



UNIVERSIDAD NACIONAL AUTÓNOMA
DE MÉXICO

FACULTAD DE PSICOLOGÍA

PROPUESTA DE MEJORA DEL EXAMEN DE
ADMISIÓN DE LA MAESTRÍA EN CIENCIAS CON
ESPECIALIDAD EN ADMINISTRACIÓN
DE NEGOCIOS DE LA ESCA,
UNIDAD SANTO TOMÁS, IPN.

INFORME DE PRÁCTICAS
QUE PARA OBTENER EL TÍTULO DE
LICENCIADA EN PSICOLOGÍA
P R E S E N T A:
ANAHID TENORIO HERRERA

ASESORA: DRA. MARÍA DEL CARMEN GERARDO PÉREZ.



FACULTAD
DE PSICOLOGÍA

CIUDAD UNIVERSITARIA, JULIO DEL 2008.



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

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

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



UNAM.188

2008

ej. 2

M.

TPs.

Agradecimientos

Un día de repente desperté, y empecé a conocer lo malo y gracias a ello, lo bueno; en mi camino hubo tropiezos y fallas, sin embargo, aprendía de ello y seguía adelante, día a día hago mi mejor esfuerzo, y éste ha sido un camino que sin ellos hubiese sido difícil y sin su apoyo aún más. Hoy culminó un año importante que habla del trabajo y dedicación que puse en ello, no obstante esto no hubiese sido posible sin la ayuda y apoyo de dos seres que amo, y a los que ni el más grande tesoro recompensaría lo que me han dado, y lo que han logrado que sea.

Este agradecimiento es a ellos, esos ángeles que fueron elegidos para tenerme y cuidarme, son mis padres y para ustedes es a quien dedico principalmente este trabajo.

*¡Gracias papá, gracias mamá!
¡Mil gracias por todo!*

A mis hermanos Josfi y Patilú

Hoy, este se convierte en un logro que quiero compartir con ustedes, para recordarles que cuando te propones algo y te aferras a lograrlo, lo alcanzas, lo cual es a base de trabajo, dedicación y motivación. Ustedes son una parte muy importante y a pesar de los problemas siempre habrá una esperanza. Gracias por su cariño y por las peleas, pues es donde nos damos cuenta de lo mucho que nos queremos, así que nunca se detengan, siempre luchen por lo que quieren y recuerden siempre que:

¡Los quiero mucho!

A mis chiquillas...

Ustedes más que nadie saben que la carrera en psicología ha sido la mejor que pudimos tener, la experiencia más gratificante, que nos sirvió para conocernos y para dar la oportunidad a otros de entrar en nuestra vida. Nunca me arrepentiré de estar en el lugar y carrera indicada, es decir en la UNAM, puesto que conocí a personitas que se volvieron más que mis amigas, gracias chicas, que de ustedes tome el ejemplo y seguí hasta terminar el trabajo; gracias por todos los momentos que pasamos juntas, por las fiestas, los días de examen y por tantos momentos de felicidad, pero sobre todo, mil gracias por abrir su corazón y dejarme entrar en él.

¡Las quiero!

A Robert

A esa personita que a pesar de ya no estar tanto tiempo juntos, sigue siendo una parte importante para mí, a mi amigo y casi hermano (nada más por el apellido y papás); tú siempre me apoyaste y en los momentos difíciles has estado conmigo, nunca sabré como agradecerte todo en lo que me has apoyado, y este espacio es una manera de hacerlo. No olvides que el tiempo o la distancia nunca serán un impedimento u obstáculo para romper la amistad que nos une, te quiero amiguito y de verdad: ¡Gracias por ser mi amigo!

A ti amorcito (Fer)

Pensaste que te olvidaba?, claro que no, tú siempre has estado a mi lado, aguantando todo, demostrándome tantas veces tu amor y ofreciéndome siempre tu apoyo, quizá nunca termine de agradecerte todas las enseñanzas que he obtenido de ti y que gracias a ti hoy culmino con este proyecto.
Te amo, y se que siempre estarás conmigo para apoyarme y si fuese necesario para hacerme ver mis errores, pues también te has convertido en mi amigo y compañero.
Gracias por todo cariño, pues sabes que eres parte importantísima para mí y ni todas las palabras, alcanzarán para demostrarte mi amor y mi agradecimiento.

¡Te amo!

A mis profesores, sinodales y quienes fueron parte importante en esto

Dra Carmén Gerardo

Debo decir muchas gracias, por su apoyo y su enseñanza y principalmente por no haberme dejado en el camino, usted supo lo difícil que fue el proceso, pero, al fin, este proyecto quedó terminado.
Muchas gracias doctora, pues a la vez fue una fuente de motivación para lograr culminar con este trabajo, y de verdad:

¡Gracias maestra!

Dr. Humberto Ponce

A pesar de no haber podido terminar conjuntamente el proyecto, debo agradecerle por el apoyo que siempre tuve de usted, por alentarme a iniciarlo, por las facilidades que me aportó para la obtención de información o documentos, pero lo más importante por ser un buen profesor, dado que el trayecto en el que tuve la oportunidad de trabajar y aprender de usted, fue muy gratificante y eso, merece un reconocimiento.

¡Gracias doctor!

A mis sinodales

Lic. Alejandra García, Lic. Raúl Tenorio, Mtra. Mirna Valle y Lic. Ricardo Lozada

Es indiscutible lo importante que fueron en este trabajo, gracias por todas sus observaciones y por el tiempo que emplearon en enseñarme cosas que quizá desconocía, todos y cada uno de los elementos que me hacían ver, que fueron de gran importancia para mí, y de verdad no me queda más que agradecerles por ayudarme y a que este trabajo fuera mejorando día a día, y no me queda mas que decir. ¡Gracias a todos y cada uno de ustedes!

A la Lic. Karen Amezcua

*Esta es una manera de agradecerte todo lo que hiciste por mí y por todo el apoyo y facilidades que me proporcionaste, debo decirlo y quizá sino hubiese sido por tú ayuda, éste trabajo continuaría detenido, de verdad y sinceramente ¡Gracias!, ya que en verdad fuiste una parte importante en este proyecto y nunca olvidaré lo que hiciste por mí, pues a pesar de no ser parte directamente de mi jurado, te siento como tal.
Esta es una forma de agradecerte por todo.*

Sinceramente ¡mil gracias!

Mi último agradecimiento es para quién me dio la oportunidad de estar aquí, y me refiero a la máxima casa de estudios, ésta es la Universidad Nacional Autónoma de México, pues ahora puedo decir que estoy orgullosa de haber estudiado en esta institución y de ser parte de ella, porque los mejores aprendizajes los obtuve estando aquí y puedo decir que me siento.

*¡Orgullosamente universitaria!
¡Orgullosamente puma!*

ÍNDICE

INTRODUCCIÓN	1
--------------------	---

PRIMERA PARTE.- ANTECEDENTES Y CONTEXTO

I) ESBOZO HISTÓRICO

1.1 Orígenes y estructura de la Maestría en Ciencias con especialidad en Administración de Negocios (MAN)	5
1.2 Descripción y tipificación de la maestría.....	7
1.3 Proceso de admisión al posgrado del IPN	9
1.3.1 Base reglamentaria del proceso	9
1.3.2 El examen de conocimientos	11

SEGUNDA PARTE.- ANTECEDENTE TEÓRICO Y CONCEPTUAL

II) EVALUACIÓN EN LA EDUCACIÓN

2.1 Evaluación	14
2.1.1 Concepto de evaluación	15
2.1.2 Funciones de la evaluación	16
2.1.3 Tipos de evaluación	17
2.1.3.1 Evaluación diagnóstica	18
2.1.3.2 Evaluación formativa	19
2.1.3.3 Evaluación sumativa	20
2.1.4 Relevancia en el proceso de admisión en instituciones educativas	21
2.2 Desarrollo de instrumentos de medición en la educación	22
2.2.1 La medición en la selección de alumnos	24

III) TIPOS DE PRUEBAS PARA MEDIR CONOCIMIENTOS

3.1 Definición de conceptos	26
3.2 Principales características	27
3.3 Pruebas para evaluar conocimientos	27

3.3.1	Tests de ensayo	28
3.3.1.1	Preguntas de respuesta restringida	29
3.3.1.2	Preguntas de respuesta extensiva	30
3.3.2	Tests objetivos.....	30
3.3.2.1	Tipos de reactivos	32
3.4	Ventajas y desventajas de los tests objetivos contra los de ensayo	40

TERCERA PARTE.- ASPECTOS METODÓLOGICOS Y EXPERIENCIAS

IV) ESTRATEGIA PARA LA ELABORACIÓN DE LA PRUEBA

4.1	Planeación de la prueba	43
4.1.1	Definición del constructo	45
4.1.2	Tabla de especificaciones	47
4.1.3	Preparación de los reactivos	52
4.1.4	Ensayo de los reactivos	57
4.1.4.1	Teoría clásica de la prueba	58
4.1.4.2	En cuanto a la Teoría de respuesta al ítem	68
4.1.5	Selección de los reactivos para la prueba final	81
4.1.6	Confiabilidad	84
4.1.6.1	Métodos para estimar la confiabilidad	84
4.1.7	Validación	88
4.1.7.1	Validez de contenido	90
4.1.7.2	Validez con el criterio	91
4.1.7.3	Validación de constructo	92
4.1.8	Estandarización de la prueba	93
4.1.9	Preparación final de la prueba	96

V) EXPERIENCIAS SIMILARES

5.1	Políticas, criterios de admisión y requisitos	98
5.2	Indicadores de éxito académico	99
5.3	Nivel predictivo del examen de admisión	101

5.4 Alternativas respecto a la validación de los exámenes	102
MAPA CONCEPTUAL	105

CUARTA PARTE.- ANALISIS DE RESULTADOS Y PROPUESTA DE MEJORA

VI) ANALISIS DE RESULTADOS	
6.1 Resultados cuantitativos	106
6.2 Resultados cualitativos	110
VII) PROPUESTA DE MEJORA DEL EXAMEN DE ADMISIÓN A LA MAN	
7.1 Exposición de motivos	115
7.2 Objetivo.....	116
7.3 Metas.....	116
7.4 Planeación.....	117
7.5 Consideraciones finales.....	125
VIII) CONCLUSIONES	
8.1 Conclusiones.....	127
8.2 Sugerencias y limitaciones	130
REFERENCIAS	132
ANEXOS.....	137

INTRODUCCIÓN

En el contexto de la Educación Superior (ES), preocupa preservar la calidad educativa en las universidades, que lleva a realizar procesos de admisión en los estudiantes para responder a tal factor, por medio de una mejor preparación en los mismos y un refinamiento en el proceso de selección, a través de la creación de cursos a nivel maestría y doctorado para cubrir la demanda de especialistas de alto nivel.

Venegas (1993), indica que existen diferentes políticas de admisión, dentro de las cuales una de ellas es el examen de admisión, que dentro de este trabajo cobra relevancia, puesto que a partir de un instrumento objetivo de evaluación se puede seleccionar al candidato idóneo a una institución, de lo contrario provocaría una exclusión injustificada de los aspirantes. Las políticas de admisión a nivel superior se guían en seleccionar a los aspirantes que muestren mayor probabilidad de éxito académico, pretendiendo con ello, una menor tasa de deserción escolar.

De manera tal que las pruebas para medir el aprovechamiento escolar son una de las formas que más destaca en la medición por considerarse un elemento "realista" del rendimiento escolar e idóneo para la evaluación de los resultados de aprendizaje.

Carreño (1993) menciona que las pruebas de aprovechamiento ofrecen ciertas ventajas que en conjunto no poseen los demás instrumentos de medición, debido a que pueden ser aplicadas en el momento adecuado, en grandes grupos, etc., a lo que agregaría que con la planeación del instrumento se puede obtener así, su confiabilidad y validez, que son elementos indispensables para su calibración.

Sin embargo hacer buenas pruebas no es fácil y se debe responder a la necesidad de que se construya un instrumento de medición que cumpla con los requisitos técnicos elementales para asegurar su funcionalidad y no sólo hacer una prueba sin una previa planeación, más aún si los fines exigen la toma de decisiones, tal como lo es la selección de alumnos que desean ingresar a un posgrado.

Siguiendo con este tema, en la ES, se requieren personas que cumplan con el perfil de ingreso que pide la institución, debido a que en México se ha tratado de asegurar la calidad de la misma, tanto en políticas públicas, como en la investigación, donde la evaluación ha constituido un elemento para asegurar ésta calidad,

redefiniendo las exigencias mínimas de la profesión, y justificando los sesgos en los ingresos y la trayectoria a un mercado académico al que no se accedía (Didou, 2005).

De manera que se debe contar con exámenes generales que midan y evalúen conocimientos acorde a la formación académica y profesional, para que a través de sus egresados, se contribuya a la elevación de la calidad y a la determinación de acciones eficaces para mejorarla (Centro Nacional de Evaluación para la Educación Superior [CENEVAL], 2001).

Lo antes mencionado coincide con lo reportado en la evaluación del posgrado en la UNAM (1996) en el que se menciona que para poder acelerar el desarrollo científico y tecnológico del país y reducir el retraso tecnológico con respecto a países desarrollados, se hace necesario el contar con profesionales, altamente calificados, por medio de la modernización de programas de posgrado y el fomento de políticas de reclutamiento de estudiantes de alta calidad para el mismo.

Dado lo cuál el proceso de admisión debe enfatizar en su selección, primordialmente en los exámenes, puesto que a través de ellos se identifica el candidato idóneo a ocupar la vacante que se presenta; el proceso de selección debe ser planeado y a la vez, corresponder a un manejo adecuado de las pruebas empleadas para tal proceso, donde las más utilizadas, dadas las ventajas que presentan, son las pruebas objetivas, a lo que Carreño (1993) menciona que la utilización de pruebas de admisión a las instituciones educativas, es uno de los temas en los que se justifica su elaboración para lograr un proceso refinado a través de exactitud en las mediciones y justicia en las comparaciones.

Se han realizado una serie de estudios por diferentes autores (Venegas, 1993; Alonso, 1998; Chávez, 2000) que investigaron la validez de instrumentos educativos y encontraron que la mayoría no investigan la validez de las pruebas empleadas o que sus coeficientes de validez resultan insignificantes, mostrando limitaciones, puesto que carecen de confiabilidad y validez.

De tal manera, que en este estudio se hace un análisis detallado del examen que se aplicó en la Maestría en Ciencias con especialidad en Administración de Negocios (MAN), con el objeto principal de presentar los pormenores de su elaboración y de ésta manera presentar sustentos teóricos válidos del proceso adecuado que debe seguirse

en la elaboración de una prueba de rendimiento objetiva, puesto que como se ha explicado, las pruebas de aprovechamiento bien diseñadas pueden verse como un paso importante para responder a las exigencias mínimas de profesionistas de alto nivel que demuestren la calidad de la educación en el mercado laboral.

Por ello, resulta pertinente evaluar en primera instancia el examen de conocimientos utilizado, antes de realizar una estructuración del mismo, siendo importante realizar este análisis con la finalidad de mejorarlo y establecer propuestas que optimicen el proceso, y que permita predecir el éxito académico del alumnado.

Cabe resaltar que para elaborar propuestas, es necesario evaluar los procedimientos utilizados en la actualidad, lo que permitirá discriminar su utilidad en función del objetivo principal que tienen.

Este estudio, busca como principal alcance la elaboración de una propuesta, estructurada de manera que beneficie tanto a la sección de la MAN, como a los aspirantes que deseen el ingreso a la institución, no en el hecho de la exclusión de los mismos, debido a que es un asunto inevitable de hacer, pues existe un reglamento de admisión en donde sólo se tiene un cierto número de lugares que pueden ser cubiertos; sino más bien en utilizar una prueba válida que auxilie en el proceso de selección y que sirva como un paso para estudios futuros para poder así, predecir el éxito académico y elevar la calidad académica de la institución.

La presente investigación consta de cuatro partes:

En la primera parte, se presentan los antecedentes y el contexto, que corresponde a la situación institucional en que se desarrolló la práctica y una descripción más profunda del problema que se aborda, que es la fase del examen que se aplica para la selección de alumnos a la MAN.

En la segunda parte se continúa con los antecedentes teóricos, respecto al tema de evaluación en la educación y los tipos de pruebas que se emplean para medir conocimientos.

En la tercera parte, se presentan los aspectos metodológicos y las experiencias similares, abarcando por un lado estrategias en la elaboración de una prueba y por la otra las investigaciones realizadas por diferentes autores acerca del tema que se aborda en el estudio.

La cuarta parte, presenta el análisis de resultados del procesamiento de los datos del estudio, las conclusiones obtenidas y se culmina con la propuesta que se expone.

Finalmente se encuentran las referencias bibliográficas y se finaliza con el apartado de anexos.

PRIMERA PARTE.- ANTECEDENTES Y CONTEXTO

I) ESBOZO HISTÓRICO

1.1 Orígenes y estructura de la Maestría en ciencias con especialidad en Administración de Negocios (MAN)

En 1970, la Universidad Nacional Autónoma de México y el Instituto Politécnico Nacional (IPN) acentuaron el crecimiento de su población escolar, que llevó a establecer instituciones de ES, para dar solución a la problemática, a lo cuál la Asociación Nacional de Universidades e Institutos de Educación Superior (ANUIES), a través del Poder Ejecutivo de la Nación, realizó un estudio para dar solución a ello, que dio como conclusión que en 1973, se crearan nuevas instituciones que solventaran tal demanda, entre ellas el Colegio de Bachilleres y la Universidad Autónoma Metropolitana, aunado a un refuerzo en las Universidades existentes en las entidades federativas (Pardo, 1980).

Esta necesidad abarcaba a nivel licenciatura, como posgrados que cubrieran los requerimientos que se esperaban para satisfacer la demanda del campo laboral.

Haciendo que en 1936 se funde el Instituto Politécnico Nacional por el Gral. Lázaro Cárdenas, que agrupó a 17 escuelas pre-vocacionales (nivel medio), 4 vocacionales (medio superior) y 7 de nivel superior, además de incorporar a 7 escuelas de nivel medio superior (CECYTS) y 19 de estudios superiores.

Los estudios de posgrado en el IPN, se inician en 1961 con la Escuela de Post-Graduados anexa a la Escuela Superior de Ingeniería Mecánica y Eléctrica. El desarrollo de programas de estudios de posgrado surgió en las áreas de ingeniería, ciencias exactas, y médico-biológicas con personal de alto nivel. En 1963, se dan marcos normativos y organizativos, para establecer procedimientos para llegar a ser Maestro y Doctor; integrándose en 1995 el Consejo Consultivo de Graduados y elaborándose el primer Reglamento de Posgrado.

De las primeras escuelas del IPN con posgrado estaría la Escuela Superior de Comercio y Administración con su Maestría y Doctorado en Ciencias Administrativas,

fundados en 1962. En la actualidad el IPN tiene 110 programas de posgrado, 20 doctorados, 56 maestrías y 25 especializaciones, impartidas en 14 Secciones de Posgrado, 3 Centros de Investigación y 3 Proyectos Institucionales.

Escuela Superior de Comercio y Administración (ESCA), Unidad Sto. Tomás

Se funda en 1845 con el nombre de Escuela Nacional de Comercio que por decreto presidencial, el 2 de diciembre de 1867 cambió su denominación a Escuela Superior de Comercio, lo cual para 1869, llevo a que la Cámara de Diputados expidiera un decreto para reformar la Escuela Superior de Comercio para ser Escuela de Administración, tomando así su actual nombre.

Como parte de las 19 escuelas de educación superior del IPN, la ESCA, Sto. Tomás ofrece 4 licenciaturas, y tres maestrías en ciencias con especialidad en Administración de: Negocios (MAN) y Pública (MAP), y la Maestría en Desarrollo de la Educación (MADE), que se inició en 1975; así como el Doctorado en Ciencias con especialidad en Ciencias Administrativas (DCA). El incremento constante de la población estudiantil dio origen a que en 1974 se creara un nuevo plantel de la ESCA: la Unidad Tepepan (www.escasto.ipn.mx).

Sección de Estudios de Posgrado e Investigación (SEPI)

Se forma en 1961, con cursos en América Latina de Maestría y Doctorado en Ciencias Administrativas, iniciando con 42 alumnos inscritos. El campo de acción de los posgrados se ubicaba en planear y crear estructuras administrativas en general; modificar o corregir las deficiencias administrativas de los negocios y el Estado, dirigir permanentemente empresas y fungir como funcionario del Estado, investigador científico y profesor de alto nivel

En 1967 los programas de estudio se modifican debido a las necesidades administrativas, con planes que permitieran el desarrollo de habilidades para las tareas de investigación (Instituto Politécnico Nacional, Escuela Superior de Comercio y Administración, 1995).

Como resumen se puede destacar que los estudios de posgrado en la ESCA-IPN buscan responder a necesidades del desarrollo social a través de problemas de interés nacional, para establecer normas modernas para la planeación, coordinación, organización, operación y desarrollo de tales estudios, brindando mayor participación en diversos sectores de la comunidad politécnica en el cumplimiento de sus finalidades (Instituto Politécnico Nacional [IPN], 2006).

1.2 Descripción y tipificación de la maestría

Como se mencionó en 1974 la MAN toma su actual nombre, con labores en la ESCA, IPN, plantel Santo Tomás, ubicado en Prolongación de Carpio No. 471 Esq. Plan de Agua Prieta, Casco de Santo Tomás; C. P. 11340. La coordinación de MAN se ubica en el edificio "H", en el segundo piso, la MAN se orienta a la formación de especialistas capaces de ubicar y resolver problemas administrativos de la empresa. En 1967 su plan de estudios fue modificado, con cambios sustanciales hasta 1992, cuando se empezó a estudiar su estructura curricular a fin de que sus egresados respondieran eficazmente al manejo de la compleja problemática que caracterizaba al país. Este plan de estudios se reestructuró en 1994 (Instituto Politécnico Nacional, Escuela Superior de Comercio y Administración, 1995).

La currícula de la Maestría se integra por tres módulos: homogeneización, fundamentación y especialización. El primero pretende que el estudiante cubra áreas débiles según el examen diagnóstico. La finalidad de la fundamentación es que el estudiante sea un verdadero administrador, mediante programas de estudio, apoyados por el área cuantitativa e investigación. El módulo de especialización comprende materias optativas, para orientar al alumno a campos de la administración y que, respondan a la demanda de la sociedad; está integrado por materias optativas que el alumno elige, y la guía de un consejero de estudios (www.escasto.ipn.mx).

Cabe resaltar que el programa de la MAN es el de mayor matrícula escolar de las tres maestrías; en el semestre 1-2004 la sección contaba con 453 alumnos de posgrado, de los cuales el 87% (397) correspondía a alumnos del programa MAN

(información del Departamento Control Escolar, SEPI), que hace que interesados en la maestría cumplan con el perfil de ingreso que se solicita (ver p.135).

Las materias que cubre el plan de homogeneización, son las correspondientes a las evaluadas en el examen de selección: el cuál abarca matemáticas aplicadas a la administración, informática administrativa, teoría de la administración, análisis económico financiero de la gestión empresarial e investigación. El modulo de fundamentación abarca cuatro semestres en los que se imparten las materias que ayudarán al alumno en su formación profesional y finalmente en la especialización, se tienen las optativas, en donde se encuentran materias como promoción y administración de pequeñas y medianas empresas, administración de la calidad total, consultoría de empresas, desarrollo organizacional, entre otras.

Anualmente cubre dos períodos de inscripción, el primero en el mes de Octubre para el período de Enero – Junio y el segundo en el mes de Abril para el semestre de Agosto – Diciembre.

En cada carrera de licenciatura, maestrías y doctorado hay una academia de profesores que tratan asuntos relacionados con el proceso de enseñanza- aprendizaje de los alumnos, para cada uno de los programas. En el caso de la coordinación MAN, las principales funciones son las de participar e informar a la academia de profesores de sus reuniones; de asesorar a los alumnos en el proceso de permanencia en la maestría, así como brindarles apoyo para la obtención del grado; y programar la estructura académica para cada semestre.

De manera tal que la finalidad de los estudios de posgrado en la ESCA- IPN buscan responder a las necesidades de desarrollo social con la finalidad de formar:

- ❖ Investigadores capaces de trabajar inter y multidisciplinariamente en la solución de problemas administrativos.
- ❖ Personal docente competente que facilite los procesos de aprendizaje de la disciplina.
- ❖ Especialistas competentes que dispongan de habilidades prácticas necesarias y acordes a su área de especialidad.

1.3 Proceso de admisión al posgrado del IPN

Se empezará por hablar de las pruebas, dado que serán efectivas en tanto den información necesaria y que midan para los fines a los que se dirige. Karmel (1978), indica que el programa de exámenes debería orientarse a las necesidades de la escuela, y que su utilidad es brindar al profesor, estudiante, padres y a la escuela, información significativa; ya que si se administran sólo porque es una práctica que se realiza, o porque se den los fondos para efectuarlo, se convertirá en un ejercicio inútil.

De manera que realizar una planeación ayude a:

- a) Obtener información sobre los conocimientos de la población que desee el ingreso a la institución, mismos que se encuentren relacionados con las necesidades académicas y profesionales.
- b) Poder comparar los conocimientos que se presenten en alumnos de diferentes carreras e incluso instituciones.

1.3.1 Base reglamentaria del proceso

Es de conocimiento de la comunidad, el uso que se ha hecho de las pruebas, ya sea para la medición de algún aspecto de la personalidad, para selección, o para otro tipo de evaluación, en el que para realizarlo se recurre al método de aplicación de pruebas, que en el área educativa se emplean para medir el rendimiento escolar, donde los más empleados son los test objetivos, de los que se hablará en la segunda parte.

Fermín (1971, en Venegas, 1993; p.30) describe que “los exámenes son medios pedagógicos para estimular la actividad intelectual de los estudiantes y corregir periódicamente los posibles defectos de su formación. Como instrumentos auxiliares de evaluación, en ellos debe atenderse, más que a la repetición o memorización de las materias tratadas durante el curso, al aprovechamiento que demuestre el alumno mediante la comprensión del saber recibido”.

Pardo (1980), menciona que desde 1980 comenzaron a desarrollarse y generalizarse exámenes calificados de manera objetiva para aplicarse uniformemente a las respuestas dadas por los alumnos, que llevó a que en 1915 aparecieran las pruebas

estandarizadas de aprovechamiento tal como el Test de Aritmética de Stone y el de Ortografía de Buckingham. En 1930, éste tipo de pruebas se ideó para medir destrezas escolares y para áreas de contenido de los programas escolares. De manera que surgió el uso de pruebas en el área educativa, con el objetivo de evaluar al individuo (alumno), con el fin de diagnóstico, que se refiere al conocimiento que tenían en algún área, antes del curso; con fines de discriminación, en donde se les clasifica de acuerdo al nivel de competencia alcanzado y su fin es la promoción; y de pronóstico, que tiene como fin el de hacer predicciones sobre el alumno y orientación para los mismos.

Una vez realizado este esbozo histórico, en lo que refiere a la utilización de exámenes, se aludirá a aquellos aspectos del reglamento que para el fin del trabajo son importantes, llevando a describir que el programa de estudios de posgrado, es “el programa que opera de acuerdo al diseño curricular establecido en su plan de estudios con el propósito de formar especialistas, maestros y doctores, aprobado por el Consejo General Consultivo del Instituto” (IPN, 2006., p. 31).

El proceso para seleccionar a los aspirantes a la Maestría en Ciencias con Especialidad en Administración de Negocios es el siguiente:

- ❖ Exámenes: se evalúan cinco materias: matemáticas aplicadas a la administración, informática administrativa, teoría de la administración, análisis económico financiero de la gestión empresarial y habilidades de investigación.
- ❖ Guías: Se elabora una para cada una de las materias, y se publican en Internet.
- ❖ Los aspirantes presentan el examen de conocimientos y al concluir se les pide que evalúen el proceso de admisión, a través de un cuestionario escrito.
- ❖ Examen de Inglés: Se solicita al Centro de Estudios Nacionales de Lenguas Extranjeras – Zacatenco los resultados obtenidos por los aspirantes.
- ❖ Lista de preceptados: Se evalúan los resultados de los 4 exámenes aplicados por MAN, además del inglés para obtener la lista de preceptados que continúan el proceso. Esta lista se publica en Control Escolar.
- ❖ Entrevista con investigadores: A los preceptados se les asigna un entrevistador, así como el día y a la hora de la misma.

◀ Lista de aceptados: Se elabora una lista con el nombre de los aspirantes seleccionados para cursar la Maestría, misma que se publica en Control Escolar, en donde se solicita la documentación para la inscripción formal, los resultados también pueden consultarse en Internet.

Existen otros instrumentos de evaluación en la educación, como son los registros observacionales y las entrevistas, sin embargo dado la naturaleza de este estudio, se da una concentración especial sólo en las pruebas o test, que si bien no suplen a éstas, son consideradas el método más adecuado y de mayor facilidad para medir el grado de conocimientos de los alumnos.

1.3.2 El examen de conocimientos

Los aspirantes a la MAN, deberán reunir como mínimo los siguientes requerimientos, tener licenciatura en áreas afines a las ciencias sociales y/o físico-matemáticas y de preferencia con experiencia laboral en aspectos de la administración de empresas y "contar con las actitudes y valores necesarios para responsabilizarse de su proceso formativo y asumir una posición activa frente al estudio y el desarrollo de los proyectos y trabajos requeridos, coincidentes con el ideario y principios del IPN" (www.sepi.escasto.ipn.mx).

Otro punto a considerar es que el examen que elaboren para el ingreso de los aspirantes a la maestría debe realizarse en base al perfil y conocimientos de cada una de las materias afines a las ciencias sociales, mismos que deben reflejarse en preguntas acorde a cada área de conocimiento de la matrícula estudiantil.

Las materias que se evaluaron en el examen de selección del período B- 2006 (año en el que se intervino) fueron cinco.

Tabla 1

Descripción de las cinco materias contenidas en el examen B- 2006 de la MAN.

Materia	Tipo de formato	Total de reactivos	Números de reactivos	Descripción de las preguntas		
				Presentan respuestas como "Ninguna de las anteriores", "sólo a y b" o "Todas las anteriores"	La redacción da pie al descarte de opciones	Pide definición de términos
Análisis Económico Financiero	Opción múltiple	18	1 -9, 11-14,16-20	√	√	
	Ver/Falso	2	10 y 15			
	Completa.	1	21			
Teoría de la Admón.	Relación de columnas	2	1 y 2			√
	Op. múltiple	3	4-6	√		
	Ver/Falso	1	3			√
Informática	Relación de columnas	1	8		√	√
	Opción múltiple	7	1 - 7	√	√	√
Habilidad de Investig.	Opción múltiple	4	1 - 4			
	Ver/Falso	5	5 - 9			
Álgebra	Op. múltiple	12	5, 7-10,12,13 y 36-40			
	Ver/Falso	14	15-19,24-27, 31-35			
	Completa.	14	1-4,6,11,14,20-23 y 28 - 30			

El examen, no tiene una planeación adecuada ni la asesoría de un experto en elaboración de pruebas, por lo que difícilmente contará con indicios de confiabilidad y validez para el empleo subsiguiente del instrumento.

Por lo tanto, se explicará el problema en el que se trabajó, y en el siguiente apartado se presentarán los lineamientos teóricos para la elaboración de una prueba de

aprovechamiento objetiva, pretendiendo que se tenga una mejor planeación y elaboración de los exámenes de admisión que se realizan a la MAN y que se pueda demostrar su confiabilidad y validez, de manera que se posea una prueba valida para seleccionar a los alumnos que deseen ingresar a dicha maestría.

El motivo por el que la prueba debe dar información válida, es por el valor que tiene dentro del proceso, dado que conforme a los resultados del examen, es que se decide los aspirantes que serán pre- aceptados y las materias de homogeneización que deben cursar, siendo de utilidad para la sección el elaborar una adecuada prueba de conocimientos para poder hablar así de un instrumento de diagnóstico, que ayude a detectar deficiencias académicas que pudieran dificultar el aprendizaje.

Con el trabajo expuesto se presente dar un conocimiento amplio de lo que un psicólogo puede realizar, debido a que la mayoría de las veces se tiene una noción general de terapeuta y se olvida o desconoce de las técnicas y herramientas con las que se cuenta y debido que no se tiene un psicólogo en la sección, su ayuda sería de utilidad para éste y otros asuntos relacionados con el mejoramiento de la calidad en la evaluación.

SEGUNDA PARTE.- ANTECEDENTE TEÓRICO Y CONCEPTUAL

II) EVALUACIÓN EN LA EDUCACIÓN

2.1 Evaluación

El año de 1990 se considera la década de la evaluación y aseguramiento de calidad en Educación Superior (ES), tanto en las políticas públicas como en la investigación; haciendo que la evaluación impulse el reordenamiento de éste sistema, sirviendo para exigir mayor responsabilidad a las instituciones y a los individuos, constituyendo en el país uno de los sistemas más complejos y diferenciados en cuanto al aseguramiento de la calidad, así como abarcando a individuos, programas, procesos de investigación, docencia e incluso administración (Didou, 2005).

La evaluación ha estado presente en los propósitos y actividades de las instituciones de educación superior, así como en la Asociación Nacional de Universidades e Institutos de Educación Superior, desde su formación en 1950, siendo hasta 1983 cuando aparecen políticas gubernamentales, para elevar la calidad de la educación superior (CENEVAL, 2001).

En 1984, la evaluación arrancó con la inauguración del Sistema Nacional de Investigadores (SNI), con la finalidad de allegar investigadores de tiempo completo, que requirió un sistema de evaluación basado en mediciones, lo que contribuyó a redefinir las exigencias de la profesión, justificando sesgos en ingresos y trayectorias y otorgando legitimidad a sus integrantes (Didou, 2005).

El reconocimiento de deficiencias académicas, ha llevado a realizar evaluaciones de los aspirantes a ingresar a escuelas profesionales, para su selección y clasificación. Aguilar, Lizárraga, & Pérez, (1981) mencionan que dichas evaluaciones abarcan por un lado el área psicométrica (aptitudes intelectuales, intereses vocacionales y rasgos de personalidad), y por el otro el área académica, que comprende áreas curriculares, pero que fuera de esta inclusión en la mayoría de los casos las investigaciones han revelado

que los coeficientes de validez no son significativamente mayores que cero, haciendo que los resultados sean mejores al emplear pruebas de conocimientos. En la siguiente sección se mencionan conceptos de evaluación y el papel que tiene dentro de lo que se denomina aprendizaje.

2.1.1 Concepto de evaluación.

La evaluación educativa “es un proceso integral, sistemático y continuo que valora los cambios en la conducta de los estudiantes” (Alonso, 1998).

Existen otras definiciones según Bloom (1977):

- a) Un método para adquirir y procesar evidencia para mejorar el aprendizaje del estudiante y la enseñanza.
- b) Una ayuda para aclarar metas y objetivos importantes de la educación y como proceso para determinar la evolución de los estudiantes.
- c) Un sistema de control de calidad que permite que en cada etapa se determina la eficacia del proceso de enseñanza- aprendizaje, y si es o no eficaz, los cambios que deben hacerse para lograr la eficacia.
- d) Un instrumento de la práctica educativa que permite establecer si determinados procedimientos alternativos son igualmente eficaces o no para alcanzar las metas educativas.

Debe tenerse presente además el hecho de que la evaluación debe ser controlada, en base a la relación entre una norma, un modelo educativo y los fenómenos a comparar, visto como un proceso interminable, que estará en continuo cambio (Amezcuza, 2004). Esto es razonable debido a la sociedad en la que se vive, en un ambiente en el cuál el sistema se encuentra en continuo cambio y que debe adaptarse a las nuevas situaciones, aún más dependiendo de las modificaciones e innovaciones educativas que día a día acontecen.

Se puede dar la tendencia a ver el proceso de evaluación, como valoración de los fines a los que se destina el procedimiento, tal como si el hecho de contar con buenos insumos garantice la calidad y de esta manera el que a través de la evaluación se garantizará el cumplimiento de los propósitos, sin considerar los elementos centrales

que influyen en dicha evaluación; Abreu (2003) es un autor que enfatiza en esto, indicando la implicación de un proceso de transformación en el sistema educativo, así como la renovación de la transferencia de créditos entre instituciones, movilidad académica, crear currículas flexibles, etc., de manera que se pueda apoyar en el proceso de evaluación del aprendizaje y así lograr medir con objetividad en cuanto a la autodirectividad, originalidad, trabajo en equipo y la capacidad de valorar habilidades interpersonales y que de esta manera, tras los resultados de la evaluación se logre cumplir los fines y propósitos planteados.

Por lo que se puede decir que la evaluación es un proceso implicado en todo ámbito educativo, que ayudará a determinar hasta que punto se están alcanzado los objetivos planteados inicialmente, a través de los resultados de aprendizaje logrados por los educandos, mismos que responden a las necesidades de mejora continuo de la educación, lo que hace necesario que los profesionistas en la disciplina cuenten con las herramientas necesarias para la obtención de confiabilidad en los datos, y de esta manera fundamentar las propuestas de cambio hacia los problemas relacionados.

En los apartados siguientes se hablará de su funcionalidad y tipos de evaluación que se manejan, lo cuál dependerá del momento en que se desee realizar ésta.

2.1.2 Funciones de la evaluación.

Como toda etapa del proceso educativo, la evaluación cumple con ciertos propósitos, Alonso (1998) menciona que una de las funciones que tiene se relaciona con la contribución al aprendizaje de los estudiantes; de esta manera la evaluación en el ámbito educativo desde el punto de vista de la aplicación de exámenes de conocimientos, va a depender de la concepción que se tenga del aprendizaje, lo cuál no es un asunto fácil debido a la complejidad del tal proceso, pero se puede ayudar con herramientas psicométricas, como las pruebas de aprovechamiento escolar.

Los exámenes y otros procedimientos de evaluación se utilizan para adoptar decisiones críticas, sobre cada estudiante y su futuro en la educación; Bloom (1977) retoma que el propósito de la evaluación es la clasificación de los estudiantes, a descubrir aquellos que han fracasado o que han tenido éxito, y los que se han limitado

a salir del paso, donde las pruebas y otras formas de evaluación contribuyen de manera reducida al mejoramiento de la enseñanza y el aprendizaje y raras veces para asegurar su aprendizaje, pero que sin lugar a duda puede verse como el primer paso para ir mejorando el proceso.

Para Venegas, (1993), el objeto de la evaluación debe ser el de indagar sobre el proceso de aprendizaje de un sujeto o un grupo, que permita detectar las características de este proceso y buscar una explicación de las mismas, rebasando el atender sólo a los resultados del aprendizaje.

Otra función es relacionar las habilidades de cambio en el estudiante con la enunciación de los cambios buscados, decidiendo los que son posibles y los que son deseables, debido a que el propósito primario de la evaluación es describir estos cambios e influir sobre ellos para comprender la causa y el efecto de las variables que sean importantes para obtenerlos en los alumnos (Bloom, 1977).

Rosado (1998) menciona que dentro de las funciones de la evaluación, se atiende a la necesidad de verificar los objetivos, en cuanto a la responsabilidad social de la institución educativa, otorgando certificados por la conclusión de un ciclo y en cuanto a la actividad educativa del profesor y los estudiantes, en donde la evaluación funge como proceso de realimentación y de motivación.

Sin embargo, no debe tomarse como un punto sin importancia, puesto que es un asunto sobresaliente para determinar avances o deficiencias en cualquier organización; y en la educación, su funcionalidad estará alrededor de los cambios propicios para la currícula educativa y en la planeación de la transferencia del conocimiento, logrando que los aprendizajes sean propicios. Por lo cuál se puede ver a la evaluación como un paso para lograr un cambio en el proceso de enseñanza-aprendizaje, de manera que al detectar errores en el proceso, se piense en soluciones plausibles y así superarlo.

2.1.3 Tipos de evaluación

Existen tres tipos de evaluación que son la sumativa, diagnóstica y formativa. Las metas de la evaluación sumativa son calificar a los estudiantes, juzgar la eficiencia docente y comparación de currícula. La formativa apunta hacia los problemas que

necesitan una solución durante la formación del estudiante, a fin de obtener beneficios en estudios subsiguientes, y tiene que ver con unidades más pequeñas. En cambio la diagnóstica tiene que ver más con el realizar juicios para colocación de los estudiantes en cierto ámbito. Sin embargo se hace necesario el dar una explicación más amplia de las mismas, por lo que a continuación se describirá cada tipo de evaluación.

2.1.3.1 Evaluación Diagnóstica

Esta evaluación se encamina hacia la ubicación de los estudiantes en un curso en base a un diagnóstico, describiendo y clasificando las conductas de los estudiantes; forjando que los propósitos sean dobles, por una parte el de ubicar al estudiante al comenzar un curso, o el descubrir las causas de las deficiencias en su aprendizaje a medida que avanza la instrucción; y por el otro establecer si el estudiante posee el dominio de ciertos objetivos importantes en un curso, y que provoque su inserción en uno más avanzado. Se puede clasificar a los estudiantes de acuerdo a características como el interés que tiene, la personalidad, antecedentes, aptitudes, habilidades y su instrucción previa (Bloom, 1977).

Tiene como principal función, la determinación de causas que subyacen a deficiencias del aprendizaje de un estudiante que no ha respondido a la instrucción; se relaciona con la agrupación de los estudiantes; existen tres tipos de instrucción (Yales, 1996; en Bloom, 1977). La primera se refiere a asignación de estudiantes a diferentes tipos de escuelas; la segunda se da dentro de la escuela, que también comprende la ubicación de los estudiantes en un grupo de instrucción acorde con el nivel de conductas iniciales requeridas y finalmente la tercera coloca a los estudiantes en conjuntos amplios, como los llamados alumnos "brillantes" o los que se puedan colocar en un grupo adecuado.

La evaluación diagnóstica puede ser diferente o relacionarse con otro tipo de evaluación, como la formativa, que estará determinada por la función a cumplir, donde se puede tratar de colocar al estudiante en un punto inicial más adecuado, convirtiendo al diagnóstico en una operación anterior al proceso y convirtiéndose en un proceso continuo en conjunción con la evaluación formativa. De igual manera, la evaluación

diagnóstica al relacionarse con la evaluación formativa determina si factores no relacionados con la instrucción son causas de fracaso en el aprendizaje enseñado (Bloom, 1977).

También se presenta una relación de la evaluación diagnóstica con las evaluaciones al final del curso (sumativa), puesto que una de las dificultades inherentes a la utilización de calificaciones sumativas para la ubicación es que emplear una calificación no describiría adecuadamente el desempeño del estudiante o el dominio que en cada uno de los objetivos; sin embargo pueden rescatarse beneficios inmediatos que dan una imagen clara de los puntos fuertes del aprendiz en sus progresos y cómo ayudar para mejorar.

Bloom (1977) menciona que otra forma de diagnóstico implica la determinación del grado en el cuál el estudiante ya ha dominado los objetivos planteados en el curso.

De forma tal que se puede ver a la evaluación diagnóstica como un proceso de detección de necesidades, puesto que se realizará previo a un curso o programa de instrucción, lo que facilitará la colocación del alumno, con lo que ya sabe y necesita saber y planear el arranque de la enseñanza a partir de lo que el estudiante posee, procurando un aprendizaje más productivo.

2.1.3.2 Evaluación Formativa

La evaluación formativa implica evidencias durante la elaboración y prueba de un nuevo currículo para basar en ello revisiones del mismo. Este tipo de evaluación tiene lugar durante la formación, en el curso del aprendizaje, para lo que se deben obtener evidencias útiles en el proceso, buscar el método más eficaz para informar de estas evidencias y encontrar formas de reducir el efecto negativo asociado a la evaluación. Se puede centrar en el proceso del estudiante y así determinar la instrucción que ha de recibir (Brown, 1992)

Un aspecto importante de éste tipo de evaluación, es la selección de una unidad de aprendizaje, que representarán "los temas de una asignatura que deben ser aprendidos a lo largo de un determinado período" (Bloom, 1977, p. 174)

La evaluación formativa debe tener todos los elementos incluidos en la tabla de especificaciones (ver definición en apartado de elaboración de pruebas), Bloom (1977) menciona que el aspecto más importante de este tipo de evaluación es la ayuda que puede dar al estudiante al aprender algún tema y para la especificación de las conductas para cada unidad de aprendizaje, estableciendo metas para el mismo y plazos para cada unidad; ayudando al estudiante al dividiendo la secuencia de aprendizaje en unidades más pequeñas y estimulándolo a que tenga una mejor preparación mientras aprende una unidad en particular.

En la aplicación de una prueba formativa, se debe dar una retroalimentación al estudiante acerca de aquellos reactivos que ha contestado de manera acertada y los que no, para que puedan obtener información de lo que posee y de lo que necesita revisar. Este tipo de pruebas pueden ayudar al docente y al estudiante a determinar la calidad del aprendizaje, pudiendo así predecir los exámenes sumativos con cierto grado de anticipación.

2.1.3.3 Evaluación Sumativa

Son las evaluaciones que se realizan al final del curso, para determinar si los estudiantes han dominado los objetivos de instrucción, con la finalidad de obtener un índice del grado de aprendizaje del alumno (Brown, 1992).

El propósito es precisar el grado de dominio de cierta tarea de aprendizaje y señalar la parte de ésta que no se domina, no califica al alumno, por el contrario se orienta hacia los resultados alcanzados en todo el curso o en una parte de él. Un examen sumativo debe elaborarse en base a una tabla de especificaciones, que es una matriz con las conductas y contenidos a ser medidos, su uso va desde la escuela elemental hasta los postgrados, siendo base para la asignación de calificaciones, con el fin de distribuir a los estudiantes en función de la cantidad o nivel de aprendizaje en relación con otros. En ocasiones una prueba de esta índole, en cuanto a orientación académica tiene como función la predicción del éxito en un curso posterior relacionado con la materia (Bloom, 1977).

Es importa resaltar el papel que juega la evaluación en el ámbito educativo y aún más en éste estudio, por lo cuál en el apartado siguiente se mencionará la normatividad de su uso en la admisión educativa.

2.1.4 Relevancia en el proceso de admisión en instituciones educativas.

La evaluación en las instituciones universitarias, ha provocado que éstas desarrollen criterios y pautas para regular el ingreso de los aspirantes; Alonso, (1998) relaciona esto con el que la normatividad incluye condicionar su acceso a las calificaciones de estudios anteriores, hasta exigir exámenes psicológicos y médicos, aunque en la mayoría de los casos, es el de aprobar un examen escrito, con diversas pautas para su calificación. Por lo tanto, mediante un proceso de admisión, se rechaza a un número considerable de tales aspirantes, que hace que el proceso necesite ser evaluado; este autor concluye que la admisión a la enseñanza superior es un tema de atención y preocupación para las universidades, quienes a pesar de diferir en los instrumentos que emplean para la selección de los alumnos, tienden a implementar diferentes procesos de admisión con la finalidad de dar respuesta a la demanda que la población estudiantil ejerce al intentar el ingreso a estudios superiores.

Aunque cabe resaltar que en el caso de los posgrados, el ingreso se ve condicionado a admitir a personas que desean un mayor grado de estudios y que por lo tanto la admisión se centra en aquellos que cuentan con un conjunto de competencias necesarias para su desarrollo, más que centrarse en la problemática de la demanda a las universidades.

De cualquier forma, dicho proceso de admisión debe ser efectivo, lo cual puede lograrse a través de la validación de los instrumentos que se emplean en él, de lo contrario, se estará excluyendo injustificadamente a un gran número de aspirantes que deseen ingresar a las instituciones educativas.

La importancia del proceso de admisión, sobresale porque en éste se introduce al aspirante a estudios de nivel superior, o se le excluye en base a los datos obtenidos a través de las actividades del propio proceso de evaluación (Alonso, 1998)

Al hablar de la inclusión estudiantil, se incluye el tema de evaluación, misma que implica la medición de rasgos o grado de conocimiento en los alumnos a ingresar a alguna institución, donde cobra relevancia conceptos como los confiabilidad y validez, por lo que se hace indispensable el hablar de tales tema con la finalidad de mostrar su importancia dentro de la implementación de un proceso de admisión.

2.2 Desarrollo de instrumentos de medición en la educación

La evaluación, incluye realizar mediciones, que en psicología comenzaron al identificar elementos del mundo real con las construcciones de un sistema lógico abstracto, al que se denominaba modelo, que se usa para formar hipótesis y así dar explicación a los eventos paradójicos de tales mediciones. Morales dice (1975), "es el procedimiento mediante el cuál asignamos números (calificaciones, medidas) a las propiedades, atributos o características de los objetivos (o de unidades experimentales), estableciendo las reglas específicas sobre las que se fundamentan tales asignaciones" (p. 11)

Dentro de la medición, en la Teoría Clásica pueden presentarse errores de medida, los cuales se refieren a los problemas de estandarización, de interpretación, el error variable, los errores personales y el constante; los cuales se definirán a continuación, que de acuerdo a Morales (1975) son:

- *Problema de estandarización:* las calificaciones de un individuo no tendrán significado si no se comparan en función de las ejecuciones de un grupo.
- *Error de interpretación,* se refiere al grupo con el que se comparó el sujeto o por la forma en que se llegó a la comparación entre el sujeto y el grupo. Para lo cual existen dos medios de comparaciones, comparar al sujeto con el grupo conveniente, donde no se puede clasificar al individuo si no se tienen un estándar de su propio grupo; y en segundo término, compararlo con un grupo al cuál desea formar parte, comparándolo con los individuos sobresalientes o "brillantes" que forman el grupo de comparación, en vez de hacerlo con sus mismos compañeros.

- *Errores variables*: surgen por los “accidentes” que pueden ocurrir en el momento de la aplicación de las pruebas, y por la ejecución de las personas o porque las calificaciones de una persona son diferentes al ser medidas.
- *Errores personales*: se producen por las diferencias en las formas de comprender, manejar e interpretar el instrumento, es decir referente a la persona examinadora.
- *Error constante*: se da porque las medidas de los atributos de las personas son indirectas, que se refiere a no medir el fenómeno en sí, sino su expresión. Este error se encuentra íntimamente relacionado con la validez.

Sin embargo, la medición en la educación, se lleva a cabo a través de la aplicación de pruebas como una práctica para evaluar el grado de conocimientos de los alumnos o de la conducta de los mismos, en donde para la evaluación de los conocimientos, la herramienta más empleada son los exámenes, que son pruebas que no suelen cumplir con los procedimientos básicos de elaboración, que es un tema a tratar más adelante.

Díaz (1987, en Venegas, 1993) señala: “El examen es un instrumento para realizar la evaluación y no es el más apropiado para obtener las evidencias que llevarán a la acreditación” (p. 28), agregando que existen otros métodos de evaluación que de acuerdo a Pardo (1980) son “el método más adecuado para la medición del grado de conocimientos” (p. 4)

Es decir que se deben emplear otros instrumentos de evaluación además de la prueba de conocimientos para que el proceso no sólo se base en los resultados del examen, sino tener más evidencias que reduzcan al máximo la incertidumbre. Fermín (1971, en Venegas, 1993) menciona que los exámenes son medios pedagógicos para estimular la actividad intelectual de los estudiantes y corregir periódicamente defectos de su formación. Como instrumentos auxiliares de evaluación, en ellos debe atenderse, más que a la memorización de las materias, al aprovechamiento que demuestre el alumno mediante la comprensión del saber recibido.

Debido a las finalidades del estudio, en la siguiente sección se hablará de una clasificación de las pruebas y del proceso a seguirse en su elaboración.

2.2.1 La medición en la selección de alumnos.

Hasta el momento las diferentes teorías del aprendizaje no han desarrollado ningún modelo explicativo del comportamiento ni de su interrelación con los aspectos cualitativos y cuantitativos (Venegas, 1993), lo que hace pensar que se debe hacer una buena planificación de los instrumentos de medición, que puedan ayudar a realizar explicaciones objetivas de lo que se pretende medir.

Dalhaus (1990, en Alonso, 1998), atribuye a James Mckenn Catell el empleo de los primeros exámenes destinados a universitarios, en su artículo de 1980 en el que los test que aplicaban en una universidad norteamericana para conocer su nivel intelectual, se encaminaban a medir energía muscular, velocidad del movimiento y del oído, discriminación de peso, tiempo de reacción y otros.

Sin embargo, en lo referente a la medición para la selección de alumnos, el instrumento debe constituir un buen diseño de medición, cumpliendo con criterios como son: estandarización, confiabilidad, objetividad y validez, ocurriendo que una prueba sea poco confiable a pesar de estar bien estandarizada y ser altamente confiable, pero con una validez poco útil. La falta de confiabilidad, reflejará falta de objetividad, y por lo tanto carecerá de validez. (Morales, 1975)

En 1900, se generalizaron y aplicaron exámenes objetivos de aplicación uniforme respecto a las respuestas dadas por los alumnos, sin embargo, el antecedente más específico del examen de admisión a las universidades remiten al Collage Entrance Examination Borrard en USA en 1900; donde para 1926 aplicó el primer examen de admisión al Collage, el Scholastic Aptitude Test (SAT), la Prueba de Aptitud Académica (PAA), que tuvo auge a partir de la Segunda Guerra Mundial (Alonso, 1998).

Sin lugar a duda la importancia de elaborar un buen instrumento de medición, no sólo tendrá repercusiones en el ámbito de diagnóstico, sino también en el de la evaluación de candidatos, Chávez (2000) menciona que es importante para la evaluación de candidatos a ocupar un puesto, debido a que lo que se busca es seleccionar al mejor individuo para su incorporación en la organización.

El constructor de instrumentos de medición debe tener presente si el instrumento es apropiado para la situación particular, así como los ámbitos donde se debe emplear

la evaluación, para determinar la eficacia de la línea base de donde se parte y así determinar si los objetivos han sido alcanzados o no.

Aguilar, Lizárraga & Pérez (1981) menciona que el reconocimiento de deficiencias académicas, ha llevado a realizar evaluaciones de los aspirantes a ingresar a escuelas profesionales, para su selección y clasificación. Dichas evaluaciones abarcan por un lado el área psicométrica (aptitudes intelectuales, intereses vocacionales y rasgos de personalidad), y por el otro el área académica, que comprende áreas curriculares; y menciona que estas evaluaciones “no investigan la validez de las pruebas empleadas y que los resultados son mejores al emplear pruebas de conocimientos, por tal motivo en este estudio el interés central girará alrededor de las pruebas de conocimiento como método para la medición de éste.

Sin embargo antes de continuar con el tema de pruebas, habrá que marcar la diferencia que existe entre evaluación, medición y pruebas, debido a que en algunos casos puede llegar a verse como procesos iguales, por lo que se presenta el siguiente diagrama para entender más este aspecto.



Figura 1. Relación producida entre los conceptos de evaluación, medición y prueba, tomada de Carrera & Romo (1983).

Por lo que se debe entender que “medir es asignar números a las cantidades de las propiedades de los objetos de acuerdo con reglas y cuya validez puede probarse empíricamente” (Carrera & Romo, 1983, p. 117)

En cambio por prueba se entiende un instrumento con un conjunto de preguntas que se han de contestar, que provocan reacciones registrables, de toda naturaleza en cuanto a complejidad, forma, duración, significado y expresión (Carrera & Romo, 1983), por lo tanto se debe tener claro que al realizar una evaluación, nunca medimos personas, sino a sus características o propiedades, y esto dependerá de los fines a los

que se destine tal medición, para que de esta manera se proceda a elegir la prueba adecuada para cubrir esos objetivos. De tal manera que en este estudio se aborda el tema de pruebas para medir conocimientos y sus características, las cuales se citan a continuación.

III) TIPOS DE PRUEBAS PARA MEDIR CONOCIMIENTOS

3.1 Definición de conceptos

Antes de comenzar, cabe mencionar algunas definiciones que se dan a lo que es un “test” el cuál tiene varios significados; Martínez (2000) menciona que se le puede ver como nombre, donde su significado es de prueba, reactivo, etc.; como verbo, entendido como ensayar, probar o comprobar, aunque en realidad es una palabra proveniente del latín “testis”, que significa testigo, testimonio, etc.

La mayoría de las veces en el lenguaje usual se le llama simplemente prueba, sin embargo se exponen algunas definiciones citadas por algunos autores.

- ✚ Anastasi (1998): lo define como una medida objetiva y estandarizada de una muestra de conducta.
- ✚ Cronbach (1971; en Martínez, 2000., p. 31): “Técnica sistemática para comparar la conducta de dos o más personas”
- ✚ Yela (1980; en Martínez, 2000., p. 32): “Un test es una situación problemática, previamente dispuesta y estudiada a la que el sujeto ha de responder siguiendo ciertas instrucciones y de cuyas respuestas se estima, por comparación con las respuestas de un grupo normativo (o un criterio u objetivo), la calidad, índole o grado de algún aspecto de su personalidad”

De lo que se puede agregar que un test o prueba es un instrumento de medición de algún aspecto de la conducta que se desea evaluar, en donde a partir de estas definiciones se pueden extraer sus principales características, mismas que serán descritas en el siguiente apartado.

3.2 Principales características

Dentro de lo que conforma un test, se extraen las siguientes características tomadas de Martínez (2000).

- a) Medida objetiva, referente a la precisión que debe tener la medida.
- b) Muestra de conductas; medir en los sujetos, rasgos o características de los mismos, eligiendo elementos representativos del rasgo o conducta a medir.
- c) Técnica sistemática, que se vincula a la situación problema que ha sido estudiada, a la que el sujeto debe responder de acuerdo a ciertas instrucciones que se han fijado anteriormente.
- d) Comparar conductas, compara la respuesta del sujeto con un grupo normativo para determinar el grado del rasgo o aspecto a medir.
- e) Predicción o inferencia, que es el hecho de hacer predicciones de conductas importantes que van más allá de las que se han observado durante la ejecución de la prueba.

Como se puede observar estas características se podrían ver como requisitos que independientemente de la finalidad de la prueba, deben cumplirse, y precisamente hablando de los tipos de pruebas, una de las más empleadas en el área de la educación son los test de rendimiento (examen), en donde se debe tomar en cuenta que para su elaboración, se debe parte de las necesidades de acuerdo a planes de estudio y tener claro el constructo que se está midiendo; este tema se explica más a detalle en el capítulo de planeación de pruebas.

3.3 Pruebas para evaluar conocimientos

En general son instrumentos que evalúan los logros alcanzados por el alumno en distintas áreas, resultan de gran utilidad si para su construcción se basó en constructos psicométricos.

Este tipo de pruebas requieren validez de contenido y se refieren a exámenes para cursos individuales de instrucción en todas las escuelas y niveles, y que son

vistos como medidas estandarizadas de aprovechamiento usadas en la instrucción (Nunnally & Berstein, 1997).

De dichos test se desprenden dos tipos principalmente, que son de ensayo o llamados de composición y los test objetivos, que es de los que más se hará énfasis por su importancia en el tema. Nunnally & Berstein (1997) menciona que la construcción de los reactivos requiere de cierta habilidad, por lo que en los apartados siguientes se mencionan los aspectos para su construcción.

3.3.1 *Test de ensayo*

Son los que se usan en las escuelas como exámenes escritos, en donde suelen utilizarse preguntas amplias en las que el alumno tiene toda la libertad para responder, organiza sus respuestas con una mayor libertad, utilizando sus propias palabras y suele ser manuscrita (Cortada, 1999).

Su uso se limita a la medición del aprovechamiento complejo, en donde Cortada (1999) coincide con Gronlund (1985) en el hecho de que el rasgo distintivo de las preguntas de ensayo es la libertad de la respuesta que se permite al educando, aunque menciona que precisamente ésta libertad hace que este tipo de preguntas encarezca su valor como una medida de aprovechamiento complejo, debido a las dificultades que representa su calificación.

Thorndike & Hagen (1973) indican que la tarea esencial que se requiere por parte del alumno se resume en organizar sus propias respuestas, utilizando sus propias palabras y estilo de escribir en un pequeño número de preguntas, lo cuál acarrea problemas de falta de confiabilidad en la calificación y en el hecho de que el alumno debe recordar los temas para otorgar respuestas adecuadas y no en el hecho de reconocerlas, tratando de medir procesos mentales de nivel superior, aunque esta "ventaja" no ha sido confirmada experimentalmente.

Por otro lado en los estudios acerca de la calificación de las respuestas a estas pruebas (Gosling, 1966; Scannell & Marshall, 1966; Marshall, 1967; en Thorndike & Hagen, 1973) han obtenido que las puntuaciones se ven afectadas por la calidad de la expresión escrita, ortografía, la gramática e incluso caligrafía, lo cuál indica que un

alumno con conocimiento, pero sin habilidad para expresarse de manera escrita, obtendrá una calificación baja, que lleva a que se califique más la habilidad de escritura que el dominio del conocimiento.

De manera que se recomienda que las preguntas de ensayo se usen para medir productos de aprendizaje que no puedan ser medidos a través de pruebas objetivas, haciendo que tales productos se refieran a seleccionar, organizar, integrar, relacionar y el evaluar ideas (Gronlund, 1985), sin embargo serviría de mucho el definir claramente los parámetros de calificación para tratar de hacerlas más confiables, (James, 1927; Cureton, 1958; Payne, 1968 & Myers, McConville & Coffman, 1966; en Karmel, 1978).

Por lo que al momento de redactar las preguntas se debe definir la tarea y redactar los reactivos de manera clara, siendo un número pequeño para poder ser respondidos por los examinados, con una estructura tal en la que los expertos estén de acuerdo en la calificación de las respuestas y cuidar que los examinados respondan a cada reactivo en una hoja por separado (Aiken, 2003).

Karmel (1978) indica problemas en el establecimiento de la validez en este tipo de pruebas, dado que se utiliza un pequeño número de preguntas, serán pequeños los objetivos y esto tiende a disminuir la validez, así como el problema de evaluar más de un área, debido a que entra en duda si realmente mide o no lo que se pretende, por lo cuál Karmel concluye que las pruebas de ensayo pueden usarse como un instrumento útil de enseñanza para facilitar el aprendizaje.

A la vez, las de ensayo se dividen en las de pregunta restringida y extensiva.

3.3.1.1 Preguntas de respuesta restringida

Se requiere que el alumno suministre la respuesta, se limita el contenido y la forma de respuesta del alumno; en cuanto al contenido se da en el hecho de restringir el alcance del tópico por examinar, las limitaciones en la forma se indica en la expresión de la pregunta; otra manera de restringir las respuestas es que las preguntas se basen en problemas específicos.

La utilidad de este tipo de preguntas es cuando se pretende medir productos de aprendizaje que requiere la interpretación y aplicación de datos de un área en

específico, lo que hace posible la medición de productos de aprendizaje de manera específica y definida; aunque esta limitación hace que no se puedan expresar aspectos como la integración, organización y la originalidad por parte del alumno (Gronlund, 1985), para lo cuál suele utilizarse la pregunta de respuesta extensiva, que proporciona una mayor libertad de respuesta, midiendo productos del aprendizaje de índole más general, pero que a la vez presenta una mayor cantidad de limitaciones que la pregunta de respuesta restringida, ya que la respuesta extensiva proporciona un amplio intervalo de libertad en la respuesta; esto se revisa con mayos énfasis en el apartado siguiente.

3.3.1.2 Preguntas de respuesta extensiva

Gronlund, (1985), menciona que da al alumno libertad de seleccionar información de lo que considere pertinente, que organice su respuesta e integre y evalúe ideas para darlas; demostrando su competencia y capacidad para seleccionar, organizar, integrar y evaluar ideas, pero esta libertad hace ineficaz a la pregunta en cuanto a productos de aprendizaje más específicos, así como dificultades que se introducen para su calificación, al mismo tiempo que restringe su uso como instrumento de medición, provocando complejos que no pueden ser medidos con otros medios, que hace menos confiable la calificación, limitando su uso al hecho de dispositivos para la enseñanza.

3.3.2 Test objetivos

El elaborar pruebas objetivas surge en respuesta a las desventajas de las de ensayo, y entre sus características se encuentra el hecho de que el proceder en la calificación se establece al momento de redactar los reactivos, lo que constituye una desventaja en los de ensayo. Este es un test estructurado, en donde a cada persona se le presenta el mismo problema y la forma de contestación es la misma, el estudiante reconoce la respuesta correcta, no recuerda; y permite emplear una mayor cantidad de reactivos que ayudan a la vez a aumentar la confiabilidad y la validez (Karmel, 1978). Este tipo de prueba se puede emplear con fines diagnósticos, formativos o resumidos y ser tipificada, es decir elaborada con todos los requisitos técnicos, probada con tablas,

con el rigor de la elaboración, y con personal calificado, de hecho Carreño (1993) menciona que un caso en el que se justifica la elaboración de una prueba tipificada, es el de las pruebas de admisión a instituciones educativas, con "el fin de asegurar mayor exactitud en las mediciones y justicia en las comparaciones" (p. 22).

Como lo menciona Cortada (1999), este tipo de pruebas suponen diversas formas, aunque todas siempre tendrán una sola respuesta que será la correcta y que se ha determinado previamente al elaborarse el test. Constituyen preguntas utilizadas con una frecuencia mayor para evaluar el aprendizaje teórico como el práctico, en donde lo que se pretende es el realizar una evaluación con pruebas integradas, a partir de una serie de preguntas objetivas, mismas que son llamadas de respuesta restringida, debido a que el alumno selecciona una opción dentro de un limitado de posibilidades, permite eliminar las dudas al calificar y podrá ser calificada por cualquier persona que tenga la clave de las respuestas que constituyan las correctas para la prueba (Quesada, 1991).

En este tipo de pruebas el alumno trabaja en una tarea estructurada, seleccionando respuestas de cantidad limitada, contesta a una muestra grande de reactivos y recibe un puntaje para cada respuesta (Cortada, 1999).

Una prueba objetiva puede ofrecer más seguridad de la validez que la acompaña, además que de acuerdo a Adkins (1981) se deberían tomar en cuenta factores importantes de éstas pruebas, como lo es la confiabilidad en la calificación, la economía en la misma, la accesibilidad de los resultados a los estudiantes en breve tiempo, lo que en el factor motivacional tiene aplicaciones importantes; otro factor mencionado por Adkins es el muestreo adecuado de contenido, debido a que permite muestrear el universo de contenido y las acciones que constituyen las metas de la unidad en particular del programa.

Los reactivos que se elaboren deben ser claros, precisos y gramaticalmente correctos, utilizando un lenguaje adecuado para el nivel de lectura de las personas a las que se dirige, incluyendo información necesaria en el reactivo para ser contestado, sin que con esto se incluyan claves para las respuestas correctas.

Al elaborar preguntas objetivas se debe evitar el redactar reactivos interrelacionados, en donde uno sea señal para la respuesta del otro o entrelazados, en

donde se hace necesario conocer la respuesta de un reactivo para saber la respuesta correcta del otro (Aiken, 2003), este punto cobrará mayor importancia en el momento del ensayo de los reactivos dentro del marco de la Teoría de Respuesta al Ítem, que se trata más adelante.

3.3.2.1 Tipos de reactivos

Dentro de los test objetivos, las preguntas más conocidas son las de opción múltiple, verdadero- falso, jerarquización, correspondencia y completamiento. Quesada (1991), menciona que cada una de estas preguntas tiene ciertos requisitos para su formulación, aunque da algunas recomendaciones que son plausibles de emplearse, independientemente de la forma que adopten las preguntas, tal como se expone a continuación.

- ✿ Usar un lenguaje comprensible y claro para los alumnos.
- ✿ Omitir preguntas capciosas.
- ✿ Ser directo.
- ✿ Evitar redactar en forma negativa.
- ✿ Presentar información para su contestación, sin datos innecesarios.
- ✿ Evaluar sólo conocimientos necesarios.
- ✿ Preguntar una sola cosa cada vez.
- ✿ Evitar dar ayudas a la respuesta.

Sin embargo existen algunos parámetros a considerar para la elaboración de reactivos, dentro de los que se recomiendan los siguientes:

Tabla 2

Criterios que deben ser considerados para elaborar reactivos de tipo selección de respuesta, adaptados de Haladyna (1999; en Hogan, 2004)

Sobre el contenido	Sobre el formato	Sobre el estilo	Sobre la redacción	Sobre las opciones
Se debe basar en lo específico, siendo independiente uno del otro.	Eludir los de verdadero/falso y los de opción múltiple complejos.	Se deben editar y probar todos los reactivos.	Las instrucciones para las preguntas deben ser muy claras.	Cuatro opciones parecen ser adecuadas (Quesada, 1991)..
La información no debe ser muy general o muy específica.	El reactivo se debe formar en posición vertical y no horizontal.	La redacción debe ser sencilla en función del grupo al que se someterá la prueba.	La idea principal debe estar en la pregunta y no en las opciones.	Sólo una de las opciones debe ser la correcta.
Evitar los reactivos que se redactan en base a opiniones.			La pregunta se debe formular en positivo.	Las opciones se deben colocar en orden lógico o numérico.
Evitar los reactivos engañosos.				Deben ser independientes de las otras y homogéneas en el contenido.
				La extensión debe ser del mismo tamaño.
				Todos los distractores deben ser verosímiles.
				Evitar las opciones graciosas.
				Evitar las negaciones, así como excepto a y b, sólo c y d, etc.

En los reactivos en los que se tiene que seleccionar la respuesta son considerados más familiares en lo que respecta a la evaluación de las capacidades y el desempeño; a continuación se dará una descripción de cada tipo de reactivo y requisitos que deben cumplir para una elaboración adecuada.

a) Preguntas de opción múltiple

Muestran un enunciado que representa la pregunta y varias opciones de respuesta. Se puede usar para medir resultados del conocimiento y varios tipos de capacidad intelectual (Gronlund, 1985).

Adkins (1981) indica que una de sus características es que contienen en sí mismos las normas, partiendo de la que se puede escoger la respuesta mejor. La alternativa correcta es en la que se han puesto de acuerdo autoridades en la materia. Es un reactivo flexible, ya que puede dirigirse hacia el mismo conocimiento, en sí, tiene todas las ventajas de objetividad en la calificación y carece de las desventajas de los de verdadero-falso.

Son los más versátiles, pueden usarse para medir logros de aprendizaje simple y complejo en todos los niveles y en todas las áreas temáticas. Para responder de manera correcta a éste tipo de reactivos, se necesita habilidad para discriminar y no sólo para reconocer o recordar la respuesta (Aiken, 2003).

Pardo (1980) menciona que este tipo de reactivos disminuyen las probabilidades de que el examen sea contestado al azar, facilitando la evaluación para los diferentes niveles de la taxonomía y que abastecen la calificación, a lo que Hogan (2004) describe que éste tipo de reactivos son más familiares para la evaluación de las capacidades y el desempeño.

Sin embargo un punto importante de éste tipo de reactivos deriva en el hecho de la elaboración de buenos distractores en las opciones de respuesta, para determinar su efectividad. Se elaboran en base a un factor racional, en el que la persona que construye la prueba juzga aquellos distractores que serán eficaces o en base a un criterio empírico, en el que de entre las respuestas incorrectas más populares, se seleccionan los reactivos que pueden servir de distractores (Aiken, 2003).

Entre lo que se recomienda para una elaboración adecuada se tiene:

- I. El número de respuesta recomendado es de cuatro o cinco opciones (Quesada 1991); aunque algunos autores recomiendan menos (Haladyna & Downing, 1993; en Aiken, 2003) para los fines del estudio se emplea lo mencionado por Quesada.
- II. Se prefiere el formato de pregunta, pero si se usa una afirmación incompleta, el espacio en blanco se debe colocar al final de la misma (Karmel, 1978; Aiken, 2003).
- III. La extensión de las respuestas debe ser similar (Quesada 1991; Karmel, 1978).
- IV. Usar un formato eficaz de ítem, presentando las opciones en forma de lista, para facilitar la lectura y la comparación de las opciones (Gronlund, 1985).
- V. Las opciones falsas deben parecer como plausibles (Quesada 1991).
- VI. La mayor parte del reactivo debe estar en la pregunta (Karmel, 1978; Aiken, 2003)
- VII. Que exista correspondencia gramatical entre las opciones y la pregunta (Quesada 1991).
- VIII. Si las opciones son fechas o edades, colocarlas en orden natural, o alfabéticamente sin que se procure dar la respuesta correcta (Aiken, 2003).
- IX. Evitar palabras tales como "todas o ninguna de las anteriores" (Quesada 1991), o usar expresiones negativas como "no" en el tronco de las opciones (Karmel, 1978; Aiken, 2003).
- X. Usar números para los reactivos y letras para las opciones, procurando que se de un número correcto de reactivos para el nivel o grado de edad que se pondrá a prueba (Aiken, 2003).
- XI. El lugar de la respuesta correcta debe asignarse al azar (Karmel, 1978; Quesada 1991).
- XII. Hacer que el nivel de dificultad sea tal que el porcentaje de examinados que respondan correctamente sea aproximado a la mitad del % al azar y el 100%: $\% \text{ correcto} = 50(K+1)/k$, siendo k el número de distractores por reactivo (Aiken, 2003).

Nunnally & Berstein (1997) mencionan que los reactivos de opción múltiple se consideran los mejores para la mayoría de los propósitos, debido a que son fáciles de administrar y de calificar, lo que los hace más confiables que los de las pruebas de ensayo, ayudando a una mayor amplitud de muestreo de los temas.

b) Verdadero- falso

Consiste en un enunciado con sólo dos opciones de respuesta. Se caracteriza porque sólo son posibles dos respuestas (Gronlund, 1985).

Se le considera un reactivo sencillo de elaborar, pero es el que más problemas plantea, puede escribirse y leerse con rapidez. Entre las críticas que recibe se encuentra el que suelen copiarse expresiones literales de un texto, pueden alentar la memorización, son ambiguos y no son útiles para medir objetivos instrucciones complejos (Adkins, 1981; Aiken, 2003).

A pesar de que Ebel (1979; en Aiken, 2003) defiende este tipo de reactivos, esta defensa es cuestionada.

Entre los requisitos a cumplir en este tipo de reactivos se encuentran.

- i. El enunciado debe incluir sólo una idea y que plantee asuntos importantes y no triviales (Aiken, 2003).
- ii. El enunciado o pregunta debe ser sencillo y breve (Quesada 1991).
- iii. Se debe redactar de manera precisa, que haga que se le juzgue inequívocadamente como verdadero o falso, es decir, que la falsedad o veracidad no se deduzcan obviamente (Gronlund, 1985).
- iv. Las afirmaciones verdaderas o falsas, deben ser aproximadamente de la misma extensión y cantidad (Aiken, 2003), evitando que las aseveraciones sean demasiado largas y complejas (Karmel, 1978).
- v. Evitar el uso de la doble negación, debido a que puede causar confusión (Gronlund, 1985).
- vi. En la afirmaciones de opinión, citar la fuente (Aiken, 2003).
- vii. Combinar al azar las proposiciones falsas o verdaderas (Carreño, 1993).

- viii. Evitar claves ajenas a las respuestas, tales como “siempre”, “nunca”, “todo”, “ninguno” y “solamente” debido a que tienden a hacerlas falsas y “usualmente”, “acaso” y “a veces” tienden a ser verdaderos los enunciados (Karmel, 1978; Gronlund, 1985; Carreño, 1993).
- ix. Plantear que las respuestas erróneas sean más atractivas, de manera que la lógica sugiera que las respuestas erróneas son correctas (Aiken, 2003).
- x. Redactar las proposiciones en términos afirmativos (Carreño, 1993).
- xi. Suprimir los datos accesorios o ajenos al aprendizaje que se desee medir (Carreño, 1993).

Sin embargo Karmel (1978), así como Adkins (1981) exteriorizan que este tipo de reactivos deben de usarse lo menos posible, debido a que presentan varios defectos tales como el propiciar la adivinación y en el hecho de que es muy difícil hacer aseveraciones que sean definitivamente verdaderas o falsas.

Adkins (1981) menciona que otra objeción es que pueden conducir a reacciones emocionales, de manera que los estudiantes tienden a completar estos reactivos con la sensación de que la prueba ha sido injusta en la estimación de lo que realmente saben o pueden realizar.

c) Jerarquización / ordenamiento

Se presenta un conjunto de elementos donde la tarea del alumno será el de ordenarlos. Adkins (1981) menciona que la tarea es indicar, para los elementos que se le presentan, el orden apropiado, que puede ser cronológico, lógico, etc., pero que sin embargo puede dar lugar a problemas en la calificación.

Carreño (1993) sugiere no usar menos de cinco ni más de diez cuestiones para ordenar y lo recomienda para contenidos que se refieren a objetivos en que la conducta esperada sea de discriminación, ordenamiento, establecimiento de secuencia u otra similar. Carreño menciona que presenta problemas en la calificación, aludiendo a dos en particular, una sería la forma injusta en la que se considera un acierto al ordenamiento correcto en la totalidad y un error cuando una parte o la totalidad de la

secuencia se realiza de manera incorrecta. La manera justa, es computar un acierto a cada par bien ordenado de elementos.

Las recomendaciones para este tipo de reactivos se enumeran a continuación.

- i. Los datos que se proporcionen deben ser plausibles de ser ordenados (Quesada 1991).
- ii. No mezclar en una misma proposición o enlistado elementos de diferente naturaleza (Carreño, 1993).
- iii. Los datos que se usen deben corresponder a la misma naturaleza (Quesada 1991).
- iv. En caso de existir material suficiente, formar de tres a cuatro grupos o series como máximo (Carreño, 1993).
- v. Usarlos sólo cuando se exploren objetivos referentes a discriminar posiciones, ordenar, jerarquizar, etc. (Carreño, 1993)

d) Correspondencia

Presenta dos columnas que se pueden relacionar, enlistándose en una columna una serie de pies, o premisas, y las respuestas se enlistan en otra columna. Este formato proporciona una forma más compacta para medir el mismo conocimiento (Gronlund, 1985), aunque Aiken (2003) menciona que son relativamente sencillos de elaborar, sólo miden la memorización de acontecimientos.

Se requiere que las instrucciones sean claras para poder establecer dichas relaciones, y se recomiendan si se desea actividades como: relacionar, vincular, clasificar, aplicar principios, inferir, etc. (Carreño, 1993)

Recomendaciones para su elaboración:

- i. Los elementos de las columnas deben ser del mismo tópico (Quesada 1991)
- ii. Ordenar las premisas en la columna izquierda y las opciones de respuesta en la columna derecha (Aiken, 2003).
- iii. Una de las columnas debe tener una mayor cantidad de elementos (Quesada 1991).

- iv. Se debe incluir material homogéneo en cada ítem de correspondencia, para hacer razonables las respuestas (Gronlund, 1985)
- v. Usar números o letras para identificar los elementos de cada columna de manera que se pueda observar su diferencia (Quesada 1991).
- vi. Procurar que todo el reactivo se coloque en una sola página (Aiken, 2003).
- vii. Especificar en las instrucciones las bases de la relación, e indicar si una respuesta puede usarse sólo una vez, más de una o no usarse (Gronlund, 1985).
- viii. Evitar que la última o últimas de las relaciones se establezcan por eliminación de las opciones ya resueltas, incluyendo de uno a tres elementos en la columna donde debe buscarse la relación (Carreño, 1993).
- ix. Incluir en cada lista, un mínimo de cinco y un máximo de doce relaciones (Carreño, 1993).

e) Completamiento o respuesta corta

Consiste en un enunciado que se encuentra incompleto y con espacios en blanco, los cuales representan la respuesta a llenar. Aiken (2003) menciona que en este tipo de reactivos se plantea una tarea, en la que el examinado debe llenar el espacio en blanco que presenta, dando una respuesta breve a la pregunta, éste tipo se encuentra entre los más sencillos de elaborar y es útil cuando se desee evaluar el conocimiento de la terminología, pero no son útiles cuando se desee medir objetivos instruccionales complejos, y la calificación no siempre es objetiva.

Carreño (1993) mencionan que son un medio útil para explorar aprendizajes simples, a nivel de conocimiento y en ocasiones para comprensión y aplicación, aunque no es tan recomendable pues se cae en pretender la adivinación.

Este tipo de reactivos suelen relacionarse con el test de ensayo, porque se produce la respuesta correcta en vez de reconocerla, pero a la vez es objetivo por el hecho de que se puede seleccionar una respuesta preparada antes del examen. Kamel

(1978) menciona que es útil para conocimientos referentes a hechos como nombres o fechas, aunque es difícil hacer reactivos claros que no confundan a los estudiantes.

Este tipo de reactivos puede ser preguntas de respuesta corta, que puede ser una palabra o una frase pequeña, aunque su calificación no es tan de rutina y no representa objetividad en su grado máximo.

Se deben seguir las siguientes recomendaciones:

- I. Señalar los espacios en blanco para las palabras que sean las más importantes (Quesada 1991).
- II. Las preguntas directas son más preferibles que las incompletas (Karmel, 1978; Aiken, 2003).
- III. Los espacios en blanco deben ubicarse al final del enunciado (Quesada 1991).
- IV. No se debe incluir muchos espacios en blanco (Quesada 1991).
- V. Las respuestas correctas deben limitarse a las metas reales de aprovechamiento (Karmel, 1978).
- VI. No dar indicaciones de la respuesta al variar el tamaño de la respuesta a través del espacio en blanco, de acuerdo con el número de letras, sino que los espacios deben tener una longitud fija (Karmel, 1978).
- VII. Se deben incluir instrucciones precisas que no dejen lugar a duda en la interpretación de la respuesta (Carreño, 1993).

3.4 Ventajas y desventajas de los test objetivos frente a los de ensayo

Ya mencionada la descripción de cada tipo de test, se hace un resumen de lo que puede ser considerado como lo rescatable para cada uno de ellos, tal resumen puede observarse en la tabla 3. Al terminar la tabla se puede observar un mapa conceptual, donde se ejemplifican los puntos referentes a las pruebas y los tipos de reactivos que existen para medir conocimientos.

Tabla 3

Diferencias y similitudes existentes entre los test de ensayo y los de tipo objetivo.

Test objetivos	Test de ensayo
El alumno elige entre opciones de respuestas predeterminadas.	Exige que el estudiante planee su respuesta y la exprese como él la entiende.
Son una serie de preguntas específicas que requieren respuestas cortas, lo que da una mayor confiabilidad a la calificación.	Sólo se plantean preguntas generales que requieren respuestas extensas, lo que no proporciona confiabilidad en la calificación y requiere mucho tiempo.
Se usa más la memoria de reconocimiento que la de evocación.	El tiempo que se emplea es más en pensar y escribir.
La calidad del test estará determinada por el constructor de la prueba y la calidad de los reactivos puede ser determinada estadísticamente.	La calidad será determinada por quien lee las respuestas del alumno, sin poder obtener datos concretos de la calidad de las pruebas escritas.
Da libertad al constructor para expresar sus conocimientos, pero limita al examinado a lo poco o mucho que sabe.	Da mayor libertad al alumno para expresar su individualidad en las respuestas, y al lector de la prueba en el momento de la calificación.
Exige que se tome una decisión que no resulta fácil entre las alternativas de respuesta.	Proporciona una medida directa de los productos complejos del aprendizaje, que no pueden ser medidos con otros medios.
Exige que se desarrollen modos de pensar para poder establecer la base de elección de sus alternativas.	Si se usa una tabla de especificaciones será más fácil construir las preguntas acorde con lo especificado.
Requieren menos tiempo para responder al reactivo, lo que permite un muestreo más amplio del contenido que en las pruebas de ensayo (Aiken, 2003).	Da énfasis a la integración y aplicación de habilidades del pensamiento y resolución de problemas.
Otra ventaja es la eficacia en la calificación, debido al ahorro de tiempo que se invierte en ello (Hogan, 2004).	Restringe su uso eficaz en un muestreo limitado, miden sólo algunas áreas.
	No deben usarse cuando el conocimiento o habilidad puedan ser evaluados por reactivos objetivos (Aiken, 2003).

Fuente: Elaboración propia.

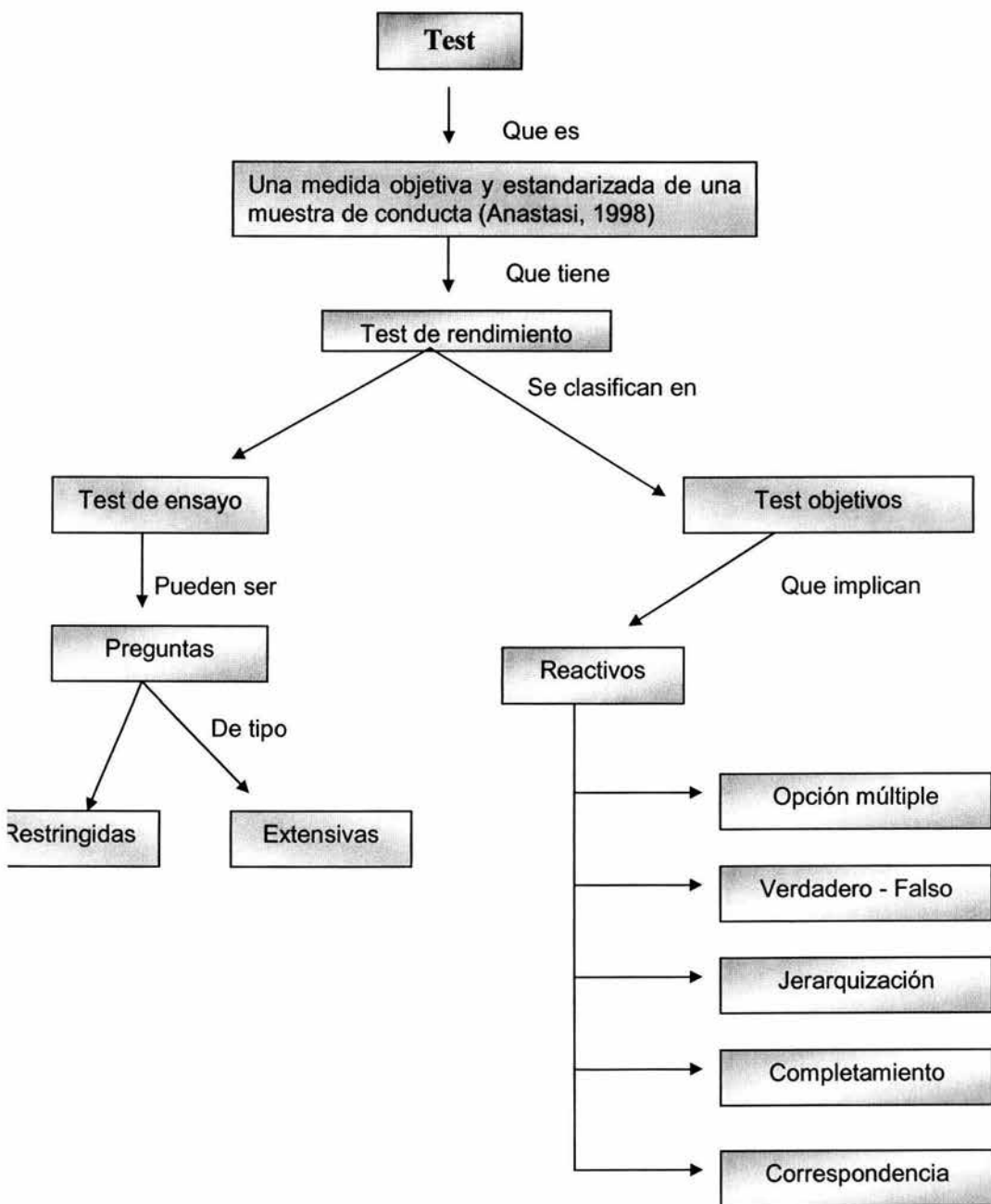


Figura 2. Representación grafica de los tipos en que se divide la prueba para medir conocimientos.

TERCERA PARTE.- ASPECTOS METODOLÓGICOS Y EXPERIENCIAS

IV) ESTRATEGIAS PARA LA ELABORACIÓN DE LA PRUEBA

4.1 Planeación de la prueba

La planeación es la clave para la elaboración de un test de aprovechamiento, dando mayor seguridad en cuanto a la medición de los resultados de aprendizaje y con una muestra representativa de conductas. Al planificar la prueba se pueden tomar una diversidad de métodos, pero al final los elaboradores de test coincidirán en algunos pasos que se mencionarán adelante.

Las pruebas se pueden utilizar para medir los resultados de aprendizaje o para su facilitación, lo que dependerá del tipo de evaluación que se realice, lo que se puede hacer a través de pruebas de rendimiento objetivas, para esto se debe considerar antes el método que se empleará; Brown (1992) hace notar que puede ser a través del método relacionado con el criterio o con las normas. En este apartado se definirán de manera general, como algo representativo del tema y al finalizar la parte de ensayo de reactivos se explica su importancia al momento de elaborar la prueba final.

- Método relacionado con los criterios

Consiste en definir los objetivos (las metas y resultados deseados) para el curso o una unidad y estos objetivos definen el universo o dominio que abarcará el examen y deben ser expresados en términos conductuales, en donde se exprese lo que se espera que el alumno sepa (Mager, 1962; en Brown, 1992); a lo que Thorndike & Hagen (1973) llaman definición operacional y que de esta manera se proceda a redactar los reactivos de manera que se logre medir esos objetivos, a través de una o más preguntas para cada uno de ellos.

Cohen & Swerdlik. (2000) menciona que en una prueba con referencia al criterio, puede resultar que el mismo patrón de resultados de la prueba ocurra para cierto reactivo, en donde las personas con calificaciones altas tienen correcto un reactivo en particular y quienes tienen puntuaciones bajas en la prueba equivocan en ese mismo reactivo. Implica un estudio piloto con al menos dos grupos: un grupo que domine el conocimiento y otro que no lo ha dominado, de manera que los reactivos que discriminen mejor entre estos grupos se considerarán "buenos".

- Método relacionado con las normas

Es el más común, especifica el contenido y las capacidades que se cubren en el curso o unidad, se construye el instrumento para muestrear este contenido y capacidades, a través de los tipos de reactivos que existen. Un buen reactivo de una prueba de rendimiento con referencia a la norma es aquél al que responden en forma correcta quienes obtienen calificaciones altas en la prueba, de manera que responden incorrectamente a ese reactivo las personas que obtienen calificaciones bajas en la prueba (Cohen, & Swerdlik, 2000).

El constructor del test debe detallar los contenidos y habilidades a medir, una guía es la Taxonomía de Bloom (ver más adelante), muestreando las áreas de contenido y habilidades y sentando la base de la planeación (Brown, 1992).

No se trata de hacer una prueba sólo por urgencia de evaluar, como sucede en los resultados de estudios publicados, donde las pruebas elaboradas por los maestros, son instrumentos defectuosos debido a la falta de planeación anterior a la redacción de preguntas (Thorndike & Hagen, 1973). Por lo que planear la prueba implica el contenido o las habilidades (objetivos) que cubrirá, expresado en su objetivo, de manera que esto representará una base para construir los reactivos y así la validez de contenido de la prueba (Brown, 1992); logrando que la muestra de preguntas refleje los objetivos de la enseñanza, a lo que Thorndike & Hagen, (1973) dicen que será la finalidad de la planeación de la prueba.

Aiken (2003) menciona que el primer paso en la planeación es la definición de las variables o constructos que van a medirse, la descripción de las personas a

examinar, especificar las condiciones bajo las que se aplicará la prueba, así como la manera de calificar, la interpretación de los resultados y el uso que se le dará; a lo que de acuerdo a Cohen & Swerdlik (2000) también se debe pensar en el objetivo principal de la prueba, la necesidad de elaborarla, el formato que debe tener, las formas que deben elaborarse, si será necesario la capacitación de personas para su aplicación, las personas que resultarán beneficiadas con la aplicación de la prueba y el significado que se dará a las calificaciones.

Otro punto importante es la finalidad de la prueba, donde Carreño (1993) menciona que las pruebas diagnósticas se usan para explorar aprendizajes que el alumno debiera poseer como antecedentes para el curso. En las pruebas formativas se exploran aprendizajes que se esperan conseguir con las experiencias suministradas y en las sumarias se explora aprendizajes derivados de todo un curso o parte de éste, de tal manera que se comenzará con el primer paso en la elaboración de una prueba o test, que es el referente a la definición clara del constructo que se desea medir.

4.1.1 Definición del constructo

García (1993) menciona que se debe comenzar con diferenciar los constructos que se piense son relevantes de aquellos que se pueden definir operacionalmente, tal como lo mencionaba Thorndike & Hagen (1973), debido a que sólo serán susceptibles de medirse aquellos que se hayan definido operativamente. La medida será válida en cuanto sea mayor la relación entre el constructo teórico y su operacionalización (Torgeston, 1958; en García, 1993), lo que nos regresa a que el primer paso será precisar y aclarar el significado del atributo a medir, conceptual y operacionalmente, a lo que Thorndike (1996) señala que al hacerlo se cubren dos objetivos; primero, describir el nombramiento del atributo y segundo dar una idea de los procesos a través de los que se evaluará el atributo, que se convierte en un procedimiento que facilitará esta primera parte.

Por lo cual se debe tener clara las características del objeto a ser medido y dar una contestación a ¿qué se va a medir? de manera operativa, es decir con conductas específicas que el alumno debe mostrar, expresado con el nombre que se le da al

atributo, como “conocimiento en aritmética”, etc., a lo que Carrera y Romo (1983) aluden que si una definición operacional es posible, el producto podrá ser medido, de lo contrario sería imposible demostrar su importancia.

El definir la finalidad de la prueba suele hacerse en un solo enunciado y de manera sencilla, sin embargo, al elaborar una nueva tiende a convertirse en algo difícil, prolongado y costoso, Hogan (2004) señala que elaborarla, puede deberse a una necesidad práctica, teórica o por una adaptación a una prueba ya existente. Por tal motivo se ha de considerar que la planeación se hace con vistas a un grupo de sujetos en particular a los que se aplicará y a criterios definidos sobre las decisiones de los resultados de la prueba, como son las de selección, para admitir o excluir en un grupo, en donde el papel de la prueba será admitir a candidatos prometedores, por la limitante de vacantes (Thorndike, 1996).

Para facilitar las conductas que deberán ser evaluadas desde el inicio se ha prestado atención a los sistemas de clasificación de los objetivos cognoscitivos, afectivos y psicomotrices de la instrucción, poniendo una mayor atención en los cognoscitivos, mediante la *Taxonomía de objetivos educativos: el dominio cognoscitivo* (Bloom & Krathwohl, 1956; en Aiken, 2003; Gronlund, 1998). Los verbos correspondientes para cada nivel de la taxonomía se indican en el apartado siguiente, en la preparación de la tabla de especificaciones.

Rosado (1998) indica que en tales dominios se incluyen acciones para la aplicación de uno u otro nivel, para que en el dominio cognoscitivo se den procesos mentales o intelectuales del aprendiz y el reconocer o evocar recuerdos, en el de comprensión alude a interpretar, resumir, parafrasear o traducir un material determinado; en aplicación a reconocer y recordar hechos aprendido en un momento dado; en el de análisis a separar un todo complejo en sus partes hasta quedar claras las relaciones existentes, de habilidades y destrezas; en el nivel de síntesis a elementos para formar algo nuevo o distinto y en evaluación a tomar decisiones o juicios en base a un conjunto de criterios o pautas.

En resumen, el definir el constructo no es algo mecánico, sino deliberar si es factible de medirse, lo que se facilita a través de una definición operacional de lo que se pretende que el alumno exhiba, y mismo en el que para facilitar la tarea de elaborar el

enunciado se puede tomar mano del dominio cognoscitivo de Bloom, y así tener una clara definición del constructo a medir, y tener mayor seguridad de que el test, medirá realmente el resultado del aprendizaje deseado.

4.1.2 Tabla de especificaciones

Al planear una prueba se debe especificar con detalle los temas y procesos que se incluirán, siendo útil construir una tabla de especificaciones de dos vías, en donde los objetivos conductuales a evaluarse se presentan en el encabezado del renglón y los objetivos del contenido, se presentan como encabezado de columna.

No se debe olvidar que la elaboración de la tabla depende de factores como la naturaleza específica del curso, objetivos logrados en cursos previos, filosofía de la escuela, necesidades especiales del estudiante y factores locales a considerarse del programa instruccional (Gronlund, 1998).

En el cuerpo o casillas de la tabla, se escriben las descripciones de los reactivos que se consideren apropiados para el renglón y columna, (Aiken, 2003). En cada casilla hay un número que representará la cantidad de reactivos importantes para cada una, anotada esta cantidad se sumará cada renglón y se anotará el resultado en la columna a la derecha, que reflejará la importancia para cada contenido. Las columnas también se suman e indicarán el peso que se está dando a cada conducta (Hills, 1982).

Su propósito es asegurar que la prueba medirá una muestra representativa de los resultados de aprendizaje y del contenido de la materia (Gronlund, 1980) y evitar el hecho de medir al azar y el que se mida sólo datos y no el conocimiento del alumno, lo cuál ocurriría si no se realiza una planificación adecuada (Hills, 1982); recordando que entre más alto sea el nivel de cognición, serán reactivos más difíciles de idear. Thorndike (1996) menciona que justamente las taxonomías se usan para expresar los conocimientos a medir, guiando a los elaboradores a desarrollar reactivos que obligan al examinado a hacer más que recordar hechos. El número de ítemes asignados a cada celda está dado por la importancia que se da a cada resultado de aprendizaje y cada área de la materia. La asignación del peso a las áreas de contenido y a los objetivos, se hace al asignar porcentajes a cada área de manera que, el total de porcentajes de

áreas de contenido ascienda a un 100%, al igual que los objetivos del proceso (Thorndike & Hagen, 1973).

No es que se deba tener un número exacto de reactivos o comportamientos para cada casilla, de hecho Hills (1982) menciona que no hay un número correcto de categorías de conductas o contenidos, e incluso de ítem para cada casilla, sólo que los ítem por casilla, conducta y tema deben reflejar la importancia que el maestro asigna a cada elemento; dado que el hacer este boceto del contenido de la prueba, se espera un reflejo de lo enseñado o lo que debía haberse enseñado en un período escolar. Thorndike (1996) menciona que el constructor debe realizar un juicio acerca de cuántos reactivos dedicar a cada aspecto.

Al realizar de manera adecuada la planeación de la prueba facilitará la elaboración de la tabla, puesto que al tener una definición clara de lo que se desea medir, agilizando el proceso de asignación de ítem a las casillas, en donde puede ser que existan diferencias entre el número de ítem disponible y el fijado en el plan original, que como menciona Hills (1982), es un proceso normal. Hills describe que la importancia de planear esta tabla es asegurar la *validez* de la prueba, determinando si mide lo que debe. Un tipo de validez importante aquí es la de *contenido*, en la que expertos deciden si los reactivos cubren los comportamientos y contenidos a evaluar, si son relevantes y la relación entre los temas y comportamientos, a partir de la tabla de especificaciones (para mayor información de validez, consultar el apartado correspondiente).

Thorndike & Hagen (1973) refieren que el número total de reactivos debe ser grande para dar una muestra adecuada de la conducta del alumno en relación con las áreas de contenido y con los objetivos del proceso, que han quedado manifestado en la tabla de especificaciones. En esta parte además del esbozo llevado a cabo en la tabla, se considera el tiempo de que se dispone para la prueba, que dependerá del tipo de ítem utilizado; la edad y el nivel de escolaridad del alumno; su nivel de capacidad, ya que no todos dominan por igual los temas; la longitud y la complejidad del ítem; el objetivo del proceso que se está estimando, que se refiere a los ítems que requieren sólo la identificación de la respuesta correcta, que son más rápidos o aquellos en los

que se requiere la aplicación de un conocimiento a otras situaciones (Thorndike & Hagen, 1973).

Para determinar la cantidad de tiempo de la prueba se usa una regla empírica, en la que Thorndike & Hagen mencionan que un alumno característico necesitará de 30 a 45 segundos para leer y responder a un ítem sencillo de opción múltiple o verdadero – falso y de 75 a 100 segundos para leer y resolver uno de opción múltiple medianamente complejo.

Para establecer el número total de ítems de cada área de contenido y objetivo conductual, se asigna mediante un porcentaje asignado a cada objetivo del proceso y se multiplica por el número total de ítems para determinar el número que deben redactarse para medir cada proceso, de manera que se obtenga un número que se anota en la columna de la extrema derecha del borrador (ver el resumen en la p. 50)

La extensión en una lista depende de los reactivos y naturaleza de la prueba, de manera que en la planeación para describir las tareas de conocimiento que un estudiante ha dominado (pruebas con referencia al criterio) es comprensible elaborar sólo una lista. Una prueba que será usada para un rango de estudiantes en orden de lo que han alcanzado (pruebas con referencia a las normas) se suele incluir especificaciones de la ejecución para clarificar cuales de los estudiantes han alcanzado los resultados previstos (Gronlund, 1998). Para clarificar más el tema, se presenta un ejemplo de lo revisado desde la definición del constructo del hasta la elaboración de la tabla de especificaciones, mismo que se describe a continuación.

- ✦ El primer punto a realizar es identificar el o los objetivos y contenidos
 - ❖ ¿Para qué se aplica la prueba?
 - ❖ ¿Qué aprendizaje exploraremos con ella?

- ✦ Determinar los objetivos de aprendizaje a evaluar en cada materia para poder realizar un desglose en contenidos específicos de cada uno y así asignar un porcentaje de importancia relativa, lo cuál es preferente sea mediante el acuerdo de varios profesores de la asignatura, con objeto de evitar influencias de opiniones personales; la suma de porcentajes debe ser igual al 100%.

- ✦ Calcular el número de reactivos que a cada materia corresponde a través de una correlación entre el peso relativo de cada una y el total de reactivos que se ha pensado contendrá la prueba.

En el ejemplo expuesto, supone una prueba de 80 reactivos, donde se evaluarán 5 materias, la distribución sería la siguiente.

Tabla 4. Ejemplo de la asignación de reactivos para cada materia en una prueba de 80.

Materia	Peso relativo %	Reactivos para cada materia	Ajuste de reactivos
1	20	$20 \cdot 80 = 16.00$	16
2	10	$10 \cdot 80 = 8.00$	8
3	20	$20 \cdot 80 = 16.00$	16
4	30	$30 \cdot 80 = 24.00$	24
5	20	$20 \cdot 80 = 16.00$	16
Total	100		80

- ✦ Se efectuará el desglose de los objetivos de aprendizaje de cada materia en subtemas o contenidos, recurriendo a la clasificación de contenidos de aprendizaje de Bloom del dominio cognoscitivo.

Tabla 5. Ejemplo de la tabla de especificaciones de acuerdo a los lineamientos recomendados para su elaboración, los cuales fueron tomados de Carreño (1993).

Materia	Peso relativo %	# de objetivos de aprendizaje a evaluarse	Conocimiento	Comprensión	Aplicación	Análisis	Síntesis	# total de reactivos
1	20	5	2/8	3/8				16
2	10	4	4/8					8
3	20	4	1/4	2/4	1/8			16
4	30	6	3/10	2/12	1/2			24
5	20	5	3/10	2/6				16
	100%		40	30	10			80

Así para el objetivo de aprendizaje de la materia 1, al que se asigna un peso de 20%, debe ser examinado con 16 reactivos en una prueba de 80; ésta materia incluye 5 objetivos a evaluarse, de los cuales 2 son del nivel de conocimiento y 3 de comprensión, por lo que de los 16 reactivos que le tocan, 8 se redactarán a nivel de

conocimientos y 8 del de comprensión, lo cuál debe percibirse de la misma manera para las demás materias, recordando que para cada nivel se deben emplear ciertos verbos específicos, los cuáles fueron tomados de Gronlund (1998) y que posteriormente se ejemplifica.

Tabla 6. Ilustración de los verbos para definir los objetivos en el dominio cognoscitivo de la taxonomía.

Categoría de la taxonomía	Verbos para cada resultado de aprendizaje específico
Conocimiento	Identificar, nombrar, definir, listar, marcar, seleccionar
Comprensión	Clasificar, explicar, resumir, convertir, predecir, distinguir entre
Aplicación	Demostrar, ordenar, organizar, modificar, resolver, ordenar, relacionar
Análisis	Diferenciar, diagramar, estimar, separar, inferir, ordenar, subdividir
Síntesis	Combinar, crear, formular, designar, componer, construir, revisar
Evaluación	Criticar, comparar, justificar, concluir, discriminar, suponer

Ejemplo del desglose para la materia 1

- Objetivos de aprendizaje en términos de la categoría de la taxonomía

Objetivo 1. Conoce ...
Objetivo 2. Conoce ...

8 ítem

Objetivo 3. Comprende ...
Objetivo 4. Comprende ...
Objetivo 5. Comprende ...

8 ítem

- Redacción de ítems de acuerdo a los verbos correspondientes para cada nivel

Objetivo 1. Conoce ...

1.1 Identifica ...

1.2 Selecciona ...

1.3 Marca ...

Objetivo 2. Conoce ...

2.1 identifica ...

2.2 Lista ...

2.3 Marca ...

2.4 Selecciona ...

2.5 Identifica ...

Una vez que se ha diseñado este bosquejo, se procede a elaborar los reactivos para medir el grado en que los estudiantes han alcanzado los objetivos para cada tema, sin embargo se debe tomar los que sean apropiados a lo que se está midiendo, Aiken (2003) ejemplifica que los reactivos de respuesta corta y de completamiento son adecuados para evaluar el conocimiento de terminología, pero no para habilidades cognitivas de orden superior, como es la síntesis o evaluación, por lo que recomienda se revise la tabla de especificaciones y se considere el costo, nivel al que se medirá cada objetivo de aprendizaje, dado que el tipo de reactivos a emplearse dependerá también del nivel de la taxonomía en la que se planteo cada objetivo.

4.1.3 Preparación de los reactivos

Los reactivos que se elijan usar en la prueba deben de ser adecuados para explorar conductas que den información sobre el constructo sometido a medir, y en tal proceso se debe tener un amplio margen de los mismos. Anstey (1976; en García, 1993) recomienda un mínimo de tres veces el número de reactivos de los que se planeen para el test, a lo que Hogan (2004) coincide mencionando que en la práctica se deben preparar entre dos o tres veces de los necesarios para la prueba final.

Al hablar de reactivos, es útil mencionar las partes que los conforman, dentro de lo que Hogan (2004) describe en primer lugar el *estímulo* al que responde el examinado, mismo que puede ser una pregunta, una frase incompleta, incluso una imagen; el *formato* de respuesta, se refiere a si serán reactivos de construcción de respuesta o de ensayo o si en su caso serán reactivos objetivos de tipo opción múltiple; las *condiciones* de la forma de respuesta al estímulo, tales como el límite de tiempo y la manera de contestar y el dónde; y por último los *procedimientos para calificar* las respuestas, que en el caso de la mayoría de las pruebas de aprovechamiento suelen calificarse como correcto o incorrecto

Una vez elaborado este bosquejo, se elaboran los reactivos para tales fines, para que de esta manera se proceda a realizar su análisis estadístico, pero considerando bajo que supuesto se hará, que bien puede ser por medio de la Teoría clásica o de la Teoría de Respuesta al Ítem (TRI), los cuales se tratarán más adelante.

a) Determinación del tipo de formato a utilizarse

En la elaboración de las pruebas objetivas se ha preferido que el alumno elija su respuesta dentro de varias opciones proporcionadas por el elaborador de la prueba, que suelen ser los reactivos de tipo objetivo, otro tipo que se emplea es donde el alumno aporta su propia respuesta, conocidos como de ensayo.

El tipo de reactivo a usar dependerá del objetivo del proceso a medirse o de la habilidad del constructor para elaborarlos (Thorndike & Hagen, 1973).

Es sabido que en una prueba de ensayo, el número de ítem será reducido por las implicaciones que resultan para su calificaciones y otros factores (Thorndike & Hagen, 1973); mientras que en las pruebas objetivas, se suele preparar más del 20% de reactivos de lo que en realidad se necesita (Aiken, 2003)

El reactivo que se seleccione debe medir el resultado de aprendizaje de manera directa, lo que obliga al uso de un tipo específico de éste, que tenga calidad, dentro de lo cuál el ítem que tiene una alta calidad, lo conforma el de opción múltiple, ya que proporciona una medida más adecuada que los otros tipos y miden una gran variedad de resultados de aprendizaje, desde simples hasta complejos, lo que ocasiona que desempeñen un papel muy importante en la aplicación de test objetivos para resultados del conocimiento (Gronlund, 1985), a lo que Thorndike (1996) también refiere en el hecho de que los de opción múltiple es el más usado debido a la rapidez en la calificación y a la flexibilidad del formato para medir diversidad de contenidos y procesos mentales.

En el capítulo de Test descrito anteriormente se examina con una mayor precisión las características, ventajas y limitaciones de ambos tipos de pruebas y se proporciona una mayor información de los diferentes tipos de reactivos y algunas sugerencias para su elaboración y por consiguiente de su empleo.

b) Estructurar el test

Aiken (2003) recomienda que una vez que se ha elegido el tipo de reactivo a utilizarse y una vez redactados, se hace que los revisen y editen personas

conocedoras, mismas que detecten errores y hagan sugerencias para mejorarlos, en donde antes de conformar la prueba se decida si la longitud que tiene es apropiada para los límites de tiempo, si los reactivos se deben agrupar en las páginas del cuadernillo, si se utilizarán hojas de respuestas, cómo será su reproducción y el tipo de instrucciones que deben incluirse en la prueba.

Por otra parte Hogan (2004) recomienda que en esta edición y revisión de los reactivos, lo primero que se debe hacer es revisar su claridad, la gramática utilizada y que estén acorde con las reglas para la redacción de los mismos, de tal manera que se pueda proceder a hacer una revisión de la corrección del contenido (se hace con los expertos en cada tema) y finalmente la revisión debe buscar posibles sesgos de género, raciales o étnicos, que si es el caso deben hacerlas personas que sean miembros del grupo de referencia al que se alude.

A continuación se da una explicación de cada uno de los aspectos a considerar respecto a los reactivos mencionados por Aiken (2003).

i. Extensión de la prueba

El número de reactivos dependerá del tiempo que se disponga, el grado, nivel de lectura de los examinados, y extensión y dificultad de los reactivos, de manera que aquellos que sean cortos o pidan memorización, se responderán más rápido que los que sean de extensión más larga. Una regla empírica es la de conceder un minuto por cada reactivo de opción múltiple y un minuto por cada dos ítem de tipo verdadero / falso; aunque estos límites de tiempo deberán extenderse o reducirse, dependiendo del grado en el que se examina a los alumnos.

Se puede pedir a los examinados que permanezcan en sus asientos una vez terminada la prueba, de manera que no existan diferencias sustanciales en el tiempo que se necesita para resolverla o que se suscite cierta presión en los examinados al notar que algunos han terminado.

ii. Ordenamiento de los reactivos

Aiken (2003) recomienda ordenar las respuestas de los reactivos de opción múltiple al azar, sin un patrón específico para que no exista preferencia en los examinados al elegir cierto tipo de respuestas. En el caso de los de apareamiento es conveniente que las premisas y opciones se coloquen en la misma página.

En cuanto a la conformación de la prueba como un todo se recomienda agrupar los reactivos del mismo tipo o tema, para ayudar a la preparación, aplicación y calificación de la prueba, recomendando ordenan subconjuntos de reactivos del más fácil al más difícil, lo que asegura el uso de tiempo del examinado, mejora su motivación y su calificación; Karmel (1978) a su vez recomienda ordenar las preguntas de la más sencilla a la más difícil o estableciendo grados de dificultad, como muy difícil, difícil, promedio, fácil, muy fácil, además de procurar el no mezclar distintos formatos.

iii. Hojas de respuesta

Es conocido el que utilizar hojas por separado en las que se marquen las respuestas, facilitan el proceso de calificación, pudiéndose elaborar plantillas o calificarlas con ayuda de un lector óptico, para que el proceso resulte más fortuito y provechoso en cuanto al tiempo.

Aiken (2003), recomienda su uso a partir de los últimos años de la escuela elemental. En dichas hojas los examinados responden colocando en un círculo o en un espacio al lado el número del reactivo el número o letra correspondiente.

Además de las facilidades que proporciona su empleo, se pueden reproducir a través de fotocopiadoras, calificándose manualmente, por computadora, o a través de un lector óptico, que son de gran utilidad cuando se examinan grupos grandes.

iv. Instrucciones

Esta es una parte importante que no debe olvidarse al momento de la conformación final de la prueba, sin embargo se colocan para el ensayo de manera que se pueda cerciorar de que son claras o las fallas que pueden tener para la prueba final,

de manera que al aplicarse a un grupo de personas deben ser colocadas al frente de la prueba y en caso de utilizarse instrucciones específicas para cada parte de la prueba, colocarlas antes de la parte respectiva.

Aiken (2003) recomienda que para no pasar por alto las instrucciones, ponerlas en negrita y que el examinador lea en voz alta las instrucciones a los examinados y en forma clara y lenta. Independientemente si se dan de manera escrita o se lean, las instrucciones deben informar el propósito de la prueba, la forma en que deben indicarse las respuestas, la ayuda que pueden esperar en caso de no entender algo, el tiempo del que disponen para contestar la prueba, la forma en que se calificará y cómo deben corregir una respuesta en caso de cometer un error.

Las instrucciones sirven para que un sujeto sepa claramente lo que debe hacer y que este motivado para hacerlo. Con respecto a la motivación se debe provocar que el sujeto haga un esfuerzo real sin producirle una ansiedad que lo debilite, es un equilibrio difícil por las diferencias de las personas, sin embargo depende mucho del contexto; Thorndike (1996) indica que si se trata de calificar para la admisión a estudios superiores, el motivar al sujeto no es tan difícil, lo que puede ser un problema es la ansiedad, por lo que recomienda que la redacción de las instrucciones tomen en cuenta la reducción de las tensiones, evitando enunciados amenazantes y exagerados, con uno o más reactivos como ejemplos, la manera de responder la prueba, la distribución del tiempo, y la calificación.

Una vez realizado el paso anterior se procede a considerar los reactivos que serán incluidos en la prueba final, a través de procesos determinados para la creación del banco de reactivos, el cuál facilita la conformación final de la prueba.

Un banco de reactivos, es un conjunto de ítems que miden el mismo dominio de conocimientos y para elaborarlo se hace a través de la Teoría de Respuesta al Ítem (TRI), debido a que es por medio de ésta como se logra colocar en una misma escala la estimación de los parámetros, lo que no se puede hacer por medio de la Teoría Clásica de la Prueba (TCP).

Muñiz (1996) cita que para la construcción de un banco de reactivos en base a la TRI se necesita estimar los parámetros de los reactivos y de los individuos, el grado de ajuste de los datos al modelo de respuesta que se elija, para eliminar los reactivos que

muestren desajustes, llevado a cabo a través del análisis de reactivos y con el índice de dificultad del reactivo, para proceder a su integración en un programa computarizado a través de paquetes disponibles.

La TCP toma en consideración el índice de homogeneidad, validez, dificultad del ítem, poder discriminativo y el coeficiente de confiabilidad del test (García, 1993). Los índices de dificultad y discriminación se conocen como estadísticos del reactivo tradicionales (Hogan, 2004).

Es conveniente mencionar los procedimientos de la TCP y TRI, los cuales son empleados a la hora del ensayo de los reactivos, representándose así las técnicas o modelos empleados para la estimación de la dificultad y validez de los reactivos, para proseguir con la selección de aquellos reactivos que conformarán la prueba final.

4.1.4 Ensayo de los reactivos

Los reactivos que se hayan diseñado para medir un rasgo, deben ensayarse, para determinar su dificultad y su capacidad para diferenciar entre los alumnos, para lo cual se debe determinar el nivel de edad o grado en el que se ensayarán los materiales, que deben ser lo más parecidos a la población meta; así como el tamaño de la muestra, que entre más grande sea, más seguras serán las estimaciones, a lo que Thorndike (1996) recomienda que para cada grupo de edad o grado se emplee un mínimo de 200 a 300 reactivos; determinar la extensión de las pruebas de ensayo, y el tiempo que debe darse, ya sea libre para determinarse a través de la ejecución de la prueba final o con límite. Thorndike indica que una vez hecho esto se hacen los arreglos para llevar a cabo el proceso.

Dado lo anterior, el ensayo de los reactivos se hace en base a la teoría clásica y para superar las limitaciones de ésta, se recurre al análisis en base a la Teoría de Respuesta al Ítem (TRI), puesto que no son excluyentes una de la otra, más bien complementarias, donde una de las finalidades de la TRI es el crear un banco de reactivos que facilitará el proceso de la conformación de la prueba final y que supera los problemas planteados por la Teoría Clásica.

4.1.4.1 Teoría clásica de la prueba (TCP)

Dentro de la TCP se maneja un ensayo informal y otro formal, en el informal se manejan entre 5 y 10 individuos a los que se les aplica la prueba con la finalidad de que comenten los reactivos, su redacción, el tiempo de aplicación, posibles dificultades en la calificación, instrucciones, etc., con el objeto de que el elaborador detecte anomalías o posibles problemas en la misma.

En la prueba formal se aplican los reactivos a muestras representativas de la población meta, donde se siguen algunos de los siguientes procedimientos.

- I. Estudio independiente: la finalidad es hacer un análisis de reactivos, en donde los examinados responden a ellos en condiciones parecidas a la prueba final. Se busca que tengan la oportunidad de responder a todos los reactivos, por lo cual la prueba puede variar o no tener límite de tiempo.
- II. Acoplamiento: se incluyen reactivos de prueba en la aplicación de una prueba existente, los cuales se manejan a los examinados como reactivos experimentales, sin indicar cuáles son, al mismo tiempo que no cuentan para la calificación de la prueba; se usan como evaluaciones psicométricas a nivel nacional, para futuras aplicaciones en las cuales se tiene ciertas fechas para su realización.
- III. Análisis continuo: exige el uso reiterado de una prueba o reserva de reactivos al que se suman nuevos reactivos en cada uso, en estos nuevos reactivos se realiza un análisis de reactivos, para su modificación o eliminación y en cuanto a la puntuación puede o no contar para la prueba (Hogan, 2004).

El puntaje obtenido por una persona en una prueba será el número de reactivos que resuelve de manera correcta, en la que la confiabilidad y validez de los datos dependen de las propiedades de los reactivos de forma individual y estos dos elementos serán los que determinan el valor de la prueba, lo que se obtiene a través del análisis de reactivos para satisfacer la demanda de la confiabilidad y la validez de la prueba (Magnusson, 1978).

Dentro de la Teoría Clásica de la Prueba, se incluyen los procedimientos tradicionales de análisis de reactivos, como punto de la prueba formal, que depende de: el índice de dificultad y de discriminación del reactivo.

a) Índice de Dificultad del reactivo (ID)

Se refiere al porcentaje de examinados que responden de manera correcta a un reactivo; autores como Hogan (2004) y Muñiz (2001) mencionan que en realidad la dificultad se refiere a la facilidad del reactivo; debido a que a medida que aumenta el número de personas que lo contestan el reactivo será más fácil. Los niveles de dificultad del reactivo se llaman, *valores p*, donde *p* es la proporción, de manera que si un reactivo tiene un valor *p* de .87 será sencillo, debido a que 87% de los examinados lo respondieron y si es de .23, querrá decir que es muy difícil, ya que sólo un 23 % lo respondieron acertadamente. Muñiz (2001) expresa esto en la siguiente fórmula:

$$ID = A / N$$

A: número de personas que aciertan el reactivo

N: número de sujetos que han intentado resolverlo

Se recalca que la media de la prueba será igual a la suma de los índices de dificultad de los reactivos. Lafourcade (1969) menciona que un buen examen será el que haya sido superado por más de la mitad de los examinados y en el cuadro que se muestra a continuación, señala que los reactivos muy fáciles y los muy difíciles deben ser descartados y analizar las causas de esa facilidad o dificultad.

Tabla 7

Parámetros de aceptación de los índices de dificultad de los reactivos, para ser evaluados, de acuerdo a los datos que se presentan.

INDICE DE DIFICULTAD	DESCRIPCIÓN
Reactivos superados por el 85%	Muy fáciles
Reactivos superados por el 50 al 85%	Relativamente fáciles
Reactivos superados por el 15 al 50%	Relativamente difíciles
Reactivos superados por el 0 al 15%	Difíciles o muy difíciles

Una limitación del índice de dificultad en la TCP, es el no constituir una propiedad interna del reactivo, sino que el valor depende el tipo de sujetos al que se aplique, ya que si muestran ser muy competentes, el reactivo resultará muy fácil o viceversa, por lo que Muñiz (2001) menciona que es de esperar que las propiedades de las pruebas no dependan de los objetos medidos, para lo que una solución adecuada son los modelos de TRI. En el caso de que los reactivos sean de opción múltiple, se muestra una fórmula clásica, para calcular el índice de dificultad, corrigiendo la influencia del azar.

$$ID = \frac{A - E}{(K-1) / N}$$

A = # de sujetos que aciertan el reactivo

E = # de sujetos que fallan el reactivo

K = # de alternativas del reactivo

N = # de sujetos que intentan resolver el reactivo

Thorndike (1996) indica que deben eliminarse los reactivos que no son útiles y aquellos que muestre una discriminación inversa o negativa, debido a que el uso del índice de discriminación es excluir a los reactivos defectuosos; este tema de la discriminación de los reactivos se trata a continuación, al mismo tiempo que se explica la manera de obtenerlo y los parámetros que se deben considerar para la elección de los reactivos que conformarán la prueba final.

b) Discriminación del reactivo

Se refiere a la capacidad de un reactivo para diferenciar estadísticamente entre los grupos de examinados, de aquellos que poseen más cierto conocimiento de aquellos que poseen menos, dicho de otra manera, de aquellos alumnos que puntúan alto y bajo en la prueba, lo que significará que tales reactivos constituirán un buen instrumento, ya que identificaron a los grupos que poseen más o menos del mismo atributo, y éste proceso puede realizarse a través del método externo o interno.

⬇ Método externo: se identifica al grupo de manera externa a la prueba, y se espera que los reactivos discriminen o diferencien entre dos grupos, pero el método más empleado es el interno, por lo cuál se explicará más a fondo.

⬇ Método interno: la identificación del grupo se encuentra dentro del mismo instrumento, se usa para crear grupos que tienen más o menos del rasgo que se mide, para lo que se califica toda la prueba, se identifica a los que tuvieron una puntuación alta de los que obtuvieron una baja y se determina hasta que punto un reactivo diferencia entre ambas calificaciones. En este método una de las divisiones que se emplea es comenzar con las distribuciones de las puntuaciones en la prueba y definir a los grupos "altos" y "bajos" como la mitad superior e inferior de esta distribución; otra división es utilizar el 27% de la parte superior e inferior, donde Truman & Kelley (1939; en Lafourcade, 1969) demostraron que en la comparación de estos grupos se desea averiguar la correlación del reactivo con el puntaje total.

El reactivo debe poseer cierto poder discriminativo, y deberá ser en positivo, dado que dicho índice de discriminación será el índice de validez del reactivo; así el mejor reactivo será aquél cuya correlación con el puntaje total sea el más alto. De manera que si una prueba separa a los examinados en diversos niveles de rendimiento, se puede decir que es un instrumento con muy buen índice de discriminación (Lafourcade, 1969).

El grado de discriminación se presenta con D significa diferencia o discriminación, o r , que es la correlación entre el desempeño en el reactivo y el criterio externo o puntuación total en la prueba; las r en el uso común incluyen la r biserial (r_{bis}) y la r biserial puntual (r_{pbis}), que son considerados métodos para determinar el índice de discriminación o de validez del reactivo (Hogan, 2004), donde para su aceptación o eliminación, se recomienda ver la tabla 7 y 8.

Tabla 8

Valores presentados para realizar evaluaciones de los reactivos, de acuerdo al índice de discriminación, tomado de Lafourcade (1969).

INDICE DE DISCRIMINACIÓN	
.40 y más	Muy buenos reactivos
.30 a .39	Razonablemente buenos, sujetos a mejorarse
.20 a .29	Regulares
Menos de .19	Deficientes, deben descartarse

El tipo de correlación que se emplee dependerá de las características de las variables, que en éste caso se refiere al reactivo y a la prueba. Muñiz (2001) menciona algunas descripciones de éstos coeficientes de correlación.

b.1. Correlación biserial- puntual (r_{bp})

Se refiere a una aplicación de la correlación de Pearson, es una medida de asociación entre dos variables continuas (representa características sin saltos, ni rupturas, tal como la edad cronológica de una persona y se presenta por números fraccionarios y enteros), de manera que una de las variables es dicotomizada (sí-no; falso-verdadero; entre otros), suponiéndose así discreta o discontinua (representa rupturas y se presenta en números enteros) para poder relacionarse con la que sigue continua; para lo cuál se emplea el coeficiente de correlación biserial puntual (r_{bp}), que también suele emplearse para determinar la correlación entre el reactivo y la calificación total en la prueba y así poder determinar el índice de confiabilidad. La utilización más común de esta correlación, se encuentra en la construcción de pruebas, más en la determinación de la validez (Padua, 1978), cuya formula se presenta a continuación:

$$r_{rbp} = \frac{n(\sum fx) - n_c(\sum x)}{\sqrt{(n_c n_i) [n \sum fx^2 - (\sum fx)^2]}}$$

donde:

n= número de sujetos.

f= frecuencia de ocurrencia de las puntuaciones obtenidas por los sujetos

x= puntuaciones obtenidas por los sujetos

fx = producto de la frecuencia por las puntuaciones

f_c = número de sujetos que obtuvieron exactamente las puntuaciones x

$n_c = \sum f_c$ = total de sujetos que obtuvieron exactamente las puntuaciones x

f_i = número de sujetos que no obtuvieron exactamente las puntuaciones x

$n_i = \sum f_i$ = número total de sujetos que no obtuvieron las puntuaciones x

fcx = producto del número de sujetos que obtuvieron exactamente las puntuaciones x por la puntuación x de cada sujeto representativo

$\sum fx^2$ = suma de los productos de la frecuencia absoluta por las puntuaciones x al cuadrado

$(\sum fx)^2$ = cuadrado de la suma del producto de la frecuencia por las puntuaciones x .

Para comprender un poco más la aplicación de esta correlación, se muestra un ejemplo, el cuál fue tomado de Elorza (2001, p.613)

A 100 estudiantes de una escuela preparatoria se les aplica una prueba de conocimientos x , que consta de 40 reactivos, así como la frecuencia con que se distribuyen los 100 estudiantes.

Tabla 9

Resultados obtenidos por los 100 estudiantes en la prueba de conocimientos, en la que se presenta el número de reactivos y la frecuencia de respuesta en cada rango de ellos.

x	f
40	2
38	4
37	6
36	12
31	10
30	12
28	10
27	10
25	4
24	4
22	3
20	3
18	3
16	2
12	2
10	1
$n =$	100

Al elaborar la prueba se hizo hincapié en el reactivo número 33; el criterio para evaluarla será incorrecto o correcto. Se calculará el coeficiente r_{bp} mediante las puntuaciones obtenidas por los 100 estudiantes en la prueba x y la respuesta al reactivo 33, que se dicotomizó como "incorrecto-correcto".

1. Se pasa en un cuadro la información sobre los resultados de la prueba y la respuesta dada al reactivo 33, lo que se muestra en la tabla 8.

Tabla 10

Resultados obtenidos en la prueba y la respuesta dada al reactivo 33, las puntuaciones obtenidas (1), su frecuencia (2), el número de estudiantes que respondieron correctamente al reactivo (3) y resta de las puntuaciones de la columna 3 a la 2, que resulta en la frecuencia de estudiantes que responden incorrectamente al reactivo 33.

(1)	(2)	(3)	(4)
x	f	fc	f_i
40	2	2	0
38	4	4	0
37	6	5	1
36	12	10	2
32	12	9	3
31	10	8	2
30	12	7	5
28	10	6	4
27	10	7	3
25	4	1	3
24	4	1	3
22	3	1	3
20	3	1	2
18	3	0	3
16	2	1	1
12	2	0	2
10	1	0	1

2. Se obtiene los productos de los valores de la columna (1) multiplicada por la (2), es decir fx , lo cual forma la columna (5). Por último, se suman estos resultados para obtener $\sum fx$.

Tabla 11

Valores correspondientes al producto de la columna 1 por la 2.

(5)	(5)
(fx)	(fx)
80	100
152	96
222	66
432	60
384	54
310	32
360	24
280	10
270	fx= 2 932

- Se eleva al cuadrado las puntuaciones x de la columna 1 y se multiplica por cada valor f de la columna 2; lo que origina la columna (6). Por último, se suman todos los valores y se obtiene $\sum fx^2$.
- Se obtiene la columna (7) al multiplicar las puntuaciones x de la columna 1 con los valores fc de la columna (3); después se suman todos estos resultados y se obtiene $\sum fcx$.

Tabla 12

Valores correspondientes a cada uno de los elementos que conforman la fórmula de la correlación biserial – puntual, para ser sustituidos en la misma.

(1)	(2)	(3)	(4)	(5)	(6)	(7)
x	f	Fc	f_i	fx	Fx^2	Fcx
40	2	2	0	80	3200	80
38	4	4	0	152	5776	152
37	6	5	1	222	8214	185
(1)	(2)	(3)	(4)	(5)	(6)	(7)
36	12	10	2	432	15552	360
32	12	9	3	384	12288	288
31	10	8	2	310	9610	248
30	12	7	5	360	10800	210
28	10	6	4	280	7840	168
27	10	7	3	270	7290	189

25	4	1	3	100	2500	25
24	4	1	3	96	2304	24
22	3	1	2	66	1452	22
20	3	1	2	60	1200	20
18	3	0	3	54	972	0
16	2	1	1	32	512	16
12	2	0	2	24	288	0
10	1	0	1	10	100	0
	100	63	37	293	89898	1987
	Σf	Nc	n_i	Σfx	Σfx^2	Σfcx

5. Los resultados obtenidos en la tabla anterior se sustituyen en la fórmula total.

$$r_{rbp} = \frac{n(\sum fcx) - nc(\sum fx)}{\sqrt{(nc \cdot n_i)[n \sum fx^2 - (\sum fx)^2]}}$$

$$r_{rbp} = \frac{(100)(1987) - 63(2932)}{\sqrt{(63)(37)[100(89898) - (2932)^2]}}$$

$$r_{rbp} = 13984 / \sqrt{(2331)(393176)}$$

$$r_{rbp} = 13984 / 30274 = 0.46$$

$$r_{rbp} = 0.46$$

Sin embargo en la actualidad se puede evitar realizar esta fórmula de manera manual, empleando el programa de computadora SPSS en el cuál se pueden ingresar los datos, de acuerdo a lo que se planea realizar y el programa efectuará la operación, provocando que solamente se proceda a la interpretación de los datos.

b.2. Correlación biserial (p_b)

Se usa cuando una de las variables es medida a nivel nominal, es decir que emplea números como parte de objetos, como designar 0 para femenino y 1 para masculino y la otra variable se presenta a nivel intervalar, en la que se establece que las distancias entre número y número son iguales como pueden ser incrementos de cada 5 números, partiendo de 0 a 5 ó de 5 a 10. En esta correlación se asume que la variable nominal puede ser dicotomizada y que subyace una variable continua en la curva normal (Padua, 1978), y se emplea la siguiente fórmula.

$$p_b = \frac{\mu_p - \mu_x}{\sigma_x} \cdot p / y$$

μ_p = media en la prueba de los sujetos que aciertan el reactivo

μ_x = media de la prueba

σ_x = desviación típica de la prueba

p = proporción de sujetos que aciertan el reactivo

y = ordenada correspondiente al valor de la puntuación típica de la curva normal que deja por debajo un área igual a p (en tablas correspondientes)

Magnusson (1978) menciona que supone que los puntajes de los individuos en variables psicológicas tienden a una distribución normal, asumiendo una dicotomía porque la prueba consiste en reactivos que se califican con 1 si es correcta y 0 si la respuesta es incorrecta, lo que provoca una división en dos de la curva, los individuos que están en la parte superior de ejecución se colocan en una categoría y los que se encuentran debajo de ese nivel se colocan en la otra.

b.3. Coeficiente *phi* (ϕ)

Se usa cuando ambas variables a correlacionar son dicótomicas. El coeficiente *phi* (ϕ) es una aplicación del coeficiente de correlación de Pearson (Muñiz, 2001).

b.4. Correlación tetracórica

Se usa cuando la prueba y el reactivo son variables dicotomizadas y ambas se asumen con una distribución normal (Muñiz, 2001).

En las correlaciones biserial y biserial- puntual se uso para ver si los reactivos miden la misma variable que los demás en la prueba, sin embargo en la tetracórica se estima el grado en que reactivos diferentes miden la misma variable, en donde el coeficiente da una estimación de la correlación momento – producto entre las distribuciones de los puntajes dentro de las que se supone que los reactivos han diferenciado (Magnusson, 1978).

Otro índice que Muñiz menciona (2001) es el índice basado en la proporción de aciertos, el cuál no se basa en la correlación del reactivo con la prueba sino en el hecho de la diferencia entre los sujetos que puntúan en la prueba por encima de la mediana y que aciertan en la prueba (competentes) y la proporción de sujetos que puntúan por

debajo de la mediana y también aciertan en el reactivo (incompetentes), en donde se utilizan grupos con criterio externo (27%), lo que se refiere en la formula siguiente.

$$d = P_c - P_i, \text{ donde:}$$

P_c = sujetos competentes que aciertan el reactivo en la prueba

P_i = sujetos incompetentes que también aciertan el reactivo en la prueba

c) Análisis de distractores

El análisis de reactivos también se emplea para el análisis de distractores, el cual es una opción no correcta o no preferida en un reactivo y se aplica cuando el reactivo cuenta con más de dos opciones; éste análisis puede generar información sobre la comprensión o dudas en un reactivo, ya que una opción incorrecta podría comportarse en realidad como una correcta; lo que significaría que más personas del grupo alto eligen esa opción (Hogan, 2004).

De los reactivos que quedan, los que correlacionan alto con la puntuación total, son los mejores (superiores a 0.20), si en ese nivel hay más reactivos de los necesarios, se procede a investigar la confiabilidad, pero si la cantidad de reactivos es menor, se deben elaborar más reactivos, empezando con el proceso para lograr una cantidad más amplia. De manera que para la confiabilidad, se ordenen los reactivos por jerarquías de acuerdo con las puntuaciones totales, se seleccionan las series sucesivas y se calcula el coeficiente de confiabilidad para cada conjunto, en donde por lo regular se requiere de 30 reactivos para tener una alta confiabilidad; de manera que el siguiente paso es la determinación de la distribución de frecuencias, la cuál si es satisfactoria, la selección de reactivos se completa, sin dejar de lado las inclinación de las frecuencias en la curva normal.

4.1.4.2 En cuanto a la teoría de respuesta al ítem (TRI)

En la sección anterior se revisaron estadísticos de reactivo tradicionales en cuanto a la TCP, en la TRI también se llevan a cabo, sólo que son procedimientos un tanto diferentes e ideados para la mejor confección del proceso.

En la Teoría de Respuesta al Ítem (TRI) se toman los criterios mencionados por Hambleton & Swaminathan, (1985; en García, 1993), en donde se considera la dificultad del ítem, discriminación, bondad de ajuste del ítem al modelo y el grado de la función de información del test, esto por la relación que tiene con el número de ítem. El rasgo fundamental de la TRI es la Curva Característica del Ítem (CCI), la cuál será del tipo "S" y para definir las se debe tomar en cuenta los tres parámetros que se mencionan más adelante; la CCI relaciona el desempeño en un reactivo o probabilidad de aprobarlo con la condición en el rasgo, lo que indica que los modelos asumen una relación funcional en el valor de la variable medida por el reactivo y la probabilidad de acertarlo, lo cuál dependerá de los valores de dicha variable (Muñiz, 1997).

De manera que la CCI será una función matemática que indique la relación entre las puntuaciones de las personas en la variable medida (prueba) y la probabilidad que tienen de aprobar el reactivo, lo que es visto como una ventaja sobre el enfoque clásico, en el que no se establece esta conexión.

La TRI proporciona un marco teórico apropiado para trabajar con los bancos de reactivos y proporciona ventajas para su aplicación en el hecho de que muestra flexibilidad en el campo de la educación, proporciona un uso eficiente de las respuestas de los sujetos a los reactivos y que el hecho de contar con el banco de reactivos permite construir pruebas de gran calidad, debido a que los reactivos que se incluyen en el banco fueron depurados por su falta de utilidad.

La construcción del banco de reactivos dentro de la TRI implica determinar la invariabilidad de los parámetros de los reactivos y de los sujetos, estudiar el grado de ajuste de los datos al modelo de respuesta elegido, para eliminar los que muestren desajustes; una vez realizado esto, aquellos con índices de discriminación bajos, demasiado altos e incluso negativos, para el grupo al que se destinará la prueba y una vez que los reactivos seleccionados se calibran en una escala común, se procede a integrarlos en un sistema computarizado para conformar el banco de reactivos que constituirá la prueba (Muñiz, 1996).

El rendimiento en el reactivo es visto como la posibilidad de aprobarlo, o dar la respuesta correcta a la prueba. La construcción del banco de reactivos involucra que estén estructurados en cuanto a contenidos y que sean fácilmente identificados, esto es

que los reactivos del banco deben medir el mismo dominio de conocimientos. Se debe disponer de una o varias muestras a las que se aplique el conjunto de reactivos y de un modelo psicométrico para la calibración de los reactivos, identificando los de baja calidad para eliminarlos, de manera que las estimaciones de los parámetros de los reactivos deben estar dentro de la misma escala (Muñiz, 1996). De manera que los modelos de TRI constituyen un intento de superar el enfoque clásico y mejorar así la medición, lo que se logrará una vez que los modelos partan de supuestos más restrictivos y fuertes, en lo que la CCI, la unidimensionalidad y la independencia local serán supuestos claves en la TRI.

Entre los modelos que tiene más utilización en la TRI, se encuentra el **parámetro de dificultad** del reactivo que refiere al punto en que la CCI atraviesa la marca del 50% de posibilidad de aprobar el reactivo, que es similar al índice de dificultad del reactivo en la TCP, sin embargo el parámetro de dificultad en la TRI se representa con el valor θ ; también se maneja el término de **pendiente** de la curva del reactivo, la cuál indica la diferencia del reactivo entre las personas que presentan capacidades distintas, los cuales son los valores θ , de manera tal que la pendiente de la CCI sería el índice de discriminación del reactivo (Hogan, 2004).

En la TRI el parámetro de pendiente o índice de discriminación ("a"), índice de dificultad ("b") y el de adivinación ("c"), originan los modelos IP, 2P, y 3P (parámetros). El modelo de un parámetro (1P) sólo considera la dificultad (b), asumiendo que todos los reactivos tienen la misma pendiente, sin considerar la adivinación, dentro de éste, el modelo más utilizado es el de Rasch, aunque precisamente éste modelo resultaría de cierta manera falso en su premisa, ya que al elaborar la prueba, se notaría que no todos los reactivos tienen el mismo poder de discriminar. En el de 2P considera la dificultad (b) y la discriminación (a), aunque no la adivinación y finalmente en el de 3P, involucra los tres parámetros "a", "b" y "c". Muñiz (1997) indica que el parámetro a, indica que cuanto mayor sea la pendiente, mayor será el índice de discriminación y aunque el nombre alude al enfoque clásico, su valor es diferente, de manera que si el valor de "a" cuando θ (nivel de habilidad) se distribuye en la curva normal con media de 0 y desviación de 1 $N(0,1)$, sin aciertos al azar ($c = 0$), dará la correlación biserial reactivo-prueba, es decir, el índice de discriminación en el modelo de la Teoría Clásica.

El parámetro b , es el valor de θ del punto de máxima pendiente de la CCI. En tanto que el parámetro “ c ” está dado por la probabilidad de acertar al reactivo cuando no se sabe nada, lo cuál en la Teoría Clásica sería la probabilidad de acertar al reactivo por azar; de manera que cuando se definen estos tres parámetros, se define la CCI, indicando que entre más se ubique a la derecha en el eje de las abscisas, indicará que los reactivos son más difíciles, pues “ b ” aumenta y el poder discriminativo “ a ” se indicará por las pendientes de CCI.

El marco teórico más empleado es la TRI en cuanto a las pruebas de aprovechamiento y personalidad, considerado el más adecuado para la construcción de un banco de reactivos, para el cuál todos los reactivos del banco deben medir el mismo dominio de conocimiento (Van Der Linder & Eggen, 1986; Reckase, Ackerman & Carlson, 1988; en Muñiz, 1996), sin embargo en la práctica se debe elegir algún modelo, estimar los parámetros de los reactivos, la competencia de las personas (θ) y comprobar el ajuste de los datos al modelo, para lo que deben haberse realizado ya algunos pasos como:

- Definición de la variable a medir.
- Elaboración de los reactivos.
- Aplicación a muestras representativas de la población, obteniendo los índices de los reactivos del enfoque clásico.
- Comprobar los supuestos de la TRI.
- Elegir uno de los modelos de la TRI.

De manera que éstos dos últimos puntos se describirán en esta parte, donde tras los datos obtenidos de la aplicación de los reactivos se evalúa el ajuste al modelo seleccionado y que estimaciones de los parámetros de los reactivos se encuentran en la misma escala; Muñiz (1996) indica que esto implica tres fases, antes de la conformación final del banco de reactivos.

- a) Elección del diseño
- b) Evaluación del ajuste de datos al modelo
- c) Calibración de los reactivos

a) Elección del diseño

Se refiere a la forma de aplicación de los reactivos de acuerdo al número y a las muestras de que se dispone, lo que permite poner las estimaciones de los parámetros en una escala común y de esa manera poder comparar las estimaciones. Los diseños que se emplean son de un solo grupo, grupos equivalentes o de anclaje.

a.1) De un solo grupo

Si la cantidad de reactivos es poco numerosa, se necesitará de una sola muestra, por lo que se le suele llamar de un solo grupo.

a.2) Dos grupos equivalentes

Si se dispone de un gran número de reactivos, de éste conjunto total se divide a la vez en dos conjuntos de reactivos, de manera que para éste tipo de diseño se emplean dos muestras aleatorias de una misma población, de tal forma que a cada una de ellas, se le aplica uno de los dos conjuntos de reactivos.

a.3) De anclaje

También se emplea cuando se dispone de un gran conjunto de reactivos y se habla de anclaje de reactivos o de sujetos, la elección de uno u otro dependerá de las circunstancias. En el caso de anclaje de reactivos, se utiliza para la construcción del banco de reactivos, en este caso se necesitará de más de una muestra, donde cada una responderá a un conjunto de reactivos, pero a la vez a algunos reactivos comunes que fungirán con el papel de reactivos de anclaje, éstos reactivos de anclaje deberán ser de al menos 15 o 25 (Vale, Maurelli, Gialluca Weiss & Ree, 1981; en Muñiz, 1996), se consideraba emplear 5 reactivos, pero son demasiado pocos. En éste caso para estimar sus parámetros se hace una comparación para determinar la ecuación que ayude a ubicar todos los reactivos en una escala común, la ventaja de éste es que no se necesita tomar la equivalencia de los grupos o del conjunto de reactivos.

En el caso de anclaje de sujetos, del conjunto de reactivos se subdividen, para que sean aplicados a dos o más muestras de sujetos, y en éste caso se debe disponer de un grupo de sujetos que serán los de anclaje, que de acuerdo a Vale, Maurelli, Gialluca Weiss & Ree (1981; en Muñiz, 1996) el grupo de anclaje de sujetos debe ser

de al menos 30; mismos que deberán responder a un conjunto completo de reactivos; una vez hecho esto se comparan las estimaciones del parámetro de habilidad de cada sujeto del grupo de anclaje, de manera que esto permita encontrar la ecuación que ayude a situar los parámetros de todos los reactivos en una escala común, la ventaja en este diseño es igual a la de anclaje de reactivos, pero presenta una desventaja que se refiere a la disposición que deben tener el grupo de personas que fungen como de anclaje, debido a que deben de contestar todo el conjunto de reactivos.

d) Evaluación del ajuste de datos al modelo de la TRI

En esta fase se verifican los datos tras la aplicación de los reactivos que constituirán el banco de reactivos, para observar si se ajustan a alguno de los modelos de la TRI, lo que permitirá interpretar las puntuaciones de las pruebas dependiendo del ajuste de los datos al modelo que se haya seleccionado.

Los modelos que de acuerdo a Muñiz (1996) son los más utilizados corresponden a los de 1, 2, y 3 parámetros, por el ajuste que representa cuando los reactivos puntúan de manera dicotómica.

A la hora de elegir un modelo se debe investigar si los datos cumplen con el a) supuesto de unidimensionalidad, b) ausencia de velocidad en la administración de la prueba, c) inexistencia de aciertos al azar e igualdad en los índices de discriminación. Los modelos de 1, 2 y 3 parámetros asumen el supuesto de unidimensionalidad y ausencia de velocidad en la administración de la prueba; el de inexistencia de aciertos al azar es asumido por el modelo logístico de 1 y 2 parámetros y el de igualdad en los índices de discriminación por el de 1 parámetro

b.1) Unidimensionalidad

Es un supuesto de los más empleados y se considera básico en la Teoría de la medición (Lumsden, 1976; Hattie, 1984, 1985; Reckase, 1985; Cuesta & Muñiz, 1994; en Muñiz, 1996). Lo que se busca con éste modelo es garantizar que el conjunto de reactivos que formarán parte del banco, miden una única área de conocimiento o

habilidad. Si el modelo es correcto, la probabilidad de acertar el reactivo sólo dependerá de θ (nivel de habilidad), lo que indica que los reactivos que se destinan a medir la variable θ constituye una dimensión, es decir son unidimensionales.

La importancia de que una prueba cumpla con este supuesto se refiere a la necesidad de que las respuestas de los sujetos, en una prueba que quiera medir el nivel que presentan éstos sujetos en una variable determinada, no se encuentre contaminada por el nivel que éstos posean en otra variable y por el hecho de que si una prueba mide más de un rasgo y se está utilizando para establecer diferencias individuales, puede ser que dos sujetos obtengan la misma calificación por caminos diferentes, lo que dificultaría la comparación entre ellos a la hora de las interpretaciones (Stout, 1987; en Muñiz, 1996).

Muñiz (1997) indica que para comprobar que el conjunto de reactivos representa una sola dimensión es a través del análisis factorial, en el que entre más varianza indique el primer factor, más unidimensionalidad existirá, lo que se verá como cuestión de grado. De manera que si se cumple este supuesto, indicará que existe independencia local entre los reactivos, lo que indicará que la respuesta a uno de los ítems no estará influida por sus respuestas en los otros. Este punto es resaltante en el hecho de que en la TRI no es posible elaborar reactivos encadenados, debido a que de lo contrario indicaría que carece de independencia local, dado que el rendimiento de un sujeto depende del rendimiento de los otros.

De manera que en la unidimensionalidad la puntuación de una persona en la prueba sólo depende del nivel que tenga en la variable medida, mientras que la independencia local indicará que el modelo asume que las respuestas de las personas a un reactivo son independientes de las respuestas a otros reactivos (Muñiz, 1997).

b.2) Ausencia de velocidad en la administración de la prueba

Se refiere a las condiciones de aplicación de la prueba, en el hecho de que no se haya realizado en condiciones de velocidad, de manera que en el caso de que los sujetos hayan dejado sin responder algunos reactivos, no sea debido a la falta de tiempo, sino al nivel que presentan en el rasgo que se evalúa con los reactivos. Aunque de acuerdo a Hambleton y Swaminathan (1985, en Muñiz, 1996) éste es un supuesto

que no suele evaluarse, debido a que se encuentra implícito en el de unidimensionalidad, por el hecho de que si la aplicación del conjunto de reactivos se diera en condiciones de velocidad, el rendimiento y la actuación de los sujetos se vería afectada, debido a la habilidad evaluada y al factor de rapidez, a lo que Cronbach & Warrington (1951; en Muñiz, 1996) recomiendan analizar las relaciones que existan entre las puntuaciones de los sujetos que contestan la prueba con tiempo limitado y las que se obtienen cuando la prueba se aplica con tiempo libre.

b.3) Inexistencia de aciertos al azar

Como su nombre lo indica, en este supuesto se asume que las respuestas de los sujetos no son debidas al azar, éste es asumido por el modelo de 3 parámetros (3PL) dada la posibilidad que un sujeto con un mínimo de habilidad o conocimiento responda correctamente un reactivo por influencia del azar, en cuanto al modelo de 1PL y 2PL implican la satisfacción de éste supuesto.

Otra forma de evaluar el supuesto es por medio de las respuestas de los sujetos que se encuentran en la parte más baja de la escala del conocimiento de habilidad, ya que tienen una baja probabilidad de responder correctamente a los reactivos, lo que permite analizar su actuación en los reactivos más difíciles, de manera que si la probabilidad de acierto en este tipo de reactivos es nula, permite concluir que los datos que se obtengan satisfacen éste supuesto.

b.4) Igualdad en los índices de discriminación constante

Los modelos de 2PL y 3PL asumen una discriminación variable en los reactivos, sólo el modelo de 1 parámetro (1PL) asume éste supuesto. Para evaluar el supuesto, Hambleton, Swaminatham & Rogers (1991; en Muñiz, 1996) proponen analizar la variabilidad en la correlación del reactivo con la prueba, si ésta es pequeña se acepta la verificación del supuesto. Otra manera es verificar el porcentaje de reactivos, en los que la correlación del reactivo con la prueba exceda el intervalo de la correlación media del reactivo – test; de tal manera que para verificar el supuesto deberá disminuir el porcentaje de reactivos en donde la correlación reactivo – prueba caiga fuera de este intervalo definido.

Una vez que se ha elegido el modelo en base a las evidencias que se tienen para la evaluación de sus supuestos, se deben comprobar las ventajas de su utilización, que dependerán del grado de ajuste de los datos al modelo elegido.

Muñiz (1997) indica que se puede hacer una elección antes para elegir el modelo más adecuado, que será dado por las características de los reactivos, de manera que si se sospecha que tras la obtención de los índices del enfoque clásico, los de discriminación no son iguales, es poco razonable emplear el de IP; así mismo si C es diferente a 0 el modelo de 2P, tampoco es plausible, también si son de opción múltiple pues tienen cierta probabilidad de acertar al azar.

Hulin, Drasgow & Pearson (1983; en Muñiz, 1996) mencionan otro camino para evaluar el ajuste, el cuál es a través de las curvas características de los reactivos, en la cuál una diferencia grande en la curva que se obtiene empíricamente y la definida por el modelo, indicará que no fue el más adecuado.

Otro método consiste en comparar las puntuaciones que se observan y las que se pronosticaron, por medio de la Función Característica de la Prueba (FCP) o del reactivo (FCR). La primera da para cada nivel de la escala de habilidad o del conocimiento, una puntuación verdadera que obtendría en la prueba un sujeto con ese nivel de habilidad, lo que hace que las puntuaciones verdaderas se encuentran en la misma escala que las observadas, lo que facilita que la comparación entre ellas sea de manera directa. La FCR, da para cada nivel de la escala de habilidad, la posibilidad de que con ese nivel de habilidad, el sujeto responda de manera acertada al reactivo; si para cada reactivo se realizan las distribuciones de las puntuaciones estimadas y las pronosticadas, se pueden analizar mediante la Chi – cuadrada (Muñiz, 1996).

En esta fase de evaluación del ajuste de los datos al modelo elegido, permitirá la eliminación de los reactivos iniciales del banco, debido a que en el camino se habrán depurado aquellos que mostrasen desajuste al modelo.

e) Calibración de reactivos

Tras la evaluación del ajuste de los datos, se habrán eliminado del conjunto inicial aquellos reactivos del conjunto inicial que no cumplan con los parámetros del modelo seleccionado, o que no mostraban este ajuste, de manera tal que la siguiente

fase es la referente a la calibración de los reactivos, la cuál cumple dos funciones: la primera se refiere a la estimación de los parámetros de los reactivos y la segunda al hecho de convertir estas estimaciones a una misma escala, de forma que las respuestas dadas por los sujetos a los diferentes subconjunto de reactivos puedan ser comparadas.

c.1) Estimaciones de los parámetros

Una vez que se ha elegido el modelo se procede ha estimar los parámetros de los reactivos y la competencia de cada persona en la variable medida a partir de los datos obtenido al aplicar los reactivos a una muestra amplia de sujetos, lo que será antes que la estimación, en donde para elegir los valores de los parámetros, se toman aquellos que maximicen la probabilidad de que ocurran los datos, que se han dado en las respuestas de los examinados (Muñiz, 1997)

Estas estimaciones se hacen por aproximaciones sucesivas y su cálculo es laborioso, por lo que se prefiere hacerlo a través de ciertos programas de ordenador, de manera que para el modelo de 1PL se utilizan programas como el BICAL (Wright, Mead & Bell, 1979; en Muñiz, 1996) y el METRIX (Renom, 1992; en Muñiz, 1996). En cuanto a los modelos logísticos de 1PL, 2PL y 3PL, se utiliza el BILOG (Mislevy & Bock, 1984; en Muñiz, 1996) y LOGIST (Wingersky, Barton & Lord, 1982; en Muñiz, 1996), lo cuál puede realizarse con el apoyo de una persona experta en estadística.

c.2) Equiparación de las estimaciones

Esta fase implica hacer correcciones en los parámetros de los reactivos, donde la equiparación depende del modelo elegido. Si el modelo fue de un solo grupo, donde los sujetos respondieron a todos los reactivos, no es necesario llevar a cabo este proceso, debido a que las estimaciones de los parámetros se encuentran ya en la misma escala.

En el caso de que el número inicial de reactivos haya sido muy amplio, requirió de otro tipo de modelo, se deben llevar a cabo estas estimaciones de los parámetros, para que los datos se sitúen en una métrica común.

En el modelo de anclaje de reactivos, la estimación de los parámetros se hará para cada una de las muestras, pero de forma independiente y después se realizará la

equiparación o transformación a partir de las estimaciones obtenidas de los reactivos comunes; en el caso de ser un diseño de anclaje de sujetos, el proceso de equiparación se realizará a partir de las estimaciones de la habilidad de los sujetos (θ) obtenidas por los sujetos comunes.

Muñiz (1996) reduce el problema de la equiparación a determinar la relación de los valores entre la pendiente de la recta (β) y la ordenada al origen (α), para lo que existen ecuaciones que permiten llevar a cabo el proceso de equiparación en un diseño de anclaje de reactivos o sujetos.

- i. Dificultad $b_y = \beta (b_x) + \alpha$
- ii. Discriminación $a_y = a_x / \alpha$

En donde para realizar los cálculos de β y α , Hambleton & Swaminathan (1985; en Muñiz, 1996) recomiendan los de regresión, media y desviación típica, media y desviación típica robustas y los métodos de la curva característica.

f) Creación del banco de reactivos

En esta etapa se habrán estimado los parámetros de los reactivos en función del diseño elegido y equiparado en una métrica común, para proceder a la creación del banco de reactivos, de manera organizada y estructurada y ser utilizada para la construcción de la prueba; con información acerca del contenido del reactivo tal y como aparecerá en la prueba, en segundo lugar la información que permita acceder al reactivo, como su contenido, el área curricular al que pertenece, lugar que ocupa en el banco u otra información que sea útil. Otras características de los reactivos pueden ser el número de opciones de respuesta, la respuesta correcta, y sus características psicométricas, como los parámetros estimados, dependiendo del modelo, la muestra que se utilizó en la calibración, el valor de la escala de habilidad, número de veces que ha sido utilizado el reactivo, así como la fecha de la última vez de utilización de cada reactivo (Muñiz, 1996).

Para finalizar esta parte se hace un breve resumen de las limitaciones que se tienen en la TC y se presenta una tabla para tener una clara concepción de los dos enfoques antes mencionados, se ejemplifican los elementos de la TC y de la TRI, a través de una tabla comparativa para cada uno de ellos, con la finalidad de observar algunas de sus directrices.

Limitaciones de la Teoría Clásica:

- Parte de un modelo lineal simple, en el que se asume que las puntuaciones de los sujetos pueden verse afectadas por un error aleatorio, situado en el sujeto, el ambiente, instrumento o en el proceso de medición.
- Invariabilidad en la medición referente a la prueba empleada, debido a que los resultados no se encuentran en una misma escala, por lo que se plantea el expresar las puntuaciones en función de un grupo normativo, lo que no es del todo satisfactorio en la práctica.

“La longitud de los objetos depende del tipo de regla con la que se mide” (Muñiz, 1996; p.45)

- Invariabilidad en el instrumento respecto del objeto a medir, en el que los índices de dificultad y de discriminación de los reactivos, depende del tipo de personas a las que se aplica el instrumento y que se utilizaron para el cálculo de tales índices.



Tabla 13

Diferencias entre la Teoría Clásica y la Teoría de Respuesta a los Ítem, tomada y adaptada de Muñiz (1997)

Elementos	Teoría Clásica	Teoría de Respuesta al Ítem
Modelo	Lineal: la puntuación empírica es igual a la verdadera, afectada por un error aleatorio.	No lineal: la puntuación empírica con la verdadera se da en base a una función curvilínea, adoptada por el modelo.
Asunciones	Débil: son generalizables y son fáciles de cumplir por los datos empíricos.	Fuerte: sacrifican la generalidad para lograr precisión en la predicción, los datos deben cumplir con los supuestos.
Invariabilidad en las mediciones	No: los resultados dependen del tipo de instrumento con el que se miden los objetos.	Si: permite mediciones invariantes del instrumento empleado, que se deriva de sus modelos.
Invariabilidad de las propiedades de la prueba.	No: los índices que se obtienen dependen de la muestra de personas que se empleen para estimarlos.	Si: se derivan de los modelos, pero se comprueban empíricamente, a mayor variabilidad de las muestras, mayor precisión en los parámetros.
Escala de puntuaciones	Entre 0 y la puntuación máxima en la prueba, misma que se transforma en otras (normas)	Entre $-\infty$ y $+\infty$ para la estimación de las puntuaciones en las personas, que da una escala con todas las mediciones, independientes de la prueba.
Énfasis	Prueba: para comparar las mediciones de dos personas y se les aplica la misma prueba o dos formas paralelas.	Reactivo: dan una medición en la misma escala común, se pueden comparar los resultados a partir de los reactivos.
Relación reactivo-prueba	No se conoce la probabilidad de acertar un reactivo en la prueba, aunque se sepa se puntuación en la misma.	Al conocer la puntuación de una persona, el calcular la probabilidad de acertar el reactivo es inmediato a través de la CCI.
Descripción de los reactivos	Índice de dificultad y discriminación, dependiente de la muestra.	Parámetros a, b y c, independientes de la muestra utilizada. El parámetro "b" está expresado en la misma métrica que los parámetros θ de los sujetos.
Errores de medida	Error típico: común para toda la muestra, sin tomar su nivel en la variable medida.	Función de información: varía según el nivel de aptitud de las personas en la variable medida.
Tamaño muestral	Funciona con 200 sujetos	Se utilizan muestras de 500 sujetos, aunque depende del modelo.

Aunque es mejor la combinación de ambos enfoques, la TC y la TRI, en función de las circunstancias, que es de acuerdo a Muñiz (1997) lo más recomendable en la práctica.

4.1.5 Selección de los reactivos para la prueba final

Es la fase final en el proceso de análisis de reactivos, dentro de la TCP, de manera que se obtengan aquellos reactivos que conformarán la prueba final que será estandarizada; tal selección considera la finalidad y diseño original de la prueba, las especificaciones de la prueba y los datos del análisis de reactivos, Hogan (2004) recomienda ciertas características para tal proceso.

- a) En cuanto a la cantidad total de reactivos, una prueba corta reduciría su confiabilidad y como regla general se puede seguir el hecho de que a manera que se aumente el número de reactivos, también se hará en la confiabilidad.
- b) El grado de dificultad promedio de una prueba está en función de sus valores p , por lo tanto la puntuación promedio de la prueba es la suma de estos valores, de manera que una prueba fácil representará valores p bajos, sin embargo el que una prueba sea fácil o difícil dependerá de la finalidad para la que se construyó.
- c) En general se busca que los reactivos tengan índices de discriminación elevados (80 o superior) debido a que contribuyen a la medición del rasgo en cuestión. Sin embargo índices de discriminación entre .30 y .50 son aceptables para un conjunto de reactivos, ya que un solo reactivo tendría una confiabilidad muy limitada. Los índices de discriminación negativos deben evitarse, igualmente que los cercanos a cero, por no contribuir a la medición del rasgo.
- d) Ciertos reactivos podrán incluirse para satisfacer los fines de las especificaciones de contenido del instrumento y así garantizar la validez de contenido, lo que indica que los estadísticos de la prueba deben usarse moderadamente al momento de realizar la selección de reactivos, e incluso pueden incluirse reactivos por fines de motivación, lo que indicaría que en tal caso se comenzará con algunos que sean sencillos e ir aumentando la dificultad de éstos. Landsheere (1978) menciona que el ordenamiento puede ser por orden de dificultad creciente o por orden cíclico de dificultades, lo cuál constituye que los reactivos se clasifiquen de manera fácil – difícil – fácil – difícil, etc., lo cuál es un aspecto que ánima al sujeto a tratar de resolver todos los reactivos.

Dentro de la TRI la construcción de una prueba a partir del banco de reactivos implica que el banco de reactivos este bien diseñado, para que facilite la construcción de la prueba, lo que se facilita con un sistema computarizado lo que garantiza la estandarización de las instrucciones, la presentación de los reactivos, el registro de las respuestas, la corrección y la puntuación de la prueba y el análisis estadístico de los datos (Prieto, 1991; en Muñiz, 1996).

De manera que una vez que el banco de reactivos se encuentra en un sistema computarizado, la fase de construcción de la prueba a partir de un banco de reactivos se facilita, debido a que en el programa computarizado se especifica las características de los reactivos a utilizar y el computador busca en el banco los reactivos que cumplan con las especificaciones requeridas y los resultantes se pueden seleccionar para formar parte de la prueba. La utilización del banco de reactivos se ha puesto de manifiesto a la hora de construir pruebas referidas al criterio y a la medida (Muñiz, 1996).

● Pruebas con referencia al criterio

Anteriormente se ha mencionado que en la planeación de la prueba se debe elegir bajo que método se construirá, de manera que si es relacionada con los criterios o con el contenido, implicó la definición de los objetivos para un curso o una unidad y redactar los reactivos para lograr esos objetivos. Hogan (2004) menciona que tanto en la TCP como en la TRI lo que se busca es medir las diferencias entre las personas, pero no debe ser el único propósito, de manera que en estas pruebas se busca si en realidad se ha cumplido con cierto criterio, dentro de lo cuál el análisis de reactivos sirve, debido a que cuando se presentan diferencias en las puntuaciones, se genera información útil sobre el reactivo.

Muñiz (1996) menciona que el objetivo es verificar en qué medida los sujetos dominan ciertas áreas de conocimiento, conocido como dominio, donde para analizar el grado de conocimiento se utiliza el porcentaje de reactivos contestados correctamente, considerando que al no superar cierto porcentaje, es indicio de que no se ha alcanzado el nivel adecuado, que hace que el porcentaje marque una puntuación de corte a

superar por los sujetos para considerar que se han alcanzado los criterios de los contenidos que han sido objeto de estudio.

La ventaja del banco de reactivos en este tipo de pruebas es la unificación de la escala de habilidad (θ) para las pruebas destinadas a evaluar un mismo dominio, pero si la utilización de la prueba tiene como finalidad la clasificación de los sujetos, el banco facilita la selección de aquellos que muestren una discriminación máxima en el nivel de la escala de habilidad asociada con el porcentaje del punto de corte , donde una vez seleccionados, se elige los que den mayor información en ese nivel de la escala, aumentando así la precisión y medición de la medida (Hambleton, 1983, 1984; en Muñiz, 1996).

● Pruebas a la medida o con referencia a normas

También son llamadas pruebas adaptadas a los sujetos. En el caso de construir pruebas construidas a la medida diferentes, que midan la misma variable aplicadas a diferentes sujetos, se necesita decidir los reactivos a utilizar y colocar los resultados en una escala común para ser comparados, lo que se facilita si se cuenta con un banco de reactivos. Las pruebas a la medida permiten extraer más información de un sujeto en cada reactivo, lo que permite mayor precisión en la medida y ahorro de tiempo, debido a que la cantidad de reactivos que se emplean puede ser hasta de un 50% a diferencia de una prueba no adaptada, debido a que da la misma información sobre el nivel de habilidad del sujeto y con una menor cantidad de reactivos (Brown y Weis, 1977; Warm, 1978; Weiss, 1983; Kiely, Zara y Weiss, 1983; Kent y Albanese, 1987; Olsen, 1990; en Muñiz, 1996).

En resumen se puede marcar que en la planeación de la prueba se debe indicar el contenido y los objetivos que cubre, así como el hincapié que se hace en cada una de estas categorías, lo que da una base para la construcción de los reactivos y ayuda a la determinación de la validez del instrumento.

4.1.6 *Confiabilidad*

Se refiere a la exactitud o consistencia de las puntuaciones, lo que indica hasta que punto pueden atribuirse a errores de medida, las diferencias entre las puntuaciones del test (Alonso, 1998).

Para Arias (1990), la confiabilidad se refiere al grado en el cual se eliminan los errores de medición, para lo que se hace preciso determinar el grado de exactitud con que se mide la habilidad, interés o estilo que pretende medir el instrumento, lo que se traduciría en el grado en el que su aplicación repetida al mismo sujeto u objeto arroje resultados iguales (Sampieri, et al., 2003)

La confiabilidad tiene por objeto determinar si el instrumento permanece constante o si cambia, debido a factores ajenos a los que interesa evaluar, para lo que se hace necesario verificar la consistencia interna de los predictores. La confiabilidad parte del supuesto de homogeneidad de los ítems, los cuáles deben correlacionar sustancialmente entre sí (Alonso, 1998); siendo uno de los requisitos esenciales de cualquier instrumento de medición.

La confiabilidad en los puntajes de una prueba se refiere a la consistencia en que los resultados ponen a un estudiante en la misma posición en relación con otros, si es que la prueba se toma repetidamente (Carrera y Romo, 1983). Sin embargo la confiabilidad es una condición necesaria pero no suficiente, dado que una prueba podrá ser confiable pero no necesariamente válida, la confiabilidad se puede determinar a través de varias técnicas, que se enuncian a continuación.

4.1.6.1 Métodos para estimar la confiabilidad.

El tipo dependerá del tiempo en que se hacen las mediciones, los instrumentos empleados y su comportamiento interno, recordando dos características: la estabilidad temporal y consistencia interna. La primera se refiere a que tras el empleo en dos ocasiones diferentes, con los mismos sujetos, bajo mismas circunstancias y de forma estandarizada arroje resultados semejantes. Para tal fin se emplean los procedimientos descritos en seguida.

- Método de estabilidad o test – retest.

Se evalúa a un grupo de personas (mínimo 30) con el instrumento que se va a emplear, se deja que transcurra un tiempo y bajo las mismas condiciones que en la primera ocasión, se vuelve a aplicar el instrumento a los mismos sujetos. Con ambas aplicaciones, se procede a calcular el índice de correlación (Pearson, Spearman, Kendall) y si la correlación es alta y positiva, se considera confiable. Entre el resultado se aproxime más a 1, mayor será la confiabilidad; se aceptan índices superiores a 0.65; si la cifra es menor, se hará necesario la revisión de los reactivos del instrumento (Arias, 1990)

- Método de formas paralelas o equivalentes

Se administran dos versiones equivalentes del instrumento de medición, se aplica a un mismo grupo de personas en un período relativamente corto; las versiones deben ser similares en cuanto a instrucciones, contenido, duración y otras características, el instrumento resultará confiable si tras la correlación de los resultados de la aplicación de ambas versiones a los mismos individuos, resulta ser positiva de manera significativa (Sampieri et al., 2003).

Establecida la equivalencia, pueden administrarse una detrás de otra (inmediatamente después de la primera) o separadas por un intervalo de tiempo, se considera la más adecuada para las pruebas educativas, aunque resulta ser costosa y tomar mucho tiempo (Karmel, 1978).

En cuanto a la segunda característica, que es la consistencia interna, se refiere al comportamiento del instrumento en su estructura interna, a través de la consistencia interna de los reactivos que lo constituyen, y se subdivide en:

- Método de división a la mitad y pares o nones

Si se desea realizar un muestreo de la consistencia del contenido sin considerar las variaciones de las respuestas de forma individual, se puede usar esta técnica

(Karmel, 1978). Se requiere sólo un aplicación de la prueba, Sampieri et al., (2003) menciona que del conjunto de reactivos, se divide en dos mitades y se comparan las puntuaciones, de manera que el instrumento resultará confiable si los resultados de ambas mitades se correlacionan de manera significativa.

Otra manera consiste en dividir la prueba a la mitad, en donde una mitad tendrá los reactivos con números pares y la otra con los nones, se procede a recalificar la prueba, obteniéndose dos puntuaciones para el sujeto, y se correlacionan, obteniendo así el coeficiente de consistencia interna. En la prueba se colocan equilibrando los pesos, de acuerdo al grado de dificultad.

- Coeficiente alfa de Cronbach

Se emplea cuando los instrumentos tienen tres o más opciones de respuesta, se requiere de una sola administración de la prueba, produciendo valores que van de 0 a 1, lo único que se hace es aplicar la prueba y calcular el coeficiente. Arias (1990) indica que el instrumento en cuestión se aplica a un grupo no menor de 30 personas y se calcula el coeficiente alfa de Cronbach para determinar el nivel de significancia, el cuál puede ser por medio del programas de computadora (como el Statistic Package for the Social Sciences, SPSS).

Sin embargo si este coeficiente no adquiere un 0.85 o más, se debe calcular la correlación reactivo con la calificación total, a través del coeficiente de correlación Producto Momento de Pearson, para detectar los reactivos que causan problemas, eliminarlos y volver a calcular el coeficiente alfa.

- Método KR-20 de Kuder & Richardson (1937; en Sampieri et al., 2003)

Se utiliza cuando el instrumento cuenta con dos opciones de respuesta, presenta la suposición de que cada reactivo mide los mismos factores que los demás, y lleva a un coeficiente de confiabilidad, que se interpreta de la misma manera que el de pares y nones, aunque no es bueno para las pruebas de velocidad y no mide las variaciones individuales de un momento a otro (Karmel, 1978). La interpretación se lleva a cabo de la misma manera que el método del coeficiente alfa. Al igual que el coeficiente alfa se

emplean para verificar la consistencia interna de la prueba, aunque se usa más para los inventarios de personalidad. Si la magnitud no alcanza ha ser de 0.85, se determina la correlación entre el reactivo y la calificación total, a través de la correlación biserial – puntual para detectar los reactivos que causan problemas, eliminarlos y así volver a calcular el coeficiente Kuder.

No es indispensable calcular todos los tipos de confiabilidad, pero en la elaboración de un instrumento se debe señalar cuál se obtuvo se obtuvo, su magnitud y su significancia, así como aquellos reactivos que se eliminaron, y el porqué, además de las características de la muestra empleada para la determinación de la confiabilidad.

Sampieri et al. (2003), alude elementos que afectan la confiabilidad y la validez.

- ✱ Improvisación: se refiere a tomar a tomar a la ligera el desarrollar un instrumento de medición, que es un proceso complicado que no debe darse en la evaluación.
- ✱ Utilizar pruebas desarrolladas en el extranjero y no ser estandarizadas para la población meta a la que se destina la aplicación.
- ✱ El no tomar en cuenta las diferencias como edad, capacidad de respuesta, el nivel de conocimiento, etc.
- ✱ Las condiciones en las que se aplica el instrumento, tanto físicas, psicológicas e incluso mecánicas, en cuanto a errores en las instrucciones, que falten páginas, etc.

A continuación se hace un resumen ilustrativo de los métodos para estimar la confiabilidad, para que de esta manera se esclarezca un poco más estos conceptos.

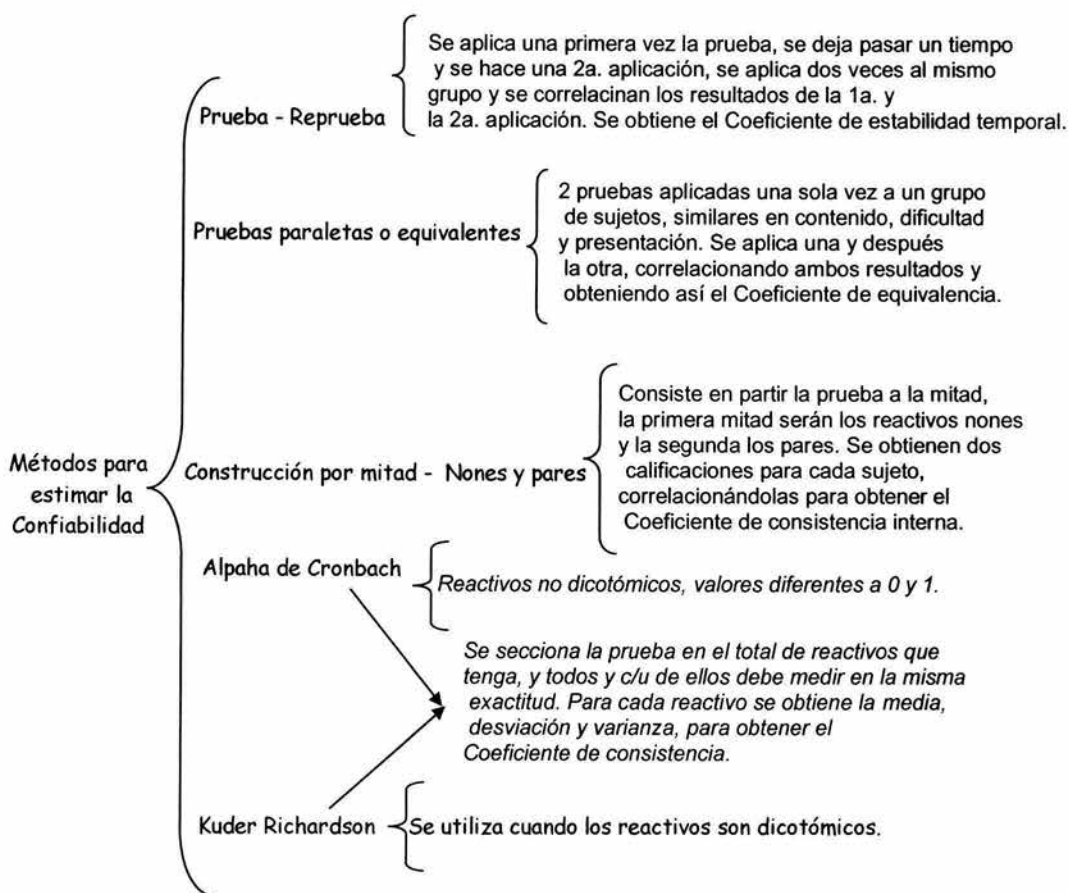


Figura 4. Resumen de los métodos que existen para obtener la confiabilidad de una prueba, así como en qué consiste cada uno de ellos.

4.1.7 Validación

Cualquier instrumento de medición utilizado como registro para obtener información sobre ciertas variables, debe cubrir dos requisitos esenciales: la confiabilidad y la validez; para González y de León (1987 a; en Alonso, 1998) un instrumento de evaluación será válido cuando “mide lo que pretende medir”.

Para evaluar la exactitud del instrumento se demanda el conocer el rasgo que se desea pronosticar después de cierto período (Alonso, 1998).

Sampieri, Collado & Lucio (2003) definen la validez como “el grado en que un instrumento realmente mide la variable que pretende medir” (p. 346). En los procesos de evaluación, se habla de predictores que deben correlacionar de manera significativa con los criterios relativos al desempeño, donde dicho predictor como los exámenes de conocimientos empleados o las pruebas psicológicas, deben ser a la vez también válidos y confiables. .

Validar la prueba es verificar que mide lo que debe medir, siendo de conocimiento la importancia que este apartado tiene para cualquier instrumento psicométrico. En la prueba de conocimientos lo que se pretende medir se relaciona con los objetivos a cubrir. Landsheere (1978) menciona que el destino de una prueba es cumplir con tres funciones: pronosticar, diagnosticar y describir .

b) Función pronosticadora

La forma de hacerlo es observar si el pronostico se está realizando, controlando los comportamientos o los logros de los examinados en un tiempo considerable para confirmar o invalidar el pronostico y de esta manera calcular eventualmente el coeficiente de validez (correlación).

c) Función diagnosticadora

Si el destino de la prueba es para descubrir los puntos del programa que están impidiendo el progreso del alumno, el hecho de implementar programas para su mejoramiento académico se está basando en el diagnóstico, lo que permitirá hacer conclusiones respecto a éste último.

d) Función descriptiva

Se refiere a las pruebas utilizadas para describir un hecho o situación, como el nivel de conocimientos, aunque es didáctico e incluye fines pronóstico y diagnóstico, y a continuación se mencionan los tipos de validez más empleados.

4.1.7.1 Validación del contenido

Sampieri et al. (2003) indica que se refiere al grado en el que la prueba muestra un dominio en específico del contenido que se mide, tal como el caso de ser una prueba de conocimiento de los presidentes de México que omita a una de ellos, no tendrá este tipo de validez, pues debe incluir a todos, lo que indicaría que requiere que se representen prácticamente a todos los reactivos del dominio de contenido de las variables a medir. Dado lo cuál decimos que si las metas del curso se encuentran representadas en la prueba, se dice que es válida.

De lo cual Carrera y Romo (1983) indican que una manera de operacionalizar la definición es a través de varias etapas:

- a) Describir el programa de la materia de interés.
- b) Decidir el modelo taxonómico a usar.
- c) Elaborar la tabla de especificaciones.
- d) Listar los objetivos específicos de la materia en consideración de la tabla de especificaciones.
- e) Elaborar los reactivos en función de los objetivos de la etapa anterior.

Karmel (1978) menciona que este tipo de validez es vital para las medidas de aprovechamiento, considerando a una prueba con validez de contenido cuando ésta representa las metas curriculares de los usuarios de la prueba. Además Karmel indica que en el manual de la prueba deben buscarse datos, para evaluar la validez de contenido, tales como la descripción de la asignatura considerada y como se realizó el muestreo, así como un resumen del currículo de especialistas que evaluaron las preguntas y las formas de calificación.

Después de lo cuál se puede decir que la prueba debe tener una gran cantidad de reactivos, que representen una muestra representativa de los conocimientos que se desean medir y de las capacidades que deben reflejar tales conocimientos.

4.1.7.2 Validación relacionada con el criterio

Establece la validez de la prueba comparando ésta con algún criterio externo, el cuál es un estándar con el que se juzga la validez del instrumento (Wiersma, 1999; en Sampieri et al. (2003), éste criterio puede ser las calificaciones escolares y su propósito radica en la predicción, siendo útil para la selección de aspirantes o para empleados, Karmel (1978) indica que éste tipo de validez se establece por correlación entre la puntuación de la prueba y una medida de criterio, en lo cuál una alta correlación positiva será indicio de validez relacionada con el criterio. Sampieri et al. (2003) menciona una división de esta validez en concurrente y predictiva, en el primer caso, los resultados de un instrumento se correlacionan con el criterio en el mismo momento, por ejemplo si se desea conocer las preferencias por cierto partido, se puede aplicar un cuestionario días antes de la elección y compara estos resultados con los arrojados en la elección. Implica el usar métodos estadísticos, llamado coeficiente de correlación, el cuál expresa el grado de relación de la puntuación de la prueba y de un criterio externo, éste coeficiente puede tener valores de +1.0 pasando por cero hasta -1.0. El +1.0 indica que la correlación es positiva, -1.0 indica una relación perfecta pero negativa y de cero indica una relación lineal entre el test y las calificaciones (Karmel, 1978)

Sin embargo un problema puede ser el criterio a utilizar, en general un criterio común es una medida de éxito como las calificaciones en programas educativos, pero aún así, ninguna medida de criterio es perfecta; Magnusson (1978) afirma que “la única forma de hacer más validos los datos del criterio está en refinar el análisis de las variable a medir así como, relacionar la medida de criterio con lo que se considera el mejor criterio” (p. 127).

En el caso de la validez predictiva, el criterio se fija en un futuro, de manera que al tener los resultados del criterio a medir, se compararía por ejemplo con el futuro desempeño de las personas (Sampieri et al., 2003). En este caso se habla de un coeficiente de validez, que es la correlación entre el test y el criterio, donde dicha estimación de la validez da un índice total de la validez de la prueba. De manera que entre más grande sea el coeficiente de validez menor será la probabilidad de error (Karmel, 1978).

Sin embargo existe otro tipo de validez, en la que se investiga los rasgos que mide la prueba, es decir lo que sus puntuaciones nos dicen acerca de un examinado, la cuál es la validez de constructo, que se revisa a continuación.

4.1.7.3 Validación de constructo

Se refiere al grado en el cuál una medición se relaciona consistentemente con otras mediciones, se considera de las más importantes; dado que el constructo es la variable medida. Sampieri et al. (2003) mencionan que este tipo incluye tres etapas: la primera se refiere a especificar la relación teórica entre los conceptos, en la base del marco teórico; la segunda se refiere a correlacionar ambos conceptos y analizarla y por último interpretar la evidencia empírica con lo observado en la validez de constructo de la medición en particular, de tal manera que para llevar a cabo este tipo de validación debe existir un marco teórico que soporte la variable en relación con otras variables.

Karmel (1978) indica que este tipo de validez se lleva a cabo para verificar la teoría en que se funda la prueba. De manera que un instrumento de medición representará las variables que pretende medir, cuanto más evidencia tenga de validez de contenido, criterio y constructo, lo que se expresa en la figura siguiente.

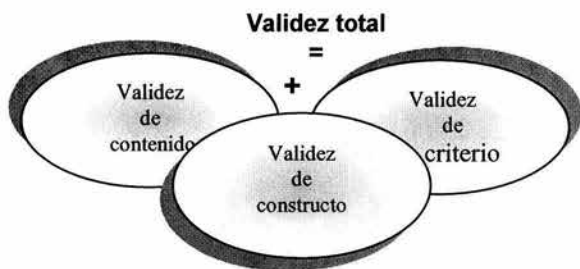


Figura 3. Representación de los tipos de validez y el cómo la obtención de los mismos, ayuda a obtener la validez total de una prueba.

Dentro de esto concluiría que la validez de contenido es importante en pruebas de aprovechamiento; la de criterio para las de inteligencia o aptitud escolar, para la predicción del éxito; y la de constructo en pruebas de personalidad, para corroborar la base teórica que las sustenta, más en técnicas proyectivas.

4.1.8 Estandarización de la prueba

Este proceso dependerá a la vez de la dirección que se le diera a la prueba, esto es si se construyo con referencia a normas o al criterio.

La estandarización de las pruebas de ejecución ha sido usada en las escuelas como medida de lo que están haciendo, las cuales son pruebas con referencia a normas, en las que se compara la ejecución de estudiantes locales con la de una muestra representativa de estudiantes pertenecientes a un grupo normativo de alguna región, nación, o estado (Gronlund, 1998)

Se suele conocer como programa de estandarización o normalización, el cuál es la continuación a la fase de selección de reactivos, la prueba final estandarizada será la misma que habrá de publicarse al final con aspectos establecidos como son las instrucciones, reactivos, mínimo de tiempo, etc., donde Landsheere (1978) refiere a que debe ser el mismo día, hora, temperatura, disposición de materiales, silencio, mismos grado de familiaridad de los sujetos con la prueba (incluyendo las hojas de respuesta), mismas instrucciones y duración si la prueba es limitada en tiempo.

En la estandarización, un paso importante es el establecimiento de las normas, mismas que son calificaciones o ciertos puntajes, que sirven como punto de referencia para poder interpretar el puntaje de un sujeto en particular, pero este proceso de normalización implica trabajo, como Nunnally & Berstein (1997) mencionan; esto es debido a que debe haber un aseguramiento de que la muestra de estudiantes sea representativa y el realizar análisis estadísticos.

La estandarización a través de las normas, puede ser por medio del análisis de las puntuaciones de la prueba ya sea por género, raza, edad, región geográfica o algún otra clasificación demográfica.

En esta fase pueden llevarse a cabo varios tipos de estudio sobre confiabilidad. Las normas son válidas para grupos más o menos extensos y definidos; de manera que el grupo normativo puede ser una muestra representativa a nivel nacional, que reproduzca de manera exacta las características del país entero, a lo que Thorndike (1996) sugiere que al decidir la población con la que se harán las comparaciones se debe primero diseñar un procedimiento de muestreo para reproducir las características

de la población en forma fiel y después llevar a cabo el diseño con la menor “dispersión” posible, para que de esta manera sea un proceso más practicable.

Gronlund (1998) menciona que en las pruebas de ejecución con referencia a normas, la construcción de la estandarización, debe tener los siguientes rasgos:

- El contenido de la prueba es usado en base a los libros y guías del programa.
- Los reactivos de la prueba se escribirán por expertos en la materia y en base a la tabla de especificaciones.
- Los reactivos de la prueba son analizados y revisados, obteniendo su poder de dificultad y discriminación, y de están revisión algunos serán eliminados.
- El conjunto final de reactivos seleccionados, será la base de las especificaciones hechas en la prueba.
- La dirección para la administración y calificación de la prueba se prescribirán rígidamente.
- La administración de la prueba para la selección de un grupo de estudiantes para un proceso nacional, regional o de un estado, se realiza para obtener las normas, que servirán para la interpretación de las puntuaciones en la prueba.
- La versión final de la prueba es publicada junto con el manual de la misma, el cuál describe las técnicas de calidad de la prueba, el procedimiento de administración, calificación, interpretación y el uso de los resultados.

De manera que la interpretación de normas en pruebas con referencia a éstas y la comparación de las puntuaciones en las pruebas estandarizadas, requiere comprensión de varios tipos de puntuaciones que serán usadas y que describen la ejecución de los estudiantes en las pruebas, llevado por procesos tales como rangos percentilares, puntuaciones de grado o estándar, para lo que se toma frecuencias en la curva normal, en notas “Z” (Landsheere 1978), sin embargo en la planeación se debe incluir decisiones sobre los tipos de calificaciones convertidas que se requerirán, tales como lo son los percentiles, las puntuaciones estándar o puntuaciones básicas en una escala básica del atributo latente (Thorndike, 1996).

Nunnally & Berstein (1997) indican que las normas se expresan en puntajes estándar transformados y en percentiles. En el primer caso se convierten los puntajes brutos a una distribución con una media de 500 y una desviación estándar de 100. En el

caso de los percentiles se refiere al porcentaje de sujetos de la muestra normal que se encuentra debajo de un puntaje en particular, su cálculo es por medio de la división del número total de estudiantes que se examinaron entre el número parcial de los que obtuvieron un puntaje inferior a cierto nivel, más la mitad de los que obtuvieron esa calificación, de manera tal que si el 60% de estudiantes logran una calificación inferior a 100, alguien que obtenga una calificación de 100, se encontrará en el percentil 60.

En el caso de que se haya diseñado una prueba con referencia al criterio, la interpretación en el proceso de estandarización es diferente, recordando el hecho de que sólo se obtienen normas cuando se trata de pruebas con referencia a las normas.

Gronlund (1998) indica que en la interpretación con referencia al criterio en las pruebas de ejecución ha sido adicional para algunos test estandarizados; lo que incluye el puntaje de porcentaje correcto, el uso de estándares de ejecución, y la interpretación por un grupo de reactivos que representan un área de contenido, una habilidad o en su caso un objetivo.

De manera tal que en el caso de las pruebas con referencia a un criterio, se requiere comprobar los objetivos, contenidos y habilidades de la prueba para igualar a los del programa instruccional, de manera tal que si después de la interpretación de la prueba con referencia al criterio, se obtiene un número suficiente de reactivos para cada tipo de interpretación, se procede a determinar los estándares de ejecución.

Uno de los métodos para este tipo de pruebas, resulta del puntaje de porcentaje correcto, con el reporte del porcentaje de reactivos en la prueba, o subtest de respuestas correctas. El reporte puede ser por estudiantes individual, por la clase, escuela o del distrito de escuelas enteras, de manera que el puntaje que se tome de tal porcentaje de respuestas correctas para varios grupos de escuelas son comparados con el puntaje de porcentaje de respuestas correctas en simples normas nacionales, que representa una base para evaluaciones a las escuelas (Gronlund, 1998).

4.1.9 Preparación final de la prueba

La etapa final es la publicación de la prueba, el cuál tiene que ver con las instrucciones de aplicación, interpretación, los manuales técnicos, los reportes de puntuaciones, y materiales suplementarios. El manual técnico es la fuente de información sobre la finalidad, fundamentos y estructura de la prueba y debe contener información de la confiabilidad, validez y procedimientos de normalización del instrumento, en caso de ser una prueba con referencia a las normas. Los materiales suplementarios pueden ser instrumentos de localización que ayudan a los aplicadores a determinar el nivel de una y determinar la adecuada para un examinado específico (Hogan, 2004).

Thorndike (1996) indica que si la prueba se prepara para que la usen otras personas, se deben preparar manuales para ellas con los siguientes datos.

1. Indicar los usos a los que se destina.
2. Incluir las instrucciones completas para la aplicación.
3. Las tablas de normas para realizar las conversiones de las puntuaciones y así poder hacer las interpretaciones.
4. Los valores de la confiabilidad de la prueba.
5. Evidencias de la validez.
6. Guías y sugerencias para el uso de los resultados.

Otro punto es el referente a las condiciones de aplicación, en la que se encuentran los factores físicos, psicológicos, la motivación e incluso la ansiedad.

Dentro de las condiciones físicas Thorndike & Hagen (1973) enumeran cuatro condiciones óptimas para un examen:

- a) Los sujetos deben estar cómodos físicamente y emocionalmente tranquilos
- b) No tener que sufrir interrupciones o algún tipo de distracción
- c) Contar con un espacio para poder manipular sus materiales
- d) Estar a cierta distancia uno de otro para evitar que copien

Anastasi (1998) sugiere un salón en el que no se produzca distracción, con ventilación y luz adecuada. Karmel (1978) menciona que las condiciones psicológicas son más importantes que las físicas y que dependen de la capacidad del aplicador para

establecer "rapport" (crear un clima de confianza con el examinado). El examinador debe mostrarse tranquilo y animado para evitar que los sujetos se sientan amenazados por la prueba y procurar decirles que deben dar su mejor esfuerzo en la prueba. Por lo que Karmel (1978) recomienda lo siguiente:

- Programar una sesión de orientación dedicada a los examinados, señalando los propósitos del examen.
- El aplicador debe mostrar una actitud amistosa y tranquila en el examen y leer las instrucciones con voz clara, pero a la vez ser objetivo, sin mostrarse indiferente.
- Intentar que haya ventilación e iluminación adecuadas.
- En el caso de ser niños pequeños, tratar de usar un salón de clases.
- Colocar anuncios en las puertas indicando que se aplica examen.
- Arreglar con la dirección para evitar interrupciones.
- Preguntar si han tenido oportunidad de ir al baño, para que se evite salir durante la aplicación del examen.
- Verificar que cada examinado tenga dos lápices y tener de sobra otros.
- Disponer de materiales, escritorios o mesas que faciliten la manipulación del material.

Además de los puntos anteriores, se debe tomar en cuenta el factor de la motivación, debido a que si se está tratando de mantener las condiciones de uniformidad en el examen, se debe motivar al examinado para que realice su esfuerzo máximo, aunque un estudiante altamente motivado, puede estar muy ansioso por hacerlo bien que puede llegar al grado de obstaculizar su desempeño.

V) EXPERIENCIAS SIMILARES

5.1 Políticas, criterios de admisión y requisitos.

El objetivo principal en la admisión de alumnos, es el poder escoger a los alumnos más capacitados, por medio de un proceso objetivo de selección, lo cuál supone que permitirá el predecir un buen rendimiento académico; para lo cuál las diferentes instituciones educativas piden criterios, políticas o requisitos, los cuáles deben distinguirse, tal como se menciona en Alonso, (1998), quien refiere a que las políticas de admisión serán los propósitos y objetivos de la institución, que se traducen en criterios de selección, como el examen de selección; y los requisitos serán los que se necesitan presentar para el proceso, como presentar el certificado, etc.

Sánchez (1985, en Alonso, 1998) expone que las condiciones sociales y las posibilidades educativas, es lo que determina el proceso de admisión y selección de estudiantes colombianos. Este autor concluye, que este proceso es el resultado de una decisión tomada en el mejor caso, en relación con habilidades y características personales de los aspirantes, mismos que se asumen como válidos para su condición, así como referentes al contexto en el que se ubican.

Otro autor que distingue entre políticas, requisitos y proceso de admisión es Lobo (1985, en Alonso, 1998), para quien políticas serán las exigencias y experiencias contribuidas por parte del estudiante, la enseñanza que a recibido y atributos del lugar donde la ha recibido. Los requisitos, dependerá de la exigencia de estudios anteriores como certificados o el propio promedio de estudios anteriores. Finalmente menciona que el proceso de admisión dependerá de la institución, quien debe presentar información clara, su función, el tiempo que se debe emplear en el proceso, la metodología que sigue, etc., de tal manera que la persona que desee ingresar este informado de tales aspectos.

Por otra parte Dobles & Fallas (1983, en Alonso, 1998) hace constatar que los criterios de admisión que se utilizan hoy en día son muy variados, pero todos deben responder a sus objetivos, deberán ser evaluables y tener implícito algún tipo de selección, los cuales son aquellos que “dejan una libre entrada”, en donde se hace

imposible que no se cuente con criterios de selección explícita, esto debido a los cambios económicos y sociales y por la necesidad de profesionales capaces.

Los explícitos son más utilizados, y contiene al “liceo” que es el terminar el grado anterior al que se pretende ingresar como condición obligatoria para su ingreso en el nivel superior, el promedio del último año del ‘liceo’, el que se encuentren trabajando, el social – individual, el equilibrio emocional, y el más utilizado: concurso por medio de pruebas, ya sea de habilidades, aptitudes, destrezas y conocimientos. Todos estos criterios deben poseer confiabilidad, validez, objetividad, aceptabilidad y practicabilidad.

Como se puede observar, existen diferencias entre los aspectos a cubrir o tomar en cuenta para el proceso de selección, lo cuál dependerá de la escuela donde se lleve a cabo el proceso de admisión de alumnos, estos requisitos, criterios o políticas deben de especificar el peso que tendrán dentro del proceso y el propósito que tendrán; por otra parte la mayoría de las instituciones utilizan criterios de selección explícita, en el que tienen reglamentos que deben seguir y un cierto número de vacantes a cubrir, lo que hace que no se excluya por que sí, sino por una serie de realidades a las que Alonso (1998), atribuye que pueden ser “personales, políticas, sociales, educativas y económicas imposibles de ignorar”.

En el caso de la Maestría en ciencias con especialidad en Administración de Negocios (MAN), los criterios de admisión son el promedio académico de licenciatura, el cuál debe ser mínimo de ocho, así como el resultado de los exámenes de admisión, del examen de inglés y de las entrevistas realizadas individualmente.

5.2 Indicadores de éxito académico.

En la actualidad y debido a los antecedentes de procesos de selección, se han citado ciertos indicadores concebidos como elementos que predicen el éxito académico en los alumnos, los cuales son considerados como antecedentes para el logro de objetivos en el alumno y a nivel institucional; lo cuál remite a comenzar hablando del rendimiento académico como uno de éstos indicadores. El rendimiento académico se refiere a los logros de eficiencia terminal de la universidad y las variables que afectan al alumno universitario en su rendimiento (Alonso, 1998), lo que hace implícito el predecir

este indicador, mediante posibilidades de éxito de los aspirantes, dentro de lo que sobresalen diversas variables que a continuación se presentan y que a través de diversos estudios, han demostrado ser predictores de éxito en los aspirantes, con lo cual también se buscaría eliminar la deserción y el rezago escolar.

Alonso (1998) hace referencia a la Prueba de Aptitud Académica (PAA) que resultó ser el mejor predictor de rendimiento en humanidades en el primer año de la carrera de Medicina. El mismo autor concluye que las variables de rendimiento académico previo, es considerado un buen predictor; los test o pruebas de admisión son consideradas como un segundo grupo de variables que predicen éxito académico, pero al mismo tiempo que se ha encontrado en ellas fallas en cuanto a confiabilidad y validez; por otra parte la inteligencia es considerada también, pero no muy significativa y finalmente concluye que la variable de personalidad, tiene una validez predictiva menor que la inteligencia y propone que esta variable no se considere dentro del proceso de selección.

Para Touron (1985, en Alonso, 1998), las aptitudes diferenciales, junto con la inteligencia, encontró que ocuparon un tercer lugar en cuanto a su validez predictiva, en lo que concierne al rendimiento académico, incluyendo además entre las variables que funcionan como predictoras del éxito escolar, las relacionadas al rendimiento académico previo al ingreso a la universidad; y las pruebas de admisión correlacionando con el rendimiento al término del primer año, que coincide con los estudios de Escudero (1984; en Venegas 1993) confirmando que el rendimiento previo es el mejor predictor del rendimiento posterior.

Venegas (1993) encontró que aparentemente resultaron buenos predictores el examen de conocimientos, promedio de preparatoria y el examen de funciones lógicas.

En cuanto a la variable edad y género, autores como Malstron & May (1984, 1981, en Venegas, 1993) no encontraron un valor predictivo.

Venegas concluye que las variables que poseen poder predictivo son: rendimiento académico, considerado por los autores como el mejor predictor; los test o pruebas de admisión aunque con limitaciones; en tercer lugar la inteligencia, en la que refiere que los autores revisados afirman que se debe desmitificar la eficacia de esta variable dentro de las pruebas de acceso a las instituciones y en cuanto a personalidad

se observa con una validez predictiva menor, sin olvidar la importancia que podría tener para determinadas carreras.

Como se puede observar, en este apartado se incluyeron aquellas variables que según investigaciones de autores anteriores son consideradas como predictores de éxito académico, concluyendo que con los resultados de dichos estudios sobre el poder predictivo de algunas variables, las más sobresalientes son el rendimiento académico previo, como segundo grupo las pruebas de admisión, aunque se ha encontrado que en la mayoría la validez tiende a presentar limitaciones y no es significativa, lo que se convierte en un problema que reclama ser atendido. En cuanto a la variable inteligencia se considera significativa pero no alta, siendo de menor significancia la de personalidad, aunque en ciertas áreas, ciertos rasgos de personalidad pueden servir como predictoras de rendimiento previo, de manera que al tener presente esta información puede ser que se puedan detectar situaciones específicas que en un futuro representen problemas para los alumnos, tal como la deserción, reprobación o incluso en un proceso de toma de decisiones para los alumnos.

5.3 Nivel predictivo del examen de admisión.

Autores como Salazar & Flores (1995; en Alonso, 1998) realizaron un estudio al examen de selección a las escuelas incorporadas a la Universidad Autónoma de Nuevo León, y observaron beneficios al aplicar este examen para ingresar al nivel medio superior, entre los que se encuentra una menor deserción en los estudios, una mejor atención de maestros si se emplean grupos pequeños y homogéneos y un mayor rendimiento académico en cuanto a evaluaciones indicativas, con una eficiencia terminal mayor, que sin el examen.

Por otra parte Aguilar, Lizarraga & Pérez (1981) afirman que las pruebas de conocimientos además de su valor predictivo, dan información valiosa acerca de los conocimientos básicos del estudiante, para lo cual Venegas (1993), señala que en estos estudios si los exámenes de conocimientos son buenas variables predictoras, se debe establecer un mínimo de calificación como requisito de admisión, para que de esta manera, contribuya al éxito académico del alumno.

Venegas (1993), menciona que en cuanto a los resultados de su estudio, se observó una relación en cuanto al promedio del examen de conocimientos y el promedio obtenido de primero a octavo semestre, y que en función de esto es conveniente hacer una revisión de dicha prueba para validarla, debido a que cuenta con limitaciones, esperando que se eleve su capacidad predictiva, si se perfecciona tal examen.

De tal manera que en base a la información revisada, se pretende realizar un análisis descriptivo de la prueba utilizada en el proceso de selección, para conocer el poder de discriminación de la misma y determinar la objetividad del proceso de admisión, debido a que la prueba no ha sido sometida a análisis para conocer su confiabilidad y validez, e incluso no se reportan datos de ello, lo cuál es un elemento importante para proseguir con la elaboración de la propuesta en función de los resultados obtenidos, debido a que tras los resultados sobre las pruebas de aprovechamiento, éstas han resultado un buen predictor de éxito académico.

En el caso de la MAN. se tiene la limitante de que se desconoce la confiabilidad de la prueba de conocimientos empleada, e incluso nunca se han llevado a cabo estudios sobre la predicción de variables sobre el rendimiento académico, mucho menos sobre la validez de la prueba empleada, sin embargo, para elaborar la propuesta se hace necesario evaluarla para determinar su utilidad en el proceso de admisión de alumnos a la maestría y así poder establecer las pautas para obtener su validación.

5.4 Alternativas respecto a la validación de los exámenes

Desde el 2003, el Consejo Nacional de Rectores (en Otalvaro & Jacome, 2003) sostuvieron que los exámenes servían de instrumento, sin ser el único, para obtener una aproximación de los logros alcanzados en el proceso formativo.

Venegas (1993) cita a Fortier, quien propone desarrollar y utilizar métodos confiables de evaluación académica, ofrecer servicios eficientes de orientación educativa, sistemas competentes para el análisis y atención de necesidades verdaderas para estudiantes y un acceso a oportunidades educativas, los cuales menciona que deben ser los cuatro grandes puntales de un sistema de admisión.

Por otra parte Venegas (1993), en su investigación realizada en la Universidad Autónoma de Querétaro encontró como limitante, que las pruebas aplicadas carecen de confiabilidad y validez, y que para formular propuestas, se hace necesario el evaluar el procedimiento que se utilizó, para poder discriminar su utilidad en el sentido de los objetivos trazados por el proceso de admisión, puesto que su organización y desarrollo implica trabajo y tiempo.

González & León (1986; en Venegas, 1993), en un estudio realizado en la Universidad de Costa Rica respecto a la validez predictiva de los puntajes de admisión contra el de éxito y fracaso, señalan la importancia que tiene el proceso de admisión, para escoger a aspirantes que presenten una mayor probabilidad de éxito académico, para lo que proponen, la selección, diseño, construcción y validación de instrumentos psicoeducativos pertinentes.

Otra investigación acerca del análisis del examen de conocimientos, es la investigación realizada por Chávez (2000), quien a través de los resultados obtenidos en su estudio, propone que en la elaboración de un buen instrumento para evaluación de algún candidato, se debe basar en "la necesidad de identificar al mejor individuo para incorporarlo a la organización" (p. 103).

En el programa de evaluación realizada por Carrera & Romo (1983) llevado a cabo en la Universidad Iberoamericana, respecto al examen de admisión, se centro en la preocupación de tener instrumentos de selección de los aspirantes a primer ingreso, que fueran rigurosos y confiables, lo que llevó a una revisión de validación de los instrumentos, con finalidades de ayuda a los estudiantes y a la propia institución. Para lo cuál se eligieron pruebas psicopedagógicas y psicométricas; los resultaron llevaron a modificar el examen de admisión, sustentando que necesitaba un estudio de la validez predictiva que corroborara los datos encontrados. En cuanto a la calidad de una prueba de conocimientos recomiendan el utilizar una prueba de tipo objetiva y de selección múltiple, justificando el porque de su utilización.

Backhoff & Tirado (1992 b, 1993, 1995; en Venegas, 1993) plantean como necesidad de instituciones de educación superior, el conocer por un lado el nivel académico con el que llegan los estudiantes para su selección y por el otro diagnosticar

también las habilidades y conocimientos que poseen para detectar sus deficiencias y corregir ese aspecto.

Aguilar, et al. (1981) en una investigación diagnóstica realizada a los estudiantes que ingresan a Psicología, en la Facultad de Psicología de la UNAM, proponen que las pruebas diagnósticas a escuelas profesionales evalúen habilidades del lenguaje y dominio de conceptos de asignaturas del ciclo anterior relacionado con el área profesional correspondiente, puesto que es evidente que las investigaciones que se han hecho sobre la validez de las pruebas de admisión tienen serias limitaciones, y como es de esperarse, fallas en cuanto a la confiabilidad y la validez.

Por lo que se puede concluir el hecho de emplear medidas confiables y factibles de medición para responder a necesidades del sistema educativo y lograr que el proceso de selección que se realice en cada una de las instituciones, sea objetivo y acorde al cumplimiento de objetivos de ambas partes, para lo que propondría que aquellos que fueran excluidos del proceso, se les proporcionará una orientación hacia los lugares en donde si podrían fomentar su desarrollo personal y profesional a través de un diagnósticos preciso.

Para finalizar con éste apartado, a continuación se presenta un mapa conceptual, en el que se resume los puntos que se han abordado hasta el momento, mismos que se emplean para la elaboración de la propuesta con la cuál se comienza el apartado siguiente.

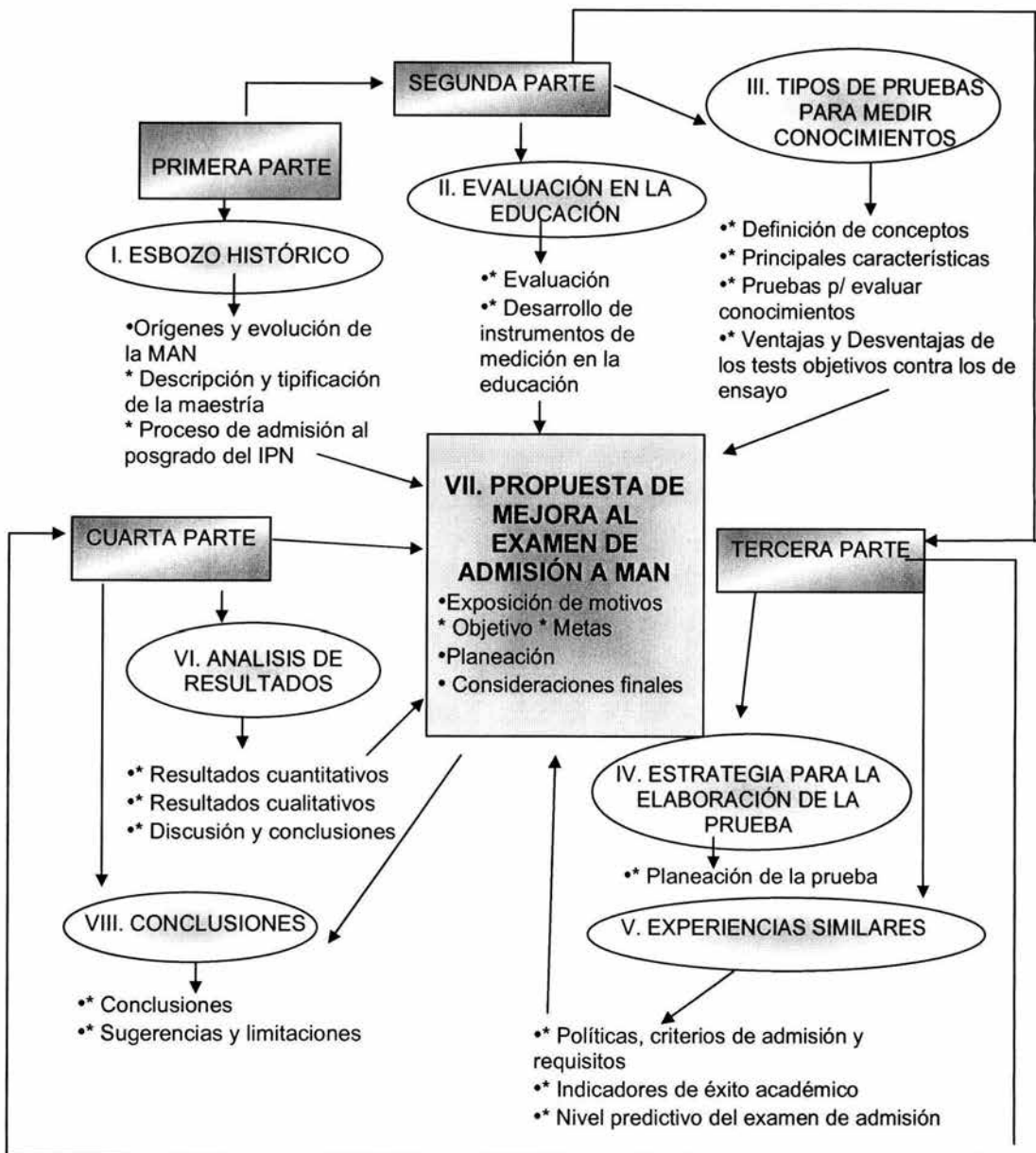


Figura 5. Mapa conceptual.

CUARTA PARTE.- ANALISIS DE RESULTADOS Y PROPUESTA DE MEJORA

VI) ANALISIS DE RESULTADOS

6.1 Resultados cuantitativos

Para fines de este apartado, en el análisis de resultados hecho al examen de admisión correspondiente al periodo 2006-B, se tiene lo siguiente.

Los resultados que se analizaron corresponden a una población de 123 alumnos, que deseaban ingresar a la Maestría en Ciencias con especialidad en Administración de Negocios. La población original era de 155, de los cuales 33 no se presentaron el día del examen, quedando los 122 que se analizaron.

De manera que en la presentación de los resultados se muestran:

- Los resultados que los aspirantes obtuvieron en el examen.
 - a) Los resultados globales, que corresponden a todos los aspirantes.
 - b) Los resultados por materia.

- Se presentan a la vez datos estadísticos como la media en cada una de las materias.

- Finalmente se presenta una descripción en lo referente a datos demográficos, tales como edad, instituciones de procedencia, etc.

El examen constó de 5 materias, cada sección con un número diferente de reactivos que se contestaron en la hoja de respuesta, misma que presentó una numeración consecutiva. Se debe mencionar que no se realizó análisis estadístico de los datos, puesto que sus características de elaboración no permitieron hacerlo.

Simplemente para fines del estudio, se resalta que la prueba empleada para el proceso de selección de la MAN, fue aplicada sin reportar índices de confiabilidad y validez, y no se recurrió a obtenerlos, debido a que se detectaron fallas en la elaboración, siendo nulo el conseguir tales índices puesto que la prueba carece de

homogeneidad en los reactivos, dificultando el establecer su consistencia interna y por lo tanto llevaría a que la prueba carezca del índice de confiabilidad y por lo tanto de validez, tanto de contenido, al no especificar el constructo que se mide; de criterio, al no dar una definición operacional del constructo a medir y por lo tanto no existiría de constructo porque no se cuenta con un marco teórico que sustente a la variable y por lo tanto se remontaría al mismo hecho.

De manera que el trabajo se centró en la metodología de elaboración para poder en primera instancia diseñar y después elaborar la prueba de aprovechamiento, debido a que si desde el principio no se establece el constructo que se desea medir, no se sabrá siquiera cuál será su finalidad.

a) Resultados globales

En la figura 6 se muestran las calificaciones obtenidas por la población total en el examen de admisión, en donde se obtuvo una media de 56.83, lo cuál representaría un alto índice de reprobación, que pondría en duda la confiabilidad de lo que se mide en el examen. Las figuras 6 a 10 muestran las calificaciones que cada alumno obtuvo por separado en cada una de las materias, en las que se presenta la media en cada una, a través de una línea horizontal marcada más fuerte.

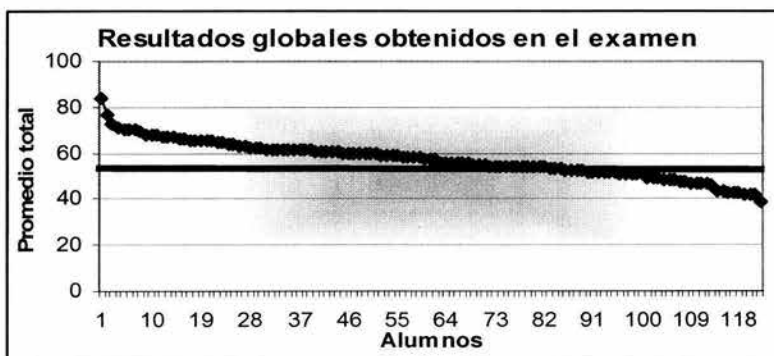


Figura 6. Resultados globales obtenidos contabilizados en las cinco materias, por los 123 alumnos que presentaron el examen correspondiente al período 2006-B de la MAN.

Se observa que la calificación máxima se situó en 83.80 y la mínima en 36.40, mostrando que en la lista de pre-aceptados se tomó como índice de aceptación a partir de 52.00, lo cuál se presenta por debajo de la media obtenida (ver anexo 4)

b) Resultados por materia

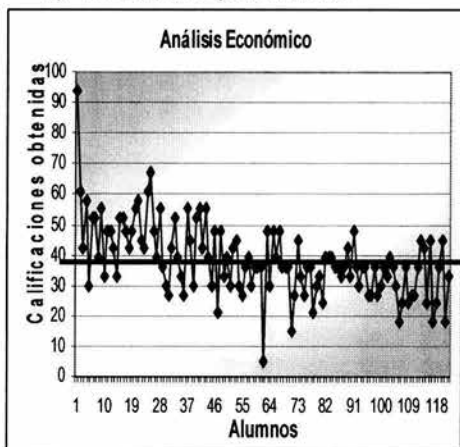


Figura 7. Calificaciones obtenidas por cada uno de los alumnos en la materia de Análisis Económico Administrativo.

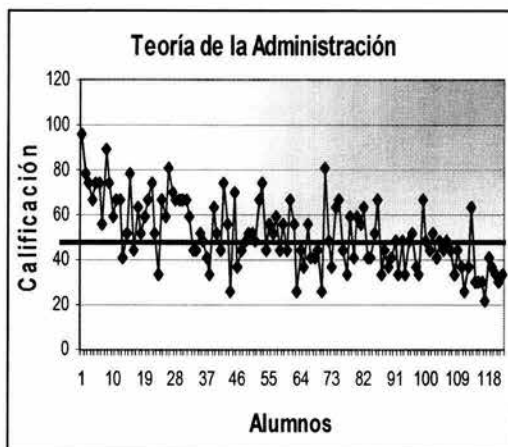


Figura 8. Calificaciones obtenidas por cada uno de los alumnos en la materia de Teoría de la Administración.

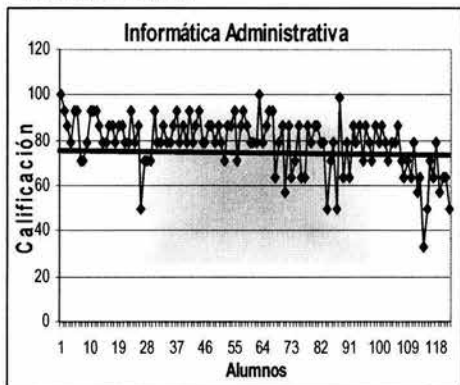


Figura 9. Calificaciones obtenidas por cada uno de los alumnos en la materia de Informática Administrativa.

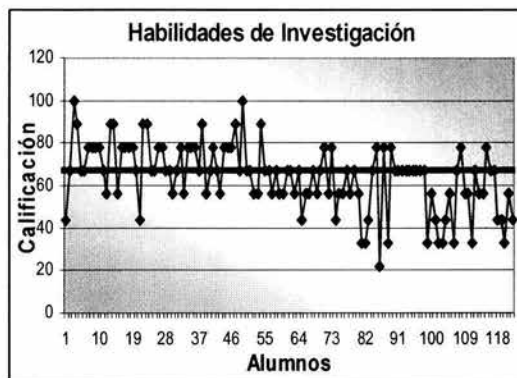


Figura 10. Calificaciones obtenidas por cada uno de los alumnos en la materia de Habilidades de Investigación.

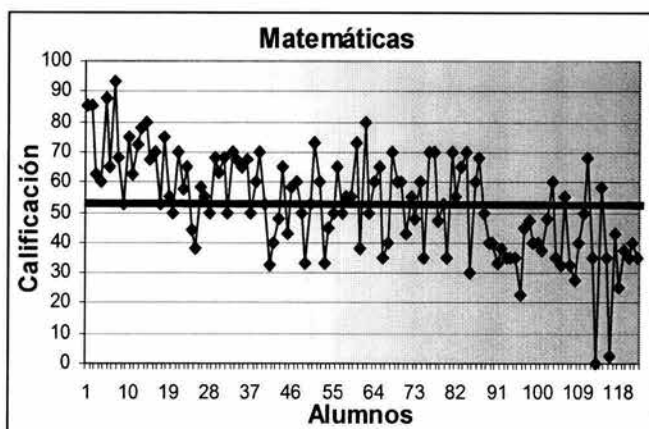


Figura 11. Calificaciones obtenidas por los alumnos en la materia de Matemáticas.

En la tabla 14 se muestra la media aritmética obtenida por los alumnos en cada una de las materias que conformaron la prueba. En esta tabla se puede apreciar que en la materia de Informática Administrativa se encontró que la media de aciertos fue mayor en comparación con la de Análisis Económico Administrativo, que fue la que menos aciertos presentó. La materia de Habilidades de Investigación siguió a la de Informática en cuanto a cantidad de aciertos.

Tabla 14

Media aritmética obtenida por los aspirantes de la MAN en cada una de las materias que conformaron la prueba.

Materia	Cal. Máxima	Cal. Mínima	Media Aritmética
Análisis Económico Administrativo	94	5	38,35
Teoría de la Administración	96	26	51,28
Informática Administrativa	100	33	78,44
Habilidades de Investigación	100	22	64,04
Matemáticas	93	23	52,90

A continuación se presentan los resultados de datos demográficos obtenidos tras la aplicación del examen.

Tabla 15

Datos correspondientes al género, edad, escuela de procedencia y situación laboral de los alumnos que presentaron el examen de MAN, B-2006.

Datos	Descripción	Porcentaje
Género	Masculino	52.9%
	Femenino	25.8%
Edad	21 a 25 años	24.5 %
	26 a 30 años	25.8 %
	31 a 35 años	16.8 %
	36 a 40 años	6.5 %
	Más de 40	5.2 %
Escuela de procedencia	Particular	10.3%
	UNAM	6.5 %
	IPN	61.9 %
Área a la que pertenecen	Ciencias exactas	28.4 %
	Ciencias médico-biológicas	21.9 %
	Ciencias sociales	49.7 %

Por lo cuál se puede decir que los datos demográficos descritos hablan de algunas características de la muestra estudiada y de la manera en como se comportaron en cada uno de los reactivos. Estos resultados se presentan de manera cuantitativa con fines de representación de lo que se obtuvo tras la aplicación del examen de admisión 2006 – B, dicho lo cuál estos resultados necesitan estar sustentados por un estudio de validez de la prueba empleada, para que de esta manera se realicen estudios acerca de la relación que tendrían los datos con variables definidas como predictoras del éxito académico; una vez realizado esto se podría ampliar aún más la interpretación de tales resultados.

6.2 Resultados cualitativos

Se analizaron las 5 secciones que conforman el examen de acuerdo a las recomendaciones hechas por diferentes autores, mismos que se han mencionado en la sección de tipos de reactivos, dado lo cuál se tiene.

Materia: Análisis Económico

- ❖ Usa preguntas capciosas.
- ❖ Usa un lenguaje no tan claro para los alumnos.
- ❖ Tiende a presentar información para su contestación.
- ❖ Combina el formato de opción múltiple, con el de completamiento y el de verdadero /falso, el cuál se recomienda eludir.
- ❖ Los reactivos no fueron probados.
- ❖ Las instrucciones no son muy claras.
- ❖ En algunos reactivos la idea principal tiende a repartirse en las opciones.
- ❖ Presentan 5 opciones en el tipo de opción múltiple, lo que implica un número grande que implica más tiempo de contestación.
- ❖ La extensión de las respuestas en las preguntas es diferente.
- ❖ Existen respuestas que contienen parte de la pregunta.
- ❖ Usan frases como "Todas o Ninguna de las anteriores", "Sólo c y d", etc.
- ❖ En los reactivos de tipo verdadero /falso tienden a causar confusión al emplear la palabra "no" en la pregunta.
- ❖ En cuanto a las preguntas de completamiento, son recomendables para conocimientos que tienen que ver con hechos, tal como fechas o nombres y el empleo que se le da tiende a dirigirse hacia la memorización de conceptos.
- ❖ En reactivos de completamiento la calificación no representó objetividad máxima.
- ❖ Se usan preguntas incompletas, siendo preferibles las directas.

Materia: Teoría de la Administración

- ❖ Utiliza combinación de formatos: correspondencia, verdadero /falso y de opción múltiple.
- ❖ En cuanto a los reactivos de correspondencia, la manera de presentar el formato puede ocasionar confusión, debido a que se coloca una columna arriba y la otra debajo.

- ❖ También se presenta diversas formas de presentación, a las que no se incluyen instrucciones acerca del cómo contestarse.
- ❖ En los reactivos de verdadero /falso, algunas preguntas pueden verse como de opinión o triviales y se requiere de memorización de conceptos para su contestación. Las preguntas tienden a extenderse provocando que sean largas y complejas.
- ❖ En las preguntas de opción múltiple, se repite parte de la idea en las respuestas, cuando la mayor parte debe estar en la pregunta.

Materia: Informática Administrativa

- ❖ Utiliza el formato de opción múltiple y de correspondencia.
- ❖ En cuanto a las preguntas de opción múltiple, en algunas de ellas no existe correspondencia gramatical entre la pregunta y las opciones, dando lugar al descarte de opciones.
- ❖ Las opciones no son de la misma extensión y emplean en ellas frases como "Ninguna de las anteriores".
- ❖ Algunas preguntas tienden a utilizar información muy específica, que hace pensar que se tomó de manera textual.
- ❖ Presenta preguntas con información innecesaria para la materia evaluada.
- ❖ En los reactivos de correspondencia no se presentan instrucciones previas sobre cómo deben ser contestadas o si los números pueden repetirse. Las columnas presentan la misma cantidad de interrogantes, alentando la contestación por eliminación de opciones.
- ❖ Las preguntas tienden a ser muy extensas, alentando la memorización.

Materia: Habilidades de Investigación

- ❖ Presenta combinación del formato de opción múltiple con el de verdadero /falso, mismos a los que no se agregó instrucciones para saber el cómo sería la manera de contestar.

Materia: Matemáticas (álgebra y aritmética)

- ❖ No se nota objetividad en la calificación, por el hecho de que las respuestas son abiertas en algunos casos, incluso se contestaron en el propio examen. En algunos reactivos de opción múltiple se presentaron instrucciones, pero no son claras, lo que pudo ocasionar confusión.
- ❖ El número de opciones es diferente en cada caso.
- ❖ Los reactivos carecieron de una edición y de ser probados.
- ❖ En especial en esta sección, existe falta de validez debido a la forma de presentación de las preguntas y contestación de las respuestas.

Resultados Cualitativos respecto a los datos demográficos y de frecuencias

De acuerdo a la tabla 14 de la pág. 110, se puede apreciar que las calificaciones máximas se obtuvieron en la materia de informática administrativa y de habilidades de investigación, seguida por la de matemáticas, teoría de la administración y finalmente análisis económico.

En cuanto a los datos demográficos de la tabla 15 de la pág. 111, se puede ver que la maestría en Negocios de la MAN es más solicitada por el sexo masculino, representando el doble que el sexo femenino.

Con respecto a la edad, se observa que personas de 21 a 30 años, son los que más piden el ingreso a la maestría, lo que se relaciona con el hecho de que la mayor parte de la población de dicha maestría es gente joven.

Dentro de la escuela de la que proceden, se observa que el 61.9% son de escuelas del propio IPN. Un tanto ligado a la escuela de procedencia, se encuentra el área a la que pertenecen los alumnos, es decir, la carrera que estudiaban, dentro de lo que se encuentra que los del área de las ciencias sociales son los que mayor demanda presentaron, lo que se relaciona con carreras administrativas; se sigue con el área de las ciencias exactas y finalmente con la médico- biológicas, en donde resalta un porcentaje representativo del 21.9%.

Una vez revisados y analizados los datos arrojados de los resultados cuantitativos y cualitativos, se procede a ligar estos datos con la bibliografía presentada, con la finalidad de elaborar así una propuesta que ayude a mejorar las fallas que presente el examen, de manera tal, que en el siguiente apartado se presenta la propuesta realizada con la finalidad de que no se discurra más en errores que pueden ser solucionados, tomando en cuenta que todo proceso puede tenerlos, pero que depende de las personas encargadas de dicho proceso, él poder modificarlo; para culminar con las conclusiones a las que se ha llegado, seguido de las sugerencias que se hacen para estudios futuros y las posibles limitaciones a las que me he enfrentado durante el proceso.

VII) PROPUESTA DE MEJORA AL EXAMEN DE ADMISIÓN DE MAN

7.1 Exposición de motivos

Se debe tomar en cuenta algunas consideraciones previas, debido a que la política educativa se da en función de la evaluación de la calidad y en la ES se contemplan aspectos para lograrla, tal como los Exámenes de Calidad de la Educación Superior (ECAES), la acreditación de programas y los estándares de calidad, visto lo cuál, autores como Otalvaro & Jácome (2003) indican que no sólo los ECAES son suficientes para valorar la calidad de la educación superior, también para valorar los aprendizajes, para hacer necesario el reconocer coherencias en los procesos; por otro lado los resultados de los estudios publicados, así como los datos obtenidos por el examen indican que la prueba elaborada es un instrumento defectuoso, que por razones obvias se restringe la presentación del mismo.

Algunas de las fallas que presenta, se muestran a continuación.

No se planea en base a los objetivos señalados en el perfil de ingreso.

- El contenido suele dirigirse hacia lo trivial, eliminando cosas importantes.
- No permite estimar el proceso mental o los contenidos que mide la prueba.
- Se utilizan diferentes tipos de reactivos, no pensado en la eficacia de estos para medir ciertos tipos de objetivos, sino más bien en el hecho de presentar una diversidad de formatos.
- La mayor parte de las preguntas son ambiguas, dado que las expresiones no son claras, haciendo que se obligue a los alumnos a adivinar.
- Las características de la prueba son inadecuadas para los fines con los que se desea utilizar los resultados, sin que se exprese el propósito por el que se elaboró de esa manera la prueba.
- No se realizó una definición operacional y los reactivos elaborados no reflejan los objetivos a los que debería dirigirse.

De manera que el examen de conocimientos empleado para la selección de alumnos que deseen ingresar a la MAN, debe considerar lo descrito en el artículo 32 que habla del Ingreso y Revalidación en el Reglamento General de Estudios del IPN

(IPN, 2006), en el que se hace constar los lineamientos a considerar para diseñar el examen de admisión, entre lo que destaca el perfil del aspirante de acuerdo al plan de estudios vigentes (ver anexo 1), y debido a que es uno de los elementos que aparecen en el reglamento, se sugiere el considerarlo, especificando además y de manera clara, lo siguiente:

- a. Objetivos del ciclo que se cubrirá en la maestría.
- b. Estructura del ciclo.
- c. Objetivos de enseñanza – aprendizaje para cada una de las materias.
- d. Perfil de ingreso del aspirante, redactado en términos de las habilidades y conocimientos necesarios a cubrir.
- e. Contenidos temáticos de cada una de las materias que se evaluarán y que han de ser representados en cada uno de los reactivos que conformaran la prueba final.

7.2 Objetivo

Objetivo general:

Diseñar una propuesta que contribuya al mejoramiento y optimización de elaboración del examen de admisión a la Maestría en Ciencias con especialidad en Administración de Negocios, basado en una planeación estratégica destacando su confiabilidad y validez.

7.3 Metas del programa

1. Identificar las capacidades académicas de los aspirantes que desean ingresar a la MAN, con la finalidad de evitar el rezago estudiantil.
2. Contar con elementos de juicio con el propósito de que el alumno organice su proyecto de maestría y así lograr la eficiencia Terminal.

Para ello se considera el hecho de que el examen debe aplicarse en respuesta a las necesidades de demanda de la población estudiantil, hacia aquellos que muestren las competencias necesarias para su desarrollo dentro de la maestría, por lo cuál también se pretende ayudar a la institución a:

- a) Contar con criterios para la implantación de un proceso de selección de alumnos para la toma de decisiones.
- b) Proporcionar información académica útil a la coordinación y al personal docente.

Para lograr que se obtengan los resultados esperados, es de gran importancia seguir el orden que se propone en este estudio, el cuál podrá variar en cuanto al tiempo, de acuerdo a las necesidades que surjan en el proceso.

7.4 Planeación

Antes que nada, se plantea que el examen debe ser utilizado como instrumento auxiliar de evaluación del aprovechamiento del alumno a través de la comprensión del saber recibido, e indicar que las pruebas son una forma de evaluación que contribuye al mejoramiento del proceso de enseñanza - aprendizaje, de manera que la prueba a realizar debe ser con fines diagnósticos, que funja como proceso de decisión para la colocación de los estudiantes en los cursos correspondientes.

De manera que la programación en la elaboración de la prueba, se toma de Thorndike (1996) quien recomienda la forma de trabajar "hacia atrás" desde la fecha en que la prueba debe estar lista para usarse, menos meses, semanas o días en los que deba terminarse cada paso preliminar, para facilitar el desarrollo del programa.

Día del examen	(1)	Prueba final disponible para aplicación a los aspirantes de la MAN, así como los materiales auxiliares, manuales y hojas de respuesta impresas.
30 días antes	(2)	Copias finales de la prueba, así como las hojas de respuesta.
60 días antes	(3)	Análisis estadístico terminado para las pruebas y entregadas al elaborador de la prueba para su uso en la preparación final de la prueba.
100 días antes	(4)	Proceso terminado y hojas de respuesta en manos de los analistas para la obtención de la confiabilidad y de la validez.

130 días antes	(5)	Forma final de la prueba e instrucciones para su aplicación, entregados a los aplicadores para realizar el proceso de confiabilidad y validez (3er. ensayo).
140 días antes	(6)	Copias de la forma final de la prueba y materiales auxiliares.
160 días antes	(7)	Banco de reactivos terminado, así como los materiales auxiliares, entregados a la instancia correspondiente para su revisión.
200 días antes	(8)	Todos los análisis estadísticos junto a información del ensayo en manos del elaborador de la prueba, para selección de los reactivos que se someterán a la elaboración del banco de reactivos de la TRI (2do. ensayo).
230 días antes	(9)	Ensayo informal, así como resultados entregados para posibles cambios en el examen o en su forma.
260 días antes	(10)	Material para el primer ensayo con las instrucciones, entregado a los aplicadores para su aplicación.
270 días antes	(11)	Copias de las pruebas y hojas de respuesta entregadas. Capacitación de los aplicadores.
280 días antes	(12)	Reactivos revisados en el contenido y forma entregados al revisor con la finalidad de ser incluidos en la prueba para su ensayo informal, así como hojas de respuesta.
310 días antes	(13)	Borradores de reactivos a los revisores para evaluación del contenido.
340 días antes	(14)	Asignación de elaboración de reactivos a los redactores.
370 días antes	(15)	Diseño y especificaciones de la prueba y del tipo de formato de reactivo a ser utilizado.
400 días antes	(16)	Inicio de la planeación de la prueba: definición de objetivos y elaboración de la tabla de especificaciones.

Presentadas las fases recomendadas por las que debe discurrir el proceso, a continuación se irá dando explicaciones de cada uno de los pasos, de acuerdo también a lo que se menciona en la primera parte del marco de referencia teórica.

Fase 16 – Inicio de la planeación de la prueba

Cabe indicar que la planeación es en base a una prueba que refiere a un criterio, que servirá para describir y verificar las tareas de conocimiento y la medida en que un estudiante ha dominado ciertas áreas de conocimiento.

- a) Definición del constructo a medir, identificando los objetivos y contenidos de aprendizaje.

Con el examen se mide el nivel de conocimientos que tengan los aspirantes a cursar la MAN, en las cinco materias que conforman la prueba.

Finalidad 1: El examen de admisión a la MAN es un instrumento educativo de aplicación colectiva que sirve para medir conocimientos en el área de matemáticas aplicadas a la administración, análisis económico, teoría de la administración, informática administrativa y habilidades de investigación, estrechamente relacionadas con el aprovechamiento académico, para proporcionar mediciones confiables y rápidas del rasgo que se mide y que cubre el perfil de ingreso de dicha maestría.

Finalidad 2: El examen de admisión a la MAN es un instrumento educativo que sirve para medir cinco áreas principales o conocimientos en ellas. En conjunto, las cinco áreas de conocimientos permiten hacer una evaluación completa del nivel que poseen los aspirantes que desean ingresar a la maestría; siendo un instrumento de aplicación colectiva, que proporciona medidas confiables y validas.

- b) Aspectos de diseño preliminares
 - Modo de aplicación: colectiva debido a la cantidad de examinados.
 - Duración: prolongada (aproximadamente 3 horas).
 - Formato de los reactivos: tipo objetivo de forma opción múltiple.
 - Informes de las puntuaciones: de manera escrita a través de carteles y de forma electrónica en la página de Internet de la maestría, proporcionando la puntuación total en la prueba y del desempeño en cada una de las materias evaluadas.

- Capacitación de aplicadores: se necesitarán personas a las que se les capacitará, acerca de la forma de aplicación y calificación del instrumento, lo cuál se puede realizar a través de una sola sesión con un horario a definir.

Los temas de cada materia quedarán representados en la tabla de especificaciones que se indica a continuación, recordando que se deben elaborar el triple de lo que se tiene pensado que abarcará en la prueba final.

Un aspecto importante a considerar en esta parte es el referente a la lista de resultados de aprendizaje a medir, con la finalidad de facilitar la elaboración de la tabla de especificaciones, por lo cuál se recomienda el siguiente formato.

Áreas a evaluarse	Objetivos por cada área	Nivel de la taxonomía
Es lo referente a cada una de las materias que se evalúan en el examen.	Se deben enunciar aquellos que se consideren importantes para el examen por cada área, considerando también la filosofía y el perfil de ingreso de la maestría.	Será tomando en cuenta lo que implica la utilización de cada nivel, enlistando las conductas a ser evaluadas en los alumnos de acuerdo a los verbos correspondientes a cada objetivo.

c) Tabla de especificaciones

Debe recordarse que la cantidad de preguntas para cada área de contenido debe corresponder a la importancia que se da a cada tema y la cantidad de reactivos para cada objetivo debe corresponder a la importancia que según el maestro tiene pensado dar al proceso para el que se redactará.

A continuación se presenta un bosquejo de un tipo de tabla de especificaciones que puede diseñarse para facilitar la planeación del número de reactivos que contendrá la prueba. Este paso debe diseñarse una vez que se hayan enlistado los temas que abarcará el examen, lo que se llevará a acabo en coordinación con cada experto en la materia.

Desglose de los objetivos en subtemas, asignando un peso relativo a cada uno, de acuerdo a la jerarquía que tengan, lo cuál debe realizarse en colaboración con los

profesores de cada asignatura, para obtener el número de reactivos que abracará cada materia, de acuerdo al siguiente cuadro.

Materia	Peso relativo %	Reactivos para cada materia	# total de reactivos
1		Peso relativo * # total de reactivos /100	
2			
3			

Posteriormente los datos que se obtengan se emplearán en la tabla de especificación que se muestra a continuación.

Materia	Peso relativo %	# de objetivos de aprendizaje a evaluarse	Conocimiento	Comprensión	Aplicación	Análisis	Síntesis	# total de reactivos
1								
2								
3								
4								
5								
	100%							# de reactivos pensados para la prueba

Cabe recordar que el nivel de la taxonomía para cada resultado de aprendizaje dependerá del nivel al que se desee medir, lo cuál se facilitará al enunciar cada objetivo, empleando los verbos de cada nivel de la taxonomía, para lo cual se recomienda ver el ejemplo de la p. 51.

Fase 15 - Preparación de los reactivos

Incluye la redacción y su revisión, recordando los elementos de cada reactivo, entre los que destaca el estímulo o pregunta, el formato, las condiciones y la forma de calificación, los que pueden revisarse en el apartado correspondiente; en esta parte se ha escogido una prueba objetiva y de selección múltiple, puesto que ofrece mayor

seguridad de la validez. En la **fase 14** se debe elegir a los maestros que redactarán los reactivos para cada materia, mismos que habrán enlistado los temas para cada una. Por definición, en una prueba en la que las preguntas presentan opciones de entre las que se puede identificar la respuesta correcta que ha sido previamente preestablecida, y en la que la calificación no se ve afectada por el juicio de quien califica, sino que se vuelve más bien en algo mecánico, sin tener que llegar a estas diferencia de juicio u opiniones, y en las que el tiempo para responder es menor que en una prueba "subjetiva", como en las de ensayo, y por el hecho de permitir el acceso a los resultados por parte de los estudiantes en menos tiempo, punto que influye en la motivación de los aspirantes, hace resaltar los beneficios de su utilización, aún más cuando se diseña para un número grande de personas, que hace más factible su utilización en cuanto a un muestreo más adecuado del contenido.

En cuanto al formato, se ha indicado que se propone el de opción múltiple, dado que tras la revisión de la bibliografía correspondiente, y estudiadas las ventajas y limitaciones de cada formato, se concluye en su utilización por el hecho de medir resultados del conocimiento y varios tipos de capacidad intelectual, por otro lado el de contener en sí mismo las normas y en el hecho de que permite medir los logros de aprendizaje, desde el más simple hasta el más complejo, lo que hace que se sustente aún más la utilización de éste formato, con el cuál también se logrará homogeneidad en los reactivos.

En cuanto a las opciones de respuesta, a pesar que tres opciones parece ser un límite adecuado y de ser el número recomendado por diferentes autores, también se considera el hecho de que entre menos opciones, mayor probabilidad existe de ser contestadas al azar, por lo que se propone lo recomendado por Quesada (1991) a lo que refiere el emplear cuatro opciones, lo cuál se considera un número adecuado y considerado el de mayor uso en la práctica.

Además de las consideraciones anteriores, se debe tomar en cuenta que como regla empírica, se deben preparar entre dos y tres veces más de los reactivos que se necesitarán para la prueba final (Hogan, 2004).

¶ Esto es lo correspondiente a la **fase 13** de la programación, en esta fase se tendrá el borrador de los reactivos elaborados y se mandarán a los revisores para su

edición y corrección, para conformar la prueba que será sometida al primer ensayo (**fase 12**). Los reactivos que se redacten deben ser sometidos a la fase informal, en la que se aplicarán los reactivos a unos 5 o 10 sujetos pertenecientes a la maestría, con la finalidad de llevar a cabo el proceso de revisión en función de su claridad, corrección gramatical, así como las instrucciones. Además se debe hacer una revisión de la corrección del contenido, la cuál la realizan los expertos en el campo temático; lo que se hará con la finalidad de identificar los reactivos que causen problemas, para mejorarse o eliminarse, y una vez realizados estos cambios se procede a realizar el ensayo de los reactivos.

En este paso se recomiendan las siguientes instrucciones:

Esta prueba dura --- horas. Consta de --- preguntas. Usted debe marcar sólo una respuesta para cada pregunta. La respuesta que piense que es correcta deberá señalarla relleno sobre la letra que corresponde en la hoja de respuestas(ver el ejemplo que aparece abajo). Tal procedimiento deberá seguirlo en cada uno de los problemas presentados. No haga otra clase de marcas sobre la prueba. No pierda su tiempo tratando de contestar preguntas que se le hagan muy difíciles, es mejor que no se detenga y que continúe respondiendo hasta terminar, para que una vez que haya acabado el examen regrese a las preguntas que dejó sin contestar. Su puntuación va a depender del número de respuestas correctas que marque.

¶ También se deben tener preparadas las hojas de respuestas, para que una vez conformada la prueba, se mande a fotocopiar junto con el examen, para ser aplicado al grupo de sujetos del ensayo (**fase 11**). Un formato propuesto de hoja de respuesta aparece en el anexo 3, aclarando que el número de opciones dependerá de los contenidos planteados en la tabla de especificaciones, por lo cuál, dicha hoja presenta un formato sujeto a cambios.

¶ Fase 10: una vez terminado, se debe mandar el material a los aplicadores, mismos a los que se les dará una capacitación, anterior a este 1er. ensayo, en la que se les indicará la forma de aplicación, el como establecer “rapport” adecuado, como comportarse antes, durante y después de la aplicación, cómo ordenar las pruebas, forma de calificar, el registro de aspirantes, e indicaciones auxiliares. Cabe mencionar que sea desde este punto, con la finalidad de que se logre familiaridad con el proceso y se facilite en ocasiones subsecuentes.

¶ Fase 9: se procede a realizar el 1er. ensayo en una muestra pequeña de sujetos, debido a que tras la aplicación de esto, se espera obtener las observaciones hechas y las posibles correcciones al formato del examen, para proceder al segundo ensayo.

¶ Fase 8: una vez que se obtienen los resultados de la fase 9, se procederá a realizar el 2do. ensayo, en el que se obtendrá el banco de reactivos, destacando que será dentro del procedimiento de la Teoría de Respuesta al Ítem, debido a que tras lo reportado en la información bibliográfica, se observa que supera las desventajas que se dan al emplear la TCP. Para este paso se requerirá de muestras grandes (alrededor de 500 sujetos), lo cuál se hará siguiendo las recomendaciones de Muñiz (1996, 1997, 2001), dadas las ventajas de contar con un banco de reactivos.

¶ Fase 7: los resultados se entregarán a la coordinación de MAN para ser revisados y conformados en la prueba que será aplicada a los aspirantes.

¶ Fase 6: se fotocopiaran para ser administrados a una muestra más pequeña de sujetos que pertenezcan a la maestría y así llevar a cabo el 3er. ensayo que es el correspondiente a la **fase 5**, para la obtención de la Confiabilidad y de la Validez.

En cuanto a la Confiabilidad por tratarse de una prueba de conocimientos, es preferible la confiabilidad a través de pruebas paralelas, debido a que es más adecuada para los fines de la prueba y por el hecho de que al contar con un banco de reactivos, se facilita el que ambas formas presenten el mismo contenido y dificultad. En cuanto a la consistencia interna del instrumento, como presentará preguntas con cuatro opciones de respuesta, se recomienda el coeficiente alfa de Cronbach. Cabe mencionar que no

es requisito indispensable el calcular todos los tipos de confiabilidad, más bien elegir aquella que el instrumento necesite.

Una vez obtenido este elemento, se procedería a evaluar si el instrumento mide lo que debe medir, de manera que se debe evaluar si contiene validez de contenido, debido a que una prueba de aprovechamiento, requiere de éste tipo de validez, tal como lo mencionan Nunnally & Berstein (1997), paso que deben realizar los expertos en el tema, a través de un acuerdo interjueces. Además debido a que el objetivo del estudio, es el de elaborar un instrumento que mida un constructo, se debe probar la validez de construcción y se concluirá con la **fase 4**.

✚ **Fase 3:** una vez que estos análisis psicométricos se encuentren terminados, se entregarán los resultados al elaborador de la prueba para completar el manual, sin olvidar informar las características de la muestra empleada para la obtención de la confiabilidad y la validez.

✚ **Fase 2:** se tendrán las copias de la prueba final a ser aplicadas, así como las hojas de respuesta, ambas deben estar ya fotocopiadas y foliadas en orden para cada uno de los alumnos aspirantes a la maestría, lo cuál correspondería a la **fase 1**, que sería con el material completo y finalizaría así el proceso de elaboración del examen.

Con estos puntos se habrá culminado la elaboración del examen de admisión a la MAN, sin embargo, es bueno tomar en cuenta otras consideraciones importantes.

7.5 Consideraciones finales

☐ Las instrucciones deben aparecer en negrita y el aplicador debe leerlas a los examinados, preguntando si ha quedado alguna duda al respecto.

☐ Se recomienda que en el momento del registro de los aspirantes, se les de o haga llegar por correo electrónico, el formato correspondiente a la hoja de respuesta e información general del examen, con la finalidad de evitar la falta de familiaridad con los formatos, o de algún tipo de angustia o tiempo perdido por su desconocimiento.

■ También se recomienda que para que la prueba sea aplicada de manera estandarizada, desde el primer ensayo se coloquen hojas en las puertas con la leyenda “No molestar, se está aplicando examen”, para evitar interrupciones; así como preguntar antes del examen si ya han ido al baño o tienen algo pendiente, debido a que durante la aplicación no se permitirán salidas.

■ Es conveniente que en caso de que el aplicador necesite algo, se le proporcionen hojas que puedan pegar en la puerta, donde escriban que necesitan y que haya una persona que vigile los pasillos y que este al pendiente de estas peticiones.

■ Citar a los aplicadores 30 minutos antes de la hora del examen, para que cada uno revise el material que se le entrega y que debe entregar al finalizar éste, checar que los materiales estén completos, así como para dar posibles indicaciones finales.

■ Proporcionar información a los alumnos que no han sido aceptados, acerca de aquellos factores por lo cuales no pudieron ingresar, con la finalidad de que si desean volver a realizar el proceso, recuperen esas áreas en las que pueden mejorar.

■ Establecer una correspondencia entre los contenidos que abarcará el examen y los citados en la guía de estudios, con la bibliografía correspondiente.

■ Realizar una pequeña platica informativa a los aspirantes a la maestría, acerca del proceso que deben seguir para ingresar e información del propio examen, que servirá como proceso de inducción.

VIII) CONCLUSIONES

8.1 Conclusiones

En este apartado se abordan las principales ideas derivadas de la información revisada y los resultados observados, lo que llevó a la elaboración de la propuesta planteada, para culminar con esta sección.

Por lo tanto se puede concluir que las calificaciones en el examen, en su mayoría representan un alto índice de reprobación, debido a que la media máxima en el examen global se encuentra muy por debajo del 80 que se pide para el ingreso, lo que ocasionó que se aceptarán en ese período (2006-B) alumnos que obtuvieron calificaciones bajas (a partir de 52), lo anterior representa una problemática puesto que los aspirantes no cumplieron con la calificación establecida, lo que puede atribuirse de cierta manera a la falta de correspondencia entre sus conocimientos y lo que se planteaba en el examen.

Siguiendo con la descripción de los datos, se puede observar que en cuanto al género, en su mayoría corresponde al sexo masculino, lo que puede indicar que los hombres tienen una mayor preferencia por la maestría en el ámbito de los negocios, sin embargo, esto se convierte en una hipótesis que puede ser probada por estudios futuros y que a la vez pudiese representar una línea de investigación; lo mismo ocurre con la edad, ya que se puede pensar que las personas más jóvenes prefieren seguir con sus estudios al egresar de la carrera.

Con respecto a la escuela de procedencia, se observa que existe una marcada inclinación en aquellos que egresaron de escuelas del propio IPN, lo que se convierte en un punto importante que indicaría que prefieren seguir en el mismo instituto del que egresaron, indicando que la mayor parte son de carreras de áreas sociales, lo que liga al hecho de que de cierta manera presentan los conocimientos mínimos para el ingreso a la maestría, haciendo que surja una correspondencia entre lo que se estudió y lo que desean seguir estudiando.

El punto anterior alude al hecho de que si la mayoría provenía del área administrativa quiere decir que de cierta manera cuenta con los conocimientos básicos en correspondencia con el perfil de ingreso y sin embargo las calificaciones globales

fueron bajas, hecho que pone en duda éste antecedente o el que quizá se pueda atribuir a la falta de planeación y conocimiento en la elaboración de pruebas, que además, fue un proceso sin una preparación previa o que se hace necesario la capacitación del personal encargado de llevar a cabo el proceso de admisión, de manera que se tenga un mejor nivel de conocimientos en el tema.

De manera tal, que al armar un conjunto de las limitaciones que presentó el examen, se tiene de manera general, dificultades en cuanto al contenido, formato, estilo, redacción y opciones, lo que marca pautas a considerar al elaborar un reactivo de selección de respuesta (Haladyna, 1999; en Hogan, 2004).

Además se hace referencia al hecho de que no se haya podido realizar un análisis estadístico, debido a las características propias del examen, tales como la falta de homogeneidad en el formato y el de incluir preguntas de tipo ensayo, lo cuál disminuyó aún más la objetividad de la calificación y al ser una herramienta que sirve para la toma de decisiones, se toma en consideración lo que menciona Morales (1975) en el hecho de que la falta de confiabilidad, refleja falta de objetividad y por lo tanto carencia de validez, que son elementos de vital importancia en las pruebas, de lo contrario se aceptarían a personas que a pesar de ingresar a la maestría no cumplieran con los demás requisitos solicitados en el perfil de ingreso, por lo cuál es favorable indicar que un examen bien elaborado puede fungir como un indicador de éxito académico, tal como se indica en los estudios de Aguilar, Lizarraga & Pérez (1981); Venegas (1993) y Alonso (1998).

El hecho de que la prueba no presente elementos de validez hace difícil el concebir su aplicación, sin embargo, este es un punto que se ha manifestado en muchos otros estudios realizados sobre el tema, tal como se indica en Aguilar et. al. (1981); Carrera & Romo (1983) y Venegas (1993), en los que tras los resultados de sus estudios se observó que las pruebas de aprovechamiento carecían de validez, aspecto que dentro del área de evaluación educativa debe representar un asunto de interés y a la vez una problemática dentro del área de la educación, a la que se debe prestar una mayor atención.

En gran medida el formato del examen aumentó la probabilidad de que este fuese contestado casi al azar y el que su calificación se tornara difícil, que lo convierte

en una prueba deficiente para los fines a que se destina, hecho que se vio reflejado en los propios resultados de la prueba, que sin lugar a dudas da mucho para discutir, sin embargo se hace necesario remarcar que al no tener bien especificado desde el principio el constructo que se desea medir, se está cayendo en un gran error, que concluirá en una prueba mal elaborada y carente de validez. También es importante indicar que no se pueden emplear reactivos que antes no han sido editados por expertos en la materia y probados en poblaciones similares a la que será meta, de lo contrario, no se tendrá el conocimiento sobre si realmente discriminarán.

Por tal motivo, al realizar esta investigación, se llegó a la conclusión de que el examen de admisión a la MAN del período B – 2006 mostró grandes deficiencias dado los resultados arrojados en el estudio en relación con la estructuración del mismo, lo cuál me llevó a mostrar bases sólidas que indican que el hecho de no realizar una adecuada planeación y elaboración en un examen acarrea grandes problemas, de entre los que destaca la falta de validez e inexactitud en los conocimientos que debe cubrir en base al perfil de ingresos y en cuanto a las mediciones hechas a través de los estudiantes. Por lo que este estudio presenta un área de mejora para la maestría, aspecto que llevó a elaborar una propuesta donde se incluyen los pormenores de la elaboración de la prueba y algunas recomendaciones para hacerlo, en base a la bibliografía y los resultados estudiados y analizados.

Realizar una prueba no es un proceso fácil, es algo que demanda tiempo y esfuerzo, personal capacitado, así como la colaboración de expertos en cada tema de estudio. Sin embargo, considero que daría buenos resultados hacerlo, sobre todo por el hecho de que se destina a un proceso de gran importancia, que es la selección de alumnos y por que se tiene que cumplir con lo estipulado en el Reglamento de estudios del IPN, en el apartado del Ingreso y Revalidación de alumnos, para que de esta manera se pueda seleccionar a las personas indicadas para la maestría indicada, hecho que a futuro puede llevar al aseguramiento de la calidad educativa requerida en el mercado laboral.

8.2 Sugerencias y limitaciones

Este estudio se diseñó a partir de una necesidad observada en el escenario de trabajo, lo cuál se logró a través de una detección de necesidades organizacionales, llevado a cabo por medio de entrevistas inestructuradas y registros propios acerca de las posibles áreas de mejora en las que se podía intervenir, lo cuál se plantea a la vez como una sugerencia debido a que al realizar una evaluación en la población u organización en la que se intervenga, ayudará a determinar las áreas en donde se podrá trabajar y así realizar las adecuaciones necesarias a considerar.

Se sugiere a la vez que en estudios futuros se considere el tiempo de la intervención para trabajar en una sola área de mejora, debido a que se puede caer en el error de querer hacer más, pero el tiempo dentro de una intervención no lo permite, lo cuál en este estudio fue la principal limitante.

Otra sugerencia es el continuar con el estudio debido a que daría grandes resultados y así, poder relacionar distintas variables, como por ejemplo las demográficas con el promedio de ciclos anteriores, etc.

Otro punto importante es el que las escuelas abran sus puertas a los psicólogos que buscamos oportunidades, para ayudar y a la vez aprender, en el caso de la MAN considero que existen muchas otras áreas que necesitan de esta ayuda, tal como lo es la elaboración y realización de las entrevistas, así como cursos de capacitación a personal de la misma en el tema de elaboración de reactivos y del proceso de reclutamiento para los alumnos, la selección de los mismos y su inducción.

Como se indicó, dentro de la intervención la principal limitante fue el tiempo, puesto que hubiese sido gratificante el poder presentar el examen terminado y validado, pero como ya se ha mencionado, es un proceso que requiere de más sesiones, sin embargo; se ha logrado estructurar una propuesta, que pueda en gran medida reducir el tiempo y así lograr el cumplimiento del objetivo planteado, para que pueda concluirse el estudio de manera satisfactoria.

Considerando que este es un proceso que requiere dedicación se recomienda el que un experto intervenga junto con la colaboración del personal y profesorado de la

Maestría en ciencias con especialidad en Administración de Negocios, que lejos de ser algo arduo, puede convertirse en algo productivo, satisfactorio y sobre todo justo.

Finalmente, se ha de considerar que no es un proceso rígido, pero si importante, de manera tal que se debe tomar en cuenta que el método o proceso que se decida seguir sea el que mejor resultados produzca y sobre todo tener siempre presente que se está tratando con personas, mismas que piden un lugar dentro de la MAN.

REFERENCIAS

- ▶ Abreu, L. (2003). Los nuevos paradigmas para la planeación del posgrado. *OMNIA. Revista de la Dirección General de Estudios de Posgrado, UNAM*. Número especial (19), 49-54.
- ▶ Adkins W. D. (1981). *Elaboración de test. Desarrollo e interpretación de los test de aprovechamiento*. México: Trillas.
- ▶ Aguilar, V. J., Lizárraga, R. S. & Pérez, P. L. (1981). Investigación diagnóstica de los estudiantes que ingresan a la carrera de Psicología. *Enseñanza e Investigación en Psicología*, 7 (1 y 2), 298- 306.
- ▶ Aiken, L. (2003). *Psychological Testing and Assessment*. Undécima. Edición. México: Prentice Hall.
- ▶ Alonso, A. J. (1998). *Validez predictiva de los instrumentos utilizados en el proceso de selección y admisión en la Facultad Mexicana de Medicina*. Tesis para obtener el grado de Maestría. UNAM. México, D.F.
- ▶ Amezcua, K. (2004). *Evaluación docente por medio de los estudiantes en el nivel de enseñanza superior, una experiencia en la Facultad de Psicología, UNAM*. Tesis de Licenciatura no publicada. Facultad de Psicología, UNAM. México, D.F.
- ▶ Anastasi, A. (1998). *Tests psicológicos*. 7ª. ed. México: Prentice may.
- ▶ Arias, R. C. (1990). *Psicometría: Teoría de los tests psicológicos y educativos*. España: Síntesis Psicológica.
- ▶ Bloom, B. S. (1977). *Evaluación del aprendizaje*. Buenos Aires: Troquel.

- ✦ Brown, F. (1992). *Principios de medición en psicología y educación*. México: Manual moderno.
- ✦ Carreño, F. (1993). *Instrumentos de medición del rendimiento escolar*. 2ª. Edición.
- ✦ Carrera, S. C. & Romo, C. M. (1983). Examen de admisión a la Universidad Iberoamericana. Un programa de evaluación para los aspirantes de primer ingreso. *Revista de la Educación Superior*. 47. 11-159.
- ✦ Centro Nacional de Evaluación para la Educación Superior, A. C., CENEVAL, (2001).
- ✦ Chávez, L. L. (2000). *Estudio sobre la estructura, resultados y modificaciones del examen de conocimientos utilizado en la evaluación del personal bibliotecario administrativo de base de la UNAM*. Tesis de Licenciatura no publicada. Facultad de Filosofía y Letras. UNAM, México, D. F.
- ✦ Cohen, R. & Swerdlik, M. (2000). *Pruebas y evaluación psicológicas. Introducción a las pruebas y a la medición*. 4ª. Edición. México: Mc Graw Hill.
- ✦ Cortada de Kohan, N. (1999). *Teorías psicométricas y construcción de tests*. Buenos Aires, Argentina: Lugar editorial.
- ✦ Didou, A. S. (2005). ¿Medir para conocer o para regular?: evaluación y acreditación de la educación superior en México. *Avance y perspectiva (México)*. No. 1, (24), 31-35.
- ✦ Elorza, H. (2001). *Estadística par las ciencias sociales y del comportamiento*. 2ª. Edición. México: Oxford.
- ✦ García, E. (1993). *Introducción a la psicometría*. Madrid, España: Siglo XXI de España Editores.
- ✦ Gronlund, N. (1980). *Elaboración de tests de aprovechamiento*. México: Trillas.

- ✦ Gronlund, N. (1985). *Medición y valuación de la enseñanza*. México: Pax – México.
- ✦ Gronlund, N. (1998). *Assessment of Student Achievement*. 6ª. Edición. United States of America: Viacom Company
- ✦ Hills, J. (1982). *Evaluación y medición en la escuela*. México: Colección Actualizada Pedagógica.
- ✦ Hogan, T. (2004). *Pruebas psicológicas. Una introducción práctica*. México: El Manual Moderno.
- ✦ <http://escasto.ipn.mx/escuela/historia/historia.html>
- ✦ <http://sepi.escasto.ipn.mx>
- ✦ Instituto Politécnico Nacional, Escuela Superior de Comercio y Administración (1995). *POCHTECA, ESCA, 150 años de vida (1845- 1995)*. México: Dirección de publicaciones del IPN.
- ✦ Instituto Politécnico Nacional, IPN (2006). *Gaceta Politécnica del 31 de Julio del 2006*, N°. Extraordinario 633, Año XLI, (9), 29- 66.
- ✦ Karmel, L. (1978). *Medición y Evaluación escolar*. México: Trillas.
- ✦ Lafourcade, P (1969). *Evaluación de los aprendizajes*. Buenos Aires, Argentina: Kapelusz.
- ✦ Landsheere, G. (1978). *Los test de instrucción*. Barcelona, España: Oikos- tau, S. A. Ediciones.
- ✦ Magnusson, D. (1978). *Teoría de los tests*. México: Trillas.
- ✦ Martínez R. (2000). *Psicometría. Teoría de los tests psicológicos y educativos*. Madrid, España: Síntesis psicología.
- ✦ Morales, M. (1975). *Psicometría aplicada*. México: Trillas.
- ✦ Muñiz, J. (1996). *Psicometría*. Madrid, España: Universitas, S. A.

- ▶ Muñiz, J. (1997). *Introducción a la Teoría de Respuesta a los Ítems*. Madrid, España: Pirámide.
- ▶ Muñiz, J. (2001). *Teoría Clásica de los test*. Madrid, España: Pirámide.
- ▶ Nunnally, J. C. & Berstein, N. (1997). *Teoría psicométricas*. México: Mc Graw Hill.
- ▶ Otalvaro, C. & Jácome, S. (2003). El tránsito de los propósitos a los efectos en la evaluación masiva de la Educación Superior. *Pedagogía y Saberes*. 19. 95–104.
- ▶ Padua, J. (1978). *Paquete estadístico para las ciencias sociales (SPSS). Oferta y conclusiones para su utilización e interpretación de resultados*. México: El colegio de México.
- ▶ Pardo, A. (1980). *Elaboración de un examen de clasificación a universitarios*. Tesis de Licenciatura no publicada. Facultad de Psicología. UNAM. México, D.F.
- ▶ Quesada, R. (1991). *Guía para evaluar el aprendizaje teórico y práctico*. México: Limusa.
- ▶ Rosado, M. A. (1998). *Hacia un modelo de Evaluación de la docencia en la Educación Superior*. Tesis de maestría no publicada. Facultad de Psicología, UNAM. México, D. F.
- ▶ Sampieri, R., Collado, C., & Lucio, P. (2003). *Metodología de la investigación*. México: Mc Graw Hill.
- ▶ Thorndike, R. (1996). *Psicometría Aplicada*. 2ª. Edición. México: Limusa.
- ▶ Thorndike, R. & Hagen, E. (1973). *Tests y técnicas de medición en psicología y educación*. México: Trillas.

- ✦ Venegas, F. M. (1993). Evaluación del proceso de admisión a la carrera de Psicología de la Universidad Autónoma de Querétaro. Tesis de Maestría no publicada. Facultad de Psicología. UNAM. México, D. F.

ANEXOS
ANEXOS

Anexo 1

Visión del Instituto Politécnico Nacional

“Una institución educativa innovadora, flexible, centrada en el aprendizaje; fortalecida en su carácter rector de la educación pública tecnológica en México, con personalidad jurídica y patrimonio propios, con capacidad de gobernarse a sí misma; enfocada a la generación, difusión y transferencia del conocimiento de calidad; con procesos de gestión transparentes y eficientes; con reconocimiento social amplio por sus resultados y sus contribuciones al desarrollo nacional; con una posición estratégica en los ámbitos nacional e internacional de producción y distribución del conocimiento” (www.ipn.mx).

Misión del Instituto Politécnico Nacional

“El Instituto Politécnico Nacional es la institución educativa laica y gratuita de Estado, rectora de la educación tecnológica pública en México, líder en la generación, aplicación, difusión y transferencia del conocimiento científico y tecnológico, creada para contribuir al desarrollo económico, social y político de la nación. Para lograrlo, su comunidad forma integralmente profesionales en los niveles medio superior, superior y posgrado, realiza investigación y extiende a la sociedad sus resultados, con calidad, responsabilidad, ética, tolerancia y compromiso social” (www.ipn.mx).

Perfil de ingreso, permanencia y egreso de la MAN.

Perfil de Ingreso: Quienes aspiran a cursar esta maestría deben poseer: poder de abstracción, capacidad creativa, inclinación por la investigación, espíritu de liderazgo, necesidades de logro o éxito, solidez de conocimientos generales, y manejo del idioma inglés.

Perfil de permanencia: La duración de la maestría es de cuatro semestres como mínimo y ocho máximo para cumplir con los créditos establecidos en el programa de estudios; es decir que debe cubrir con 4 cursos

de homogenización que no cuentan con créditos, y 4 semestres con un total de 82 créditos, de los cuales 66 son de las materias curriculares, 10 obligatorias y 3 optativas, y 16 del examen de grado; la calificación mínima aprobatoria es de ocho, y el alumno que repruebe una materia deberá cursarla otra vez; quien repruebe dos materias, o la misma dos veces, es dado de baja.

Perfil de egreso: El egresado de MAN tendrá los conocimientos que, de acuerdo con su perfil e intereses, le permitan ubicarse como: directivo en organizaciones productivas y de servicios, especialista en algunas de las áreas funcionales que conforman las organizaciones, empresario independiente, consultor/asesor de alto nivel que oriente a funcionarios y directivos al mejor logro de los objetivos en sus organizaciones, investigador en instituciones educativas, catedrático de licenciatura y posgrado capaz de transmitir conocimientos, mejorar habilidades y fomentar actitudes críticas y emprendedoras (www.sepi.escasto.ipn.mx).

Gaceta



Acuerdo por el que se expide el Reglamento

Acuerdo por el que se expide el Reglamento de
Estudios
de Posgrado del IPN

Acuerdo por el que se expide el Reglamento para
la Operación, Administración y Uso de la Red
Institucional de Cómputo y Telecomunicaciones
del IPN



70

Aniversar
io
versaria
1936 -
2006

Ayyw spn
nrv<

ACUERDO

I. En el caso de las especialidades y de las especialidades médicas, contar con un mínimo de cinco profesores asistentes, de asignatura o colegiados;

II. En el caso de los programas de maestrías en ciencias o maestría en ingeniería con orientación científica, contar con una planta docente mínima de ocho profesores colegiados, con al menos cinco doctores;

III. En el caso de los programas de maestrías en ingeniería o en administración, contar con una planta docente mínima de seis profesores colegiados;

IV. En el caso de los programas de doctorado en ciencias y doctorado en ingeniería, contar con una planta docente de al menos nueve doctores;

V. Tratándose de programas de maestría y doctorado en ciencias o maestría y doctorado en ingeniería integrados, contar con una planta docente mínima de doce profesores colegiados, con al menos nueve doctores;

VI. En todos los casos, mantener vigente una matrícula de por lo menos un alumno por cada profesor de tiempo completo, y

VII. Ser evaluados en plazos no mayores a tres años, con base en los lineamientos y los indicadores de desempeño que para el efecto establezca la Secretaría.

El incumplimiento de alguno de los requisitos señalados en el presente artículo implicará la intervención de la

Secretaría y del colegio académico, para analizar la viabilidad del programa correspondiente y para definir las medidas que conducirán a su recuperación.

La Secretaría realizará un seguimiento del programa y si no se aprecia una mejoría en sus indicadores, podrá dictaminar que éste entre en receso de un semestre, durante el cual no podrán admitirse alumnos de nuevo ingreso. La propuesta de receso del programa deberá contar con el dictamen favorable del colegio académico.

El programa respectivo se cancelará si del análisis que se realice al término del periodo de receso, según dictamen del colegio académico, la Secretaría advierte que no se cumplen las condiciones para mantenerse vigente. Dicho dictamen será sometido a consideración del Consejo General Consultivo del Instituto, a través de la Secretaría, para que determine lo conducente.

Artículo 27. Las actividades de actualización podrán realizarse previa aprobación del colegio de profesores correspondiente y de la Secretaría y la vigencia de las asignaturas respectivas será de un año, pudiendo renovarse previo acuerdo de dicho colegio.

Capítulo Tercero DEL INGRESO Y LA REVALIDACIÓN

Artículo 28. Para ser admitido como alumno en los programas de especialidad, especialidad médica y maestría, el aspirante deberá cumplir con los siguientes requisitos:

ACUERDO

- I. Poseer título profesional o certificado oficial de terminación de estudios;
- II. Aprobar el proceso de admisión diseñado para tal efecto;
- III. Acreditar el examen de comprensión de lectura y traducción del idioma inglés o de aquel que el colegio de profesores considere adecuado, ya sea a través del Centro de Lenguas Extranjeras del Instituto o el equivalente en otro tipo de examen reconocido nacional o internacionalmente y aprobado por el colegio académico;
- IV. No haber causado baja en algún posgrado del Instituto, salvo que le haya sido revocada por el colegio académico, y
- V. Cubrir los derechos y cuotas correspondientes.

Sólo en caso justificado los coordinadores de los programas de especialidad podrán solicitar a la Secretaría que se dispense el requisito señalado en la fracción III anterior.

Artículo 29. Para ser admitido como alumno en los programas de doctorado en ciencias o doctorado en ingeniería, el aspirante deberá satisfacer los siguientes requisitos:

- I. Poseer título de licenciatura;
- II. Poseer el grado de maestro;
- III. Aprobar el proceso de admisión para tal efecto;
- IV. Obtener un promedio global aprobatorio en el examen de

dominio de por lo menos dos habilidades del idioma inglés o de aquel que el colegio de profesores considere adecuado, ya sea a través del Centro de Lenguas Extranjeras del Instituto o el equivalente en otro tipo de examen reconocido nacional o internacionalmente y aprobado por el colegio académico;

V. No haber causado baja en algún posgrado del Instituto, salvo que le haya sido revocada por el colegio académico, y

VI. Cubrir los derechos y cuotas

Artículo 30. Para ser admitido al doctorado en ciencias o doctorado en ingeniería sin haber realizado estudios de maestría, además de ser propuesto por el colegio de profesores correspondiente, se deberá cumplir con los siguientes requisitos:

- I. Del aspirante:
 - A. Poseer un promedio mínimo de ocho en escala de cero a diez, en los estudios de licenciatura, o su equivalente en otras escalas, y
 - B. Cumplir con los requisitos previstos por el artículo 29 del presente Reglamento, con excepción de lo dispuesto en la fracción II del mismo.
- II. Del programa:
 - A. Que cuente con registro en el Consejo Nacional de Ciencia y Tecnología o que cumpla con las fosas de graduación requeridas por dicho organismo y contar con un cuerpo académico calificado reconocido por e) Sistema Nacional de Investigadores, y

ACUERDO

B. Que ofrezca líneas de investigación de actualidad
En todos los casos los directores de tesis deberán ser profesores colegiados y pertenecer al programa de que se trate y acreditar que cuentan con fondos suficientes para sostener la investigación del alumno.

El tema de tesis del alumno deberá ser de tal calidad que permita la publicación de los resultados en una revista de la especialidad de circulación internacional, con registro en el ISI *Data Base*, o su equivalente, avalado por la Secretaría, a propuesta del colegio correspondiente.

Artículo 31. Un estudiante de maestría podrá solicitar su migración a un programa de doctorado, siempre que se cumplan los siguientes requisitos:

I. Contar con el dictamen favorable y autorización del cuerpo académico y del co85legio de profesores correspondiente del programa donde se encuentra inscrito;

II. Haber concluido su programa de asignaturas de maestría en un plazo máximo de tres semestres;

III. Poseer un promedio de nueve o superior en su programa de maestría, y

IV. No haber causado baja en algún posgrado del Instituto, salvo que le haya sido revocada por el colegio académico.

Artículo 32. El proceso de admisión para ingresar a un programa de posgrado deberá garantizar que el aspirante cuenta con los atributos necesarios para desarrollar

adecuadamente el perfil de egreso definido por el programa y comprenderá, cuando menos, lo siguiente:

- I. Exámenes escritos o asignaturas propedéuticas;
- II. Entrevistas colegiadas;
- III. Análisis del currículo del aspirante, y
- IV. Comprobación documental de los requisitos.

El proceso se iniciará a partir de la emisión de una convocatoria general aprobada por el colegio correspondiente, en la que se establecerán los requisitos particulares de cada programa.

Los cuerpos académicos que corresponda establecerán la matrícula máxima que se podrá admitir en cada programa, tomando en consideración el número de estudiantes que cada profesor está en posibilidad de atender, cuidando la proporción entre dicha matrícula y el número de profesores de tiempo completo, de manera que se garantice una eficiencia terminal satisfactoria.

El proceso de admisión será responsabilidad de la comisión de admisión nombrada para cada programa de posgrado y estará coordinado por el director de la unidad académica o el jefe de sección correspondiente, según sea el caso, quienes definirán los lineamientos generales de admisión.

Artículo 33. La comisión de admisión de un programa podrá exigir al interesado la aprobación de asignaturas previas para su ingreso, sin que implique ser considerado como alumno del posgrado.

ACUERDO

El colegio de profesores deberá solicitar el dictamen académico favorable de la Secretaria cuando el aspirante a una especialidad, maestría o doctorado, posea un título otorgado por la institución de enseñanza superior ajena al instituto, antes de su ingreso.

Los documentos académicos expedidos e el extranjero deberán ser validados de conformidad a los acuerdos vigentes.

Artículo 34. Cuando la comisión de admisión de un programa dictamine favorable la admisión de un aspirante, procederá a elaborar el acta de admisión y el programa de estudios que el aspirante tendrá que cumplir para obtener su grado.

Dicho programa de estudios deberá cumplir con los lineamientos y requisitos establecidos en el plan de estudios del programa aprobado e incluir las actividades académicas a realizar por el alumno. En el programa de estudios se especificará el consejero al alumno.

Artículo 35. La revalidación de los estudios de posgrado realizados en el instituto o de otros, se llevará a cabo sobre las asignaturas acreditadas y no sobre los programas como tal.

En la revalidación de las asignaturas de posgrado se considerará el nivel de equivalencia bajo los siguientes criterios:

- I. El tiempo transcurrido entre el momento en que se cursó la asignatura y aquel en que se solicita la revalidación;

- II. El nivel de la asignatura y el currículo del profesor que lo impartió;
- III. La equivalencia en el número de horas de las asignaturas cursadas;
- IV. El contenido y nivel de profundidad de la asignatura cursada respecto de las asignaturas vigentes en el plan de estudios del programa, y
- V. Otras que el programa o el colegio de profesores considere convenientes por la orientación del programa.

El colegio de profesores de la unidad académica que corresponda dictaminará las asignaturas revalidadas, previa propuesta de la comisión de admisión y del cuerpo académico.

El máximo de créditos revalidables en cada programa será decidido por el colegio de profesores correspondiente.

En cada caso se levantará un acta que indicará cuáles asignaturas del plan de estudios se revalidaron y los antecedentes en que se fundamentó dicho dictamen. El análisis deberá realizarse sobre documentos originales y el acta de revalidación deberá acompañarse de una copia cotejada de dichos documentos.

Artículo 36. Para ser admitido como alumno de actividades de actualización, el aspirante deberá satisfacer los siguientes requisitos:

- I. Poseer título profesional o carta oficial de terminación de estudios:

ACUERDO

II. Aprobar el procedimiento de admisión diseñado para tal efecto,

III. Cubrir los derechos y cuotas correspondientes.

Capítulo Cuarto DE LA ACTIVIDAD ACADÉMICA

Artículo 37. Las actividades de desarrollo de la tesina para alumnos de especialidad y de la tesis para alumnos de especialidades médicas, maestría y doctorado son obligatorias, salvo en los casos en que les especifique así el presente Reglamento.

El colegio de profesores respectivo designará, a propuesta del coordinador del programa correspondiente, a un consejero de estudios que asumirá la función de director de tesis, cuando ésta sea registrada.

En el caso de alumnos de tiempo completo, se designará al director de tesis o de tesina, según corresponda, a más tardar al término del primer semestre y en el caso de alumnos de tiempo parcial, al término del segundo semestre a partir de su inscripción como alumno. Los directores de tesis asumirán, a partir de su designación, el papel de consejeros de estudios.

La elaboración de tesis será considerada como una asignatura y no tendrá valor en créditos.

Artículo 38. El colegio de profesores designará, en la modalidad que estime conveniente y para cada alumno en programas de maestría o de doctorado, así como para cada alumno en programas de especialidad, un comité tutorial que se integrará

cuando se designe director de tesis o tesina para el estudiante, con la finalidad de dar seguimiento al cumplimiento de su programa individual de actividades y al desarrollo de su tesis o tesina.

Dicho comité estará conformado, en lo posible, por cuatro profesores con derecho a voz y voto y con al menos tres profesores colegiados, entre los cuales uno de los directores de tesis o tesina formará parte del comité. El cuarto miembro del comité tutorial será preferentemente externo al programa.

El comité tutorial será propuesto por el coordinador del programa respectivo y aprobado por el cuerpo académico del programa y entrará en funciones a más tardar al inicio del segundo semestre después de haber ingresado el alumno y terminará éstas al designarse el jurado para el examen de grado.

Los miembros del comité tutorial, los directores de tesis o tesinas y los consejeros de estudio podrán cambiar durante el desarrollo del trabajo del estudiante, previo análisis, justificación y aprobación del cuerpo académico y conocimiento del colegio de profesores de la unidad académica de que se trate.

Artículo 39. Los comités tutorales tendrán a su cargo las siguientes funciones:

- I. Determinar el número de sesiones en que el alumno deberá presentar sus avances, considerando un mínimo de una sesión por semestre, donde una de ellas podrá ser en el seminario del programa;



INSTITUTO POLITÉCNICO NACIONAL
ESCUELA SUPERIOR DE COMERCIO Y ADMINISTRACIÓN
UNIDAD SANTO TOMÁS
MAESTRÍA EN CIENCIAS EN ADMINISTRACIÓN DE NEGOCIOS



HOJA DE RESPUESTAS

INSTRUCCIONES: Marque en cada número su opción de respuesta, rellenando completamente el círculo.

<i>Teoría de la Administración</i>	<i>Análisis Económico Financiero</i>	<i>Matemáticas Aplicadas a la Administración</i>	<i>Informática Administrativa</i>	<i>Investigación</i>
1 (A) (B) (C) (D)	21 (A) (B) (C) (D)	41 (A) (B) (C) (D)	61 (A) (B) (C) (D)	81 (A) (B) (C) (D)
2 (A) (B) (C) (D)	22 (A) (B) (C) (D)	42 (A) (B) (C) (D)	62 (A) (B) (C) (D)	82 (A) (B) (C) (D)
3 (A) (B) (C) (D)	23 (A) (B) (C) (D)	43 (A) (B) (C) (D)	63 (A) (B) (C) (D)	83 (A) (B) (C) (D)
4 (A) (B) (C) (D)	24 (A) (B) (C) (D)	44 (A) (B) (C) (D)	64 (A) (B) (C) (D)	84 (A) (B) (C) (D)
5 (A) (B) (C) (D)	25 (A) (B) (C) (D)	45 (A) (B) (C) (D)	65 (A) (B) (C) (D)	85 (A) (B) (C) (D)
6 (A) (B) (C) (D)	26 (A) (B) (C) (D)	46 (A) (B) (C) (D)	66 (A) (B) (C) (D)	86 (A) (B) (C) (D)
7 (A) (B) (C) (D)	27 (A) (B) (C) (D)	47 (A) (B) (C) (D)	67 (A) (B) (C) (D)	87 (A) (B) (C) (D)
8 (A) (B) (C) (D)	28 (A) (B) (C) (D)	48 (A) (B) (C) (D)	68 (A) (B) (C) (D)	88 (A) (B) (C) (D)
9 (A) (B) (C) (D)	29 (A) (B) (C) (D)	49 (A) (B) (C) (D)	69 (A) (B) (C) (D)	89 (A) (B) (C) (D)
10 (A) (B) (C) (D)	30 (A) (B) (C) (D)	50 (A) (B) (C) (D)	70 (A) (B) (C) (D)	90 (A) (B) (C) (D)
11 (A) (B) (C) (D)	31 (A) (B) (C) (D)	51 (A) (B) (C) (D)	71 (A) (B) (C) (D)	91 (A) (B) (C) (D)
12 (A) (B) (C) (D)	32 (A) (B) (C) (D)	52 (A) (B) (C) (D)	72 (A) (B) (C) (D)	92 (A) (B) (C) (D)
13 (A) (B) (C) (D)	33 (A) (B) (C) (D)	53 (A) (B) (C) (D)	73 (A) (B) (C) (D)	93 (A) (B) (C) (D)
14 (A) (B) (C) (D)	34 (A) (B) (C) (D)	54 (A) (B) (C) (D)	74 (A) (B) (C) (D)	94 (A) (B) (C) (D)
15 (A) (B) (C) (D)	35 (A) (B) (C) (D)	55 (A) (B) (C) (D)	75 (A) (B) (C) (D)	95 (A) (B) (C) (D)
16 (A) (B) (C) (D)	36 (A) (B) (C) (D)	56 (A) (B) (C) (D)	76 (A) (B) (C) (D)	96 (A) (B) (C) (D)
17 (A) (B) (C) (D)	37 (A) (B) (C) (D)	57 (A) (B) (C) (D)	77 (A) (B) (C) (D)	97 (A) (B) (C) (D)
18 (A) (B) (C) (D)	38 (A) (B) (C) (D)	58 (A) (B) (C) (D)	78 (A) (B) (C) (D)	98 (A) (B) (C) (D)
19 (A) (B) (C) (D)	39 (A) (B) (C) (D)	59 (A) (B) (C) (D)	79 (A) (B) (C) (D)	99 (A) (B) (C) (D)
20 (A) (B) (C) (D)	40 (A) (B) (C) (D)	60 (A) (B) (C) (D)	80 (A) (B) (C) (D)	100 (A) (B) (C) (D)

Anexo 4

FOLIO	ANÁLISIS ECONÓMICO ADMINISTRATIVO	TEORÍA DE LA ADMINISTRACIÓN	INFORMÁTICA ADMINISTRATIVA	HABILIDADES DE INVESTIGACIÓN	MATEMÁTICAS	PROMEDIO
130	94	96	100	44	85	83,80
55	61	78	93	67	85	76,80
127	42	74	86	100	62,5	72,90
154	58	67	79	89	60	70,60
65	30	74	93	67	88	70,40
85	52	74	93	67	65	70,20
27	52	56	71	78	93	70,00
9	39	89	71	78	68	69,00
36	55	74	79	78	53	67,80
148	33	59	93	78	75	67,60
81	48	67	93	67	62,5	67,50
141	48	67	93	56	72,5	67,30
41	42	41	86	89	78	67,20
54	33	52	79	89	80	66,60
103	52	78	79	56	67,5	66,50
153	52	44	86	78	70	66,00
47	48	63	86	78	53	65,60
116	42	52	79	78	75	65,20
124	48	59	86	78	55	65,20
38	55	67	86	67	50	65,00
89	58	74	79	44	70	65,00
107	45	52	79	89	57,5	64,50
22	42	33	93	89	65	64,40
28	61	67	79	67	44	63,60
19	67	59	86	67	38	63,40
73	48	81	50	78	58	63,00
112	39	70	71	78	55	62,60
147	55	67	71	67	50	62,00
7	36	67	71	67	68	61,80
37	30	67	93	56	63	61,80
11	27	67	79	67	68	61,60
33	42	59	79	78	50	61,60
53	52	44	86	56	70	61,60
115	39	44	79	78	67,5	61,50
150	33	52	79	78	65	61,40
118	27	48	86	78	67,5	61,30
30	55	41	93	67	50	61,20
97	45	33	79	89	60	61,20
86	30	63	86	56	70	61,00
82	52	52	79	67	52,5	60,50
111	55	44	93	78	32,5	60,50
122	42	74	79	67	40	60,40
3	55	56	86	56	48	60,20
117	39	26	93	78	65	60,20
57	30	70	79	78	43	60,00
68	48	37	79	78	58	60,00
77	21	44	86	89	60	60,00
125	48	48	86	67	50	59,80

32	33	52	79	100	33	59,40
66	39	52	86	67	53	59,40
78	30	48	79	67	73	59,40
40	42	67	71	56	60	59,20
17	45	74	86	56	33	58,80
152	30	44	86	89	45	58,80
135	27	56	93	67	50	58,60
10	36	52	71	67	65	58,20
99	39	59	86	56	50	58,00
110	30	44	93	67	55	57,80
145	36	56	86	56	55	57,80
72	36	44	79	56	73	57,60
1	36	67	79	67	38	57,40
31	5	56	79	67	80	57,40
20	48	26	100	56	50	56,00
45	30	44	79	67	60	56,00
83	48	37	86	44	65	56,00
46	39	56	93	56	35	55,80
51	48	41	93	56	40	55,60
109	36	41	64	67	70	55,60
13	36	44	79	56	60	55,00
67	36	26	86	67	60	55,00
21	15	81	57	78	43	54,80
48	27	48	86	56	55	54,40
58	45	37	64	78	48	54,40
5	33	63	71	44	60	54,20
64	27	67	86	56	35	54,20
16	36	44	64	56	70	54,00
49	36	33	64	67	70	54,00
100	21	59	86	56	47,5	53,90
126	30	41	79	67	52,5	53,90
23	33	59	86	56	35	53,80
24	24	56	86	33	70	53,80
105	39	63	79	33	55	53,80
15	39	41	79	44	65	53,60
14	39	41	50	67	70	53,40
155	36	52	71	78	30	53,40
39	36	67	79	22	60	52,80
4	33	33	50	78	68	52,40
120	36	44	99	33	50	52,40
142	42	37	64	78	40	52,20
35	33	41	79	67	40	52,00
42	48	48	64	67	33	52,00
79	36	33	86	67	38	52,00
93	30	48	79	67	35	51,80
106	36	33	86	67	35	51,40
146	36	48	71	67	35	51,40
62	27	52	86	67	23	51,00
98	27	37	79	67	45	51,00
80	36	33	71	67	47,5	50,90
101	27	67	86	33	40	50,60
104	30	48	79	56	40	50,60

113	36	44	86	44	37,5	49,50
6	33	52	79	33	48	49,00
71	39	41	71	33	50	48,50
29	36	48	79	44	35	48,40
138	30	44	79	56	32,5	48,30
76	18	48	86	33	55	48,00
91	24	44	71	67	32,5	47,70
151	36	33	64	78	27,5	47,70
44	24	44	71	56	40	47,00
61	27	37	64	56	50	46,80
12	27	26	79	33	68	46,60
59	36	37	57	67	35	46,40
43	45	63	64	56	0	45,60
25	42	30	33	56	58	43,80
114	24	30	50	78	35	43,40
129	45	30	71	67	2,5	43,10
52	18	22	64	67	43	42,80
136	24	41	79	44	25	42,60
90	36	37	57	44	37,5	42,30
119	45	33	64	33	35	42,00
121	18	30	64	56	40	41,60
140	33	33	50	44	35	39,00

SUMA 4679 6257 9570 7813 6454 6954,60
PROMEDIO 38,35245902 51,28688525 78,44262295 64,0409836 52,9016393 57,00