

UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO

FACULTAD DE PSICOLOGIA



**ELABORACION DE UNA PRUEBA DE CONOCIMIENTO DEL
LENGUAJE PARA SER CONSIDERADA COMO FACTOR QUE
INFLUYE EN LA COMPRESION DE LA LECTURA.**

T E S I S
Que para Obtener el Título de :
LICENCIADO EN PSICOLOGIA
P R E S E N T A :
ELKE KORT DE CAPELLA

MEXICO, D. F.

1977.



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

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

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

Reconocimientos:

Al Dr. Serafín Mercado D., por el
asesoramiento general.

A la Lic. A.M. Cardero García por
asesorar el aspecto lingüístico.

A la Coordinación de Apoyos
Académicos de la Unidad Xochimilco
de la UAM por las facilidades pres
tadas.

I N D I C E

	Pág
INTRODUCCION	1
CAPITULO I	
Descripción de la prueba de conocimiento del lenguaje	4
Criterios para estudiar el lenguaje	7
Revisión de investigaciones que relacionan el conocimiento del lenguaje con la comprensión de la lectura	12
CAPITULO II	
Bases psicométricas para la elaboración de una prueba de rendimiento	25
Clasificación de las pruebas psicológicas	27
Desarrollo histórico de los tests o pruebas de rendimiento	30
Criterios para la construcción y valoración de los instrumentos de medición	32
1. Estandarización	32
2. Objetividad	34
3. Análisis de los reactivos	35
4. Confiabilidad	39
5. Validez	45
6. Normas y puntajes derivados	51

CAPITULO III

Método	53
Sujetos	53
Procedimiento	55
Piloteos de la prueba	57
Análisis de los datos y resultados	66
Validez de contenido	80
RESUMEN	82
CONCLUSIONES	83
SUGERENCIAS	85
APENDICE	88
BIBLIOGRAFIA	97

INTRODUCCION

Este trabajo consiste en la elaboración de una Prueba de Cono cimiento del Lenguaje para ser aplicada a estudiantes univer sita rios. La prueba está planeada para formar parte de un instrumento de medida de la lectura que evaluará la velocidad y la comprensión de la lectura de estudiantes universitarios (Serafín J. Mercado D. y otros, 1976).

Refiriéndose a investigaciones hechas en lengua inglesa, se ha estudiado como una variable importante en la comprensión de la lectura el manejo del lenguaje, es decir la gramática (consistiendo en sintaxis y morfología), la semántica, la fonología y la normativa. Se pretende que esta prueba servirá posteriormente para comprobar hipótesis al respecto.

En la actualidad no existe información acerca de las capacidades reales de lectura de los estudiantes universitarios mexicanos. La deseabilidad y necesidad de incrementar las habilidades tanto de velocidad como de comprensión de la lectura de los estudiantes universitarios en general parecen obvias. Sin embargo, no existen datos empíricos que podrían guiar la planeación de curricula. Conociendo el rendimiento de los estudiantes al inicio de una carrera universitaria se podría planear material realista por un lado (Serafín J. Mercado y otros, 1976) y por el otro, proponer programas de entrenamiento para que los estudiantes alcancen un nivel mínimo en la comprensión de la lectura que les permita seguir sus estudios con éxito.

Tomando en cuenta las áreas del estudio del lenguaje, el instrumento para evaluar el conocimiento del lenguaje fue construido seleccionando los reactivos de la prueba bajo los criterios de gramática, semántica, fonología y normativa.

Como los estudiantes, a los que se aplicará la prueba, proceden de diferentes sistemas de educación, se tratará de examinar conocimientos realmente básicos que se adquieren en la enseñanza primaria y media.

El material para formular los reactivos fue escogido de libros de textos de lenguaje actuales.

La población que se escogió para los piloteos de la prueba fue de diversas facultades de la UNAM. El último piloteo fue realizado durante los meses de agosto y septiembre de este año - con alumnos de las facultades de Filosofía y Letras, Ciencias Políticas, Psicología, Química, Ingeniería y Arquitectura todos cursando el segundo semestre de su carrera.

La confiabilidad de la prueba fue estimada tanto con el método de Spearman-Brown como el método de Kuder-Richardson. Se obtuvo el nivel de dificultad de cada reactivo. Los datos fueron procesados mediante un programa escrito en Fortran en la computadora B 6700 del Centro de Cálculo de la UNAM.

CAPITULO I

DESCRIPCION DE LA PRUEBA DE CONOCIMIENTO DEL LENGUAJE

La prueba fué diseñada para ser aplicada a estudiantes universitarios. Se pretende medir conocimientos básicos que se deben haber adquirido en los años de enseñanza primaria y secundaria.

El término conocimiento se define igual que Bloom, 1974 y se cita textualmente:

"El conocimiento, tal como se lo define aquí, incluye aquellos comportamientos y situaciones de examen que acentúan la importancia del recuerdo de ideas, materiales o fenómenos, ya sea como reconocimiento o evocación. La conducta que se espera de un estudiante en situación de evocación es similar a la que se esperó de él durante el aprendizaje original, cuando se confía en que almacene en su mente el aprendizaje original, cuando se confía en que almacene determinada información. Posteriormente se le pide que recuerde esa misma información. Aun cuando pueden darse ligeras modificaciones del material que debe ser recordado, este hecho desempeñará un papel de menor importancia en el comportamiento o el examen de conocimientos.

El orden de los objetivos de conocimiento abarca desde las conductas más específicas y relativamente concretas a las más complejas y abstractas. Así, el conocimiento de hechos específicos se refiere a los tipos de información que pueden ser aislados y recordados separadamente, mientras que el conocimiento de universales y abstracciones subraya las interrelaciones y los esquemas en que puede organizarse y estructurarse la información".

A continuación se presenta el esquema de la prueba que señala el número de reactivos de cada área del estudio del lenguaje bajo los criterios mencionados en el Capítulo I.

Criterio de Morfosintaxis:

Reactivos 1 al 10, 17, 18, 25, 26, 28.

Criterio semántico:

Reactivos 11-16, 19 al 24, 27.

Criterio fonológico:

Reactivos 29 al 33.

Criterio de normativa:

Reactivos 34 al 38.

El material para formular los reactivos fué seleccionado de libros de texto actuales para la enseñanza del lenguaje en las escuelas primarias y secundarias.

Los diversos tipos de respuestas que se le piden a los estudiantes comprenden reactivos de opción múltiple y de tipo suministro que se limitan a respuestas breves, de unas pocas palabras, de subrayar y colocar signos de puntuación y acentuación.

La prueba será aplicada (prueba de aplicación colectiva) en grupo y se leen a los sujetos las instrucciones que están impresas en la prueba.

La prueba es de rendimiento. Como prueba de potencia o ejecución máxima no se fija un tiempo límite para las respuestas. El registro del tiempo para que todos los estudiantes puedan intentar contestar todas las preguntas, se llevó a cabo durante el piloteo e indica que 40 minutos son el límite superior de tiempo.

En el Apéndice se incluye la prueba y la clave de respuestas.

CRITERIOS PARA ESTUDIAR EL LENGUAJE

La prueba de conocimiento del lenguaje se elaboró tomando los siguientes criterios como base para el estudio del lenguaje (Lingüística 1, 1975):

ESTUDIO DEL LENGUAJE	
<u>Ciencias del signo lingüístico</u>	
MORFOSINTAXIS	{ Sintaxis → estudia las funciones de las palabras Morfología → estudia las formas de las palabras, es decir, su estructura interna
SEMANTICA	Estudia los significados de las palabras
FONOLOGIA	Estudia los fonemas
FONETICA	Estudia los sonidos
<u>Normativa</u>	
GRAMATICA	Enseña los usos correctos relacionados con la Gramática y la Semántica
ORTOLOGIA	Enseña la correcta pronunciación de los sonidos
ORTOGRAFIA	Enseña la correcta escritura de los grafemas.

Las unidades del nivel superior o primera articulación, los morfemas y signos lingüísticos complejos, se estudian en la GRAMÁTICA o MORFOSINTAXIS y en la SEMÁNTICA. En la Gramática se estudian las relaciones de los signos lingüísticos, y en la Semántica, sus significados.

Los signos lingüísticos, al encadenarse unos detrás de otros, forman signos lingüísticos cada vez más complejos, por ejemplo:

León	—————→	es un signo lingüístico
El león	—————→	es un signo lingüístico más complejo
El león corre	—————→	es un signo lingüístico más complejo aún
El león corre mucho	—————→	es más complejo todavía, etc.

Una parte de la gramática es la SINTAXIS, que estudia la función de las palabras; otra parte es la MORFOLOGÍA, que estudia la estructura interna de las palabras.

Las unidades del nivel inferior o segunda articulación, los fonemas, se estudian en la FONOLOGÍA. La realización de los fonemas, es decir, los sonidos, se estudian en la FONÉTICA.

El uso correcto del idioma nos lo enseña la **NORMATIVA**. Una parte de la normativa es la Gramática de los Usos, que se ocupa de los usos correctos relacionados con la Gramática o Morfosintaxis y la Semántica.

Otra parte de la normativa es la Ortografía, que se ocupa del uso correcto de las letras. Y la Ortología es otra parte de la normativa que se ocupa de la correcta pronunciación de los sonidos.

A continuación se presentan ejemplos de los reactivos de la Prueba de Conocimiento del Lenguaje representativo de cada uno de los criterios para estudiar el lenguaje:

MORFOSINTAXIS

a) La Sintaxis

En la siguiente expresión: No abras la puerta a nadie.
La parte "a nadie" funciona
como complemento indirecto.

Estudiar las funciones y relaciones de las palabras es estudiar la lengua con criterio sintáctico.

b) La Morfología

En las siguientes expresiones: cualquiera - cualesquiera
album - álbumes

Las palabras cualesquiera
y álbumes presentan el plu
ral de las palabras album
y cualquiera.

Estudiar la forma de las palabras y los cambios que pueden tener es estudiar la lengua con criterio morfológico.

SEMANTICA

Cuando se trata de descubrir el significado de las palabras y la abundancia de su uso (Polisemia y Homonimia) se está estudiando la lengua con criterio semántico.

FONOLOGIA

El siguiente ejemplo muestra el criterio fonológico.

Roma, la Ciudad Eterna, atrae los ojos de todo el mundo por su grandeza histórica y artística. Aquí se pide que se coloo

quen las comas en los debidos lugares. La colocación de los signos de puntuación hace cambiar el tono de voz en diferentes lugares de la oración. Las distintas entonaciones pueden llegar a producir cambios de significado e intención.

NORMATIVA

La normativa, que se llama también "gramática de los usos", dice cuando un uso es "gramaticalmente" correcto y cuando es incorrecto. En el siguiente ejemplo se pide que se seleccione de las opciones dadas, la correcta:

Debemos _____ las verduras antes de servir las.

Las opciones son: cocer - coser - cueser - cuecer.

REVISION DE INVESTIGACIONES QUE RELACIONAN EL CONOCIMIENTO
DEL LENGUAJE CON LA COMPRESION DE LA LECTURA.

Haciendo una revisión de la literatura de las investigaciones en lengua inglesa acerca de la importancia de la variable "conocimiento del lenguaje" David H. Russel y Henry R. Fea citan en el Handbook of Research on Teaching (1963)

(22) los siguientes investigadores e investigaciones:*

Según Cole (1946, p. 40) "con toda probabilidad, un vocabulario inadecuado es la mayor causa aislada para el fracaso de leer con comprensión, tanto en los campos de conocimientos generales como en los técnicos".

Se encontró que la mayoría de los estudios realizados han si do llevados a cabo con poblaciones de estudiantes del nivel de primaria. Estos estudios están relacionados sobre todo con los métodos de instrucción en la enseñanza de la lectura.

J.A. Holmes (1953) sugirió "que en el caso de niños que se inician en la lectura de la discriminación de las palabras, el sentido de las mismas, y el tiempo de reconocimiento con tribuyen en un 56% a la varianza en la velocidad y el poder de la lectura, mientras factores como el conocimiento del vocabulario dentro de un contexto, inteligencia y la percepu

* Ver Notas de Referencias (Pág. 101 y 102)

ción de relaciones verbales representan el 77% de la varianza en la capacidad de la lectura."

Revisando la literatura sobre la afirmación citada de Cole (1946), se observan los siguientes resultados en los estudios acerca de cómo el incremento en el vocabulario se relaciona con la mayor comprensión de la lectura:

Dunkel (1944), concluyó "que la habilidad de determinar el significado preciso de las palabras estaba relacionado con la habilidad de leer con comprensión." En investigaciones hechas por Hunt (1953), Reed y Pepper (1957), se reportaron aproximadamente el mismo resultado. En cuanto al problema de cómo incrementar el vocabulario no se presenta evidencia experimental. Únicamente se señala "que el dominio del vocabulario y la comprensión están estrechamente relacionados."/

Investigaciones directas sobre el valor de la lectura para aumentar el vocabulario han dado resultados negativos."

Traxler (1938), encontró "que la lectura en sí no parecía formar el vocabulario en alumnos de enseñanza media."

Sachs (1943), encontró "lo mismo con alumnos de primer ingreso en la universidad." McCullough (1957 b, p. 477) hizo én-

fasis "en la importancia del estudio del vocabulario, sobre todo en la enseñanza media (high school)". Ella propone la enseñanza directa de vocabulario, definiéndola como sigue:

"El enfoque directo significa estudiar listas de palabras u oraciones precisamente para el desarrollo del dominio de las palabras...". Cole, McCullough y D.H. Russel (1961) se refieren "a investigaciones y práctica apoyando la instrucción directa en vocabulario" y Alm (1957), advierte "que hay poco valor en el estudio de palabras en cuanto a que se están escribiendo definiciones de palabras de listas de palabras dadas fuera de cualquier contexto.

Hay que escoger palabras que el alumno pueda utilizar en un contexto."

Ha habido bastantes intentos para incrementar el vocabulario enseñando la historia de las palabras. Moir (1953), elaboró "una unidad sobre la historia del idioma inglés para el octavo grado." Funk y Levis (1954), y Jespersen (1955), compilaron material similar. Serre (1953), revisando los resultados de las investigaciones concluyó "que el método en el cual el maestro proporciona una definición, explica y da ejemplos es superior a un método en el cual el alumno usa el diccionario sólo."

McCullough (1958), haciendo una revisión de las investigaciones sobre el aprendizaje de la palabra en un contexto concluyó, "que los alumnos deben comprender que los indicios de un contexto se extienden más allá de la oración en la cual aparece una palabra desconocida.

En algunos casos la naturaleza de todo un libro debe ser estudiado para comprender el significado específico de una palabra desconocida utilizandola en un contexto específico dentro de un texto.

Aparte del significado transmitido por la selección del vocabulario, el idioma transmite significado por medio de la sintaxis, la gramática y la puntuación."

Pei (1949), explicó "que en el idioma inglés la oración es la unidad del pensamiento, porque tanto el vocabulario, como la morfología y la sintaxis todos llevan el peso del significado."

Cook 1959, criticó "los métodos de enseñanza de la gramática, sintaxis y puntuación en las escuelas. El enfatizó el hecho de que se ha comprendido el papel que juega el lenguaje como instrumento de comprensión, sin embargo, la gramática es to-

davía enseñada como un conjunto de reglas en vez de un factor en la interpretación del significado."

Existe muy poca investigación sobre los efectos del conocimiento de la sintaxis y la gramática sobre las habilidades de lectura. Strom (1956), señaló "la frecuencia con la cual se hace notar la importancia del conocimiento de la gramática para la comprensión en el contexto. Ella analizó la sintaxis y la gramática de prosa literaria y poesías que estudiantes deberían leer como texto de comprensión. Ella encontró poca relación entre la comprensión de los estudiantes y su habilidad de clasificar los elementos del texto en cuanto a gramática y la sintaxis. Sus conclusiones implican que si la gramática y la sintaxis fueran enseñadas como formas auxiliares para obtener significado de las lecturas, entonces su experimento hubiera demostrado una relación significativa entre el conocimiento de la gramática y la comprensión."

Shores (1943), encontró "en alumnos de noveno grado que una buena capacidad de lectura en materias de ciencias se relacionaba con el nivel de comprensión de la lectura y con el conocimiento del vocabulario pero no con las habilidades de encontrar información o la comprensión del sentido muy general del pasaje."

John R. Bormuth (1969), revisa las investigaciones sobre la comprensión de la lectura y consta "que hay poca investigación y de poco valor sobre este tema." El autor señala "la importancia de dar un nuevo enfoque, una nueva definición de qué es la comprensión." /Bormuth propone "una definición operacional. Hasta entonces no se había tomado en cuenta las habilidades del lenguaje, que son la base de las habilidades de comprensión. Tanto investigadores como maestros saben que los alumnos que comprenden el lenguaje necesitan llevar a cabo determinadas tareas, como modificar sustantivos, verbos, etc., Sin embargo, las definiciones tradicionales de la comprensión nunca expresan esto." Bormuth propone la siguiente definición del término comprensión:

" La habilidad de comprensión es un conjunto de habilidades generalizado de adquisición de conocimientos que permite adquirir y exhibir información ganada como resultado de leer el lenguaje escrito".

"Esta definición afirma que la comprensión es una respuesta al sistema del lenguaje. Esto significa que el lenguaje es un sistema que puede tomar un infinito número de formas para señalar cualquier información semántica y el conocimiento de cómo funciona este sistema permite a las personas adquirir

cualquier información codificada en el lenguaje.

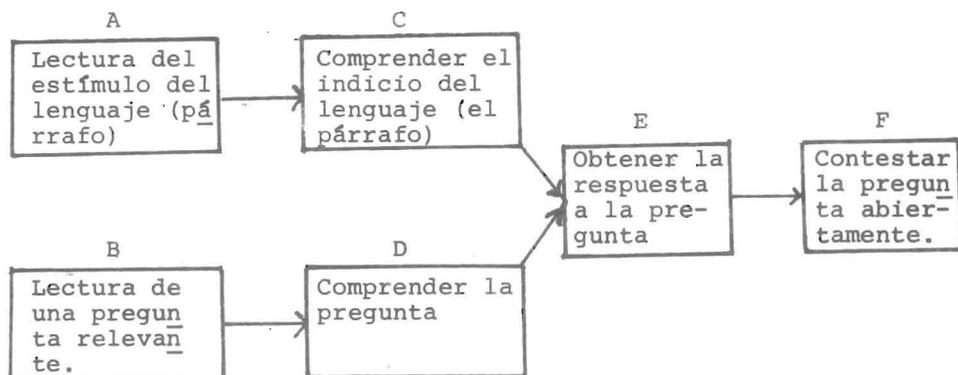
Así la comprensión no es únicamente un conjunto de procesos mentales que pueden ser definidos independientemente del lenguaje. Sino más bien es un conjunto de procesos que opera sobre aspectos específicos del lenguaje. La instrucción en todos los niveles de enseñanza, por lo tanto, debería ser diseñada sistemáticamente para que los alumnos aprendan aquellos sistemas a través de los cuales el lenguaje indica significado.

Algunos de estos sistemas de dar señales serían:

- Los significados semánticos de las palabras.
- La forma en que los prefijos influyen sobre el sentido semántico y las funciones sintácticas de las palabras, etc."

Bormuth (1969), presenta el siguiente esquema* para ilustrar lo anterior, indicando "como pueden surgir dos fuentes independientes de dificultad que influyen sobre una respuesta en cualquier tarea dada:"

* Journal of Educational Psychology, Vol. 61, N° 5, (pág. 350)



Análisis de los procesos principales que constituyen los componentes para responder correctamente a una unidad instruccional. (Bormuth, 1969).

Revisando la publicación de la International Reading Association, 1969, "Reading-What can be measured" que recopila información acerca de todas las publicaciones en el campo de la lectura se obtuvieron los siguientes datos acerca de las investigaciones llevadas a cabo:*

"En primer lugar habría que mencionar un aspecto general que se encuentra en todas las recopilaciones de investigaciones revisadas, es decir, el énfasis en las precauciones con las

* Ver Notas de Referencias (Pág. 103)

que se deben de tomar los datos encontrados por diferentes es
cuelas y diferentes investigadores.

Hay muchas áreas a investigar, algunas han recibido más aten
ción que otras, pero inclusive aquellas que han recibido bas
tante atención, dejan más preguntas abiertas que contestadas.

Todavía no existe un acuerdo sobre como definir la lectu
ra y menos la comprensión. En cuanto a la definición de la lectu
ra en términos psicolinguísticos se incluirían habilidades
como el uso de la sintaxis, conocimiento de palabras utili-
zadas en el contexto y conocimiento de como utilizar indicios
fonológicos (Goodman, 1969)."

Haciendo referencia a la construcción de tests para medir la
lectura, se comenta "que todavía no se han construido tests
que tomen en cuenta los elementos considerados centrales por
los psicolinguistas. Sin embargo, aunque los investigadores
estén muy lejos de estar de acuerdo si alguna habilidad espe
cífica o alguna combinación de habilidades es esencial para
la lectura, hay un acuerdo general de que existen ciertas
habilidades que están relacionadas con la lectura."

Bormuth (1965a, 1965b, 1966) utilizó el "cloze procedure" en
varios estudios para investigar algunos de los factores gra-

maticales que están relacionados con el grado de dificultad de un texto.

Más de un 90% de las pruebas que miden (en idioma inglés) la habilidad de la lectura incluyen una parte para medir el conocimiento de vocabulario. De hecho ha sido propuesto por varios investigadores (Karlin 1964, Wilson 1967) "que el puntaje en vocabulario permite a los maestros darse una idea de las habilidades en la lectura de sus estudiantes."

V.H. Hughes (1953) correlacionó "los puntajes de 332 alumnos de 5° grado (fifth grade) en pruebas de significado de palabras y comprensión con puntajes en pruebas de otros aspectos de habilidades de lenguaje como ortografía, puntuación, utilización del lenguaje, y sentido de las oraciones." Hughes encontró "que hay un alto grado de traslape entre todos los tests de habilidades del lenguaje."

Todavía hay mucha confusión acerca de como medir el vocabulario y las revisiones de las investigaciones hechas en este campo llevan a la conclusión que falta evidencia para afirmar que se ha encontrado una forma válida y confiable con la cual se pueda medir el conocimiento del vocabulario como sub-área de la lectura.



En vista de lo anterior, se propone enfocar las investigaciones hacia procedimientos para obtener más datos acerca de la habilidad de la lectura en los estudiantes.

Acerca de los métodos para evaluar el rendimiento en la lectura no hay duda de que los tests estandarizados, formales e informales proporcionan información, sin embargo, hay muy poco acuerdo entre los especialistas en problemas de lectura sobre la naturaleza de esta información y el valor que pueda tener para el maestro de lectura. Tests estandarizados obviamente indican cómo un estudiante ejecuta en relación a otro estudiante, en un momento dado, pero poco o casi nada dicen acerca del porqué de la ejecución. Sin embargo, estos tests permiten comparar confiablemente a los estudiantes en términos de un aprovechamiento general en la lectura. Aparte los tests pueden ser utilizados como un medio de determinar la necesidad y el tipo de asesoría indicado.

La comprensión del vocabulario en un área académica determinada ha sido propuesta como un buen índice de éxito en este campo de contenido. Mary E. Johnson (1952) construyó un test de vocabulario que consiste en 150 items de selección múltiple para alumnos de 5° grado en comprensión del vocabulario en 6 campos de contenido; aritmética, geografía, historia, ciencias, salud y literatura. Las palabras fueron to

mas de libros de texto de 5° grado que los estudiantes utilizaban en sus estudios diarios. Dependiendo de los puntajes obtenidos, se indicó a los estudiantes que necesitaban un programa para incrementar el vocabulario. Programas parecidos se reportaron para estudiantes de high school (Wyatt & Rideway, 1958, Dunlop, 1951).

En resumen, todavía las investigaciones no dan una respuesta acerca de algún método específico para evaluar el rendimiento en la lectura de los estudiantes. Pero ante la necesidad práctica de obtener datos sobre las habilidades, capacidades y necesidades de los estudiantes, tanto pruebas estandarizadas como pruebas informales pueden ayudar para agrupar estudiantes con fines de instrucción, para determinar niveles de lectura y diagnósticos del rendimiento en la lectura.

Se termina esta revisión bibliográfica con una cita de Frank Smith, 1975 que dice "que el lenguaje es vital. La naturaleza del lenguaje puede ser considerada como un reflejo de como desarrollan y organizan los seres humanos sus pensamientos, la manera como se aprende el lenguaje nos cuenta bastante acerca del aprendizaje en general."

Frank Smith (1975), al considerar el papel que juega la gra-

mática en la comprensión hace notar "que la gramática, o por lo menos la sintaxis, podría ser considerada como el puente entre la estructura de superficie y la estructura profunda, aunque la sintaxis no se presenta directamente en la estructura de superficie del idioma (estructuras de superficie son las características "físicas" del lenguaje y las estructuras profundas son el significado)." El autor sugiere "que, por un lado, en el proceso de la comprensión podemos comprender oraciones cuya gramática desconocemos, oraciones que nosotros mismos no podríamos producir. Por el otro lado, la gramática podría ser considerada el centro de la comprensión. Igualmente que el que habla o escribe utiliza la gramática para producir estructuras de superficie que representan su significado, también el que escucha o lee debe utilizar la gramática para comprender."

CAPITULO II

BASES PSICOMETRICAS PARA LA ELABORACION DE UNA PRUEBA DE RENDIMIENTO

INTRODUCCION

Definiciones de las pruebas o tests:

Una prueba puede definirse "como una situación estandarizada diseñada para tomar una muestra del comportamiento de un individuo y darle un puntaje." (Nunnally, 1964).

"Una prueba psicológica es una técnica sistemática que compara la conducta de dos o más personas." (Anastasi, 1973).

Kerlinger (1973), da la siguiente definición: "Una prueba es un procedimiento sistemático, durante el cual se presenta el individuo examinado ante un conjunto de estímulos construidos, a los que responde. Las respuestas le permiten al examinador asignar al examinado un puntaje, a partir del cual se pueden hacer inferencias si el examinado posee aquello que se supone que la prueba está midiendo."

Las pruebas psicológicas que miden habilidades, rendimiento y aptitudes se construyen con reactivos que permiten únicamente dos categorías de respuestas, correctas o incorrectas. A estos reactivos se les asigna comúnmente una calificación de 1 cuando la respuesta es correcta y una calificación de 0 cuando la respuesta es incorrecta. Estos reactivos son llamados reactivos dicotómicos. El puntaje que obtiene una persona en una prueba con un cierto número de reactivos es simplemente el número de estos contestados correctamente. (Kerlinger, 1973).

Se ha desarrollado un cuerpo sustancial de teoría y técnicas estadísticas que tratan de resolver los problemas de la construcción, la estandarización, la precisión de la medida y la validación de estas pruebas.

Ese cuerpo de teoría y técnicas se llama comúnmente teoría de los tests psicológicos o tests mentales. (Ferguson, 1971).

A continuación se presentarán algunos aspectos elementales de la teoría psicométrica en la que se basó la construcción de la Prueba de Conocimiento del Lenguaje, motivo de este trabajo.

CLASIFICACION DE LAS PRUEBAS PSICOLOGICAS

Desde el punto de vista de su objetivo, las pruebas pueden ser: (Morales, 1975).

de potencia o de ejecución máxima
de rasgos o peculiaridades (pruebas proyectivas)

Las pruebas de potencia o de ejecución máxima exigen al examinado su máximo rendimiento en la tarea que se le pide que ejecute, tanto en lo que se refiere a lo que "puede" hacer, como a lo que se "mide" en el momento del examen.

Estas pruebas incluyen las de inteligencia, de habilidades o aptitudes múltiples y específicas, de rendimiento (intelectual o de cualquier otro tipo).

Bajo este rubro se incluye la Prueba de conocimiento del Lenguaje elaborada.

Las pruebas de rasgos (pruebas proyectivas) incluyen las de ajuste, temperamento, intereses, valores, actitudes, preferencias, escalas de personalidad, inventarios de personalidad, técnicas proyectivas, etc.

La calificación que se lleva a cabo en pruebas de este tipo es vaga y poco confiable por su carácter subjetivo en la mayoría de los casos.

Por la forma o tipo de respuestas que exigen, las pruebas se clasifican en:

- pruebas objetivas
- pruebas subjetivas

Por la forma de administración, las pruebas se clasifican en:

- pruebas de aplicación individual
- pruebas de aplicación colectiva
- pruebas autoadministradas

En función de la libertad de ejecución las pruebas se pueden clasificar en:

- pruebas de poder
- pruebas de velocidad

Por la forma de dar las instrucciones, pruebas orales o pruebas escritas.

Por el material que se emplea las pruebas pueden ser

de lápiz-papel

verbal completamente

de ejecución

de aparatos especiales

de combinación de los tres primeros

DESARROLLO HISTORICO DE LOS TESTS O PRUEBAS DE RENDIMIENTO

En 1945, en las escuelas públicas de Boston surgió la preocupación por desarrollar exámenes escolares con ciertas mejoras técnicas. Se sustituyeron por exámenes escritos los exámenes orales realizados por inspectores.

Los primeros tests estandarizados para la medida de los resultados de la instrucción escolar comenzaron a aparecer a partir del comienzo del siglo. Encabezados por las investigaciones de E.L. Thorndike, estos tests utilizaban los principios de medición desarrollados en los laboratorios psicológicos. Como ejemplo se pueden citar las escalas para estimar la calidad de la caligrafía y las redacciones, así como pruebas de ortografía, cálculo y razonamiento aritmético. Más tarde aparecieron las baterías de rendimiento, con la publicación de la primera edición del Stanford Achievement Test en 1923. Estas pruebas proporcionaban medidas comparables de la actuación en diferentes asignaturas escolares, valoradas en relación con un grupo normativo único.

Al mismo tiempo, se demostraba cada vez más el desacuerdo existente entre los profesores con respecto a la calificación de los exámenes de tipo clásico. Hacia 1930 se admitía generalmente, que los exámenes corrientes no sólo supu

nían una mayor pérdida de tiempo, tanto para los alumnos como para los examinadores, sino que sus resultados eran menos seguros que los obtenidos con el nuevo tipo de elementos objetivos. A medida que estos últimos se fueron empleando cada vez más en las pruebas estandarizadas de rendimiento, se produjo una tendencia cada vez mayor a programar elementos para medir la comprensión y la aplicación de conocimientos y otros objetivos educativos amplios.

A medida que un número creciente de psicólogos con conocimientos de psicometría participaba en la elaboración de las pruebas de rendimiento, los aspectos técnicos de las pruebas de rendimiento fueron pareciéndose cada vez más a los tests de aptitudes e inteligencia. Los procedimientos para la elaboración y la valoración de todos ellos tienen mucho en común. El esfuerzo creciente por preparar pruebas de rendimiento que midieran la consecución de objetivos educativos amplios, en contraposición a la insistencia en hechos sin importancia, contribuyó también a que su contenido se pareciera más al de los tests de inteligencia. En la actualidad la diferencia entre estos dos tipos de pruebas se cifra principalmente en el grado de especificidad del contenido y en la medida en que el test presuponga un curso expreso de instrucción previa. (Anastasi, 1973).

CRITERIOS PARA LA CONSTRUCCION Y VALORACION DE LOS INSTRUMENTOS DE MEDICION

El diseño de cualquier instrumento psicológico se guía por los siguientes criterios que deben ser examinados detenidamente para decidir si un instrumento que se pretende utilizar constituye un buen diseño de medición.

1. Estandarización
2. Objetividad
3. Análisis de los reactivos
4. Confiabilidad
5. Validez
6. Normas y puntajes derivados

1. Estandarización

Al definir una prueba como una situación estandarizada; el adjetivo estandarizada o tipificada significa que todos los procedimientos del examen son planteados antes, de tal manera que todos los sujetos son examinados con las mismas preguntas o problemas de la misma manera (Adkins, 1971).

Una prueba es estandarizada o tipificada únicamente si dos examinadores de la prueba trabajando independientemente podrían obtener los mismos resultados con el mismo grupo de sujetos.

Algunos de los puntos esenciales de la estandarización son:

- a) todos los sujetos deben contestar las mismas preguntas
- b) las instrucciones deben ser claras y las mismas para todos los sujetos
- c) ningún sujeto debe recibir alguna ventaja que no reciban los demás
- d) debe ser aplicado un sistema predeterminado de calificar uniformemente las respuestas de todos los estudiantes

La única variable independiente suele ser el individuo mismo al que se le aplica la prueba.

Otro paso importante en la estandarización es el establecimiento de normas. Sin normas no es posible interpretar las puntuaciones de las pruebas. La puntuación de un individuo solo puede valorarse comparándola con las obtenidas por otros. Como su nombre indica, una norma es la actuación media o normal. La puntuación directa obtenida en una prueba no tiene ninguna significación hasta que se valora en función de un adecuado conjunto de normas.

En el proceso de la estandarización de una prueba, este debe aplicarse a una gran muestra que represente el tipo de sujetos a los que se destina. Este grupo, conocido como muestra de

estandarización, sirve para establecer las normas, que permiten la designación de la posición del individuo respecto de la muestra normativa o de tipificación.

Si no se cumplen estos criterios de una medida estandarizada, no se puede hablar de una prueba o de un test psicológico (Magnusson, 1972).

2. Objetividad

En el párrafo anterior se describieron algunos aspectos de la objetividad de las pruebas.

Un máximo grado de objetividad se halla íntimamente relacionado con la estructura del instrumento, la determinación del grado de dificultad de los reactivos de la prueba, su validez, su confiabilidad (que se basan en procedimientos empíricos y objetivos) y especialmente con los criterios de calificación.

Así la aplicación, la puntuación y la interpretación de las puntuaciones son objetivas en cuanto son independientes del juicio del examinador.

3. Análisis de los reactivos

La media y la varianza de los reactivos de una prueba

Tomando como ejemplo una prueba con n reactivos aplicada a una muestra de N sujetos, se puede obtener el número de reactivos contestados correctamente por cada sujeto.

Los puntajes obtenidos por N sujetos se denominan $X_1, X_2, \dots, \dots, X_n$. Aparte se puede obtener el número de sujetos que contestan correctamente cada uno de los n reactivos. Estos números se denominan P_1, P_2, \dots, P_n y normalmente se dividen entre N para obtener las proporciones p_1, p_2, \dots, p_n . En las pruebas de rendimiento, estas proporciones representan los índices de dificultad de los reactivos.

Así, cuando pocos sujetos de la muestra N contestan correctamente el reactivo, por ejemplo, $p_i = 0.05$ ó $p_i = 0.10$, el reactivo es un reactivo difícil.

Cuando una proporción alta de la muestra contesta correctamente el reactivo, por ej., $p_i = 0.80$ ó $p_i = 0.90$, se puede considerar que el reactivo es fácil.

La proporción se relaciona inversamente con la dificultad, a mayor proporción le corresponde un reactivo más fácil. La cantidad p_i representa la media aritmética de los reactivos individuales.

Cada uno de los N sujetos obtiene un puntaje de 1 ó 0 para cada ítem, y p_i es la media del ítem i .

La varianza de cada reactivo se define como $S^2 = \frac{\sum (X - \bar{X})^2}{N}$

El índice de dificultad de un reactivo, es decir, la media del reactivo, no es independiente de la varianza del reactivo. La varianza es máxima con un valor de $p_i = .50$.

La varianza se aproxima a cero cuando p_i se aproxima a 1.00 o cero.

Correlación punto-biserial

La correlación punto-biserial es un método de selección de los reactivos que se utilizan en la construcción de una prueba de rendimiento. Cuando no se tiene un criterio externo para comprobar la validez de los reactivos, entonces se usa normalmente el criterio de la consistencia interna. El cons

structor del test selecciona del conjunto de reactivos construidos, aquellos que exhiben mayor consistencia.

Ese método de selección, que se aplica en ausencia de un criterio externo, proporciona soluciones aproximadas al problema de selección. El conjunto de reactivos seleccionados no es generalmente el de una óptima consistencia interna, pero debe ser uno de los de alta consistencia interna.

"El coeficiente de correlación punto-biserial es una medida de la capacidad de la prueba para discriminar entre dos grupos. Este estadístico se puede interpretar como una medida del grado de la tendencia con la cual una variable continua (las calificaciones totales) diferencia o discrimina entre las dos categorías de la variable dicotómica (respuestas correctas y respuestas incorrectas)." (Ferguson, 1971).

En resumen, el análisis de las correlaciones es útil para especificar la forma y el grado de relaciones funcionales imperfectas.

Cuando una variable dicotómica está correlacionada con una variable continua se utiliza la fórmula "punto-biserial" que es la siguiente:

$$r_{pb\dot{i}} = \frac{M_p - M_q}{s_t} \sqrt{pq}$$

donde:

M_p = calificación promedio de los sujetos que contestan el reactivo correctamente

M_q = calificación promedio de los sujetos que contestan el reactivo incorrectamente

s = desviación estándar de las calificaciones

p = proporción del grupo que contestó correctamente el reactivo

q = (1-p)

$r_{pb\dot{i}}$ = coeficiente punto-biserial para cada reactivo

Cualquier reactivo que tenga una correlación punto biserial cerca de 0 o negativa con el puntaje total de la prueba debe ser investigado con cuidado (Nunally, 1967). En una prueba de rendimiento es posible que un reactivo correlacione cerca de cero con el puntaje total y de todos modos represente un

reactivo válido, pero esto sucede muy pocas veces. Es mas probable que este reactivo sea excesivamente difícil o fácil, es ambiguo, o de hecho tenga muy poco que ver con el contenido de la prueba. Generalmente, este tipo de reactivo debe ser eliminado.

Entre los reactivos restantes, los reactivos que correlacionan más alto con los puntajes totales son, generalmente, los mejores reactivos. Son probablemente menos ambiguos, no pueden ser de una dificultad extrema, tanto en una como en otra dirección y tenderán a hacer la prueba altamente confiable.

"Las correlaciones se encuentran normalmente entre cero y alrededor de .40. Correlaciones arriba de .20 se consideran buenas." (Nunnally, 1967).

4. Confiabilidad

El concepto de confiabilidad de las pruebas se define generalmente como "la precisión o consistencia con la que mide una prueba lo que pretende medir". (Guilford, 1965).

Como una medición pobre puede invalidar cualquier investigación científica, tanto la validez como la confiabilidad deben

ser considerados la base para la construcción de un instrumento de medición.

La confiabilidad es una condición necesaria pero no suficiente para la validez de una prueba (Nunnally, 1967).

La confiabilidad de una prueba se refiere a la consistencia de las puntuaciones obtenidas por los mismos individuos cuando son examinados con la misma prueba en diferentes ocasiones y bajo otras condiciones variables del examen.

"Las medidas que se obtienen al llevar a cabo experimentos, están sujetas a errores de mayor o menor grado. Los errores de medida pueden ocurrir sistemáticamente o al azar. Una medida es considerada confiable según el grado en que los errores de medida sean pequeños." (Morales, 1975).

"En otras palabras, la confiabilidad puede ser definida como la relativa ausencia de errores de medida en un instrumento de medición; con más errores, menos confiabilidad; con menos errores mayor confiabilidad. Un índice de 1.00 indica una confiabilidad perfecta. En las pruebas de rendimiento se considera un buen índice de confiabilidad a partir de un valor de .80" (Nunnally, 1967).

Métodos para estimar la confiabilidad

"El coeficiente de confiabilidad (r_{tt}) puede ser estimado a partir de diversos procedimientos prácticos que se mencionan a continuación:" (Ferguson, 1971).

1. Método test-retest

El mismo instrumento de medida se aplica en dos ocasiones a la misma muestra de sujetos. Cuando el instrumento es una prueba psicológica, la prueba se aplica dos veces a una muestra de sujetos y los puntajes se correlacionan.

2. Método de formas paralelas

Se pueden aplicar formas paralelas o equivalentes de una prueba al mismo grupo de sujetos, y correlacionar las observaciones. Se necesitan criterios para establecer que existe paralelismo.

3. Método de división por mitades

Este método es apropiado cuando una prueba homogénea puede ser dividida de alguna manera en dos mitades y se pue

den obtener dos calificaciones. Estos dos puntajes se correlacionan. Este método es un procedimiento común para pruebas psicológicas con las que se pueden obtener calificaciones para los reactivos pares o impares.

Para la construcción de la Prueba de Conocimiento del Lenguaje se utilizó "el método de división por mitades para establecer el coeficiente de confiabilidad debido a que se trata de una prueba homogénea." (Guilford, 1965).

El procedimiento es el siguiente:

Al computar los datos obtenidos en los piloteos, se divide la prueba en reactivos pares e impares. Se obtienen puntajes para las dos mitades y estas se correlacionan.

La correlación pares-impares se computa con el coeficiente de correlación de Pearson de la siguiente manera:

$$r_{xy} = \frac{N (\sum XY) - (\sum X) (\sum Y)}{\sqrt{N (\sum X^2) - (\sum X)^2} \sqrt{N (\sum Y^2) - (\sum Y)^2}}$$

X = reactivos pares
 Y = reactivos impares
 N = número de sujetos

El coeficiente de confiabilidad para todo el test se computa a partir del coeficiente de correlación pares-impares de Pearson de la siguiente forma:

(Fórmula de Spearman-Brown de Confiabilidad)

$$r_{tt} = \frac{2r_{xy}}{1+r_{xy}}$$

$x =$ reactivos pares
 $y =$ reactivos impares

r_{tt} representa la confiabilidad de toda la prueba.

Otro método fue desarrollado por Kuder-Richardson cuya fórmula la K-R 20 toma en cuenta todas las posibles maneras de subdividir una prueba cuando se tiene reactivos dicotómicos (Guilford, 1965):

Fórmula de Kuder-Richardson K-R 20 de confiabilidad

$$r_{tt} = \left(\frac{n}{n-1}\right) \frac{s_t^2 - \frac{\sum_{i=1}^n p_i q_i}{n}}{s_t^2}$$

en donde:

$n =$ número de reactivos

$s_t^2 =$ varianza de los puntajes en una prueba, se define como $(X-\bar{X})^2/N$

$p_i q_i =$ producto de las proporciones de respuesta correctas e incorrectas para el reactivo i

$$\sum_{i=1}^n p_i q_i = \text{suma de estos productos para } n \text{ reactivos}$$

El coeficiente r_{tt} computado con esta fórmula tomará valores entre cero y la unidad.

"Si se satisfacen todos los supuestos del método pares-impares para estimar la confiabilidad, la fórmula de división por mitades y la de K-R 20 darán resultados idénticos (Ferguson, 1951). Pero como estos supuestos son satisfechos muy raras veces, en la práctica se obtienen diferencias en los coeficientes. Un problema con el método de división por mitades consiste en que la prueba puede ser dividida de muchas maneras, obteniéndose muchos valores diferentes para r_{tt} .

Se puede demostrar que si una prueba es dividida de todas las maneras posibles, el promedio de todos los coeficientes de confiabilidad obtenidos con la fórmula Spearman-Brown será la fórmula de Kuder-Richardson 20. Este coeficiente tiene un valor único para una prueba determinada.

La fórmula Kuder-Richardson 20 es una medida de la consistencia interna, y homogeneidad de una prueba. Si los reactivos en una prueba tienen una alta intercorrelación y miden aproximadamente el mismo atributo, entonces el coeficiente de con-

fiabilidad será alto. Cuando las intercorrelaciones son bajas, debido a que los reactivos miden atributos diferentes o porque existe error, entonces el coeficiente de confiabilidad será bajo." (Nunnally, 1967).

5. Validez

Se define la validez como "el grado con el que una prueba mide lo que intenta medir" (Tyler, 1972). La validez de una prueba se refiere "a lo que ésta mide y a cómo lo mide. Solamente es posible definir qué es lo que mide una prueba determinada mediante el examen de los criterios específicos y otras fuentes objetivas de información, utilizadas para establecer su validez" (Anastasi, 1950). Además no puede informarse sobre la validez de una prueba en términos generales. No puede decirse de ninguna prueba que tiene una validez alta o baja. en abstracto. Su validez debe determinarse respecto al uso particular para el que consideramos la prueba.

La validez es cuestión de grado y no una propiedad que se posee o no, y la validación es un proceso que jamás termina. Mientras las medidas de longitud y otros atributos físicos sencillos pueden haber probado sus méritos y nadie las cambiaría por otras medidas, la mayoría de las medidas deberían

ser vigiladas constantemente para observar si se conducen como deberían. Tanto nuevas evidencias como el desarrollo de enfoques nuevos podrían sugerir que se hagan modificaciones a los instrumentos de medida.

Lo que se valida en un sentido estricto no es el instrumento de medida, sino el uso que se haga del instrumento.

Las medidas psicológicas sirven para tres propósitos principales:

representación de un universo de contenido específico

establecimiento de una relación funcional con una variable en particular y,

explicación de las diferencias individuales

A estos tres propósitos le corresponden tres tipos de validez:

Validez de contenido

Validez predictiva o empírica (relacionada con un criterio)

Validez de constructo



Validez de contenido

En la construcción de la Prueba de Conocimiento del Lenguaje se empleó el procedimiento de validación de contenido que se utiliza normalmente para valorar las pruebas de rendimiento.

"La validez de contenido es la representatividad o el muestreo adecuado del contenido, la sustancia, los tópicos, el tema de un instrumento de medida. La validación de contenido se guía por la pregunta: "¿Son la sustancia o el contenido de este instrumento de medida, representativos del contenido del universo o del contenido de la propiedad que se está midiendo?" "Cualquier propiedad psicológica o educativa tiene un universo teórico de contenido que consiste en todas las cosas que pueden ser posiblemente dichas u observadas acerca de esta propiedad. Los miembros de este universo pueden ser llamados reactivos. El universo tiene un número infinito de miembros. Una prueba con alta validez de contenido sería teóricamente una muestra representativa del universo. Si fuera posible escoger reactivos al azar y en número suficiente, entonces cualquier muestra de reactivos formarían probablemente una prueba de alta validez de contenido" (Kerlinger, 1973).

Normalmente, y desgraciadamente, no es posible sacar muestras de reactivos al azar de un universo de contenido. Este uni-

verso existe únicamente en teoría. Si es posible y deseable hacer muchos reactivos, sobre todo en el área de rendimiento, y escoger muestras al azar de estas colecciones de reactivos para propósitos de examen. Sin embargo, la validación de contenido de tales colecciones, no importa que tan "buenos" sean los reactivos, siempre se podrá cuestionar.

"Si no es posible satisfacer la definición de validez de contenido hay que tratar de alcanzar un grado razonable de validez de contenido. La validación del contenido consiste esencialmente en un "juicio". El constructor de una prueba, sólo o con otros, juzgará la representatividad de sus reactivos. Los reactivos de una prueba deben ser estudiados examinándolos con respecto a la representatividad del universo al que pertenecen. Esto significa que cada reactivo debe ser juzgado en cuanto a su relevancia supuesta para la propiedad que se está midiendo, lo cual no es una tarea fácil. Normalmente se acude a algún experto "competente" que juzgue el contenido de los reactivos. El universo del contenido debe, si es posible, ser definido explícitamente, es decir, al experto se le deben proporcionar instrucciones específicas para que pueda dar un dictamen y valorar lo que está juzgando." (Kerlinger, 1973).

Validez predictiva o empírica

Esta validez se estudia comparando puntajes en una prueba o escala con una o más variables externas, o criterios que se supone miden el atributo que se está estudiando. Se pone énfasis en el criterio y en lo que predice el criterio.

La validez predictiva indica la eficacia de una prueba en la predicción de la conducta del individuo en situaciones específicas.

"La validez predictiva se asocia generalmente con problemas y soluciones prácticas. No interesa tanto el porqué de una ejecución sino como se puede utilizar para ayudar a resolver problemas prácticos y a tomar decisiones " (Magnusson, 1972).

Validez de constructo

Desde el punto de vista científico, la validez de constructo es uno de los adelantos más significativos en la teoría y práctica de la medida.

Al preguntar sobre la validez de constructo de una prueba lo que interesa generalmente es saber cuáles propiedades psicoló

gicas u otras pueden "explicar" la varianza de una prueba. Se trata de explicar las diferencias individuales en el puntaje de una prueba. El interés gira alrededor de la propiedad que se ha medido más que en la prueba en si misma.

El investigador empieza generalmente con los constructos o las variables que entran en relaciones funcionales. Lo que le interesa es descubrir el porqué de la existencia de estas relaciones, qué es lo que hay detrás.

Se observa que la validez de constructo y el planteamiento de investigaciones científicas están muy relacionadas. No se trata simplemente de validar una prueba. Se trata de intentar validar la teoría detrás de la prueba. Morales, 1975, cita a Cronbach quien señala "que hay tres partes en la validez de constructo: suponer cuales son los constructos que posiblemente son la causa del resultado de la ejecución de una prueba, derivar hipótesis de la teoría relacionada con el constructo, y probar las hipótesis empíricamente".

La validez de constructo contrasta con los enfoques empíricos que definen la validez de una medida únicamente por su éxito en predecir un criterio.

6. Normas y puntajes derivados

Además de la confiabilidad y la validez, debemos considerar las diversas unidades en que se expresan los puntajes para poder interpretarlos y evaluar las pruebas. Un puntaje que nos indica únicamente cuántas preguntas contestó correctamente una persona no nos dice mucho sobre ella a menos que tengamos un estándar de comparación.

"Los puntajes derivados o transformados que se emplean en pruebas estandarizadas se han diseñado para posibilitar las comparaciones entre el puntaje de un individuo y las normas del grupo. Para derivar o transformar puntajes existen diversos métodos. El más sencillo de ellos es la obtención de normas percentiles para el grupo y la transformación del puntaje de cada individuo en un ámbito percentil correspondiente." (Ferguson, 1971).

El otro método comúnmente empleado para transformar los puntajes en sus equivalentes que indiquen el lugar de un grupo donde se encuentra un individuo, utiliza, como base de las normas, la media y la desviación estándar.

Además de los rangos percentiles y los puntajes estándar hay otros medios para dar significado a los puntajes; al

hacer comparaciones interindividuales. "Las diversas unidades de medición existentes en psicometría se han elaborado para cubrir los casos específicos de medición de que se trate; es decir, ¿Que se va a medir?, ¿A quiénes se va a medir?, ¿Para qué y cuándo se va a efectuar la medición?, ¿Que fin persigue, y cómo se van a utilizar tales datos?." (Nunnally, 1964).

El primer paso para obtener normas para una prueba es la definición de una población normativa.

Cualquier norma, sea cual fuere el modo en que se exprese, se halla limitada a la población normativa determinada de la cual se ha derivado. El usuario de una prueba nunca debe perder de vista la manera en que se han establecido las normas. Las normas de los tests psicológicos en modo alguno son absolutos, universales o permanentes. Representan meramente la ejecución en el test de los sujetos que constituyen el grupo normativo. Al escoger esa muestra, se suele hacer un esfuerzo por obtener una sección transversal representativa de la población a la que se destina la prueba.

CAPITULO III

M E T O D O

Sujetos

Como la prueba está destinada a estudiantes universitarios, los sujetos fueron de la población estudiantil de la UNAM. Se llevaron a cabo 5 piloteos de la prueba con un total de 745 sujetos.

El primer piloteo se llevó a cabo con 44 sujetos de la Facultad de Psicología de la UNAM.

El segundo piloteo se realizó con 104 sujetos de la Facultad de Psicología de la UNAM.

El tercer piloteo se llevó a cabo con 75 sujetos de la Facultad de Psicología de la UNAM.

El cuarto piloteo se llevó a cabo con 285 sujetos de diversas facultades de la UNAM. (Ingeniería, Medicina, Química, Filosofía, Psicología).

El quinto y último piloteo se llevó a cabo con estudiantes de segundo semestre de la UNAM. El total de 237 sujetos corresponde a las siguientes facultades:

1. Facultad de Filosofía y Letras	40 sujetos
2. Facultad de Ciencias Políticas	40 sujetos
3. Facultad de Ingeniería	38 sujetos
4. Facultad de Arquitectura (Diseño Industrial)	39 sujetos
5. Facultad de Química	40 sujetos
6. Facultad de Psicología	40 sujetos

Procedimiento

Se hicieron aplicaciones colectivas de la prueba en salones de las distintas facultades. Se pidió a los maestros que cedieran una de sus horas de clases explicando tanto a ellos como a los estudiantes el objeto de la investigación.

Se hacía énfasis en que se iban a obtener datos valiosos acerca de la población estudiantil y que la prueba era anónima. En todos los grupos se logró motivar a los estudiantes y el 100% cooperó.

Se repartía el material de la prueba con las instrucciones a la vista, pidiendo a los sujetos que no pasaran la hoja hasta no instruírseles.

Se leían las instrucciones y se preguntaba si no tenían dudas, se procedía a dar la instrucción de iniciar las contestaciones de la prueba poniendo en la siguiente hoja los datos pedidos acerca del sexo, edad, etc. Se les pedía que no anotaran nombres. Se vigiló, hasta donde fué posible, que no se comunicaran entre ellos.

Se les daba todo el tiempo requerido para que terminaran la prueba. Se les agradecía su colaboración y en todos los

grupos los estudiantes pedían que se les aclarara sus dudas acerca de las respuestas correctas de la prueba.

Piloteos de la prueba

Se llevaron a cabo un total de cinco piloteos con la prueba.

En cada piloteo se obtuvieron los siguientes datos:

1. Número de sujetos a los que se aplicaba la prueba.
2. Correlación punto-biserial.
3. Dificultad de cada reactivo.
4. Índice de confiabilidad.
5. Media.
6. Dispersión.

Cada piloteo se realizó con muestras de la población estudiantil a la que será aplicada la prueba.

El primer piloteo se realizó con 44 sujetos de la Facultad de Psicología. No se tomó en cuenta el semestre que cursaban. Se construyó la prueba con 43 reactivos del conjunto de reactivos elaborados.

Para el segundo piloteo que se llevó a cabo con 104 sujetos de la Facultad de Psicología se modificó la primera versión de la prueba sustituyendo los reactivos de correlación punto-biserial negativa (3 reactivos) y menores de .200/ por otros (17 reactivos) reactivos elaborados. No se tomó en

cuenta el semestre que cursaban. (Nunnally, 1967).

El tercer piloteo se llevó a cabo con 75 sujetos de la Facultad de Psicología modificando la segunda versión de la prueba haciendo sustitución de 1 reactivo de correlación punto-biserial negativo, y 9 reactivos con correlación punto-biserial menor de .200. No se tomó en cuenta el semestre que cursaban los estudiantes.

El cuarto piloteo se realizó con 285 sujetos de diversas facultades (como se especificó anteriormente) haciendo sustitución de 6 reactivos, dos de correlación punto-biserial negativos, 4 con una correlación punto-biserial menor de .200 quedando un total de 38 reactivos. No se tomó en cuenta el semestre que cursaban los estudiantes. El quinto y último piloteo fué realizado con estudiantes de segundo semestre de diversas facultades de la UNAM, con un total de 237 sujetos.

En esta última versión no se presentaron ni correlaciones negativas ni menores de .200.

En las figuras N° 1, 2, 3, 4, 5, 6 y 7 que aparecen a continuación, se representa la ejecución de los sujetos de cada facultad.

FIGURA N° 1

DISTRIBUCION DE FRECUENCIAS ABSOLUTAS

5° piloteo con el total de 237
estudiantes de segundo semestre
de la UNAM

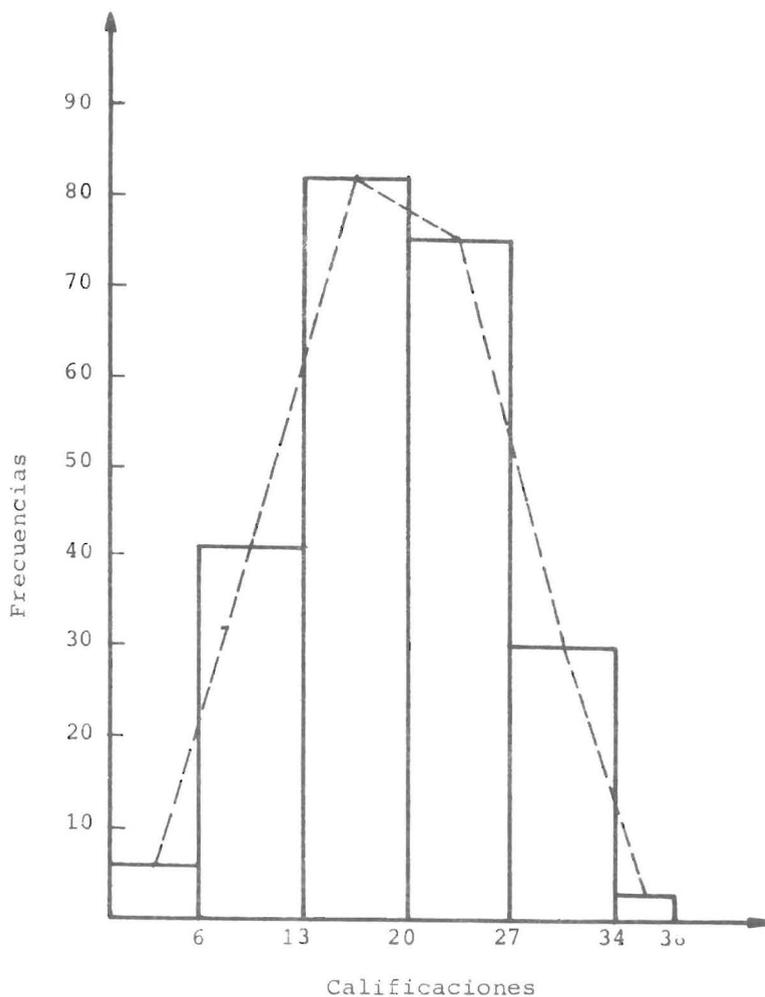


FIGURA N° 2

DISTRIBUCION DE FRECUENCIAS ABSOLUTAS

5° piloteo con 40 estudiantes del
segundo semestre de la Facultad
de Filosofía y Letras.
Fecha: 25/VIII/1976

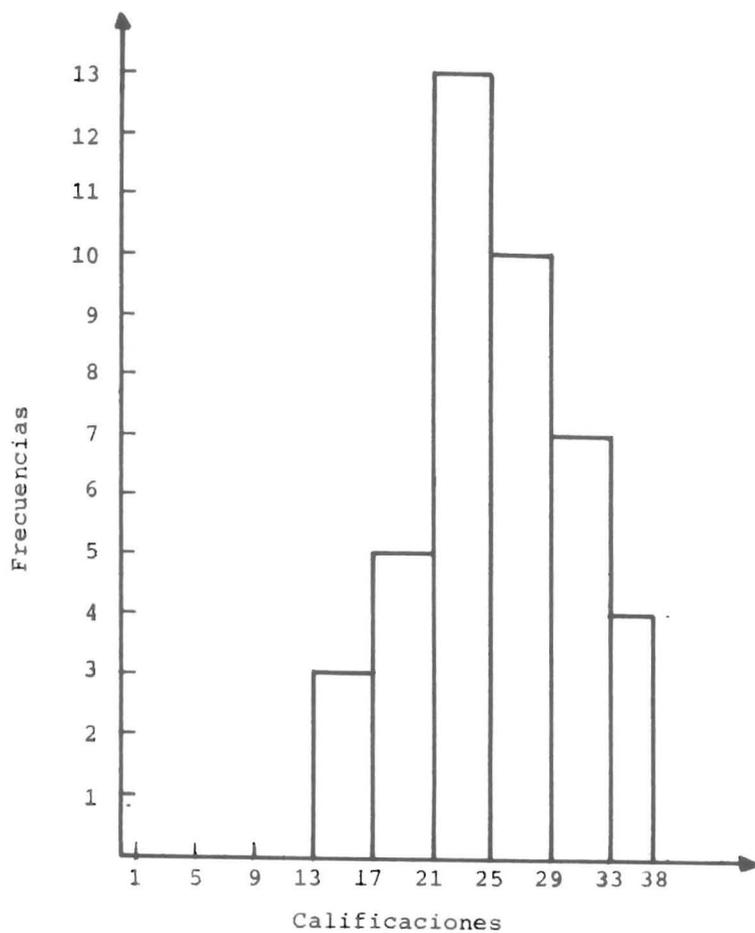


FIGURA N° 3

DISTRIBUCION DE FRECUENCIAS ABSOLUTAS

5° piloteo con 40 estudiantes
del segundo semestre de la Fa-
cultad de Ciencias Políticas y
Sociales.
Fecna: 23/VIII/1976

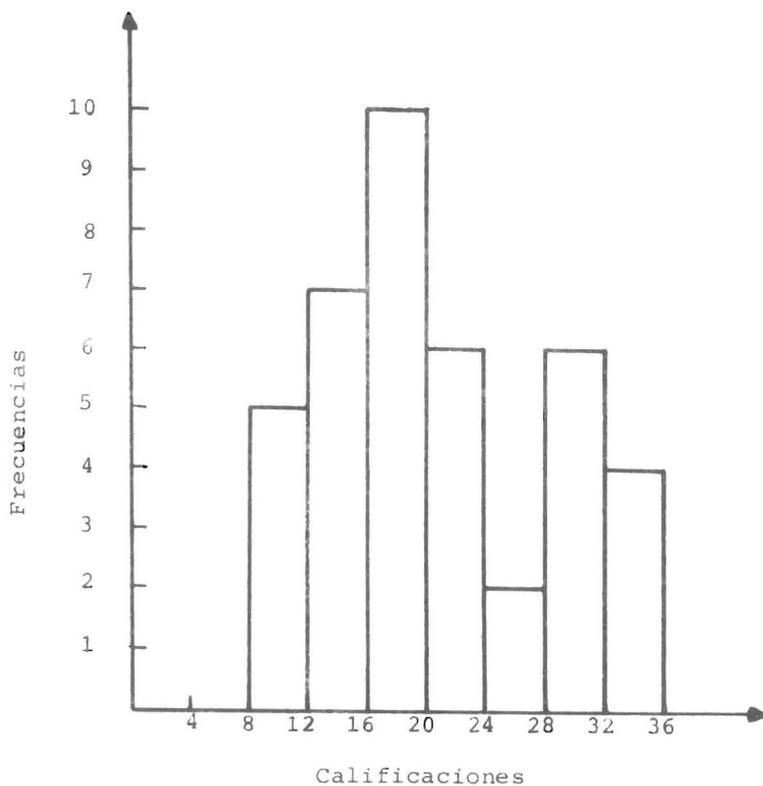


FIGURA N° 4

DISTRIBUCION DE FRECUENCIAS ABSOLUTAS

5° piloteo con 38 estudiantes
de segundo semestre de la Fa-
cultad de Ingeniería.
Fecha: 8/IX/1976

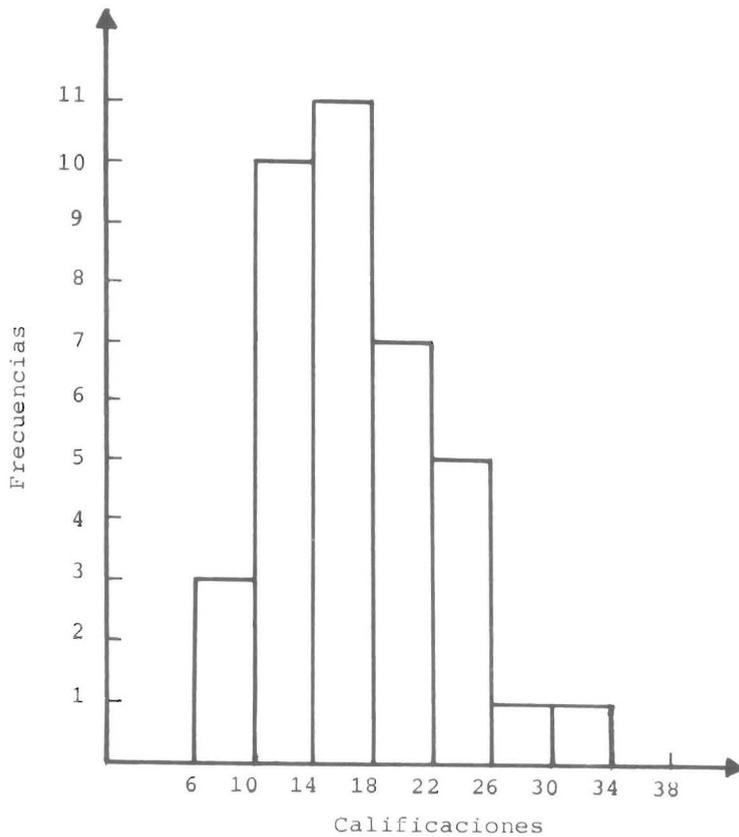


FIGURA N° 5

DISTRIBUCION DE FRECUENCIAS ABSOLUTAS

5° piloteo con 39 estudiantes
del segundo semestre de la
Facultad de Arquitectura (Diseño Industrial).
Fecha: 2/IX/1976

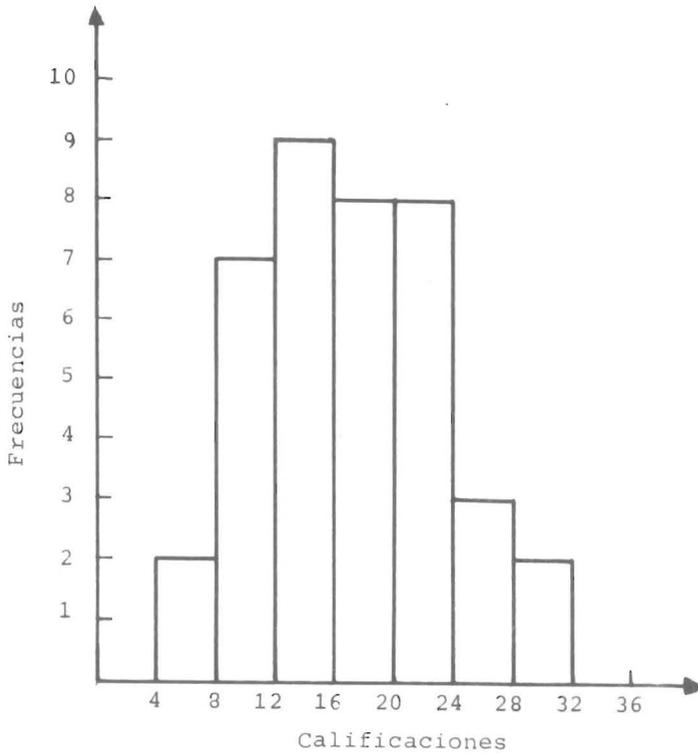


FIGURA N° 6

DISTRIBUCION DE FRECUENCIAS ABSOLUTAS

5° piloteo con 40 estudiantes
del segundo semestre de la fa
cultad de Química
Fecha: 3/IX/1976

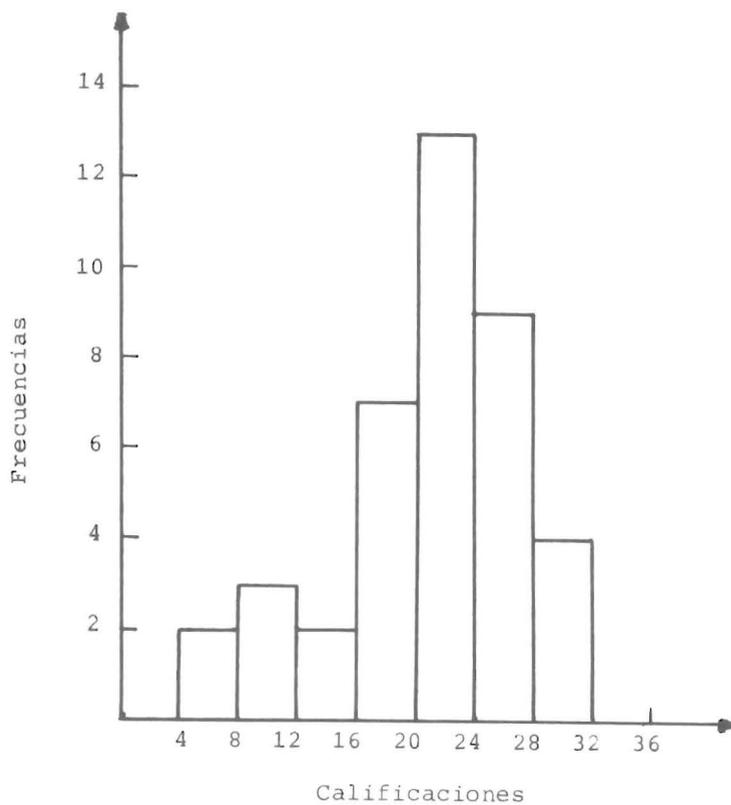
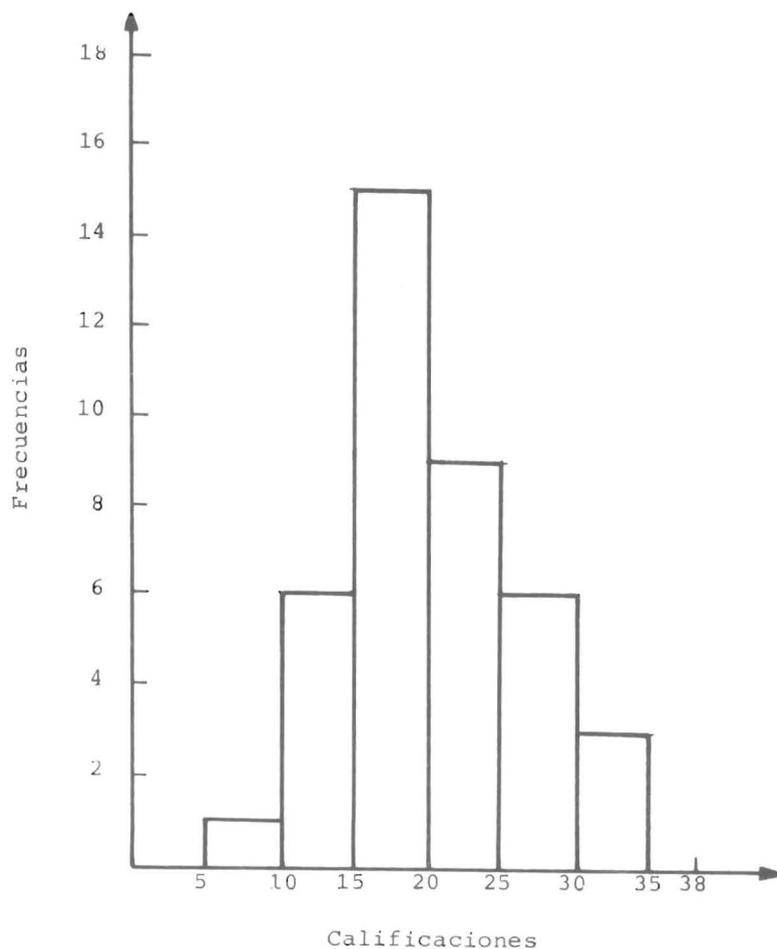


FIGURA N° 7

DISTRIBUCION DE FRECUENCIAS ABSOLUTAS

5° piloteo con 40 estudiantes
de la Facultad de Psicología
de segundo semestre.
Fecha: 18/VIII/1976



Análisis de los datos y resultados

En las Tablas I, II, III, IV y V se especifican los resultados del análisis de los reactivos para los cinco pilotos.

Se efectuó un análisis de reactivos haciendo una correlación punto-biserial entre cada reactivo y la calificación total de la prueba utilizando la fórmula:

$$r_{pbi} = \frac{M_p - M_q}{s_x} \sqrt{pq}$$

M_p = media para los sujetos que contestan el reactivo correctamente

M_q = media para los sujetos que contestan el reactivo incorrectamente

s_x = desviación estándar de todos los reactivos

p = proporción del grupo que contestó correctamente

q = proporción del grupo que no contestó correctamente (1-p)

Se obtuvo el índice de dificultad de cada reactivo, dividiendo el número de sujetos que obtuvieron los reactivos correctos entre el total de los sujetos.

Se obtuvo la confiabilidad de la prueba por bipartición de pares impares para cada piloteo utilizando la fórmula Spearman-Brown y la de Kuder-Richardson KR 20.

Correlación Pares-Impares: Coeficiente de correlación de

$$\text{Pearson} = r_{xy}$$

$$r_{xy} = \frac{N(\sum XY) - (\sum X)(\sum Y)}{\sqrt{N(\sum X^2) - (\sum X)^2} \sqrt{N(\sum Y^2) - (\sum Y)^2}}$$

X = reactivos pares
Y = reactivos impares
N = número de sujetos

Confiabilidad Fórmula Spearman-Brown

$$r_{tt} = \frac{2r_{xy}}{1+r_{xy}}$$

X = reactivos pares
Y = reactivos impares

Estimación de la confiabilidad Kuder-Richardson

$$r_{tt} = \left(\frac{n}{n-1}\right) \frac{\sigma^2_t - \sum pq}{\sigma^2_t}$$

n = número de ítems

p = proporción de respuestas correctas

q = 1-p proporción de respuestas incorrectas

σ^2_t = varianza

La correlación pares-impares y la confiabilidad en los cinco piloteos fue la siguiente:

<u>Correlación pares-impares</u>	<u>Confiabilidad</u>
1. $r_{xy} = 0.616$	$r_{tt} = 0.7622$
2. $r_{xy} = 0.707$	$r_{tt} = 0.8284$
3. $r_{xy} = 0.771$	$r_{tt} = 0.8704$
4. $r_{xy} = 0.715$	$r_{tt} = 0.8339$
5. $r_{xy} = 0.779$	$r_{tt} = 0.8755$ KR = 0.8252

Las medias y las dispersiones de cada piloteo fueron las siguientes:

<u>Medias</u>	<u>Dispersiones</u>
1. $\bar{x} = 20.023$	$s = 5.428$
2. $\bar{x} = 19.558$	$s = 5.913$
3. $\bar{x} = 18.827$	$s = 6.56$
4. $\bar{x} = 20.600$	$s = 6.306$
5. $\bar{x} = 19.692$	$s = 6.809$

La confiabilidad obtenida de la prueba es alta y se puede considerar como una medida estable y de poder discriminativo.

Los datos fueron procesados mediante un programa escrito en Fortran en la computadora B/6700 del Centro de Cálculo de la UNAM.

TABLA I

Primer piloteo

Los resultados del análisis de reactivos fueron los siguientes:

ITEM	CORRELACION	DIFICULTAD DEL ITEM
1	0.021	0.659
2	0.461	0.432
3	0.167	0.409
4	-0.091	0.727
5	0.197	0.818
6	-0.006	0.682
7	0.635	0.318
8	0.416	0.614
9	0.273	0.159
10	0.533	0.273
11	0.360	0.818
12	0.507	0.477
13	0.440	0.500
14	0.223	0.114
15	0.186	0.886
16	0.476	0.364
17	0.367	0.227
18	-0.012	0.432
19	0.243	0.409
20	0.180	0.523
21	0.465	0.409
22	0.073	0.409

TABLA I (Continuación)

ITEM	CORRELACION	DIFICULTAD DEL ITEM
23	0.148	0.455
24	0.410	0.159
25	0.342	0.159
26	0.197	0.477
27	0.306	0.500
28	0.081	0.614
29	0.298	0.545
30	0.093	0.841
31	0.187	0.818
32	0.262	0.341
33	0.036	0.250
34	0.423	0.364
35	0.277	0.277
36	0.282	0.682
37	0.471	0.705
38	0.616	0.386
39	0.191	0.727
40	0.259	0.182
41	0.377	0.341
42	0.168	0.227
43	0.188	0.364

La confiabilidad por bipartición pares-impares, con la corrección de Spearman Brown fue de $r_{tt} = 0.7622$

La media fue de 20.023 reactivos correctos de 43, con una dispersión de 5.428.

TABLA II

Segundo piloteo

Los resultados del análisis de reactivos fueron los siguientes:

ITEM	CORRELACION	DIFICULTAD DEL ITEM
1	0.246	0.750
2	0.288	0.837
3	0.177	0.510
4	0.118	0.615
5	0.016	0.663
6	0.254	0.452
7	0.464	0.644
8	-0.034	0.923
9	0.396	0.404
10	0.123	0.058
11	0.439	0.519
12	0.419	0.471
13.	0.020	0.346
14	0.273	0.375
15	0.358	0.731
16	0.536	0.365
17	0.045	0.192
18	0.120	0.731
19	0.200	0.385
20	0.106	0.202
21	0.437	0.846
22	0.452	0.683

TABLA II (Continuación)

ITEM	CORRELACION	DIFICULTAD DEL ITEM
23	0.432	0.625
24	0.517	0.385
25	0.374	0.471
26	0.517	0.327
27	0.302	0.692
28	0.487	0.558
29	0.494	0.356
30	0.137	0.279
31	0.399	0.702
32	0.322	0.712
33	0.401	0.375
34	0.349	0.308
35	0.351	0.365
36	0.413	0.423
37	0.338	0.288
38	0.468	0.471
39	0.344	0.317
40	0.471	0.202

La confiabilidad por bipartición pares-impares, con la corrección de Spearman Brown, fue de $r_{tt} = 0.8284$

La media fue de 19.558 reactivos correctos de 40, con una dispersión de 5.913.

TABLA III

Tercer piloteo

Los resultados del análisis de reactivos fueron los siguientes:

ITEM	CORRELACION	DIFICULTAD DEL ITEM
1	0.308	0.560
2	0.470	0.213
3	0.477	0.187
4	0.449	0.800
5	0.415	0.840
6	0.312	0.360
7	0.499	0.733
8	0.340	0.773
9	0.412	0.827
10	0.086	0.147
11	0.316	0.587
12	0.274	0.493
13	0.298	0.373
14	0.118	0.133
15	-0.016	0.160
16	0.293	0.640
17	0.481	0.480
18	0.403	0.400
19	0.419	0.800
20	0.475	0.267
21	0.418	0.787
22	0.304	0.613

TABLA III (Continuación)

ITEM	CORRELACION	DIFICULTAD DEL ITEM
23	0.556	0.520
24	0.427	0.307
25	0.478	0.440
26	0.506	0.200
27	0.465	0.707
28	0.387	0.400
29	0.423	0.307
30	0.472	0.720
31	0.432	0.720
32	0.472	0.347
33	0.315	0.227
34	0.406	0.413
35	0.584	0.507
36	0.449	0.373
37	0.016	0.267
38	0.181	0.347
39	0.247	0.227
40	-0.025	0.627

La confiabilidad por bipartición pares-impares, con la corrección de Spearman-Brown, fué de $r_{tt} = 0.8704$

La media fué de 18.827 reactivos correctos de 40, con una dispersión de 6.562.

TABLA IV

Cuarto piloteo

Los resultados del análisis de reactivos fueron los siguientes:

ITEM	CORRELACION	DIFICULTAD DEL ITEM
1	0.334	0.477
2	0.399	0.186
3	0.415	0.172
4	0.277	0.874
5	0.282	0.860
6	0.262	0.340
7	0.329	0.814
8	0.359	0.846
9	0.392	0.902
10	0.431	0.460
11	0.322	0.526
12	0.166	0.544
13	0.382	0.526
14	0.461	0.389
15	0.308	0.505
16	0.184	0.477
17	0.409	0.474
18	0.403	0.421
19	0.360	0.751
20	0.518	0.396
21	0.457	0.863

TABLA IV (Continuación)

ITEM	CORRELACION	DIFICULTAD DEL ITEM
22	0.473	0.744
23	0.421	0.681
24	0.543	0.544
25	0.328	0.474
26	0.460	0.453
27	0.304	0.754
28	0.355	0.428
29	0.439	0.365
30	0.339	0.670
31	0.332	0.723
32	0.524	0.344
33	0.450	0.326
34	0.265	0.319
35	0.362	0.467
36	0.373	0.435
37	0.268	0.302
38	0.151	0.768

La confiabilidad por bipartición pares-impares con la corrección de Spearman-Brown, fué de $r_{tt} = 0.8339$.

La media fué de 20.600 reactivos correctos de 38, con una dispersión de 6.306.

TABLA V

Quinto piloteo

Los resultados del análisis de reactivos fueron los siguientes:

ITEM	CORRELACION	DIFICULTAD DEL ITEM
1	0.355	0.540
2	0.403	0.485
3	0.340	0.797
4	0.352	0.861
5	0.362	0.738
6	0.439	0.414
7	0.392	0.300
8	0.400	0.684
9	0.378	0.692
10	0.497	0.333
11	0.496	0.236
12	0.502	0.224
13	0.288	0.848
14	0.208	0.831
15	0.245	0.287
16	0.327	0.793
17	0.450	0.730
18	0.508	0.249
19	0.459	0.827
20	0.312	0.726
21	0.437	0.599



TABLA V (Continuación)

ITEM	CORRELACION	DIFICULTAD DEL ITEM
22	0.501	0.409
23	0.322	0.439
24	0.481	0.304
25	0.349	0.325
26	0.446	0.392
27	0.493	0.464
28	0.504	0.591
29	0.349	0.511
30	0.285	0.519
31	0.301	0.359
32	0.279	0.291
33	0.311	0.422
34	0.418	0.485
35	0.427	0.485
36	0.422	0.523
37	0.409	0.430
38	0.247	0.549

La confiabilidad por bipartición pares-impares con la corrección de Spearman-Brown fué de $r_{tt} = 0.8755$.

La media fué de 19.692 reactivos correctos de 38, con una dispersión de 6.809.

La confiabilidad por el método de Kuder-Richardson fué de $KR 20 = 0.8252$.

Validez de contenido

El contenido de esta prueba fué elegido del dominio de la enseñanza del lenguaje en la enseñanza primaria y secundaria. Se acudió para asesoramiento a una experta en letras españolas con experiencia en la enseñanza media de lengua y literatura española.

Obviamente, la prueba no es validada en términos de validez predictiva, porque la prueba pretende medir directamente la ejecución.

Los reactivos fueron escogidos basándose en los criterios actuales para el estudio del lenguaje. Los libros de textos actuales citados en la bibliografía sirvieron de base para la selección de los reactivos. La colección de reactivos representa una muestra de las áreas estudiadas en la enseñanza del lenguaje.

En el Capítulo I se describió con detalle los criterios utilizados para la construcción de la prueba y el número y tipo de reactivos que corresponde a cada criterio.

La dificultad de escoger reactivos para una población que

viene de muy distintos sistemas de educación se trató de resolver diseñando reactivos que cubran conocimientos realmente básicos que deberían ser cubiertos bajo cualquier sistema de educación básica.

RESUMEN

Se estructuró un instrumento de medición del conocimiento del lenguaje para estudiantes universitarios.

Se obtuvo una confiabilidad de .8755 con la fórmula de Spearman-Brown y de .8252 con la fórmula de Kuder-Richardson.

Se realizó un análisis de reactivos. Se obtuvieron las medias y dispersiones de la población de 745 sujetos estudiados en 5 piloteos de la UNAM.

Se pretende integrar esta prueba a un instrumento que mida velocidad, comprensión y habilidad gramatical de los estudiantes universitarios. Para la utilización de la prueba en estudios de campo se necesita obtener normas y puntajes estandarizados.

CONCLUSIONES

Se obtuvo un instrumento de medida con un alto coeficiente de confiabilidad (0.8755 y 0.8252). La prueba presenta consistencia interna y es una prueba homogénea. Los reactivos tienen alta intercorrelación y miden aproximadamente el mismo atributo según el alto coeficiente de confiabilidad obtenido. Se considera que la prueba constituye un instrumento estable y de poder discriminativo.

Se considera, que presenta validez de contenido y que se haya alcanzado un grado razonable de validez de contenido. Los reactivos fueron juzgados y valorados (Kerlinger, 1973) "con respecto a su representatividad del universo al que pertenecen, por una persona experta" en la investigación y la ense

ñanza media de la lengua española (Cardero García, A.M., 1975).

Las calificaciones obtenidas por los estudiantes universitarios de la UNAM, y las opiniones recogidas de la mayoría de los maestros que prestaron sus grupos para la aplicación de la prueba, ponen en evidencia que el manejo del lenguaje es, en general, deficiente.

Sobre todo en las carreras técnicas, como en la Facultad de Arquitectura y en la Facultad de Ingeniería, la preocupación por este hecho fué notable y los maestros señalaban que existe una necesidad de planear cursos de redacción y de comprensión de la lectura, tanto como repasos de la gramática y entrenamientos en la precisión para el empleo del lenguaje, para el curriculum de la carrera desde el inicio de la misma.

SUGERENCIAS

El instrumento de medida sirve para dos propósitos principales, primero representa un universo de contenido específico y segundo, será empleado para plantear hipótesis acerca de las relaciones entre el conocimiento del lenguaje, la comprensión de la lectura y la velocidad de la lectura y probarlas empíricamente.

Lo que interesa investigar son las relaciones funcionales en el idioma español que se han encontrado entre el conocimiento del lenguaje en el idioma inglés (vea revisión bibliográfica) y descubrir el porqué de la existencia de estas relaciones. Habrá que investigar si este instrumento representa una variable o constructo que se supone influye decisivamente sobre

la comprensión de la lectura. El paso siguiente sería plantear esta hipótesis e investigarla. Después de utilizar la prueba experimentalmente, se podrá investigar la validez de constructo de la prueba. La validez de constructo contrasta con los enfoques empíricos que definen la validez de una meda únicamente por su éxito en predecir un criterio. El interés se desplaza hacia la propiedad que se pretende medir y se aleja del interés por la prueba en si misma.

Hay una necesidad práctica de obtener datos sobre las habilidades, capacidades y necesidades de los estudiantes y un instrumento como este podrá ayudar para agrupar estudiantes con fines de instrucción, cursos de remedio, entrenamiento y para determinar niveles de conocimiento y rendimiento. Se podrá utilizar para la detección de problemas y labor de orientación. Se podrá comparar el rendimiento de los estudiantes que vienen de diferentes escuelas y de diferentes partes del país y hacer el análisis de la población estudiantil.

Se sugiere la aplicación de esta prueba, por ejemplo, a los estudiantes de la Universidad Abierta, porque los estudiantes que estudian bajo este sistema, dependen para su éxito o fracaso, fundamentalmente, de sus habilidades en la lectura y la comprensión. En este caso, se podría hacer una labor de orientación a los estudiantes después de haber en-

contrado una relación funcional entre el rendimiento académico y el rendimiento en la prueba. También existe la posibilidad de utilizar la prueba para poder estudiar las relaciones que pueden existir entre el conocimiento del idioma materno español y el éxito en la comprensión de una segunda lengua.

Se hace énfasis, en que no se pretende con esta prueba "marcar" a los estudiantes, injustamente segregarlos en grupos con habilidades mayores o menores, restringiendo los rangos dentro de los cuales los estudiantes podrán crecer y cambiar, sino que se intenta obtener información.

Las pruebas pretenden adquirir información y hay que evitar que esta información se use inadecuadamente. Cuando hay que hacer decisiones importantes, hay que acudir a toda la información al alcance.

Para poder utilizar esta prueba posteriormente en investigaciones de campo, el siguiente paso sería aplicarla a una población representativa de todos los estudiantes del país, como población normativa.

A P E N D I C E

PRUEBA DE CONOCIMIENTO DEL LENGUAJE .

(Elke Kort de Capella)

PRUEBA DE CONOCIMIENTO DEL LENGUAJE

NOMBRE _____

SEXO _____

EDAD _____

UNIVERSIDAD _____

FACULTAD O DEPARTAMENTO _____

AÑO, SEMESTRE O TRIMESTRE QUE CURSA _____

FECHA DE APLICACION _____

I N S T R U C C I O N E S :

Cuando el examinador dé la señal, da la vuelta a la hoja y comienza a contestar las preguntas a continuación.

En cada caso fíjate bien en la respuesta que se te pide, hay respuestas que exigen que subrayes, complementes, etc.

PRUEBA DE CONOCIMIENTO DEL LENGUAJE

Señala los complementos indirectos en las siguientes oraciones:

1. No abras la puerta a nadie
2. Fernando compró un libro para su padre.

Antepón un prefijo (derivativo) para negar el significado de las siguientes palabras:

3. _____ humano
4. _____ moral

Escribe el plural de las siguientes palabras:

5. cualquiera _____
6. album _____
7. lord _____

Escribe los dos participios correspondientes a cada verbo.

Ejemplo: confundir - confundido - confuso

8. abstraer _____

9. corromper _____
10. convertir _____

Forma sustantivos abstractos de los siguientes adjetivos:

11. áspero _____
12. redondo _____

Escribe los antónimos de las siguientes palabras:

13. siempre _____
14. bajo _____

Encuentra un sinónimo para las siguientes palabras:

15. aburrirse _____
16. aborrecer _____

Escribe el superlativo de los siguientes adjetivos:

17. bueno _____
18. célebre _____

Sustituye las palabras subrayadas por el adjetivo correspondiente. Ejemplo: Una enfermedad del corazón.
Una enfermedad cardiaca.

19. Navegación por el mar. Navegación _____.
20. Las armas que ocasionan la muerte. Las armas _____.
21. El rocío de la mañana. El rocío _____.
22. Las instalaciones del puerto. Las instalaciones _____.
23. Navegación por los ríos. Navegación _____.
24. Yacimientos de oro. Yacimientos _____.

Subraya en la siguiente oración la (s) preposición (es) que encuentres:

25. El parque de la ciudad, bueno para pasear, se extiende hasta el río.

Subraya en las siguientes oraciones las conjunciones que encuentres.

26. Házme caso, siquiera por una vez.
27. Ayer salió porque hizo buen tiempo.
28. Comeremos peras o manzanas.

Acentúa o no, según convenga, la palabra marcada en cada ejemplo:

29. Llamaron al velador, el cual acudió rápidamente.
30. No sé, en fin, que debo hacer.

Acentúa o no, según convenga, las siguientes palabras:

31. el naufrago
32. el examen
33. la tesis

Las siguientes palabras están incompletas. Completa el espacio con g o j, y señala la palabra que se escribe con la letra g.

34. A. Te_er (_____)
- B. Cru_ir
- C. Gru_ir
- D. Fin_ir

Escribe la opción correcta: (_____)

35. Es necesario que nosotros los intereses del nuevo jefe.
- A. conscamos
 - B. conózcamos
 - C. conóscamos
 - D. conozcamos

En las siguientes oraciones coloca las comas en los debidos lugares:

36. Es preferible no obstante que vigile la puerta.
37. Roma la Ciudad Eterna atrae los ojos de todo el mundo por su grandeza histórica y artística.

Selecciona la palabra correcta y anota la letra correspondiente en el espacio en blanco:

38. Debemos(_____) las verduras antes de servir las.
- A. cocer
 - B. coser
 - C. cueser
 - D. cuecer

CLAVE DE RESPUESTAS

1. a nadie
2. para su padre
3. in
4. in/a
5. cualesquiera
6. álbumes
7. lores
8. abstraído - abstracto
9. corrompido-corrupto
10. convertido-converso
11. aspereza
12. redondez
13. nunca/jamás
14. alto
15. fastidiarse/hastiar
16. odiar/detestar
17. buenísimo/bonísimo/óptimo
18. celeberrimo
19. marina/marítima
20. mortíferas/mortales
21. matutino/matutinal/matinal
22. portuarias
23. fluvial
24. áureos/auríferos
25. de para hasta
26. siquiera
27. porque
28. o
29. el cual
30. qué
31. naufrago
32. examen
33. tesis
34. fingir
35. conocamos
36. _, no obstante, _
37. _, la Ciudad Eterna, _
38. cocer

BIBLIOGRAFIA

1. Adkins Wood, Dorothy.: *Elaboración de Tests*, Edit. Trillas, México, 1971.
2. Anastasi, Anne.: *Test Psicológicos*, Aguilar, S.A. de Ediciones, Madrid, 1973.
3. Bello, Andrés.: *Gramática de la Lengua Castellana*, Edit. Sopena, Argentina, 1975.
4. Bloom Benjamin, S. y Colaboradores.: *Taxonomía de los objetivos de la educación. La Clasificación de las metas educacionales*. Edit. "El Ateneo", Buenos Aires, 1974.
5. Bormuth, John R., Carr, J., Manning, J., & Pearson, D., 1970. Children's comprehension of between-and within-sentence syntactic structures. *Journal of Educational Psychology*, Vol. 61, N° 5, 349-357.
6. Bormuth, John R.: *Psycholinguistics and the Teaching of Reading*, International Reading Association, Newark, 1972.
7. Cardero García, A.M., Rull R. Angeles., 1975. *Linguística*, Libro de Consulta 1, Educación Media Básica, Nuevas Técnicas Educativas, S.A. México, D.F.

8. Ejercicios ortográficos (Teoría y Práctica de la Ortografía). Colección de Manuales Didácticos, Edit. Esfinge, S.A. México, D.F., 1971.
9. Español, -Quinto grado- Ejercicios, Secretaría de Educación Pública, México, D.F., Comisión Nacional de los Libros de Texto Gratuitos, 1975.
10. Español, -Sexto grado- Ejercicios, Secretaría de Educación Pública, México, D.F., Comisión Nacional de los Libros de Texto Gratuitos, 1975.
11. Farr, Roger.: Reading-What can be Measured, International Reading Association, Newark, 1969.
12. Ferguson, G.: Statistical Analysis in Psychology and Education, Third Edition, McGraw-Hill, New York, 1971.
13. Guilford, J.P.: Fundamental Statistics in Psychology and Education, 1965, McGraw-Hill, Inc.
14. Kerlinger, Fred, N.: Foundation of Behavioral Research, Second Edit. 1973. Holt, Rinehart and Winston, Inc.

15. Magnusson, David.: Teoría de los tests. Edit. Trillas, México, D.F., 1972.
16. Mercado, D. Serafín, J., A. Suárez, C. Ramírez, T., E. Kort, & L. E., Zendejas, 1976. La lectura de los estudiantes de la Facultad de Psicología de la UNAM de México, medida con un instrumento integrado de velocidad, comprensión y habilidad gramatical. Revista de la Educación superior, Asoc. Nal. de Universidades e Institutos de Enseñanza Superior (en prensa).
17. Morales, Ma. Luisa.: Psicometría aplicada, Edit. Trillas, México, 1975.
18. Nunnally, Jum C.: Educational Measurement and Evaluation, 1964, McGraw-Hill, Inc.
19. Nunnally, Jum C.: Psychometric Theory, 1967. McGraw-Hill, Inc.
20. Pleyán, C.: Verbo. Lengua Española 2°. Edit. Teide, S.A. Barcelona, 1972.
21. Pleyán, C.: Nombre. Lengua Española 1°. Edit. Teide, S.A. Barcelona, 1972.

22. Russel, David H., y Fea, Henry R.: Research on Teaching Reading, Capt. 16, Handbook of Research on Teaching, Gage N.L., Editor, Rand McNalley & Co., Chicago, 1963.
23. Smith, Frank.: Comprehension and Learning: A conceptual Framework for Teachers, 1975. Holt, Rinehart and Winston, New York.
24. Tyler, L.E.: Pruebas y medición en psicología, Edit. Prentice-Hall International, Englewood Cliffs, New Jersey, 1972.

NOTA DE REFERENCIAS:

Las referencias citadas en (22) Handbook of Research on Teaching, cap. 16, Research on Teaching Reading, 1963,

y que se mencionan en este trabajo, son las siguientes:

1. Alm, R.S. Teaching reading is our business. Engl. J., 1957, 1961.
2. Cole, Luella. The elementary school subjects. New York, Rinehart, 1946.
3. Cook, Luella. Language factors involved in interpretation. Reading Teachers, 1959, 12, 152-157.
4. Dunkel, H.B. Testing the precise use of words. Coll. Engl., 1944, 5, 386-389.
5. Funk, W., & Lewis, N. 30 days to a more powerful vocabulary. New York, Funk, 1954.
6. Hunt, L.C. Can we measure specific factors associated with reading comprehension? Journal of Educational Research, 1953, 51, 161-72.

7. Mc. Cullough, Constance M. What does research reveal about practices in teaching reading? Eng. J., 1957, 46, 475-490. (b)
8. Mc Cullough, Constance M. Context aids in reading. Reading Teacher, 1958, 11, 225-234.
9. Moir, W. A world of words. Engl. J., 1953, 42, 153-155.
10. Pei, M. The story of language. Philadelphia: Lippincott, 1949.
11. Reed, J.C., & Pepper, R.S. Interrelationship of vocabulary, comprehension and rate among disabled readers. J. exp. Educ., 1957, 25, 33-337.
12. Russel, D.H., Children learn to read. (2nd ed.) Boston, Ginn, 1961.
13. Sachs, H.J. The reading method of acquiring vocabulary, J. educ. Res., 1943, 36, 457-464.

Las referencias citadas en (11) Reading-What can be measured?
1969.

y que se mencionan en este trabajo son las siguientes:

1. Goodman, K.S. The psycholinguistic nature of the reading process. In K.S. Goodman (Ed.), *The psycholinguistic nature of the reading process*. Detroit: Wayne State University Press, 1968. Pp. 13-26.
2. Hughes, V.H. A study of the relationships among selected language abilities, *Journal of Educational Research*, 1953, 47, 97-106.
3. Johnson, Mary E. The vocabulary difficulty of content subjects in grade five. *Elementary English*, 1952, 29, 277-80.
4. Karlin, R. *Teaching reading in high school*. New York, Bobs-Merrill, 1964.
5. Wilson, R.M. *Diagnostic and remedial reading*. Columbus, Ohio, Charles E. Merrill Books, 1967.
6. Wyatt, Nita M., & Rideway, R.W. A study of the readability of selected social studies materials. *University of Kansas Bulletin of Education*, 1958, 12, 100-05.