Para la calificación, el puntaje máximo es de 28 puntos, se debe tomar en cuenta ésta diferencia al interpretar las normas.

#### Subescalas de Ejecución:

- . Figuras Incompletas. En el contenido se eliminaron 3 reactivos, incluyendo el peine, la zorra y la sombrilla para mejorar la confiabilidad, quedando 23 reactivos. Se cambió el orden de estos. En cuanto a aplicación y calificación no hubo cambios.
- . Ordenación de Dibujos. En contenido, los niños de 6 a 8 años de edad, empiezan con el reactivo 1 y los de 9 a 16 años con el reactivo 2. No hubo cambios en aplicación y calificación.
- . Diseño con cubos; No hubo cambios en el contenido. En la aplicación, los niños entre 6 y 9 años, empiezan con el reactivo 1 y los de 10 a 16 años con el 2.
- . Composición de Objetos. No hubo ningún cambio.
- . Claves. No hubo ningún cambio.
- . Laberintos. (complementaria) No hubo ningún cambio.

Los coeficientes de confiabilidad de las Escalas se obtuvieron mediante la fórmula de Kuder-Richardson Brown (Verbal .90, Ejecución .89, Total .94); también las subescalas son

factorias ya que cuentan con los siguientes coeficientes promedio: Verbales entre .81; Ejecución entre .66 y .82; solo 10 coeficientes se encuentran abajo de .60.

La consistencia interna, se midió por el procedimiento de división por mitades; y la relación entre los cocientes intelectuales y las subescalas se obtuvo según la fórmula Pearson (media de 100 y desviación estándar de 3). (WISC-RM, 1984:4)

Los errores estándar de medida (EEM) "indican el valor límite de error que circunda el puntaje de una subescala (...) Los EEM de 1.14 para aritmética a una edad de 6 1/2 años, indican una probabilidad de alrededor de 2-3 de que los niños obtengan un puntaje normalizado dentro de 1.14 puntos como un puntaje "correcto"; aun más, la posibilidad de alrededor de 19-20 de que su puntaje normalizado se encuentre dentro de 2.28" (WISC-RM Manual, 1984:14).

Un puntaje correcto, es el puntaje promedio obtenido por un niño examinado en algunas ocasiones; entre más pequeño sea el EEM, menos se desconfía del puntaje; pero se deben tener presentes las variaciones de los EEM de una subescala a otra, dentro de un mismo grupo de edad o de una edad a otra en una misma subescala.

Por ejemplo, "a la edad de 11 1/2, las variaciones de EEM van de .84 en Diseño con formas a 1.88 en Comprensión. Para vocabulario, los EEM varían de 1.00 a la edad de 16 años a 1.81 a la edad de 6 1/2." (WISC-RM Manual, 1984:14)

En cada una de las subescalas, la distribución de puntajes naturales en cada nivel de dificultad se obtuvo una media de 10 y una desviación estándar de 3. En cada subescala fue determinada la progresión del puntaje normalizado de un nivel a otro; los cuales van de 1 a 3 desviaciones estándar a cada lado de la media. Al igual que en el WPPSI y en el WISC-R, en el WISC-RM fueron utilizados esos valores.

Para construirse los cuadros de CI, se obtuvieron las sumas de los puntajes normalizados de las Escalas Verbal, Ejecución y Total. En cada grupo de edad se calculó la media y la desviación estándar de las 3 sumas de los puntajes normalizados para los grupos de edad combinados y, se tomaron como base para construir los equivalentes de

En los 3 CIs de las escalas, la media y la desviación estándar de la suma de su puntaje normalizado se igualó a 100 y 15 respectivamente. En " las escalas verbal y de ejecución, los CI se extienden a 3 1/2 desviaciones estándar a ambos lados de la media (...). Para la suma total, los CI se extienden a 4 desviaciones estándar a cada lado de la media (...) y los sujetos de la muestra de estandarización estuvieron dentro de esos límites." (WISC-RM Manual, 1984:11)

Un niño de cualquier edad, no puede obtener un puntaje normalizado total menor a 5, al puntaje natural de 0 se le asigna 1 punto normalizado en cada subescala. Si no se tiene un puntaje natural por encima de 0, por lo menos en 3 subescalas verbales y en 3 ejecución, no se debe calcular ningún CI, ni parcial ni total.

Para que sea significativa la diferencia entre puntajes normalizados de 2 subescalas verbales, dependerá de cuales sean las que se comparen y de su confiabilidad.

Por ejemplo, a los 6 1/2 años al compararse los puntajes normalizados de Semejanzas Verbales (confiabilidad .86) y Laberintos (confiabilidad .62) es necesaria una diferencia mayor para ser significativa que si se comparan Aritmética (confiabilidad .86) y, Laberintos (confiabilidad .86) ya que las primeras tienen una confiabilidad menor.

El nivel de confianza de 15% es el mejor para hacer comparaciones entre subescalas, existe el 85% de posibilidades de que la diferencia sea suficientemente grande para ser significativa. Una diferencia de 1 o 2 puntos, puede deberse al azar.

Para que una diferencia se considere significativa en cada grupo de edad entre CI Verbal y de Ejecución al nivel 15, deberá ser de 10 puntos; y al 5% de confianza de 14 años. (WISC-RM, Manual, 1984:17).

En 1984 a partir de la estandarización del WISC-RM, Gómez Palacio realizó una investigación con el propósito de conocer si en las Escuelas Primarias Oficiales existían niños con un potencial intelectual muy superior.

Constituyó la muestra por 331 niños de 6 a 9 años de edad, de ambos sexos, inscritos en 1o y 3er grado de primaria, los cuales obtuvieron calificación de 10 en el período escolar anterior, y además sus maestros opinaron que eran muy inteligentes.

Utilizó como instrumentos al WISC-RM (SEP, 1982) y al cuestionario demográfico y profesional para maestros de educación básica, elaborado por Reyes Lagunes. Los resultados obtenidos fueron los siguientes:

El que un niño haya obtenido un promedio de 10 en un ciclo escolar, no garantiza que él sea un niño con un potencial superior; el WISC-RM permitió detectar a los niños con un potencial muy superior. Las características de los niños fueron las siguientes: 15 niños de 1er grado de primaria y 15 niños de 3er; 20 niños del sexo masculino y 10 del femenino, los cuales mostraron superioridad en el desarrollo de su capacidad de abstracción y, se consideraron adaptados socialmente. Además de que "están sujetos en la escuela a los mismos estímulos y recompensas que otros niños de inteligencia promedio" (Gómez Palacio, 1984:74)

## **COMO INSTRUMENTO DE PSICODIAGNOSTICO**

Los Lineamientos para la Evaluación Psicológica, de la DGEE (1884), establecen que el WISC-R debe ser utilizado conforme a la estandarización hecha en México en 1983. (WISC-RM)

El WISC-RM es aplicado a niños de 6 a 17 años de edad. Por medio de él se realiza un análisis cuantitativo para obtener el Coeficiente Intelectual y un análisis cualitativo de la interpretación de Kaufman.

Con los niños que cuentan con la edad requerida, pero que por algún motivo no se encuentran capacitados para responder satisfactoriamente en el WISC-RM, se utiliza la Escala de Terman-Merrill (1960).

Cuando un niño en el WISC-RM obtiene puntajes de CI entre 55 y 70, pero se sospecha que cuenta con mayor potencial de aprendizaje, se sugiere interpretar la prueba conforme a las normas de la Batería SOMPA (1982).

En las entrevistas realizadas al personal directivo de los Centros dependientes de la SED, se señaló que tanto el WISC-R como el WISC-RM, son consideradas buenas pruebas para medir la inteligencia; siendo ambas utilizadas indistintamente, ya que no todos los Centros cuentan con los Manuales y el material necesarios.

## **BATERIA DE EVALUACIÓN INTELECTUAL SECUENCIAL-SIMULTANEA KAUFMAN. MÉXICO, D.F.**

### **COMO INSTRUMENTO DE INVESTIGACIÓN**

La Bateria de Evaluación Secuencial-Simultánea Kaufman, se estandarizó y adaptó a partir de los estudios realizados en el período 1981-82, publicada en 1985 por el Departamento de Investigación de la Dirección General de Educación Especial.

Anteriormente no se tiene conocimiento de que ésta Bateria haya sido utilizada en algunas Instituciones. Las razones por las cuales se eligió la estandarización de este instrumento (Gómez Palacio, Rangel y Padilla, 1985-87) fueron:

1. Separa el conocimiento adquirido en la experiencia de los niños y la habilidad para resolver problemas poco familiares.

2. Proporciona resultados que permiten planear intervenciones educativas específicas.

3. Por no contar, hasta entonces, con una prueba estandarizada en México para la medición psicológica de niños en edad preescolar. (3 a 5 años)

4. Por ser una prueba fácil de aplicar y contar con criterios objetivos de calificación.

Después de revisar la Bateria Original, los investigadores (Gómez Palacio, Rangel y Padilla. Informe sobre teoría y resultados;1986) decidieron llevar el proyecto de estandarización en 6 fases:

1. Preparación del material en donde se decidió: eliminar una de las subpruebas por no corresponder al contexto social de México; perfeccionar la traducción de las instrucciones para su administración y, respetar los lineamientos establecidos por los autores en cuanto a las instrucciones de aplicación.

2. Capacitación de los investigadores de campo.

3. Selección de la muestra: La Bateria fué originalmente diseñada para ser aplicada a niños de 2.6 a 12.6 años, por tanto se decidió abordar tres tipos de población de niños residentes en el D.F. Comprendía a: niños inscritos en guarderías de mercados de cada una de las delegaciones políticas y niños inscritos en escuelas primarias oficiales, lo cual arrojó un total de 1,100 niños en la muestra.

4. Se trabajó en un total de 40 guarderías y 70 primarias, se aplicó la Prueba Visual de PEABODY con el propósito de validación.

- . Trabajo de campo. Efectuado por 10 psicólogos experimentados.
  - . Análisis estadístico. En donde se contó con la colaboración de American Guidance Service, en el Estado de Minnesota, Estados Unidos, supervisado directamente por el Dr. Alan S. Kaufman.
- Este análisis tuvo como propósitos fundamentales:
- . Depurar la Batería.
  - . Obtener normas de calificación adecuadas a la población infantil mexicana.
- . Reporte de resultados.

Como consecuencia del análisis de los datos de la muestra, la versión mexicana de la batería Kaufman contiene las siguientes modificaciones:

A continuación se presenta un cuadro general con las modificaciones que se hicieron a una de las subpruebas de la Batería original del Kaufman. (Gómez Palacio, y Padilla, Informes sobre teoría y resultados, DGEE:1986)

SUBPRUEBA	REACTIVOS ORIGINALES	REACTIVOS ELIMINADOS	REACTIVOS DE LA V. MEX.
Ventana Mágica	15	2	13
Conocimiento de Caras	18	3	15
Movimiento de manos	26	5	21
Cierre Gestalt	28	8	20
Memoria de Números	19	4	15
Triángulos	21	3	18
Denamamiento de Palabras	23	3	20
Analogías de Matriz	28	8	20
Memoria Espacial	24	4	20
Serie de Fotografías	22	5	16
Vocabulario	27	3	24
Aritmética	47	11	36
Adivinanzas	43	10	33
Lectura *			
1) Descodificación	51	20	31
2) Comprensión	35	10	25

\* La subprueba de Lectura es en la versión mexicana, una sola subprueba compuesta por 5 reactivos.

revisión estudio inicial se podría aplicar en general a niños de 3 a 12.6 años, tomando en consideración que las normas de calificación son el producto del análisis de datos de los que acuden a las instituciones anteriormente citadas.

En 1985, en México se realizó una investigación por parte de Gómez Palacio, Hinojosa y García Cedillo acerca de la validación de constructo y utilidad en el diagnóstico diferencial en problemas de niños inscritos en programas de educación especial. (DGEE, 1988)

El objetivo de llevar a cabo ésta investigación fué el de proveer al psicólogo clínico, de instrumentos de evaluación estandarizados y confiables, determinando el nivel de validez de constructo del K-ABC, así como buscar indicadores que pudieran ser útiles en el diagnóstico diferencial de niños que requieren de educación especial.

La muestra para llevar a cabo éste estudio fué de 746 niños cuyo rango de edad osciló entre 3 y 11 años, se trabajó con 4 muestras de niños inscritos en alguno de los programas de seguimiento de la DGEE: Niños con problemas de aprendizaje, deficiencia mental, con alteraciones de lenguaje y con trastornos de audición.

Se decidió trabajar con estos niños por dos razones: "primero, porque son ellos quienes se benefician de los programas que ofrece la DGEE, y segundo, porque la Batería Kaufman es válida en la evaluación de estos niños, precisamente" (Reporte de Investigación, DGEE, 1988:4).

A todos los sujetos que participaron se les aplicó un cuestionario socioeconómico.

Con base en lo anterior, se puede decir que no hay hasta la fecha trabajos relacionados con la Batería Kaufman y, a pesar del esfuerzo realizado por la DGEE el mencionado instrumento no ha logrado cumplir con uno de sus objetivos, que fué el de contar con una prueba estandarizada en México pues su utilización es excepcional en la evaluación de niños con problemas de aprendizaje.

## **LA BATERIA KAUFMAN COMO INSTRUMENTO DE PSICODIAGNOSTICO**

En los Centros de Orientación, Evaluación y Canalización (DGEE), la Batería de Evaluación Kaufman según el manual de procedimientos, forma parte de la Batería básica de pruebas de diagnóstico que se aplican a los niños canalizados al departamento de psicología.

Ésta Batería se aplica a niños desde 3 años hasta 12 años de edad, con el fin de obtener información sobre el estilo de procesamiento cognitivo en términos de

esamiento secuencial y procesamiento simultáneo que enriquecerán el estudio psicológico" (Lineamientos para la Evaluación Psicológica, DGEE:1984).

in embargo, en las entrevistas realizadas con el personal que labora en esa institución, así como con el personal directivo de uno de los Centros, se menciona que la Bateria Kaufman no es un instrumento que se utilice regularmente, a menos que la docente encargada del grupo de tratamiento la solicite, previo análisis del caso.

Las razones principales se refieren al tiempo prolongado de aplicación de la prueba, como a las dificultades para su interpretación.

En cuanto a los CAPEP, se informa que la Bateria Kaufman no forma parte de la batería de pruebas seleccionadas para el diagnóstico psicológico, a pesar de que parte del personal de psicología recibió la capacitación para aplicarla.

## **SISTEMA DE EVALUACIÓN MULTICULTURAL Y PLURALÍSTICO. SOMPA.**

### **COMO INSTRUMENTO DE INVESTIGACIÓN**

En México se realizó la estandarización de la Batería de pruebas SOMPA, a través del Tratamiento de Investigación de la DGEE en 1982. La investigación quedó a cargo de la Dra. Margarita Gómez Palacio, aprobada por sus autores, Lewis y Mercer.

El Manual de Aplicación de la Batería, es la traducción de los Manuales originales, que trata los fundamentos teóricos e incluye los datos necesarios para su aplicación en el campo.

La elección en México de ésta Batería, para su estandarización, obedece a:

Posibilita una evaluación integral de sujeto a través de sus 3 modelos (Médico, Psicológico y Pluralístico), lo que implica tomar en cuenta la salud, conducta, medio ambiente y rendimiento escolar del niño.

Su Modelo Pluralístico desarrolla varios criterios de calificación, de acuerdo a los contextos socioculturales en los que el niño se encuentra.

Con ésta estandarización se aporta, según Gómez Palacio (1982) instrumentos adaptados a la cultura mexicana, proporcionando criterios de calificación que permitan evaluar al sujeto en estudio, con criterios derivados de un grupo de individuos representantes de su mismo contexto sociocultural.

Esta estandarización se basó también en que:

Cada vez que se hace el diagnóstico de los niños en cuanto a su rendimiento intelectual, se da una etiquetación por carecer de normas especiales para ello. Es decir, no consideran sus orígenes culturales, oportunidades de aprendizaje, medio socioeconómico, historia médica, etc..

Los psicólogos mexicanos evalúan al niño con pruebas que, en su mayoría, son versiones literales de diseños extranjeros. Además los resultados obtenidos por los niños en México son comparados con los obtenidos por niños de otro contexto sociocultural.

El primer paso para la estandarización de la Batería fué que en junio de 1980, se hicieron listas de las Escuelas Oficiales, Primarias y Secundarias (generales y especiales), en el D.F. (ciclo escolar 1980-1981).

En julio de 1980, la Dirección General de Estadística de la SEP, indicó que la matrícula escolar era aproximadamente de 1,834,238 estudiantes inscritos en 2046 primarias y 535 secundarias.

De éstas listas se eligieron al azar 50 primarias y 50 secundarias (turnos matutino y vespertino), de las cuales se extrajo un promedio de 11 sujetos por plantel, dando un total de 1100; el rango de edad fluctuaba entre los 6 y 16 años; 550 varones y 550 mujeres.

Una vez obtenida la autorización en los planteles, se inició el trabajo de campo (septiembre de 1980). Éste trabajo se realizó en dos etapas:

Primero, una vez seleccionados los sujetos, el paso siguiente fué ponerse en contacto con los padres de la familia a los que se les aplicaron las Escalas Socioculturales.

Después de aplicadas las Escalas se administró el WISC-R; únicamente a los alumnos de 6 años a 6 meses se les aplicó el Bender.

En la segunda etapa de investigación; el procedimiento fué el mismo que el de la primera etapa, únicamente la muestra de alumnos varió de 10 a 15 sujetos, el rango de edad fué de 6 a 12 años y se incluyó el Inventario de la Conducta Adaptativa, a todos los sujetos se les aplicó el WISC-R y el Bender.

En las dos etapas de investigación, se pidió a los maestros su cooperación para responder a un cuestionario que permitiera conocer la evaluación de los profesores hacia sus alumnos en cuanto a rasgos de personalidad. Esto permitió hacer cálculos de validez de los instrumentos implicados. Así como también proporcionar las calificaciones de los alumnos en el trimestre correspondiente a la fecha de aplicación de las pruebas.

Los instrumentos que se estandarizaron en la Batería fueron:

Escala de Inteligencia Wechsler para niños. WISC-R.

Prueba Gestáltico Visomotora Bender.

Escalas Socioculturales.

Inventarios de Conducta Adaptativa. ABIC.

## **. Escala de Inteligencia Wechsler para niños. WISC-R:**

Para la estandarización del WISC-R dentro del SOMPA, fué empleada la misma prueba utilizada para la estandarización de toda la Batería. Además de que los resultados obtenidos se correlacionaron con los del Bender.

La confiabilidad de la prueba se obtuvo mediante la fórmula de Spearman Brown; siendo en la Escala Verbal un coeficiente de confiabilidad de .90, en la Escala de Ejecución de .89 y en la Escala Total de .94. Los coeficientes promedio de las subescalas varían entre .79 y .89.

La correlación entre los CI y la subescalas se obtuvo por la fórmula de Pearson, dando puntajes en la Escala Verbal de .26 a .83, en la Escala de Ejecución de .33 a .76, en la Escala Total de .37 a .76. Para la Escala Total, se obtuvo una media de 10 y una desviación estándar de 3.

El WISC-R es uno de los instrumentos que se estandarizó dentro de la Batería de pruebas SOMPA, dando lugar al WISC-RM que es utilizado dentro del Modelo Social y Pluralístico; conforme a las normas establecidas dentro de la Batería SOMPA así como la interpretación cualitativa de Kaufman la cual pretende centrarse en la persona y aprender porque los chicos obtuvieron una puntuación en particular y no en destacar la habilidad de su ejecución." (Kaufman, 1982:3).

El WISC-RM se aplica dentro del SOMPA cuando un sujeto arroja puntajes de CI de 55 y 70 y se sospecha que cuenta con un mayor Potencial de Aprendizaje.

### **Modelo Social:**

El WISC-RM se utiliza dentro de éste modelo como una medida de logro y de aprendizaje, detectándose además que tanto ha aprendido el niño de la cultura de la edad donde vive. También por medio de él se puede observar el grado de adecuación y desviación del comportamiento del niño, según las normas del rol de estudiante de escuela pública.

### **Modelo Pluralístico:**

Dentro de éste modelo con el WISC-RM se predice el desempeño del sujeto como estudiante. Además se estima el Potencial de Aprendizaje (evaluando los puntajes por medio de normas específicas), utilizando las Escalas Socioculturales como base para la interpretación de la ejecución de un niño en el WISC-RM.

Al evaluar un sujeto con ésta prueba, en éste modelo, se realizan dos tipos de comparaciones:

Se compara su rendimiento con las Normas Generales de la población.

Se compara su rendimiento con el grupo de sus mismas características socioculturales.

#### Estimación del Potencial de Aprendizaje. (EPA)

Dentro del SOMPA, se considera que el WISC-RM, evalúa la conducta aprendida y el nivel de funcionamiento escolar, por lo que bajo ciertas condiciones, los CIs del WISC-RM sirven para Estimar el Potencial de Aprendizaje de un sujeto, realizándose la estimación como: "Perfil del Nivel de Funcionamiento Escolar".

Las condiciones que se tienen que tomar en cuenta para obtener el EPA son:

Oportunidades o experiencias similares, motivaciones y recompensas para aprender el material de la prueba.

Ausencia o presencia de problemas emocionales al momento de ser evaluado.

Si no se toman en cuenta las condiciones antes mencionadas y solo se basa la estimación en los CIs dados por el WISC-RM, no se podrá inferir el Potencial de Aprendizaje.

Cuando se interpreta la evaluación del WISC-RM dentro del SOMPA entre los dos niveles, se debe tener en cuenta cuando los puntajes caen en área de riesgo, ya que eso sugiere que el niño tal vez tiene o tendrá dificultades en los programas educativos escolares, necesitando así de Educación Especial.

#### Prueba Gestáltico Visomotora Bender.

La prueba Bender forma parte del modelo médico de evaluación. Su utilización tiene como propósito "evaluar la madurez perceptual del niño y su posible daño neurológico" (Investigación SOMPA, 1980:114)

El Bender fué aplicado a una muestra de 700 sujetos de ambos sexos cuyo rango de edad fué de 6.3 a 12.8 años de edad.

a distribución de la muestra fué por edad y sexo. La administración fué realizada por psicólogos en forma independiente. Se utilizó el Manual de Puntuación de Koppitz (1963) para la calificación, lo cual permitió elevar los índices de Confiabilidad.

través del procesamiento estadístico de los puntajes se obtuvieron normas de calificación para los estudiantes mexicanos en el D.F. inscritos a las escuelas públicas.

a puntuación media para cada nivel de edad se obtuvo de la comparación entre los puntajes medios de la población mexicana y los obtenidos por E. Koppitz con población americana.

se utilizaron dos ecuaciones de regresión para determinar las relaciones entre edad y puntaje así como entre edad y desviación estándar con los puntajes obtenidos en la prueba. (Investigación SOMPA, 1982:122).

Como resultado anterior, proporciona medidas estimativas para generar medias y desviaciones estándar predictivas.

En la evaluación del Bender dentro de la Batería SOMPA se obtienen puntajes crudos comparando los obtenidos en las nueve figuras. El puntaje crudo del Bender se convierte en puntaje escalar apropiado a la edad del niño para lo cual se consultan las tablas de correspondientes. (Investigación SOMPA, 1980:123)

Las Normas de calificación obtenidas en ésta investigación, así como la comparación con el Sistema propuesto por Koppitz se presentan a continuación:

Puntajes medios en la prueba Guestáltica Visomotora de L. Bender.

EDAD	KOPPITZ 1963	KOPPITZ 1974	MÉXICO 1981
5.0 a 5.5	13.2	13.7	—
5.6 a 5.11	10.2	9.7	—
6.0 a 6.5	8	8.6	7.7
6.6 a 6.11	6.4	7.2	7.2
7.0 a 7.5	5.1	5.8	6.7
7.6 a 7.11	4.2	4.6	6.5
8.0 a 8.5	3.4	4.2	6
8.6 a 8.11	2.7	3	5.5
9.0 a 9.5	2.2	2.8	5
9.6 a 9.11	1.8	2.3	4.6
10.0 a 10.5	1.5	1.9	4.2

10.6 a 10.11	1.2	-1.8	3.7
11.0 a 11.5	—	1.4	3.2
11.6 a 11.11	—	1.4	2.8
12.0 a 12.5	—	—	2.3
12.6 a 12.11	—	—	1.8

En la investigación realizada por la DGEE las normas obtenidas para el Bender son parte de la creación de Normas Pluralísticas, de tal manera que los puntajes obtenidos en la Batería se comparan únicamente con los puntajes obtenidos por el grupo al que pertenecen y con los que tengan una combinación de características socioculturales similares. (Investigación SOMPA, 1980: 8).

### **. Escalas Socioculturales.**

En México, las Escalas tienen la misma estructura que las de Estados Unidos, lo que modifica son palabras, por ejemplo: parentesco por relación, fuente de ingresos por actividad principal, jefe de casa por jefe de familia, etc..

Más que comparar la familia del niño con familias de un grupo étnico como lo hace el modelo original de la Escala, en la estandarización realizada en México se suele comparar las familias de determinado estatus socioeconómico. (Estandarización SOMPA, 1982: 79).

En la estandarización solo dos Escalas tienen significancia: Escala de tamaño de la familia y Escala de aculturación urbana. La muestra estuvo constituida por 1,100 padres de familia, de estas el 73.4% tuvo una organización familiar estructurada y el tamaño de familia era de 2 a 10 miembros aproximadamente.

En la Escala de Estrato Socioeconómico, "la media estadística nos indica que en la muestra, el promedio de las familias presentan características: el jefe de familia sostiene hogar por medio de un empleo asalariado. Además la familia no recibe ayuda económica adicional." (Estandarización SOMPA 1982:72).

En cuanto a los datos que se obtienen en la Escala de Aculturación Urbana, el puntaje más alto es en cuanto a la participación de los padres.

Referente a los datos de escolaridad se obtuvieron porcentajes en donde los padres tenían similares grados de escolaridad.

on respecto a la procedencia de los padres de familia se indica que el 51% procedían D.F.

Por último, en la utilización del idioma español, en la familia los datos reportaron que 70% de la muestra habla español.

### **Inventario de Conducta Adaptativa. ABIC.**

La muestra para la estandarización, estuvo constituida por 600 niños de 6 a 11 años de edad residentes del D.F.

Se realizaron 242 preguntas del formato original, previamente revisado en cuanto a validez y adaptación a la cultura mexicana. La adaptación de éste instrumento consistió en el cambio de fechas de festividades, fiestas nacionales, sustitución de nombres de localidades, por los nombres actuales de México. "Ésta adaptación fué mínima debido a que la traducción en español ya había sido efectuada para la estandarización del ABIC en Estados Unidos, con una población hispanoparlante" (Estandarización del ABIC, 1982:50).

Los procedimientos de calificación son los mismos que en el original. Los puntajes del ABIC ayudan a identificar el foco de la problemática. Así un bajo puntaje indica que el niño rinde menos de lo esperado para su edad cronológica y, si los puntajes escalares caen en la zona de riesgo, indican la necesidad de ayuda especial.

Después del estudio estadístico, el ABIC para México es de 238 preguntas que corresponden a los puntajes de los niños mexicanos.

Las calificaciones son la evaluación de la conducta del niño a través de la opinión de los padres en 6 "subsistemas sociales: la familia, comunidad donde habitan, relaciones con sus compañeros, la escuela, la economía y el cuidado propio" (Estandarización del ABIC, 1982:99).

Se cuenta con una escala de veracidad la cual permite detectar si la información reportada es errónea.

Se calcularon coeficientes de confiabilidad para cada nivel de edad por medio del procedimiento de división por mitades: ítems pares, ítems nones. Se utilizó la fórmula de Spearman Brown.

La confiabilidad para los puntajes crudos en las subescalas del ABIC es satisfactoria. La confiabilidad de los coeficientes está por encima de .80. Los coeficientes de confiabilidad son

simadamente los mismos para cada una de las subescalas". (Estandarización IPA, 1982:106)

Algunas de las diferencias estadísticas favorecen a las niñas, en la muestra de los es de edad de 6 a 11 años obtuvieron puntajes más elevados que los niños en 17 ítems. Los niños obtuvieron puntajes más elevados que las niñas en 10 reactivos.

La validez del ABIC, se realizó por medio del estudio de correlación entre los 5 ítems de la prueba. Concluyendo, que la intercorrelación entre la ejecución del rol y los sistemas sociales es alta, de .84 o superior. "Lo anterior justifica la combinación de los puntajes para cada Escala, en un puntaje total" (Estandarización SOMPA, 1982:117).

En México los Inventarios de la Historia de la Salud son 45 reactivos, al igual que en el formato original. Las preguntas se hacen a los padres, no se necesita convertir los ítems brutos en puntajes escalares. El perfil se realiza con los puntajes crudos.

Según se reporta en el Manual de Aplicación del SOMPA en México, que los inventarios son un medio para identificar a niños cuyas enfermedades, operaciones o accidentes pasados, pueden estar influyendo en su rendimiento académico actual, y en la capacidad para funcionar de manera adecuada en roles no académicos.

A lo largo de la estandarización de la Batería, el trabajo de campo "tuvo como escenario el plantel escolar al que asistía el niño, no obstante en ocasiones, se tuvo que ir al hogar para tener acceso a los sujetos" (Estandarización SOMPA, 1982:5)

Otros problemas a los que se enfrentó el equipo de psicología, fueron el encuentro del psicólogo con el director de la escuela, surgiendo confusión debido al desconocimiento de la investigación, o a la desconfianza que emanaba acerca de cual era realmente el rol del psicólogo. También hubo problemas con el mobiliario, condiciones generales de ambientación, luz, espacio, etc.

## **COMO INSTRUMENTO DE PSICODIAGNOSTICO**

En los CAPEP, la Batería de pruebas no forma parte del estudio para el diagnóstico.

El SOMPA es un instrumento normado por la DGEE que forma parte de la Batería de pruebas seleccionadas en el diagnóstico de los problemas de aprendizaje.

El psicólogo aplica SOMPA en los casos en que el WISC-RM "arroje puntajes de CI entre los 55 y 70 y se sospeche de un mayor Potencial de Aprendizaje" (Lineamientos metodológicos de la DGEE, 1986:27)

Aun embargo a pesar de que ésta batería ha sido estandarizada para la población mexicana, con toda una metodología y marco teórico estructurados, no es utilizada en

regular, ya que su tiempo de aplicación es aproximadamente de 5 horas, lo cual no considera operativo debido a la demanda de atención en ésta institución.

***CUADROS DE ANALISIS Y  
COMENTARIOS***

RUEBA: ESCALAS EVOLUTIVAS DEL DESARROLLO

UTOR: ARNOLD GESELL

UGAR: ESTADOS UNIDOS

ÑO: 1938

REVISIONES Y MODIFICACIONES	PSICODIAGNOSTICO	COMENTARIOS
<p>- Diagnóstico del Desarrollo Normal y Anormal del Niño. TRADUCCION: Bernardo Serebrisky.</p> <p>- Reimpresión de la Traducción de 1979. José Penhos 1985, México.</p> <p>ASPECTOS PSICOMETRICOS: No ha sido estandarizada en México.</p>	<p>En los CAPEP este instrumento actualmente no forma parte de la Batería de pruebas de Psicodiagnóstico. Se considera que es un instrumento para detectar problemas severos en el desarrollo, por tanto dejo de utilizarse.</p> <p>En las Escuelas de Educación Especial y Centros de Intervención Temprana de la DGEE, la prueba forma parte de la Batería que se utiliza, para valorar a los niños que presentan una problemática severa en el desarrollo, o en alguna de sus áreas.</p>	<p>- Se está empleando una prueba no adaptada a la población mexicana, es decir los coeficientes que se utilizan para evaluar el desarrollo de los niños mexicanos es inadecuado, el Diagnóstico que se obtiene es cuestionable.</p> <p>- La prueba fué diseñada en 1938 y a la fecha no se ha revisado, por tanto pensamos que es una prueba no actualizada para la evaluación del desarrollo infantil.</p> <p>- La versión que se utiliza es una Traducción Argentina, lo que nos lleva a cuestionar su utilidad.</p> <p>- El personal que utiliza la prueba no se pregunta sobre su efectividad, además de que los materiales son elaborados por ellos mismos, lo que redundo en su confiabilidad.</p> <p>- En las investigaciones realizadas en México, se menciona un trabajo de Estandarización que ha servido como apoyo para otras investigaciones, sin embargo el trabajo mencionado no fué posible localizarlo, lo que demuestra que no hay difusión en la investigación.</p>

RUEBA : GESTALTICO VISOMOTORA

UTOR: LAURETTA BENDER

UGAR: ESTADOS UNIDOS

ÑO: 1938

REVISIONES Y MODIFICACIONES	PSICODIAGNOSTICO	COMENTARIOS
<p>- Elizabeth Koppitz. 1964, Estados Unidos.</p> <p>- Elizabeth Koppitz. 1975, Estados Unidos.</p> <p>- "El Test Gestáltico Visomotor para Niños". E. Koppitz. TRADUCCION: María Julia García. 1980, Argentina.</p>	<p>En los Centros CAPEP y DGEE, el Bender es una prueba básica en el Psicodiagnóstico.</p> <p>Se aplica a niños desde 5 años en adelante.</p> <p>El Manual que utilizan en ambas Instituciones es la TRADUCCION Argentina de 1980. Sin embargo la Tabla para obtener las Normas de Calificación no es la propuesta en éste Manual.</p>	<p>- A pesar de que se utiliza invariablemente en el Diagnóstico, no parece cuestionarse el Sistema de Calificación, lo cual nos lleva a pensar que se desconoce gran parte del contenido de la prueba.</p> <p>- El Sistema de Calificación se altera y modifica sin que se de ninguna justificación.</p> <p>- Las investigaciones señalan la necesidad de revisar las Normas de Calificación ya que existen diferencias altamente significativas entre las poblaciones Norteamericana y Mexicana, sin embargo parece no tomarse en cuenta.</p>
<p>ASPECTOS PSICOMETRICOS: No ha sido estandarizada en México.</p>		

RUEBA: MONTERREY

UTOR: GOMEZ PALACIO Y COLS. DGEE.

UGAR: MONTERREY, NVO. LEON. MEXICO.

ÑO: 1974

REVISIONES Y MODIFICACIONES	PSICODIAGNOSTICO	COMENTARIOS
<p>A la fecha, la prueba Monterrey ha tenido cinco revisiones, siendo vigente la realizada en 1981.</p> <p>ASPECTOS PSICOMETRICOS: Se desconocen, ya que únicamente se cuenta en los centros, con Instructivos de Aplicación y Calificación. No hay información acerca de la Metodología y Estandarización.</p>	<p>Esta prueba se utiliza en las Unidades de Grupos Integrados dependientes de la DGEE para la Detección y Selección de los niños con problemas en el proceso de adquisición de la lecto-escritura y el cálculo matemático, así como</p> <p>para los alumnos repetidores del primer y segundo grado de educación primaria.</p>	<p>A pesar de ser considerada como una prueba básica de los Grupos Integrados, no se reportan en el manual los criterios técnicos y metodológicos de su estandarización lo que nos sugiere que no se ha realizado hasta el momento.</p> <p>Ante la dificultad para obtener información relacionada con este instrumento, se cuestiona:</p> <ul style="list-style-type: none"><li>- La utilidad de éste instrumento.</li><li>- La validez y confiabilidad en la detección y selección de los problemas escolares.</li></ul>

RUEBA : ESCALA DE INTELIGENCIA STANFORD-BINET

AUTOR: TERMAN-MERRILL

PAIS: ESTADOS UNIDOS

AÑO: 1905

REVISIONES Y MODIFICACIONES	PSICODIAGNOSTICO	COMENTARIOS
<ul style="list-style-type: none"> <li>- Alfred Binet. 1905, Francia.</li> </ul>	<p>En los Centros CAPEP y DGEE, se utilizan el Manual de la Traducción Española, para el Diagnóstico de los problemas de aprendizaje.</p>	<ul style="list-style-type: none"> <li>- A pesar de ser una de las pruebas con múltiples revisiones, y la más utilizada para el psicodiagnóstico en los Centros, no existen revisiones o modificaciones actuales de la escala.</li> </ul>
<ul style="list-style-type: none"> <li>- Alfred Binet. 1908, Francia.</li> </ul>	<p>La prueba se aplica generalmente a niños de 4 a 6 años de edad.</p>	<ul style="list-style-type: none"> <li>- En ambas Instituciones, se sabe del manejo incorrecto de las Tablas Normativas, sin embargo no se cuestionan el porque, ni se promueven cambios.</li> </ul>
<ul style="list-style-type: none"> <li>- Binet y Simon. 1911, Francia.</li> </ul>	<p>En los CAPEP se emplean las Tablas Normativas propuestas por Pinneau en 1960.</p>	<ul style="list-style-type: none"> <li>- La Escala original es altamente costosa por lo que se recurre a las reproducciones, afectando con esto su confiabilidad y validez.</li> </ul>
<ul style="list-style-type: none"> <li>- Stanford-Binet. 1916, Estados Unidos.</li> </ul>	<p>En la DGEE se emplean las Tablas Normativas revisadas en 1972 por Pinneau.</p>	
<ul style="list-style-type: none"> <li>- Terman-Merrill. 1937, Estados Unidos.</li> </ul>		
<ul style="list-style-type: none"> <li>- Terman-Merrill. Forma LM. 1960, Estados Unidos.</li> </ul>		
<ul style="list-style-type: none"> <li>-Revisión de Tablas Normativas R. Pinneau. 1960-1972, Estados Unidos.</li> </ul>		
<ul style="list-style-type: none"> <li>- Revision y Traducción de la Escala de 1960: J. Germain. 1975, España.</li> </ul>		
<p><b>ASPECTOS PSICOMETRICOS:</b> No ha sido estandarizada en México.</p>		

RUEBA: PRUEBA DEL DIBUJO DE UN HOMBRE

UTOR: FLORENCE GOODENOUGH

UGAR: ESTADOS UNIDOS

ÑO: 1926

REVISIONES Y MODIFICACIONES	PSICODIAGNOSTICO	COMENTARIOS
<p>-Prueba de la Inteligencia Infantil por medio del Dibujo de la Figura Humana TRADUCCION. María Luisa Ferrando de Cobanera. 1951, Argentina.</p>	<p>En los CAPEP, la prueba de Goodenough se utiliza como parte de la Batería de Diagnóstico, con objeto de obtener un Coeficiente Intelectual.</p>	<p>- La utilidad de la prueba difiere de acuerdo a cada Institución. La utilizan según los lineamientos establecidos, sin cuestionarse el porque de esta diversificación, lo que propicia diferentes interpretaciones.</p>
<p>- Prueba del Dibujo de un Hombre. Goodenough. REVISION Y EXTENSION: Dale Harris. 1963, Estados Unidos.</p>	<p>En la DGEE, la interpretación del Dibujo de la Figura Humana se realiza en base a los criterios de Koppitz, obteniéndose datos generales sobre el niño, sin reportar Coeficiente Intelectual.</p>	<p>- Siendo éste, un instrumento básico y de enlace entre los maestros y los Centros, se cuestiona la vigencia de las Normas de Calificación puesto que la prueba se diseñó hace 68 años.</p>
<p>- El Test de Goodenough. REVISION, AMPLIACION Y ACTUALIZACION: Dalia Ares. 1981, España.</p>	<p>En algunos centros dependientes de la DGEE se utiliza la Escala de Harris-Goodenough como una prueba de tipo proyectivo, tomándose en cuenta la escala cualitativa.</p>	<p>- Las investigaciones realizadas en México son numerosas, sin embargo todas ellas utilizan Normas de Calificación extranjeras, poniendo en duda los resultados obtenidos, ya que no hay investigación sobre el instrumento mismo.</p>
<p>- Prueba del Dibujo de la Figura Humana. REVISION: E. Koppitz. 1967, Estados Unidos.</p>		<p>- Por ser una prueba muy solicitada, de fácil aplicación, bajo costo y mínimo tiempo, no se le ha dado la importancia suficiente para estandarizarla a la población mexicana.</p>
<p>- El Dibujo de la Figura Humana en Niños. TRADUCCION: María Julia García y M. Braile. 1973, Argentina. ASPECTOS PSICOMETRICOS: No ha sido estandarizada en México.</p>		

**PRUEBA: ESCALA DE INTELIGENCIA PARA LOS NIVELES PREESCOLAR Y PRIMARIA.**

**WPPSI.**

**AUTOR: DAVID WECHSLER.**

**LUGAR: ESTADOS UNIDOS.**

**AÑO: 1967.**

REVISIONES Y MODIFICACIONES	PSICODIAGNOSTICO	COMENTARIOS
<p>Escala de Inteligencia para los Niveles Preescolar y Primaria. WPPSI-Español.</p> <p>TRADUCCION: Hinojosa Elsa y Treviño Margarita. 1981, México.</p> <p>ASPECTOS PSICOMETRICOS: No ha sido estandarizada en México.</p>	<p>En los CAPEP se utiliza la Traducción de 1981. El objetivo es medir la inteligencia de los preescolares. A pesar de que la Escala está contemplada como una prueba básica su utilización es mínima, empleándose en su lugar la Escala de Inteligencia Terman-Merrill.</p> <p>En la DGEE la prueba se llega a utilizar en forma esporádica, en su lugar se emplea con mayor frecuencia la Escala Terman-Merrill.</p>	<p>- A pesar de que en el Manual de la prueba se especifica que las normas establecidas no son adecuadas para poblaciones mexicana y latina, esto pasa inadvertido, obteniéndose así Diagnósticos que no corresponden totalmente a la población.</p> <p>- A diferencia del WISC, en el WPPSI no se ha realizado ninguna revisión o adecuación técnica desde su creación.</p> <p>- En diversos trabajos de Tesis se ha comentado, entre otras cosas la necesidad de revisar las normas de calificación de la prueba, sin embargo esto no se ha tomado en cuenta, lo cual nos lleva a pensar en el poco interés que se le presta a esta área de la educación.</p> <p>- Por otro lado, si el WPPSI estuviera adaptado a la población mexicana, sería una mejor opción para medir el nivel intelectual de los niños, ya que es una prueba más actual que la Escala Terman-Merrill.</p>

RUEBA: ESCALA DE INTELIGENCIA PARA EL NIVEL ESCOLAR. WISC.

UTOR: DAVID WECHSLER.

UGAR: ESTADOS UNIDOS.

ÑO: 1949.

REVISIONES Y MODIFICACIONES	PSICODIAGNOSTICO	COMENTARIOS
<p>- Escala de Inteligencia para el Nivel Escolar Revisada. WISC-R. David Wechsler. 1974, Estados Unidos.</p> <p>- Escala de Inteligencia para el Nivel Escolar Revisada. WISC-R Español. TRADUCCION: Elke Kort de Capella. 1981, México.</p> <p>- Escala de Inteligencia para Nivel Escolar revisada para México WISC - RM ADAPTACION Y ESTANDARIZACION: Gómez Palacio y Cols. DGEE.1982, México</p>	<p>El WISC-RM forma parte de la Bateria de pruebas, utilizadas para la evaluación psicológica, en los centros de la DGEE.</p> <p>Se aplica a niños desde los 6 a los 17 años de edad. La interpretación cualitativa se obtiene a través de la propuesta de Kaufman.</p> <p>El WISC-RM, se puede interpretar dentro de la Bateria SOMPA como una alternativa para los niños que obtienen un CI bajo, haciendo la observación de que únicamente podrá ser aplicada para la población para la cual fué estandarizada o población similar a la de la muestra.</p>	<p>- Por lo anterior, se considera que esta prueba reúne los requisitos necesarios DE CONFIABILIDAD Y VALIDEZ para ser aplicada a la población infantil mexicana del D.F.</p> <p>- Es una prueba que cuenta con suficiente bibliografía, accesible a todo especialista que trabaja en el campo de la educación en México.</p> <p>Sin embargo, esta restringida por los aspectos técnicos de la estandarización a los niños del D.F., limitando el acceso a la población indígena, rural y conurbada.</p>
<p>ASPECTOS PSICOMETRICOS: Muestra aleatoria de 1,100 niños, asistentes a escuelas primarias y secundarias públicas del D.F. La CONFIABILIDAD Y VALIDEZ se sustentaron en la revisión de 1974. Así como a través de la Fórmula de Spearman-Brown; Escala Verbal de .70 a .81; Escala Ejecutiva de .66 a .82; CIM=100 y D.S.±15. CONSISTENCIA INTERNA a través de División por Mitades. CORRELACION ENTRE LOS CIs y las subescalas por la fórmula de Pearson M=10 y D.S.±3.</p>		<p>- Por el elevado costo de los materiales de aplicación, se utilizan en los Centros, reproducciones que pudieran afectar su confiabilidad.</p>

RUEBA : BATERIA DE EVALUACION INTELECTUAL SECUENCIAL SIMULTANEA.  
KAUFMAN.

UTOR: A. KAUFMAN.

UGAR: ESTADOS UNIDOS.

ÑO: 1982.

REVISIONES Y MODIFICACIONES	PSICODIAGNOSTICO	COMENTARIOS
<p>- Revisión, Traducción y Estandarización: Gomez Palacio y Cols. 1985, México.</p> <p>ASPECTOS PSICOMETRICOS: Muestra aleatoria de 1,100 niños. De 2.6 a 12.6 años de edad. Asistentes a guarderías y primarias de las Delegaciones Políticas del D.F. (40 guarderías y 70 primarias). Se eliminó el rango de edad de 2.6 por no mostrar suficiente confiabilidad.</p> <p>Se obtuvieron Coeficientes de Confiabilidad de Consistencia Interna por Mitades Separadas. Se aplicó en primer lugar el Test de Vocabulario Visual de Peabody (Traducción) con propósito de Validación, seguido por la Bateria.</p>	<p>En la DGEE la Bateria esta incluida como instrumento básico en el Psicodiagnóstico. Es aplicada a niños entre 3 y 12 años de edad, con el fin de obtener información sobre el estilo de procesamiento cognitivo que presentan.</p> <p>Dada su complejidad, es una prueba que debe ser aplicada solamente por personal capacitado.</p> <p>En los CAPEP, únicamente se capacitó a un grupo de psicólogos en la utilización de esta prueba. Sin embargo no es utilizada ya que no esta contemplada como un instrumento básico para el diagnóstico.</p>	<p>- Dado que la prueba Peabody no es una prueba equivalente al Kaufman por no estar estandarizada en México, se cuestiona el propósito para el cual fué empleada.</p> <p>- No se reportan datos técnicos relacionados con los criterios de validez y margen de error empleados en la estandarización.</p> <p>- Su aplicación se limita a la población del D.F. restringiendose a zonas rurales, indígenas y conurbadas.</p> <p>- Al utilizarse la prueba esporadicamente, se pone en evidencia el alto costo de la investigación y la poca productividad que tiene el instrumento.</p>

RUEBA: SISTEMA DE EVALUACION MULTICULTURAL Y PLURALISTICO. SOMPA.

AUTOR: JANE MERCER.

LUGAR: ESTADOS UNIDOS.

AÑO: 1978.

REVISIONES Y MODIFICACIONES	PSICODIAGNOSTICO	COMENTARIOS
<p>- REVISION, TRADUCCION Y ESTANDARIZACION: Gómez Palacio y Cols. DGEE. 1982, México.</p> <p>ASPECTOS PSICOMETRICOS: En general la muestra estuvo formada por 1,100 sujetos asistentes a escuelas primarias y secundarias del D.F. Turnos matutino y vespertino.</p> <p>BENDER: Muestra de 700 sujetos entre 6.3 y 12.8 años. Se obtuvieron medias para cada nivel de edad. Se utilizaron ecuaciones de regresión proporcionando medias estimativas y D.S. predictivas. Se utilizó el Manual de Puntuación de 1963.</p> <p>ESCALAS SOCIOCULTURALES: Muestra de 1,100 padres de familia. Se modificó la conceptualización de algunas palabras. Solo dos de las Escalas tienen significancia estadística para México. los resultados se presentaron en porcentajes.</p>	<p>La Bateria SOMPA se utiliza en los Centros de la DGEE solo en casos especiales, es decir, cuando el niño obtiene un CI entre 57 y 70 y se sospecha de un mayor potencial de aprendizaje.</p> <p>En los CAPEP, la Bateria no está incluida en la selección de pruebas para el psicodiagnóstico.</p>	<p>- La Bateria de prueba SOMPA es una de las pocas pruebas que ha sido estandarizada para la población infantil mexicana. Sin embargo una de sus carencias es la muestra tomada para la estandarización. El número es reducido en rangos de edad y se limita a población exclusivamente del D.F.</p> <p>- En el Manual no se especifica el Coeficiente de Validez ni el Margen de Error.</p> <p>- Una de las limitaciones de ésta prueba es que a pesar de estar estandarizada en el D.F., solo se utiliza excepcionalmente, ya que el tiempo de aplicación es aprox. de 5 hrs. Por lo tanto, a pesar del extenso trabajo realizado resulta ser una prueba poco práctica por la duración y dificultad en la aplicación e interpretación.</p> <p>- La Estandarización de esta Bateria en México, llama la atención, ya que la prueba original ha sido muy criticada en Estados Unidos, principalmente por la eficacia del esquema del modelo tripartito, la falta de normas nacionales y la falta de pruebas de Validez de sus medidas.</p>

**INVENTARIO DE LA  
CONDUCTA ADAPTATIVA  
PARA NIÑOS. ABIC:**

Muestra de 600 niños de 6 a 11 años de edad. Los Coeficientes de Confiabilidad para cada nivel de edad se obtuvieron por el procedimiento de División por Mitades (Spearman-Brown). La mitad de los coeficientes esta por encima de  $.80 + 5$ . El Estudio de Correlación entre los cinco sistemas del ABIC es de  $.84$  o superior.

**WISC-RM:**

La muestra fué de 1,100 sujetos entre los 6 y los 16 años. La Validez y la Confiabilidad de Consistencia Interna para cada subescala se obtuvo a través de División por Mitades (Spearman-Brown) Escala Verbal entre  $.70$  y  $.81$ ; Escala de Ejecución entre  $.66$  y  $.82$ . La Media del CI es de 100 y la D.S. de  $+ 15$ . La Correlación entre los CIs y las subescalas Media de 10 y D.S  $+3$  (Fórmula Pearson )

## ***CONCLUSIONES***

### **Como instrumentos de Investigación:**

Los trabajos de investigación en las Instituciones dedicadas al estudio de las pruebas psicológicas, así como los trabajos de Tesis, son numerosos, sin embargo la difusión de los resultados es escasa o nula, en la mayoría de los casos, ya que en México existen pocos libros y revistas que permitan dar a conocer los diversos trabajos que se han realizado.

En ocasiones los psicólogos dedicados al Psicodiagnóstico no cuentan con mas información que la proporcionada por los manuales de Normas y Lineamientos de cada institución, que por un lado limita la aplicación de diferentes Instrumentos de Medición que no se encuentran normados y por el otro, no permiten que se cuestione la pertinencia o no de la utilización de las pruebas seleccionadas.

Ante la carencia de Instrumentos de Medición Psicológica diseñados en nuestro País, se recurre a la utilización de pruebas extranjeras traducidas, la mayoría en otros países, como es el caso de las Escalas Evolutivas del Desarrollo; la Prueba Guestáltico / isomotora de L. Bender; la Escala de Inteligencia Stanford-Binet y la prueba del Dibujo de un Hombre.

Esto trae como consecuencia que el vocabulario y los conceptos que se exploran en cada reactivo de las pruebas difieren en cada País y en cada región y lo mas importante, las normas con las cuales la población infantil mexicana es evaluada pertenezcan a otros contextos socio-culturales.

Autores como Salvia (1981), Mehrens (1982), Anastasi (1971), entre otros, señalan que la Normalización de las pruebas únicamente sirve para la población para la cual fueron diseñadas, por tanto es necesario tomar en cuenta ésta observación en la interpretación del Psicodiagnóstico al utilizar pruebas no estandarizadas para la población mexicana.

En los trabajos realizados por el Dr. Díaz Guerrero (1994), se menciona constantemente que las pruebas psicológicas responden al desarrollo cultural de un País, por lo que es importante señalar la necesidad apremiante de revisar y estandarizar los instrumentos que en la actualidad se utilizan en los Centros de Atención Especializada, para que respondan, siguiendo a Díaz Guerrero, al desarrollo cultural de los niños de México.

Algunas de las investigaciones que se han llevado a cabo a través de la D.G.E.E., si bien han pretendido incorporar al Psicodiagnóstico nuevos Instrumentos de Medición, esto se ha visto limitado por varios factores como:

- Desde la Psicometría los Instrumentos no cumplen con las características técnicas de las pruebas.
- Las muestras poblacionales no han sido representativas de las diferentes regiones del País.
- Los costos que implica el uso de las prueba y el tiempo de aplicación, no corresponden a los presupuestos y demanda de atención de las Instituciones por ejemplo en el S.O.M.P.A., se requieren de 5 horas y de 3 especialistas para su aplicación.

### **Como Instrumentos en el Psicodiagnóstico:**

Las pruebas utilizadas para la detección y diagnóstico de las dificultades que interfieren en el proceso de la enseñanza-aprendizaje son consideradas como elementos indispensables para la Toma de Decisiones.

Las pruebas Psicológicas no determinan el problema de aprendizaje en sí, informan acerca de las capacidades y habilidades con las que cuenta el sujeto hasta el momento de aplicación de la prueba, éstas habilidades y capacidades serán contrastadas con lo que se espera de el sujeto, de acuerdo al Marco Teórico o expectativa que sustente cada prueba, siendo esto el punto de partida para la elaboración del plan de tratamiento.

Por las razones anteriores se hace indispensable que las pruebas se encuentren estandarizadas de acuerdo a las características de las población a la cual se va a evaluar.

Las pruebas aquí revisadas cumplen en su mayoría con los requisitos de Validez y Confiabilidad establecidos, sin embargo el hecho de no estar estandarizadas para la población infantil mexicana, evidencian el comparar a un niño con un grupo en el cual no se haya representado, ocasionando la obtención de diagnósticos poco certeros como lo señalan Salvia (1981), Brown (1980), Forns (1980).

Los psicólogos dedicados a la aplicación de los Instrumentos Psicológicos, ante los problemas de traducción, realizan adaptaciones en algunos reactivos de las mismas con objeto de que las consignas sean más comprensibles para el niño mexicano, utilizando para ésto su criterio lo cual resta Validez al instrumento pues carece de los criterios técnicos para realizar éstas modificaciones, como lo señalan Landsheere (1978), Salvia (1981), Thorndike (1978) entre otros. Un ejemplo de ello se observa en la prueba Stanford-Binet, en la cual se cambia la palabra melocotón por durazno ó naranja sin ningún estudio anterior.

Las pruebas psicológicas utilizadas en la actualidad para la detección y diagnóstico fueron creadas décadas atrás y si bien miden lo que pretenden, no han sido revisadas ni actualizadas en ésta época en la cual los niños son estimulados de manera diferente como la prueba de F. Goodenough, diseñada hace 68 años, evaluada actualmente de acuerdo al diseño original; también la Escala de Stanford-Binet, elaborada en 1905 y que a pesar de las múltiples revisiones de que ha sido objeto en otros países, en México, no se ha promovido una revisión y actualización de la misma. A pesar de lo anterior ambas pruebas son instrumentos básicos en las Instituciones sujetas a la presente investigación.

De las pruebas revisadas, se encontró que la Escala de Inteligencia WISC-RM estandarizada por la D.G.E.E., cubre con los requisitos indispensables de Confiabilidad y Validez. Es una prueba utilizada regularmente para los niños de edad escolar, sin embargo está restringido su uso para la población del D.F., exclusivamente, debido a que la estandarización se realizó únicamente con una muestra representativa de éste núcleo.

Podemos concluir que en México existe déficit en cuanto a la utilización de instrumentos de Medición en el campo educativo.

Esto obedece en buena medida a los limitados recursos que se destinan a la investigación educativa.

## ***SUGERENCIAS***

- Fomentar la comunicación entre las Instituciones dedicadas a la revisión y estudio de las pruebas psicológicas y los Centros de Atención Especializada, con el objeto de actualizar y renovar los instrumentos de medición psicológica utilizadas en el psicodiagnóstico.
- Difundir los trabajos de investigación que se lleven a cabo a través de revistas y publicaciones, con el objeto de mantener el interés y la actualización de los expertos en el campo de la psicología.
- Es importante fomentar la investigación en el ámbito educativo, comisionando equipos interdisciplinarios en la revisión y estandarización de las pruebas psicológicas, con el objeto de actualizar y renovar los instrumentos que se utilizan para el Diagnóstico Psicológico, beneficiando de ésta manera a la población infantil.
- Generar recursos económicos que permitan contar con Instrumentos de Medición Psicológica confiables y realizados para llevar a cabo un Psicodiagnóstico de acuerdo a las características del niño mexicano.
- Considerar los diferentes tipos de población; urbana, rural e indígena a fin de llevar a cada una estandarización que permita el uso de las pruebas psicológicas en las diferentes regiones del País, en especial con los grupos marginados que es donde se detectan los problemas más severos en éste campo.
- Revisar las investigaciones realizadas por la D.G.E.E., referentes a la estandarización de las pruebas: SOMPA y KAUFMAN para hacerlas más funcionales.
- Proponer y diseñar Instrumentos de Medición Psicológica que respondan a la realidad de las características del niño mexicano.
- Promover y organizar: cursos, seminarios, talleres, conferencias, etc., sobre éstos aspectos para que la información llegue al especialista, con el fin de mantenerlo en el ejercicio de su función.
- Los resultados expresados en éste trabajo abren la posibilidad de retomar diversas líneas de investigación, generando así nuevos proyectos que satisfagan las necesidades aquí surgidas.

# ***ANEXO***

## ANEXO 1

ABLA 6 . Datos normativos para la Escala de maduración del Bender.

Distribución de las medidas y desviación standard

Edad	N	Media	Desviación Standard	+/- D. S.
5.0 a 5.5	81	13.6	3.61	10.0 a 17.2
5.6 a 5.11	128	9.8	3.72	61.1 a 13.5
6.0 a 6.5	155	8.4	4.12	4.3 a 12.5
6.6 a 6.11	180	6.4	3.76	2.6 a 10.2
7.0 a 7.5	156	4.8	3.61	1.2 a 8.4
7.6 a 7.11	110	4.7	3.34	1.4 a 8.0
8.0 a 8.5	62	3.7	3.6	0.1 a 7.3
8.6 a 8.11	60	2.5	3.03	0.0 a 5.5
9.0 a 9.5	65	1.7	1.76	0.0 a 3.5
9.6 as 9.11	49	1.6	1.69	0.0 a 3.3
10.0 a 10.5	27	1.6	1.67	0.0 a 3.3
10.6 a 10.11	31	1.5	2.1	0.0 a 3.6
TOTAL	1,104			

Distribución por grados de los puntajes medios del Bender

comienzo del año	N	Edad (Media)	Puntaje	Desviación Standard	Ubicación al +/- Desviación Standard
J. de Infantes	38	5.4	13.5	3.61	9.9 a 17.1
1er. Grado	153	6.5	8.1	4.41	4.0 a 12.2
2do. Grado	141	7.5	4.7	3.18	1.5 a 7.9
3er. Grado	40	8.7	2.2	2.03	0.2 a 4.2
4to. Grado	39	9.8	1.5	1.88	0.0 a 3.4

## ***BIBLIOGRAFIA***

- Acle Tomasini, Guadalupe. Estudio exploratorio sobre la relación entre la Escala de Inteligencia Wechsler para niños y la Prueba de Goodenough-Harris en un grupo del Instituto Patria. Tesis de Licenciatura. México. UNAM. 1971.
- Adaptación Y Estandarización de la Batería de Evaluación Intelectual Secuencial- Simultánea Kaufman en México. México. D.E.D.G.E.E., 1986
- Ahmann, J. Stanley. Los Test psicológicos y su aplicación en la escuela. 2a edición. Argentina, Ed. Troquel. 1969.
- Ahumada Rodríguez, Rene. Estudios Psicológicos de validación cruzada del Wechsler WISC) y el Goodenough en escolares mexicanos. Tesis de Licenciatura. México, UNAM, 1965.
- Ahumada Rodríguez, Rene. Desarrollo intelectual del escolar mexicano. Tesis de Doctorado. México, UNAM, 1969.
- Alcazar J. y Galván Millan. Guía Para la aplicación de la Escala de Desarrollo Normal del Niño de Gesell y Amatruda. Trabajo de Especialidad. México, UNAM, 1981.
- Anastasi, Anne. Test Psicológicos. México, Ed. Aguilar, 1971.
- Anderson, Richard y Faust, Gerald. Psicología Educativa. México, Ed. Trillas, 1981.
- Ausbel, David. Psicología Educativa un punto de vista cognoscitivo. México, Ed, Trillas 1980.
- Arias Galicia, Fernando. Los Test. México, Ed. Trillas, 1981.
- Beloff, John. Las Ciencias Psicológicas. Una revisión de la Psicología Moderna. Ed. México, El Manual Moderno, 1979.
- Bender, Lauretta. Test Guestáltico Visomotor. Usos y Aplicaciones Clínicas. Buenos Aires-Barcelona, Ed. Paidós, 1982.
- Brown, Frederick. Principios de la medición en Psicología y Educación. México, Ed. El Manual Moderno, 1980.
- Canudas González, Hilda M. Validez Predictiva de las pruebas: Dettroit Engel: Rutgers Drawing: Frostig: Goodenough y Koppitz. Tesis de Licenciatura. México, UNAM. 1976.

- Cantón Arjona, V. y Tadeo Madrazo, Z. Consideraciones teórico prácticas sobre el desarrollo evolutivo del niño de 0 a 5 años. Construcción y presentación de las Escalas Evolutivas de A. Gesell . México, UPN, 1990.
- Caparoso Chávez, S. y Carrillo Rodríguez, O. Estudio exploratorio entre los problemas de conducta y la disfunción cerebral mínima evaluado a través del Bender, WPPSI y el Dibujo de la Figura Humana. Tesis de Licenciatura. México, Universidad Iberoamericana, 1987.
- Carrasco Altamirano, Alma C. El Psicólogo profesional de apoyo al trabajo de los grupos integrados "B". Tesis de Licenciatura. México, UNAM, 1983.
- Castelán Martínez, M. Normalización de la Escala de Inteligencia Wechsler para el Nivel Preescolar y Primario (WPPSI) en una muestra de la Cd. de México. Tesis de Licenciatura. México, UNAM, 1990.
- Castellanos Pérez, S., Servin O. y Oriostegui P. El nivel de desarrollo de los menores que trabajan de acuerdo a la teoría de Arnold Gesell. Tesis de Licenciatura. México, UNAM, 1984.
- Cerda, Enrique. Psicometría General. Barcelona, Ed. Herder, 1978.
- Chadwick, Clifton. "Evaluación Educacional" en: Evaluación de la práctica docente. México, U.P.N. (Antología), 1987.
- Collin, G. Selección de tests al servicio de la psicología infantil. 2a. edición. Buenos Aires. Ed. Kapelusz, 1959.
- Coto Villa, M. y Llano de la Torre, A. Estudio sobre la traducción, adaptación, confiabilidad y validez del test para preescolares y primaria de Wechsler (WPPSI) en un grupo de niños de nivel socioeconómico medio alto. Tesis de Licenciatura. México, Universidad Iberoamericana, 1981.
- De la Parra Aguilera, Guadalupe. Escala de Inteligencia para Preescolares de Wechsler. Un estudio comparativo por nivel socioeconómico. Tesis de Licenciatura. México. UNAM, 1977.
- Diccionario Enciclopédico de la Educación Especial. México, Ed. Diagonal Santillana, 1986.
- Dorsch, Friedrich. Diccionario de Psicología. 5ta. edición, Barcelona, Ed. Herder, 1985.
- Eco, Humberto. Como se hace una tesis. España, Ed. Gedisa, 1977.

- Ehrlich Quintero, Nelva. Investigación sobre los resultados del WISC en un grupo de preescolares mexicanos de nivel socioeconómico bajo. Tesis de Licenciatura. México, UNAM, 1969.
- Escala de Inteligencia Revisada para el Nivel Escolar WISC-R Español. (Manual). Wechsler, David. Traducción: Elke Kort de Capella. México, Ed. EL Manual Moderno, 1981.
- Escala de Inteligencia para los Niveles Preescolar, Primaria, WPPSI - Español (Manual). Wechsler, David. Traducción: Hinojosa, Elsa y Treviño, Margarita. Ed. El Manual Moderno, México, 1981.
- Escala de Inteligencia Wechsler Adaptada para Niños Mexicanos de niveles de edad de 6 a 16 años. (WISC-R Mexicano). Manual Wechsler, David. Adaptación: Gómez Palacio M., Padilla E., Roll S. México, S.E.P.-D.G.E.E., 1983.
- Fernández Bañuelos, Lucía E. Correlación de la Escala Ejecutiva del WISC y los Tests de Figuras de Torrance "Pensando positivamente con dibujos", en niños de 7 a 9 años alumnos de Escuelas Oficiales del D. F. Tesis de Licenciatura. México, Universidad Iberoamericana, 1980.
- Flores Flores, Alberto. Análisis correlacional de los subtest esencialmente verbales del WAIS como indicadores confiables del coeficiente intelectual. Tesis de Licenciatura. México, UNAM. 1984.
- Flores Luna, Ma. de Jesús. Estudio Correlativo entre la Escala de Inteligencia Wechsler para escolares (WISC) y el Test de Habilidades Psicolingüísticas de Illinois (ITPA), en niños que asisten a una Primaria. Tesis de Licenciatura, México, UNAM, 1986.
- Forns, María. Psicodiagnóstico y Estadística. Barcelona, Ed. CEAC, 1980.
- Fuchs Modiano, Esther. Estudio Experimental de la Frustración en un grupo de niños preescolares a través del Test "Fuchs - Lara". Una comparación al sexo. Tesis de Licenciatura. México, UNAM, 1969.
- Fuentes Montañés, M. y Martínez Arévalo, M. Normalización del Test Guestáltico Visomotor de Lauretta Bender en Escolares Mexicanos. Tesis de Licenciatura. México, Universidad del Valle de México, 1990.
- Galicia Cadena, Irma. El Test de Lauretta Bander como auxiliar para el maestro especialista, para evaluar y diagnosticar al niño con problemas de aprendizaje. Tesis de Licenciatura. México, Escuela Normal de Especialización, 1986.

- Gómez Almada, Carlos F. Estudios Psicológicos del Goodenough Harris a nivel Preescolar mexicano. México, Tesis de Licenciatura. U.N.A.M., 1966.
- Garza Mercado, Ario. Manual de Técnicas de Investigación para estudiantes de C.S. 4a. edición. México, El Colegio de México. Ed. Harla, 1988.
- Georgia, Sanchs A. Medición y Evaluación. México, Ed. Herder, 1975.
- Gesell, Arnold y Amatruda, C. Diagnóstico del Desarrollo Normal y Anormal del Niño. Buenos Aires, Paidós, 1979.
- Gesell, Arnold y Amatruda, C. Diagnóstico del Desarrollo Normal y Anormal del Niño. México, Ed. Paidós, 1985.
- Gómez Palacio, Margarita y Rangel Hinojosa, E. Estandarización de la Batería de Pruebas SOMPA en México D.F.: Informe sobre teoría y resultados. Adaptación. México, D.G.E.E., S.E.P.- O.E.A., 1982.
- Gómez Palacio, Margarita y Rangel Hinojosa, E. Identificación de la capacidad intelectual "Muy Superior" por medio de la Escala de Inteligencia WISCRM (en estudiantes de escuelas primarias públicas en México), México, D. F. D.G.E.E., S.E.P. - O.E.A., 1984.
- Goodenough, Florence. Test de inteligencia por medio del dibujo de la Figura Humana. 2da. edición. Argentina, Ed. Paidós, 1957.
- Gutiérrez Ruiz, M. y López López, I. Programa de Entrenamiento en el Diagnóstico de Gesell y Amatruda a asistentes educativas pertenecientes a un Centro de Desarrollo Infantil. Tesis de Licenciatura. México, U.N.A.M., 1982.
- Harris, Dale B. El Test de Goodenough: Revisión, Ampliación y Actualización 1er. edición. Castellana. España, Ed. Paidós 1981.
- Hasbach Soto, M y Traulsen Nuñez. Ejecución del Test de Bender en niños y niñas entre once y doce años de dos niveles socioeconómicos. Tesis de Licenciatura. México, Universidad Iberoamericana, 1981.
- Hegerdahl G., Gloria. Evaluación psicológica de acuerdo a la prueba de DFH para niños, de Koppitz a grupos Escolares Mexicanos. Tesis Profesional. México. Universidad Femenina de México., 1979
- Hernández Sánchez, Soledad. La Psicomotricidad del Niño Down de 1 a 3 años, con un análisis comparativo y las características marcadas por Arnold Gesell en su Teoría de Desarrollo en niños anormales. Tesis de Licenciatura, México, UPN, 1989.

- Hurtado de Mendoza, Ma. de los Angeles. Estudio sobre el subtest de Comprensión de la Escala Verbal del Test de Inteligencia para niños de Wechsler. Tesis de Licenciatura. México. UNAM, 1966.
- Jarquín Fagoaga, Ofelia. Estudio piloto para la estandarización de la Escala de Wechsler para niños (Investigación Psicológica en sujetos mexicanos de 5 a 15 años). Tesis de Licenciatura. México. UNAM, 1965.
- Karmel, Louis J. Medición y Evaluación Escolar. México, Ed. Trillas, 1979.
- Kaufman, Alan S. Psicometría razonada con el WISC-R. México, Ed. El Manual Moderno, 1982.
- Koppitz, Elizabeth. El Dibujo de la Figura Humana. 8a. edición. Buenos Aires, Ed. Guadalupe, 1963.
- Koppitz, Elizabeth. El Test Guestáltico Visomotor para Niños. 8a. edición. Buenos Aires, Ed. Guadalupe, 1980.
- Landsheere, G. Los Test de Instrucción. Barcelona, Ed. Oikostau, 1978.
- Lanz Mora, Gladys E. Revisión de la Prueba del Dibujo de la Figura Humana de Elizabeth Koppitz en un Grupo de Niños Mexicanos. Tesis de Licenciatura, México, UNAM, 1983.
- Lineamientos Teóricos para la Elaboración del Perfil del Niño que asiste a CAPEP. México, Subsecretaría de Educación Elemental. D.G.E.P., S.E.P., 1991.
- Lineamientos para la Evaluación Psicológica. México, Dirección Técnica, Departamento de Programación Académica. S.E.P., DGEE, 1984.
- Livas González, Irene. Análisis e Interpretación de los resultados de la Evaluación Educativa. México, Ed. Trillas, 1988.
- Magnusson, David. Teoría de los Test. México, Ed. Trillas, 1978.
- Manual de Aplicación del Sistema Multicultural y Pluralístico (SOMPA). México, SEP - DGEE. 1982.
- Manual de Calificación del Sistema de Evaluación Multicultural y Pluralístico (SOMPA). México, SEP - DGEE., 1982.
- Manual de Normas y Procedimientos del Área de Psicología. México, Departamento de Operación de Centros de Atención Psicopedagógica de Educación Preescolar. S.E.P., D.G.E.P., 1998

- Mass Bustamante, M. y Nuñez Amezcua, M. Detección y Desarrollo del Aprendizaje a través de la LectoEscritura, mediante cuatro Pruebas Psicométricas: Bender, Frostig, ITPA y WISC-R. Tesis de Licenciatura. México, Universidad Iberoamericana, 1981.
- Mehrens, William y Lehman, Irvin. Medición y Evaluación en la Educación y en Psicología. México, Ed. CECSA, 1982.
- Miller Monroe, David. Resultados de las Pruebas Psicológicas. México, Ed. Limusa, 1975.
- Molina Castro, Jorge A. Ponencia: Batería de Evaluación Intelectual Secuencial ~ Simultánea Kaufman. Su aplicación en México. Oaxtepec, Morelos. Encuentro de Educación Especial, 1988.
- Morales, María Luisa. Psicometría Aplicada. México, Ed. Trillas, 1975.
- Nilo, Sergio U. "Temas de Evaluación", en: Evaluación de la Práctica Docente. (Antología). México, U.P.N. 1987.
- Nunnally, Jum C. Teoría Psicométrica. México, Ed. Trillas, 1987.
- Ortega M., Pablo y López Matías. Estudio de Estandarización de la Prueba Colectiva de Inteligencia General. Forma "A" de L.M. Terman. México, SEP. Instituto Nacional de Psicopedagogía, 1940.
- Padilla Alvarez, Ma. Francisca. Sistema de Evaluación Intelectual para Niños, denominado SOMPA. Adaptado a niños residentes en el D.F. Oaxtepec, Morelos, Encuentro de Educación Especial, 1988.
- Palemonte Salgado, Delia. El Dibujo de la Figura Humana como Instrumento para predecir el Rendimiento Escolar en niños de 6 años. Tesis de Licenciatura, México, UNAM, 1983.
- Pallares Díaz, Ana María. Estudio exploratorio sobre la confiabilidad del Test Goodenough Harris en un grupo de niños del Instituto Patria. Tesis de Licenciatura, México, UNAM, 1971.
- Parrodi García, Stella y Vázquez Arzaba, A. Estudio Diagnóstico Comparativo entre tres grupos psiquiátricos mediante la Prueba de Bender. Tesis de Licenciatura. México, UNAM, 1982.
- Patrón Castillo, Ma. del Carmen. Estudio de Validación Cruzada de Goodenough, Wechsler (WISC) y Harris, en el Escolar Mexicano. en: Estudio Normativo del WISC y Goodenough - Harris en 660 escolares

- mexicanos de tres niveles socioeconómicos. Por Appendini Tazer, Ida Lorenza y Cáceres Rivera. Tesis de Licenciatura, México, UNAM, 1968.
- Pérez Velasco, Gabriela C. y Sánchez Carlos, Ma. Teresa. Estudio comparativo entre niños de familias integradas y niños de madres solteras. Tesis de Licenciatura. México, UNAM, 1987.
- Phillips, Ray Cook. Evaluación y Educación. Buenos Aires, Ed. Paidós, 1974.
- Planchard, Emile. Iniciación a la Técnica de los Tests. Ed. EUDEBA, Buenos Aires, Manual de Psicología, 1970.
- Plan Nuevo León, SEP - DGEE, México, Gobierno del Estado de Nuevo León. Monterrey, Nuevo León. .
- Prueba Monterrey (para grupos integrados, 5ta. Versión). Gómez Palacio y Cols. SEP- DGEE. Centro de Información y Documentación.
- Prueba Psicopedagógica de Detección de Problemas de Aprendizaje. (Diseño Monterrey). México, SEP - DGEE. Centro de Información y Documentación.
- Ramírez López. Transtornos de Aprendizaje: Su entrenamiento a través del Método Frostig. Tesis de Licenciatura. México, UNAM, 1978.
- Ray Bazan y Alvarez Páramo. La utilidad del Bender y otras pruebas de Evaluación para discriminar entre niños con problemas específicos de aprendizaje de alto y bajo riesgo. Un estudio de subtipificación. Guadalajara, Jal., Instituto Tecnológico y de Estudios Superiores de Occidente. Tesis de Licenciatura. 1982.
- Reguera R., Ma. Lilibiana y Rentería, Concepción. Estudio comparativo del Test Guestáltico Visomotor de Lauretta Bender en niños y niñas de 5 a 6 años de edad. Tesis de Licenciatura. México, Universidad Iberoamericana, 1979.
- Reyes Lagunes, Isabel. El Wechsler para niños en México. Consideraciones Psicológicas sobre adaptación. Tesis de Licenciatura. México, UNAM, 1965.
- Rincón Gallardo, P. Carmen. Impacto del encuentro temprano en índices de desarrollo infantil y percepciones materna. Tesis de Licenciatura. México, UNAM, 1989.
- Rodríguez Sierra, Aurora. Correlación de Vías Colinérgicas Centrales con Aprendizaje, Memoria y Retardo Mental. México, Tesis de Maestría. UNAM, 1984.

- Rojas Soriano, Raúl. Guía para realizar investigaciones Sociales. 5ta. edición. México, UNAM, 1980.
- Rubio Traspeña, Ma. Irma. Manual de Aplicación para la Prueba Monterrey. Tesis de Licenciatura. México, UNAM, 1984.
- Salvia, John. Evaluación en la Educación Especial. México, Ed. El Manual Moderno, 1981.
- Sattler, Jerome M. Evaluación de la Inteligencia Infantil y Habilidades Especiales. 2da. edición. México, Ed. El Manual Moderno, 1988.
- Serrano Mora, Fernando. Curso de Psicometría aplicada a la Educación. México, Ed. Librería Imagen Editores, 1984.
- Soberanis Quintero, Ma. Estela y Tabora Mears, Gloria B. Estudio Comparativo entre 3 Escalas de Inteligencia (Terman - Merrill, WISC Y WPPSI). Tesis de Licenciatura. México, UNAM, 1984.
- Stufflebeam y Shinkield, H. Evaluación Sistemática. Barcelona, Ed. Paidós, 1987.
- Szekely, Bela. Los Test. Manual de Técnicas de Exploración Psicológica. 5ta. edición. Buenos Aires, Ed. Kapelusz, 1966.
- Tavella M., Nicolas. Los Test en la Escuela. 4a. edición. Argentina, Ed. Universitaria de Buenos Aires, 1972.
- Tecla, J., Alfredo. Teoría, Métodos y Técnicas en la Investigación Social. 14ava. edición. México, Ed. Ediciones Taller Abierto, 1993.
- Terman M, Lewis y Merrill A., Maud. Medida de la Inteligencia. Método para el empleo de las pruebas del Stanford - Binet. 3a. revisión. Traducción: José Germain. 9a. edición. Madrid, Editorial. Espasa, Caple. 1975.
- Tenorio Bahena, Jorge. Técnicas de Investigación Documental. 3ra. edición. México, Ed. Mc GrawHill, 1988.
- Thorndike, Robert y Hagen, Elizabeth. Test y Técnicas de Medición en Psicología y Educación. México, Ed. Trillas, 1978.
- Uribe Velázquez, Ana María A. Indicadores Emocionales en niños hipoacúsicos de 5 a 6 años por medio del DFH. Tesis de Licenciatura, México, UNAM, 1991.

- Valderrama, P. y Cois. Evolución de la Psicología en México. México, Ed. El Manual Moderno, 1994.
- Vallejos Dellaluna, Marla. Estudio Comparativo de las Pruebas de Inteligencia Raven y Goodenough en niños de 6 a 7 años. Tesis de Licenciatura. México, UNAM, 1968.
- Vargas Oseguera, Catalina. Un estudio Comparativo de la Percepción Visual entre niños de 5 a 6 años de edad en dos niveles socioeconómicos diferentes. Tesis de Licenciatura. México, Universidad Iberoamericana, 1981.
- Vázquez Becil, Julieta. Impacto de la asistencia a un Centro de Desarrollo Infantil en el conocimiento perceptual. Tesis de Maestría. México, UNAM, 1987.
- Venegas Padilla, Eda Lorna. Alteraciones detectadas a través del Bender y el Dibujo de la Figura Humana evaluación por el método de Elizabeth Koppitz en niños de la Ciudad de México Tepito. Tesis de Licenciatura. México, UNAM, 1985.
- Watkins, Ernest. Sistema de Puntuación de Watkins para el Test Gestáltico de Bender (Manual). Buenos Aires, Argentina, Ed. Panamericana, 1980.
- Xolalpa González, Salvador. La utilidad del Test Bender en la Educación Especial. Tesis de Licenciatura. México, Escuela Normal de Especialización, 1987.
- Zimbardo, Phillip. Psicología y Vida. México, Ed. Trillas, 1986.