

**UNIVERSIDAD PEDAGÓGICA NACIONAL  
UNIDAD AJUSCO**

**ACADEMIA DE PSICOLOGIA EDUCATIVA**

**“VALIDACIÓN DE UNA PRUEBA DE RENDIMIENTO:  
CIVISMO 1º. GRADO DE SECUNDARIA.**

**T E S I S**

**QUE PARA OBTENER EL TITULO DE LICENCIATURA**

**P R E S E N T A N**

**MARIA YOLANDA DIAZ RODRÍGUEZ  
ANA YELY SANTILLAN FLORES**

**ASESORA: Dra. Rosa María González Jiménez**

**México, D.F. 1998**

# **AGRADECIMIENTO**

## **A Dios**

Por el rayo de esperanza que nos ilumina día con día, por las alegrías y satisfacciones, pero sobre todo por la FE que tenemos en TI.

## **A nuestros Padres:**

Quienes con su confianza, cariño y apoyo sin escatimar esfuerzo alguno nos han convertido en personas de provecho ayudándonos al logro de una meta mas: nuestra carrera profesional, por compartir tristezas y alegrías, éxitos y fracasos, por todos los detalles que nos han brindado durante nuestra vida y por hacer de nosotras lo que somos.

## **A nuestros Hermanos-as:**

Como un testimonio de infinito aprecio y eterno agradecimiento, por el apoyo y consejo, deseamos de corazón y que nuestro triunfo como profesionistas lo sientan como suyo.

## **A una gran amiga Luz Ma.:**

Por el apoyo, comprensión y confianza que nos brindó en los momentos más difíciles. A nuestros Familiares y Amigos: Que directa e indirectamente han formado parte de nuestro crecimiento como seres humanos y han contribuido a la realización de esta meta aportando su apoyo. A nuestra Asesora: Dra. Rosa Ma. González, por el apoyo, consejo y confianza pero sobre todo por aportar su conocimiento durante la realización de esta meta. A la Universidad Pedagógica Nacional: Por habernos brindado el espacio y recursos y quien hizo posible este anhelo.

## **A nuestros Profesores:**

Con admiración y respeto por haber contribuido sus conocimientos en el transcurso de nuestros estudios. No es fácil llegar, se necesita ahínco, lucha y deseo, pero sobre todo apoyo como el que hemos recibido durante este tiempo. Ahora más que nunca se acrecenta nuestra confianza e ilusión. Gracias por lo que hemos logrado.

**Yolanda y Ana Yely**

# INDICE

## Introducción

### Capítulo 1. La Escuela Secundaria en México

- 1.1. Creación y evolución del ciclo de secundaria
- 1.2. El currículo de educación secundaria (ES)
  - 1.2.1. Modificaciones curriculares
  - 1.2.2. Planes y programas de secundaria
  - 1.2.3. Los contenidos
- 1.3. Procesos de enseñanza aprendizaje
  - 1.3.1. Enseñanza-aprendizaje de las ciencias histórico-sociales
  - 1.3.2. Enseñanza-aprendizaje del civismo
- 1.4. Calidad educativa en secundaria
  - 1.4.1. Baja calidad educativa en el ciclo de educación secundaria (CES)
  - 1.4.2. El fracaso escolar en el ciclo de educación secundaria (CES)
  - 1.4.3. Articulación entre niveles
  - 1.4.4. Condiciones en cuanto a las proporciones; profesor-alumno
- 1.5. Evaluación Educativa (EE)
  - 1.5.1. La problemática de la evaluación
  - 1.5.2. Modelos Evaluativos
    - A) Ralph W. Tyler
    - B) Lee J. Cronbach
    - C) Michel Scriven
    - D) Daniel Stufflebean
  - 1.5.3. Investigación evaluativa
  - 1.5.4. Funciones de la EE
    - A) Evaluación Sumativa
    - B) Evaluación Formativa
  - 1.5.5. Los objetos de la EE

- 1.5.6. Técnicas de evaluación
  - A) Técnicas cualitativas
  - B) Técnicas cuantitativas de Rendimiento
- 2.1. Antecedentes
- 2.2. Desarrollo de las pruebas de rendimiento
  - 2.2.1. Desarrollo de las PR en México
- 2.3. Aspectos generales de las PR
  - A) Usos de las PR
  - B) Cuestionamientos a las PR

### **Capítulo III. Aspectos Psicométricos de las pruebas**

- 3.1. Medición
  - 3.1.1. Escalas de medición
  - 3.1.2. Errores de medición
- 3.2. Confiabilidad
- 3.3. Validez
  - 3.3.1. Validez de contenido
  - 3.3.2. Validez de criterio
  - 3.3.3. Validez de constructo

#### **Metodología**

#### **Resultados**

#### **Conclusiones**

#### **Anexo 1**

#### **Anexo 2**

#### **Anexo 3**

#### **Anexo 4**

#### **Bibliografía**

## INTRODUCCION

La presente investigación se refiere al proceso de diseño, construcción y validación de una Prueba de Rendimiento de Civismo 1o. grado de Secundaria, y se inscribe en el tema más general de la evaluación sumativa.

En México se reflejan escasas investigaciones en el campo educativo; además de falta de recursos humanos y materiales algunos autores señalan como limitaciones la ausencia de técnicas e instrumentos validados para la población mexicana.

En este sentido, este trabajo pretende ser un aporte técnico para el ciclo de secundaria, a través del diseño y validación de una prueba de rendimiento en la materia ya mencionada.

Las pruebas de rendimiento son instrumentos que miden el dominio y la destreza en distintas áreas del conocimiento (Ary y Cheser, 1990). Para el diseño y construcción de la prueba se diseñó un plan de prueba, en el cual se redactaron los reactivos, teniendo como referencia los libros de texto de la materia y grado.

La validación del instrumento se realizó con una muestra de 120 alumnos-as de 6 escuelas secundarias técnicas oficiales del turno matutino, ubicadas en el D. F .

En el primer capítulo se presenta una panorámica general del origen y evolución del CES. Se presenta un apartado acerca del currículo en el cual se menciona el manejo que se hace en este nivel en cuanto a planes y programas, también se hace una breve reseña sobre la enseñanza-aprendizaje de las Ciencias Histórico-Sociales, dándole mayor importancia a la enseñanza-aprendizaje del Civismo; y para concluir con una breve revisión acerca de la calidad educativa en este nivel.

En el siguiente apartado, se abordó la evaluación desde algunos aspectos como: la problemática de la evaluación desde los siguientes puntos: modelos evaluativos, investigación evaluativa, funciones de la evaluación, y los objetivos de la evaluación

educativa; por otra parte, las técnicas de evaluación en las cuales se hace referencia a las técnicas cualitativas y cuantitativas, en estas últimas se incluyen las pruebas de rendimiento.

En el capítulo II, analiza, con mayor detalle, el desarrollo de las pruebas de rendimiento a nivel general y en México, así como los usos y el cuestionamiento de estas pruebas.

El siguiente capítulo se refiere a una explicación del proceso de validación en este capítulo se abordaron los siguientes temas: medición, confiabilidad y validez, en este se explican algunos tipos de validez: de contenido, de criterio y de constructo.

Por último, se describe la metodología que se siguió para la construcción y validación de la prueba, los resultados y conclusiones del presente trabajo.

# **CAPITULO 1.**

## **LA ESCUELA SECUNDARIA EN MEXICO**

Este capítulo comienza presentando los orígenes y evolución del Ciclo de Educación Secundaria {CES). También se aborda el tópico curricular, analizando los dos últimos planes y programas de estudio de la escuela secundaria {ES) nacional. Por último se describen algunos factores relacionados con el fracaso escolar en este ciclo, temáticas muy relacionadas con la evaluación.

Desde este momento es necesario señalar una situación reconocida referente a la escasa producción de investigaciones en el ciclo de secundaria, como lo señalan Quintanilla y Galván (1995):

"La educación secundaria no ha sido estudiada; lo mismo podría decirse sobre la enseñanza técnica agrícola e industrial, la educación preescolar y la privada, así como de otras modalidades de enseñanza formal, como la militar."

### **1.1. CREACION Y EVOLUCION DEL CICLO DE SECUNDARIA**

A principios de siglo la juventud mexicana que deseaba continuar estudiando al término de la primaria. solo podía optar por dos alternativas: la Escuela Nacional de Maestros (ENM) y la Escuela Nacional Preparatoria {ENP). La primera como opción terminal para formarse como normalistas, la segunda como opción propedéutica para continuar estudios universitarios.

Recién concluida la revolución. se crea la Secretaría de Educación Pública (SEP), que procuró una organización más definida en cuanto a los ciclos escolares, para atender la demanda creciente de alumnos que concluyen la primaria. así como se plantea una política mas definida que busca ampliar el nivel educativo de la población. a la vez que reconoce necesidades específicas de la adolescencia.

Un primer intento de creación de la ES, como un ciclo intermedio entre la primaria y la preparatoria, se realizó en Veracruz en el Congreso Pedagógico del Estado. Se instituye a través de una Ley estatal expedida en Octubre de 1915.

Durante el gobierno de Plutarco Elías Calles. siendo Secretario de Educación Pública, el Dr. José María Puig, se consumó la Reforma Educativa que dió lugar al Ciclo de Educación Secundaria (CES), mediante dos decretos inspirados por el creador de este ciclo: Moisés Sáenz. Aún en este año en Veracruz continuaban funcionando algunas secundarias.

Los dos decretos inspirados por Moisés Sáenz, mencionaban lo siguiente: el primer decreto del 29 de agosto de 1925 autoriza ala SEP la institución de escuelas secundarias a nivel nacional; el segundo del 1o.de enero de 1926, niega la aceptación de alumnos egresados de primaria a la ENP, dándose una separación orgánica: el ciclo preparatorio que constaba de 5 años de estudio queda en 2 años dependiendo de la Universidad; por su parte la secundaria se instituye con una duración de 3 años. dependiendo de la Dirección General de Educación Secundaria de la SEP (Mejía, 1976).

Los principios que definen la pertenencia de la secundaria a la SEP se relacionan con un criterio democrático que menciona que la educación en Cualquier nivel debe ser para todos los ciudadanos. Este principio de igualdad se conserva como un "ideal". .ya que en la realidad grandes sectores de la población no concluyen la primaria. En el caso de la ES, en la actualidad, es considerado este ciclo como básico y obligatorio.

Los objetivos que se plantearon para la escuela secundaria fueron (Sáenz, 1928):

- Ampliar y perfeccionar la educación primaria.
- Preparar al alumno como futuro ciudadano.
- Capacitarlo como trabajador de la producción y distribución de riquezas.
- Cultivar la personalidad independiente y libre del adolescente.



Por lo tanto, Sáenz insistía en la formación de la personalidad del adolescente, así como se buscaba educar para la vida, este objetivo se lograba por medio de talleres en los cuales se aprendía un oficio. Se pretendía que la ES, junto con la primaria, formara a los niños y jóvenes como ciudadanos responsables, poseedores de una cultura universal y orgullosos de ser mexicanos.

La ES en sus inicios se encontraba restringida a las zonas con mayor desarrollo urbano y captaba a grupos minoritarios de la población que contaba con los recursos suficientes. Ya que en 1930 se reporta a nivel nacional 21,757 alumnos inscritos en secundaria y 109 escuelas.

## **1.2. EL CURRÍCULO DE EDUCACION SECUNDARIA (ES)**

En este apartado se presentan los problemas relativos al currículo del CES en México. Los incisos se ordenan analizando los planes y programas de estudio, los contenidos de los programas y los procesos de enseñanza-aprendizaje; en especial en el área de las Ciencias Sociales, resaltando en estas la asignatura de Civismo. Esto es fundamental para la construcción de la Prueba de Rendimiento, de ésta investigación, ya que su validez es de contenido basándose en el programa de estudios de secundaria.

El currículo se refiere a la planeación educativa que va desde los aspectos más generales, de la política educativa, a los más específicos, que orientan su instrumentación llevada a cabo por las escuelas. Hay una cantidad de problemáticas que atraviesan el campo curricular, nuevamente se recomienda la necesidad de promover investigaciones que permitan conocer sus implicaciones para la ES nacional. Así mismo, este es uno de los objetivos de la presente investigación. A continuación se comentan algunas problemáticas que presenta el campo del currículum en general, a nivel secundaria: .Los fines que persigue la escuela y el quienes los determinan (organización, contenidos, estrategias de enseñanza ).

- El currículum debe tener en cuenta las demandas de las instituciones (familia, sistema educativo, comunidad, políticos, la industria, la religión, grupos ecologistas, de salud, feministas, etc.)

- El definir cuales son los objetivos educativos y quienes los determinan, (grupos de profesores, administración central o federación)
- El problema de la obligatoriedad en las escuelas secundarias no es ocasionado por el sistema educativo, sino porque pequeños sectores no realizan su inscripción o no egresan de este nivel.
- Según Quiroz ( 1994 ), de cada 100 niños que se inscriben en primero de primaria solo 28 concluyen la secundaria.
- Es evidente que el problema no esta en la satisfacción de la demanda, sino en las posibilidades del sistema educativo de retener a los estudiantes hasta la conclusión de sus estudios secundarios.
- Sumado a todos estos factores se encuentra el problema de la deserción que puede ser ocasionada por circunstancias externas como internas.

Algunos de estos problemas se abordaran durante el desarrollo de este capítulo.

Es evidente que la evaluación educativa tiene, o debe tener, gran participación durante la solución de éstas problemáticas, y durante ésta investigación se presenta explícitamente una forma de evaluación, sin embargo, debe quedar claro que no es la única y no es posible solucionar las problemáticas, a nivel secundaria, con una sola forma de evaluación.

### **1.2.1. MODIFICACIONES CURRICULARES**

Al currículum de educación básica se le han hecho varias modificaciones en los últimos 35 años, ya que el plan de estudios de 1957 estaba organizado por asignaturas: de Historia, Geografía y Civismo. En los 70's se organizaron los programas por áreas y se elaboraron los libros de texto gratuitos. En 1974 con el Congreso de Chetumal se restauró la organización curricular por asignaturas con las materias de Historia, Geografía y Educación Cívica, esta restauración se implantó en el sistema educativo de secundaria.

Nuevamente en 1983 se volvió a hacer una modificación, cambiando la organización curricular por áreas, sin embargo muchas escuelas conservaron el programa por asignaturas.

En esta misma década en el país, la enseñanza se empezó a tomar en cuenta dentro de la investigación educativa.

Dentro de las instituciones que han hecho Investigaciones sobre el aspecto curricular se encuentran: UPN (Sinaloa), UAS, Facultad de Psicología de la UNAM, UAG, ENP de la UNAM, UAM de Azcapotzalco y el Departamento de Investigaciones Educativas.

El conocimiento generado en este ámbito no es el resultado sumado de investigaciones puntuales, sino de una serie de estudios donde las estrategias se pueden y se complementan en la persecución de un mismo objetivo, y por tanto es una función de los objetivos y de las perspectivas, comparativas por un grupo determinado de investigadores.

En opinión de Díaz Barriga (1995), las nociones relativas al campo de la Investigación curricular, en la década de los 80's se vincula con la concepción de currículum como producto, como proceso o bien como práctica social educativa.

Tomando en cuenta estas circunstancias en 1993 surge otra modificación en el currículum de nivel básico, volviendo de nuevo a la organización por asignaturas.

A continuación, se expondrán las modificaciones realizadas a los planes y programas de ES.

### **1.2.2. PLANES Y PROGRAMAS DE SECUNDARIA**

Los planes y programas de secundaria en México han sido reformulados sucesivamente en los años 1932, 1936, 1941, 1945, 1947,1953, 1960, 1974,1991 y 1994. Si bien no se ha investigado con mayor detalle los cambios experimentados en las diferentes modificaciones, diversos autores (Mejía, 1976; Muñoz Izquierdo, 1983, Guevara, 1992), han señalado que hasta los setenta los programas de estudio de secundaria solo sufrieron modificaciones menores, en términos de contenidos y horas asignadas a diferentes materias; al respecto Mejía (1976; p.153) comenta:

"...los diez planes de estudio vigentes durante 50 años en la escuela secundaria mexicana, los cambios registrados consisten en simples barajamientos de asignaturas y actividades, en

función del aumento o disminución de las sesiones de clase a ellas destinadas".

Las modificaciones que se realizaron en las décadas 30's y 40's solamente cambiaron en la duración que se le dedicaba a cada materia.

Producto de las "Resoluciones de Chetumal" en 1974 se adoptó la organización de los estudios por áreas de conocimiento (Ciencias Sociales y Ciencias Naturales), como la alternativa al plan de estudios por asignatura. Hasta 1992 coexistieron los dos planes de estudio: el 75% de las escuelas seguía el plan de áreas y el 25% continuaba el plan por asignaturas. (Guevara Niebla, 1992).

En referencia a los planes y programas de estudio vigentes hasta 1992, Guevara ( 1992) menciona:

“No se ha hecho hasta ahora una evaluación sistemática que permita medir con precisión las ventajas de uno u otro sistema curricular, aunque existe la idea generalizada de que el estudio por áreas de conocimiento a repercutido en un empobrecimiento de la educación secundaria...”

Como antecedentes de los nuevos programas de estudio, y producto del PPME, en 1989 se instituye un proceso de consulta, diagnóstico y elaboración del plan de estudios para secundaria y los programas que lo constituyen; así mismo, se intentó identificar los principales problemas educativos del país, a fin de precisar las prioridades y definir estrategias para su atención.

El plan de estudios establece como prioridad esencial (SEP, 1993, p.13):

"... contribuir a elevar la calidad de la formación de los estudiantes que han terminado la educación primaria, mediante el fortalecimiento de aquellos contenidos que responden a las necesidades básicas de aprendizaje de la población joven del país y que sólo la escuela puede ofrecer".

En mayo de 1992, la SEP inició la última etapa de la transformación de los planes y programas de estudio. El nuevo plan de estudios se estructura por asignaturas y establece cinco prioridades:

1. Asegurar que los estudiantes profundicen y ejerciten su competencia para utilizar el español en forma oral y escrita, desarrollando la capacidad de expresión y claridad.
2. Ampliar y consolidar los conocimientos y las habilidades matemáticas y las capacidades para aplicar la aritmética, el álgebra y la geometría.
3. Fortalecer la formación científica de los estudiantes y superar los problemas de aprendizaje que se presentan en este campo. (Referente a las Ciencias Naturales).
4. Profundizar y sistematizar la formación de los estudiantes en Historia, Geografía y Civismo, al establecer cursos por asignaturas que sustituyen a los de áreas de Ciencias Sociales.
5. El aprendizaje de una lengua extranjera (Inglés o Francés) destacando los aspectos de uso más frecuentes en la comunicación.

En el plan actual se conservan actividades para la formación integral del estudiante: la expresión y la apreciación artísticas, la educación física y la educación tecnológica. Así mismo, se establece la obligatoriedad de este ciclo.

Para el primero y segundo año este nuevo plan se empieza a aplicar en el año escolar 1993-1994. Mientras que en tercer grado se aplica en el siguiente año escolar .

Para el diseño y adecuación de los programas de estudio vigentes en las escuelas secundarias se realizó una prueba piloto llamada " Prueba Operativa" en 1990, a fin de probar su pertinencia y viabilidad, sin embargo, no se han hecho públicos los resultados de la evaluación.

En relación con el mapa curricular (Ver cuadro 2) Quiroz (1995), comenta que es muy

parecido a la modalidad de asignaturas de las II Resoluciones de Chetumal". Las diferencias son que se agrega una hora semanal a Español y Matemáticas; en Física y Química se fusionan los dos cursos uno introductorio en primer grado y se agrega una hora en cada materia en 20. y 30; el curso de tercer grado de Civismo se sustituye por uno de Orientación Educativa de 3 horas por semana y en el tercer año se agrega una asignatura opcional que se decidirá en cada entidad federativa. Habría una carga semanal de 35 horas en lugar de las 32 horas del plan anterior .

En cuanto a las orientaciones didácticas, comenta Quiroz (1995), aparecen en la parte introductoria de los programas y. frecuentemente articuladas al enfoque pedagógico de la disciplina. Los contenidos escolares son presentados como temas a estudiar, a diferencia de los objetivos particulares y específicos que aparecían en los programas anteriores.

Asimismo, este programa pretende que el maestro tenga la posibilidad de seleccionar aquellos objetivos y contenidos que considere más convenientes, otorgándoles un lugar de mayor relevancia.

### **1.2.3. LOS CONTENIDOS**

Los contenidos de un programa designan el conjunto de saberes o formas culturales cuya asimilación y apropiación por los alumnos se considera esencial para su desarrollo. Es importante mencionar que se conciben dos formas de contenidos que forman un programa de estudios. Hasta los años setenta la educación escolar se concebía como la realización de una serie de aprendizajes de contenidos específicos, sistemáticamente planificados, .gracias a los cuales los alumnos incorporan e interiorizan los conocimientos a los que la sociedad concede importancia en un momento determinado. Esta orientación ha estado históricamente acompañada con una interpretación de la enseñanza y el aprendizaje en términos de pura transmisión y recepción.

A partir de Piaget y el constructivismo se considera de otra manera la educación escolar, ya que en esta se aseguran las condiciones óptimas para que los alumnos desplieguen sus potencialidades y capacidades cognitivas, afectivas, sociales y de aprendizaje. En pocas

palabras, esta orientación le da mayor importancia a la actividad del alumno por medio de la creatividad y del aprendizaje por descubrimiento; relativizando la importancia de los contenidos y concibe al profesor como un facilitador más que como un transmisor del saber constituido.

En cuanto a las formas de aprendizaje que se busca propiciar en el educando, se distinguen dos modalidades: el pensamiento convergente y el divergente. El primero evoca ideas y trata de encadenarlas para llegar a una conclusión predeterminada, así como se busca una respuesta correcta. El segundo se caracteriza por la producción de soluciones alternativas, estimulando la búsqueda y creación de nuevos saberes (Rodríguez, 1990; Stephen, 1988; Ausubel, 1978).

En educación básica otro aspecto relevante en relación con los contenidos tiene que ver con su orientación. Antiguamente los contenidos se concebían desde la lógica de la especialidad; en años más recientes se privilegia una lógica pedagógica para su enseñanza. El principal argumento es que el alumno de secundaria todavía no tiene definida una especialidad futura y en consecuencia la formación que se ofrezca debe ser la que corresponde a una persona común, no especialista.

En relación con los contenidos el PPME señala que se debe promover el paso de contenidos informativos que suscitan aprendizajes fundamentalmente memorísticos, a aquellos que aseguren la asimilación y recreación de valores, el dominio y uso cada vez más preciso y adecuado tanto de los diversos lenguajes de la cultura contemporánea como de métodos de pensamiento y acción que han de confluir en el aprendizaje. En pocas palabras, se pretende un aprendizaje significativo.

Un aspecto importante, que no permite que este objetivo se cumpla es la preparación que tienen los maestros, ya que según González (1994), menciona:

“ La función primordial de la escuela y el maestro es dirigir, orientar y facilitar el aprendizaje de los alumnos-as; el grado de eficiencia con que se efectúe la evaluación,

podrá determinar en buena medida el éxito o fracaso de la enseñanza" ..

Por lo tanto, si los maestros no tienen una buena preparación, no podrán proporcionar un aprendizaje significativo.

También es posible que este hecho no se haya realizado, ya que no se han especificado las propuestas concretas para conseguirlo, un ejemplo de esto lo encontramos en la materia de Civismo que se enseña de manera memorística, por lo tanto, no se ha llevado un cambio relevante en cuanto a contenido se refiere.

Por tal motivo, en la construcción de la prueba que se realizó, durante esta investigación, se plantearon reactivos que pertenecen al nivel cognoscitivo de comprensión y aplicación, además de conocimiento, para poder responder en cierto grado al aprendizaje significativo, pero estos reactivos representan un porcentaje bajo durante el contenido de la prueba.

### **1.3.PROCESOS DE ENSEÑANZA-APRENDIZAJE**

El proceso de enseñanza-aprendizaje es generalmente definido como un conjunto de fases sucesivas en que se cumple el fenómeno intencional de la educación y de la instrucción. El concepto moderno conforme a las actuales tendencias pedagógicas determina que enseñanza y aprendizaje no son factores independientes, sino que existe una interacción, ya que el tipo de enseñanza dependerá de cuál sea el aprendizaje del que recibe dicha enseñanza. (Bachillerato Pedagógico No.2, 1989).

Los procesos de enseñanza-aprendizaje han ido modificándose a través de los tiempos, producto de las reflexiones y cuestionamientos de carácter curricular, además del cambio de paradigmas en tomo a las diferentes corrientes de pensamiento en el campo de la educación. Anteriormente se tenía en cuenta, casi exclusivamente, la enseñanza como un aspecto general de habilidades y técnicas específicas a ser aplicadas por el maestro a los educandos. En las dos últimas décadas se han realizado una serie de investigaciones que



han dado lugar a propuestas referentes a los contenidos específicos por áreas y materias .

### **1.3.1. ENSEÑANZA-APRENDIZAJE DE LAS CIENCIAS HISTORICO -SOCIALES**

Los cambios que han surgido en el área de las Ciencias Naturales han influido en el campo de las Ciencias Sociales, por lo tanto, a continuación se presenta una breve reseña de lo que han sido las modificaciones en las Ciencias Naturales:

Inicio (1860-1920); no hubo prácticamente investigación que apoyara los aspectos curriculares de la enseñanza de las ciencias;

El utilitario-libros de texto (1920-1957), se empieza a cuestionar los contenidos de los programas de estudio; y como estrategias de enseñanza predominaban la memorización y el uso del libro de texto.

La primera revolución en la enseñanza de las Ciencias Naturales (1957-1978); en este período es notable un esfuerzo en el campo de la innovación curricular ya que se pretende que la enseñanza de las ciencias tenga una orientación práctica y que la ciencia. se vea como un proceso. Surge como estrategia de enseñanza-aprendizaje el "aprendizaje por descubrimiento".

La segunda revolución en la enseñanza de las Ciencias Naturales (1980); en este período se da la tradición Piagetiana, mejor dicho la escuela constructivista. Se le da al alumno un papel más activo en el aprendizaje relacionando su ambiente exterior con el ambiente escolar.

Sin embargo Taboada (1995), señala la existencia de un reducido número de trabajos sobre la enseñanza y el aprendizaje de las Ciencias Histórico Sociales en comparación con otros campos que multiplican este número; en la educación básica y media. También es

importante mencionar que la comunidad científica de este campo esta en proceso de formación.

### **1.3.2. ENSEÑANZA-APRENDIZAJE DEL CIVISMO**

En este apartado se presentará una breve reseña de lo que es el Civismo, aunque sus investigaciones son muy antiguas y escasas, como ya se mencionó anteriormente. Después se presentará un breve esbozo sobre los Derechos Humanos, ya que es un tema básico en la asignatura del Civismo. Por último, se mencionarán algunas características que son relevantes para la enseñanza de este último.

Según Barrera Valdez (1974), la palabra Civismo se deriva del latín "civis" que quiere decir ciudadano y fue usada por primera vez durante la época de la Revolución Francesa.

Por su parte, Ramírez (1950), menciona que el Civismo debe entenderse en dos sentidos claramente diferenciados, uno de los cuales es eminentemente social, en tanto que el otro es esencialmente político. Por mera comodidad de expresión se habrá de decir, pues, que hay un Civismo social y un Civismo político, por más que se entienda que ambos se integran y funden, constituyendo una cosa únicas pues tomada la palabra en un sentido o en el otro, la enseñanza del civismo ha de conducir siempre a una sola meta que es la buena ciudadanía.

Mientras que para García Ruíz (1960), el Civismo es una de las fórmulas principales de actividad que debe de adoptar la escuela. El Civismo no puede ser en forma alguna una materia de memorización, ni práctica viciosa de acumulamiento de preceptos de urbanidad, sino por el contrario, un sentido de acción para la niñez. Su enseñanza debe consistir en una serie de actividades auténticas y para ser más preciso, la vida misma en sus más ricas y fecundas manifestaciones, dentro de las aulas. Es la escuela la continuación del hogar y de ellos se deriva una responsabilidad: la de ser una escuela de verdadero Civismo, es decir , escuela de integración y de progreso.

Por lo tanto, se puede decir que el Civismo es esencialmente importante para favorecer la formación de los ciudadanos que necesita la sociedad, ya que, con el Civismo no sólo se apoya a la sociedad, sino también al núcleo familiar y este es uno de los objetivos que tiene la escuela.

Sin embargo, como ya se mencionó al inicio de este apartado siguen sin existir avances informativos de esta asignatura, con esta investigación se espera aportar de cierta manera un avance en la información de la asignatura.

De las investigaciones mencionadas por Taboada (1995), señala que los temas que se abordan son:

Las concepciones del nacionalismo en los niños y adolescentes y el de la formación para la paz y los derechos humanos. Este último es de sumo interés para la asignatura del Civismo.

La Ley General de Educación establece como uno de los fines del sistema educativo propiciar el conocimiento y respeto de los derechos humanos (artículo 70. fracción VI).

Fuentes Molinar (1994), menciona la relación entre la educación básica y los derechos humanos:

“... la educación básica como derecho de los niños y la experiencia escolar como espacio de apropiación y de práctica de los derechos humanos”.

Siguiendo con Fuentes Molinar ( 1994 ), se "señala una marginación "precoz" de la escuela viéndose con esto cancelado el derecho de la educación para todos. En cuanto al terreno curricular está de acuerdo en que se de la formación cívica y ética tomando como base los derechos humanos. Sin embargo, propone evitar que el tema se convierta en una asignatura y aparezca disperso en el currículum como forma de mirar y hacer preguntas a otros contenidos, no sólo los de formación cívica, sino en la historia y los estudios sociales, en las ciencias de la naturaleza y en el lenguaje y la literatura. También menciona que estos

conocimientos se deben de enseñar de forma que los niños y jóvenes los puedan aplicar dentro y fuera de la escuela .

Por su parte, Ramírez Hernández (1994), señala que en la actualidad el término "Derechos Humanos" aparece en los nuevos planes de estudio y su tratamiento se circunscribe exclusivamente, en el programa de Educación Cívica, a través del conocimiento y comprensión de los derechos y deberes. También comenta que en México esta preocupación de relacionar la educación con los derechos humanos es reciente .

Gil Cantero (1994), señala que en la educación secundaria existen dos aspectos importantes para la enseñanza de los derechos humanos:

- Incrementar la información sobre los derechos humanos;
- Ampliar las experiencias con el entorno social hacia situaciones más reales.

De esta manera se señala que la Educación Cívica debe ser constructiva sobre todo en el tema de los derechos humanos para poder aplicar estos conocimientos en la vida cotidiana, así como innovadora y actualizada, ya que la escuela es el primer lugar donde deben de practicarse los derechos humanos.

Por último se mencionarán algunas características relevantes en la enseñanza del Civismo: Aunque algunos autores (García Ruíz, 1960), mencionan que la enseñanza del Civismo no debe aprenderse memorísticamente, en la realidad este aprendizaje si es de memoria. Se aprenden los derechos sin razonarlos.

En muy pocos casos utilizan material didáctico (solo en exposiciones).

Los libros de texto sólo se quedan en contenidos de conocimiento.

Diversidad de contenido en los libros de texto.

Los libros de texto en algunos casos contienen pocas actividades de comprensión y aplicación, y en otros estas son nulas.

El programa es extenso.

La evaluación en ésta asignatura, como en otras teóricas, en su mayoría es por medio de exámenes escritos, en los cuales las preguntas solo son de contenido de conocimiento, no se llega a la comprensión ni a la aplicación; y además se toman muy en cuenta las tareas (Quiroz, 1994).

Con todo lo anterior, se puede decir, que la enseñanza del Civismo se encuentra limitada únicamente al conocimiento sin llegar a una reflexión, por lo tanto, el sujeto no sabe en que circunstancias aplicar sus derechos .

Así pues, con esta investigación se pretende que el Civismo ya no sea sólo un aprendizaje meramente memorístico, sino crítico y significativo, por lo tanto, en el instrumento que se realizará se tomarán en cuenta, además del nivel cognitivo de conocimiento, los niveles de comprensión y de aplicación, estos se especificarán más adelante. Pero debe quedar claro, que no es el único medio para que la enseñanza cambie, es importante la participación del profesor, como ya se mencionó anteriormente.

#### **1.4. CALIDAD EDUCATIVA EN SECUNDARIA**

Hasta la década de los setenta había la firme convicción de que la educación escolar era un bien para el individuo, el mejor camino para la movilidad social y que constituía un motor de la prosperidad en un país. Bajo un razonamiento económico el desarrollo educativo era un factor clave en la generación de riqueza, en este contexto se favoreció el incremento al presupuesto, que permitió la incorporación de grandes sectores de la población a la educación, en especial la básica. Sin embargo, debido a las sucesivas crisis económicas se fueron frustrando las esperanzas puestas en la escuela.

Lo anterior se fundamenta con lo que menciona la Organización de las Naciones Unidas (ONU), que considera que la educación que se ofrece a la población, es la mejor inversión que puede realizar un país. La evidencia histórica y experiencias cercanas demuestran la alta correlación que existe entre una educación básica de calidad y las posibilidades de desarrollo nacional (Programa Nacional de Acción, 1994 ).

En México se ha señalado reiteradamente la baja calidad de los servicios que se ofrecen en el sistema nacional de educación. Esto es esencialmente cierto en el llamado ciclo básico (preescolar, primaria y secundaria), la SEP (1991) reporta una eficiencia terminal promedio del 60%; al respecto de la calidad.

Por otra parte, (y con Ivan Llich a la cabeza) investigadores de todo el mundo (Apple, 1987; Delval, 1989) empezaron a señalar a la institución educativa como la principal responsable de las deficiencias que los menores presentaban. Los antes llamados "problemas de aprendizaje" atribuidos al alumno o a su contexto sociofamiliar, empezaron a denominarse "fracaso escolar" , fracaso de la escuela y de toda la sociedad .

Según Gutiérrez ( 1994 ), el fracaso escolar es el aprovechamiento escolar que tiene como consecuencias, la repetición de una materia escolar, la interrupción de estudios, el traslado a una rama de educación especial y la deserción.

En relación al fracaso escolar Schmelckes ( 1995 ), analiza tres factores que influyen en deterioro de la calidad educativa: la demanda, la oferta y la interacción entre demanda y oferta.

En cuanto a la demanda estaría el nivel socioeconómico de la familia. Los alumnos con bajos recursos tienen menos probabilidades de aprender y-o permanecer en el sistema educativo para lograr una educación satisfactoria. El capital cultural de la familia es otro aspecto relevante de esta categoría; se ha observado que la mayor educación de los padres se relaciona de manera positiva con el rendimiento de sus hijos.

En relación con la oferta la autora señala la relación entre relevancia de los contenidos, la falta de práctica pedagógica al interior del aula, la inversión en actualización docente, material didáctico, apoyo bibliotecario y el tiempo invertido por el docente a su práctica escolar.

El tercer gran rubro es la interacción entre la oferta y la demanda, por un lado, la relación entre la escuela y la comunidad; por el otro, entre el sistema educativo y la sociedad. Es en el proceso educativo mismo donde se define la escolarización y el éxito o fracaso de los niños. Dentro de este proceso intervienen varios factores entre los que destacan los aspectos socioculturales: el nivel socioeconómico, los conflictos familiares, la delincuencia, la drogadicción, etc.; y por el otro lado, la dinámica al interior de la institución escolar: carencia en la planeación educativa, falta de profesionalismo por parte del profesor, y un currículo que no se cumple.

En cuanto al deterioro de la calidad educativa en la ES, en la medida del menor interés por parte de autoridades e investigadores, es posible suponer que es muy significativo. Aunque son pocas las referencias hacia este ciclo (en casi todos los casos se refieren a primaria), a continuación se presenta un panorama muy general referido al fracaso escolar en México, así como las causas que influyen en este.

Para los propósitos de este trabajo habría que resaltar algunos aspectos a evaluar de la calidad educativa.

#### **1.4.1 BAJA CALIDAD EDUCATIVA EN EL CICLO DE EDUCACION SECUNDARIA (CES)**

Los fines de la ES, siguen siendo los mismos señalados por su iniciador Moisés Sáenz. Pero en un diagnóstico del SEN presentado por el PPME, se enfatiza que:

"La estructura académica vigente ha dejado de responder a las expectativas sociales de la población, ya que la secundaria actual, salvo la modalidad técnica, no ofrece su incorporación a la vida productiva..."

Por lo tanto, aunque se sigue tomando en cuenta los fines señalados por Moisés Sáenz, actualmente no se cumplen dentro de la ES.

Como consecuencia a esta posición, en los últimos 10 años se han abierto mayoritariamente escuelas de tipo técnico o telesecundarias.

Una nueva finalidad de la ES, apoyada en el neoliberalismo, es fomentar la competitividad. Como lo menciona Ornelas (1995) actualmente se requieren obreros calificados, técnicos, gerentes y profesionales para ocupar los puestos de trabajo bien pagados y adaptarse al cambio tecnológico.

No deja de sorprender que el tema de la educación pasó de ser una ocupación más vinculada a campos del saber filosófico y pedagógico, para convertirse en prioridad del ámbito económico.

La política que sucesivos gobiernos han impulsado, buscaba garantizar espacios educativos para toda la población (en especial para la primaria). Sin embargo, no se agota el problema con garantizar un espacio para los menores, es necesario precisar para que se está en ese espacio; hay que pretender también educar y definir sus fines, aspecto crucial en la definición curricular y en las formas en que éste espacio se evalúa.

El SEN argumenta sus fines de acuerdo a una serie de propósitos que consideran diversos grupos debe de llevar a cabo la escuela, como la familia que pretende un desarrollo armónico de sus hijos, y una formación de calidad que los capacite para la vida futura.



También están los grupos sociales que buscan vía escolar conformar ciudadanos responsables y participativos en su comunidad; otros sectores demandantes son los grupos religiosos, preocupados por lo que consideran una formación moral; y los grupos de empresarios que pretenden ciudadanos capacitados para ingresar a la planta laboral.

Por medio de la evaluación educativa se pretende elevar la calidad del servicio de la ES. Para lograr esto la evaluación se tiene que enfrentar a diversas problemáticas que son:

- Delimitar y definir las formas de evaluación a nivel secundaria (no sólo por medio de exámenes).
- Modificar el manejo de los talleres, para que realmente capaciten para un posible trabajo.
- Uniformar, por medio de una modificación curricular, el programa de estudios para que la calidad que ofrecen cada una de las distintas instituciones de ES para que ofrezcan las mismas oportunidades de ingresar al nivel medio superior, a los alumnos.
- No sólo se debe valorar los aspectos del rendimiento escolar, sino también se debe de valorar el contexto social e individual del alumno.
- El fracaso escolar (el cual se explicará más adelante).

Más adelante se mencionarán las técnicas de evaluación que ayudarán a solucionar estas problemáticas.

#### **1.4.2. EL FRACASO ESCOLAR EN EL CICLO DE EDUCACION SECUNDARIA (CES)**

Según Guevara Niebla (1991), CEMEDIN (1991).y López, E.O. (1990), el fracaso escolar en México en la educación básica se considera una situación "crítica", si se atiende a los siguientes indicadores:

- El 43% de los alumnos no concluyen educación primaria;
- El 25% de los alumnos o terminan la enseñanza secundaria;
- Sólo el 66% de los alumnos terminan la primaria satisfactoriamente;
- Sólo 3 de cada 10 alumnos terminan la educación primaria;
- El 50% reprueba o deserta en los primeros 3 años.

Las estadísticas del sector educativo no presenta indicadores directos de la calidad de los servicios que ofrece, aunque si existen datos que indirectamente son útiles para medir la calidad: reprobación, deserción y eficiencia terminal.

A falta de información "oficial" sobre el fracaso escolar se recurre a fuentes públicas. De acuerdo a una investigación hecha por Guevara Niebla ( 1991 ), en la que se aplicaron dos exámenes nacionales, uno en escuelas primarias y otro en secundarias, para medir el aprovechamiento de los alumnos. Para obtener la información el estudio comprendió cuatro áreas: Matemáticas, Español, Ciencias Sociales y Ciencias Naturales.

El promedio general obtenido en la prueba fue de 3.97 sobre diez. Sólo 3.8% de los alumnos obtuvo calificaciones promedio superiores a 6. únicamente el 17% superó, en promedio, el 5 de calificación. Las escuelas particulares (4.58 puntos) calificaron por encima de las públicas (3.88), aunque la diferencia relativa fue menor .

Dichos resultados, según el autor, fundamentan la hipótesis de la baja calidad de la educación en el país, por lo tanto urge tomar medidas para su progreso.

Otras investigaciones. (Tirado, 1986), coinciden con Guevara en cuanto a los bajos rendimientos obtenidos por los alumnos de las ES.

Por otro lado, Quiroz (1994), menciona que de cada 100 estudiantes 26 reprueban cuando menos una materia por año escolar, esto lo relaciona con la forma de evaluación entre las asignaturas:

"La diferencia estriba en que las materias con más altos índices de reprobación (Matemáticas, Física y Química), el peso de los exámenes es mayor en la asignación de las calificaciones; en las materias de índices bajos de reprobación (Civismo, Historia y Geografía), el cumplimiento de las tareas y la participación en clase tendrían un peso mayor".

También Quiroz (1994), señala que se dan diferentes índices de reprobación en los grados, por ejemplo en la materia de Civismo 10% en el primero, 9% en segundo y 5% en tercero. Es decir, la tasa de reprobación es menor en el tercer grado. Esto debido a que en los dos primeros grados hay posibilidad de que el alumno apruebe en el siguiente año escolar; mientras que en el tercer grado puede perder un año escolar, por lo que se presiona al maestro por parte del alumno, de los padres de familia y en algunos casos hasta por parte de los directivos.

En opinión de Quiroz (1994), existen tres elementos que combinados tienen el mayor peso de influencia sobre la reprobación en este nivel:

a) Características del currículum formal

- La extensión en cuanto al tiempo, no se concluyen los programas escolares.
- Los contenidos especializados y alumnos que aún no han definido su especialidad.
- La fragmentación doce maestros, doce materias lo que complica la adaptación a los estudiantes, ya que se les deja muy poco tiempo para un trabajo más profundo de los contenidos.

b) Las condiciones materiales del trabajo escolar: los bajos salarios de los maestros, su contratación desventajosa (la mayoría por horas clase), la fragmentación de su tiempo entre una gran cantidad de grupos y alumnos y el número de alumnos por grupo.

c) Las tradiciones académicas del magisterio:

La sobrevaloración del saber de su especialidad, trae como consecuencia el excesivo trabajo para los estudiantes.

La importancia atribuida a la evaluación formal que trae como consecuencia la aplicación de exámenes frecuentes ( de 6 a 7 exámenes por semana).

Otro aspecto importante en la calidad educativa es la deserción. Quiroz ( 1994), menciona que este factor en la escuela secundaria es atribuible tanto a situaciones externas como internas de la dinámica académica.

Todo parece indicar que el aprovechamiento que tienen los alumnos de este ciclo, evaluado con pruebas de rendimiento, es muy deficitario. A fin de profundizar un poco más en los factores que están influyendo en estos resultados, a continuación se analizan algunos aspectos que la OCDE ( 1991 ), ha señalado como muy significativos en relación a la calidad educativa..

### **1.4.3. ARTICULACION ENTRE NIVELES**

La OCDE ( 1991 ), señala que. es conveniente alternar. los ciclos escolares por la necesidad general de una continuidad entre un ciclo de aprendizaje y el siguiente.

La transición del ciclo primario al cm secundario representa un problema singular para los alumnos, sobre todo para aquellos que presentan dificultades de aprendizaje. Por lo tanto, se plantean sus recomendaciones como son: promover un ambiente de seguridad y estimulante, la existencia de una comunicación entre los profesores del ciclo primario y secundario, así como la existencia y la entrega de un expediente de cada alumno y la existencia de un curso introductorio en el que se señale los beneficios y procedimientos de la ES.

Actualmente en nuestro país, la duración del ciclo primario es de 6 años, mientras que en el ciclo secundario es de 3 años. Existiendo una marcada diferencia entre un ciclo y el siguiente: de tener un solo maestro pasan a contar con un maestro por materia, existe la figura del "prefecto" encargado de la disciplina escolar. En cuanto a la continuidad y seguimiento que se debería tener la información que se conserva de la historia académica o son sus calificaciones. Por lo tanto, se entiende que no existe comunicación entre los profesores de cada ciclo, ni un expediente que ofrezca información de los antecedentes escolares de cada alumno. En algunas secundarias si existe el curso introductorio. Lo anterior solamente se refiere a las escuelas de la zona urbana.

#### **1.4.4. CONDICIONES EN CUANTO A LAS PROPORCIONES PROFESOR-ALUMNO**

El tamaño de una escuela es un factor significativo en la determinación del entorno de enseñanza y aprendizaje. Los estudios realizados muestran que las escuelas pequeñas poseen la ventaja de crear una atmósfera cordial y una sensación de comunidad, mientras que las escuelas grandes brindan una más amplia elección de materias y oportunidades para actividades al margen del currículo.

El clima escolar es una cuestión de la mayor significancia pedagógica. Los estudios han revelado que la dedicación del personal docente y su cooperación activa en el seno de las escuelas constituyen factores relevantes para distinguir las escuelas buenas de las menos eficaces (OCDE, 1991 )

Otro aspecto importante es el hecho de que las escuelas deberían precisar claramente sus normas y objetivos, los cuales deben ser aceptados por los profesores, los alumnos y sus padres.

El número de alumnos es otro aspecto crucial en los procesos de aprendizaje: cuanto menos numeroso sea el grupo. mayor será el logro de cada alumno.

Por su parte, Sandoval (1993) y Quiroz (1994}, mencionan que las condiciones del ciclo de secundaria son las siguientes: instalaciones deficientes, sobrecupo de alumnos, bajos salarios a los maestros (por lo que se tienen dos trabajos o más), contratación desventajosa (en la mayoría por horas clase), la fragmentación de su tiempo entre una cantidad de alumnos y grupos (un maestro puede atender hasta 15 grupos y 750 alumnos), identificación de sus objetivos. tendencia al enciclopedismo en el contenido de las materias del mapa curricular, contenidos programáticos abundantes y sin relación entre las materias, elevados índices de pasantes de licenciatura sin formación pedagógica, falta de apoyos técnicos, etc.

Otro aspecto que no se tomó en cuenta es el tiempo que el profesor invierte en preparar sus clases y en el que se traslada de un lado a otro.

Como ya se ha señalado, las acciones que se propone el presente gobierno para mejorar la calidad en el SEN son: ampliar la cobertura nacional y dar impulso a la formación y actualización docente. Otro aspecto que marcan como prioritario es iniciar una política de evaluación en todo el SEN. Con este propósito se ha creado una instancia encargada de la evaluación de la educación básica en el país: el Centro Nacional de Evaluación. En el siguiente capítulo se desarrollará este polémico tema.

Todos estos señalamientos que se hacen acerca de la calidad de la educación están en estrecha relación con el tema del currículo escolar. ya que el modo en que este se ha definido, planificado y evaluado influye decisivamente en la calidad de la educación.

## **1.5. EVALUACION EDUCATIVA (EE::)**

En este apartado se realizará una breve revisión de lo que es la Evaluación Educativa (EE}, empezando con la problemática que existe en la EE, los modelos existentes en este ámbito, la Investigación evaluativa, sus funciones, sus objetos de estudio y sus técnicas.

### **1.5.1. LA PROBLEMÁTICA DE LA EVALUACIÓN**

Aunque se mencione que la EE debe estar presente durante todo el proceso educativo, se sabe que se toma como objeto fundamental la evaluación de los alumnos. Este es uno de los principales problemas que se presentan en el ámbito de la EE..

Por lo tanto, a continuación se abordan los Modelos Evaluativos, la Investigación Evaluativa, funciones y objetos de la EE, para conocer que debe abarcar ésta en el ámbito educativo.

### **1.5.2. MODELOS EVALUATIVOS**

En este apartado se reseña el desarrollo de algunos de los modelos de la EE. presentando la propuesta de sus principales representantes .

**A) Ralph W. Tyler:** el iniciador de la evaluación educacional.

A este autor se le considera el iniciador de la evaluación educacional, ya que propuso, describió y aplicó un método ya desarrollado para la evaluación, algo que nadie había hecho antes; y su metodología ha sido tan penetrante como influyente.

A su método se le denomina "Modelo de Evaluación por Objetivos", el cual consistía en la comparación de los resultados del aprendizaje de los alumnos con los objetivos previamente determinados; y tenía como intención, si se alcanzaban los objetivos, se tomarían las decisiones en una dirección concreta, sino se alcanzaban, o sólo se alcanzaban parcialmente, debían tomarse decisiones distintas.

Además Tyler, proporcionó medios prácticos para la retroalimentación de las prácticas educativas. La evaluación debería proporcionar un programa útil que permitiera la reformulación o redefinición de los objetivos.

En este sentido menciona:

"El proceso de evaluación significa fundamentalmente, determinar en que 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).

Con 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, 1973, p.126).

También destacó la necesidad de establecer, clasificar y definir los objetivos en términos de rendimiento, como una etapa inicial del estudio evaluativo.

**B) Lee. J. Cronbach:** evaluación para la toma de decisiones.

Comienza sus trabajos sobre la EE, reflexionando acerca de las formas y propósitos de la evaluación, argumentaba que después de 1930 las pruebas se aplicaron casi exclusivamente para evaluar individuos, al respecto comentaba:

"Yo sostengo que en las evaluaciones de los cursos no se requieren calificaciones individuales precisas" (Cronbach, 1982, pp. 36).



Según Cronbach (Citado por Rosales, C. 1990), la metodología de la evaluación ha de procurar un equilibrio entre los procedimientos experimentales y naturales, en función de su mejor aplicación a cada situación escolar.

Su modelo de evaluación tenía como objetivo principal la recolección y uso de la información para tomar decisiones acerca de un programa educativo.

Por lo tanto, él considera que la evaluación consiste en la búsqueda de información y en su comunicación a quienes han de tomar decisiones sobre la enseñanza.

Una de las importantes preocupaciones para Cronbach lo constituye la comunicación de los datos de la evaluación. La cual se manifiesta en características como las siguientes (citado por Rosales, C., 1990):

- Claridad, comprensible a quienes se destina;
- Oportunidad, disponible en el momento en que se necesita;
- Exactitud, que distintos observadores han de percibir la misma realidad de igual manera;
- Validez, los contenidos de la evaluación se deben corresponder con la realidad;
- Amplitud, debe proporcionar posibilidades diferentes para la adopción de diversas alternativas.

Cronbach sugiere que para la toma de decisiones se tome en cuenta a los miembros de la comunidad, sólo a nivel de consulta, la decisión preliminar se le atribuye a quienes contratan la evaluación.

### **C) Michel Scriven:** Evaluación sin objetivos.

Para M. Scriven (citado por Rosales, C., 1990). 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.

Propone que la evaluación se realice sin referencia a objetivos, que el evaluador no tome en consideración, e incluso desconozca deliberadamente los objetivos que se hayan propuesto para la enseñanza.

La intención de este es juzgar el valor de algo, basándose en las necesidades de los usuarios, este es un punto debatible de su modelo.

Scriven fue el que introdujo la diferencia entre evaluación formativa y sumativa:

La primera (citado por Rosales, C., 1990), 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.

La segunda, se centra en el estudio de los resultados, calculando el valor del objeto una vez que ha sido desarrollado. Por medio de este estudio se puede ayudar a los responsables de la toma de decisiones a sacar conclusiones en relación a un programa o producto ya concluido.

Cuestionaba que Cronbach no tomará en cuenta la evaluación sumativa, y por lo tanto la evaluación resultaba parcial, ya que no tomaba en cuenta el objeto ya concluido.

También distingue entre evaluación intrínseca (al propio proceso) y evaluación final (los resultados del proceso ):

- La primera evalúa las cualidades de una medición, sin tener en cuenta los efectos sobre los usuarios (metas, estructura, metodología, calificación y actitudes del personal, recursos).
- La segunda no se ocupa de la naturaleza del programa, de los resultados, etc. sino más bien de los efectos del objeto evaluado sobre los usuarios. En esta evaluación se incluyen las Pruebas de Rendimiento, las cuales son el punto principal de esta investigación.

Otro aspecto importante que propone Scriven es la metaevaluación (evaluación de la evaluación) como una forma de control para los propios evaluadores, ya que considera que estos últimos tienen la obligación profesional de que las evaluaciones terminadas estén sujetas a una evaluación competente.

Por todo lo anterior, este autor es considerado como uno de los que mayores contribuciones ha tenido en el desarrollo del campo de la EE. Sin embargo, su modelo ha sido poco utilizado.

**D) Daniel Stufflebeam:** Evaluar para perfeccionar.

Para Stufflebeam la evaluación es:

"el proceso de identificar, obtener y proporcionar información útil y descriptiva acerca del valor y el mérito de las metas, la planificación, la realización y el impacto de un objeto determinado, con el fin de servir de guía para la toma de decisiones, solucionar los problemas de responsabilidad y promover la comprensión de los fenómenos implicados" (Stufflebeam, D. y Shinkfield, A., 1987).

Esta definición postula tres propósitos evaluativos:

- servir de guía para la toma de decisiones,
- proporcionar datos para la responsabilidad,
- promover la comprensión de los fenómenos implicados.

Por otra parte, el autor menciona que el propósito más importante de la evaluación no es demostrar sino perfeccionar. También hace la recomendación de no identificar la evaluación con la valoración, el juicio personal, la investigación experimental o la coincidencia entre objetivos y resultados.

Stufflebeam empezó a buscar un método más relevante y factible que los ya existentes, como resultado de esta búsqueda, fue el llamado Modelo de Evaluación "CIPP" (contexto, entrada-input-, proceso y producto).

Para perfeccionar la base conceptual del trabajo evaluativo se especificaron cinco problemas que necesitaban ser solucionados:

1. Definición. Debe juzgarse si las definiciones son vagas, erróneas o imperfectas en su orientación metodológica.

2. Toma de decisiones. Se pensaba que las decisiones que debían tomarse en cuenta a los cambios no estaban muy claras.

3. Valores. Se creía que muchas evaluaciones reflejaban erróneamente una orientación que no atendía a los valores, denunciando la escasa literatura acerca de como conocer y tratar con múltiples perspectivas de valor, como conseguir los criterios e instrumentos apropiados y cómo asignar valores a la información obtenida.

4. Niveles. Los evaluadores no reconocían las distintas audiencias. Por lo tanto, se necesita una planificación de evaluaciones que sirviera de guía a los evaluadores para identificar las audiencias, valorar sus necesidades informativas y planificar la recopilación de datos y las actividades informativas.

5. Diseño de la investigación. Se consideraba que tanto la investigación como la evaluación tenían propósitos distintos y por lo tanto requerían diferentes procedimientos.

Dentro del modelo CIPP, propuesto por Stufflebeam, se proponen cuatro tipos de evaluación:

1. Evaluación del contexto: identificar el objeto de estudio (institución, población o persona) y valorar sus necesidades, reconocer que oportunidades satisfacen estas necesidades y si los objetivos propuestos son coherentes con las necesidades.

2. Evaluación de entrada: su orientación es ayudar a valorar un programa procediendo a identificar y examinar críticamente los métodos utilizados, el propósito de este tipo de evaluación es ayudar a los clientes en la consideración de estrategias alternativas.

3. Evaluación del proceso: es la comprobación continua de la realización de un plan que permite contar con información acerca de hasta que punto las actividades del programa siguen las líneas que se habían planificado .

4. Evaluación del producto: recopila las descripciones y juicios acerca de los resultados, relacionándolos con los objetivos y la información proporcionada por las anteriores evaluaciones para interpretar su valor y su mérito.

### **1.5.3. INVESTIGACION EVALUATIVA**

Aunque Stufflebeam recomendaba no identificar la evaluación con la investigación, Rosales (1990), menciona que:

“La evaluación se identifica con la investigación y utiliza los mismos métodos que ella en su tarea de recogida de datos”.

Por lo anterior , surge la interrogante ¿Qué tipo de investigación evaluativa se utilizará? . Primeramente se tendría que tomar conciencia de que la enseñanza es extensa, por lo cual su metodología es variable. Según Bazinet (1984), un método de evaluación del rendimiento es esencialmente un instrumento de medición que funciona comparando las cualidades de la persona que ejecuta sus resultados, con determinados criterios.

Por otro lado. Bazinet (1984) señala que la diversidad de las técnicas existentes no se debe sólo al progreso de la ciencia, sino también a que se concibieron para resolver distintas necesidades.

Hacia los años sesenta y parte de los setenta las prácticas evaluativas retornaron el procedimiento que se privilegiaba en las Ciencias Sociales (CS): el método experimental. Este último obtiene la información con base a unos cuantos indicadores, privilegiando las pruebas objetivas.

Otros evaluadores han cuestionado el uso de este modelo para la EE, ya que la validez es totalmente ausente en este procedimiento.

Dentro de las críticas que se le hacen al modelo experimental destacan las siguientes:

- 1) Requiere que se mantenga constante el programa, en vez de perfeccionarlo continuamente.
- 2) Es útil para tomar decisiones sólo después de que un proyecto ha recorrido un ciclo y no durante su planeación y procedimiento.
- 3) Trata de controlar demasiadas condiciones, lo que ocasiona que el programa no sea general y no se aplique en distintas regiones.

A consecuencia de lo anterior se propone una evaluación holística (abarca todo) que tome en consideración todos los componentes de la enseñanza: proceso, resultados y contexto. Según Gimeno y Pérez (1991), la evolución reciente "del sistema educativo y del pensamiento sobre el currículum lleva a una ampliación constante de sus contenidos y en consecuencia a considerar más objetivos y cualidades susceptibles de ser evaluados.

Los enfoques que dominaron en los años sesentas y setentas ayudaron a mantener la idea de la evaluación holística, ya que defendían la necesidad de especificar los objetivos en campos diversos y en múltiples categorías, estos enfoques fueron las taxonomías de BLOOM (1975), GAGNE (1979), y GAGNE, BRIGGS y WAGER (1988) y surgieron para ser utilizadas como procedimientos que ayudaran a concretar instrumentos de evaluación más precisos.

La taxonomía del primer autor se citará en el capítulo tres.

Por lo tanto, actualmente se considera que se debe tomar en cuenta la metodología cuantitativa como la cualitativa para realizar una investigación.

Por su parte Gimeno (1989), menciona que la polémica de los métodos cualitativos frente a los cuantitativos, con los sistemas de calificación de alumnos que les son propios, es otro ejemplo de cómo, en un sistema de evaluación, ha de juzgarse desde la perspectiva del tipo de información que selecciona como potencialmente relevante para el evaluador y en coherencia con qué tipo de rendimiento ideal se emplean de hecho esas técnicas.

Por consiguiente, es importante especificar los métodos cualitativos y los cuantitativos. Los primeros hacen referencia al abordar el objetivo de estudio desde dentro, tratando de conocer las razones de su evolución y la finalidad de sus cambios. Para ello adopta estrategias variadas, utilizando diferentes perspectivas, y se dota de la suficiente flexibilidad para modificar las técnicas que se están utilizando si la situación así lo requiere. La metodología concreta que se utiliza más a menudo es la observación y la entrevista. (Marchesi, A., Echeita, G., Martín, E., 1993). Este método no se reduce a informar sólo procedimientos numéricos o de categorías verbales, ya que tiene más capacidad de información para quienes lo reciben.

Mientras que los métodos cuantitativos, conceden valor de objetividad a la expresión numérica de las cualidades y estados de aprendizaje. Son sistemas cómodos para el profesor. Es un modelo sencillo de realizar la evaluación para el profesor, que se ajusta, en

cuanto a comodidad de manejo, a la economía de procesos psicológicos que ocurren cuando ese profesor tiene que realizar la operación psicológica de emitir opiniones formalizadas sobre sus alumnos (Gimeno, 1989).

Al hacer referencia a los métodos cuantitativos y cualitativos, se denotan algunas deficiencias en estos, se sugiere la complementariedad de ambos. En los últimos tiempos se habla de métodos múltiples en la investigación evaluativa, éstos métodos comprenden diversas alternativas:

- a) La conducción de análisis de tipo cuasi-experimental.
- b) La combinación de métodos cualitativos y cuantitativos.
- c) La combinación de metas múltiples de evaluación como procesos y productos.

En esta investigación, tomando en cuenta lo anterior, se puede decir que la utilización de métodos múltiples facilita y enriquece el proceso de recogida de información y por lo tanto, enriquece la EE.

#### **1.5.4. Funciones de la EE**

Hacia 1967 Scriven introdujo la distinción entre evaluación formativa y evaluación sumativa, siendo estas las dos funciones principales de la EE.

A continuación se conocerán las características de cada una de ellas.

##### **A) Evaluación Sumativa**

A partir de los años sesentas se apoyó mucho a la educación, por lo cual se incrementaron los recursos destinados a ésta, los sectores de usuarios reclamaron el justificar el incremento de estos recursos, a consecuencia de esto, aparece una política evaluativa de rendición de cuentas (accountability).



Según Gimeno (1989), al ser presionado el sistema escolar por la sociedad y la administración, se reclama un proceso de sobre simplificación, y de esta manera aparece la llamada evaluación sumativa. Esta se refiere al juicio final global de un proceso terminado, emitiendo una valoración global. Su enfoque es retrospectivo, sanciona lo que ha ocurrido, mirando desde el final de un proceso.

Retornando a Stufflebearn y su modelo. se puede decir. que este tipo de evaluación (sumativa) es semejante a la evaluación del producto, ya que las dos evalúan los resultados para emitir un juicio de valor .

Por otra parte, Marchesi, A., Echeita. G.. Martín. E. (1993), mencionan que las personas que se dedican a la Evaluación Sumativa deben ser observadores externos que valoren la eficiencia de los resultados de acuerdo con los criterios previamente establecidos.

La evaluación sumativa tiene como finalidad seleccionar y jerarquizar a los alumnos por sus resultados obtenidos.

Según Gimeno y Pérez ( 1991 ), la evaluación sumativa también puede servir para decidir si se continúa o no un proceso, pero no se puede incidir sobre el proceso concluido. porque la información que ofrece se refiere, como mucho, a la calidad del resultado. No puede contestar a la pregunta de cómo ésta aprendiendo el alumno, sino saber lo que ya ha aprendido.

Este tipo de evaluación suele expresarse mediante un método cuantitativo. es decir, en una nota o en una escala numérica o en una categoría verbal. Suele presentarse en períodos largos de tiempo, realizándose en diversos tipos de pruebas. Los resultados que arrojan estas pruebas pueden quedarse en la relación profesor-alumno, pero después suelen proyectarse al exterior, por medio de informes, o en las boletas de calificaciones. Si se quisiera realizar esta evaluación de forma continua se entraría en una dinámica de aplicación constante de pruebas y calificación de alumnos.

Tanto el sistema escolar como la presión social reclaman el tipo de evaluación sumativa final de un proceso, básicamente. Esto puede justificarse debido a que el modelo tyleriano sigue vigente y ésta evaluación toma en cuenta su principio de comparar los resultados con los objetivos.

## **B) Evaluación Formativa**

La evaluación formativa tiene como propósito favorecer al proceso educativo. Esta evaluación tiene sentido realizarla mientras el proceso transcurre. Su enfoque es proactivo. se proyecta hacia lo que queda por realizar. Se tiene que realizar de esta forma, para detectar las insuficiencias e intentar corregirlas en el momento.

A diferencia de la evaluación sumativa ésta si responde a la pregunta de cómo están aprendiendo y progresando los alumnos.

Esta evaluación se realiza de forma constante. puede realizarse de manera informal por los profesores.

Según Gimeno y Pérez (1991 ), señalan:

“ El propósito mismo de la evaluación formativa aplicada a contextos naturales de aula nos lleva a asociarla con una evaluación de carácter continuo, realizada a través de procedimientos informales o no muy elaborados, dependiendo más de las capacidades diagnósticas de los profesores que de pruebas o instrumentos técnicos complicados. No porque no los admita, sino porque la dinámica normal de una clase hace inviable una aplicación constante de instrumentos o pruebas formales”.

Por lo anterior, este tipo de evaluación es similar al modelo de Stufflebeam. ya que en su modelo la evaluación tenía el propósito de perfeccionar el ámbito escolar. en particular es semejante a la evaluación del proceso, porque es en la que aparece el carácter continuo de la evaluación.

Por otra parte. Gimeno (1989), comenta que la evaluación formativa no se ha integrado en el sistema pedagógico por falta de asimilación de los profesores y de formación en este tipo de evaluación. En otra publicación más reciente Gimeno y Pérez (1991) comentan que:

"La evaluación formativa con la intención de diagnosticar procesos es una aspiración tan exigente que requiere una relación de tipo tutorial entre el profesor y cada alumno en particular. Pero éste (el tiempo en que se aprende) no es accesible a los profesores sino cuando siguen el proceso.

Por lo tanto, es recomendable analizar las condiciones de trabajo así como la formación que se posee, para realizar cualquier tipo de evaluación.

Por tal motivo, las condiciones para realizar esta investigación son de forma externa, ya que no es posible estar presente durante todo el proceso de enseñanza-aprendizaje del Civismo, por lo tanto, se hará uso de la evaluación sumativa.

#### **1.5.5. LOS OBJETOS DE LA EE**

Por lo general, tanto los profesores como los alumnos han sido el objeto principal de la EE. Pero actualmente existen evaluaciones de los programas de estudio, de los centros escolares, de las instituciones de las que dependen la educación, de los materiales didácticos, etc.

A continuación se especificará cada una de ellas.

\* Los alumnos. Se considera que la evaluación entorno a este objeto ha sido excesiva casi siempre las actividades de aprendizaje se han sometido a pruebas para la obtención de

calificaciones en torno a las que gira la actitud que el alumno tomará ante las personas que lo rodean según sea el caso positivo o negativo.

Por otra parte, Rosales (1990), comenta:

"En el caso de los alumnos las investigaciones sobre el aprendizaje escolar y el desarrollo madurativo ponen de relieve la insuficiencia de los resultados de carácter intelectual y subrayan la necesidad de evaluación del ámbito actitudinal, de integración socioafectiva. Se manifiesta asimismo el interés de la evaluación del ámbito psicomotor de la conducta. Dentro del mismo terreno intelectual se trata de proyectar la evaluación más sobre las habilidades, las técnicas, las estrategias de conocimiento, que sobre el dominio de la información"

Es decir, Rosales sugiere que la evaluación no nada más quede en el nivel intelectual, sino se expanda a los niveles psicoafectivo y psicomotor, así como sobre habilidades, las técnicas, las estrategias de conocimiento.

Esta evaluación también es práctica y útil para el profesor.

. \*Profesores. En el caso de los profesores según Rosales (1990):

" Tradicionalmente centrado en las características personales bajo el supuesto de que la manera de ser repercute en la eficacia profesional, a partir de los años sesentas y en función del desarrollo de la Psicología conductista y del proceso de tecnificación general en los ámbitos social y laboral, la evaluación del profesor pasa a centrarse en su conducta observable, en sus formas de actuación ante los alumnos".

Es decir, si el profesor se preocupa por el proceso de aprendizaje de sus alumnos, estaría ejerciendo una evaluación de tipo formativa. También es importante tener presente como el profesor lleva a cabo el programa de estudio.

“Programa de estudio. Después de ser objetos principales de evaluación el profesor y los alumnos, ocupa un lugar primordial a la evaluación de los programas de estudio, la cual se desenvuelve en tres planos de estudio: a) planificación de programa (en esta se evalúan las condiciones y necesidades que presenta la institución y las personas a quienes va dirigido el programa), b) proceso del programa (evalúa la eficacia del desarrollo de las actividades establecidas), y c) valoración de resultados o alcance del programa (evalúa los resultados comparándolos con los objetivos planteados).

Si bien esto se evalúa conforme a las condiciones que el centro escolar o la institución ofrezca para realizar el programa.

“Centros escolares. Como ya se mencionó en el primer capítulo, el clima o el ambiente que se desenvuelve en una institución escolar es importante para la calidad de la enseñanza-aprendizaje. Por tal motivo, la evaluación ya no es sólo de personas sino también de microinstituciones donde se realiza el aprendizaje.

Esta modalidad tiene su aparición en la década de los ochentas.

Al evaluar el centro escolar es importante evaluar los recursos didácticos con que este desarrolla las actividades escolares.

\*Recursos didácticos. En cuanto a este objeto de evaluación Rosales (1990), menciona:

"Su evaluación ha progresado desde el interés por el efecto global del medio sobre el aprendizaje del alumno, al estudio específico de la influencia de determinados atributos como el lenguaje sobre el desarrollo de habilidades y aprendizajes específicos en el alumno, todo ello dentro de una amplia serie de relaciones de interacción con factores como la situación y contenidos con que se utiliza, la preparación del profesor y las características psicopedagógicas de los alumnos".

Es decir, en este apartado no sólo importan las características de los recursos didácticos sino lo que verdaderamente se evalúa es la preparación que el profesor tiene para lograr un buen uso de estos y la influencia que todo esto tenga en el alumno para desarrollar sus habilidades y aprendizajes específicos.

Chávez y Zamora ( 1989), mencionan que los recursos didácticos se presentan en distintas formas y medios, entre ellos destacan: los libros de texto, libros de trabajo, material auto-instrucción, lecturas complementarias, programas de TV educativa, diapositivas, filminas, cintas grabadas y equipos.

También sugieren que cada proyecto curricular debe determinar que materiales utilizar .

Estos programas curriculares a su vez están determinados por una institución que es macrosocial.

\*Instituciones. En este caso la evaluación corresponde a las instituciones responsables de normar y coordinar la enseñanza, tomando en cuenta las situaciones histórica, económica y sociopolítica en que se desenvuelven estas instituciones y también se valoran los resultados que se obtengan a través del impacto social y el costo de sus programas.

Sin dejar de tomar en cuenta a otro de los objetos de la evaluación el cual es de suma importancia se hará referencia a la interacción.

\*Interacción. La evaluación no puede limitar únicamente a cada uno de los objetos antes mencionados sino que es necesaria la consideración de la interacción de todos los objetos. Ya que todos estos se relacionan en el proceso de aprendizaje.

Es decir, la evaluación no sólo se refiere a personas, ya que, actualmente cualquier cosa presente en el ámbito educativo puede ser un objeto de evaluación. Sin embargo, es importante considerar que cualquier evaluación debe tener una planificación previa que tome en cuenta las necesidades del objeto de evaluación para determinar la orientación teórica que deberá tomar .

### **1.5.6. TECNICAS DE EVALUACIÓN**

Como ya se mencionó anteriormente en la EE, en la actualidad predominan los métodos múltiples, ya que se ha señalado que estos son más convenientes, para no dejar de lado algún aspecto importante de una evaluación determinada. Los métodos múltiples permiten elegir una variedad de técnicas adecuadas al objeto de estudio. Estas técnicas son útiles para la recogida de datos, éstas técnicas se encuentran clasificadas en dos rubros:

\*Técnicas cualitativas: entrevista, observación no participante, diario del investigador, análisis de documentos y pruebas no estandarizadas.

\*Técnicas cuantitativas: listas de control, escalas, cuestionarios y pruebas estandarizadas.

A continuación se especificará cada una de ellas.

#### **A) Técnicas Cualitativas**

Entre estas técnicas la metodología que se utiliza más a menudo es la observación y la entrevista, se hablará más adelante de estas y de otras .

\*Observación. Según Casanova (1992), se entiende por observación el examen atento que un sujeto realiza sobre otro u otros sujetos o sobre determinados objetos y hechos, para llegar a un conocimiento mayor sobre ellos mediante la obtención de una serie de datos, generalmente inalcanzables por otros medios.

Actualmente se distinguen dos tipos de observación: la participante, en la que el observador se integra en mayor o menor medida en el grupo al que debe observar; y la no participante en la que el observador se mantiene al margen de las actividades y relaciones del grupo y no se integra en él.

\*Entrevista. Según Casanova (1992), puede definirse como una conversación intencional. Dada su similitud con el cuestionario, también se conceptúa como un cuestionario, más o

menos estructurado o abierto, planteado y respondido de forma oral, en situación de comunicación personal directa.

Es una técnica que suele utilizarse con dos objetivos principales: la investigación y la orientación. Por ello las funciones que cumple son de carácter diagnóstico, evaluador, investigador, orientador y terapéutico.

La entrevista se clasifica de la siguiente manera:

a) Entrevista formal (previamente preparada por el entrevistador). A su vez, esta puede presentar distintos grados de estructuración:

-Estructurada, seguimiento de un cuestionario preparado con antelación, que servirá de guión.

- -Semiestructurada, seguimiento del cuestionario establecido, pero modificando las preguntas necesarias en función de las respuestas.
- Libre o abierta, entrevista preparada de antemano, pero en la que las preguntas no están predeterminadas; la conversación hará aparecer aquellas que se consideren más oportunas.

b) Entrevista informal, sin preparación previa, en la que el entrevistador, recoge la información relevante para su estudio, en situaciones de encuentro espontáneo.

\*Diario del investigador. El evaluador registra notas de lo que considera más relevante para el seguimiento del estudio, tanto de la situación como de los sentimientos, pensamientos y actitudes que le surgen en ese momento.

\*Análisis de documentos, también denominado análisis de contenido. Según Holstin (citado por Chávez y Zamora, 1989), el análisis de contenido se define como el método que permite, indicar juicios de valor, con base en la identificación objetiva y sistemática de las características específicas del plan de estudios.



Dicho análisis comprende las siguientes fases:

- a) Determinación de amplitud y análisis (muestra o población completa).
- b) Concreción de la unidad de análisis (palabra, oración, frase, texto, tema...).
- c) Identificación, selección y estructuración de las categorías de análisis.
- d) Codificación de los datos, siempre y cuando se trate de numéricos se tabulan y cuantifican.
- e) Análisis estadístico, en su caso.

\*Pruebas no estandarizadas. Son utilizadas para la valoración de los resultados de los alumnos, pretenden analizar los errores más que los aciertos, en cuanto al desempeño de los alumnos. Es una actividad esencialmente comprensiva que busca corregir las fallas en el proceso de razonamiento de los estudiantes.

## **B) Técnicas Cuantitativas**

Este tipo de técnicas buscan realizar una recogida de datos que permitan abarcar grandes poblaciones.

\*Listas de control. Este instrumento consiste en un registro de conductas observables previamente establecidas en la que se marcan las conductas u objetos cada vez que ocurren en una unidad de tiempo determinado. Siendo una tabla dicotómica en la que la conducta observable puede estar presente o ausente es una tabla de doble entrada.

\*Escala. Ary y Cheser (1990), las definen como un conjunto de valores numéricos asignados a sujetos, comportamientos con el propósito de cuantificar y medir sus cualidades. Sirven para medir actitudes e intereses.

Las escalas pueden clasificarse de la siguiente manera:

- 1) Escalas de actitudes.
- 2) Escalas numéricas.
- 3) Escalas gráficas.
- 4) Escalas descriptivas.

\*Cuestionarios. Son un conjunto de preguntas y respuestas anteriormente planteadas que se aplican por escrito a un determinado número de sujetos. Con frecuencia se aplica en muestras amplias.

Según Casanova ( 1992), existen tres tipos de cuestionarios:

a) los cuestionarios cerrados, plantean contestarse con "sí" o "no", o señalando un apartado determinado con el signo que se indique en las instrucciones.

b) los cuestionarios abiertos, plantean las preguntas, pero no sugieren alternativas de respuestas, sino que simplemente dejan un espacio para que el sujeto exprese todo lo que crea oportuno.

c) los cuestionarios mixtos, están compuestos por preguntas de ambos tipos, de manera que la mayor parte resulte cuantificable y una parte pequeña, es tratada e interpretada cuestionario por cuestionario.

\*Pruebas estandarizadas. Consisten en un conjunto de preguntas destinadas a producir cierto tipo de conductas que al ser evaluadas bajo condiciones estandarizadas producen puntuaciones con propiedades psicométricas. Según González, (1994): "la mayoría de las pruebas estandarizadas son elaboradas por equipos de investigación en universidades o por editores profesionales quienes cuentan con expertos en medición y en planes de estudio. En países que no han logrado un desarrollo tecnológico en el campo, se utilizan traducciones y-o adaptaciones de los instrumentos".

Por otro lado, Casanova ( 1992), recomienda que las pruebas estandarizadas utilizadas en diseños cuantitativos se empleen como complemento de un modelo cualitativo de evaluación.

No sólo en las pruebas estandarizadas se pretende complementarlas con técnicas cualitativas sino en todas las técnicas cuantitativas, por eso el objetivo de esta investigación es hacer una combinación del cuestionario ( cuantitativo) y pruebas no estandarizadas ( cualitativas ).

## **CAPITULO II. PRUEBAS DE RENDIMIENTO**

El presente capítulo iniciará con una pequeña definición del término "prueba", para pasar a una reseña de los antecedentes de las Pruebas Psicológicas en general, posteriormente se atenderán las, Pruebas de Rendimiento (PR), en esta última parte se abordará su desarrollo general y su desarrollo en México.

El término " prueba" se deriva del latín testis, que significa en inglés test, siendo empleada sin traducción en diversos países del mundo (Forns, 1980).

Resumiendo lo que muchos autores dicen sobre el significado del término "prueba" se puede decir que es un instrumento que contiene un conjunto de tareas o preguntas que permiten observar cierto tipo de conductos y medirlas en forma objetiva, para determinar ya sea el desarrollo mental de un individuo, el nivel escolar, la existencia y el grado de una aptitud dada.

Por tal motivo, existen diferentes fines de una prueba:

- **Motivación:** es un elemento primordial tanto para el maestro como para el alumno, principalmente para este último para conocer los resultados de sus pruebas lo más pronto posible, ya que así tiene la oportunidad de analizar su rendimiento y, con

base en el mismo, adoptar la conducta que considere pertinente para mejorarlo.

- **Diagnóstico:** tiene por objeto localizar dificultades específicas del aprendizaje de algún alumno en particular o del grupo en su conjunto. Una vez detectadas las fallas o limitaciones, se puede trazar un plan para solucionar esos problemas de rendimiento.
- **Pronóstico:** es conveniente que el profesor, antes de empezar la enseñanza, cuente con información que le permita prever como se desarrollará el procedimiento de aprendizaje.
- **Promoción:** tiene por objeto controlar el ingreso al grado inmediato superior a otro nivel de enseñanza, con apoyo en los datos aportados por ellas. Las pruebas aplicadas con este fin de promoción percibir si el estudiante tiene conocimientos para continuar con estudios más avanzados o si es necesario repetir el ciclo.
- **Agrupamiento:** responde a la necesidad de organizar los grupos de acuerdo con los conocimientos de los alumnos y poder homogeneizar, así, la futura labor docente.
- **Selección:** son utilizadas cuando., por diversos motivos, no poseen plazas suficientes para toda la población escolar que demande servicios, o bien, cuando se necesita medir los conocimientos que posee el aspirante a ingresar en dicha institución.
- **Investigación:** se recomienda utilizarla en aquellos casos en los que el instructor incluye una o más variables en el proceso enseñanza-aprendizaje y desea conocer los efectos de estas, ( Hurtado de Mendoza, 1980).

Otro aspecto que es importante mencionar, es que una prueba es una técnica cuantitativa, por lo tanto no es posible realizar la evaluación del proceso escolar con sólo esta técnica.

## **2.1. ANTECEDENTES**

Las primeras formas de evaluación psicológica abordan rasgos de la personalidad entre estas se encuentran: el horóscopo y una prueba de selección para los altos puestos del gobierno (220 años a. C. a 206 d. C.).

A partir de principios de siglo, la psicología cuantitativa ha ido desenvolviéndose a lo largo de dos corrientes principales. Una de ellas, la que hace relación a las diferencias individuales, se originó de un problema de astronomía; la otra, partió de la filosofía (Keats, 1974).

Por otro lado, según Fernández, (1981), se reconocen cuatro corrientes que han influido en los instrumentos de evaluación psicológica: 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.

Por tal motivo, los instrumentos de evaluación psicológica han ido evolucionando: Francis Galton (1882-1991), Mc. Keen-Cattell (1861-1934) y Alfred Binet (1875-1911). Los instrumentos de estos autores contribuyeron a la Psicología Diferencial. Dichos autores se interesaban en las diferencias individuales y la teoría evolucionista: los instrumentos de los dos primeros autores eran "test de inteligencia", ya que tenían la preocupación por los procesos psicológicos, por medio de sus escalas a partir de la psicometría se da el desarrollo del psicodiagnóstico, para dar inicio a la Psicología Clínica.

Según González, ( 1994 ), en las dos primeras décadas del siglo XX, en varios países se expande el intento por encontrar medidas psicológicas semejantes a las mediciones de las ciencias fisiconaturales, donde es posible reconocer dos planteamientos. Por una parte, la Psicología Experimental busca las leyes o principios generales que regulan los fenómenos psicológicos y, por la otra, la Psicología Diferencial o Correlacional, ésta interesada en constatar las diferencias individuales.

Para 1912 W. Stern introdujo el concepto de cociente mental o coeficiente intelectual (C.I.), a finales de esta década Arnold Gesell idea una Escala de Madurez, con base en la observación de niños y la detección de las pautas de comportamiento claves en el desarrollo infantil.

Ya en la década de los 30's; se desarrollaron nuevos instrumentos con diversos objetivos entre ellos destacan: la prueba Gestáltica Visomotora de Bender; la escala Wechsler - Bellevue, la primera prueba estandarizada para adultos. Mientras que para la década de los 40's los instrumentos tenían una influencia psicoanalista entre los que destacan: el Rorschach, el Test de Apreciación Temática (T.A.T) y la prueba de frustración de Rosezweg.

Por otra parte, la obra de Piaget, tuvo también influencia en los instrumentos psicológicos, planteando un diagnóstico de pensamiento lógico.

Todas éstas pruebas se han clasificado en tres tipologías:

- Pruebas de Aptitudes o de Inteligencia, son instrumentos para evaluar las capacidades verbales o no verbales de un individuo (Ary y Cheser, 1990).
- Pruebas de Personalidad, son instrumentos que miden los intereses, actitudes, comportamientos, sentimientos, etc., de un individuo.
- Pruebas de Rendimiento o de Aprovechamiento son instrumentos que miden el dominio y la destreza en distintas áreas del conocimiento (Ary y Cheser, 1990). En este tipo de pruebas se basa la presente investigación. Por lo tanto, a continuación se presentaran algunos aspectos relevantes sobre las PR.

## **2.2. DESARROLLO DE LAS PRUEBAS DE RENDIMIENTO**

Las PR aparecieron a finales del siglo XIX, cuando Fisher aplicó escalas que pretendían medir habilidades en ortografía, caligrafía, cálculo y dibujo. la primera prueba objetiva de rendimiento en los E.U.A., fue elaborada en 1895 por Rice. Su principal aportación a lo que se denomina actualmente "pruebas estandarizadas de rendimiento", radica en el enfoque objetivo y científico de su método para evaluar los conocimientos de los alumnos (Mehrens y Lehman, 1982).

Pero debido a que se tomaban más en cuenta las pruebas de inteligencia, las PR dejaron de usarse; ya principios de los 70's, debido a que se cuestionaba la eficacia que las pruebas de inteligencia ofrecían al ámbito escolar, aparecen de nueva cuenta las PR.

Al volver a aparecer las PR, se correlacionaban con las pruebas de aptitudes, según Ahmann (1969):

"Este problema surge del hecho de que quienes preparan los tests suponen comúnmente que la mejor manera de determinar el potencial de alguien para una capacidad futura dada es examinando su rendimiento presente en ese campo o en campos afines".

Por lo tanto es importante determinar que evalúan las PR. Según Mehrens y Lehman (1982) mencionan:

"... Una prueba de rendimiento se utiliza para medir el nivel presente de conocimientos, habilidad o ejecución de un individuo; ..., las pruebas de rendimiento miden lo que, un alumno aprendió (o actividades previas de aprendizaje)..."

De tal forma, que los niveles de complejidad cognoscitiva que valoran estas pruebas son (Bloom, Hastings y Madaus, 1975):

- 1) Conocimiento, se refiere principalmente al recuerdo. Es la operación cognoscitiva más simple.
- 2) 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) Aplicación, implica el uso de abstracciones en situaciones particulares y concretas. las abstracciones pueden presentarse bajo la forma de ideas generales, reglas de procedimientos y métodos generalizados .

Por lo tanto, según Gronlund (1980), el grado en que las PR contribuyen a mejorar el aprendizaje y la enseñanza está determinado en gran parte por los principios en que se funda su elaboración; estos principios se enlistan a continuación:

1.- Las PR deben medir resultados de aprendizaje claramente definidos y en armonía con los objetivos educacionales.

2.- Las PR deben medir una muestra adecuada de los resultados del aprendizaje y. del contenido de la materia de estudio incluida en la enseñanza.

3.- Las PR deben comprender los tipos de ítems más adecuados para medir los resultados que se desean del aprendizaje.

4.- Las PR deben diseñarse según el uso particular que se dará a los resultados.

5.- Las PR deben hacerse lo más confiables posibles y se deben interpretar con sumo cuidado.

6.- Las PR deben utilizarse para mejorar el aprendizaje del estudiante.

Sin embargo, como señala Schmelkes (1995) la evaluación:

“... no debe de ser utilizada para determinar la suerte de los individuos. Debe ser utilizada para determinar la suerte de colectivos: escuelas, conjuntos de escuelas, grupos de maestros o investigadores para orientar los apoyos especiales y diferenciales que la situación tan inequitativa como la descrita reclama, si lo que pretendemos es hacer valer el derecho a la educación ”.

Por lo tanto, las conclusiones de una evaluación no pueden basarse exclusivamente en los resultados de las pruebas, aunque se sigue considerando los resultados que arrojen éstas, un indicador útil.



### 2.2.1. DESARROLLO DE LAS PR EN MEXICO

En el caso de México, aún es muy escasa la utilización de PR, esto se fundamenta con una revisión que realizó González, (1994), sobre el uso de las pruebas psicológicas en el campo educativo de México, en la que se puede llegar a las siguientes conclusiones:

- La mayoría de las pruebas que se utilizan oficialmente por la SEP son creadas en su mayoría en otros países. (Escala evolutiva del desarrollo de Arnold Gesell, Bender, Escala de inteligencia de Stanford-Binet (Terman-Merril), DFH, Escalas de inteligencia Weschsler, Batería de evaluación intelectual Secuencial-simultánea. Kaufman, Sistema de evaluación Multicultural y pluralístico (SOMPA).
- Algunas de estas son adaptadas a la población mexicana, sin embargo algunas son adaptaciones de hace 20 años y no se han vuelto a renovar. Asimismo no cuentan con normas para la población mexicana.
- La mayoría son utilizadas para la población infantil.
- Celas pruebas revisadas solamente tres fueron creadas en el país: la Monterrey, Inventario de Ejecución Académico (IDEA), y el Test de Diagnóstico en Habilidades Básicas (TEHBA); la primera de estas no fue estandarizada, solamente adecuada a los cambios curriculares; la segunda no se sabe si se adecuó al presente plan de estudios; y la tercera ésta estandarizada para la población infantil durante la citada revisión, se hizo su adaptación para la población adulta analfabeta.
- De todas estas pruebas solamente dos son de rendimiento y son utilizadas para la población infantil y , adulta.
- En su mayoría las pruebas utilizadas son de aptitudes o de inteligencia y de personalidad.

Por otro lado, la primera evaluación en secundaria utilizando PR con base en normas fue realizada a mediados de los ochentas .

Sin embargo, se sabe que en el año de 1994 el PPME, la Dirección General de Educación Especial y la Dirección General de Educación Primaria {DGEP), dejaron de utilizar las pruebas de aptitudes para utilizar las PR y atender a los alumnos con dificultades de aprendizaje.

Las PR no pueden ser traducidas y adaptadas en otros países, por que el contenido de los reactivos está directamente relacionado con los programas de estudio. Latapí (1994), recomienda que para mejorar la calidad de la educación en México, es necesario que se destinen recursos para el desarrollo tecnológico de la educación.

A consecuencia de todo lo anterior el objetivo de la presente investigación es construir un instrumento que evalúe el rendimiento escolar en la población adolescente, por lo menos en el área de Civismo primer grado.

### **2.3. ASPECTOS GENERALES DE LAS PR**

Las PR o de aprovechamiento también suelen llamarse objetivas según Severino (1985):

"Las pruebas objetivas son instrumentos de evaluación constituidos por una serie de preguntas o ítems claros y breves que exigen respuestas también breves, claras y precisas, a fin de evaluar el rendimiento o aprovechamiento instructivo de los alumnos de la forma más objetiva y ecuánime posible, mediante la corrección y calificación de las respuestas obtenidas con independencia de las personas que las califiquen",

Cuando estas pruebas están tipificadas con normas y baremos para su aplicación previa a cada muestra se les denomina también "tests de instrucción. o de "aprovechamiento".

Por otro lado, las PR se emplean para valorar y evaluar el desarrollo de las habilidades de un sujeto-grupo en cuanto a contenidos de aprendizaje. Por lo cual, su función en el ámbito escolar es esencial. Las PR se clasifican de la siguiente manera:

a) Baterías de evaluación especial: se refieren a la evaluación de todo un ciclo escolar: en algunos países se utilizan para certificar las habilidades de un profesionista o bien como requisito para ser aceptado en algún programa de postgrado.

b) De materias aisladas: se refieren a la evaluación de un curso específico, suelen ser utilizadas como exámenes.

c) Diagnósticas: están diseñadas para resaltar las destrezas e inhabilidades en el desarrollo de capacidades. Son utilizadas para detectar aspectos específicos de un programa que un alumno no haya logrado aprender, a fin de aplicar programas correctivos.

Las características generales de las PR son:

- Evalúan aspectos específicos del aprendizaje escolar .
- Identifican las habilidades que un sujeto ha logrado adquirir de un programa escolar
- No predicen lo que será la vida futura del sujeto.
- Las habilidades se especifican en términos conductuales.

Como se mencionó anteriormente, las PR al aparecer de nueva cuenta se correlacionaban con las Pruebas de Aptitudes, por lo cual, comparten algunas características entre las que destacan:

- Pertenecen a las medidas de ejecución máxima, ya que se interesan en obtener información del mejor rendimiento posible de un individuo, mientras que las de

ejecución típica se interesan por la conducta usual o habitual.

- La interpretación de las pruebas se basan en normas y criterios.

Con base en normas se pretende interpretar cada uno de los resultados individuales sobre la base de valores de tipo estadístico (las normas), referidas a la distribución de los resultados de la prueba en relación con un grupo de referencia (grupo normativo). Mientras que las pruebas con base en criterios tienen como objetivo analizar a cada individuo de acuerdo a un objetivo de comportamiento establecido (el llamado criterio).

A partir de la diferenciación que hace Scriven (1967) entre evaluación formativa y sumativa, se empieza a nombrar a las pruebas con base en normas como sumativas, y las diseñadas con base en criterios como formativos.

La prueba a construir en esta investigación se encuentra dentro de la clasificación de materia aislada, ya que solamente se hará referencia al primer grado de Civismo; y cuenta con la característica de ser sumativa, ya que sus resultados serán analizados estadísticamente. En el siguiente capítulo se conocerán algunos aspectos importantes para poder analizar los resultados estadísticamente.

## **A) Usos de las PR**

Los tests de aprovechamiento desempeñan un importante papel en la escuela. Las autoridades administrativas utilizan versiones estandarizadas para dividir los niveles, colocar estudiantes transferidos, evaluar los currículos, identificar niños excepcionales y aplicar a la comunidad lo que acontece en la escuela. Los orientadores los utilizan para ayudar a los estudiantes a tomar decisiones educativas y vocacionales más realistas. No obstante su función más significativa está en el programa educativo de la escuela. Aquí, los tests de aprovechamiento se vuelven parte integral del proceso de enseñar y aprender y se dejan sentir más directamente en el desarrollo del estudiante.

Se utilizan tests estandarizados y otros elaborados por el propio maestro para medir el aprovechamiento en clase, aunque estos últimos son los que más se emplean (Gronlund, 1980).

El uso de estas pruebas se puede clasificar según Brown (1991):

\* Usos para estudiantes:

- Proporcionarles realimentación a los estudiantes sobre la eficiencia de su aprendizaje.
- 
- Para impulsar a los alumnos a estudiar .
- Cuando se usan pruebas de rendimiento con fines de diagnóstico, la meta es identificar las causas de las dificultades de los estudiantes e indicar las acciones correctivas posibles.
- Hay también momentos en que las PR pueden ser útiles en el proceso de asesoramiento, le permite al estudiante discernir sobre sus intereses, sus hábitos de estudio o sus motivaciones.
- Las PR se pueden utilizar para predecir el éxito académico futuro.

\* Usos para los instructores:

- Proporcionar realimentación al instructor.
- Proporcionar también indicios respecto a los aspectos de la enseñanza que conducen a un aprendizaje incompleto.
- Como base para la asignación de calificaciones.

- Como prueba previa o de diagnóstico antes de iniciar la instrucción
- Verificar los estudios de los alumnos.
- Para estimular los debates
- \* Usos administrativos:
  - Como instrumentos de selección.
  - Ayudar a tomar decisiones de colocación o clasificación.
  - También para saber si una persona ha alcanzado el nivel mínimo de competencia que se necesita para dedicarse a alguna actividad.
- \* Evaluación de la instrucción:
  - El empleo de las medidas de rendimiento ayudan a evaluar la eficacia de los maestros, los planes de estudios o los métodos de enseñanza.
  - Aunque, su empleo en éstas se debe de abordar con precaución.

## **B) Cuestionamiento a las PR**

De las críticas que se le han hecho a este tipo de pruebas sobresalen las siguientes:

- Las críticas relativas a los reactivos ya la construcción de estas pruebas. Generalmente este tipo de pruebas se basan en reactivos de elección múltiple, estos reactivos tienen ciertas imperfecciones, pueden ser ambiguos o castigar a los alumnos que tengan conocimientos más profundos. También la dificultad que

supone su preparación adecuada, con los necesarios requisitos de validez, confiabilidad y eficacia práctica.

También se ha criticado que estas pruebas sólo miden aprendizajes memorísticos, pero esto es incierto, ya que como se mencionó anteriormente, estas pruebas valoran diferentes niveles cognitivos (conocimiento, comprensión y aplicación).

Lo que sí es cierto es que evalúan aspectos cognoscitivos restringidos de tipo convergente. Los referidos al pensamiento creativo, valores y actitudes no se consideran. Por ello se recomienda utilizar otras técnicas que permiten obtener información relativa a este tipo de aprendizaje.

- Muchas críticas se basan en las decisiones que se toman con base a los resultados de estas pruebas, entre las que se encuentran; que si determinan los planes de estudio y si los maestros enseñan para el examen, así como los métodos de instrucción y los sistemas escolares.
- Otra de las críticas que se les hacen es que se basan en un modelo de discriminación o de diferencias individuales. Ya que se reconoce como meta de las pruebas obtener conocimientos sobre el dominio de las materias importantes y no el de hacer distinciones entre los alumnos. Asimismo, se debe hacer hincapié en la destreza y no en las diferencias individuales.

### **CAPITULO III. ASPECTOS PSICOMETRICOS DE LAS PRUEBAS**

En este último capítulo se señalará la forma en como se valida una prueba, tomando en cuenta para esto los conceptos de validez y confiabilidad.

Por lo tanto, es importante comenzar a explicar lo referente a la medición.

### **3.1.MEDICION**

La medición es un concepto ampliamente relacionado con la evaluación, porque permite cuantificar los resultados de esta última. Ya que según Mehrens y Lehmann (1982), mencionan que:

"la medición y la evaluación son imprescindibles en la toma de decisiones en la educación".

La medición se refiere según Brown (1991), al:

"asignar números a la conducta, de acuerdo a reglas preestablecidas".

La medición puede realizarse por medio de observaciones, escalas de evaluación o cualquier otro instrumento que nos permita obtener información en una forma cuantitativa.

En el caso de la medición del aprovechamiento de los alumnos, que es un punto base de esta investigación, se utiliza la construcción de PR.

Lo que da el carácter de instrumento de medición a estas pruebas, es su contenido y lo que se propone medir, y para que se conviertan en escalas de medida se recurre a los métodos estadísticos.

#### **3.1.1. ESCALAS DE MEDICION**

Retornando el capítulo II, de la definición de escalas se puede decir que es un conjunto de valores numéricos asignados a sujetos, objetos o comportamientos con el propósito de cuantificar y medir sus cualidades. Sirven para medir actitudes, valores e intereses (Ary y Cheser,1990).



Existen varios tipos de escalas, sin embargo, comúnmente se utilizan cuatro:

- Escalas Nominales: Esta escala se considera la más primitiva. Consiste en asignar un número o nombre distinto a cada categoría para clasificar el fenómeno a estudiar
- Escalas Ordinales: Este tipo de escala pretende ordenar en función de una dimensión las características de un determinado fenómeno.
- Escalas de Intervalos: Pretende interpretar las distancias entre las diferentes categorías o puntuaciones. Son comúnmente utilizadas por psicólogos y educadores.
- Escalas de Razón: Cuenta con las características de las escalas anteriores (intervalos), pero además cuenta con un punto cero absoluto en su origen. No son utilizadas.

### **3.1.2. ERRORES DE MEDICION**

La medición en Psicología, se encuentra con algunas dificultades, llamadas errores de la medición.

Según Brown (1991) el error es:

"... cualquier efecto irrelevante para los fines o los resultados sobre las faltas de consistencia en la medición".

Esta definición tiene dos sentidos, por un lado, se puede considerar que una variable introduce error si: 1) si no tiene que ver con las finalidades de la medición (errores constantes) y 2) produce falta de consistencia en las calificaciones de una situación a otra (errores al azar). Por otra parte, el aspecto de irrelevancia se ocupa primordialmente del error sistemático y se vea como el aspecto de validez.

Entre los errores al azar sobresalen:

- Errores dentro de la prueba: Cualquier aspecto que haga que un sujeto responda a un reactivo sobre bases distintas de los conocimientos de la respuesta "correcta." o apropiada. También serán fuente de error reactivos muy difíciles que propician que se opte por adivinar.
- Errores en la aplicación de la prueba: Las condiciones en las que se aplica la prueba entre las que se encuentran: la falta de comprensión en las instrucciones, interrupciones inesperadas, condiciones climatológicas que afectan el proceso. .
- Errores por causas del examinado: Este tipo de errores son los más difíciles de controlar, y obedecen a aspectos como motivación ante la situación de prueba, maduración, salud, fatiga, ansiedad, etc.

Por otra parte, entre los errores constantes o sistemáticos, se encuentran:

- Errores de Halo: Se define como la tendencia a evaluar aspectos parciales de la persona, guiándose por la impresión general que se tiene de la misma.
- Error de Severidad: Se define como la tendencia general a conceder puntuaciones bajas a los individuos en todas las características observables.
- Error de Indulgencia o Generosidad: Consiste en conceder puntuaciones demasiado altas.
- Error de Tendencia Central: Se define como la tendencia a evitar los juicios extremos ya buscar el punto intermedio de la escala de clasificación .

- Error de Lógica: Tendencia a calificar en forma similar conductas que según el observador están lógicamente relacionadas. Este tipo de error tiene dos vertientes:
  1. Error de consenso: Tendencia a calificar al sujeto según la impresión u observación de terceras personas.
  2. Error de proximidad: Tendencia a dar una calificación similar a aspectos que están próximos en la pauta de observación .

Por otro lado, es importante hacer mención de la consistencia, por lo tanto, a continuación se hablará de la confiabilidad y de la validez, que explicarán a que se tiene la consistencia

### **3.2. CONFIABILIDAD**

Después de mencionar los errores que se producen en las pruebas o en cualquier medición a continuación se dará a conocer las características que debe poseer un instrumento de medición, que son: la confiabilidad y la validez.

La confiabilidad se refiere al grado de consistencia o exactitud de un instrumento o prueba. Lo conveniente sería que los resultados obtenidos en dos aplicaciones, en condiciones ligeramente distintas, fueran semejantes (Gronlund, 1980; Severino, 1985; Rodríguez y García, 1992, entre otros).

La necesidad de consistencia en las mediciones compete primordialmente al ámbito de aspectos psicológicos o educativos, ya que en éstos las mediciones son de manera característica mucho menos confiables que las mediciones físicas, (peso longitud, etc.) la variación entre una medición y otra son mínimas lo que produce resultados bastantes precisas, cosa que no siempre ocurre con la conducta humana.

La variación en las puntuaciones de una persona es una característica llamada "varianza de error" y las fuentes de esta variación se conocen como "fuentes de error". (Mehrens y Lehmann, 1982).

La confiabilidad se ocupa básicamente en dar respuesta a dos tipos de problemas:

La consistencia de las calificaciones: a) relación entre las calificaciones obtenidas en diversas condiciones de medición; b) la diferencia que podría ser aceptable al ser sometido a la prueba en dos momentos; c) la diferencia entre las calificaciones obtenidas y las "reales".

Las causas de la discrepancia entre las calificaciones de la prueba: a) los factores que producen calificaciones poco consistentes; b) la magnitud relativa a los afectos de calificaciones poco consistentes.

Para estimar el grado o coeficiente de confiabilidad de una prueba, se utilizan generalmente tres métodos:

\* Estabilidad: Una medida de estabilidad, frecuentemente llamada estimación test-retest de confiabilidad. se obtiene aplicando una prueba a un grupo de individuos, volviéndola a aplicar posteriormente a los mismos individuos y correlacionando las dos series de puntuaciones alcanzadas (Mehrens y Lehmann, 1982).

El tiempo que media entre las dos aplicaciones puede ser de algunos minutos o de varios años. Así, es posible obtener diferentes valores del coeficiente de confiabilidad, dependiendo del tiempo que transcurre entre las aplicaciones de las pruebas. Sin embargo, no es el tiempo propiamente dicho el que causa los cambios en los resultados, sino las experiencias que haya tenido el individuo entre los exámenes.

En el caso de las PR, el aprendizaje escolar suele variar tiempos relativamente cortos. en este sentido este tipo de confiabilidad no es adecuado para estos instrumentos. Al respecto Nunnally y Berinstein (1995) señalan que:

“Las pruebas escolares. están diseñadas para ser temporalmente inestables si no lo fueran, la intervención educativa sería inútil”.

- Equivalencia: Con este método se administra al mismo grupo; durante la misma sesión, dos formas equivalentes de un test. Las formas del test son equivalentes en tanto que están elaboradas para medir las mismas capacidades (es decir, con arreglo a la misma tabla de especificaciones), pero, para determinar la confiabilidad, también es importante que estén elaboradas independientemente.

Sin embargo, por los costos que implica construir una forma alterna se acostumbra utilizar por mitades que consiste en dividir el test en dos partes, una de ellas con las preguntas numeradas pares y la otra las preguntas numeradas nones. La correlación entre las puntuaciones de los reactivos pares y los numerados nones da un coeficiente de confiabilidad del test-completo (Karmel, 1978).

La ventaja del método por mitades radica en que se necesita únicamente una fórmula de la prueba y una sola aplicación de la misma para poder estimar su confiabilidad.

Al respecto Nunnally y Berinstein ( 1995) señalan que la confiabilidad en formas alternas y por mitades están generalmente entre .75 y .90 para pruebas con reactivos de opción múltiple.

- Homogeneidad. Por ende, se puede definir la homogeneidad como la consistencia en la ejecución en todos los reactivos de una prueba psicológica (Loevinger, 1947; Anastasi, 1961; citados por Brown, 1991 ).

Este procedimiento indaga acerca de la consistencia interna dada por la relación entre reactivos, los índices de homogeneidad hacen un hincapié primordial sobre la estructura interna de la prueba. Este procedimiento se ubica dentro de la teoría de dominio-muestra que presupone que, asociado a cada rasgo, hay un universo hipotético de reactivos que miden esa construcción en particular.

Las fórmulas que se utilizan para estimar la confiabilidad por medio de este método son: Alpha de Cronbach y el Kuder-Richardson 20 y 21.

En el caso de pruebas y aptitudes es fundamentalmente determinar la consistencia intensa; para las pruebas de rendimiento, los índices de confiabilidad suelen ser bajos (Nunnally y Berinstein, 1995), ya que 00 se miden rasgos de la persona, sino sus ejecuciones.

En la construcción de pruebas de rendimiento es importante considerar dos factores que influyen en la confiabilidad del instrumento: la variabilidad de las puntuaciones y la longitud de la prueba.

La variabilidad de un reactivo le da el tipo de población a la que se aplica, entre más homogénea sea la población, menor variabilidad se encontraran en las respuestas, y entre más heterogeneidad de la población mayor será la variabilidad de las respuestas.

El tipo de reactivos también afecta la variabilidad de las respuestas. Reactivos muy fáciles (más de .85 de grado de dificultad) permite que casi todos los contesten y por lo tanto, no habrá variaciones en los resultados de las pruebas; otro tanto ocurre con los reactivos muy difíciles (-30 o menos). Mehrens y Lehman (1982) consideran que los reactivos deberían idealmente tener un grado de dificultad de .50, o acercarse a este valor .

Por otra parte, la longitud de la prueba afecta la confiabilidad. Entre más larga sea la prueba (con reactivos que presenten variaciones) mayor será la confiabilidad.

### **3.3. VALIDEZ**

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

El término validez se refiere al grado en que una prueba mide lo que pretende medir realmente; es decir, si la prueba sirve en realidad al propósito para el cual fue hecha (Rodríguez y García, 1992).

Tanto la validez como la confiabilidad son términos numéricos que van estrechamente relacionados, no podría hacerse referencia a uno, dejando de lado al otro. La validez de una prueba estará siempre, limitada por su confiabilidad, los índices de confiabilidad baja afectan la validez de la prueba; en este sentido, la confiabilidad es una condición necesaria - pero no suficiente- para la validez de una prueba. Ya que para que una prueba sea válida o veraz, debe antes que nada ser confiable.

En términos estadísticos la validez se define como la proporción de varianza real que es relevante para los fines del examen. La relevancia se atribuye a las variables que mide la prueba, en este sentido el error de medición significativo para la validez, es el que introduce cualquier variable irrelevante para sus fines (Brown, 1991).

La validez se puede definir en varios niveles y diversos modos; dependiendo de las características que se pretenden medir, así como el objetivo para el cuál se construye una prueba cobrará mayor importancia un determinado tipo de validez. El procedimiento para validar una prueba de aptitudes o una de rendimiento es diferente; así mismo, la validación de una prueba selectiva difiere de la validación de una prueba diagnóstica. La validez, al igual que la confiabilidad, es siempre específica de la situaciones.

Por lo tanto, existen varios tipos de validez, pero entre ellos destacan tres; 1) validez de contenido, 2) validez de criterio y 3) validez de constructo.

La construcción de pruebas de rendimiento escolar deben privilegiar la validez de contenido y la de criterio, en ese orden (Nunnally y Berinstein, 1995); a continuación se describen los diferentes tipos de validez:

### **3.3.1. Validez de Contenido.**

Este tipo de validez se refiere, como su nombre lo indica, al contenido de la prueba. El contenido significa los constituyentes sustantivos de la "materia" que se pretende medir; por consiguiente, en la construcción de la prueba, se hará hincapié, primordialmente, en el área temática que se cubra.

La definición de contenido establece que los reactivos de la prueba deben ser una muestra representativa del universo de las conductas o contenidos posibles, esto se fundamenta con lo señalado por Gronlund (1980):

"El deseo natural de cada uno de nosotros es que cualquiera que sea la prueba de aprovechamiento que construyamos, o seleccionemos, arroje resultados que sean representativos de lo tópicos y comportamientos que deseamos medir".

Por tanto, la validez de contenido reviste particular importancia en el caso de las pruebas de rendimiento. En este caso interesa la precisión con que el test mide el contenido de la materia y los temas vistos durante el período de enseñanza.

Por otra parte, el procedimiento para lograr la validez de contenido comprende:

a) Identificar los temas de la materia y los resultados conductuales que se habrán de medir.



- b) Revisar el programa de estudios.
  
- c) Diseñar un plan de prueba en el que se especificarán por unidad y niveles cognitivos, cantidad y tipo de reactivos.
  
- d) Revisar los libros de texto de la materia para saber como se aborda en los de texto cada contenido del programa.
- e) Redactar un banco de reactivos, de acuerdo al plan de prueba especificado se procede a redactar estos reactivos que deben ser adecuados a las orientaciones del programa y de los libros de texto.
  
- f) Diseñar una prueba preliminar la cual será juzgada por especialistas y maestros (jueces expertos).
  
- g) Modificar reactivos de acuerdo a la validación de jueces expertos.
  
- h) Prueba preliminar se aplicará a una muestra representativa de la población a la que se pretende evaluar (piloteo).
- i) Se analizarán los reactivos en cuanto al nivel de dificultad y grado de discriminación.

El índice más importante es el del poder discriminativo de los reactivos. Este análisis estadístico indica la amplitud con la que el reactivo mide lo que debe medir.

El análisis de discriminación, se usa para construir todos los tests estandarizados a fin de determinar el grado en que cada reactivo de prueba mide la misma cosa que la prueba completa en que está incluida. Primero, se encuentra el 25% superior de estudiantes (puntuaciones totales de test) y el 25% inferior. luego se hace un análisis de cada reactivo contando el número de estudiantes, superiores e inferiores, que contestaron correctamente el reactivo. El porcentaje del grupo inferior se resta del porcentaje del grupo superior.

El dato resultante indica hasta qué punto el reactivo dado cumple con su función de discriminar. Cuanto más grande la diferencia tanto mejor el reactivo, (Karmel, 1978).

Mientras que el nivel de dificultad de los reactivos se refiere al porcentaje de personas que responden correctamente (Karmel, 1978; Brown, 1980).

j) Seleccionar los reactivos, con base en el análisis anterior , para la prueba definitiva.

k) Diseñar la prueba definitiva.

### **3.3.2. Validez de Criterio.**

Este tipo de validez se refiere a la comparación de las Puntuaciones de un test con un criterio o criterios externos, como las calificaciones escolares o el éxito en el trabajo.

Según Mehrens y Lehmann, ( 1982), señalan que algunos autores establecen una distinción entre tipos de validez de criterio:

La validez concurrente y la validez predictiva. El propósito de cada una de éstas es:

- Validez predictiva: se interesa por la utilidad de la puntuación lograda en la predicción de alguna futura ejecución.
- Validez concurrente: se pregunta si la puntuación lograda en la prueba puede o no sustituirse por alguna forma menos eficiente de reunir datos criterio.

Así, el que si la validez de criterio deberá expresarse como concurrente o predictiva, depende de si está más interesado en predecir o en evaluar una condición presente.

Además, ésta forma de validez se establece con una tabla de expectativas o, en la mayoría de los casos por correlación entre la puntuación del test y una medida de criterio.

Las medidas de criterio, como cualquier otras medidas, deben poseer ciertas características para considerarse adecuadas (Brown, citado por Mehrens y Lehmann, 1982):

- Deben ser relevantes, la medición criterio debe en realidad, reflejar los aspectos más importantes del criterio conceptual. El grado de relevancia de una medida criterio implica un juicio de valor y no todos coincidirán en un caso específico.
- Otra característica, es que sea confiable.
- Una tercera característica de la medida de criterio es que ésta debe estar libre de predisposición o contaminación (Brogden y Taylor, citados por Mehrens y Lehman, 1982).

### **3.3.3. Validez de Constructo**

Este tipo de validez se utiliza para el estudio de los rasgos psicológicos y sus manifestaciones (inteligencia, capacidad de razonamiento aptitud verbal, etc.), de un determinado sujeto, estableciendo las diferencias individuales.

Según Brown (1991), la validación de construcción se lleva a cabo mediante una acumulación de evidencias respecto al rasgo que mide realmente la prueba. Las evidencias se pueden acumular de diversos modos y de varias fuentes distintas, incluyendo estudios de contenido y validez del criterio. Conforme se acumulan las evidencias se esclarece la relación entre el rasgo y la prueba, la evaluación viene a ser un proceso de juicio.

Esta validez es de gran importancia en los tests de personalidad, especialmente en las técnicas proyectivas. También se presentan en las pruebas de aptitud y en las escalas de actitudes.

## **METODOLOGIA**

La investigación para validar la prueba de rendimiento para el ciclo de secundaria, de la materia de Civismo primer grado. se realizó en dos fases: en la primera se planificó, diseñó y construyó la prueba preliminar. en la segunda fase se realizó el piloteo con una muestra de la población de referencia.

Esta investigación forma parte de una investigación más extensa realizada por la profesora Rosa María González quien coordinó la estandarización de 17 pruebas de rendimiento.

El procedimiento seguido permitió la validez de contenido y de criterio (concurrente). así como la confiabilidad por equivalencia y homogeneidad de las pruebas.

### **PRIMERA FASE. PLANIFICACION, DISEÑO y CONSTRUCCION DE LA PRUEBA PRELIMINAR**

a) Planificación de la prueba: Se realizó una revisión del Plan y los Programas de estudio de Secundaria (SEP. 1993). en términos de orientación curricular que guía la materia de Civismo primer grado, sus propósitos y contenidos.

Con base en esta revisión, se diseño un plan de prueba (anexo 1) en donde se describen los contenidos específicos por unidad. de la materia seleccionada.

Posteriormente. se realizó una revisión de los libros de texto de secundaria, específicamente de Civismo primer grado, acordes a los contenidos expuestos en el plan de prueba..

En el mercado existen alrededor de diez editoriales que publican libros de texto para las diferentes materias de Secundaria. Se realizó una revisión de los libros de texto de las diferentes editoriales más vendidas en el mercado; de acuerdo con entrevista a los responsables de las librerías: "La Casa del Libro" "Librería del Sótano de Coyoacán" y "Librería de Cristal", quienes informaron que "Editorial Trillas", "Editorial Planeta" y "Fernández Editores" son los más solicitados. Se pudo constatar que la forma y extensión en que cada texto desarrolla los contenidos es disparado, y que hay textos que no abordan algunos contenidos del programa.

Dentro de las especificaciones para las pruebas se determinó que los reactivos serían de opción múltiple (una respuesta correcta y 4 distractores) con un total de entre 50 y 70 reactivos, que es el tiempo aproximado de una sesión de clase (se calcula un minuto para la respuesta de cada reactivo).

b) Diseño de prueba preliminar: De acuerdo con el plan de prueba se redactaron reactivos suficientes que abarcaran los contenidos señalados para cada una de las materias. Para la redacción de reactivos se siguieron las orientaciones que diversos autores dan al respecto en términos de cantidad, claridad, pertinencia, adecuación en términos de edad y aspectos culturales (Gronlund, 1977; Mehrens y Lehman. 1982), que en este caso son alumnos de escuelas secundarias técnicas públicas ubicadas en el Distrito Federal.

Una vez redactados los reactivos se diseñó un formato de prueba preliminar y un cuestionario para jueces expertos a fin de pasar a revisión, fungiendo dos especialistas en currículo y dos maestros que imparten la materia.

En el cuestionario para especialistas en currículo (anexo 2) se valoró la claridad de los reactivos, la pertinencia en términos de relevancia dentro del programa de estudios y el nivel cognitivo que mide cada reactivo.

c) Construcción: Con base en las correcciones y sugerencias de los especialistas se procedió a rectificar aquellos reactivos mal diseñados a fin de editar y reproducir la prueba preliminar. Dando como resultado en total de 120 reactivos.

## **SEGUNDA FASE. PILOTEO DE PRUEBA PRELIMINAR y EDICION DE PRUEBA DEFINITIVA**

Para el análisis psicométrico de la prueba de rendimiento se procedió a pilotearla con una muestra de estudiantes inscritos en el primer grado de secundarias públicas ubicadas en el Distrito Federal del turno matutino del ciclo escolar 1995- 1996.

Se decidió pilotear la prueba con alumnos de secundarias técnicas exclusivamente por una cuestión de recursos.

### **SUJETOS**

Para la selección de la muestra se procedió por el método de muestreo aleatorio tritápico por conglomerados (Scheaffer, Mnedenhall y Ott, 1997), considerando a la Delegación como el conglomerado, la escuela como unidad primaria y el-la alumno-a como unidad última.

Para la selección de sujetos se procedió a elegir aleatoriamente seis Delegaciones Políticas: Azcapotzalco, Xochimilco, Coyoacán, Benito Juárez, Tlalpan e Iztapalapa. Posteriormente se seleccionó una escuela de cada Delegación en la que se eligió un grupo y en éste se seleccionaron aleatoriamente 22 alumnos-as.

Con fines de reposición (por mortandad de la muestra), se incluyeron un 5% de sujetos.

## **ESCENARIO**

Una vez gestionados los permisos respectivos ante la Dirección General de Secundarias Técnicas.

Las características generales de éstas escuelas son:

- En general, las seis escuelas presentaron condiciones más o menos similares: espacio reducido para un grupo de 50 alumnos, mala ventilación e iluminación.
- E.S. T .No.37 , la escuela se encuentra ubicada en una zona de clase baja.
- E.S.T. No.43, se ubica en una zona residencial.
- E.S.T. No.14, se encuentra ubicada en una zona de clase media.
- E.S. T .No.27 , se ubica en una zona residencial de clase media.
- E.S.T. No.39, la aplicación se realizo en el taller de electrónica
- E.S. T .No.94, se ubica en una zona rural a las afueras de Xochimilco.

En la primera sesión a los noventa minutos se suspendió la prueba y se indicó que el día siguiente se continuarían. En la segunda sesión los alumnos fueron entregando de acuerdo como iban terminando. La mayoría de los alumnos terminaron antes del tiempo máximo destinado (120 minutos) para contestar la prueba.

## **ANALISIS DE LA INFORMACION**

Se procedió a calificar y codificar cada prueba para conformar una base de datos en un ordenador, utilizando el programa WINSPSS a fin de determinar:

## **a) ANALISIS DE REACTIVOS**

El grado de dificultad de los reactivos se calculó a través de la media de cada reactivo.

El nivel de discriminación de los reactivos se determinó por medio de la correlación del reactivo con la suma del total de reactivos.

Análisis de distractores. Se realizó el porcentaje de respuesta de cada una de las opciones (1 respuesta correcta y 4 distractores).

Para la selección de los reactivos que conformaron la prueba definitiva se siguieron los siguientes criterios, en orden de importancia:

- Que se cubrieran todos los contenidos del plan de prueba de acuerdo a las observaciones de los jueces expertos.
- Que el nivel de discriminación fuera mayor de .15 (ver cap. III).
- Que el grado de dificultad estuviera comprendido entre .80 y .30 (ver cap. III).

Que los distractores atrajera el 5 % y el 17 % de las respuestas.

## **b) VALIDEZ**

A fin de garantizar la validez de contenido aquellos reactivos que presentaron los mejores índices de dificultad y discriminación se incorporaron a una tabla de especificaciones. En caso de quedar descubierto algún subcontenido, se incorporaron reactivos con índices menos adecuados, a fin de cubrir la totalidad de contenidos del programa de estudio.

Para determinar la validez de criterio (concurrente) se seleccionaron dos variables: promedio de calificación en la materia (medición intervalar) e interés en la materia con tres dimensiones: mucho, regular, nada (medición ordinal). Para determinar los coeficientes de



validez de criterio se correlacionaron la suma del total de reactivos de la prueba con el promedio y el interés por la materia.

### **c) CONFIABILIDAD**

Para determinar la equivalencia de las pruebas se procedió a utilizar el método por mitades utilizando la fórmula de Spearman- Brown (Brown, 1991). Para determinar la homogeneidad de las pruebas se utilizó el coeficiente alpha de Cronbach (Brown 1991).

## **RESULTADOS**

Adjunto a este trabajo se presenta la prueba de rendimiento definitiva de primer grado en la asignatura de Civismo de nivel secundaria (anexo 3).

La prueba incluye la carátula, la prueba misma, hoja de respuestas, claves de respuestas correctas, tabla de especificaciones de cada uno de los reactivos en cuanto a nivel cognoscitivo (conocimiento, comprensión y aplicación), el análisis de cada reactivo en relación con el grado de dificultad y el nivel de discriminación.

Como se ha mencionado, la validación de una prueba de rendimiento implica un proceso que incluye la validez del contenido, la validez de criterio, el análisis de reactivos y la confiabilidad de la prueba. A continuación se describe la muestra con la que se validó la prueba de rendimiento para secundaria, así como los resultados psicométricos obtenidos.

### **MUESTRA**

Para facilitar la interpretación de los resultados obtenidos en el proceso de validación de la prueba de rendimiento se transformaron los puntajes brutos en puntajes estándar, a través de dos tipos de calificaciones: percentiles (medida ordinal que facilita la comprensión de las calificaciones), y calificaciones. Z (medida intervalar que se requiere para trabajar estadísticamente los puntajes).

El promedio de calificación, en la materia de Civismo de primer grado, es de 51.4 reportado en percentiles.

Para identificar posibles diferencias entre hombres y mujeres se procedió a establecer cinco rangos de calificación (muy deficiente, deficiente, regular, bien y muy bien), con base en los puntajes transformados (calificación Z). En el gráfico 2 se presentan los porcentajes por sexo, por lo que se puede decir que las diferencias entre niños y niñas son mínimas, sin embargo, ellas superan ligeramente a los varones.

### **VALIDEZ DE CONTENIDO**

La validez de contenido se refiere a los constituyentes sustantivos de la materia que se pretende medir, y se determina por la valoración cualitativa en relación con los contenidos del Programa de estudio señalados a través de un plan de prueba previamente diseñado. Al final de la prueba se presentan los contenidos que cubren la prueba de Civismo de primer grado.

En el diseño de la tabla de especificaciones se cuidó contemplar el total de contenidos del Programa de Estudio aprobado en la última Reforma Educativa de la materia de Civismo primer grado, cosa que se cumplió en un 90 % aproximadamente. En cuanto a los niveles cognoscitivos se procuró que el diseño de los reactivos se integrara algo más que el conocimiento, redactando reactivos de comprensión y aplicación, como lo señalan las nuevas orientaciones curriculares.

Como se podrá apreciar en la Tabla antes citada el resultado es disparejo, ya que la prueba de Civismo de primer grado contiene 25 reactivos de conocimiento, 24 de comprensión y 11 de aplicación.

En el plan de prueba original se tenían contemplados mayor cantidad de reactivos de comprensión y conocimiento en la materia, reactivos que los expertos en currículo aceptaron como relevantes; sin embargo, cuando los maestros calificaron los reactivos,

valoraron negativamente algunos de comprensión y muchos de aplicación, argumentando que no los incluía el libro de texto.

Esto se pudo verificar en la revisión que se realizó para elegir los libros de texto que respondieran al plan de prueba en lo que se pudo observar una gran cantidad de contenidos que solamente se referían a conocimiento, ya que en su mayoría los contenidos se basaban en artículos constitucionales como en la "unidad VIII, las libertades"; o les faltaban actividades de comprensión y aplicación. En algunos casos como la "unidad II, los derechos de niños y jóvenes"; que contenía de dos a tres artículos y era superficial en cuanto a los objetivos del programa oficial.

El número de reactivos que contenía la prueba preliminar era de 120, el número de reactivos de la prueba definitiva fue de 60, los reactivos eliminados fueron 60. La eliminación de reactivos no es una actividad sencilla, se tiene que conjugar una serie de factores (validez de contenido, dificultad y discriminación de reactivos, coeficientes de confiabilidad y longitud óptima de la prueba) para decidir que reactivos quedarán integrados en la "mejor" prueba definitiva posible.

### **VALIDEZ DE CRITERIO (CONCURRENTE)**

La validez de criterio se refiere al uso de un instrumento para estimar alguna conducta criterio que es externa al mismo, en la que se busca establecer una relación estadística entre los resultados de la prueba con un criterio predeterminado.

Como ya se mencionaba en el capítulo III, rara vez excede de .3 a .4 el resultado de la correlación prueba-criterio. Para la validación de criterio de la prueba se tomaron dos variables: 1) promedio de calificaciones (.36) y 2) interés por la materia (.09), presentados en el gráfico 3. Para intentar comprender la baja correlación que se presentó en esta materia, se procedió a analizar el interés que los alumnos expresaron con base a tres niveles: mucho, regular y nada. En el gráfico 4 se presentan los resultados.

Al analizar las fluctuaciones integradas (gráfico 4) en cuanto a interés promedio por la materia, se observa que la mayoría de los alumnos manifiestan un interés regular (70.3 %), esto se relaciona con lo que menciona Quiroz (1995):

"Civismo tiende a considerarse una materia "fácil" y poco significativa en cuanto a los índices de reprobación".

## **CONFIABILIDAD**

La confiabilidad se refiere a la consistencia de las mediciones. La validez de una prueba, así como el grado de dificultad, el nivel de discriminación y la longitud influyen en la determinación de sus coeficientes de confiabilidad. Para las pruebas de rendimiento la consistencia se obtiene a través de la equivalencia y la homogeneidad de la prueba; a continuación se presentan estos resultados.

## **GRADO DE DIFICULTAD Y NIVEL DE DISCRIMINACION DE LOS REACTIVOS**

Al final de la prueba se describen el grado de dificultad y nivel de discriminación de cada uno de los reactivos que componen la prueba (anexo 4). En el capítulo III, se señalaba que un reactivo es adecuado cuando tiene un grado de dificultad entre .30 y .85, y un nivel de discriminación de .20 o mayor. La prueba de Civismo primer grado obtuvo un grado de dificultad de .56 en promedio y un nivel de discriminación de 0.28, (gráfico 5), por lo que se puede decir, en términos generales los reactivos que componen la prueba de rendimiento de Civismo primer grado tiene índices adecuados en cuanto a grado de dificultad de los reactivos y nivel de discriminación.

## **COEFICIENTE DE EQUIVALENCIA Y HOMOGENEIDAD**

En el capítulo III, se señalaba que una prueba de rendimiento debería tener coeficiente de confiabilidad mínimos de .75 o mayores. Consultando el gráfico 6 es posible afirmar que la prueba de Civismo primer grado, presenta adecuados índices de confiabilidad, en cuanto a

equivalencia y homogeneidad, ya que presenta una equivalencia de .80 y una homogeneidad de .86.

El proceso de construcción de la prueba implica un proceso en donde constantemente se está valorando cuales son los mejores reactivos que cumplan tanto con aspectos de validez de contenido como la confiabilidad. El tener un mayor banco de reactivos piloteados ayuda a seleccionar "los mejores".

## CONCLUSIONES

El ámbito escolar mantiene una gran demanda por las pruebas de rendimiento a nivel internacional, sin embargo, en México no existen pruebas de rendimiento que puedan ser utilizadas por los diferentes grupos que se interesan en la investigación evaluativa de tipo sumativa. por lo cual, éste trabajo pretende dar una alternativa a estos grupos.

A continuación se analiza el proceso de validación de la prueba.

El promedio general en cuanto a calificaciones en la materia de Civismo primer grado, es de 51.4. Cualquier instrumento psicométrico necesita explicitar el significado de las cifras que arroja; en educación, a través de las pruebas que el maestro aplica, tradicionalmente se ha establecido una escala del 0 al 10 en donde se ha marcado como criterio de reprobación el 5 o cualquier cantidad menor. Este criterio es convencional, ya que un examen con reactivos con menor grado de dificultad daría por resultado un gran número de alumnos aprobados, el caso Contrario (reactivos más difíciles) presentaría un gran número de reprobados.

Una prueba de tipo normativo selecciona reactivos con grado de dificultad de .50 en promedio. Un principio de construcción de estos instrumentos se basa en una distribución normal, por lo que el promedio de calificaciones rondarán este valor, es decir, siempre habrá mejores y peores resultados en menor proporción, y el grueso de los sujetos quedará comprendido en el término medio, en este sentido las calificaciones que se presentan no

significa que la mayoría de los alumnos estén reprobados, simplemente señala el grado de dificultad de los reactivos.

Se considera que el hecho es importante señalarlo, ya que diversas pruebas de tipo selectivo que se aplican a la población escolar (vg. EXANI I y II) frecuentemente argumentan que los alumnos rechazados tienen calificaciones reprobatorias; ante esto es necesario solicitar cual es el grado de dificultad de los reactivos, ya que de éste dependerá la proporción de alumnos que aprueben-reprueben.

Al analizar los resultados por sexo, no se encontraron diferencias significativas, para valorar con más detalle esta información se procedió a formar 5 grupos:

- En el grupo “muy deficiente” se conformó en su mayoría por niñas.
- El grupo “deficiente” se conformó en su mayoría por niños.
- En los grupos “regular” y bien manifiestan un equilibrio entre niñas y niños.
- grupo “muy bien” se conformó en su mayoría por niñas. (ver gráfico 2).

Diversas investigaciones señalan que el interés que los adolescentes muestran por sus materias tienden en promedio a estar sesgados genéricamente: ellas se orientan principalmente por áreas de conocimiento memorístico y de servicio. y ellos por conocimientos que no les implica personalmente.

En cuanto a validez de contenido, se puede decir que aunque el plan de prueba original pretendía contener en igual número reactivos de conocimiento, de comprensión y de aplicación, al ser revisados por expertos en currículo y profesores en la materia valoraron negativamente algunos reactivos de comprensión y muchos de aplicación; y al ser aplicada y calificada se descartaron algunos otros reactivos debido a su grado de dificultad. quedando un total de 60 reactivos en la prueba definitiva (25 de conocimiento, 24 de comprensión y 11 de aplicación).

En resumen, se puede decir que los problemas que se detectaron para la redacción de éstos reactivos son los siguientes:

1. Los reactivos se diseñaron de acuerdo con el programa de estudio de Civismo primer grado y este incluye una cantidad extensa de contenidos, así como objetivos que marcan el orden y los límites que se deben seguir, lo que no permite profundizar la comprensión y aplicación a la vida cotidiana por parte de los alumnos.
2. Los reactivos se redactaron siguiendo los contenidos de los libros de texto que en su mayoría son solo de conocimiento, en algunos casos contienen algunas actividades de comprensión y de aplicación, y en otros casos éstas son nulas; así mismo existe diversidad en sus contenidos, ya que algunos libros son de nueve unidades, otros de ocho y otros de seis, sin embargo, solamente fueron utilizados aquellos que respondieran a los objetivos del programa de estudio de la materia.
3. En su mayoría fueron descartados reactivos que no cumplían con el grado de dificultad o el nivel de discriminación adecuado siendo éstos de comprensión y aplicación, lo que hace pensar que los profesores continúan promoviendo el aprendizaje memorístico.

Lo anterior sugiere que las formas de enseñanza-aprendizaje en éste ciclo continua privilegiando aspectos memorísticos.

Al revisar los reactivos de otras pruebas de secundaria: Examen de Admisión a Bachillerato (Sánchez Azura et al, 1992), 1001 Preguntas para Ingresar al Bachillerato {UNAM, 1995) y Guía Práctica para el Examen de Admisión a Secundaria ( 1992), las cuales Presentan una composición similar a la prueba presentada por ésta investigación .

Como ya se mencionó, es importante para la validez de contenido que existan: grado de dificultad y nivel de discriminación.

En el primero de los casos los puntajes son adecuados de acuerdo con los estándares internacionales para éste tipo de instrumentos, ya que el grado de dificultad que se obtuvo fue de .56 y el nivel de discriminación fue de .28 (gráfico 5).

Un principio de distribución de éstos instrumentos se basa en una distribución normal, siempre habrá mejores y peores resultados en proporción, y el resto de los sujetos quedará comprendido en el término medio, esto no significa que la mayoría de los alumnos están reprobados, sino señala el grado de dificultad de los reactivos.

Con esto se sugiere que no solo los alumnos son responsables de su formación, sino también es responsabilidad del sistema educativo et promover formas de evaluación más complejas.

Para validar la prueba se seleccionaron dos criterios: interés por la materia y promedio de calificaciones. En el primer rango se obtuvo un porcentaje positivo pero bajo (.09); en el segundo rango se obtuvo un porcentaje bueno (.36), de acuerdo a lo mencionado en el capítulo III.

Para intentar comprender esta situación se analizó el interés que los alumnos expresaron con base a tres niveles: mucho, regular y nada, al observar los resultados expresados en el gráfico 4 se puede decir que la mayoría de los alumnos manifiestan un interés regular por la materia de "Civismo fácil" o poco significativa (Quiroz, 1995), o bien sus intereses en ésta etapa (adolescencia) se concentran en otras actividades ajenas a los contenidos escolares.

En cuanto a la confiabilidad que fue medida por medio de los coeficientes de equivalencia y homogeneidad los resultados obtenidos son adecuados, ya que como se señala en el capítulo III los coeficientes de confiabilidad deben de ser mínimos de .75 o mayores, y observando el gráfico 6 se puede detectar que son mayores .



El proceso de validación es interminable, ya que éste proceso implica una valoración constante de los mejores reactivos adecuados a cada situación. La modificación curricular o la actualización de los libros de texto hacen necesaria una revisión de la prueba.

Finalmente, es importante recalcar que éste tipo de prueba no es apropiado para hacer una valoración individual, aunque se considera útil como un indicador más para valorar grupos, instituciones o regiones.



## BIBLIOGRAFÍA

**Ahmann, J.S. (1969).** Los tests psicológicos y su empleo en la escuela. 2ª. Edición. Edit. Troquel, Buenos Aires Argentina.

**Alonso, S. I. (1994).** Civismo 1 Edit. Kapeluz Mexicana, S.A. de C.V. 2da. Edición. México, D.F. Edit. Cincel, S.A.

**Ary Y Cheseer, (1990).** Introducción a la Investigación Pedagógica. México, D.F. Edit. Mc. Graw-Hill, 2a. Edición. pp. 178-201 y 308-331.

**Apple, M. W. (1987).** Educación y poder Edit. Paídos, Barcelona España.

**Ausubel. (1978).** Psicología Educativa : un punto de visa cognoscitivo. Edit. Trillas, México.

**Bachillerato Pedagógico no.2. (1989).** Antología de Psicología Educativa 1. México" D.F. pp. 30

**Barrera, V. F. (1974).** El buen ciudadano: primer curso. Edit. Porrúa, México. pp. 2-3.

**Bazinet, A. (1984).** La evaluación del rendimiento: métodos para la evaluación de los mandos intermedios en la empresa. Edit. Herder; Barcelona España.

**Bloom, B., Hasting, J. y Madaus, G. (1975).** Evaluación del aprendizaje. Tomo 1. Ediciones Troquel, Buenos Aires. pp. 209-288.

**Brown, F. (1991).** Principios de la medición en psicología y educación. El Manual Moderno, México.

**Casanova, M. A. (1992).** La evaluación garantía de calidad para el centro educativo. Zaragoza. Edit. Luis Vives. pp. 74-105.

**CEMEDIN, (1991).** La situación de la infancia en México. México: CEMEDIN.

**Chávez y Zamora, (Coomp), (1989)** Antología de evaluación circular. Cuadernos de planeación universitaria. Año 3 (2): México, UNAM. pp. 313-361.

**Cronbach L. J. (1982).** Desingning evaluations of educational and social programs. Jossey-Bass, San Francisco.

**Delval, J. (1989).** La representación infantil del mundo social. En: TI. Elliot, Y. Enesco y J. Linaza. El mundo social en la mente infantil. Alianza. Psicología. Madrid.

**Díaz Barriga, A. (Coord.) (1995).** La investigación educativa en los ochenta perspectivas para los noventa. Procesos curriculares, institucionales y organizacionales. Vol. 5. Edit. Consejo Mexicano de Investigación Educativa. 1a. edición. México, D.F. pp. 23-140.

**Equipo de Asesoramiento de Terrassa. (1989).** Pruebas Pedagógicas graduadas para preescolar y ciclo inicial. Visor, Madrid.

**Fernández, R. (1981).** Psicodiagnóstico. Concepto y Metodología. Edit. CINCEL. Bogotá.

**Forns, M. (1980).** Evaluación Psicológica Infantil. Edit. Barcanova. Barcelona.

**Fuentes Molinar, O. (1994).** La educación básica y los derechos humanos. En: Cero en Conducta, Año 9, No.36-37. Enero-Abril. México, D.F. pp. 4-7.

**García Ruíz, (1960).** La enseñanza de la geografía, la historia y el civismo. México, D.F .

**Gil Cantero. (1994).** Cómo educar en derechos humanos. En: Cero en Conducta, Año 9., No.36-37. Enero Abril. México, D.F. pp. 44-67.

**Gimeno Sacristán, J. (1989).** El Currículum: Una reflexión sobre la práctica. 2a. edición. Ediciones MORATA. Madrid España. pp. 373-403.

**Gimeno Sacristán y Pérez. (1991).** Comprender y transformar la enseñanza ediciones MORATA. Madrid España. pp. 364-397.

**González, R. M. (1994).** Tesis Adaptación del Test Diagnóstico en Habilidades Básicas TEHBA para Población Adulta Analfabeta. UPN. México, D.F.

**Gronlund, N. E. (1980).** Elaboración de tests de aprovechamiento. Edit. Trillas. México, D.F. 3a. reimpresión. Agosto. pp. 11-25 y 139-150.

**Guevara Niebla, G. C. (1991).** México: ¿un país de reprobados? .En: Nexos. No.62. pp. 33-44.(1992). (Comp.). La catástrofe silenciosa. Fondo de Cultura Económica. México, D.F.

**Gutiérrez, R. (1994).** Módulo II. El fracaso escolar: consecuencias psicosociales en los alumnos que repiten un año escolar en la educación básica. En: Antología Seminario de Investigación. UPN. México, D.F .

**Hernández, C., Ortíz, M. y Regalado, L. (1994).** Las pruebas Psicológicas. Su uso en México Tesis de Licenciatura. U. P. N., México.

**Hurtado de Mendoza, M. A. (1980).** Pruebas de rendimiento académico y objetivos de instrucción. 1a. edición. Edit. Diana. México, D.F. pp.18. 25-27 y 38-40.

**INEGI. (1994).** Estadísticas Históricas de México. INEGI, México, D.F. Tomo I.

**Inventario de Ejecuciones Académicas. (1994).** Fac. de Psicología, UNAM, México (en prensa).

**Karmel, L. J. (1975).** Psicología Educativa. Edit. Pax. (5a. edición). México, D.F. (1978). Medición y evaluación escolar. Edit. Trillas. México. D.F. 1a. reimpresión.

Keats, J. A. (1974). Introducción a la psicología cuantitativa Edit. Limusa. México.

**Latapí, P. (1994).** La investigación educativa en México. Fondo de Cultura Económica, México. Introducción. En: Jaime Torres Bodet, textos sobre educación. Edit. CNCA, México, D.F .

**López, E. O. (1990).** Menores en situación extraordinaria México: UNICEF.

**Manzanilla López, M. A, Rojas Nava, V. y Alvarez Orozco, R. (1994).** Civismo 1. El ciudadano moderno. Edit. Prentice-Hall Hispanoamericana S.A.

**Marchesi, A., Echeita, G., y Martín, E. (1993).** Desarrollo psicológico Y educación III Edit. Alianza. 6a. reimpresión. Madrid.

**Mehrens, W. y Lehman, Y. (1982).** Medición Evaluación en la Educación en Psicología. Edit. CECSA, 1a. edición. México, D.F. pp. 13-44, 139-146 y 418-475.

**Mejía, Z. R. (1976).** Moisés Sáenz. Educador de México. Cincuentenario de la fundación del Sistema Nacional de Escuelas Secundarias Mexicana (1926-1976). Federación Editorial Mexicana, México, D.F.

**Muñoz Izquierdo, C. (1983).** Presente futuro de la educación secundaria. SEP. Centro de Estudios Educativos. México. D.F .

**Nunnally, J. C. Y Berinstein, I. J. (1995).** Teoría Psicométrica. Edit. Mc. Graw Hill, México.

**O.C.D.E. (1991).** Escuelas y calidad de enseñanza. Edit. Paídos. Barcelona.

**Ornelas, (1995).** El sistema educativo mexicano. La transición de fin de año CIDE, NAFINSA, Fondo de Cultura Económica. México.

**Poder Ejecutivo Federal. (1989).** Programa para la Modernización Educativa 1989-1994. México. pp.38, 44 y 58.

**Plan Nacional de Educación. (1980).** Programa de Educación Básica. México.

**Programa Nacional de Acción. (1994).** Tercera evaluación. Comisión Nacional de Acción en favor de la infancia. México.

**Quintanilla y Galván, (1995).** Sujetos de la educación y formación docente. En: Rueda, M, (coord.). Procesos de enseñanza y aprendizaje 1. Consejo Mexicano de investigación Educativa-Universidad Veracruzana. México..

**Quiroz, R. (1994).** La práctica educativa. Secundaria obligatoria. reprobación y realidad escolar. En: Cero en Conducta. Año 9. No.36-37. Enero-Abril. México, D.F. pp. 91-99.  
(1995); Los cambios de 1993 en los planes y programas de estudio en la Educación Secundaria. Documento DIE 40; Centro de Investigaciones y de Estudios Avanzados del IPN, México, D.F.

**Ramírez, R. (1950).** La enseñanza del civismo. Vol. X. Biblioteca de Educación. México. D.F .

**Ramírez Hernández, G. (1994).** Los caminos de la educación en derechos humanos en México. En: Cero en conducta. Año 9. No.36-37 .Enero-Abril. México, D.F. pp. 8-18.

**Rodríguez, M. (1990).** Manual de creatividad. Edit. Trillas. México, D.F .

**Rodríguez Cruz, H. M. y García González, E. (1992).** Evaluación en el aula. Edit. Trillas. 3a. edición. 1a. reimpresión. México. D.F. pp. 18-20.

**Rosales López, C. (1990).** Evaluar es reflexionar sobre la enseñanza. Madrid; Edit Narcea. pp. 20-40.

**Sáenz, M. (1928).** Memoria de los trabajos realizados por el Departamento de Enseñanza Secundaria. (1924-1928). SEP. México.

**Sandoval, E. (1993).** Escuela Secundaria y Modernización Educativa. En: Cero en Conducta, Año 8, No.35. México.

**Schmelkes, S. (1995).** Evaluación y calidad educativa. Hacia una relación virtuosa. En: Diversidad en la educación. SEP. UPN. México.

**Secretaría de Educación Pública. (1983).** Bases instructivas para la evaluación del aprendizaje. SEP .

Dirección General de Educación Secundaria. México.

(1991). Estadística básica del sistema educativo nacional. SEP. México.

(1993). Plan y programa de estudio 1993.Educación Básica Secundaria. SEP. México.

**Severino, L. S. (1985).** Evaluación del rendimiento escolar. Pruebas Objetivas para el ciclo medio de EGB. Edit. Escuela Española, S.A. pp. 19-63.

**Scriven, M. (1967).** The methodology of evaluation. En: Perspectives on Curriculum Evaluation. AERA Monograph Series on Curriculum Evaluation, No.1 Ran Mc Nally, Chicago.



**Sheaffer, R., Mendenhall, W. y Ott, L. (1987).** Elementos de muestreo. Iberoamérica, México.

**Stephen. (1988).** Psicología en el aula. Edit. Trillas. México.

**Stufflebeam, D. y Shinkfield, A. (1987).** Evaluación Sistemática. Edit. Paídos. Barcelona. pp. 91-109.

**Taboada, E. y Valentín. (1995).** La investigación educativa en los ochentas perspectivas para los noventa. Ciencias Sociales. Vol. II. Edit. Consejo Mexicano de Investigación Educativa. 1a. edición. pp. 121-207.

**Tirado, F. T. (1986).** La crítica situación de la educación básica en México. En: Ciencia y Desarrollo. CONACYT , México. Vol. XII. No.71. Noviembre-Diciembre. pp. 81-94. .

**Tyler, R. W. (1973).** Principio Básicos del Currículo. Edit. Troquel, Buenos Aires. Argentina.