



TEST DIAGNOSTICO EN HABILIDADES BASICAS:
VALIDACION POR DIFERENCIACION DE GRUPOS.

T E S I S

QUE PARA OBTENER EL TITULO DE:
LICENCIADO EN PSICOLOGIA EDUCATIVA
P R E S E N T A N :

ACOSTA VILLEGAS MARTHA BEATRIZ
ANGULO FUENTES MARISELA
CAMACHO OROZCO MARIA DEL CONSUELO

DIRECTORA: MTRA. ROSA MA. GONZALEZ JIMENEZ



MEXICO, D. F.

1997

2008/05/05

AGRADECIMIENTOS:

A la Lic. Rosa Ma. González Jiménez asesora de Tesis, por su apoyo, constancia y tolerancia que, mas brindo en todo momento durante la realización de este trabajo.

A los profesores: Carmen Ortega, Ma. Del Pilar Migez, Salvador Arciga y Cuitlahuac; lectores de tesis por su paciencia y comentarios.

A MI FAMILIA:

GRACIAS POR LOS VALORES, SACRIFICIOS Y APOYO CON QUE ME HAN ORIENTADO PARA LLEGAR A LAS METAS QUE ME HE FORZADO. Y MUY EN ESPECIAL A MIS PADRINOS CON QUIENES HE COMPARTIDO TODA MI VIDA.

A MI MADRE:

POR SER LA PERSONA QUE ME DIO LA VIDA Y A LA CUAL SIEMPRE ESTARE AGRADECIDA.

CONSUELO

A MIS PADRES:

GRACIAS PORQUE ME DIERON LA VIDA Y CON ELLA LA RAZÓN, A QUIENES LES DEBO LO QUE SOY, Y ENCAMINARON EN MI EL TRABAJO Y LA UTILIDAD EN LA VIDA.

A MIS HERMANOS:

POR SU COMPRESIÓN Y APOYO QUE EN TODO MOMENTO ME BRINDARON.

MARISELA

A MIS PADRES Y HERMANOS:

GRACIAS POR EL APOYO, EJEMPLO Y COMPRESIÓN QUE TODA LA VIDA ME HAN BRINDADO.

A MI ESPOSO E HIJO:

GRACIAS POR LA PACIENCIA Y EL TIEMPO QUE ESTE TRABAJO LES HA QUITADO.

MARTHA BEATRIZ.

El presente trabajo pretendió dar un mayor soporte psicométrico en términos de validez de la prueba al Test Diagnóstico en Habilidades Básicas TEHBA (1994), de primer grado de primaria - entendiendo por habilidades básica al conjunto de conocimientos escolares elementales (lectura, Escritura y Matemáticas) las cuales son base para la adquisición de nuevos aprendizajes-; a través de la validación por el método de diferenciación de grupos para lo cual se seleccionaron 73 sujetos perteneciendo 35 a grupo USAER y 38 grupo regular.

La aplicación se llevo a cabo de forma individual y en dos sesiones, en la primera matemáticas y en la segunda lectura y escritura.

Obteniéndose de manera general resultados de diferencia estadísticamente significativas. Aunque analizando por subpruebas se noto que en el área de matemáticas las subpruebas de lógica y Relaciones y Resolución de problemas no presentan esta diferencia.

INDICE

INTRODUCCION	1
---------------------	----------

CAP. I EVALUACION EDUCATIVA	3
------------------------------------	----------

- 1.1 Definición
- 1.2 Propósitos
- 1.3 Clasificación de la evaluación
- 1.4 Requisitos del proceso de evaluación
- 1.5 Evaluación de las habilidades básicas

CAP. II INSTRUMENTOS DE EVALUACION PSICOLOGICA	18
---	-----------

- 2.1 Antecedentes
- 2.2 Primeros instrumentos psicológicos
- 2.3 Clasificación de los instrumentos de evaluación psicológica
- 2.4 Cuestionamiento a los instrumentos de evaluación

**CAP. III LOS INSTRUMENTOS DE MEDICION:
ASPECTOS PSICOMETRICOS**

45

3.1 Principios de la medición

3.1.2 Características básicas de la medición

3.1.3 Errores de la medición

3.1.4 Confiabilidad

3.1.5 Validez

3.1.6 Normatividad

CAP. IV TEST DE HABILIDADES BASICAS: TEHBA

4.1 Conceptualización

4.2 Proceso de planeación del TEHBA

4.3 Descripción de las pruebas que integran el TEHBA

METODOLOGIA

86

ANALISIS DE RESULTADOS

89

CONCLUSIONES

96

LIMITACIONES

101

BIBLIOGRAFIA

102

INTRODUCCION

La evaluación, como una forma sistemática de constatar los efectos del proceso educativo, se ha señalado como una prioridad ante el cuestionamiento frecuente en nuestro país al Sistema Nacional de Educación referente a la baja calidad de los servicios que ofrece.

Una de las formas de evaluación remite a los aspectos cuantitativos, que generalmente se obtienen a través de pruebas estandarizadas. En 1994 un grupo de investigadores de la Universidad Pedagógica Nacional estandarizo el Test Diagnóstico en Habilidades Básicas (TEHBA) para los tres primeros grados de Primaria. La adquisición y dominio de las llamadas habilidades básicas (Lectura, Escritura y Matemáticas Elementales), se plantea como un requisito indispensable para los habitantes de países occidentales.

El aspecto más importante que debe contemplar un instrumento de medición es la Validez, entendido por esto el grado en que una prueba mide realmente lo que se desea medir. Existen diferentes formas de validación de un instrumento, una de ellas es la Validez por diferencia de grupos.

El objetivo de la presente investigación fue realizar la validación por el método de diferenciación de grupos del TEHBA de primer grado, a fin de dar mayor soporte psicométrico a las tres pruebas que la componen.

En el primer capítulo se describen aspectos generales de la evaluación, su definición, los propósitos de la evaluación, las diferentes formas de clasificación y los requisitos que debe cumplir todo proceso evaluativo.

En el siguiente capítulo se abordan los diferentes tipos de instrumentos de evaluación psicométrica que existen, iniciando con los antecedentes que dieron lugar a esta técnica, su

clasificación, describiendo las características principales de las pruebas de aptitudes y rendimiento.

En el tercer capítulo se plantean las características psicométricas de todo instrumento estandarizado: niveles de medición, error de medida, tipos de Confiabilidad y de Validez. Así como el procedimiento para obtenerlos.

En el cuarto capítulo se presenta una descripción general del Test Diagnóstico en Habilidades Básicas, el procedimiento que se siguió para su construcción y estandarización, así como las características más relevantes de las pruebas.

Se continúa con la metodología que se siguió para la validación por diferencia de grupos del TEHBA de primer grado: se procedió a la selección de dos grupos conformada por 71 sujetos en total. Un grupo que estuviera cursando el primer grado de primaria regularmente y el otro grupo compuesto por alumnos que estuvieran cursando en el programa de USAER para niños repetidores del primer grado. Con los datos codificados se procedió a estimar si había diferencia en los grupos por medio de la prueba de Mann-Whitney.

Posteriormente se presentan los resultados obtenidos y el análisis de los mismos.

Por último se presentan las conclusiones relativas a la investigación, así como las limitaciones del mismo.

CAPITULO I

EVALUACION EDUCATIVA

La calidad de la educación se supera de manera permanente en función y como resultado de la actualización de los planes y programas de estudio, del mejoramiento de las técnicas, los currícula y los procedimientos aplicados y/o utilizados en la escuela; y en ésta, los progresos alcanzados por la Ciencia de la Educación y las ciencias auxiliares que coadyuvan en ese propósito, se manifiestan en y con el perfeccionamiento de la planeación educativa y de los sistemas de organización escolar y docente, de conformidad con los principios teóricos de la Psicología del niño y del adolescente.

Una de los campos en que mayor penetración han tenido las practicas de evaluación es el educativo. Si bien las prácticas evaluativas es posible rastrearlas a través de la historia desde el año 2000 a.C., su delimitación como un **campo profesional específico** con diversas propuestas teórico/metodológicas STUFFLEBEAM Y SHINKFIELD (1987) lo ubican en los primeros años de la década de los setenta.

LA EVALUACIÓN EDUCATIVA en los años cincuenta fue definida como el proceso de determinar la magnitud en la cual los objetivos educativos habian sido alcanzados, época en que la valoración (assessment) y evaluación (evaluation) aparecían como sinónimos. Entre los años sesenta y principios de los setenta la evaluación fue definida como el proceso de proporcionar información para la toma de decisiones. Más recientemente la evaluación se concibe como el proceso para determinar el valor o mérito de una actividad, programa, persona o producto (TENBRINK, 1988)

El desarrollo de la evaluación educativa a partir de los años ochenta a permitido una nueva identidad profesional a los evaluadores, por lo que varias Universidades han abierto diversos programas de posgrado para su estudio y formalización.

A RALPH W. TYLER se le reconoce el mérito de ser el primero en formular una propuesta evaluativa más integral; con los ya clásicos estudios que realizó entre los años treinta y cuarenta se inicia un profundo y polémico debate en torno a la evaluación educativa, término que él introduce al campo educativo.

Antes de que TYLER publicara su modelo evaluatorio los estudios sobre el tema se centraban en el estudiante y los logros que éste alcanzaba en la escuela. Con su propuesta proporcionó los medios prácticos para la retroalimentación de las prácticas educativas. La evaluación debería proporcionar un programa personal con información útil que pudiera permitir la reformulación o redefinición de los objetivos planteados. En este sentido asienta:

“El proceso de evaluación significa, fundamentalmente, determinar en qué medida el currículo y la enseñanza satisfacen realmente los objetivos de la educación. Puesto que los fines educativos consisten esencialmente en cambios que se operan en los seres humanos, es decir, transformaciones positivas en las formas de conducta del estudiante, la evaluación es el proceso de determinar en que medida se consiguen tales cambios” (TYLER, 1973 p. 109)

Este autor enfatizó una amplia gama de objetivos educativos, como los currículos y los apoyos pedagógicos. Destaco la necesidad de establecer, clasificar y definir los objetivos en términos de rendimiento, como una etapa inicial del estudio evaluativo. En este sentido, la evaluación se convirtió en un proceso para determinar la comparación entre estos objetivos y las operaciones.

Consideraba que la evaluación debe de determinar la congruencia entre trabajo y objetivos. Concluyó que las decisiones acerca de los programas debería estar basadas necesariamente en la coincidencia entre los objetivos del programa y sus resultados.

La principal recomendación de TYLER era que el profesor, administrador o el que diseñaba el currículo debería emitir juicios racionales acerca de las áreas programáticas. Para él la evaluación debía ser una etapa esencial de este desarrollo curricular, además consideraba que debía existir una serie de etapas lógicas que afinzaran el proceso de evaluación.

Su procedimiento consiste en el diseño del programa en que debe definirse con claridad las metas y los objetivos de enseñanza para producir aprendizajes en los estudiantes (definido el aprendizaje como cambios en los modelos de comportamiento), formula que requiere de métodos que midan los alcances del cambio, sugiriendo las pruebas objetivas, aunque no agota el problema en este único indicador

"Puesto que todo programa de enseñanza tiene diversos objetivos, y puesto que casi todos ellos admiten varios puntajes o términos descriptivos para sintetizar la conducta de los estudiantes en relación con el objetivo, es lógico que los resultados que proporcionen los instrumentos de evaluación no serán un puntaje único ni un termino descriptivo singular, sino un perfil analizado o un conjunto amplio de términos descriptivos que pongan de manifiesto el rendimiento de los estudiantes." (TYLER, 1973, : 123)

insistía en la validez de los instrumentos utilizados.

TYLER proporcionó medios prácticos para la retroalimentación (término que el introdujo al lenguaje evaluativo) . La evaluación debía proporcionar un programa útil que permitiera la reformulación o redefinición de los objetivos. Su método sería utilizado principalmente por profesores y diseñadores del currículo, aunque también admitía a los evaluadores profesionales (STUFFLEBEAM Y SHINKFIELD, 1987).

Su concepto de evaluación apuntaba al mejoramiento

“La información derivada del proceso evaluativo racional puede producir buenos efectos, esto es, un perfeccionamiento de la educación” (TYLER, 1974, p.126)

El auge que por los años sesenta tuvo el Conductismo en los Estados Unidos imprimió toda una orientación al modelo tyleriano; la concepción de esta corriente tendió a esquematizar mucho una alternativa con muchos matices.

A decir de STUFFLEBEAM Y SHINKFIEL (1987) el modelo tyleriano ha sido aplicado con una estrechez de miras ya que hay diferencias importantes en la aplicación del método tyleriano y lo postulado por el autor. Comentan que una propuesta que introdujo el concepto de retroalimentación, se haya utilizado para valorar el logro final. También señalan que el modelo promete atender tanto a los procedimientos de enseñanza como a los estudiantes, y que los evaluadores atendieron sólo los efectos producidos en los estudiantes.

Otro autor que vino a revolucionar las prácticas evaluativas fue MICHAEL SCRIVEN, filósofo de la ciencia australiano. Trabajó en varias universidades en los Estados Unidos entre los años 1952 y 1982.

SCRIVEN ha sido, quizá, el teórico más propositivo y el crítico más agudo en el campo evaluativo.

Una buena parte de la obra de SCRIVEN no está publicada. La siguiente descripción de sus postulados se basa en algunas de sus trabajos editados y en el texto de STUFFLEBEAM Y SHINKFIELD (1987). Con Stufflebeam trabajó por diez años en diferentes programas de evaluación.

Para SCRIVEN (citado por STUFFLEBEAM Y SHINKFIELD, 1967) la evaluación es una actividad metodológica que consiste en la recopilación y combinación de datos,

mediante la definición de unas metas que proporcionen escalas comparativas o numéricas con el fin de justificar: 1) los instrumentos de recopilación de datos, 2) las valoraciones y 3) la selección de metas.

STUFFLEBEAM Y SHINKFIELD (1987 pp. 343 - 344) comentan que la evaluación para SCRIVEN consiste fundamentalmente en

“... la valoración sistemática del valor o mérito de las cosas”

SCRIVEN diferencia entre la meta de la evaluación y sus funciones. La meta es **juzgar el valor** de algo. En cuanto a las funciones que la evaluación cumple señala dos: 1) la formativa y 2) la sumativa.

1) La función formativa tiene un propósito de mejoramiento. Esta proporciona información continua para ayudar a planificar y producir el objeto. Tiende a ayudar al personal responsable a perfeccionar cualquier cosa que se este desarrollando.

Para la función formativa recomienda que la evaluación la realicen - principalmente - el personal encargado de trabajarla - con el seguimiento de un evaluador externo

2) La función sumativa pretende calcular el valor del objeto una vez que ha sido desarrollado. Con base en el juicio de valor emitido por los evaluadores acerca de un programa o producto, se puede ayudar a los responsables de tomar decisiones a sacar conclusiones en relación a un programa o producto ya concluido.

En el caso de la función sumativa sugiere que debe de ser practicada por un equipo de evaluadores externo, a fin de aumentar su objetividad. Recomienda que los resultados de las evaluaciones se hagan públicos.

Para este autor la evaluación constituye una constatación o estimación del valor de la enseñanza, considerada no sólo en sus resultados, sino también en su proceso de desarrollo, por ello insistió en diferenciar la evaluación sumativa y evaluación formativa.

La primera se centra en el estudio de resultados, mientras que la segunda constituye una estimación de la realización de la enseñanza y contiene en sí el importante valor de poder servir para su perfeccionamiento al facilitar la toma de decisiones durante la realización del proceso didáctico.

También distingue entre evaluación intrínseca (al propio proceso) y evaluación final (los resultados del proceso). La primera evalúa las cualidades de una mediación, sin tener en cuenta los efectos sobre los usuarios (metas, estructura, metodología, cualificación y actitudes del personal, recursos). La segunda no se ocupa de la naturaleza del programa, de los recursos, etcétera, sino más bien de los efectos del objeto evaluado sobre los usuarios.

Producto de sus formulaciones diversos autores se abocaron a diseñar instrumentos para la evaluación que tomaran más en cuenta el aspecto formativo (CHADWICK, 1979; BLOOM, HASTINGS Y MADAUS, 1975).

Estos instrumentos permitieron realizar evaluaciones de tipo diagnóstico y los resultados que arrojaban eran comparadas con determinados criterios establecidos. Con los años se fueron diseñando pruebas con base en normas y con base en criterios (TENBRINK, 1988).

En el país, el desarrollo teórico/metodológico/técnico en el campo de las ciencias humanas muestra retrasos importantes, en comparación con otros países. Muchas razones se han esgrimido para tal situación; la psicología aplicada al campo educativo no ha sido la excepción.

Aún está por escribirse la historia de las técnicas e instrumentos psicológicos en nuestro país; al intentar una primera sistematización de la bibliografía y documentos oficiales que tratan acerca de los instrumentos para la evaluación psicológica, se ha podido constatar lo escaso del material publicado.

1.2 PROPOSITOS DE LA EVALUACION EN EL AULA

Si se explica el concepto en otros términos, se tiene que, como la educación es parte fundamental del proceso educativo, la que a su vez contiene procesos y funciones concretas y precisas que se tratarán más adelante, queda claro cuál es y en qué consiste su propósito principal: encauzar la conducta del alumno hacia determinadas direcciones, por lo que la evaluación educativa se constituye en una estructura básica, una etapa esencial y parte dependiente del proceso de enseñanza-aprendizaje; por lo tanto, las direcciones hacia las que apunta la educación están contenidas en los objetivos que la escuela define, determina y establece previamente; esta circunstancia propicia la vinculación de la educación con los propósitos de su proceso general, llevándola, conforme al criterio de MEHERENS Y LEHMAN (1982, p. 21), a relacionar sus propósitos y funciones con los siguientes propósitos particulares:

- 1o. Identificar y definir los objetivos conforme a los cambios deseados en la conducta del alumno;
- 2o. Planear y dirigir las experiencias del aprendizaje, armonizándolos con los objetivos educacionales generales establecidos;
- 3o. Determinar el grado de progreso del alumno en relación con los objetivos educacionales; y
- 4o. Utilizar los resultados de la evaluación para mejorar y perfeccionar la enseñanza y la instrucción.

Además, del propósito principal y de los propósitos particulares se derivan los siguientes propósitos específicos de la evaluación:

i) Determinar el nivel de conocimientos y otros grados de avance del educando durante el periodo de estudios que se evalúe;

ii) Preparar y proporcionar una base adecuada para la asignación de calificaciones;

iii) Preparar y proporcionar un estudio para conocer las dificultades individuales y colectivas de los estudiantes, que sirva de punto de referencia para diseñar el futuro aprendizaje; y

iv) Propiciar, facilitar y estimular el aprendizaje en el alumno.

Por lo que respecta a las funciones principales de la evaluación educativa, que derivan de los propósitos que arriba se detallan, se exponen las siguientes;

1a. Verificación del cumplimiento de los objetivos establecidos: mediante un proceso gradual, continuo sistemático y acumulativo, la evaluación verifica constata los cambios que la enseñanza produce en el alumno.

2a. Prognosis de las posibilidades educativas del alumno: esta función permite plantear conjeturas acerca de las perspectivas educacionales del estudiante, elabora o presenta un pronóstico sobre las metas educativas a alcanzar; asimismo, conviene identificar las fuentes o causas probables y/o aparentes que pudieran estar afectando el cumplimiento de los objetivos, a fin de modificar o perfeccionar el método o la actividad de la enseñanza.

3a. Diagnósis de la problemática del proceso de enseñanza-aprendizaje (factores determinantes de su desarrollo); función que permite explorar, descubrir y ubicar las fallas y dificultades de los alumnos; además, es de naturaleza clínica; su carácter integral, su objetividad, continuidad y comparabilidad facilitan a la escuela a los educadores y alumnos, determinar los grados de avance en el cumplimiento de los objetivos previstos.

4a. Dirección del proceso educativo; esta función consiste en orientar y encauzar correctamente dicho proceso, por que una vez localizadas las fallas y dificultades de tal desarrollo, se planearán las actividades académicas tendientes a la orientación y reencausamiento de todo el proceso de la enseñanza.

1.3 CLASIFICACION DE LA EVALUACION

En este sentido, PAWLIK (1979) especifica al decir que "hay autores que han diferenciado dos tipos de evaluación, en donde, de ser utilizadas de modo adecuado, le permiten al maestro obtener un mejor conocimiento de sus alumnos y, por lo tanto, también un mayor y mejor aprovechamiento escolar"; los tipos a que hace referencia PAWLIK y que aquí se retoman, son los siguientes:

1 Evaluación formal.

Esta modalidad contiene todos los elementos necesarios e importantes que la hacen sistemática, como son: listas de comprobación, pruebas escritas, registros, Etc.; por otra parte puede proporcionar a los profesores una información "objetiva" acerca de los campos de conducta observables, que son definidos, determinados y estructurados adecuadamente, sentándose así las bases para la elección de aquellas alternativas más apropiadas para el seguimiento de un propósito educativo que se esté realizando

2 Evaluación informal.

En cambio esta otra modalidad se centra en una observación no sistemática sobre cualquier fenómeno que pudiera surgir en el salón de clases, lo que hará que el maestro tome las decisiones más convenientes para la atención y estudio de tal fenómeno; por otro lado, este tipo de evaluación varía en cantidad, además de que no se atiende a la conveniencia de diseñar y establecer controles para su registro que permitan sistematizarla; no obstante, en cuanto a calidad se refiere, contiene y aporta información valiosa utilizable en la toma de decisiones.

Ambas modalidades de evaluación tienen un peso específico importante, en función de que hacen factible que el profesor tenga una oportuna y destacada participación en el proceso de enseñanza-aprendizaje.

1.4 REQUISITOS DEL PROCESO DE EVALUACION

De conformidad con los criterios de GEORGIA (1975 : 21,22), la realización del proceso de evaluación debe cubrir los siguientes requisitos:

1.- Determinar y clarificar, prioritariamente, qué se evaluará.

La identificación y definición de los objetivos educacionales, invariablemente estarán en el primer plano de las actividades académicas; pero, por otra parte, la efectividad de la acción evaluatoria depende de una cuidadosa descripción de las partes de los programas de las asignaturas que se tienen que evaluar, y de las cualidades técnicas de los instrumentos de evaluación diseñados y utilizables; consecuentemente, el progreso de la evaluación estará determinado por la habilidad y sensibilidad del maestro para definir los aspectos de la conducta de los alumnos medibles, que se tomarán en cuenta en el proceso educativo.

2.- Seleccionar las técnicas de evaluación más apropiadas, en concordancia con los propósitos dinamizadores del proceso.

La elección se basará en la eficacia de las técnicas para medir la objetividad de sus resultados y la conveniencia de su uso, con apego a los propósitos estimuladores y animadores de la evaluación; pero hay que tener presente que existen técnicas que son apropiadas para determinados propósitos, e inapropiadas para otros.

3.- Definir y emplear varias técnicas complementarias que requiere un proceso de evaluación integral.

Es indudable que la utilización de una sola técnica no permite apreciar los progresos del alumno, en términos de los resultados que arroja cada proceso educativo.

4.- Tener conciencia de las limitaciones de las técnicas de evaluación y de su uso.

Dichos medios varían desde los instrumentos especializados, que se diseñan y desarrollan adecuadamente y cuya eficacia está debidamente comprobada, hasta los métodos más sencillos, como la observación; por otra parte, muchos errores que ocurren en los procesos de evaluación se pueden corregir o eliminar, si se atiende con cuidado al diseño, la construcción y selección de las técnicas que deberán aplicarse; en cambio, otros errores son susceptibles de controlarse si se desarrolla y mejora la capacidad de uso de tales técnicas; y por último, también contribuye en todo este intento la adquisición de capacidad y experiencia para interpretar y analizar los resultados.

5.- Considerar que la evaluación es un medio que persigue un fin, no es el fin en sí mismo.

La aplicación de las técnicas de evaluación asegura, por un lado, que el proceso es útil y, por el otro, que el educador conoce tal propósito; en consecuencia, con la evaluación educativa se obtiene, además, información necesaria para la toma de decisiones y medidas educativas. Lo antes expuesto implica que:

1o. Los procedimientos de evaluación se seleccionan de acuerdo con las decisiones que se piensan tomar, y

2o. Todo procedimiento a seguir tiene que mejorar las medidas de carácter educativo, de orientación o administrativas que se estén aplicando pero que, por sus fallas y resultados, es pertinente modificar y/o adecuar.

Estas decisiones obedecen a la conveniencia de orientar adecuadamente tanto el proceso de enseñanza-aprendizaje, como los contenidos temáticos de los programas de las asignaturas; por ende, la posible y oportuna detección de alguna alteración que limite o interfiera el normal desarrollo del aprendizaje escolar, adquiere mayor importancia.

1.5 EVALUACION DE LAS HABILIDADES BASICAS

Hay que entender por habilidades básicas al conjunto de conocimientos escolares elementales que sirven de apoyo a la formulación de nuevos aprendizajes, expresados en planes y programas de estudio.

Desde este contexto es que, no sólo el lenguaje escrito, o el conocimiento y dominio, de las operaciones matemáticas básicas, son una exigencia que sistemática y universalmente se plantea, en el sentido de que el niño que asiste a la escuela debe contener entre sus características principales, sino que es toda una serie compleja de conductas predeterminadas y elaboradas que se manifiestan o que, a veces, permanecen ocultas en el niño; esta conceptualización no es producto del azar, sino resultado de diversos factores sometidos a muchas influencias pero que es fundamental si se pretende que el sujeto de todo proyecto educativo, alcance los niveles de desarrollo general.

A.- LECTURA

Se explica como un proceso muy complejo que involucra diferentes aspectos de y en la evaluación; también, en ocasiones es considerado como un acto relativamente placentero pero que, para tal efecto, es necesario que se tenga un control acerca de las habilidades básicas para la lectura.

El acto de leer es una actividad compleja, también, en función de que presenta innumerables referentes; en la actualidad y en términos reales, ninguna prueba diagnóstica es tan útil para evaluar en su totalidad los aspectos que comprende el acto de la lectura.

Las pruebas, a lo sumo, sólo muestran ciertas conductas específicas de la lectura, o que tienen relación con ella; a propósito de este punto de vista, SALVIA (1988 : 167), sostiene que "las conductas particulares evaluadas por cualquier prueba, son aquellas cuya evaluación los autores consideran como más importantes".

En términos específicos, en las pruebas de lectura se solicitan ciertas conductas, entre las que se destacan: la identificación de letras y palabras, lectura mecánica, lectura en silencio y comprensión.

Pero también, en la evaluación del acto de la lectura se subrayan los errores más comunes, entre los que sobresale: la no lectura, el error grueso de pronunciación, la omisión, la sustitución, la adición, la inversión, el error parcial de pronunciación y aspectos gramaticales.

B.- ESCRITURA

Está considerada como un acto esencialmente creativo, que propicia y facilita la expresión subjetiva, pero que, para poder alcanzarla, se requiere tener un adecuado dominio de las habilidades y aptitudes básicas de la lectura; dicho acto es complejo, puesto que involucra varios aspectos en su evaluación; entre las conductas que se solicitan para su medición en las pruebas de escritura, destacan: el copiado, el dictado y la redacción.

Pero, por otra parte, en la evaluación que se efectúa de la escritura, se pone mayor énfasis en los errores que enseguida se especifican: la posición del sujeto, la forma de tomar el lápiz, la posición de la hoja de papel, la presión de la escritura, la uniformidad, la inclinación, los espaciamentos, los márgenes, la ausencia de la respuesta respectiva, la no identificación de letras, la omisión, la adición, la inversión, la sustitución, la ortografía, la sintaxis y la puntuación.

CAPITULO II

INSTRUMENTOS DE EVALUACION PSICOLOGICA

2.1 ANTECEDENTES

Es de suma importancia dar un espacio a la " historia " de los instrumentos de medición en Psicología, ya que, con ello - consideramos - ayudará a una mejor comprensión de la manera en que se fueron dando este tipo de técnicas, así, como a las diferentes polémicas a que se dio lugar. Para esto se siguieron de cerca algunos planteamientos que desarrollo FERNÁNDEZ en su texto: Psicodiagnóstico. Concepto y Metodología (1981).

Los antecedentes de las técnicas en Psicología, los podemos citar desde tiempos atrás. MC REYNOLDS citado por FERNÁNDEZ (1981);menciona que la obra más importante de la antigüedad, en relación con la evaluación psicológica, es el tetrabiblos perteneciente a PTOLOMEO. En el capítulo XIII de este texto se encuentra una relación detallada entre los signos del zodiaco y los rasgos psicológicos implicados en ellos. El se basa en una serie de argumentos para considerar el horóscopo como una evaluación psicológica:

- 1) Hay un sujeto a evaluar, un evaluador y unos protocolos, así como la interpretación de estos;
- 2) Existe un " marco " teórico en el cual se hace mención de hipótesis deterministas;
- 3) Hay también un sistema de categorías taxonómicas, a las cuales se hace referencia en el momento de clasificar a las personas.

El tener el signo de Libra o Leo hace que se tengan ciertas características prefijadas, que van a determinar lo que suceda en el transcurso vital del individuo (FERNANDEZ, 1981 : 18).

El Horóscopo según MC REYNOLDS, es la forma de evaluación psicológica más antigua y duradera de todos los tiempos.

Por otro lado DU BOIS menciona cómo, en la dinastía Han China (220 años a.C. a 206 d.C.), se ocuparon prácticas evaluatorias para seleccionar a los aspirantes a altos puestos de gobierno, a través de entrevistas estructuradas y rudimentarios cuestionarios (FERNANDEZ, 1981 : 19).

Siempre es difícil ubicar los inicios de cualquier disciplina y de las técnicas que utiliza. En Psicología por ejemplo son muchas y variadas sus fuentes. Continuando con la clasificación que hace FERNANDEZ (1981) se reconocen cuatro influencias determinantes en lo referente a los instrumentos de evaluación en Psicología: La Filosofía positivista, la Medicina en especial, la Psicofísica, la Psiconeurología y la Psiquiatría; la Matemática en lo referente a la probabilidad y estadística y la Biología con los estudios de LAMARK Y DARWIN.

La etapa que caracteriza la autora como " el período científico " de la Psicología, se comienza con los trabajos que WEBER publica en 1834, los cuales se refieren a la medida de los umbrales diferenciales a diferentes niveles de estimulación, que llevarán a una nueva disciplina: La Psicofísica.

FECHNER en 1860, tomando en cuenta los trabajos de WEBER; publicó sus Elementos de Psicofísica, en donde intenta demostrar la posibilidad de medir un aspecto subjetivo. Para lo cual se basó en el postulado de que " la sensación no puede medirse directamente, pero si se puede tenerse una idea aproximada de ella a través del autoinforme " (FERNANDEZ, 1981 : 25). Lo importante de la aportación de FECHNER se refiere a la realización de métodos específicos de presentación de estímulos y medición de las respuestas, con lo cual por primera vez se utiliza la lógica y los métodos de medición positiva en la Psicología.

FERNANDEZ (1981), piensa que la Psicofísica contribuyó a la consecución de la Psicometría, disciplina al servicio de los estudios sobre diferenciación individuales.

Un segundo momento importante para el desarrollo de los instrumentos de evaluación, lo forman los estudios pioneros de BERNOUILLE, de MOIVRE, LAPLACE, continuados durante el

S.XIX por Quetlet, que conformaron las bases de la actual Psicología Matemática. Adolfo Quetlet según Fernández, tuvo una decisiva influencia no sólo en Psicología, en otras disciplinas como la Sociología y la Economía. Su primordial aportación fue demostrar la posibilidad de aplicación de los métodos estadísticos al estudio del comportamiento humano.

La otra gran influencia para el desarrollo de los instrumentos de medición, se puede ubicar en los avances habidos durante el S.XIX en la Neurofisiología. Burdach (1776 - 1847), hizo un importante trabajo en su época con la disección de más de mil cerebros de enfermos mentales, en donde, en la sustancia gris se encuentran las facultades superiores y en la blanca la actividad motriz. Por otro lado Marc Dax y Broca hicieron estudios sobre las bases cerebrales del lenguaje y Ferrier logró descubrir una serie de centros psicomotrices (FERNANDEZ, 1981 : 26). Este empeño biólogo suscita una escisión de la Psiquiatría, como muestra Bondy (FERNANDEZ, 1981 : 27) en dos corrientes antagónicas: la " organicista " en donde se manifiesta un afán clasificatorio y la Psicologista en donde se encuentra la búsqueda de procedimientos evaluativos de los padecimientos mentales.

Otro paso más lo da Von Grshley (1839 - 1914), con el diseño de tres tipos de tareas (escritura, dibujo y lectura) para la exploración sistemática de los post-traumatismos craneales. Por otro lado, Rieger (1855 - 1939), establece una batería de ocho subtest (percepción, imitación, memoria asociativa, reconocimiento, etc.) para la exploración de pacientes con lesiones cerebrales (FERNANDEZ, 1981 : 28).

La cuarta corriente de pensamiento que se puede reconocer, según Fernández, es la de los filósofos y biólogos, preocupados por la teoría evolucionista formulada a fines del S.XVIII, que culminan con los trabajos de Charles Darwin en el S.XIX. Lamarck había esbozado algunas ideas referentes a la evolución, principalmente a la herencia de caracteres adquiridos y la modificación por selección natural, la supervivencia del más apto y de la continuidad en la evolución de las especies, así como la idea de que los cambios genéticos podían producirse al azar.

Como menciona Boring, después de la teoría evolucionista " se desarrollo toda la tarea de estudiar las diferencias individuales en las capacidades mentales y de establecer niveles

psicológicos por medio de los test mentales " (FERNANDEZ, 1981 : 28). Varias de las críticas que le han formulado a las pruebas psicológicas se basan en su pretensión diagnóstica, que remite al proceder médico y su afán clasificatorio.

En cuanto a los padecimientos físicos, fueron apareciendo alternativas terapéuticas de acuerdo con el tipo de enfermedad. Sin embargo en Psicología no siempre ha ocurrido así, las alternativas terapéuticas en el campo psicológico, no siempre han tenido la rapidez ni la efectividad que se hubiera deseado, así como tampoco hubo la posibilidad de utilizarlas con grandes sectores de la población.

2.2 LOS PRIMEROS INSTRUMENTOS PSICOLOGICOS.

Tres han sido los autores pioneros en el desarrollo de instrumentos de evaluación psicológica: Francis Galton (1882 - 1911), MC. Keen Cattell (1861 - 1934) y Alfred Binet (1875 - 1911).

Francis Galton fue considerado el iniciador de la Psicología diferencial. Se centro en la teoría evolucionista y su aplicación a las ciencias del comportamiento. En 1883 publicó " Inquiries Into Human Faculty and its development ", obra considerada como el principal punto de partida de los test mentales. Desarrolló pruebas de discriminación sensorial y coordinación motora; en Inglaterra estableció el primer laboratorio antropométrico con la finalidad de medir las características físicas y sensoriomotoras de las personas (MORALES, 1975 : 64).

MC. Keen Cattell, fue el que mayor influencia ejerció en su época en la Psicología norteamericana, su obra esta impregnada por el evolucionismo, él tiene el afán del estudio de las diferencias individuales.

Se marcha a Europa para estudiar con Wundt en Leipzig, sus objetivos diferencialistas le condujeron al estudio diferencial de los tiempos de reacción. Es Cattell quien en 1890 acuña el término " test mentales ". Después presenta n trabajo en la "American Psychological Association", en el que continua la obra de Galton y en donde da pruebas sobre la distribución normal de las características psicológicas y acerca del carácter genético de la superioridad intelectual (FERNANDEZ, 1981 : 32).

El tercer iniciador de las técnicas de evaluación psicológica es el francés Alfred Binet, que puede ser considerado como uno de los fundadores de la Psicología diferencial.

Binet junto con Beaurais, funda en 1889, el primer laboratorio de Psicología en " La Sorbone ". El mismo año publica, Psychologie Individuelle. Propone iniciar el estudio de formas más complejas de comportamiento: " no son las sensaciones, son las facultades psíquicas superiores las que hace falta estudiar, son ellas las que juegan el puesto más importante y la Psicología Individual deberá prestarles mucha más atención " (FERNANDEZ, 1981 : 33).

Relacionado con el Psicodiagnóstico, Binet habla de tres grandes problemas de la Psicología individual: 1) estudiar las diferencias individuales de los procesos psicológicos; 2) estudiar las diferencias psíquicas en individuos o en grupos; 3) estudiar las relaciones de diversos procesos psíquicos en un mismo individuo.

Desde 1895 hasta 1903 Binet trabaja conjuntamente con Simón en una colonia para niños deficientes en Perray-Vaucluse, este trabajo da como resultado lo que se puede considerar como el primer test de inteligencia. El supuesto básico de este test consiste en que las capacidades intelectuales se incrementan con el desarrollo.

Con la introducción de las Escalas Binet-Simón, nace el movimiento de las pruebas y se da a partir de la psicometría, el desarrollo del psicodiagnóstico, siendo paralelo al nacimiento de la Psicología Clínica (MORALES, 1975 : 52).

En 1904 el Psicólogo Inglés Spearman, desarrolló una teoría de dos factores de la inteligencia, el factor G (general) y el factor S (específico), sentando las bases para lo que se conoce como el análisis factorial (CERDA, 1978 : 21).

En las dos primeras décadas del S.XX en diferentes países se extiende el intento por encontrar medidas psicológicas semejantes a las mediciones de las ciencias fisiconaturales, donde es posible reconocer dos planteamientos. Por un lado, la psicología experimental busca las leyes o principios generales que regulan los fenómenos psicológicos y , por la otra, la psicología diferencial o correlacional, está interesada en constatar las diferencias individuales.

Ya en 1912, los avances en el campo de la psicometría eran amplios. En este año W. Stern introdujo el concepto de Cociente Mental o Coeficiente Intelectual (C.I.) (ANASTASI, 1977 :12).

En la Universidad de Yale en 1919 Arnold Gesell, ideó una escala de desarrollo de las primeras edades, la Escala de Madurez, basándose en observaciones de niños y la detección de pautas de

comportamiento son muy importantes en el desarrollo infantil.

En 1938, Laureta Bender realizó la prueba Gestáltica Visomotora que se basa en los principios de la Teoría de la Gestalt, que fue desarrollada por Wertheimer, Kolher y Kofka en Alemania (SATTLER 1988:252).

En 1939, aparece la escala Wechsler - Belleuve, iniciándose por primera ocasión, la estandarización de una prueba para adultos en función de su edad cronológica (CERDA, 1978:51).

Con el arribo de los primeros psicoanalistas en los Estados Unidos tienen influencia en el diseño de pruebas basadas en los postulados de la Psicología del Yo, algunas de éstas son el Rorschach en 1921, el Test de Apreciación Temática (T.A.T.), en 1941 la prueba de frustración de Rosezweg 1944, principalmente.

Por otro lado la obra de Piaget y su escuela, tuvieron influencia en el desarrollo de instrumentos psicológicos, mencionando el diagnóstico del pensamiento podría ser de mayor valor que el C.I.

2.3 CLASIFICACION DE LOS INSTRUMENTOS DE EVALUACION PSICOLOGICA.

La palabra " prueba " se deriva del latin testis, que quiere decir test, que ha sido utilizada en diferentes países del mundo (FORNS, 1980 : 28).

Diferentes definiciones se han propuesto del concepto de prueba. Se muestran tres de ellas como ejemplo:

1.- Una prueba se puede considerar " como un conjunto de tareas o preguntas destinadas a producir cierto tipo de conductas que al ser presentadas bajo condiciones estandarizadas producen puntuaciones que tienen propiedades psicométricas deseables " (SALVIA,1987 : 210)

2.- Una prueba es considerada " un instrumento de experimentación que permite observar cierto tipo de conductas y medirlas en forma objetiva " (FORNS, 1980 : 28).

3.- Una prueba tiene como objeto " determinar tanto el desarrollo mental de un individuo, como el nivel de su saber, la existencia y el grado de una aptitud dada " (COLLIN, 1959 : 32).

En diferentes décadas se han utilizado diversas denominaciones para describir lo que un sujeto es capaz de hacer, tales como, capacidad, inteligencia, aptitud, actitud, madurez, personalidad, etc., por lo cual se han creado diferentes instrumentos. Aunque se encuentran diferencias en las pruebas que se aplican, estas se marcan tanto por el tipo de conducta que se desea evaluar, como por la teoría en que se sustentan.

ARY y CHESER (1990) establece una diferencia entre pruebas, inventarios, técnicas proyectivas y escalas. Las pruebas son utilizadas para evaluar la capacidad del sujeto para percibir relaciones, resolver problemas y aplicar el saber en una variedad de contextos. Por otro lado el inventario viene de las teorías de rasgos, de la tipología y de la motivación. Las técnicas

proyectivas funcionan interpretativamente con base en estímulos ambiguos o no estructurados.

Por último, la escala esta referida a un conjunto de valores numéricos asignados a sujetos, objetos o comportamientos con el propósito de evaluar sus calidades.

Las pruebas se aplican por lo general en ambientes educativos, con una diversidad de propósitos. Su objetivo es dar a los alumnos/as, padres, maestros/as, psicólogos/as educativos y demás profesionistas, información que ayude a que tomen decisiones que mejoren el desarrollo educativo de los estudiantes.

SALVIA (1987) reconoce por lo menos seis razones específicas para aplicar pruebas:

1) SELECCION.

Las pruebas se pueden administrar para localizar estudiantes que son diferentes (tanto en sentido positivo, como negativo) de sus compañeros de edad que necesitan atención especial. Un ejemplo es la prueba que se les hace de un nivel escolar a otro.

2) UBICACION.

A veces se necesita que un especialista evalúe al alumno, antes de que empiece o abandone programas educacionales especiales.

3) PLANEACION DEL PROGRAMA.

En repetidas ocasiones las pruebas se aplican para ayudar a maestros y directores/as en la planeación de programas educativos para individuos o grupos de estudiantes. La información que dan las pruebas se utiliza para designar el lugar en grupos, o para asignar a los estudiantes a programas específicos compensatorios o correctivos. O también para decidir que enseñar o como hacerlo, tanto en individuos como con grupos.

4) EVALUACION DEL PROGRAMA.

Aquí ya no es tanto el estudiante el que se evalúa, sino el programa y aquí se aplican las pruebas tanto al principio como al final del año para valorar el progreso de los alumnos/as en el programa.

5) EVALUACION DEL PROGRESO INDIVIDUAL.

Está es otra razón para aplicar pruebas a estudiantes y es para ver su progreso a través de los sucesivos grados (estudio longitudinal)

6) INVESTIGACION.

Las comparaciones entre los grupos, instituciones comunidades y regiones, requieren de la aplicación de instrumentos válidos y confiables.

En la elección del tipo de prueba que se va a utilizar, es importante que se tome en cuenta el objetivo que se pretende alcanzar, también lo son las características técnicas de la prueba y la medida de las conductas muestreadas que representan.

Las pruebas se pueden clasificar de diferentes maneras, y esto es en base a lo que cada autor mide. Por ejemplo: de acuerdo a los procedimientos de aplicación, como las pruebas individuales o colectivas, también pueden ser con base en las instrucciones, como las pruebas verbales o escritas; o hay unas que se pueden sustentar en los propósitos de la prueba ó el uso de los resultados, como las basadas en normas ó las sustentadas en criterios.

Como hay mucha complejidad en las conductas, personalidad y aptitudes que se estudian de un sujeto a través de las pruebas, es muy difícil establecer clasificaciones específicas.

A finales de los años setenta se plantea un debate que pretende delimitar la validez de cada uno de los términos que pretenden describir los resultados de las pruebas, acordando aceptar solamente dos de ellos para los aspectos cognoscitivos: 1) aptitud que refiere los aprendizajes de

un sujeto a través de su vida; 2)rendimiento, referido a los aprendizajes limitados a un programa de instrucción Ebel, (1980).

De hecho en los libros editados sobre el tema, posterior a los años sesenta, las pruebas se clasifican en: Pruebas de Aptitudes, Pruebas de Rendimiento y Pruebas de Personalidad (Nunnally y Bernstein, 1995; Brown, 1991; Meherens y Lehman, 1982). A continuación se describen sus características generales.

1.-PRUEBAS DE APTITUDES

- a) Generales
- b) Múltiples
- c) Especiales

2.- PRUEBAS DE RENDIMIENTO

- a) Diagnósticas
- b) De materiales aisladas
- c) Baterías de evaluación

3.- INVENTARIOS DE PERSONALIDAD, PRUEBAS PROYECTIVAS Y ESCALAS DE ACTITUDES.

PRUEBAS DE APTITUDES

Se inicia presentando las pruebas de aptitudes, ya que son las que más se han utilizado en el campo educativo en México. Con el concepto de "aptitud" se pretendió detener - acordando un término común-una polémica que por décadas se tuvo en relación con el significado y la etiología de la inteligencia (adquirida/heredada). En los últimos noventa años han estado en la escena una serie de interrogantes relativas a la capacidad de las personas: ¿Es posible definir la palabra capacidad?; ¿es medible la capacidad? ¿tiene una persona una capacidad general para adquirir como conocimientos, o acaso existen muchas diferentes capacidades, cada una de ellas específica

a un tipo determinado de conocimientos?. Los psicólogos han utilizados varias denominaciones para identificar un concepto que consideran es de utilidad en la predicción de varios tipos de conducta: capacidad, inteligencia, potencial y aptitud.

El término de inteligencia -todavía- sigue siendo problemático. Por una parte se le concibe como una sola función global (factor g); por ejem. Terman se concentró en el pensamiento abstracto, sostenía que era la parte esencial de la inteligencia. Para otros como Binet, la inteligencia se refiere a un conjunto muy variado de cualidades como el juicio, sentido común, iniciativa y adaptación (Sattler, 1988).

En cuanto al significado de la inteligencia, se mencionan tres orientaciones:

- a) medida de capacidad
- b) conducta observada
- c) es lo que miden las pruebas específicas de habilidad cognoscitiva

En el fondo de las diferentes propuestas, de las diferentes propuestas esta presente el aspecto etiológico de la inteligencia: heredado, adquirido y/o ambos factores; así como concepciones teóricas acerca de ésta. Parece que salvo muy contadas excepciones, la idea de un innatismo exclusivo de la inteligencia, ya no se considera (Meherens y Lehman 1982; Anastasi, 1977).

Convencionalmente se utiliza el concepto de aptitud para referir se a diferentes tipos de pruebas, ya que éste no remite como el de inteligencia - al problema innatista. Al respecto Fergguson indicaba que no había detrás del estudio de la capacidad humana ninguna teoría sistemática capaz de generar hipótesis fructíferas acerca de la conducta de la conducta pragmáticamente al debate (Ebek, 1980).

Thorndike (1978 p. 37) apela al término habilidad como un sinónimo de función. Había tres ideas en torno a lo que miden las pruebas de aptitudes: a)rendimiento una tarea, b)diferencia individual, c)disposición natural. Una definición amplia del concepto de aptitud implica

“ la capacidad de un sujeto para realizar con éxito actividades que se desarrollan a nivel práctico y/o simbólico “

snow (1980) recomienda distinguir entre aptitud como constructo teórico y aptitud como indicador que ofrece una prueba.

De las primeras pruebas psicológicas que se diseñaron para evaluar aspectos conductuales se basan en buena medida en constructos teóricos como madurez, inteligencia, capacidad etcétera ejemplo de éstos son las pruebas de Gessell y de inteligencia de Weschler. A éstas se les denomina pruebas de aptitudes generales y pretenden medir aspectos amplios de la conducta humana.

Por su parte las pruebas sensoriomotoras, de capacidad numérica mecánica, memoria, atención, etcétera tienden a comprobar la presencia, o no, de estas aptitudes con el objeto de determinar la capacidad del sujeto para la realización de actividades, por ejem. De este tipo de instrumentos son el Test de Bender y el Test de Frostig. A éstas se les denomina pruebas de aptitudes específicas.

Por último, las pruebas de aptitudes especiales suelen definirse como la capacidad potencial de un individuo en un tipo muy específico de actividad, como las pruebas de vista y audición; de aptitud mecánica, de aptitud musical y artística, creatividad y todas las pruebas de razonamiento (verbal aritmético, etc.).

las características generales que se reconocen en todas las clasificadas como pruebas de aptitudes son:

a) Evalúan aspectos amplios de la historia de un sujeto.

Las pruebas de aptitudes pretenden valorar el resultado de la experiencia de vida de un individuo, en la que se incluyen: capital cultural de los padres, oportunidades culturales del niño, aspectos sociales, factores orgánicos, etc.

b) Se basan en la estabilidad de las características a evaluar.

El propósito de los resultados de las pruebas de aptitud es predecir futuras ejecuciones con base en la inteligencia o aptitudes específicas, por tanto, parten del supuesto de que dichas características se mantienen relativamente estables durante el tiempo. Existen un acuerdo casi general con respecto a que las pruebas de aptitudes que se aplican a niños (as) pequeños (de los 2 a los 12 años) tienen una variabilidad mayor que aquellos resultados que se obtienen de sujetos mayores, que atienden a ser más estables (Sattler, 1988).

PRUEBAS DE RENDIMIENTO

Las pruebas de rendimiento son instrumentos diseñados para valorar los aprendizajes referidos a los contenidos de un programa de estudios. Los aprendizajes que un sujeto va construyendo a través de su experiencia en la escuela incluyen aspectos cognoscitivos, motrices, actitudinales y valorativos. Las pruebas de rendimiento solamente valoran ciertos procesos relativos al conocimiento.

Los niveles de complejidad cognoscitiva que valoran estas pruebas son: 1) conocimiento; 2) comprensión; 3) aplicación, (Bloom, Hastings, y Madaus, 1975).

- 1) El conocimiento se refiere principalmente al recuerdo. Es la operación cognoscitiva más simple.
- 2) la comprensión se refiere a un tipo de entendimiento tal que el individuo sabe que se le comunica y puede utilizar el material o idea sin relacionarlo necesariamente con otros materiales o advertir todas sus implicaciones.
- 3) La aplicación el uso de abstracciones pueden presentarse bajo la forma de ideas generales, reglas de procedimientos y métodos generalizados.

Son varios los factores que influyen en el rendimiento escolar de un alumno: a) el currículo; b) el ambiente sociocultural del alumno; c) aspectos personales del alumno.

- a) El currículo. La puesta en práctica de un programa de estudios, su calidad y relevancia, la forma como el maestro lo pone en práctica, el tipo de instalaciones, las relaciones que se establecen en el contexto escolar influyen poderosamente en el rendimiento.
- b) El ambiente sociocultural. La escolaridad de los padres, alimentación, ambiente familiar, relaciones con compañeros, son factores que influyen poderosamente en el rendimiento.
- c) Aspectos personales. Otros factores que influyen en el rendimiento escolar son los hábitos de estudio, la autoestima del alumno y su motivación hacia el estudio.

De acuerdo a la clasificación general se presentan en tres formas:

- a) Baterías de evaluación especial, que se refieren a la evaluación que se realiza de todo un ciclo escolar, en países industrializados, es frecuente que profesionistas se sometan a exámenes que certifiquen sus habilidades, un ejem. de esto es el California Achievement Test, estandarizado en 1970 por Tieg y Clark.
- b) De materias aisladas, referido a la evaluación en un curso específico, ejem. la prueba Metropolitan Achievement Test, estandarizada por Durost, Prescott y Balow en 1971.
- c) Diagnósticas, están diseñadas para resaltar las destrezas y habilidades en el desarrollo de sus capacidades, se remite a la detección de aspectos específicos de un programa ejem. la Stanford de diagnóstico en lectura, estandarizada en 1976 por Garden.

- Inventario de Ejecución Académica (IDEA); Construida por investigadores de la Facultad de Psicología de la Universidad Nacional Autónoma de México, coordinados por la Dra. Silvia Macotela (1992). Valora aspectos relacionados con la lectura, la escritura y las matemáticas de los tres primeros grados de primaria.

A pesar de que el IDEA es una prueba que en el sector oficial aún no se utiliza, se ha incluido, ya que evalúa habilidades básicas.

- Test Diagnóstico en Habilidades Básicas (TEHBA); Su diseño fue realizado por la Profra Rosa Ma. González Jiménez, en colaboración con algunos investigadores de la Universidad Pedagógica Nacional en 1994. El TEHBA son una serie de pruebas de evaluación objetiva de los conocimientos y habilidades básicas de estudiantes con dificultades de aprendizaje en los primeros tres grados de primaria.

Las áreas que evalúa son: Matemáticas, Lectura y Escritura. Las habilidades que se exploran van de acuerdo a los objetivos establecidos en los Planes y programas de estudio de la Modernización Educativa.

El diagnóstico por esencia es individual y es adecuado para el individuo que tropieza con dificultades.

El desarrollo de una prueba de diagnóstico se basa en dos postulados:

- 1.- La capacidad de subdividir destrezas o conocimientos, en subdestrezas componentes
- 2.- La capacidad de elaborar items de prueba que permitan una medición válida de estas subdestrezas.

Una buena prueba de diagnóstico debe abarcar las capacidades componentes que son necesarias para una buena ejecución, son comunes, sobre todo en lectura y la aritmética, ya que son habilidades educativas fundamentales que se pueden dividir en habilidades componentes.

PRUEBAS DIAGNOSTICAS EN LECTURA

Estas pruebas proporcionan al maestro un análisis sistemático de las destrezas y habilidades en

lectura. Sin embargo, es importante ver el rendimiento de un alumno a la luz de las conductas muestreadas por dichas pruebas.

La evaluación con referencia a criterios es una práctica de los últimos años de la década de los sesenta. está atada a objetivos instruccionales y se diseñaron reactivos individuales para evaluar la destreza en objetivos específicos. En este tipo de prueba lo que interesa es la consistencia de las respuestas de los reactivos.

PRUEBA DIAGNOSTICA EN MATEMATICAS

Se diseñaron para identificar destrezas y debilidades específicas en el desarrollo de habilidades o áreas particulares.

Se cuenta con menos pruebas de diagnóstico en matemáticas que de lectura, pero su evaluación está mejor definida. Ya que la ejecución correcta de algunas operaciones matemáticas depende de la correcta realización de otras operaciones (ejemplo: la multiplicación depende de la suma), siendo relativamente más sencillo establecer una secuencia de desarrollo de habilidades y evaluación en matemáticas que en lectura.

El conocimiento obtenido a partir de la aplicación de una o más de las pruebas, acerca del dominio que el estudiante tiene de habilidades matemáticas específicas: junto con el conocimiento de la secuencia general del desarrollo de habilidades matemáticas, puede ayudar a los maestros a diseñar el contenido de programas para las personas.

Las características generales que se pueden reconocer en las pruebas de rendimiento son:

a) Evalúan aspectos restringidos del aprendizaje escolar.

A diferencia de las pruebas de actitudes pretenden valorar aspectos de la historia de un sujeto, las de rendimiento se centran en contenidos específicos, como puede ser un programa, un ciclo o un nivel escolar.

b) Son dispositivos de identificación.

No pretenden hacer interferencias acerca del desempeño futuro de un sujeto, sino identificar las habilidades que ha logrado adquirir conforme a un programa específico.

Dependen de las experiencias formales del aprendizaje que se adquieren en la escuela.

c) Las habilidades se pueden especificar en términos conductuales.

d) La interpretación de las pruebas se basan en normas y/o criterios.

Tanto las pruebas de aptitudes como las de rendimiento comparten una serie de características:

a) Son pruebas estandarizadas.

Están elaboradas con métodos para conseguir muestras de conducta uniforme. Estas son diseñadas por grupos de investigadores y expertos en medición por el alto costo material y humano que implica su elaboración (MEHERENS, 1982 : 39). Este tipo de instrumentos cuentan con altos niveles de confiabilidad y validez.

b) Son medidas de ejecución máxima.

Aquí interesa obtener una estimación del mejor rendimiento posible de un individuo, a diferencia de las de ejecución, en donde lo importante es la conducta usual o habitual.

c) Se ocupan de funciones cognoscitivas.

Pretenden una evaluación de las funciones cognitivas, como habilidades, capacidades y conocimientos

d) La interpretación de las pruebas se basan en normas y/o criterios.

Los puntajes " brutos " que se dan en las pruebas, necesitan de una referencia para tener significado, el cual es obtenido en base a normas o en base a criterios.

Es importante planificar el diseño de una prueba, ya que es necesario saber de que tipo de instrumento se trata, para que en base a ello y a los objetivos sabremos que tipo de procedimiento se seguirá en la construcción.

CONSTRUCCION DE LA PRUEBA.

La construcción de las pruebas de rendimiento se basan en el reconocimiento de la necesidad de alguna información educativa en particular. Dicha información puede estar en el área de habilidades básicas (lectura, lenguaje y aritmética) o en un área particular de materia. La necesidad puede ser diagnosticar las dificultades que se presentan para el aprendizaje o determinar el nivel general de aprovechamiento de los alumnos.

Pasos para la construcción de una prueba de aprovechamiento.

1.- Planear la prueba

2.- Preparar los elementos de la prueba

3 - Puesta a prueba experimental y revisión

4 - Administrar la edición que hay que uniformar

La administración de la edición que se obtenga para uniformar, suministra la base para la realización de las tablas de normas; las cuales hacen que sea posible comparar la actuación de prueba de un alumno con los otros grupos de alumnos a nivel nacional, regional o estatal.

PLANEACION DE LA PRUEBA

La planeación sistemática debe incluir la identificación de los objetivos y el contenido de instrucción por medir, determinando de esta manera el énfasis relativo que debe darse a cada uno de los objetivos y a cada área de contenido, así como a elaborar una tabla de especificaciones que sea el reflejo del énfasis que se deseado.

La planeación de estas pruebas incluye el esfuerzo en grupo de maestros, especialistas en planes de estudio y expertos en pruebas. El método para seleccionar el contenido consiste generalmente en un intento por elaborar una prueba que encaje lo más que se pueda en los programas escolares.

PREPARACION DE LOS ELEMENTOS DE PRUEBA

El contenido, el vocabulario y la complejidad de los elementos deben ser apropiados para el curso en particular y para el nivel de grado en especial para los cuales se va a realizar la prueba.

En la preparación está involucrada la redacción de los elementos que miden directamente los

objetivos y el contenido de instrucción que se debe definir en la tabla de especificaciones.

Una vez preparado un número suficiente de elementos, se edita la forma experimental de la prueba o prueba previa. La cual debe incluir instrucciones, formato de prueba, límites de tiempo y estipulaciones para llevar a cabo la calificación de lo que se desee forme parte de la edición final.

REVISION Y PUESTA EN PRACTICA EXPERIMENTAL

La edición experimental de la prueba se administra a grupos de alumnos para los cuales se ideó la prueba. El propósito es obtener datos relativos a los siguientes aspectos:

- 1.- La dificultad de cada elemento de prueba
- 2.- El poder discriminatorio de cada elemento de prueba
- 3.- La efectividad de cada elemento de distracción para cada elemento de prueba
- 4.- La equivalencia de los elementos en las diversas formas de la prueba
- 5.- La adecuación de las instrucciones, los límites de tiempo y el formato de la prueba

La efectividad, el poder discriminatorio y la efectividad de cada uno de los elementos de prueba se determinan mediante el análisis de los elementos. También son de gran importancia los informes proporcionados por los maestros que suministran las pruebas. Ya que sus críticas son útiles para poder mejorar las instrucciones que se les dan a los alumno y para incrementar la efectividad de los otros procedimientos de administración de prueba. Teniendo como resultado, el poder eliminar los elementos que sean demasiado fáciles o demasiado difíciles.

ADMINISTRACION DE LA EDICION PARA UNIFORMAR

Se administra la edición final de la prueba a un grupo representativo de alumnos con el propósito de establecer normas (la confiabilidad u otros aspectos técnicos de la prueba también se determinan en este momento). Las normas hacen posible la comparación de la calificación de la prueba que obtuvo un individuo con la que obtuvo otro (s) individuo (s), cuyas características se conocen.

El momento más importante en el proceso que se lleva a cabo para uniformar, es el de seleccionar un grupo de alumnos que sea representativo de los educandos para los cuales se desea tenga utilidad la prueba.

ELEMENTOS QUE DEBEN CONTENER LAS PRUEBAS

- Instrucciones para administrar y calificar la prueba, los cuales se redactan de manera tan precisa que los procedimientos son uniformes para los diferentes usuarios de la prueba.
- Se proporciona manual de pruebas, guías para calificar y suministrar la prueba, para evaluar sus cualidades técnicas y para interpretar y usar los resultados.
- Normas basadas en grupos representativos de individuos, como auxiliares para interpretar las calificaciones de pruebas.

3.- INVENTARIOS DE INTERESES, PERSONALIDAD Y ACTITUDES.

Aunque este tipo de instrumentos pretenden evaluar los intereses, actitudes, comportamientos, sentimientos, etc., generalmente se les conoce como " pruebas de personalidad ". Las cuales están caracterizadas porque utilizan una gran variedad de métodos para su construcción.

La personalidad, abarca un campo amplio y heterogéneo que incluye una multitud de conceptos y construcciones.

Históricamente, el campo de la personalidad ha sido centro de bastantes teorías de las cuales varias conducen a diversos métodos tanto de definición, como de medición de los conceptos. De acuerdo a la definición que hace BROWN (1991) reconocemos tres:

a) Métodos de personalidad autoadministrados.

Aquí se clasifican dos formas: 1) Inventarios de personalidad.- en donde se le pide al sujeto que se describa a sí mismo, se le puede presentar una lista de adjetivos o enunciados en donde se le pide verifique cual de ellos lo describe mejor. Un ejemplo lo es el Inventario Minessota Multifacético de la Personalidad (MMPI), 2) Escalas de actitudes.- las cuales buscan sistematizar el escalamiento de los reactivos; hay una gran variedad de este tipo por ejemplo: el método de las comparaciones binarias de Thurstone (ARCE, 1994 : 21).

b) Métodos proyectivos.

En estas técnicas se le presentan al sujeto estímulos ambiguos para que interprete su contenido. En esta técnica " se sigue la hipótesis de proyección que indica que un individuo, al enfrentarse a una situación de estímulos ambiguos, impone una estructura a los estímulos que refleja su organización particular de personalidad " (BROWN, 1991 : 448). Un ejemplo representativo de estas técnicas proyectivas es la Rorschach (prueba de las manchas de tinta).

c) Métodos situacionales.

El requisito común de estas técnicas de evaluación de la personalidad, es que el individuo se encuentre en una situación en donde se exija alguna acción o reacción, siendo el objetivo de observar y tomar nota de la conducta presentada; la cual tendrá una calificación de acuerdo a las características de personalidad que presente. Un ejemplo lo es la entrevista, que puede variar de calificaciones cuantitativas en dimensiones bien definidas (entrevista estructurada) a evaluaciones cualitativas globales (entrevista no estructurada).

2.4 CUESTIONAMIENTO A LOS INSTRUMENTOS DE EVALUACION PSICOLOGICA

SATTLER, PAWLIK, MORALES, BLOOM y FERNANDEZ, son autores que han realizado criticas acerca de las pruebas psicológicas! En este caso nos referiremos a sólo dos aspectos de estas dentro del ámbito educativo; las cuales son: 1.- Pruebas de aptitud en relación a las pruebas de rendimiento y, 2.- Pruebas con base en normas en relación a las Pruebas con base a criterios.

1.- PRUEBAS DE APTITUD CONTRA PRUEBAS DE RENDIMIENTO

Ante la gran expansión de diversas pruebas de aptitudes en diferentes campos, se inician una serie de investigaciones con respecto a su efectividad diagnóstica, primordialmente en la detección de alumnos que presentan deficiencias.

En investigaciones que se han realizado con el WPPSI y WISC, diferentes autores coinciden en que los perfiles de estos instrumentos no dan diferencias significativas en lectores aventajados de aquellos con desventaja (SATTLER, 1988 : 177).

En esté sentido hay conclusiones con respecto a las pruebas de Madurez percepto - motriz, como el Test Bender y el Test Frostig.

Las pruebas de aptitudes presuponen que los " problemas de aprendizaje " son atribuibles únicamente al menor, sin tomar en cuenta programas, métodos de enseñanza, ambiente grupal, desempeño de la maestra, factores socioeconómicos, etc.

Esta línea de marcada orientación conductista, propone como alternativa el diseño de instrumentos que evalúen las conductas y habilidades que puedan ser necesarias y observables

para los aprendizajes, y esto es lo que realizan las pruebas de rendimiento.

En países industrializados generalmente se aplican pruebas de rendimiento a la población estudiantil. En algunos casos se utilizan los dos tipos de pruebas

2.- PRUEBAS CON BASE EN NORMAS, CONTRA PRUEBAS CON BASE EN CRITERIOS

Las críticas que se han realizado a las pruebas con base en normas están centradas en que éstas presuponen diferencias individuales, lo que valoran son los éxitos académicos de unos pocos en relación con el resto del grupo de referencia.

Para las pruebas basadas en criterios se pone en juicio "HARRIS, 1974; CRAMBERT, 1975 que los contenidos no son solamente susceptibles de someterse a un control analítico de su validez, sino que además lo necesitan " citado por (PAWLIK, 1977 : 35).

La otra dificultad que muestran éstas consiste en el supuesto de la no diferencia individual, que llevaría a la escasa o inexistente dispersión de los resultados de la prueba.

Por otro lado, a la utilización de instrumentos psicológicos se les ha cuestionado la tendencia "instrumentista", que considera a la evaluación como un momento puntual y no como un proceso a seguir, como señala ANASTASI (1977 : 298) "se pone mucho énfasis en los test y muy poco en la psicología "

Por nuestro lado avalamos a la mayoría de las críticas que se han hecho, considerando que las dificultades que enfrenta el maestro (a), frente al grupo traspasan con mucho las insuficiencias propuestas teórico/metodológico que se han elaborado.

Se han descrito mucho las necesidades y carencias del sector educativo en el país, pero si pensamos que al no contar con instrumentos que apoyen - con todas sus limitaciones - el trabajo docente, es seguir dejando al maestro (a) con el gis y el pizarrón como único recurso de apoyo en la enseñanza que realiza día con día.

CAPITULO III

LOS INSTRUMENTOS DE MEDICION: ASPECTOS PSICOMETRICOS

Para efectos de diseño, estructuración y construcción de las pruebas psicológicas mediante las cuales se evaluarán y medirán los grados de avance en el aprendizaje, (entendido como un proceso mediante el cual un sujeto adquiere destrezas o habilidades prácticas, incorpora contenidos o adopta nuevas formas de conocimiento y/o acción), así como las conductas y/o actitudes entendidas como estructuras funcionales que sustenta, impulsa, orienta, condiciona, estabiliza y dan estabilidad a la personalidad; de los alumnos, se toman en cuenta ciertos principios e ideas que serán expuestos a continuación.

3.1 PRINCIPIOS DE MEDICION

En primer término, se procede a tratar el denominado principio de medición, cuya aplicación es fundamental en cualquier tipo de evaluación; cabe aclarar que la evaluación es la técnica más idónea para recopilar información relativa al aprendizaje de los alumnos en cada asignatura.

En términos generales, para medir los grados de aprovechamiento y niveles de conocimientos obtenidos durante las etapas del proceso enseñanza-aprendizaje por el alumnado, el docente elabora y/o prepara sus propias pruebas y toma en consideración lo que observa de sus educandos en el desarrollo de cada clase.

* Ver capítulo IV

A. LAS PRUEBAS Y SU CLASIFICACION

Para llevar a cabo la evaluación, el profesor, además de utilizar las pruebas comunes diseñadas y estructuradas por él mismo y que aplica en los exámenes normales de asignaturas, también recurre a otros medios especiales de medición que se explican en seguida, valiéndose de ellos con la finalidad de recopilar información más completa acerca del aprovechamiento del alumnado.

A.1 PRUEBAS PSICOLOGICAS ESTANDARIZADAS

Toda prueba psicológica, antes de su aplicación, obviamente, atraviesa por diversas fases en su diseño y construcción para poderla definir como estandarizada; en este sentido, tanto su contenido, su administración y su calificación, así como su tamaño y cantidad (que será igual a la de la población que se examinará), constituyen la base principal para que se le caracterice como estandarizada; y si además de las características de referencia, reunidas e incluidas en la formulación, se toman en cuenta las normas que facilitan la utilización de dicho cuestionario bajo criterios de validez y confiabilidad, se estará en condiciones de considerar que la prueba cumple con los requisitos que estipulan las normas de estandarización, a la que (SZEKELY, 1966 : 18), explica como " proceso mediante el cual se establecen las normas de una prueba psicológica, de modo que se le reconoce y denomina bajo el rubro de estandarización " .

Por otra parte cabe señalar que la planeación de las pruebas estandarizadas la llevan a cabo grupos de investigadores universitarios y de impresores profesionales; así mismo, también participan en este proceso, expertos en técnicas de medición (como actuarios) y en planeación académica. En otro contexto, cuando un país aún no ha alcanzado un cierto grado de desarrollo y nivel de crecimiento en materia pedagógico-educativa aceptables, entonces es frecuente que se recurra al uso de traducciones de formatos de cuestionarios que se pensaron y prepararon para ser aplicados a poblaciones estudiantiles con otras características antropológicas y

socioculturales; o bien, se le tienen que hacer adecuaciones y/o ajustes al instrumento atendiendo, en lo posible de las circunstancias, a las peculiaridades de la población en y a la que se tiene que examinar.

En síntesis, el diseño, la elaboración y estandarización de los cuestionarios de esta modalidad de pruebas, responden a los diversos factores culturales, lingüísticos y de edades de la población específica objetiva.

A.2 PRUEBAS PSICOLOGICAS ADAPTADAS

Para que una prueba se tipifique como adaptada, hay que tomar en cuenta cuatro niveles de adaptación, atendiendo a las características de cada una de ellas y que son los de:

1o. Traducción de la prueba

Hay casos en los que no se plantean y realizan análisis psicométricos sobre los reactivos de los cuestionarios, simplemente se cumple con el trámite de traducir el instrumento que posteriormente se utiliza en poblaciones estudiantiles con características muy diferentes a las de nuestros educandos.

2o. Otorgamiento de una traducción adecuada

En este caso, los índices de discriminación y los grados de dificultad asignados a las preguntas a resolver, o las incógnitas a despejar, tienen que ser aprobadas como nuevos reactivos para proceder a experimentarlos en poblaciones no representativas; lo anterior tiene que efectuarse con el fin de determinar y fijar con precisión dichos grados de dificultad y niveles de

discriminación.

3o. Impulso de una nueva normatividad, apoyado en un plan y/o esquema de una prueba preliminar o piloto.

Esta acción está pensada para cada nueva población estudiantil objeto de estudio, por lo que en este caso se estaría hablando de una nueva estandarización, pero que, sin embargo, contiene un análisis objetivo de los nuevos reactivos que se han experimentado al haberlos aplicado a una población representativa (grupos de distintos grados).

3.2 CARACTERISTICAS BASICAS DE LA MEDICION

De conformidad con la psicología correlacional, el estudio de las diferencias antropológicas e intelectuales de los individuos son materia de un análisis continuo, longitudinal; sin duda, las personas no sólo se distinguen por su aspecto físico, sino también por otras características, entre las que destacan su personalidad, sus habilidades, sus aptitudes y actitudes.

Ante tal reflexión, surgen muchas interrogantes que nos llevan a preguntarnos sobre la naturaleza y extensión de estas diferencias; por tanto y en respuesta a tales inquietudes, la psicología diferencial a la que ANASTASI (1977), considera como parte de la existencia de las diferencias en el comportamiento producidas por variables y factores internos y externos al propio individuo y/o grupo de individuos; se ha enfocado al estudio del fenómeno y a la búsqueda de aquellos elementos que permitan el establecimiento de rasgos característicos individuales y/o particulares que faciliten la descripción de las personas, las que, de común acuerdo con esta corriente del pensamiento psicológico, varían de una a otra persona.

Dentro del contexto académico se está pasando por un proceso de evaluación y, como reflejo natural de ello, la búsqueda de un instrumento de evaluación objetivo y efectivo, que no da más que sirva para la valoración de los programas y contenidos temáticos de las asignaturas, sino también que, al aplicarse, los resultados de los exámenes, den a conocer la realidad en los cambios conductuales de los educandos, es una constante esencial.

En términos de la evaluación educativa es posible utilizar diferentes métodos y técnicas, los que debidamente empleados nos permite lograr mejores resultados y, con esa información oportuna y objetiva, tomar decisiones adecuadas.

Es evidente que la medición, como se anticipa en párrafos precedentes, tiende a la adopción de ciertos números y escalas para representar aspectos sobresalientes de un determinado fenómeno objeto de estudio; por lo que, a la manera de (BROWN, 1980 : 230), la medición es " asignar

números a la conducta, de acuerdo a las reglas establecidas ". Finalmente se tiene que la medición o valoración de una característica y/o variable, conforme a nuestra hipótesis de trabajo, implica la utilización de determinados procedimientos, según normas previamente especificadas, obteniéndose como resultado la correspondencia de valores numéricos asignados a las variables, según van apareciendo en las distintas respuestas de los cuestionarios aplicados a cada alumno.

Quedando claro, en este sentido, que la Ciencia de la Educación, para que su acción sea más eficiente, se auxilia de otras ciencias y sus ramas; en consecuencia, la estadística aplicada a la Psicología Educativa, valora, en términos cuantitativos las características y/o variables que arrojan las respuestas contenidas en los cuestionarios; y para tal efecto, aporta los elementos de valoración y/o medición idóneos de dichos resultados.

Se entiende que los fundamentos de la estadística aplicada se anexan al método de pruebas, con ello se realiza un análisis de los reactivos por escalonamiento y bajo criterios de confiabilidad y validez del instrumento en cuestión, por lo que, de esta forma, se amplían los recursos evaluatorios que se adapta a los requerimientos de la evaluación y la medición; algunos de estos recursos son parte elemental de la psicometría, la cual propicia y facilita el diseño, la construcción y estandarización de mecanismos de valoración psicológica adecuados.

Por otra parte, en la Psicometría los conceptos teóricos y técnicos, de conformidad con las bases de SZEKELY, (1966:49) se articulan a los siguientes elementos:

- a) Sistemas de unidades aplicables a las funciones psicológicas y a la conducta,
- b) Diseño y construcción de pruebas,
- c) Evaluación de la prueba psicológica,
- d) Validez y
- e) Confiabilidad del instrumento.

A su vez, los principios teóricos psicológicos con los que se formula y estructura una prueba, quedan implícitos y explícitos en su contenido, en los reactivos y/o variables que se pretenden medir, por lo que de esa forma, la prueba adquiere el carácter de instrumento psicológico; pero también se debe recurrir a métodos estadísticos, que son de gran utilidad en el tratamiento de los datos, de modo que la transformación de nuestro instrumento queda, así, determinada.

En aras de comprender los grados de exactitud de las escalas de medición psicológicas y educacionales, es preciso conocer el modelo de escala utilizable; desde ese contexto, primero debe quedar claro, de conformidad con SATTLER,(1988 : 124), que la escala de medición se explica como " un sistema para asignar valores numéricos a ciertas características o rasgos", que tienen carácter de " mensurables ", es decir, son medibles.

Existen varios procedimientos para ordenar los datos; tales procedimientos y/o técnicas de medición, están considerados por diversos autores como escalas, entre las que destacan las siguientes:

1) Escalas Nominales.

A todas las características y/o variables que se van midiendo se les asigna un nombre, un número o una clasificación diferentes. En esta modalidad de escalas nominales se representan clases que se apartan entre una y otra; no se establece un orden jerárquico para su organización.

2) Escalas Ordinales.

En esta modalidad de escalas, las variables y/o características se van ordenando de manera jerárquica, de tal suerte que van apareciendo las diferencias que distinguen a unas de otras; por ejemplo, las variables de estrato social o comprensión del tema, las que pueden ser : nivel alto, medio, bajo, excelente, muy buena, buena, regular o deficiente.

3) Escalas de Intervalos.

En cambio con este tipo de escala, el procedimiento que se sigue es a base de medidas de distancia específicas en y para cualquier número; por lo tanto, la bondad principal de esta modalidad de medición consiste en que permite ordenar, jerarquizar y determinar la magnitud de la información y los resultados. Es importante no perder de vista que los trámites seguidos en estas escalas, especifican los intervalos que va marcando esta técnica de medición, pero a partir de lo cual también es posible definir los intervalos que sean idénticos, aunque en ningún momento se parte de cero.

4) Escalas de Razón.

Esta última modalidad se apoya en la técnica de intervalos idénticos y sus resultados, pero partiendo del cero absoluto; sin embargo, si se pone atención a las características de las pruebas psicológicas no es aceptable ni posible valorar un cero absoluto, por lo tanto, esta escala de medición no es aplicable a las ciencias humanas.

Si se pretende lograr efectividad, precisión y resultados satisfactorios en los estudios de casos, al ejercer y desarrollar la Psicología Educativa, las técnicas de medición que se recomiendan y que más se utilizan son las Escalas Ordinales y las Escalas Intervalares; ésta última es la que se aplica con mayor frecuencia en virtud de que no solamente abarca la clasificación, sino también la magnitud y/o cantidad de diferencias que suelen encontrarse entre los seres humanos.

Por ejemplo, es factible transformar (pasar de) la escala de intervalos (a) en una escala de unidades similares (iguales), en función de que resulta provechoso recurrir y apoyarse en un método estadístico apropiado; la utilización de dicho medio facilita la transformación y el cambio de las calificaciones intervalares a unidades, por lo que, en este caso, es correcto caracterizarlas como calificaciones estándar, las que de acuerdo con (BROWN, 1980 : 213), " Expresan la ejecución de una persona en términos de su desviación de la medida en unidades de desviación estándar ".

Es pertinente señalar que existen diversos tipos de calificaciones con los que se pueden realizar la interpretación y el análisis de las escalas ordinales y las intervalares; los más comúnmente utilizables son : a) rangos por déciles, b) rangos por centiles y c) calificaciones estándar " Z " .

3.3 ERRORES EN LA MEDICION.

Al desarrollar proyecto de investigación o realizar estudios de casos, durante la etapa de tratamiento de la información, suele suceder que al medir los reactivos o las variables contempladas en las pruebas o cédulas, se incurra en ciertos errores; y en Psicología, la medición de esos datos también implica la comisión, casi ineludible, de tales errores. Sin embargo, en nuestra ciencia se pretende y procura en lo fundamental que al medir, se tengan los mínimos errores, en el entendido de que, como asevera BROWN, (1980), " Las calificaciones observadas reflejen las calificaciones reales, con tan poco error como sea posible"; en este sentido, de nuestro tratadista aprendemos dos tipos de errores que aquí se rescatan:

1) Error al azar.

Ocurre cuando la variable afecta, en uno otro sentido, al estado en que se realiza la prueba; por lo que, para lograr una mayor confiabilidad del instrumento de medición, es necesario e importante controlar y verificar los presuntos errores; y así, por ejemplo, al aplicar una prueba en dos exámenes y fijar los tiempos límite diferentes para la realización de cada una, podría suceder que se produzca un error de este tipo.

2) Errores constantes.

Este tipo de errores produce efectos, definidos como errores de principios coordinados o sistemáticos, que repercuten en la aplicación de las pruebas; su conocimiento es muy importante, por la gran relevancia que tiene para los principios bajo los cuales se lleva a cabo la medición. Por otro lado, estos errores son una resultante de los propios instrumentos de medición y, además, son decisivos para los efectos de determinación de la validez y confiabilidad de los métodos, las técnicas y los resultados de la medición; por ejemplo, cuando un reactivo o variable pudieran ser ambiguos, confusos o poco entendibles.

A continuación se exponen las principales causas que generan estos errores:

1a. Las inherentes, o que van unidas a la prueba, como por ejemplo, la del muestreo del reactivo;

2a. Las particulares de la índole y/o forma de aplicación, de la técnica, entre las que se enuncian las siguientes : la situación física, las instrucciones, los registros de tiempos, o factores distractores; y

3ª Las debidas a las modificaciones en las características particulares del individuo, como el cambiar de ambiente, el grado de madurez emocional, la educación, etc..

Por lo tanto, una prueba educativa, específica (TAVELLA, 1972 : 9), " Es un instrumento para apreciar objetiva y cuantitativamente aspectos del fenómeno educativo en determinado nivel de medición y con cierto margen de error en las condiciones de una situación controlada ".

Finalmente, se tiene que las escalas de medición se ven afectadas por la aparición de errores que sesgan los resultados, lo cual no impide o limita su empleo en determinado momento, siempre y cuando no rebase cierta magnitud y sean identificados y controlados dichos errores. Hasta el momento no hay opiniones específicas que definan y determinen el grado de error que se tenga que aceptar como el mejor, puesto que, por ejemplo, el error estándar está relacionado con los coeficientes de confiabilidad operados en la evaluación de una prueba; en consecuencia, a un mayor grado de confiabilidad corresponderá invariablemente, un menor margen de error estándar en la medición.

3.4 LA CONFIABILIDAD

Hay quienes sostienen que el error, que va unido con la falta de confiabilidad de una prueba, provoca situaciones que pueden ocurrir en los resultados de la aplicación, por lo cual el instrumento pierde objetividad; y con mayor razón cuando no se identifican y controlan los errores.

La confiabilidad es un factor determinante e indicativo del grado de solidez de una prueba en sus reactivos y variables, en las mediciones; por eso, cuando un instrumento de medición dado arroja como resultado dos datos diferentes en un período de dos meses, de ser así, se puede asegurar sin temor de equivocación, que es poco confiable.

(TAVELLA, 1972 : 11) explica, en términos psicométricos, que "una escala es confiable" en la medida en que es "aplicada dos o más veces al mismo "alumno, obteniéndose "el mismo resultado o diferencias" que quedan dentro de los márgenes "aceptables"; pero MAGNUSSON, (1978 : 77) aclara que, si el instrumento proporciona datos confiables, entonces, en este caso, se trata de una prueba confiable, en virtud de que se obtienen "los mismos resultados" midiendo "el rasgo" una vez más "bajo condiciones similares" a las del sujeto "examinado". Nuestro autor indica que dichas condiciones hacen posible que la prueba tenga consistencia, pues de lo contrario, los resultados que arroje cada examen siempre variarían.

En términos generales, conforme a la tesis de SALVIA, (1987 : 1985), se utilizan tres métodos para evaluar el grado de confiabilidad de una prueba, los que a continuación se describen.

I) CONFIABILIDAD DE TEST-RETEST

Se trata de un estado de equilibrio cuya técnica consiste en la aplicación de la prueba a los alumnos en dos exámenes y, entre una y otra aplicación, se marca en un intervalo de tiempo no

muy largo, puesto que entre más corto sea dicho lapso de tiempo, tendrá un mayor grado de confiabilidad y la probabilidad de cambio será menor en función de lo corto del intervalo de tiempo, el cual, a su vez, estará determinado por el propio instrumento y por las condiciones en que se efectúe cada examen; y por tanto, los resultados de ambas aplicaciones se correlacionarán, obteniéndose de esta manera el coeficiente de confiabilidad.

2) CONFIABILIDAD DE FORMAS ALTERNAS O EQUIVALENTES

Este método consiste en dos pruebas que miden el mismo rango, en la misma amplitud y que son estandarizados en relación al mismo tipo de población; aún cuando las condiciones son las mismas, los reactivos no son iguales; no obstante, se parte del hecho de que las medias y las varianzas de ambas formas, supuestamente deben ser las mismas.

Ambas formas se tiene que aplicar en y a la misma población, y los resultados se deben correlacionar para, así, poder calcular el coeficiente de confiabilidad.

3) CONFIABILIDAD DE CONSISTENCIA INTERNA

En esta modalidad de método, la prueba se divide en dos; una vez que se aplican y califican, las puntuaciones de ambos grupos de las pruebas se correlacionan, dando como resultado una estimación de cada uno a la que se le denomina como cálculo de confiabilidad de mitades separadas; la división puede hacerse por pares, impares, grado de dificultad; y el procedimiento con el que se formula, es el de Spearman-Brown, el cual permite la correlación de los datos.

Hasta este punto se han tratado los factores determinantes de la solidez y/o consistencia de una prueba, pero es pertinente aclarar que también existen otros factores, tales como: a) la distancia de la prueba (corta o larga), b) el intervalo del test-retest (si es corto, el individuo tiende a acordarse de lo que contestó en su examen anterior o, por el contrario, no lo recuerda y puede originar modificaciones en sus respuestas posteriores), c) dificultad de la prueba (puede tratarse de reactivos de fácil o difícil respuesta), y d) la variabilidad durante la aplicación del examen (

como por ejemplo: cambiar las instrucciones o duplicar la aplicación de la prueba, etc.)

Una de las razones fundamentales que señala SALVIA(1987:91), por las que se obtiene el coeficiente de confiabilidad, es la del cálculo del margen de error o sesgo que con relativa frecuencia se tiene que efectuar y que se agrega al puntaje final de aciertos que logra el estudiante; es importante tomar en cuenta que, cuando una prueba es aplicada y se obtienen resultados como el apuntado antes, éstos se interpretan y analizan de conformidad con el error de medición de dicha prueba.

3.5 LA VALIDEZ

La Validez es otra de las características fundamentales que tiene una prueba; cuando llega a faltarle, no es posible saber qué es lo que ha medido o tiene que medir por lo que será difícil interpretar y analizar los resultados; por lo tanto, los conceptos genéricos de validez, solidez y confiabilidad están estrechamente relacionados, por lo que no se puede hacer referencia de uno y pasar por alto la mención de los demás, es decir, no son excluyentes sino que se interrelacionan.

En párrafos precedentes se ha explicado que la confiabilidad se basa en el margen y/o grado de error, o sea, que como afirma SATTLER, (1988 : 23,24), la validez " Es la proporción de la varianza real que es considerada relevante para los fines del examen " .

Consecuentemente, la validez de la prueba estará de acuerdo con la distancia con que se tiene que medir un rasgo; así mismo, por el vínculo que exista entre las calificaciones y las técnicas de medición que dan origen a este juicio o criterio externo.

Los principios de la Psicología estipulan que los resultados de la prueba tendrán un especial significado, la condicionante consiste en que dichos resultados tienen que interrelacionarse con otras variables que sean de gran importancia; pero lo que no se puede soslayar ni pasar por alto, en lo concerniente a la validez de una prueba y procurando estar al pendiente, es la delimitación de las variables que se estén midiendo y ver hasta qué grado el instrumento de medición usado, realmente sirve para el tipo de evaluación que se esté realizando

Según Sattler, se dispone de tres modos diferentes mediante los cuales se puede diseñar y construir una prueba bajo criterios de validez y confiabilidad, que se exponen enseguida.

1) VALIDEZ DE CRITERIO

Nuestro autor explica que la validez de criterio " se refiere a la relación de los resultados de una prueba con algún tipo de criterio o resultado "; dicho juicio debe contener características psicométricas que son de gran utilidad, ya que propician y facilitan la medición de las variables; por tanto, deben servir, en lo fundamental, a los propósitos para los cuales se construya.

Cuando no queda definida y determinada esta relación, es aventurado sostener que la prueba mide, en términos reales, los rasgos, características o variables para los cuales se haya diseñado; esta modalidad se subdivide en dos formas relativas al criterio y que a continuación se explican.

a. Validez Concurrente

Tiene lugar cuando los resultados de la prueba están o no vinculados con alguno de los criterios de medición que se esté aplicando, como puede ser el promedio de las calificaciones de exámenes parciales aplicados por la escuela de alguno de los alumnos.

b. Validez Predictiva o de Pronóstico

Esta forma consiste en llevar a cabo una correlación entre los resultados que registre la prueba, y el grado de rendimiento que se obtenga, de conformidad con un juicio y/o criterio importante, como podría ser la aprobación satisfactoria del grado escolar cursado.

2) VALIDEZ DE CONTENIDO

Con el objeto de verificar si los reactivos preparados y seleccionados para una prueba, son verdaderamente representativos de lo que se propone evaluar y medir, la evaluación que se realice de la prueba, tiene que ser sistemática.

3) VALIDEZ DE CONSTRUCCION

Este tercer criterio hace referencia a la evaluación que se efectúa a través de diversas técnicas y/o procedimientos preparados ex profeso para relacionar los elementos teóricos de los reactivos, con la pretensión de medir de esta manera es grado de evaluación del rasgo psicológico preestablecido; el propio SATTLER, (1988 : 230), propone tres factores fundamentales, que aquí rescatamos, determinantes de la validez de una prueba:

- 1o. Los relativos a la misma prueba;
- 2o. Los relacionados con el criterio; y
- 3o. La presencia y participación de sucesos imprevisibles.

4) VALIDEZ POR DIFERENCIACION DE GRUPOS

Por último, este cuarto método utilizado para determinar la validez de la prueba, según BROWN (1981 : 133), consiste en " ver si las calificaciones predictivas, cuando el examen se ejecuta bajo este criterio, registran ciertas diferencias que pudieran tener los grupos específicos".

Por ejemplo, señala BROWN se divide al grupo en dos; el primero lo formarán aquellos alumnos que tengan aptitudes y éxito en álgebra, cuyos estratos por calificaciones serán A,B ó C; y el segundo, se integra con los estratos de calificaciones D y F que han fracasado; acto seguido, se comparan las calificaciones que cada subgrupo obtenga en su examen respectivo, para detectar la existencia aparente de alguna diferencia que haya entre ambos subgrupos que, en términos estadísticos, pudiera tener cierta significación.

Planteado de otra manera, podríamos preguntarnos: ¿ Los estudiantes brillantes obtienen calificaciones significativamente más altas en sus pruebas que los que fracasan ?; como bien pudiera apreciarse en los resultados, se concluiría que ambos subgrupos obtienen calificaciones

significativamente diferentes en la prueba.

A veces, también puede aparecer otro problema con la aplicación de este criterio de validez por diferenciación de grupos, es el relativo al significado estadístico de la diferencia entre promedios de grupos, como la función del tamaño de dichos grupos; de tal manera que al aumentar el tamaño de un grupo, las diferencias más pequeñas en las calificaciones promedio que logra, adquieren un relativo significado estadístico.

Las pequeñas diferencias que aparecen en las calificaciones promedio, dada la forma en que con frecuencia se manipulan en la medición, las variables contenidas en los reactivos de las pruebas aplicadas a los grupos grandes, son a menudo significativas estadísticamente, como se ha mencionado; sin embargo, la prueba pudiera tener poco valor práctico para determinar y efectuar discriminaciones entre grupos: y para evitarse tales inconvenientes, TILFON (1937) citado por Brown 1991, recomienda que hay que determinar la magnitud de la yuxtaposición de las evaluaciones logradas por cada subgrupo. Así mismo, también se pueden evitar las interpretaciones erróneas mediante la utilización de las técnicas de medición más apropiadas, como la media, la desviación estándar o una medida de yuxtaposición, de igual importancia estadística dentro de la diferencia de las medias, en las calificaciones predictorias; por otra parte, una diferencia media de significación estadística, puede indicar o no una diferencia suficientemente grande, como para tener una aplicación práctica importante; Sin embargo, si las calificaciones de una prueba no indican una discriminación entre los subgrupos, determinada según este u otro criterio, dicha prueba no tendrá validez, de acuerdo a lo siguiente:

1o. Con los datos del ejemplo anterior, hay que definir los subgrupos; a) Los que aprueban y b) Los que reprueban; y

2o. Para determinar si se cuenta con una prueba que sea importante para efectos estadísticos, en la medición de las calificaciones que obtengan en su examen ambos subgrupos, hay que aplicar las técnicas estadísticas para el cálculo de la t ; este procedimiento permite comparar la diferencia en los resultados (la calificación de cada subgrupo), al aplicar la técnica de la media, y por la

utilización de una medida de error de muestreo.

La fórmula utilizable para efectuar el cálculo del caso de referencia, es la siguiente :

$$T = \frac{X_s - X_u}{\sqrt{(S_s / n_s) + (S_u / n_u)}}$$

X_s = Aprobados

X_u = Reprobados

El cuadro de valores t significativos demuestra que esta gran diferencia se produciría al azar, en menos de una de cada mil veces; en consecuencia, se puede decir que las calificaciones que obtengan en sus pruebas los alumnos aprobados y reprobados serán diferentes, toda vez que la prueba de razonamiento matemático predice el éxito o el fracaso en álgebra, o sea, el avance o atraso en la enseñanza - aprendizaje del álgebra.

3.6 LA NORMATIVIDAD

Como se ha mencionado en párrafos anteriores, las pruebas que se fundamentan en ciertas normas o reglas, constituyen la base, el punto de referencia para interpretar y analizar los datos y resultados; la interpretación de los datos y resultados de la prueba aplicada a un estudiante, implica el conocimiento de los parámetros bajo los cuales se harán la evaluación y la comparación, lo que, a la vez, permitirá saber cuáles normas se aplicaron para efectuar dichas acciones; y (MAGNUSSON, 1978 : 291), por su parte, explica que si se compara al sujeto (alumno) con un grupo de referencia correcto, la muestra de individuos (educandos) para lo cual ha sido considerada la norma, debe ser una representación justa de la población relevante, es decir, debe ser una muestra representativa del universo.

Pero también se ha anticipado en dichos párrafos que, con el fin de que una prueba diseñada sea significativa, debe ser confiable y válida, pero en caso de que las normas con las que se construyó no son las adecuadas, entonces los resultados no serán veraces, la prueba aplicada sufrirá un sesgo y la puntuación final se verá alterada por ese motivo.

Se subraya que las medidas psicológicas son de una gran utilidad dentro del contexto educacional, pero lo son aún más cuando las normas se obtienen de un cuerpo y grupo normativos y representativos; en el entendido de que, una muestra normativa toma en consideración las características de las personas en igualdad de circunstancias que las de la población en su conjunto; y cuando exista una cantidad suficiente de ellas para que su observancia sea uniforme y estable, y de como resultado un rango total de puntuaciones derivadas.

La muestra normativa, representativa, es una base fundamental de la evaluación y la medición, en virtud de que en ella se puede efectuar una comparación de resultados entre el grupo de estudiantes, con los del alumno al que se haya examinado; y por último, se consideran tres factores a explicar, tomando en cuenta lo expuesto antes:

1o. REPRESENTATIVIDAD

Este primer factor se refiere a las variables demográficas inherentes a la población objeto de estudio, para la cual se pretende normativizar la prueba, tomando en consideración, en primer término, que la muestra se integrará con los educandos que tengan las mismas o similares características que las de la población para las que se quiere normativizar dicho instrumento; y en segundo lugar, será necesario incluir las variables de edad, grado escolar, sexo, aculturación, inteligencia, ubicación histórica y características de la población.

2. NUMERO DE SUJETOS

Es decir, la cantidad de alumnos que se toman en cuenta para la integración de la normatividad, debe ser lo suficientemente grande para que haya estabilidad y equidad; cuando una muestra es mayor las normas son más equitativas y estables permanentes; y el número de sujetos que integran la muestra está determinado por el carácter homogéneo del grupo y además, se tienen que considerar los recursos disponibles para el desarrollo de la investigación.

Desde este contexto, BROWN, (1980 : 325), asevera que una muestra representativa utilizable a nivel nacional, conforme a la norma, no podrá ser menor de cien alumnos.

Durante la aplicación de una prueba, tiene que convivirse el conocimiento de la situación - problema, con los principios y las técnicas de medición psicológica, así mismo, se debe evaluar el grado de utilidad de dicho instrumento en cada caso particular. El propio BROWN, (1980 : 212), previene en el sentido de que “ a menos de que se tenga pleno conocimiento acerca de las variables que se están midiendo con la prueba y el grado de precisión con que se hace, no es factible interpretar y analizar los datos, es decir, las calificaciones que resultan en cada prueba con ningún grado de confianza o exactitud ”.

3o. RELEVANCIA DE LAS NORMAS

Por otra parte, BROWN clasifica las normas de ejecución; al evaluar la prueba, en dos categorías:

a) Interpretación basada en normas

Esta primera categoría se relaciona con el contenido y las normas y no es otra cosa, más que la propia ejecución por lo tanto, en una misma prueba se pueden llevar a cabo dos tipos de calificaciones.

b) Interpretación basada en criterios

Finalmente, con esta segunda categoría se puede apreciar que las características de una prueba son determinadas por: 1)El tipo de instrumento que se desea diseñar y construir y, 2)Las características psicométricas tomadas en cuenta por su diseño y construcción, tienen que ser acordes con la prueba.

Quizás cabría aclarar que las pruebas basadas en normas, son las que se basan en criterios, su validez se sustenta, principalmente, en los contenidos de las mismas.

CAPITULO IV

TEST DIAGNOSTICO EN HABILIDADES BASICAS; TEHBA

4.1 CONCEPTUALIZACION

El Test Diagnóstico en Habilidades Básicas (TEHBA), es un conjunto de pruebas con las que se pretende, principalmente efectuar una evaluación más objetiva de los conocimientos, habilidades y aptitudes básicas de aquellos alumnos que muestran dificultades en el proceso de enseñanza-aprendizaje, en donde es una actividad que esta encaminada a promover en el alumno la asimilación de la cultura.

El aprendizaje es el proceso técnico mediante el cual el alumno responde a la acción del educador. El método de enseñanza es paralelo al método de aprendizaje.

En tres etapas se da lugar este paralelismo:

- a) Comprender, b) Retener y c) Aplicar activamente.

Con éste Test se evalúan tres áreas esenciales: MATEMATICAS, LECTURA Y ESCRITURA, en las que las habilidades y aptitudes que se exploran se corresponden con los objetivos que se definen y plantean en los planes y programas de estudio vigentes, los que, a su vez, están en concordancia con los nuevos programas contenidos en el actual proyecto de Modernización Educativa.

No obstante que el objetivo fundamental del Test Diagnóstico en Habilidades Básicas (TEHBA), está pensado y orientado para la atención de alumnos con problemas de aprendizaje, sin embargo

CAPITULO IV

TEST DIAGNOSTICO EN HABILIDADES BASICAS; TEHBA

4.1 CONCEPTUALIZACION

El Test Diagnóstico en Habilidades Básicas (TEHBA), es un conjunto de pruebas con las que se pretende, principalmente efectuar una evaluación más objetiva de los conocimientos, habilidades y aptitudes básicas de aquellos alumnos que muestran dificultades en el proceso de enseñanza-aprendizaje, en donde es una actividad que esta encaminada a promover en el alumno la asimilación de la cultura.

El aprendizaje es el proceso técnico mediante el cual el alumno responde a la acción del educador. El método de enseñanza es paralelo al método de aprendizaje.

En tres etapas se da lugar este paralelismo:

- a) Comprender, b) Retener y c) Aplicar activamente.

Con éste Test se evalúan tres áreas esenciales: MATEMATICAS, LECTURA Y ESCRITURA, en las que las habilidades y aptitudes que se exploran se corresponden con los objetivos que se definen y plantean en los planes y programas de estudio vigentes, los que, a su vez, están en concordancia con los nuevos programas contenidos en el actual proyecto de Modernización Educativa.

No obstante que el objetivo fundamental del Test Diagnóstico en Habilidades Básicas (TEHBA), está pensado y orientado para la atención de alumnos con problemas de aprendizaje, sin embargo

puede ser un instrumento muy útil en la evaluación docente, puesto que le permite al profesor discriminar, de entre todas las habilidades y aptitudes exploradas y desarrolladas por un grupo, aquellas que no ha logrado dominar, facilitándole al maestro el orientar el trabajo académico frente a sus alumnos

Tanto las educadoras como el personal de apoyo se han interesado en conocer cuáles son las escalas de medición más idóneas, más efectivas, cuya aplicación se sustente en la evaluación educativa, de tal manera que les permita una interpretación clara y confiable de las dificultades que pudiera mostrar en clase un alumno.

La evaluación, en términos de la norma, propicia y facilita una descripción más objetiva de lo que es capaz de hacer un alumno con respecto a otros estudiantes; las calificaciones que se obtienen con este tipo de instrumentos, suelen definirse y presentarse como calificaciones estandarizadas, éstas son utilizables en la selección, clasificación e investigación en encuestas comparativas.

Por otro lado, la evaluación, con referencia a los criterios coadyuva en la eliminación de los inconvenientes clasificatorios de las pruebas normativas, con la intención de dar una alternativa de respuesta a la necesidad de orientar adecuadamente el proceso educativo del educando, y sin que se le someta a comparación con los demás procesos.

En este caso, el diseño y construcción de dichas herramientas se sustentan en objetivos instruccionales con la pretensión de describir ciertas fuerzas, propensiones y debilidades que manifiesta el alumno durante su formación y desarrollo educacionales; lógicamente, las pruebas se utilizan primordialmente para fundamentar el diagnóstico de los problemas de aprendizaje, problemática que sirve de base para diseñar los programas de recuperación que el maestro responsable del grupo, aplica de manera sistemática.

Las pruebas que componen el TEHBA, tienen validez de contenido; es decir que, su contenido paso por el tamiz de la evaluación mediante un escrutinio cuidadoso consistente en la discusión y emisión de juicios acerca de lo que significa y debe ser, en términos pedagógico y académicos, el

contenido de la prueba.

El criterio y validez de contenido fue determinado y quedó establecido a partir del análisis y examen de tres factores importantes, que propone (SALVIA, 1987 : 101), y los cuales se mencionan a continuación:

- 1.- Los tipos de reactivos incluidos en la prueba
- 2.- La integridad de la muestra de reactivos
- 3.- La forma en que los reactivos de la prueba evalúan el contenido.

Se podría concluir diciendo que, las pruebas constitutivas del TEHBA, son utilizables para el desarrollo de proyectos, tanto en la investigación de campo como en la de gabinete (en la recopilación de datos), puesto que se trata de cuestionarios que se sometieron a un riguroso proceso de estandarización, que los convierte en herramientas valiosas y confiables para el estudio de la población urbana y rural, en síntesis, son técnicas esenciales para el desarrollo comunitario.

4.2 PROCESO DE PLANEACION DEL TEHBA

El proceso de planeación (diseño y construcción) del TEHBA, atravesó por dos etapas a las que, a su vez siguió una serie de pasos para la formulación y presentación final de este mecanismo.

A la primera fase corresponden las siguientes acciones:

- 1a. Revisión de planes y programas de estudio de primer grado de primaria.
- 2a. Revisión de libros de texto oficiales y privados de la materia de Español y Matemáticas.
- 3a. Consulta de especialistas en planeación-programación educativa y maestros de primer grado.
- 4a. Diseño del formato de las pruebas a partir de los objetivos de enseñanza-aprendizaje.
- 5a. Redacción y selección de reactivos de las pruebas experimentales.
- 6a. Diseño y selección de nuevos reactivos, bajo criterios de grado de dificultad y nivel de discriminación.
- 7a. Examen piloto de las pruebas experimentales a una muestra representativa.
- 8a. Diseño y presentación de las pruebas definitivas.

Para la segunda fase hay que considerar las acciones a continuación indicadas:

- 1a. Aplicación de las pruebas definitivas de Matemáticas, Lectura y Escritura, a una muestra representativa de la población que se atiende.

- 2a. Capacitación de encuestadores-capacitadores.
- 3a. Consideración como población (universo), a todos los educandos que cursaron el primer grado de primaria en escuelas oficiales, turno matutino, el periodo escolar 1993-1994, ubicadas en la Delegación Benito Juárez, del Distrito Federal.
- 4a. Cálculo y determinación de la muestra mediante el método de muestreo aleatorio bietápico por racimos, donde a la escuela se le considero el racimo y el alumno, la unidad secundaria.
- 5a. Selección aleatoria de 40 escuelas de la Delegación Benito Juárez, y de cada escuela se seleccionaron dos alumnos.
- 6a. Cálculo de la muestra representativa final integrada por 75 educandos.

4.3 DESCRIPCION DE LAS PRUEBAS QUE INTEGRAN EL TEHBA

Al inicio de este capítulo, se explica que el TEHBA es un conjunto de pruebas que pretenden fundamentalmente, ser un instrumento de evaluación objetiva de los conocimientos, habilidades y aptitudes básicas de aquellos alumnos que presentan dificultades en el aprendizaje en los primeros tres grados de primaria.

Asimismo, se asienta que son tres áreas las que evalúa el TEHBA: Matemáticas, Lectura y Escritura.

Para los efectos de estudio de nuestro proyecto de investigación, se dispone de un TEHBA para cada grado: primero, segundo y tercer grado de primaria, que previamente se diseñaron.

Es obvio que por tiempo y optimización de los recursos disponibles sólo utilizamos el de primer grado de primaria.

A) Descripción de la prueba de Matemáticas

El pensamiento matemático es esencialmente abstracto y sólo alcanza entidad como tal cuando se desliga de su referencia a lo concreto.

Su formación y desarrollo exige unas formas de conocer y accionar la realidad, un determinado desarrollo intelectual, coincidente con lo que Jean Piaget ha denominado período de las Operaciones Concretas.

El niño que no ha accedido a este estadio del desarrollo intelectual, sólo razona sobre los estados y no sobre las transformaciones.

Su pensamiento se ha separado todavía de sus acciones efectivas sobre los objetos, no es reversible.

“ La reversibilidad es la característica fundamental en la que se basa todo conocimiento lógico y matemático ”

Para que haya esa reversibilidad debe darse:

- 1) Una construcción
- 2) Clasificación
- 3) Seriación

Y con esto se da una adquisición paralela con el número.

Cuando esto sucede el alumno es capaz de invertir mentalmente el proceso de un resultado que se le propone, de relacionar el estado inicial y el final.

En esto consiste las operaciones concretas, “en acciones interiorizadas ejecutadas no materiales, sino interior y simbólicamente, que pueden ser combinadas de varias maneras y en particular que son reversibles y que se refieren a objetos manipulables” (Piaget, 1975)

La prueba de matemáticas utilizada consta de 25 reactivos divididos en cinco subpruebas:

- a) Lógica y relaciones
- b) Conocimiento de números
- c) Operaciones numéricas
- d) Resolución de problemas
- e) Figuras geométricas

El material que compone esta prueba, es:

- Manual de instrucciones generales
- Cuaderno de instrucciones
- Cuaderno de respuestas
- Hoja de registro
- Protocolo

Administración de la prueba de matemáticas:

Para iniciar la aplicación de esta prueba, el examinador debe contar con el cuaderno de instrucciones, en el que se describe el procedimiento para resolver cada ejercicio.

Por lo que al cuaderno de respuestas se refiere se convino en que se le entregaría al alumno para que fuera anotando las respuestas que se le solicitaron.

Se determinó que aproximadamente se requieren de 15 a 27 minutos para la aplicación de las cinco subpruebas, aunque las diferencias individuales, respecto al estilo particular de cada niño para contestar, podrían alargar o abreviar el tiempo de aplicación: en promedio, la prueba de matemáticas se realiza en 22 minutos.

Calificación de la prueba de matemáticas.

Las instrucciones que se tienen para la calificación de esta prueba, son objetivas, en virtud de que no se requiere de una interpretación subjetiva de las respuestas del niño.

En el cuaderno de instrucciones, en el ejercicio correspondiente, se incluyeron señalamientos de los criterios de corrección; de acuerdo a estas instrucciones, se acordó que se irían calificando las respuestas que el alumno presentó en el cuaderno de respuestas.

Hay solamente dos tipos de calificación:

Positivo = 1

Negativo = 0

Se quedó en que, una vez concluido el proceso de calificación de todos los reactivos de la prueba, se vaciarían los resultados en la hoja de registros, y una vez que estuviera terminado este procedimiento, se llenaría el protocolo correspondiente, donde se señalarán los aciertos por subprueba y los aciertos totales de la prueba.

B) Descripción de la prueba de Lectura

El acto lector supone el poder descifrar unos signos gráficos y abstraer de estos un pensamiento. Estos signos tienen un equivalente sonoro, el cual está asociado con el lenguaje hablado, en donde:

- 1.- Encierran un significado
- 2.- Están colocados de modo concreto
- 3.- Acomodados en el espacio en una determinada dirección espacial.

El aprendizaje de la lectura, supone un proceso complejo en donde intervienen factores mentales, lingüísticos, perceptivo-motrices y socio-afectivos;. Los cuales deben alcanzar su nivel madurativo para que se puede producir el acceso a la lectura; en caso de que esto no fuese así, surgirían una serie de dificultades.

La lectura sería mecánica no comprensiva, se efectuaría con lentitud, con alta de ritmo, saltos o repetición de líneas no encuentran gusto por la lectura y presentan dificultades para los demás aprendizajes escolares.

Esto se puede vencer con una reeducación adecuada, aunque algunas persisten de forma atenuada en el alumno.

La prueba que se utilizo consta de 15 reactivos divididos en cuatro subpruebas:

- a. Lectura mecánica
- b. Identificación de letras y palabras
- c. Comprensión de la lectura
- d. Habilidad gramatical

El material que compone esta prueba es:

- Manual de instrucciones generales
- Cuaderno de instrucciones
- Cuaderno de respuestas
- Cuaderno de calificaciones
- Hoja de registro
- Protocolo

Administración de la prueba de lectura:

Para el inicio de la aplicación de esta prueba, es necesario que el examinador cuente con el cuaderno de instrucciones, en el que se describe el procedimiento para la realización de cada ejercicio; y también se debe contar con el cuaderno de calificaciones, en el cual se anotarán las respuestas que el niño vaya presentando para cada ejercicio.

El cuaderno de respuestas se le dará al niño para que conozca el procedimiento y vaya siguiendo las instrucciones correspondientes.

Aproximadamente se requieren de 11 a 21 minutos para la aplicación de las cuatro subpruebas, aunque las diferencias individuales, respecto al estilo particular de cada niño para contestar pueden alargar o abreviar el tiempo de realización; en promedio, la prueba de lectura tarda 16 minutos.

Su aplicación se hará siguiendo las indicaciones que se proporcionan en el cuaderno de instrucciones, y se irán registrando las contestaciones que emita el niño en el cuaderno de respuestas.

Calificación de la prueba de lectura:

El acto de la lectura es un proceso complejo que implica e involucra diversos aspectos en su evaluación.

GONZALEZ (1994), autora del TEHBA plantea; que si bien se considera a la lectura como un acto primordialmente placentero, para llegar a disfrutar de ella, es requisito que se tenga un adecuado dominio de las habilidades básicas de dicho ejercicio.

La valoración se enfoca hacia el dominio central de tales habilidades que son tomadas como una condición necesaria aunque no suficiente, para el acto lector pleno.

Por lo que al cuaderno de instrucciones se refiere, en el inciso de cada ejercicio se hacen señalamientos sobre los criterios de corrección y, conforme a las instrucciones, se irán calificando las respuestas que el alumno marque en el cuaderno de respuestas.

En esta prueba también se cuantificarán los errores que se cometan durante la demostración de las siguientes habilidades y aptitudes del alumno para: la lectura mecánica, la lectura comprensiva y la habilidad gramatical; en consecuencia, se muestran indicaciones precisas sobre el método para evaluar cada ejercicio, las cuales aparecen enunciadas en el cuaderno de calificaciones.

Una vez que se ha concluido la calificación de todos los reactivos de la prueba, se vacían los resultados en la hoja de registro; en la misma se anotarán los errores cometidos, se sumará el total de cada ejercicio y, posteriormente, se transformarán los errores en aciertos, siguiendo la

especificación que se hace en la tercer columna de la hoja.

De acuerdo a los señalamientos que se indican para el número de errores permitidos por ejercicio, podrá ser tache 0 (error) ó 1 (acierto).

Con los aciertos por subprueba y con el total de la hoja de registro, se preparará el protocolo.

C) DESCRIPCION DE LA PRUEBA DE ESCRITURA

El aprendizaje de la escritura es complejo e intervienen factores psicofísicos y determinadas condiciones ambientales.

La evolución del aprendizaje del lenguaje escrito atraviesa distintas fases:

- 1.- Aparece el garabateo con una diferenciación progresiva entre lo que es el resultado de una intención de representación a través del dibujo y lo que es una iniciación a la escritura.
- 2.- Hacia los 5 ½ años ó 6 años el niño puede copiar letras y palabras.
- 3.- Poco despues, paralelamente al aprendizaje de la lectura pasa a la etapa de transcripción de las palabras dictadas (6-7 años).
- 4.- Posteriormente accede el lenguaje escrito, y a percibir su función de expresión personal y de comunicación.

El perfeccionar el lenguaje escrito será una labor que se llevará a cabo a través de la escolaridad básica del niño.

Para que el niño pueda aprender a escribir, es necesario que haya adquirido determinado grado de "maduración" en aspectos como:

- a) Coordinación visomanual.- sin la cual no podrá realizar los movimientos finos y precisos que exigen los grafismos
- b) El lenguaje.- para comprender el paralelismo entre el simbolismo del lenguaje oral y el lenguaje escrito.
- c) La percepción.- que posibilitara la discriminación y realización de grafismos en una situación espacial determinada de cada letra dentro de la palabra, de las palabras en la línea y en el conjunto de la hoja de papel, así como el sentido direccional de cada grafismo y de la escritura en general.

En caso de que se presenten dificultades en la escritura es necesario que se den ejercicios psicomotrices, de coordinación viso-motriz, gráficos, que faciliten y mejoren el control grafo-motriz, de realización correcta de los movimientos que componen las letras, de sistematización de las escrituras y de perfeccionamiento.

La prueba de escritura que se utilizó consta de 10 reactivos divididos en tres subpruebas:

- a) Habilidad para copiar
- b) Habilidad para el dictado
- c) Habilidad para la redacción

El material que compone esta prueba es:

- Manual de instrucciones generales

- Cuaderno de instrucciones

- Cuaderno de respuestas

- Hoja de registro

- Protocolo

Administración de la prueba de escritura:

Para iniciar la aplicación de la prueba, el examinador contará con el cuaderno de instrucciones, en donde se describe el procedimiento a seguir en cada ejercicio.

El cuaderno de respuestas se le entregará al alumno, a fin de que él vaya anotando las respuestas que se le soliciten. Aproximadamente se requieren de 15 a 27 minutos para la realización de la prueba, como los alumnos muestran diferencias (rapidez mental o mente analítica), ello hará que el tiempo de examen se alargue o se abrevie, pero el promedio de aplicación de la prueba es de 22 minutos.

Dicho examen se iniciará siguiendo las indicaciones que se consignan en el cuaderno de instrucciones.

Calificación de la prueba de escritura

El acto de la escritura es un proceso complejo, que implica el involucramiento y conjunción de diversos aspectos en su evaluación.

Si bien se considera a la escritura como un acto primordialmente creativo, de posibilidad de expresión subjetiva, es importante tener un adecuado dominio de las habilidades básicas de la escritura, para alcanzar y perfeccionar este acto.

La valoración se centra en el control de dichas habilidades, a las que GONZALEZ (1994), considera como una condición necesaria aunque no suficiente, como en la lectura, para el acto de la escritura.

Tomando en cuenta las deficiencias que también presentan los alumnos en el ejercicio de la escritura, se atendieron las propuestas que, para el caso, han formulado diferentes investigadores y especialistas.

De acuerdo con los señalamientos, principalmente se evalúa: la actitud del niño al momento de escribir, al análisis de la grafía y algunos aspectos sintácticos; las instrucciones enunciadas para proceder a la calificación de la prueba no son del todo objetivas como en la de los incisos anteriores, en virtud de, para este, caso se requiere de cierta capacidad de interpretación y análisis por parte del examinador, específicamente para la prueba de redacción.

Por otra parte en el cuaderno de instrucciones, en el inciso relativo a cada ejercicio, se señalan los criterios de corrección; procediéndose a calificar las respuestas emitidas por el alumno, conforme a tales indicaciones.

En esta prueba se van contabilizando los errores que se cometen en las habilidades y/o aptitudes para: copiar, el dictado y la redacción; las instrucciones precisas sobre cómo evaluar cada ejercicio, se encuentran en el cuaderno de calificación. Y ya ejecutado este proceso en todos los

reactivos de la prueba, se anotarán los resultados en la hoja de registro.

Además, también se asentarán los errores que se hayan presentado, sumando el total por cada ejercicio; posteriormente se transformarán los errores en aciertos, siguiendo la especificación que se hace en la tercer columna de la hoja.

De acuerdo a los señalamientos establecidos para el número de errores permitidos por ejercicio, aparecerá anotado tache 0 (error) ó 1 (acierto).

Con los aciertos por subprueba y el total de la hoja de registro, se procederá al llenado del protocolo.

No debe perderse de vista que la aplicación del TEHBA es individual, que las pruebas se deben o pueden administrar selectivamente, o sea, que se puede aplicar una sola prueba o todas, dependiendo de las deficiencias que previamente se hayan detectado en el alumno a evaluar.

El tiempo promedio para la administración de estas tres pruebas, es de 60 minutos. El examinador no debe modificar la forma en que se presenten las instrucciones, ni ayudar con palabras u otorgar apoyo más allá de los límites que se señalan.

Es recomendable que la aplicación de cada prueba se realice en una sola sesión; en el supuesto caso de que se aplicarán las tres subpruebas de las tres áreas, sería favorable realizarlas en dos sesiones, siguiendo este orden:

1a. sesión: Matemáticas y Lectura

2a. sesión: Escritura.

Para llevar a cabo los exámenes, es conveniente contar con un espacio cerrado, con buena iluminación y, de ser posible libre de ruidos y distractores; sólo estarán presentes, el alumno y el aplicador de la prueba.

Previo al inicio de la aplicación, el examinador procurará establecer un clima de confianza con el alumno a evaluar.

En las pruebas que se apliquen al niño, deberá disponer de lápiz y goma para anotar las respuestas.

Todas las pruebas cuentan con hoja de registro y con un formato de protocolo, donde se pueden interpretar los resultados que arrojan las pruebas.

Hoja de registro: Una vez llenado, este documento nos permite presentar un perfil del alumno evaluado, donde se señalan las habilidades que no ha logrado dominar.

De acuerdo con este perfil, es posible diseñar un programa específico de recuperación.

Protocolo: Existen dos lecturas posibles de los resultados que se asientan en este formato. -

1) Lectura de Escala Ordinal (Por subprueba y prueba total).

- Escala ordinal para la subpruebas:

Bueno - Regular - Deficiente

Esta graduación permite hacer una apreciación de cuáles son las áreas y contenidos que deben ser recuperados.

- Escala Ordinal para la prueba total:

Excelente - Bien - Regular - Deficiente - Muy deficiente - Requiere de Educación Especial. La amplitud e intensidad de los objetivos del programa que necesita recuperar el alumno evaluado, son graduales.

Excelente: Aquí el alumno ha logrado el dominio de por lo menos el 91% de los objetivos planteados por el programa.

Bien: El alumno logró el dominio de por lo menos, el 80% de los objetivos predeterminados para el programa.

Regular: El alumno cumplió por lo menos, con el 50% de los objetivos formulados para el programa; adicionalmente, el niño requerirá de apoyo en su casa; pero por otro lado, el maestro señalará cuales son los objetivos de dicho programa que no se hayan cubierto.

Deficiente: Cuando el educando sólo ha logrado el dominio de menos de 50% de los objetivos fijados para el programa; por lo tanto requiere de apoyo especial en el salón de clase. Así como durante el periodo de vacaciones; de dicho apoyo depende que pueda aprobar el curso.

Muy Deficiente: Con esta escala se puede ver que el alumno presenta deficiencias severas, en cuanto al dominio de los adjetivos instruccionales señalados para el curso.

Consecuentemente, se necesario iniciar un programa intensivo de apoyo, en el que se conjunten esfuerzos del maestro y de los padres de familia; asimismo, se necesitará de la intervención de un profesor en Educación Especial, así como de una valoración más amplia que preparará el personal especializado.

Requerimiento de Educación Especial: El alumno presenta deficiencias que no pueden ser superadas en un grupo regular, por ello es conveniente que se canalice al niño a un servicio especializado para valoración exhaustiva.

Cuando el alumno obtiene la calificación regular, deficiente o muy deficiente, significa que precisará de un programa de recuperación que se podrá diseñar con base en la hoja de registro.

2) Escala de Intervalos:

- Rangos percentiles para subpruebas y prueba total
- Calificaciones " Z " para prueba total.

METODOLOGIA

A fin de estimar la Validez por el método de diferencia de grupos (Brown : 1991) del Test Diagnóstico en Habilidades Básicas (TEHBA) se procedió a seleccionar dos grupos independientes, con un total de 73 sujetos, debido a que en este tipo de validación, los grupos pequeños arrojan diferencias más claras y para otro tipo de validación son necesarios 100 sujetos mínimos.

Sujetos:

El primer grupo se conformó con 38 alumnos/as de la escuela Primaria Emilio Bravo, ubicada en Av. Emiliano Zapata esq. Lázaro Cárdenas Col. Cuauhtepc Barrio Bajo, perteneciente a la Coordinación No 2 de la Dirección General de Educación Especial, que cursaban el primer grado de primaria regular con una E.C. de 6.5 a 7 años.

El segundo grupo se conformó con 35 alumnos/as inscritos en el programa que ofrecía la Dirección General de Educación Especial, USAER para menores que hubieran reprobado el 1er. grado de primaria por deficiencias en la lectura, escritura y manejo de operaciones numéricas, con una E.C. de 7.5 a 9 años.

La unidad de Servicios de Apoyo a la Educación Regular (USAER), es la instancia técnico-operativa y administrativa de la Educación Especial que se crea para favorecer los apoyos teóricos y metodológicos en la atención de los alumnos con necesidades educativas especiales dentro del ámbito de la escuela regular, favoreciendo así la integración de dichos alumnos y elevando la calidad de la Educación que se realiza en las escuelas regulares.

El criterio para seleccionar los grupos fue el rendimiento que presentaron en su escolaridad.

El primer grupo considerado como grupo regular (calif. Aprobatorias)

El segundo grupo considerado como grupo especial (calif. Reprobatorias).

INSTRUMENTOS:

-Test Diagnóstico en Habilidades Básicas (TEHBA) para primer grado, el cual consta de las siguientes pruebas:

LECTURA: consta de 15 reactivos, dividida en 4 subpruebas:

- a) Lectura Mécanica
- b) Identificación de letras y palabras
- c) Comprensión de la lectura
- d) Habilidad gramatical

ESCRITURA: consta de 10 reactivos, dividida en 3 subpruebas:

- a) Habilidad para copiar
- b) Habilidad para el dictado
- c) Habilidad para la redacción

MATEMATICAS: consta de 25 reactivos y dividida en 5 subpruebas:

- a) Lógica y relaciones
- b) Conocimiento en números
- c) Operaciones numéricas
- d) Resolución de problemas
- e) Figuras geométricas

Tiempo promedio de aplicación de cada prueba:

	GRUPO REGULAR	GRUPO USAER
LECTURA	15 Min.	18 Min.
ESCRITURA	20 Min.	25 Min.
MATEMATICAS	21 Min.	25 Min.

-Cuestionario para recabar información complementario.

ESCENARIO:

La Escuela Primaria Emilio Bravo ubicada en Av. Emiliano Zapata Esq. Lázaro Cárdenas Colonia Cuauhtemoc Barrio Bajo perteneciente a la Coordinación No 2 de la Dirección General de Educación Especial, consta de las suficientes aulas y servicios necesarios (sanitarios, cooperativa, patio, canchas, dirección telefónica, conserje) para atender a la población inscrita la cual se caracteriza por tener un nivel socio-económico bajo medio.

La aplicación se llevo a cabo en tres sitios diferentes: salón, patio y dirección, los cuales eran asignados conforme a las necesidades de la escuela, donde no contaba con las condiciones idóneas para la ejecución de las pruebas

PROCEDIMIENTO:

La aplicación se realizó individualmente en dos sesiones siguiendo las instrucciones señaladas en el manual, se encontraban presentes examinador y examinado exclusivamente.

En la primera sesión se aplicó la prueba de matemáticas y en la segunda las de lectura y escritura, el tiempo promedio para cada prueba fue:

	GRUPO REGULAR	GRUPO USAER
LECTURA	15 Min.	18 Min.
ESCRITURA	20 Min.	25 Min.
MATEMATICAS	21 Min.	25 Min.

Una vez calificadas las pruebas se procedió a la codificación por medio de una computadora personal, utilizando el paquete estadístico WINSPSS. Para determinar de diferencia de grupos se utilizó el estadístico de prueba de Mann-Whitney, ya que se pretendía encontrar la diferencia entre dos muestras independientes, y el nivel de medición del TEHBA es intervalar.

ANALISIS DE RESULTADOS

En la prueba de escritura se presenta una media para el grupo USAER de 19.18 y para el grupo regular de 50.51 (ver cuadro 1a),dando como resultado una calificación Z de -6.47, lo que representa un nivel de significancia de p .05.

En la prueba de lectura se presenta una media para el grupo USAER de 21.6 y para el grupo regular de 48.4 (ver cuadro 1b) dando como resultado una calificación Z de -5.50, lo que representa un nivel de significancia de p .05.

En la prueba de matemáticas se presenta una media para el grupo USAER de 20.7 y para el grupo regular de 49.2 (ver cuadro 1c), dando como resultado una calificación Z de -5.83, lo que representa un nivel de significancia de $p .05$.

Para observar con más detalle cada una de las subpruebas que componen la prueba de escritura, se procedió a analizar cada una de estas (ver cuadro 2a) En las subpruebas de escritura se presentó una media para el grupo USAER de:

Habilidad para la copia	27.3
Habilidad para el dictado	21.1
Habilidad para la redacción	27.9

y para el grupo regular de:

Habilidad para la copia	43.5
Habilidad para el dictado	48.9
Habilidad para la redacción	43.0

(ver cuadro 2a), dando como resultado una calificación Z en cada subprueba de :

Habilidad para la copia	-3.6
Habilidad para el dictado	-5.7
Habilidad para la redacción	-3.7

Lo que representa una significancia de P .05. por lo que si hay diferencia estadísticamente significativa en cada una de ellas.

En las subpruebas de lectura se presentó una media para el grupo USAER de :

Lectura mecánica	24.4
Identificación de letras y palabras	48.7
Comprensión de la lectura	22.7
Habilidad gramatical	27.0

y para el grupo regular de :

Lectura mecánica	46.0
Identificación de letras y palabras	48.7
Comprensión de la lectura	47.4
Habilidad gramatical	43.8

(ver cuadro 2b), dando como resultado una calificación Z en cada subprueba de:

Lectura mecánica	-4.54
Identificación de letras y palabras	-5.81
Comprensión de la lectura	-5.24
Habilidad gramatical	-3.65

Lo que representa un nivel de significancia de P .05 en cada una de ellas.

En las subpruebas de matemáticas, se presentó una media para el grupo USAER de :

Lógica y relaciones	31.4
Operaciones numéricas	23.1
Resolución de problemas	32.7
Figuras geométricas	31.8

y para el grupo regular de :

Lógica y relaciones	39.9
Operaciones numéricas	47.1
Resolución de problemas	38.7
Figuras geométricas	3.5

(ver cuadro 2c), dando como resultado una calificación Z en cada subprueba de :

Lógica y relaciones	-1.9	No hay diferencia
Operaciones numéricas	-5.12	
Resolución de problemas	-1.45	No hay diferencia
Figuras geométricas	-5.56	

Para la subprueba de apreciación numérica existe diferencia estadísticamente significativa.

Para las subpruebas de lógica y relaciones y resolución de problemas no existe diferencia estadísticamente significativa

Una posible causa de la ausencia de diferencias puede ser el grado de dificultad de los reactivos.

Para lógica y relaciones los fáciles son fáciles: 0.75 en promedio (ver THEBA, p. 15).

En el caso de operaciones numéricas, por el contrario son reactivos difíciles : 0.43 en promedio (ver THEBA, p. 16).

Para la subprueba de figuras geométricas, existe una diferencia estadísticamente significativa a favor del grupo USAER. Este hecho puede deberse a que esta subprueba consta únicamente de dos reactivos, lo que dificulta establecer diferencias precisas entre ambos grupos.

Conclusiones

Las tendencias más recientes en el campo de la evaluación sugieren tomar en cuenta el proceso de aprendizaje, más que los productos terminales (Stufflebeam y Shinkfield, 1989), lo que permite tomar acciones para mejorar las practicas educativas; en ese sentido se han diseñado diversos instrumentos para la evaluación llamadas Pruebas de Rendimiento de tipo Diagnóstico, que permite valorar los aciertos y las deficiencias de los alumnos en una área de conocimientos específico.

En la medida que el maestro indica los errores que los alumnos presentan, él puede realizar acciones correctivos que permitan un conocimiento efectivo. Como se ha señalado el uso de pruebas no agota el problema de la evaluación, sin embargo, ayuda en mucho el trabajo que maestros y profesionales de apoyo realizan cotidianamente, en especial cuando no se cuenta con una formación suficiente en cuanto al problema de la evaluación de los aprendizajes escolares.

Los aprendizajes de los alumnos en la escuela se refieren as habilidades, conocimientos y destrezas específicas de acuerdo a un programa de estudios. Coll (1992) menciona además los aprendizajes relativos a valores y actitudes.

Las pruebas diagnosticas generalmente están referidas a las llamadas habilidades básicas que se han definido como aquellos conocimientos escolares elementales que sirven de punto de apoyo a futuros aprendizajes como el lenguaje escrito y las operaciones matemáticas básicas (Tenbrink; 1988).

El desarrollo de pruebas psicológicas data de finales del siglo XIX, dando lugar a la Psicometría como campo específico que marca los procedimientos para el diseño de este tipo de instrumentos. Las características psicometricas de una prueba son un validez y confiabilidad. La confiabilidad indica hasta que punto una prueba es consistente en sus mediciones.

La característica más importante de una prueba es su validez, existiendo diversas formas de validez: de Construcción, de Criterio y de Contenido. En el caso de pruebas diagnósticas se puede considerar un 4º tipo de validez, la validez por diferencia de grupos diferenciados por el criterio que pretende evaluar (Bronw, 1991).

En México la evaluación educativa ha sido más un aspecto retórico que un practica sistemática. En relación con los instrumentos para la evaluación el modelo que hasta hace poco tiempo imperaba obedecía sobre todo a orientaciones de tipo clínico, en las que dificultades del aprendizaje era problema del alumno. A esta orientación correspondía las pruebas de aptitudes como el WISC para valorar la inteligencia, o el Bender para determinar las habilidades visoespaciales que se consideraban básicos para la adquisición de la lectoescritura y las matemáticas elementales.

Ante reiteradas críticas hacia el modelo clínico, surgen nuevas orientaciones en las que se considera que las deficiencias no solamente son atribuibles a los alumnos, las deficiencias están sobre todo en el proceso de enseñanza (programas, estrategias de enseñanza, materiales específicos, etc.).

En el año de 1994 aparecieron en nuestro país dos pruebas de rendimiento de tipo diagnóstico más acorde con esta última orientación, y que permite valorar los aciertos y errores que los alumnos presentan en el aprendizaje de la lectura, escritura y matemáticas para 1o., 2o. Y 3o. grado de primaria: la de Silvia Macotela de la UNAM llamada inventario de Habilidades Básicas, y la de Rosa Ma. González de la UPN llamada Test Diagnostico en Habilidades Básicas (TEHBA).

El THEBA (1994, p. 12) fue estandarizada con una muestra representativa de la población de referencia, determinada como 'los alumnos de 1o., 2o. Y 3o grado de primaria del Distrito Federal. El instrumentos esta compuesta por tres pruebas (una por grado) y valoran tres áreas: lectura, escritura y matemáticas.

Las pruebas cuentan con índices aceptables de confiabilidad y con validez de Contenido y de Criterio. La validez de un instrumento se ha definido en tanto miden el rasgo que se pretende. La validez de Contenido se refiere a la evaluación sistemática del contenido de la prueba con objeto de determinar si los reactivos seleccionados son representativos de lo que se pretende medir. Por su parte la validez de Criterio se refiere a la relación entre los resultados de una prueba y algún tipo de criterio o resultado.

De acuerdo con la Mtra. González decidimos valorar la validez por diferencia de grupos. El criterio para diferenciar los grupos fue el conocimiento que mostraban los alumnos en cuando a los aprendizajes de las áreas que avalúa el THEBA, para lo cual seleccionamos un grupo regular y un grupo de educación especial del programa USAER, en donde canalizan a los alumnos que concluyendo el primero de primaria no alcanzan los conocimientos y habilidades suficientes en lectoescritura y/o matemáticas, de acuerdo al resultado que obtienen en la prueba Monterrey, un instrumento que como se menciono carece de indicadores para valorar características de la prueba.

La muestra estuvo por 73 alumnos que habían terminado el primer año de primaria, 38 de ellos estaban en un grupo regular y 35 clasificados como alumnos deficientes por el programa USAER, ambos grupos inscritos en la misma escuela. La escuela se ubica en una zona de clase baja media.

Para el análisis estadístico de la diferencia de los grupos se utilizo la prueba U de Mann Witney, por ser grupos independientes y el nivel de medición que arroja el instrumento es ordinal.

En el análisis de las pruebas totales de lectura, escritura y matemáticas se obtuvieron diferencias con un nivel de significancia de .05, con puntuación Z de -5.50, - 6,47 y -5.83 respectivamente.

Al realizar un análisis de cada una de las subpruebas que compone cada área se observo que tanto las subpruebas de escritura como de lectura muestran diferencias estadísticamente significativas.

En el caso de las subpruebas de matemáticas Lógicas y Relaciones y Resolución de problemas se observa que no existen diferencias con un nivel de significancia de 0.5. las calificaciones Z obtenidas en estas subpruebas son de

- 1.90 y - 1.45 respectivamente.

Una explicación posible por la cual no se obtuvieron diferencias significativas en estas subpruebas es el grado de dificultad de los reactivos que las componen. La subprueba de Lógica y Relaciones tiene un grado de dificultad mínimo (.87 promedio). En el caso de la subprueba de Solución de Problemas es contrario al anterior, está compuesto por reactivos muy difíciles (.37 en promedio).

En el caso de la subprueba de figuras geométricas, existe una diferencia estadísticamente significativa a favor del grupo USAER. Este hecho puede deberse a que esta subprueba consta únicamente de dos reactivos, lo que dificulta establecer diferencias precisas entre ambos grupos.

Por lo cual se observa que en el apartado de matemáticas no se marca tanta diferencia como en lectura y escritura.

Considerándose como posible causa el que en USAER se encontraron niños con problemas neurológicos, visomotores y de lenguaje en su mayoría considerándose como un obstáculo para la adquisición de la lectoescritura, siendo evidentes las diferencias entre lectura y escritura y no así en matemáticas.

Como ya se señaló existen una serie de cuestionamiento referidos a las pruebas objetivas, sin embargo consideramos que el diseño de pruebas de tipo formativo apoyan el trabajo del maestro y de los profesionales apoyo en educaciones básicas. En especial cuando su construcción obedece a la metodología adecuada para este tipo de instrumentos.

Por lo cual consideramos que la evaluación diagnóstica, realizada con el apoyo de pruebas como el THEBA permitirá obtener una retroalimentación que favorezca el proceso de enseñanza, tanto en grupos regulares como en grupos de educación especial.

Siendo de gran importancia el retomar la participación de los/as maestros/os para que den sus comentarios y la relevancia que para ellos, representa la aplicación de este tipo de prueba.

Así como la importante necesidad de revisar los planes y programas de estudio que se utilizaron en la Educación Básica, y el apartado de matemáticas del THEBA, debido a que no arroja los datos esperados que como ya se mencionó anteriormente puede deberse a diversas causas.

LIMITACIONES

Hubiera sido correcto seleccionar sujetos de escuelas diferentes, seleccionados aleatoriamente a fin de que la muestra tuviera un mayor nivel de representación.

Debido a limitaciones de recursos no fue posible proceder de esa forma.

REFERENCIAS BIBLIOGRAFIA

1.- ACHENBAC, Thomas M.

"Investigación de Psicología del Desarrollo"

Ed. Manual Moderno, México, 1981.

2.- AHMANN, J. Stanley.

"Los Test Psicológicos y su empleo en la escuela"

Ed. Troques, Buenos Aires, 1968 - 1969.

3.- ANASTASI, A.

"Test Psicológicos"

Ed. Aguilar, México, 1977.

4.- ARCE, C.

"Introducción a la Investigación Pedagógica"

Ed. Mc.Graw - Hill, México, 1994.

5.- ARY, D. y CHESER, J.

"Introducción a la Investigación Pedagógica"

Ed. Mc.Graw - Hill, México, 1990.

6.- AZCOAGA, J.E.

"Alteraciones del Aprendizaje Escolar. Diagnóstico, Fisiopatología y Tratamiento"

Ed. Paidós, Buenos Aires, 1985.

7.- BROWN, F.

"Principios de la medición en Psicología y Educación"

Ed. Manual Moderno, México, 1980 - 1991.

8.- BLOOM, B.; HASTING J. y MADUS, G.

"Evaluación del Aprendizaje" Tomo I.

Ed. Ediciones Troquel, Buenos Aires, 1975.

9.- CARREÑO, Fernando H.

"Enfoques y principios teóricos de la evaluación"

Ed. Trillas, México, 1986.

10.- CARREÑO, Fernando H.

"Instrumentos de medición del Rendimiento Escolar"

Ed. Trillas, México, 1985.

11.- CERDA, E.

"Psicometría General"

Ed. Herder, Barcelona, 1985.

12.- CHADWICK, Clifton.

"Evaluación Educativa"

Ed. U.P.N., México, 1970.

13.- CLARENCE, H. Nelson.

"Mediciones y Evaluación en el aula"

Ed. Kapelusz, Buenos Aires, 1971.

14.- COLLIN, G.

"Selección del Test al servicio de la Psicología Infantil"

Ed. Kapelusz, Buenos Aires, 1959.

15.- CONTRERAS, Ferto Raúl.

"Evaluación en la Escuela Primaria"

Ed. Oasis, México, 1969.

16.- CRAIC, Robert; MEHERENS, William.

"Psicología Educativa Contemporánea"

Ed. Limusa, México, 1979.

17.- FERNANDEZ, B.R.

"Psicodiagnóstico, Concepto y Metodología"

Ed. Cincel, Bogota Colombia, 1981.

18.- EBEL R.L.

"Fundamentos de la Medición Educativa"

Ed. Guadalupe Buenos Aires, 1974.

19.- FERNANDEZ, Pérez Miguel.

"Evaluación y Cambio Educativo. El Fracaso Escolar"

Ed. Morata, Madrid, 1988.

20.- FORNS, M.

"Evaluación Psicológica Infantil"

Ed. Barcanova, Barcelona, 1980.

21.- GEORGIA, S.

"Medición y Evaluación"

Ed. Herder, México, 1975.

22.- GONZALEZ, R.M.

"Test Diagnóstico en Habilidades Básicas (TEHBA)"

Ed. S.E.P. / U.P.N., México, 1994.

23.- GROUNDLUND, Norman.

"Medición y Evaluación de la Enseñanza"

Ed. Pax - México, 1973.

24.- JIMENEZ Y Coria, Laureano.

"La Evaluación del Aprendizaje y del maestro"

Ed. Porrúa, México, 1972.

25.- KARMELO, Louis J.

"Medición y Evaluación escolar"

Ed. Trillas, México, 1974.

26.- KINSBOURNE, Marcel.

"Problemas de Atención y Aprendizaje en los niños"

Ediciones Científicas, La Prensa Medica Mexicana S.A. de C.V.

27.- LAFOURCADE, Pedro D.

"Evaluación de los Aprendizajes"

Ed. Kapeluz, Argentina, 1973.

- 28.- LESTER, Tarnopool, Sc. D.**
"Dificultades para el Aprendizaje"
Ediciones Científicas. La Prensa Médica Mexicana, S.A. 1986.
- 29.- MACOTELA, Silvia y ROMAY, M.**
"Inventario en Habilidades Básicas"
Ed. Trillas, México, 1992.
- 30.- MAGNUSSON, D.**
"Teoría de los Test"
Ed. Trillas, México, 1978.
- 31.- MEHERENS, W. y LEHMAN, I.**
"Medición y Evaluación en la Educación y en Psicología"
Ed. CECSA, México, 1982.
- 32.- MERCER, CECIL D.**
"Dificultades en el Aprendizaje"
Ed. CEAC, Barcelona España.
- 33.- MORALES, M.L.**
"Psicometría aplicada"
De. Trillas, México, 1975.
- 34.- NISBET, J.D. y ENTWISTLE, N.J.**
"Metodología de Investigación Educativa"
De. Oikos-Tau. Barcelona, 1980.

35.- NUNNALLY Y BERNSTEIN

“ Teoría Psicometrica ”

Ed. Mc Graw Hill México, 1995.

36.- PAWLIK K.

“ Diagnostico del Diagnostico ”

Ed. Harder Barcelona, 1977.

37.- SALVIA, J.

"Evaluación en la Educación Especial"

De. Manual Moderno. México, 1987.

38.- SATTler, J.

"Evaluación de la Inteligencia Infantil y Habilidades Especiales".

Ed. Manual Moderno, México, 1988.

39.- SNOW.

“ Aptitudes y Rendimiento “.

Ed. Guadalupe, Buenos Aires, 1980.

40.- SZEKELY, B.

"Los Test" Manual de Técnicas de Exploración Psicológica

Ed. Kapelusz, Buenos Aires, 1966.

41.- STUFFLEBEAM, D. y SHINKFIELD, A.

"Evaluación Sistemática"

Ed. Paidós, Barcelona, 1987.

42.- TAVELLA, N.

"Los Test en la Escuela"

Ed. Universitaria de Buenos Aires, Buenos Aires, 1972.

43.- THORNDIKE, R. y HAGEN, e.

1978. "Test y técnicas de medición en Psicología y Educación"

Ed. Trillas. México, 1972.

44.- TORO, J. y CERVERA M.

1990. "Test de análisis y Lectoescritura"

Visor, Madrid.

45.- WALL. W. D.

"El fracaso Escolar"

Ed. Paidós. Buenos Aires 1970.