

UNIVERSIDAD DEL VALLE DE GUATEMALA

Facultad de Educación

10
116
1
L67c
1985

INCIDENCIA DE LA CALIDAD DE LAS PRUEBAS DE CIENCIAS NATURALES,
ELABORADAS POR EL PROFESOR DE LA MATERIA, SOBRE EL RENDIMIENTO
DE LOS ESTUDIANTES DE SEGUNDO AÑO BASICO

JESUS MARIA GONZALEZ CAMAÑO

Guatemala

1985

INCIDENCIA DE LA CALIDAD DE LAS PRUEBAS DE CIENCIAS NATURALES,
ELABORADAS POR EL PROFESOR DE LA MATERIA, SOBRE EL RENDIMIENTO
DE LOS ESTUDIANTES DE SEGUNDO AÑO BASICO

UNIVERSIDAD DEL VALLE DE GUATEMALA

Facultad de Educación

INCIDENCIA DE LA CALIDAD DE LAS PRUEBAS DE CIENCIAS NATURALES,
ELABORADAS POR EL PROFESOR DE LA MATERIA, SOBRE EL RENDIMIENTO
DE LOS ESTUDIANTES DE SEGUNDO AÑO BASICO

JESUS MARIA GONZALEZ CAMAÑO

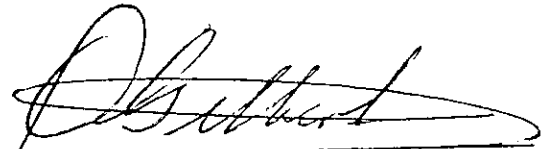
Trabajo de investigación presentado para optar al grado
académico de Maestría en Medición, Evaluación e
Investigación Educativas

Guatemala

1985

Vo. Bo.

(f)



Doctor Otto E. Gilbert A.
Asesor

Fecha de aprobación: 25 de octubre de 1985

"No creo... que los exámenes ineficaces pueden, a la larga, contri
buir a la buena formación, y de ellos infiero mi convicción de que el único camino que un cuerpo de examinadores puede tomar con acier
to es esforzarse continuamente por mejorar los exámenes, para que brinden el máximo apoyo a la buena formación".

R.A.C. Oliver

A mi esposa Carmen C. de González, inspiración constante en mi vida.

A mis hijos: Jesús María, Liseika Iliana y Yelkos Hamed.

A la memoria de mi querida madre María Antonia Camaño de González. (Q.E.P.D.)

A mi padre Jesús María González P.

RECONOCIMIENTO

Deseo expresar mi más profundo agradecimiento al Licenciado Arcadio Madrid por su desinteresada colaboración en el desarrollo de este estudio.

También a los profesores de Ciencias Naturales de segundo año básico de los institutos experimentales de la ciudad de Guatemala.

De igual manera expreso mi gratitud a la Licenciada Yetilú Iunge de Baessa por su paciencia y sabias orientaciones en el trabajo.

Mi agradecimiento especial al Doctor Ottó E. Gilbert A., Director del Programa de Maestría, a quien debo muchas enseñanzas en el campo de la investigación científica.

RESUMEN

Las pruebas, elaboradas por los docentes, que tienen por finalidad medir los resultados del aprendizaje, deben cumplir con una serie de requisitos para alcanzar un nivel de calidad que les dé mayor claridad, va lidez y confiabilidad.

En el presente trabajo se trató de establecer si existe una relación entre la calidad de los instrumentos elaborados por los docentes y el ren dimiento académico de los estudiantes.

Para lograr el objetivo del estudio, se desarrolló una fundamentación teórica del problema en base a los estudios realizados al respecto y se elaboró un instrumento que contiene criterios para la evaluación de los ítemes que integran los test de rendimiento.

Diez pruebas de Ciencias Naturales, de otros tantos institutos experi mentales, fueron evaluadas usando el instrumento de cotejo de criterios, para establecer su nivel de calidad. Estas pruebas fueron aplicadas a una muestra de 1613 estudiantes de segundo año básico.

La hipótesis planteada establece que hay una correlación estadística- mente significativa, a un nivel de $p < \alpha = 0.05$, entre el índice de calidad de los ítemes que integran las pruebas de Ciencias Naturales, elaboradas por los profesores de la materia, y el rendimiento promedio del grupo de estudiantes de segundo año básico al que se le aplicó cada una de estas

pruebas.

Se calculó un coeficiente de correlación del momento del producto de Pearson el que resultó ser de 0.02 por lo que se rechazó la hipótesis al ternativa y se aceptó la nula. Es decir, que no se encontró una relación significativa entre la calidad de los test y el rendimiento de los estu- diantes de segundo año básico, en Ciencias Naturales.

Se recomienda que los docentes elaboren ítemes siguiendo los críte- rios establecidos en el instrumento que se utilizó en este estudio y que en futuras investigaciones se incluya un mayor número de grupos experimen- tales y por lo tanto mayor número de test.

CONTENIDO

	Páginas
RESUMEN	xiii
I. INTRODUCCION	1
A. Antecedentes	1
B. Características	2
C. Justificación	3
D. Objetivos	4
II. FUNDAMENTACION TEORICA	7
A. Evaluación y medición	7
1. Evaluación	7
2. Medición	10
B. Pruebas o test	12
1. Pruebas objetivas	13
2. Pruebas de ensayo	17
3. Calidad de las pruebas	18
C. Investigaciones realizadas	19
III. METODOLOGIA	23
A. Problema	23
B. Hipótesis	23
C. Variable	23
D. Población y muestra	24
E. Instrumentos	24
F. Procedimiento	26

	Páginas
IV. RESULTADOS	29
A. Estadísticas descriptivas	29
B. Estadísticas inferenciales	33
V. DISCUSION DE RESULTADOS	35
A. Prueba de hipótesis	35
B. Conclusiones	35
C. Recomendaciones	37
VI. BIBLIOGRAFIA	39
ANEXOS	
A Instrumento para la evaluación de los test de aprovechamiento elaborados por los docentes	41
B Evaluación de los test de aprovechamiento elaborados por el profesor del curso	45
C Punteo de las pruebas de Ciencias Naturales, elaboradas por los profesores de segundo año básico en los institutos experimentales de la ciudad de Guatemala	65

LISTA DE CUADROS

Cuadro	Páginas
4.1 Estadísticas descriptivas de los test de Ciencias Naturales, elaborados por los profesores de segundo año básico en los institutos experimentales de la ciudad de Guatemala por cada sección (n = 1613)	30
4.2 Estadísticas descriptivas, por instituto, de los test de Ciencias Naturales elaborados por los profesores de segundo año básico	32
4.3 Estadísticas descriptivas correspondientes a la evaluación de los test de aprovechamiento elaborados por el profesor del curso	33
4.4 Punteos de calidad de los tes de Ciencias Naturales y la media aritmética de los punteos obtenidos por los estudiantes	34

I. INTRODUCCION

Los procedimientos de medición y evaluación, son de particular importancia para el desarrollo de los programas educativos, porque son los medios para determinar el grado en que la enseñanza ha sido eficaz.

La persona que construye una prueba se enfrenta a algunos problemas fundamentales. Uno de ellos consiste en determinar los contenidos que se ha de medir en relación a los objetivos propuestos y otro en decidir los procedimientos. La calidad de una prueba educacional depende de la manera como se resuelvan estos problemas.

Las pruebas elaboradas por los docentes revelan a menudo fallas que pueden surgir por interpretación equivocada sobre los objetivos de la enseñanza, defectos técnicos en la construcción o incapacidad para aplicar un tratamiento estadístico a los resultados obtenidos.

Si bien la validez y confiabilidad talvez sean los conceptos técnicos más importantes en la medición Mehrens y Lehmann (1982), no se puede descartar la importancia de la calidad de la elaboración de los ítemes en una prueba de aprovechamiento.

El presente trabajo se relaciona con los conceptos arriba mencionados. A continuación se presentan los antecedentes del problema, sus características, la justificación y los objetivos del estudio.

A. Antecedentes

En toda actividad dedicada a la enseñanza no se puede prescindir de

la evaluación del aprendizaje logrado por los alumnos. Para realizar dicha evaluación es indispensable contar con instrumentos que midan el rendimiento de los discentes en una forma objetiva, precisa, válida y confiable. Sobre los docentes cae la responsabilidad de evaluar a sus alumnos, hecho que implica que los educadores deben interesarse en elaborar pruebas cuyos ítemes sean de alta calidad, es decir que reúnan todas las características técnicas que son indispensables para que los instrumentos de medición sean claros, precisos, objetivos, válidos y confiables.

En la presente investigación se pretende calificar los instrumentos de evaluación, usando para ello los criterios elaborados y aceptados por autoridades reconocidas en la confección de ítemes, en medición y en evaluación educativas.

B. Características

La parte esencial de la evaluación de la enseñanza la constituyen las pruebas efectuadas en el aula, ya sea diariamente, semanalmente, mensualmente o al final de cada unidad.

Según Gronlund (1973), estas pruebas son instrumentos que contienen una serie de tareas que se usan para medir los comportamientos de una persona, o de grupos, en un momento dado.

Las pruebas elaboradas por los docentes más usadas en las escuelas son las de aprovechamiento. Dichas pruebas permiten al docente medir el progreso de sus alumnos en el curso de la enseñanza, decidir la clase de labor complementaria necesaria y evaluar su propia enseñanza.

La medición del esfuerzo educativo se cumple creando situaciones o ítemes de pruebas precisos y adecuados para evaluar en qué grado se ha alcanzado cada objetivo a nivel conceptual de comprensión y de aplicación. Dado que es indispensable que la tarea educativa resulte exitosa, es importante verificar hasta qué nivel se ha logrado el éxito esperado de acuerdo a los objetivos planteados.

C. Justificación

De la misma manera que una prueba de rendimiento educacional debe procurar medir aquello que el proceso de la educación trató de lograr Ebel (1976), en el presente trabajo de investigación se establece la relación entre la calidad de los ítemes de las pruebas de Ciencias Naturales y el rendimiento de los estudiantes en esas pruebas.

Motivó este estudio el hecho de que en Panamá son frecuentes los reclamos de padres de familia y de estudiantes, con relación al deficiente rendimiento, debido en parte, a la mala calidad de las pruebas.

Las tareas de medir y evaluar el rendimiento escolar de los alumnos se han hecho cada día más difíciles y es misión de los educadores tratar de realizarlas adecuadamente. Son numerosas las críticas de parte de los afectados hacia la calidad del producto educativo, pero es poco lo que las instituciones, los educadores y todos los que tienen participación directa o indirecta en el proceso educativo, hacen para crear instrumentos válidos y confiables para evaluar los resultados de la enseñanza.

La calidad de los test elaborados por los docentes es un aspecto de singular importancia en el proceso educacional, de allí el interés de los

mismos en elaborar pruebas de alta calidad.

Los test defectuosos ponen de manifiesto los inconvenientes del sistema.

Karmel (1974:22) dice:

"Los test bien contruidos, que reflejan la instrucción del salón de clase, pueden mejorar el aprendizaje contribuyendo a que el estudiante forme hábitos de estudio y encause su energía intelectual hacia los objetivos deseados".

En este informe se usa indistintamente los términos "pruebas" y "test" los cuales en educación, según "Webster's New Collegiate Dictionary" citado por Karmel (1974:20), son:

"Cualquier serie de preguntas o ejercicios o cualquiera otros medios que sirvan para medir la habilidad, conocimiento, inteligencia, capacidades o aptitudes de un individuo o un grupo".

D. Objetivos

El enfoque principal de la investigación se refiere al estudio de la calidad de las pruebas y cuánto afecta la calidad de éstas a las calificaciones que los estudiantes obtienen en ellas.

Los objetivos específicos del estudio y de este informe son:

1. Evaluar técnicamente la calidad de las pruebas elaboradas por los docentes incluidos en este estudio.
2. Estimular a los educadores para que construyan pruebas con ítemes de alta calidad técnica.

3. Hacer recomendaciones acerca de los criterios que tienen que tomarse en cuenta para la elaboración de pruebas.

... ..
... ..
... ..

II. FUNDAMENTACION TEORICA

Debido a que se ha considerado a la evaluación y medición como factores imprescindibles en la toma eficaz de decisiones en la educación, y que, todas las cualidades exigibles a un buen instrumento de medición dependen de las propiedades de los ítemes que lo constituyen, a continuación se expondrá las principales características de la evaluación, de la medición y de las pruebas.

A. Evaluación y medición

1. Evaluación. Para evaluar es necesario que se fije un propósito.

En educación, algunas veces, se recogen datos que no miden objetivos educativos específicos, sino que sirven para decidir las metas que hay que fijar o los procedimientos que deben emplearse en la enseñanza para alcanzar aquéllos.

Según Smith (1971), el objetivo principal del educador responsable consiste en ayudar al estudiante a adquirir una capacitación cada vez mayor en una serie de aspectos que desarrollan las aptitudes intelectuales, sociales, emocionales y físicas que exige o son apreciadas por la sociedad en donde se desenvuelve el estudiante.

Para Adams (1979), la labor del maestro como evaluador tiene dos aspectos esenciales, ninguno de los cuales puede por sí solo monopolizar su atención. Uno se orienta hacia el programa de enseñanza del grupo, la medida en que la clase, en su conjunto alcanza sus objetivos. El

otro se orienta hacia el estudio de alumnos individuales. Ambos enfoques constituyen aspectos significativos de la educación y son indispensables para una buena enseñanza.

Hay numerosas opiniones sobre los términos evaluación y medición.

La evaluación la han definido Stufflebeam y Cols, citados por Mehrens y Lehmann (1982:15), como:

"El proceso para definir, obtener y proporcionar la información indispensable para juzgar las alternativas en una decisión".

La definición citada, concuerda estrechamente con Ball y Anderson (1981), cuando afirman que la evaluación está orientada hacia la decisión.

La evaluación de los adelantos con los alumnos es uno de los aspectos primordiales del trabajo del profesor. La evaluación permite conformar una buena imagen del lugar en que se encuentra el alumno y de la manera en que está adelantando. Tiene, por tanto, importancia fundamental para la enseñanza eficaz del maestro y a la vez para el aprendizaje del alumno.

La evaluación en educación la han definido Rodríguez Cruz y García González (1982:16) como:

"El proceso completo de señalar los objetivos de un aspecto de la educación y estimar el grado en que tales objetivos se han alcanzado.

El establecimiento del 'estado actual' de la situación enseñanza-aprendizaje.

El proceso que determina si se ha logrado éxito en lo que se propuso hacer".

La evaluación del aprendizaje no se limita a la especificación de

cantidades, es el estudio profundo y dinámico de los cambios de sentimientos, necesidades, actitudes y conducta por parte del aprendiz.

Cuando los objetivos de la enseñanza están bien redactados la evaluación del proceso enseñanza-aprendizaje se convierte en una tarea relativamente sencilla. Dieppa.(1968). Cuando el maestro y los alumnos entienden qué estarán haciendo éstos últimos, al final del proceso del aprendizaje la evaluación consistirá en determinar si han adquirido nuevas formas de conducta.

Alkin (1974:45), al referirse a la evaluación afirma que:

"Recientemente, la evaluación en la educación ha alcanzado una nueva relevancia, y esa relevancia nos ha guiado a desprestigiar y reconsiderar un término al que nos hemos acostumbrado a usar.

Tal reconsideración es bienvenida pues guía a una definición más precisa de lo que expresamos cuando usamos el término evaluación y hacia una precisión más grande para delinear los componentes de la evaluación".

La evaluación puede ser una herramienta valiosa, talvez indispensable, para la responsabilidad de la educación, pero es una parte de ella y no el todo. Womer (1973).

Según Gronlund (1973:6):

"A medida que guiamos a los alumnos hacia el logro de los objetivos de la instrucción, que diagnosticuemos su dificultad para aprender, determinamos su grado de preparación para nuevas experiencias en el aprendizaje, los colocamos en grupos del salón de clases para que se dediquen a actividades especiales, les ayudamos con sus problemas para ajustarse y preparamos informes sobre el adelanto de niños para sus padres, no podemos en modo alguno escapar a la obligación de llegar a juicios evaluativos. Preciso es tomar decisiones y de alguna manera hay que actuar".

Cuando se obtiene medidas (que no sean medidas físicas, como la talla y el peso), y se registran datos sobre una muestra de resultados, la evaluación va más allá de la medición, por cuanto supone la existencia de juicios de valor. Se mide la capacidad de un alumno en varios terrenos. Al interpretar los resultados en términos de patrones correspondientes a su edad, de sus planes educacionales y vocacionales o a partir de cualquier otra base de juicio de valor, ya no se limita a "medir"; se está "evaluando" sus capacidades y progreso.

Los estudios desarrollados por Mehrens y Lehmann (1982:125), dan evidencias claras sobre el valor de la evaluación al señalar que:

"Los educadores, en particular, se han preocupado, se preocupan y continuarán preocupándose por medir y evaluar el avance de sus alumnos, el valor y la relevancia de los programas de estudios, así como la efectividad de su labor de docencia. Esta tarea se torna cada vez más difícil al irse volviendo más complejos los fines de la educación y al aumentar desmesuradamente el número de estudiantes".

Dado que es importante que la tarea educativa resulte exitosa, también es necesario esforzarse para verificar si se logra el éxito deseado. A través de la evaluación el docente puede determinar el éxito o fracaso del aprendizaje en el salón de clase.

2. Medición. La medición ha sido expresada en diferentes formas por autores como Stevens citado por Kerlinger (1982:300), quien opina de la siguiente manera:

"En su sentido más amplio, la medición consiste en asignar numerales a objetos o eventos conforme a ciertas reglas".

La medición es una herramienta que permite solucionar problemas fun-

damentales como: los que cada maestro enfrenta cuando aplica un test o califica a los alumnos. Permite establecer, afinar y aclarar metas prácticas.

Otra definición de la medición que se considera de relevancia es la que asigna Lindeman (1971:9), quien al respecto dice:

"Puede definirse la medición como el acto por el cual se establece una relación de correspondencia entre un conjunto o serie de números y otro de personas u objetos, según ciertas normas establecidas. El conjunto o serie de números depende de la característica que pretende medir el tipo de instrumento de medición empleado. La serie de objetos o individuos queda determinada por los propósitos de la medición".

Las técnicas de medición educacional son útiles para la evaluación de un programa total de enseñanza y para obtener información con respecto al progreso y desarrollo de un alumno determinado.

Es necesario tener presente que la medición no constituye un fin en sí misma, sino que sólo es una parte, aunque importante, del proceso total de la evaluación. Lindeman (1971).

En cualquier campo de actividad, según Thorndike y Hagen (1970), el medir consta de tres pasos comunes:

- a. señalar y definir la cualidad o atributo que se ha de medir
- b. determinar un conjunto de operaciones en virtud de las cuales puede manifestarse y hacerse perceptible; y
- c. establecer un conjunto de procedimientos o de definiciones para traducir las observaciones a enunciados cuantitativos de

grado o de cantidad.

Sin embargo, el proceso de la medición es secundario respecto al de definir los objetivos. Primero debe formularse claramente los propósitos que hay que alcanzar. Después podrán buscarse procedimientos de medición, que nos sirvan de instrumentos para estimar la medida en que tales objetivos se han alcanzado.

B. Las pruebas o test

La evaluación del progreso del alumno es uno de los aspectos más importantes de la función del maestro.

Las pruebas efectuadas en el aula constituyen la médula de la evaluación de la enseñanza. Tienen la ventaja de permitir al maestro medir el progreso en el curso de la instrucción, establecer las bases de la nueva enseñanza, decidir la clase de labor complementaria necesaria y evaluar su propia enseñanza. Howit (1968).

Las pruebas, como instrumentos de evaluación, deben contener ítemes claramente redactados de manera que no confundan al estudiante para lograr el objetivo deseado.

Al respecto Lafourcade (1969:21), dice:

"Si los alumnos superaron realmente los objetivos, pero la prueba, debido a su deficiente construcción, no produjo la información correcta, la línea de ajuste, como es obvio, estará apuntada en otro sentido. Saber que un grupo pequeño o grande no superó importantes objetivos de un tramo del programa, e inculpar dicho resultado al poco interés de los alumnos, a su irónica falta de aplicación al estudio, al "mal de la época", etc. y continuar el desarrollo de la asignatura como si nada malo hubiera ocurrido, no parece ser la mejor salida del problema".

Sin embargo Mager (1973), afirma que las dificultades no provienen de problemas que tengan que ver con el hecho de redactar ítemes sino de los propios objetivos. Cuando se trata de esbozar ítemes de criterio, por ejemplo, éste sólo existe de manera imperfecta o confusa. Otra dificultad según Mager, surge de la tendencia a considerar que el alumno puede realizar cualquier clase de prueba.

Los test elaborados por los profesores les proporcionan éstos indicios valiosos sobre la efectividad de la enseñanza e información acerca del progreso individual. Sin estos datos, es imposible tomar decisiones relativas a la formación de grupos, las reprobaciones y aprobaciones, y la planificación de cursos de recuperación. Los test también ayudan a evaluar al grupo en conjunto.

1. Pruebas objetivas. La contribución más importante en materia de evaluación, fue la iniciativa de Mann en 1845, al introducir un examen escrito uniforme para que todos los estudiantes contestasen las mismas preguntas.

El test objetivo se elaboró para superar alguna desventaja y limitación del test de ensayo que veremos posteriormente.

El procedimiento de calificación del test objetivo queda determinado cuando se redactan los ítemes del test. Es decir, la respuesta correcta se establece siempre antes del examen.

El test objetivo es un examen estructurado, es decir, a cada examinando se le presenta exactamente el mismo problema y debe contestarlo de la manera prescrita. No es necesario que el estudiante organice su respues-

ta como sucede en el formato de ensayo. Las pruebas objetivas piden que el estudiante reconozca, no que recuerde, la respuesta correcta. Debe señalarse, sin embargo, que este test puede elaborarse para apreciar el recuerdo y el uso de la información aprendida previamente, aunque en la mayoría de los casos el test objetivo no explore esta fuente de conocimiento. Karmel (1974).

Como lo señala Karmel (1974), un incremento de reactivos tiende a aumentar la confiabilidad del test. El test objetivo se presta más a ésto que el de ensayo porque cada ítem objetivo es breve y exige menos tiempo para ser contestado. El mayor número de reactivos del test objetivo puede muestrear más temas. Además, las pruebas objetivas facilitan la calificación y esta será la misma, independientemente de quién califique.

El test objetivo incluye varios tipos de reactivos. Los reactivos de la prueba objetiva pueden clasificarse al subdividirlos en los que requieren que el alumno suministre la respuesta y los que requieren que seleccione la respuesta, tomándola de un número dado de alternativas Gronlund, (1973).

Entre los tipos de pruebas pueden citarse:

a. Pruebas objetivas.

1) Verdadero y falso: Es uno de los tipos más populares entre los profesores. Esencialmente es un ítem de elección múltiple en que sólo se presenta una respuesta, debiendo el estudiante juzgar la veracidad o falsedad de la oración.

Tiene la desventaja de que las puntuaciones alcanzadas por los alumnos pueden ser por adivinación. Sin embargo, una de las funciones más útiles del elemento de falso verdadero es medir la capacidad del alumno para distinguir los hechos de las opiniones. Gronlund (1973).

2) Completación: Es uno de los ítemes más fáciles de construir. Tal circunstancia se debe generalmente a que los niveles que se miden con este tipo de prueba son los más bajos según la taxonomía de Bloom.

Una ventaja de los ítemes de completación es que el alumno debe suministrar la contestación, hecho que reduce la posibilidad de que el alumno obtenga la contestación correcta por adivinación. Una desventaja fundamental de éstos es que son inadecuados para medir aprendizajes complejos.

3) Ítemes de apareamiento: Se basa en la medición de información sobre hechos que están asociados. Es un método compacto y eficaz para medir productos del aprendizaje sencillo.

Al respecto Gronlund (1973:188), dice:

"Cada vez que los productos de aprendizaje hacen incapié en la capacidad para identificar la relación entre dos cosas, y siempre que se puedan obtener premisas homogéneas y respuestas en número suficiente, un ejercicio para casar es al parecer el más apropiado".

Dos de las principales ventajas de los ítemes de apareamiento es que hacen posible medir un gran volumen de material sobre hechos ligados entre sí, en un tiempo relativamente corto y otra su facilidad de construcción.

El ejercicio de apareamiento tiene sus desventajas, tales como: su limitación de información sobre hechos que se adquieren mediante la memorización y la dificultad de encontrar material homogéneo que sea significativo desde el punto de vista de los objetivos y productos de aprendizaje.

4) Los ítemes de selección múltiple: Los ítemes de selección múltiple son los más populares de los ítemes objetivos. Se adaptan fácilmente a diferentes áreas de contenidos y objetivos de enseñanzas y se les puede utilizar para medir tanto la memorización como otras habilidades de mayor complejidad. Mehrens y Lehmann (1982).

Según Karmel (1974), los ítemes de opción múltiple pueden usarse para apreciar casi cualquier objetivo, exceptuando la organización de las respuestas y la capacidad para producirlas.

El autor señala nueve áreas que miden eficazmente las preguntas de selección múltiple:

- a) Información
- b) Vocabulario
- c) Hechos aislados
- d) Relaciones de causa y efecto
- e) Comprensión
- f) Discernimiento y análisis crítico

- g) Solución de problemas
- h) Interpretación de datos
- i) Aplicación de principios

Elaborar una pregunta de selección múltiple no es una tarea fácil.* La situación se complica cuando al profesor le corresponde encontrar opciones incorrectas que sean plausibles.

Las ventajas principales de los ítemes de selección múltiple radican en el hecho de que son uno de los elementos de prueba más ampliamente aplicables para medir el aprovechamiento, pueden medir varios productos complejos del aprendizaje y otra de sus ventajas es que son de fácil calificación.

Sin embargo, los ítemes de selección múltiple presentan algunos inconvenientes porque se limitan a medir productos de aprendizaje a nivel verbal. No son apropiados para medir la capacidad para organizar y presentar ideas y además presentan la dificultad para encontrar un número suficiente de opciones no correctas, pero plausibles.

2. Pruebas de ensayo. Las preguntas de ensayo pueden medir muchos aspectos que no se logran con las pruebas objetivas.

Al respecto Gronlund (1973:246), dice:

"Entre estos se incluyen productos como la capacidad de recordar, organizar e integrar ideas; la capacidad para expresarse por escrito y la de suministrar más bien que meramente reconocer interpretaciones y aplicaciones de los datos.

Tales productos requieren de una estructuración de la respuesta

menor que la impuesta por los elementos de prueba objetiva".

Las preguntas de ensayo pueden ser de respuesta restringida o extensiva. La pregunta de respuesta restringida tiende a limitar tanto el contenido, como la forma de la respuesta del alumno. El contenido usualmente se limita, con sólo restringir el alcance del tópico para examinar.

En las preguntas de respuesta extensiva generalmente se deja al alumno en libertad de seleccionar cualquiera información sobre hechos que él considere pertinentes, para organizar la respuesta de acuerdo con su mejor criterio e integrar y evaluar ideas.

La principal ventaja de la pregunta de ensayo es que proporciona una medida directa de los productos complejos del aprendizaje que no pueden medirse por otros medios. Otra ventaja es el énfasis que se da a la integración y la aplicación de las habilidades del pensamiento y de la resolución de problemas. Las preguntas de ensayo también ofrecen facilidad en su construcción.

La limitación principal de los ítemes de ensayo es la baja confiabilidad de la calificación debido a que influye la subjetividad del calificador. A esto podría agregarse otra limitación, que no es de la prueba, sino del profesor y es el tiempo que se requiere para calificar las contestaciones. Estas fallas pueden reducirse a un mínimo mediante una esmerada elaboración de las preguntas y el establecimiento de normas específicas de calificación.

3. Calidad de las pruebas. La mayor parte de los investigadores y autores de libros en evaluación y medición han reconocido a la

confiabilidad y a la validez como aspectos principales para determinar la calidad de una prueba.

Magnusson (1972), por ejemplo, señala que la confiabilidad es la exactitud de la medición, cuantas veces se aplique la prueba, independientemente de lo que se esté midiendo y lo que se ha querido medir y que la validez es el grado hasta donde una prueba es capaz de lograr ciertos objetivos.

Ebel (1976), señala, además de la confiabilidad, otros factores importantes que es necesario tomar en cuenta al juzgar la calidad de una prueba tales como: pertinencia, equilibrio, eficacia, objetividad, especificidad, dificultad, discriminación, ecuanimidad y tiempo complementario.

Esta serie de factores que contribuyen a la calidad de una prueba pueden ser analizados estudiando la prueba misma, a partir de un examen de las especificaciones que se emplearon para elaborarla, los puntajes de la prueba y de los ítemes.

Toda prueba requiere un planeamiento previo en el cual el constructor de la prueba debe tener conocimientos sobre evaluación y construcción de pruebas educativas. Redactarla cuidadosamente cumpliendo con todos los requisitos y especificaciones de la materia, poniendo especial cuidado en que las instrucciones sean claras y concisas.

C. Investigaciones realizadas

Son numerosos los escritores como Jiménez (1979), Botacio, et al (1975), Sejas et al (1978); que han abordado el tema sobre medición y evaluación escolar, como también sobre el análisis de ítemes, pero son traba

jos que sólo se aproximan al problema que se plantea para ser investigado a través del presente estudio.

En 1975, Botacio, Castellanos y González desarrollaron el trabajo: "Diseño de un modelo para analizar pruebas objetivas de aula en forma manual" trabajo de investigación presentado para optar al grado académico de Maestría en Medición, Evaluación e Investigación Educativas de la Universidad del Valle de Guatemala. En este estudio se desarrolla un manual para analizar y evaluar estadísticamente las pruebas objetivas de aula. Para ello, los autores construyeron dos formatos para analizar ítemes. Utilizaron diversas fórmulas estadísticas tales como: índice de facilidad, índice de discriminación, media total de la distribución de puntajes, correlación punto biserial de cada ítem con el punteo total del test y confiabilidad. Una vez terminado el manual se elaboró una encuesta para obtener la opinión y evaluación de los docentes.

Francisco Jiménez Villalobos, en 1979 desarrolló el trabajo: "Comparación de los resultados obtenidos al aplicar pruebas de ensayo y selección a un mismo contenido matemático", presentado para optar al grado académico de Maestría en Medición, Evaluación e Investigación Educativas de la Universidad del Valle de Guatemala. El autor utilizó una muestra de 15 alumnos del quinto año de bachillerato de la Escuela Experimental de Aplicación Mirón Muñoz, de la Universidad Rafael Landívar de Guatemala, clasificándolos en dos grupos: grupo A con 38 y grupo B con 37 alumnos. Hizo uso de dos tipos de test: uno de ensayo y otro de selección múltiple de un mismo contenido matemático.

Aplicó la prueba de ensayo a un grupo de 38 alumnos y la de selección

múltiple a otro grupo de 37. Utilizó la comparación de medias por medio de una prueba de t de Student.

El estudio no arrojó diferencias estadísticamente significativas al medir el rendimiento con pruebas de ensayo y con pruebas objetivas en matemáticas.

El grupo de investigadores formado por Sejas, Parada y Vargas, en 1978 trabajó sobre: "Número de alternativas de ítemes de escogencia y sus efectos sobre los resultados en pruebas de rendimiento", tesis para optar al grado académico de Maestría en Medición, Evaluación e Investigación Educativas de la Universidad del Valle de Guatemala.

Por medio de un muestreo aleatorio simple seleccionaron 20 escuelas con 581 alumnos.

Los autores utilizaron dos pruebas: una constituida por 40 ítemes de opción múltiple de 4 alternativas y la otra constituida por 40 ítemes de opción múltiple de 5 alternativas. Aplicaron en primer lugar, una prueba piloto con el fin de probar los instrumentos a emplearse en la experimentación definitiva. Seleccionan una muestra de 12 escuelas con 341 alumnos cuyas características fueron similares, con la única diferencia que a los de un grupo se les aplicó una prueba que contenía ítemes con 5 alternativas y la otra contenía ítemes de 4.

César Fernández Rojas, en 1977 investigó sobre: "El proceso de evaluación aplicado según los pasos del método científico en el Instituto Nacional para Varones "Adrián Zapata" y el Instituto Experimental "Dr. José Matos Pacheco", de la ciudad de Guatemala", tesis para optar al

grado académico de Maestría en Medición, Evaluación e Investigación Educativas de la Universidad del Valle de Guatemala. En la investigación realizada, el autor utilizó una población de 58 profesores y una muestra de 30. Ambas fueron escogidas de los institutos sometidos a estudio.

El autor usó como instrumento un "Cuestionario para profesores". Después de haberse aplicado el cuestionario elaborado para el estudio a los profesores, no se encontraron resultados estadísticamente significativos.

III. METODOLOGIA

Los aspectos metodológicos utilizados en este estudio se exponen a continuación comenzando por el enunciado del problema que se investigará.

A. Problema

¿Incide la calidad de las pruebas de Ciencias Naturales, elaboradas por el profesor de la materia, sobre el rendimiento de los estudiantes de segundo año básico?

B. Hipótesis

H_1 : Existe una correlación estadísticamente significativa, a un nivel de $p\alpha = 0.05$, entre el índice de calidad de los ítemes que integran las pruebas de Ciencias Naturales elaboradas por el profesor de la materia y el rendimiento promedio del grupo de estudiantes de segundo año básico en cada una de estas pruebas.

H_0 : No existe una correlación estadísticamente significativa, a un nivel de $p\alpha = 0.05$, entre el índice de calidad de los ítemes que integran las pruebas de Ciencias Naturales elaboradas por el profesor de la materia y el rendimiento promedio del grupo de estudiantes de segundo año básico en cada una de estas pruebas.

C. Variables

Las variables involucradas en las hipótesis que se someterán a prueba

son las siguientes:

1. Variable independiente. Calidad de las pruebas: Puntuación proporcional total que se le asigna a cada prueba como resultado del análisis de cada uno de sus ítems, con base en los criterios que reúne.
2. Variable dependiente: Rendimiento académico: Media aritmética de los punteos que obtienen los estudiantes en el test.

D. Población y muestra

Para la realización de este estudio se obtuvieron 23 pruebas que fueron aplicadas a 3199 estudiantes de segundo año de los institutos experimentales de la ciudad de Guatemala.

En forma aleatoria se seleccionaron diez pruebas, escogiéndose una por cada instituto experimental con una muestra de 1613 estudiantes que presentaron los test.

E. Instrumentos

1. Pruebas o test elaborados por los docentes: Se analizaron diez pruebas que contenían secciones como selección múltiple, ensayo, apareamiento y completación. No se encontró pruebas de falso y verdadero. Algunas contenían una sección, mientras que otras tenían la combinación de dos o tres.

2. Lista de cotejo de criterios: Para evaluar las pruebas de rendimiento escritas, mencionadas con anterioridad, se usó una lista de cotejo de criterios, elaborada por el autor de esta investigación, en base a juicio de expertos. Este instrumento contiene los criterios obser



vables en los test. La lista de selección múltiple se divide en criterios para ítemes en general en opciones y en enunciados. Las listas de ensayo y completación contienen los criterios generales, para los ítemes, y se preparó, además, una lista de criterios para la sección de apareamiento.

Los expertos en medición y evaluación escolar consideran una gran cantidad de criterios para la elaboración de ítemes, pero en el instrumento que se utilizó en este trabajo sólo se anotaron aquellos que pueden ser observables y aplicables a pruebas elaboradas por los profesores.

3. Cuadro para la calificación de los ítemes: Para cada sección de una prueba se usó un cuadro. En el caso de la de selección múltiple se dividió en tres partes y en dos partes para el resto de las secciones.

Cada ítem de una sección se evaluó según el criterio obteniéndose así una proporción para cada uno de ellos. La suma de las proporciones arroja un punteo total para cada sección.

Las partes de cada sección recibieron un peso que fue determinado multiplicando el número de ítemes por los criterios, este producto se dividió entre el número total de criterios. Se obtuvo así porcentajes que, sumados, dieron la calidad del test. Cuando la prueba contenía más de una sección se sumaron los resultados de todas las secciones, dividiéndose entre el número de esas, obteniéndose así la calidad de todo el test.

a. Fórmulas para la evaluación de los test.

1) Selección múltiple

$$\left(\frac{A}{N}\right) PA + \left(\frac{B}{N}\right) PB + \left(\frac{C}{N}\right) PC = C.T.S.M.$$

En donde A es el producto de la cantidad de ítemes por los criterios de las opciones; N el número total de criterios; PA la proporción total de la sección; B es el producto de los ítemes por los criterios de los enunciados, PB la proporción total de la sección; C la suma de los criterios; PC la proporción total de la sección, y C.T.S.M. es la calidad total de la sección de selección múltiple.

2) Ensayo y completación

$$\left(\frac{A}{N}\right) PA + \left(\frac{C}{N}\right) PC = C.T.S.$$

En donde A es el producto de los ítemes por los criterios; N la suma total de criterios, PA la proporción total de la sección; C la suma de los criterios; PC la proporción de esta sección y C.T.S. la calidad total de la sección evaluada.

Para la sección de apareamiento sólo se utilizó la proporción de los criterios.

F. Procedimiento

A través del Ministerio de Educación de Guatemala se tramitaron los permisos correspondientes para el acceso a los institutos experimentales del ciclo básico en la ciudad, con el fin de obtener las pruebas de Cien

cías Naturales que los profesores elaboraron y aplicaron a los estudiantes de segundo año.

Se analizaron los ítemes de cada prueba en base a la lista de cotejo de criterios.

La calidad de las pruebas elaboradas por los profesores se determinó cotejando las características de cada ítem de la prueba con los criterios de calidad establecidos para los ítemes de cada tipo, obteniéndose así un puntaje proporcional. Este puntaje se correlacionó con la media del número de respuestas correctas que obtuvieron los estudiantes en el test, utilizándose el coeficiente de correlación del momento del producto de Pearson.



IV. RESULTADOS

En este capítulo se presentan los resultados obtenidos a través del procesamiento de datos, tanto descriptivos como inferenciales.

A. Estadísticas descriptivas

Se calculó medias, desviaciones estándares, amplitudes observadas y posibles de cada test presentado, clasificado por instituto y sección. (Vea Cuadro 4.1).

Al analizar el Cuadro 4.1 se observa que hubo institutos que aplicaron la misma prueba a varias secciones y, como consecuencia, a un mayor número de estudiantes. En otros institutos un sólo profesor atendía a todas las secciones, pero aplicaba diferentes pruebas a cada sección, de las cuales se seleccionó al azar sólo una de éstas.

Las pruebas difieren en su amplitud posible, sin embargo para efectos de las estadísticas inferenciales, fueron transformadas en proporciones sobre 100 puntos.

En términos generales la media proporcional del rendimiento de los grupos, es de 51.34.

Se puede observar que hubo desviaciones desde 1.06 hasta 22.00.

Cuadro 4.1

Estadísticas descriptivas de los test de Ciencias Naturales elaborados por los profesores de segundo año básico en los institutos experimentales de la ciudad de Guatemala, por cada sección (n = 1613)

Ins- titu- to	Estadís- tico Sección	n	\bar{X}	S	A-O	A-P
A	A	37	7.80	1.32	4.0-9.5	0.0-10.0
	B	34	7.13	1.32	4.0-9.0	0.0-10.0
	C	37	7.38	1.42	4.0-9.5	0.0-10.0
B	A	31	61.71	11.19	42-80	0.0-100
	B	34	75.89	13.96	40-100	0.0-100
	C	35	45.57	13.43	29-74	0.0-100
	D	32	44.91	10.62	25-68	0.0-100
	E	28	48.71	10.39	30-67	0.0-100
	F	30	45.63	11.97	25-61	0.0-100
	G	28	52.50	9.32	36-68	0.0-100
	H	31	47.94	10.94	30-61	0.0-100
C	A	55	3.63	5.06	4.0-7.2	0.0-8.0
	B	53	4.41	1.34	1.0-7.0	0.0-8.0
	C	56	5.25	1.20	4.0-8.0	0.0-8.0
	D	56	3.75	1.06	1.0-7.0	0.0-8.0
D	A	49	24.86	6.99	10-38	0.0-40
	B	46	24.59	7.13	10-36	0.0-40
	C	54	22.85	7.80	8-38	0.0-40
	D	48	24.08	7.39	8-36	0.0-40
	E	46	22.26	7.25	6-32	0.0-40
	F	45	27.96	7.68	12-38	0.0-40

Ins- titu to	Estadís- tico Sección	n	\bar{X}	S	A-O	A-P
E	A	42	21.12	5.74	0.0-31	0.0-50
	B	43	21.28	4.54	16-34	0.0-50
	C	41	20.97	6.27	8-32	0.0-50
	D	42	19.76	4.57	10-29	0.0-50
	E	42	21.52	6.54	5-33	0.0-50
	F	42	23.31	4.30	13-32	0.0-50
F	A	68	20.63	5.52	6-29	0.0-30
G	A	53	33.80	22.00	0-100	0.0-100
	B	53	40.17	20.89	10-85	0.0-100
H	A	36	33.50	9.07	4-47	0.0-50
	B	35	32.46	8.34	16-49	0.0-50
	C	36	27.44	11.44	9-49	0.0-50
I	A	40	69.08	14.64	22-91	0.0-100
	B	40	70.10	15.71	20-99	0.0-100
	C	40	72.65	19.70	13-99	0.0-100
J	A	49	56.14	15.59	35-98	0.0-100
	B	46	55.67	15.25	20-90	0.0-100

En el Cuadro 4.2 se calculó medias, desviaciones estándares, amplitudes observadas y posibles de los resultados obtenidos en los test aplicados a los diez institutos ($n = 1613$).

Cuadro 4.2

Estadísticas descriptivas, por instituto, de los test de Ciencias Naturales elaborados por los profesores de segundo año básico

Prueba	Estadístico Instituto	n	\bar{X}	S	A-O	A-P
1	A	108	7.44	1.39	4.0-9.5	0.0-10.0
2	B	249	53.07	15.59	25-100	0.0-100.0
3	C	220	4.41	1.40	1.0-8.0	0.0-40.0
4	D	288	24.94	10.46	6.0-38.0	0.0-40.0
5	E	252	21.33	5.58	0.0-34	0.0-50.0
6	F	68	20.63	5.52	6.0-29.0	0.0-30.0
7	G	106	37.0	21.68	0.0-100.0	0.0-100.0
8	H	107	31.12	10.04	4.0-49.0	0.0-50.0
9	I	120	70.61	16.89	20.0-99.0	0.0-100.0
10	J	95	55.92	15.43	20.0-98.0	0.0-100.0

Se puede observar en el Cuadro 4.2 que n varía de 68 a 288 estudiantes. Esto ocurre debido a las características de cada instituto y al número de alumnos que fueron sometidos a los test.

Las medias varían entre 4.41 y 70.61; las desviaciones típicas varían entre 1.39 y 21.68; las amplitudes observadas varían entre 0.0 y 100, lo mismo sucede con las posibles.

En base a los criterios establecidos en el instrumento para la eva-

luación de los test de aprovechamiento, se estableció la calidad de cada prueba mediante un punteo proporcional.

Cuadro 4.3

Estadísticas descriptivas correspondientes a la evaluación de los test de aprovechamiento elaborados por el profesor del curso

Test	1	2	3	4	5	6	7	8	9	10
Punteo	46	80	80	89	58	74	47	72	48	73

La media aritmética de los diez índices que aparecen en el Cuadro 4.3 es de 66.6 y su desviación típica es de 15.75.

Se puede inferir que la prueba con más alto puntaje resultó ser la número cuatro, que contenía 20 preguntas de selección múltiple. La más baja fue la número uno, que contenía ensayo y completación.

B. Estadísticas inferenciales

Mediante el coeficiente de correlación del momento del producto de Pearson, se calculó las estadísticas inferenciales para el contraste de la hipótesis nula planteada. Para calcularlo se utilizó como variable independiente (X) los punteos de calidad de cada uno de los test de Ciencias Naturales y como variable dependiente (Y) la media aritmética de los punteos obtenidos por los estudiantes que resolvieron ese test.

Cuadro 4.4

Punteos de calidad de los test de Ciencias Naturales
y la media aritmética de los punteos obtenidos
por los estudiantes

Test	<u>X</u> Calidad de las pruebas	<u>Y</u> Media de punteos en el test
1	46	74
2	80	53
3	80	55
4	89	62
5	58	43
6	74	69
7	47	37
8	72	62
9	48	71
10	73	56

Con base en los datos del Cuadro 4.4, se calculó un coeficiente de correlación del momento del producto de Pearson, el cual resultó ser de $r_{xy} = 0.02$.

El valor crítico del coeficiente de correlación para 8 grados de libertad y un nivel de probabilidad de error alfa igual a 0.05 es de 0.63, por lo tanto, el coeficiente encontrado no es estadísticamente significativo.

V. DISCUSION DE RESULTADOS

Este capítulo incluye la discusión de la prueba de hipótesis, las conclusiones y recomendaciones.

A. Prueba de hipótesis

La hipótesis nula planteada era que: no existe una correlación estadísticamente significativa, a un nivel de $p\alpha = 0.05$, entre el índice de calidad de los ítemes que integran las pruebas de Ciencias Naturales, elaboradas por el profesor de la materia y el rendimiento promedio del grupo de estudiantes de segundo año, en cada una de estas pruebas.

Esta hipótesis se puso a prueba calculando la correlación del momento del producto de Pearson entre la variable independiente y la variable dependiente.

El valor del coeficiente de correlación calculado entre el índice de calidad de la prueba de Ciencias Naturales y la media aritmética de los puntajes obtenidos por los estudiantes en la prueba correspondiente, fue de .02 mientras que el valor crítico para 8 grados de libertad y un nivel de $p\alpha = 0.05$ es de 0.63. Por consiguiente se tuvo que aceptar la hipótesis nula y rechazar la alternativa, pues el valor encontrado (0.02) es menor que el crítico (0.63).

B. Conclusiones

1. Los profesores de Ciencias Naturales de segundo año básico de los

institutos experimentales usaron la mayor parte de tipos usuales de evaluación escrita, a excepción del falso-verdadero.

2. Los ítemes incluidos en las pruebas elaboradas por los profesores de Ciencias Naturales de segundo año básico, exploraron menos del 10% de los contenidos enseñados.

3. El promedio de los resultados obtenidos en los punteos de los alumnos que presentaron las pruebas de Ciencias Naturales fue del 51% (véase Cuadro 4.2), promedio que el autor considera que es de un rendimiento deficiente.

4. Los índices de calidad de los test de aprovechamiento (véase Cuadro 4.3) sugiere que éstos son deficientes ($\bar{X} = 67\%$).

5. La calidad del instrumento, se vio disminuida debido a que al elaborarlo para la evaluación de la calidad de los test, se tuvo que eliminar una serie de criterios que usualmente se utilizan, pero que en este caso, no eran observables o aplicables a muchas pruebas.

6. Las limitaciones de tiempo y la cantidad de ítemes por evaluar, obligó a que sólo diez pruebas fueran estudiadas. Se considera que si se hubiera incluido un número mayor de test, probablemente la correlación hubiera aumentado.

7. La conclusión general es que en esta investigación se encontró que no existe correlación estadísticamente significativa, a un nivel de $p_{\alpha} = 0.05$, entre la calidad de los ítemes de una prueba y el rendimiento de los estudiantes en la misma.

C. Recomendaciones

Los resultados obtenidos en esta investigación, permiten hacer las siguientes recomendaciones:

1. Que los profesores elaboren ítemes siguiendo los criterios establecidos en el instrumento que se utilizó en esta investigación para evaluar la calidad de los test a fin de que la medición sea válida y confiable.
2. Que se realice otros estudios de este tipo, que incluyan mayor número de test, en los que el investigador sea quien evalúe las pruebas.
3. Que en futuras investigaciones no se incluya pruebas que hayan sido elaboradas por los profesores con el propósito de "quedar bien" con los investigadores y que no fueron elaboradas en la forma reglamentaria para evaluar a los estudiantes.
4. Que los profesores elaboren pruebas que exploren el 100% de los contenidos enseñados.
5. Que se hace evidente, según la experiencia de este estudio, la necesidad de que los docentes reciban cursos y seminarios para actualizarlos en la elaboración de pruebas objetivas de rendimiento.



VI. BIBLIOGRAFIA

- Adams G., Sachs. Medición y evaluación; en educación, psicología y "guidance". Barcelona, Herder.
1970
- Alkin, Marvin C. et al. Evaluation and decision making. Los Angeles. Center for the study of Evaluation: U.C.L.A.
1974
- Ball, Samuel y Anderson, Scarvia B. Temas profesionales en la evaluación de los programas de educación/adiestramiento. Traducción de Ella Molina. Universidad del Valle de Guatemala.
1981
- Botacio T., Faustino; et al. Diseño de un modelo para analizar pruebas objetivas de aula en forma manual. Tesis para optar al grado académico de Maestría en Medición, Evaluación e Investigación Educativas de la Universidad del Valle de Guatemala.
1975
- Dieppa, Jorge J. Técnicas para la redacción de los objetivos de la enseñanza. Universidad del Valle de Guatemala. Organización de los Estados Americanos.
1968
- Ebel, Robert L. Fundamentos de la medición educacional. Traducción de Ramón Alcalde. Buenos Aires, Guadalupe.
1976
- Fernández R., César. El proceso de evaluación aplicado según los pasos del método científico en el Instituto Nacional para Varones "Adrián Zapata y el Instituto Experimental "José Matos Pacheco", de la ciudad de Guatemala. Tesis para optar al grado académico de Maestría en Medición, Evaluación e Investigación Educativas de la Universidad del Valle de Guatemala.
1977
- Gronlund, Norman E. Medición y evaluación en la enseñanza. Traducción de Salvador Sumano. México, Pax-México.
1973
- Howitt, Lillian O. Pruebas prácticas en el aula. Traducción de Carlos Gerhard. México, UTEHA.
1979
- Jiménez, V., Francisco. Comparación de los resultados obtenidos al aplicar pruebas de ensayo y selección a un mismo contenido matemático. Tesis para optar al grado académico de Maestría en Medición, Evaluación e Investigación Educativas de la Universidad del Valle de Guatemala.
1979
- Karmel, Luis. Medición y evaluación escolar. México, Trillas.
1974

- Kerlinger, Fred N. Investigación del comportamiento. México, Interamericana, 1982
- Lafourcade, Pedro. Evaluación de los aprendizajes. Buenos Aires, Kapelusz, 1969
- Lindeman, Richard H. Tratado de medición educacional. Traducción de Elvira Rissech de Wiñas. Buenos Aires, Paidós, 1974
- Mager, Robert F. Medición del intento educativo. Buenos Aires, Guadalupe, 1975
- Magnusson, David. Teoría de los test. Traducción de Javier Aguilar. México, Trillas, 1972
- Mehrens, William A. y Lehmann, Irvin J. Medición y evaluación en la educación y en la psicología. México, Continental, 1982
- Rodríguez C., Héctor M. Evaluación en el aula. México, Trillas, 1982
- Sejas L., Elizabeth et al. Número de alternativas de ítemes de escogencia y sus efectos sobre los resultados en pruebas de rendimiento. Tesis para optar al grado académico de Maestría en Medición, Evaluación e Investigación Educativas de la Universidad del Valle de Guatemala, 1978
- Smith, Robert M. y Cols. El maestro y el diagnóstico de las dificultades escolares. Buenos Aires, Paidós, 1971
- Tavella, Nicolás M. Análisis de los ítemes en la construcción de instrumentos psicométricos. México, Trillas, 1978
- Thorndike, Robert L. y Hagen, Elizabeth. Test y técnicas de medición en psicología y educación. México, Trillas, 1970
- Womer, Frank. ¿Qué es la evaluación educativa? Traducción de Elba de Katelaar. Universidad del Valle de Guatemala, Instituto de Investigaciones Educativas, 1973
- Wood, Dorothy A. Elaboración de test; desarrollo e interpretación de los test de aprovechamiento. Traducción de Rogelio Díaz Guerrero. México, Trillas, 1976

ANEXO A

Instrumento para la evaluación de los test de aprovechamiento elaborados por los docentes

Lista de cotejo según juicio de expertos.

1. Selección múltiple

1.1 Enunciado

- 1.1.1 ¿El enunciado incluye elementos comunes a todas las opciones. (Artículos, nombres y palabras)?
- 1.1.2 ¿Evitó información innecesaria?
- 1.1.3 ¿Existe congruencia gramatical entre el enunciado y las opciones?
- 1.1.4 ¿Evitó que el enunciado proporcione pautas para encontrar la clave?
- 1.1.5 ¿El enunciado del ítem se presentó como pregunta directa, como oración incompleta o como una orden?
- 1.1.6 ¿Existe congruencia de contenido entre el enunciado y las opciones?

1.2 Opciones

- 1.2.1 ¿Las opciones son todas plausibles?
- 1.2.2 ¿Las opciones son aproximadamente iguales en longitud?
- 1.2.3 ¿Las opciones están colocadas en columna?
- 1.2.4 ¿Evitó usar opciones excluyentes entre sí?
- 1.2.5 ¿Evitó el uso de opciones superpuestas, es decir que una de ellas incluya a la otra?
- 1.2.6 ¿Hay solamente una respuesta correcta o la mejor?
- 1.2.7 ¿Evitó pautas irrelevantes para la respuesta correcta?

1.3 Prueba en general

- 1.3.1 ¿La prueba tiene instrucciones claras, suficientes y precisas?
- 1.3.2 ¿No se abusó de opciones tales como: "todas las anteriores" o "ninguna de las anteriores"?
- 1.3.3 ¿El enunciado y sus opciones están escritos en la misma página?
- 1.3.4 ¿Las preguntas aparecen en orden de complejidad (de lo sencillo a lo complejo)?
- 1.3.5 ¿La clave está colocada en cada una de las posiciones de las opciones, en un número aproximadamente igual de veces, pero sin seguir un patrón determinado?

2. Ensayo

2.1 Ítemes

- 2.1.1 ¿Evalúa aspectos de la materia que sólo pueden ser medidos con preguntas de ensayo?
- 2.1.2 ¿Estableció un marco de referencia específica dentro del cual el estudiante desarrolló la pregunta?
- 2.1.3 ¿Está la pregunta elaborada en forma clara sobre lo que se desea saber con respecto al aprovechamiento de los estudiantes?
- 2.1.4 ¿Definió y precisó el campo de conocimiento dentro del cual se debería enmarcar la pregunta?
- 2.1.5 ¿Eliminó las preguntas optativas?
- 2.1.6 ¿Evitó escribir las preguntas con frases como: "qué piensa usted", "en su opinión", "escriba todo lo que sabe"?

2.2 Prueba en general

- 2.2.1 ¿Aparecen escritas instrucciones claras y precisas sobre lo que el estudiante debe hacer?
- 2.2.2 ¿Indicó en forma escrita el tiempo suficiente en que el estudiante deberá contestar la pregunta?
- 2.2.3 ¿Aparecen consignados con anticipación qué factores se deben de considerar en la evaluación (ortografía, redacción, contenido)?

3. Apareamiento

- 3.1 ¿La prueba tiene instrucciones claras y precisas?
- 3.2 ¿Las premisas son suficientemente homogéneas?
- 3.3 ¿Las premisas y las respuestas están en una sola página?
- 3.4 ¿Cada columna tiene diferente número de elementos?
- 3.5 ¿Las premisas o enunciados son más largas que las respuestas?
- 3.6 ¿Las respuestas se presentan en un orden sistemático (numérico, cronológico, alfabético, etc.)?
- 3.7 ¿La serie incluye entre 5 y 12 premisas o enunciados?
- 3.8 ¿Las opciones o respuestas se presentan formando columna al lado izquierdo?
- 3.9 ¿Las premisas se presentan formando columnas al lado derecho?
- 3.10 ¿Los enunciados tienen un espacio o paréntesis a la derecha de las premisas para las respuestas, formando columnas?
- 3.11 ¿El título de la sección de la prueba aparece con su puntaje a evaluar?

4. Respuesta corta o completación

4.1 Ítemes

- 4.1.1 ¿Se formuló el ítem de manera que la respuesta requerida sean palabras o frases leves?
- 4.1.2 ¿El ítem es específico, claro, preciso y sin ambigüedades?
- 4.1.3 ¿Evitó la mutilación de la oración?
- 4.1.4 ¿Los espacios para las respuestas se agruparon a la derecha del ítem?

4.2 Prueba

- 4.2.1 ¿Dio instrucciones escritas claras y precisas sobre lo que el alumno debe hacer?
- 4.2.2 ¿El título de esta sección de la prueba y su puntaje están claramente escritos?

4.2.3 ¿Los espacios para las respuestas son iguales y adecuados en longitud?

4.2.4 ¿Los espacios para las respuestas se agruparon en columnas?

5. Falso y verdadero

5.1 Ítemes

5.1.1 ¿Evitó construir el ítem empleando términos negativos?

5.1.2 ¿Evitó las afirmaciones triviales (sin importancia)?

5.1.3 ¿Prescindió de oraciones largas y complejas?

5.1.4 ¿El ítem está redactado en forma específica, evitando generalidades?

5.1.5 ¿Evitó el uso de determinantes específicos como "todos", "puede ser", "a veces", "nunca", "generalmente", etc.?

5.1.6 ¿El ítem tiene una extensión aproximadamente igual a los demás?

5.1.7 ¿Evitó usar el ítem con doble respuesta (en parte verdadero y en parte falso)?

5.1.8 ¿Fue descartada la pregunta con trampas?

5.1.9 ¿Evitó incluir dos ideas en un mismo ítem?

5.1.10 ¿Evitó que la verdad o falsedad sea obvia?

5.1.11 ¿Rehusó construir el ítem en base a opiniones del examinando, excepto para identificar las opiniones de un autor o autoridad de la materia?

5.2 Prueba

5.2.1 ¿La prueba tiene instrucciones claras y precisas de lo que el estudiante debe hacer para dar las respuestas?

5.2.2 ¿La prueba contiene especificaciones de la parte que se va a evaluar y su puntaje?

ANEXO B

Evaluación de los test de aprovechamiento
elaborados por el profesor del curso

Test: No. 1

Tipo de test: Ensayo

Instituto: A

Ítemes Criterios		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	T	P	
		Criterios para ítemes	1	0	0	0	0	0												0
2	0		0	0	0	0												0		
3	1		1	1	1	1												1		
4	0		0	0	0	0												0		
5	1		1	1	1	1												1		
6	1		1	1	1	1												1		
7																				
8																				
9																				
10																				
11																				
91%	T	.5	.5	.5	.5	.5												.5	46	
C. Prueba	1	0																0		
	2	0																0		
	3	0																0		
	4																			
9%	T	0																0	0	
Total de la Sección:																		46		

Valor Total del Test No. 1: 46

Test: No. 1

Tipo de test: Completación

Instituto: A

Ítemes Criterios		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	T	P	
Criterios para ítemes	1	1	1	1	1	0												.8		
	2	1	1	0	0	0												.4		
	3	1	1	1	1	1												1		
	4	0	0	0	0	0												0		
	5																			
	6																			
	7																			
	8																			
	9																			
	10																			
	11																			
83%	T	.75	.75	.5	.5	.25												.55	46	
C. Prueba	1	0																0		
	2	0																0		
	3	0																0		
	4	0																0		
17%	T	0															0	0		
Total de la Sección:																			46	

Test No. 2

Tipo de test: Selección Múltiple

Instituto: B

Ítemes Criterios		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	T	P	
Opciones	1	1	0	0	0	0	1	1	0	1	0	1	0	0	0	0	0	.33		
	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1		
	3	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1		
	4	1	1	1	1	0	1	1	1	1	1	0	1	1	1	1	1	.87		
	5	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	.93		
	6	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1		
	7	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1		
52%	T	1	.71	.86	.86	.71	1	1	.86	1	.86	.86	.86	.86	.86	.86	.86	.88	46	
Enunciados	1	1	1	1	0	0	1	0	1	0	1	1	1	1	0	1		.67		
	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1		1		
	3	1	0	0	0	0	0	0	0	0	0	1	0	0	1	1		.27		
	4	1	1	1	0	1	0	1	1	1	1	1	1	1	1	1		.87		
	5	0	0	0	0	0	0	0	0	0	0	1	0	1	1	1		.27		
	6	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1		.94		
45%	T	.83	.67	.5	.33	.5	.5	.5	.67	.5	.67	1	.67	.83	.83	1		.67	30	
Prueba	1	1																	1	
	2	1																	1	
	3	1																	1	
	4	0																	0	
	5	1																	1	
3%	T	.8																	.8	2
Total de la sección																				78

Valor total del test No. 2: 80

Test: No. 2

Tipo de test: Completación

Instituto: B

Ítemes Criterios		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	T	P
Criterios para ítemes	1	1	1	0	1	1	1	1	1	1	1							.9	
	2	0	1	0	0	1	1	1	0	0	1							.5	
	3	1	1	1	1	1	1	1	1	1	1							1	
	4	1	1	1	1	1	1	1	1	1	1							1	
	5																		
	6																		
	7																		
	8																		
	9																		
	10																		
	11																		
91%	T	.75	1	.5	.75	1	1	1	.75	.75	1							.85	77
C. Prueba	1	1																1	
	2	0																0	
	3	0																0	
	4	1																1	
9%	T	.5																.5	4
Total de la Sección:																			81

Test: No. 3

Tipo de test: Selección Múltiple

Instituto: C

Items Criterios		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	T	P
		Opciones	1	0	0	1	0	1											
2	1		1	1	1	1													1
3	1		1	1	1	1													1
4	1		1	1	1	1													1
5	1		0	1	1	1													.8
6	1		0	1	1	1													.8
7	1		1	1	1	1													1
50%	T	.86	.57	1	.86	1													.86 43
Enunciados	1	0	0	1	0	0													.2
	2	1	1	1	1	1													1
	3	0	1	1	0	1													.6
	4	0	1	1	1	1													.8
	5	0	1	1	0	0													.4
	6	0	1	0	1	1													.6
43%	T	.17	.83	.83	.5	.67													.60 26
Prueba	1	0																	0
	2	1																	1
	3	1																	1
	4	0																	0
	5	1																	1
7%	T	.6																	.6 4
Total de la sección																			73

Test: No. 3

Tipo de test: Completación

Instituto: C

Ítemes Criterios		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	T	P
Criterios para ítemes	1	1	1	1	1	1	1	1										1	
	2	1	1	0	1	1	1	1										.86	
	3	1	1	1	1	1	1	1										11	
	4	1	1	1	1	1	1	1										1	
	5																		
	6																		
	7																		
	8																		
	9																		
	10																		
	11																		
88%	T	1	1	.75	1	1	1	1										.96	84
C. Prueba	1	0																0	
	2	0																0	
	3	1																1	
	4	0																0	
12%	T	.25																.25	3
Total de la Sección:																			87

Test: No. 4

Tipo de test: Selección Múltiple

Instituto: D

Ítemes Criterios		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	T	P		
Opciones	1	0	1	1	0	0	0	1	1	1	0	1		1	0	0	1	1	1	1	1	1	.65		
	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1		
	3	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1		
	4	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	.95	
	5	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	.95	
	6	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
	7	0	1	0	0	0	1	1	1	1	0	1	1	1	0	0	1	1	1	1	1	1	1	.65	
53%	T	.71	.71	.86	.71	.71	.86	1	1	1	.71	1	1	1	.71	.71	1	1	1	1	1	1	.88	47	
Enunciados	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1		
	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1		
	3	1	1	1	1	0	0	1	1	1	1	0	0	1	0	1	1	1	1	1	1	1	1	.75	
	4	1	0	1	1	1	1	1	1	0	0	0	1	1	1	0	1	1	1	1	1	1	1	.75	
	5	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
	6	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
45%	T	1	.83	1	1	.83	.83	1	1	.83	.83	.83	.83	.83	1	.67	1	1	1	1	1	1	1	.92	41
Prueba	1	1																					1		
	2	1																					1		
	3	1																					1		
	4	0																					0		
	5	0																					0		
2%	T	.6																					.6	1	
Total de la sección																						89			

Valor Total del Test: 89

Test No. 5

Tipo de test: Selección Múltiple

Instituto: E

Items Criterios		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	T	P
		Opciones	1	0	0														
2	1		1															1	
3	1		1															1	
4	1		1															1	
5	1		1															1	
6	1		1															1	
7	0		0															0	
45%	T	.71	.71															.71	32
Enunciados	1	0	1															.5	
	2	1	1															1	
	3	0	0															0	
	4	1	1															1	
	5	0	1															.5	
	6	1	1															1	
39%	T	.5	.83															.66	26
Prueba	1	0																0	
	2	1																1	
	3	1																1	
	4	1																1	
	5	0																0	
16%	T	.6																.6	10
Total de la sección																			68

Valor total del test: 58

Test: No. 5

Tipo de test: Completación

Instituto: E

Ítemes Criterios		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	T	P
		Criterios para ítemes	1	1	1	0	1	0	1	0	1	1	0	1	1	1	0	0	0
2	1		1	0	0	0	0	0	1	0	0	1	1	0	0	1	0	.38	
3	1		1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
4	0		0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	.06	
5																			
6																			
7																			
8																			
9																			
10																			
11																			
94%	T	.75	.75	.25	.5	.25	.5	.25	1	.5	.25	.75	.75	.5	.25	.75	.25	.52	49
C. Prueba	1	0																0	
	2	0																6	
	3	0																0	
	4	0																0	
6%	T	0															0	0	
Total de la Sección:																			49

Test: No. 6

Tipo de test: Completación

Instituto: F

Ítemes Criterios		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	T	P
		Criterios para ítemes	1	1	1	1	1	1											
2	1		0	0	0	1												.4	
3	1		1	1	1	1												1	
4	1		1	1	1	1												1	
5																			
6																			
7																			
8																			
9																			
10																			
11																			
83%	T	1	.75	.75	.75	1												.85	71
C. Prueba	1	1																1	
	2	0																0	
	3	0																0	
	4	0																0	
17%	T	.25																.25	4
Total de la Sección:																			75

Valor total del test: 74

Test: No. 6

Tipo de test: Apareamiento

Instituto: F

Ítemes Criterios		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	T	P
		Criterios para ítemes	1	1															
2	1																	1	
3	1																	1	
4	1																	1	
5	1																	1	
6	0																	0	
7	1																	1	
8	1																	1	
9	1																	1	
10	0																	0	
11	0																	0	
100%	T	.73																.73	73
C. Prueba	1																		
	2																		
	3																		
	4																		
	T																		
Total de la Sección:																			73

Test: No. 7

Tipo de test: Ensayo

Instituto: G

Ítemes Criterios		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	T	P	
		Criterios para ítemes	1	0	0	0														0
2	0		0	0														0		
3	1		1	1														1		
4	0		0	0														0		
5	1		1	1														1		
6	1		1	1														1		
7																				
8																				
9																				
10																				
11																				
86%	T	.5	.5	.5														.5	43	
C. Prueba	1	0																0		
	2	0																0		
	3	0																0		
	4																			
14%	T	0																0	0	
Total de la Sección:																			43	

Valor total del test No. 7: 47

Test: No. 7

Tipo de test: Completación

Instituto: G

Ítemes Criterios		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	T	P
Criterios para ítemes	1	1	1	1	1	1	1	1										1	
	2	0	0	0	0	0	0	0										0	
	3	1	1	1	1	1	1	1										1	
	4	0	0	0	0	0	0	0										0	
	5																		
	6																		
	7																		
	8																		
	9																		
	10																		
	11																		
88%	T	.5	.5	.5	.5	.5	.5	.5										.5	44
C. Prueba	1	0																0	
	2	0																0	
	3	0																0	
	4	0																0	
12%	T	0																0	0
Total de la Sección:																			44

Test: No. 8

Tipo de test: Selección Múltiple

Instituto: H

Ítemes Criterios		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	T	P	
		Opciones	1	0	0	1	0	1	0											.33
2	1		1	1	1	1	1											1		
3	0		0	0	0	0	0											0		
4	1		1	0	1	1	1												.83	
5	0		1	1	1	1	0												.67	
6	1		1	1	1	1	1												1	
7	0		0	0	1	1	1												.5	
51%	T	.43	.57	.57	.71	.86	.57											.62	32	
Enunciados	1	1	1	1	1	1	1												1	
	2	1	1	1	1	1	1												1	
	3	0	0	1	1	1	1												.67	
	4	0	0	0	1	0	1												.33	
	5	0	0	0	1	0	0												0	
	6	1	1	1	1	1	1												1	
43%	T	.5	.5	.67	1	.67	.71											.67	29	
Prueba	1	1																	1	
	2	1																	1	
	3	1																	1	
	4	0																	0	
	5	0																	0	
6%	T	.6																.6	4	
Total de la sección																			65	

Valor total del test No. 8: 72

Test: No. 8

Tipo de test: Completación

Instituto: H

Ítemes Criterios		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	T	P	
		Criterios para ítemes	1	1	1	1	1	1	1	1	1	1	1							1
2	0		0	0	0	0	0	0	0	1	0							.1		
3	1		1	1	1	1	1	1	1	1	1							1		
4	1		1	1	1	1	1	1	1	1	1							1		
5																				
6																				
7																				
8																				
9																				
10																				
11																				
41%	T	.75	.75	.75	.75	.75	.75	.75	.75	1	.75							.78	71	
C. Prueba	1	1																1		
	2	0																0		
	3	1																1		
	4	1																1		
9%	T	.75																.75	78	
Total de la Sección:																				

Test: No. 9

Tipo de test: Ensayo

Instituto: I

Ítemes Criterios		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	T	P	
Criterios para ítemes	1	0																0		
	2	0																0		
	3	1																1		
	4	0																0		
	5	1																1		
	6	1																1		
	7																			
	8																			
	9																			
	10																			
	11																			
67%	T	.5																.5	34	
C. Prueba	1	0																0		
	2	0																0		
	3	0																0		
	4																			
33%	T	0																0	0	
Total de la Sección:																			34	

Valor total del test No. 9: 48

Test: No. 9

Tipo de test: Completación

Instituto: I

Ítemes Criterios		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	T	P
Criterios para ítemes	1	1	1	1	1	1	1	1	1	1								1	
	2	1	1	0	1	0	1	1	1	1								.78	
	3	1	1	1	1	1	1	1	1	1								1	
	4	0	0	0	0	0	0	0	0	0								0	
	5																		
	6																		
	7																		
	8																		
	9																		
	10																		
	11																		
90%	T	.75	.75	.75	.75	.75	.75	.75	.75	.75								.69	62
C. Prueba	1	0																0	
	2	0																0	
	3	0																0	
	4	0																0	
10%	T	0															0	0	
Total de la Sección:																		62	

Test: No. 10

Tipo de test: Selección Múltiple

Instituto: J

Ítemes Criterios		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	T	P	
Opciones	1	1	1	0	1	1	1	1	0	1	1	0	1	0	0	0	0	.56		
	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1		
	3	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1		
	4	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1		
	5	1	1	1	1	1	1	1	1	1	1	0	1	1	1	0	1	.88		
	6	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	.94		
	7	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1		
53%	T	1	1	.86	1	1	1	1	.86	.86	1	.71	1	.86	.86	.71	.86	.91	48	
Enunciados	1	0	0	0	1	0	1	0	1	0	1	0	0	0	0	0	0	.25		
	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1		
	3	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1		
	4	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	.12		
	5	0	1	1	1	1	1	1	0	1	0	1	1	0	0	0	0	.51		
	6	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1		
45%	T	.5	.67	.67	.83	.67	1	.67	.83	.67	.67	.67	.67	.5	.5	.5	.5	.66	30	
Prueba.	1	1																	1	
	2	1																	1	
	3	1																	1	
	4	0																	0	
	5	1																	1	
2%	T	.8																	.8	2
Total de la sección																			80	

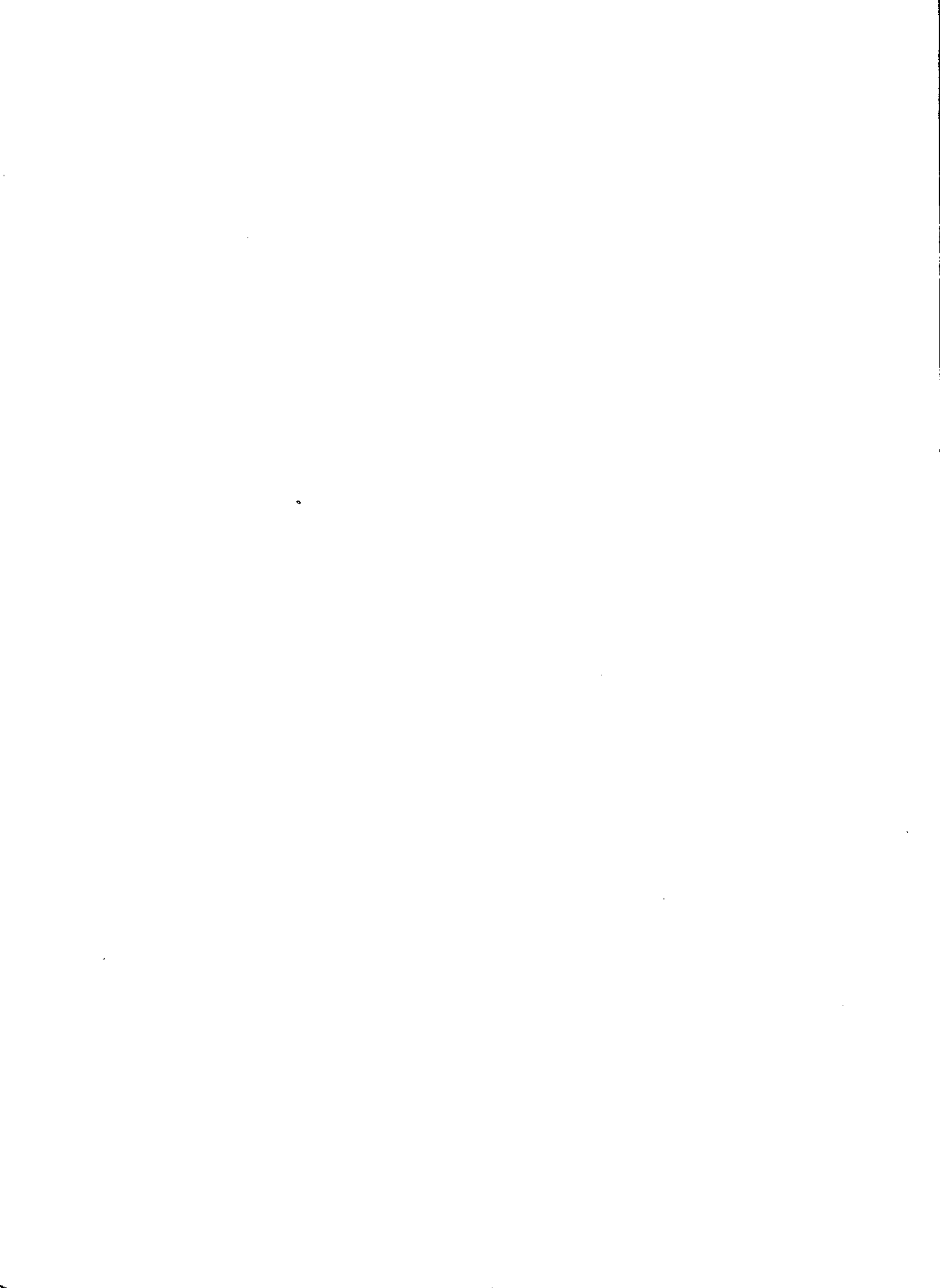
Valor total de la prueba No. 10: 73

Test: No. 10

Tipo de test: Ensayo

Instituto: J

Ítemes Criterios		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	T	P
Criterios para ítemes	1	0	0															0	
	2	1	1															1	
	3	1	1															1	
	4	1	1															1	
	5	1	1															1	
	6	1	1															1	
	7																		
	8																		
	9																		
	10																		
	11																		
	T	.83	.83															.83	66
C. Prueba	1	0																0	
	2	0																0	
	3	0																0	
	4																		
	T	0																0	0
Total de la Sección:																			66



ANEXO C

Punteo de las pruebas de Ciencias Naturales, elaboradas por los profesores de segundo año básico, en los institutos experimentales de la ciudad de Guatemala

Prueba No. 1

Instituto: A

Valor de la prueba: 10 puntos

SECCION	A				C				D			
	10	9	6	7	6	9	6.5	8	8	9	8	6
	9	8	6.5	10	6.5	8	8	7	9	6.5	9.5	7.5
	9	8	7.5	9	4.5	6.5	5.5	8.5	7.5	8	9.5	7.5
	7	8.5	8	10	5	6	4	6.5	6.5	5.5	6	5.5
	9.5	4	6	8.5	8.5	8	7.5		6	8	6	6
	8	7.5	9	7	8.5	7.5	8.5		7.5	5	8	4
	8.5	8	8	8	7.5	6	7.5		8	8	8	9
	8	8	6		9	8	8		9	5	8.5	
	7	6.5	6		6.5	7.5	8.5		7	9	9	
	9	7.5	6		7.5	4.5	8		6	9	7.5	
N	37				34				37			
\bar{X}	7.80				7.13				7.38			
S	1.32				1.32				1.12			

N = 108

TOTALES: \bar{X} = 7.44

S = 1.39

Prueba No. 2 Instituto: B

Valor de la Prueba: 100 puntos

SECCION	A		B		C		D		E		F		G		H	
	65	54	90	80	33	44	25	39	52	55	30	60	61	67	52	49
	76	66	75	100	65	31	38	36	38	55	45	31	55	56	48	34
	60	63	90	80	39	49	28	48	53	57	56	35	58	52	56	44
	50	48	90	55	48	47	50	50	30	59	51	26	49	56	51	41
	38	73	90	60	29	51	51	63	50	45	41	33	55	42	44	72
	63	71	60	60	32	35	44	64	57	48	56	60	48	36	60	31
	63	50	55	70	31	36	48	50	32	51	57	33	62	38	40	47
	53	48	80	80	47	30	50	38	35	33	61	48	39	67	60	64
	70	75	60	65	32	58	39	39	42	45	52	51	61	45	61	61
	54	68	80	65	49	64	56	61	56	29	50	48	53	42	49	31
	54	53	60	70	40	67	32	48	67		48	50	68		35	42
	42	66	90	80	52	63	29	48	56		35	59	46		55	32
	80	68	80	65	32	63	44	29	51		26		44		61	42
	69		90	80	31	74	68	39	59		70		48		48	
	80		90	95	28	34	51		67		39		67		60	
	61		40	85	31	62	36		53		54		50		45	
	52		80		55	62	45		51		25		61		41	
	80		90		51		51		38		39		44		30	
N =	31		34		35		32		28		30		28		31	
\bar{X}	61.71		75.89		45.57		44.91		48.71		45.63		52.50		47.94	
S	11.19		13.96		13.43		10.62		10.39		11.97		9.32		10.94	

N = 249

T O T A L E S : \bar{X} = 53.07

S = 15.59

Prueba No. 3 Instituto: C

Valor de la Prueba: 8 puntos

SECCION	A				B				C				D			
	3.6	2.5	6.0	4.8	3	5	4	4	5	5	5	4	5	6	4	1
	5.6	3.6	2.5	4.8	6	5	6	4	8	6	5	6	4	4	4	3
	0.6	6.0	4.0	6.4	6	4	3	6	5	7	5	4	4	6	4	3
	4.8	3.5	4.0	6	3.6	6	5	3	4	5	6	4	3	4	4	2
	0.4	5.6	2.5	5.2	5	3	6	5	5	4	5	6	2	3	4	4
	3.6	5.6	4.8	6	6	4	2	4	7	4	6	6	5	4	2	4
	5.2	5.6	2	4.0	4	4	1	2	4	4	4	4	4	4	4	3
	0.4	5.5	4.0	4.8	7	4	4	3	5	6	5	4	3	5	4	3
	0.4	4.8	5.6	5.6	6	4	4	4	4	7	4	6	4	4	4	4
	4.4	4.8	4.8	4.0	6	4	4	4	4	5	4	6	4	6	4	4
	5.2	7.2	2.5	4.4	6	3	6	4	5	6	4	6	7	4	3	4
	4.0	4	2.8	5.6	6	1.6	3		6	6	4	4	4	3	4	3
	5.6	4.5	3.2	4.0	5	6	4		6	6	7	8	2	4	2	3
	4.0	2	4.0		6	5.6	4		8	5	6	6	4	4	3	3
N	55				53				56				56			
\bar{X}	3.63				4.41				5.25				3.75			
S	5.06				1.34				1.20				1.06			

N = 220

T O T A L E S : \bar{X} = 4.41

S = 1.40

Prueba No. 4 Instituto: D

Valor de la Prueba: 40 puntos

SECCION	A	B	C	D	E	F
	14 26 20 28	20 28 20 34	32 20 32 20	24 28 22 16	22 30 16 14	34 18 30 24
	14 28 30 22	22 32 28 10	16 30 16 22	26 14 28 10	18 26 6 8	26 14 12 16
	30 38 30 22	30 22 30 34	34 14 28 20	28 26 26 22	18 16 32 26	28 18 36 34
	20 24 30 26	32 32 36 30	36 12 22 8	30 22 26 34	32 30 20 16	22 14 36
	36 36 20 24	18 20 22	24 24 28 16	24 22 34 12	18 22 26	30 32 34
	30 34 22 20	28 26 20	18 22 20 36	28 20 26 16	28 22 26	24 40 28
	24 30 26 22	24 18 20	24 26 30 26	30 34 22	32 20 22	20 40 32
	22 28 34	16 22 22	16 22 28 30	18 18 22	30 8 26	20 32 38
	24 10 18	20 24 32	18 20 12 10	36 12 24	30 16 6	32 34 26
	26 40 20	24 28 14	16 28 38 26	8 34 14	26 14 24	32 38 36
	10 14 26	30 36 30	24 18 26 10	22 32 34	14 26 12	24 28 34
	24 28 20	26 24 24	20 38 30 26	36 20 28	26 32 24	36 30 14
	22 34 32	22 30 22	32 16 12	34 30 26	28 26 30	38 24 24
	28 20 12	20 30 32	34 20 8	24 8 26	30 20 30	28 30 16
N	49	46	54	48	46	45
\bar{X}	24.86	24.59	22.85	24.08	22.26	27.95
S	6.99	7.13	7.80	7.39	7.25	7.68

N = 288

T O T A L E S: \bar{X} = 24.94

S = 10.46

Prueba No. 5 Instituto: E

Valor de la Prueba: 50 puntos

SECCION	C	D	E	F	G	H
	27 28 18 19	20 17 16 28	19 16 25 14	17 21 24 27	23 27 28 27	22 18 22 31
	16 26 19 23	21 25 18 12	27 27 26 17	18 10 15 17	24 8 29 30	27 32 22 22
	21 18 20 19	17 27 21 18	10 25 20 16	21 19 23 14	14 26 27 19	27 25 26 18
	16 31 30 20	34 22 17 19	18 21 19 23	26 24 19 20	19 24 16 31	24 17 20 20
	23 19 11 19	22 20 21 21	30 21 8 11	17 13 14 13	22 19 5 16	25 22 25 14
	16 26 20 28	18 29 23 23	16 23 26	21 18 25 15	24 28 21 9	24 25 18 16
	25 14 31	24 15 15 22	31 21 21	18 14 29	24 33 31	25 23 28
	22 16 16	31 21 15	27 25 11	19 17 17	5 19 18	13 25 26
	21 23 25	27 16 22	21 28 17	19 25 18	25 27 20	23 18 30
	26 18 0	22 20 20	27 32 20	19 18 28	20 17 21	27 22 24
	24 20 19	27 21 17	10 17 25	23 16 27	25 18 23	22 22 29
	29 22 23	23 23 25	24 13 32	23 23 26	21 21 20	24 27 23
N	42	43	41	42	42	42
\bar{X}	21.12	21.28	20.97	19.76	21.52	23.31
S	5.74	4.54	6.27	4.57	6.54	4.30

T O T A L E S :

N = 252
 \bar{X} = 21.33
S = 5.58

Prueba No. 6 Instituto: F

Valor de la Prueba: 30 puntos

SECCION	A					
	20	18	18	6	22	28
	28	28	22	18	14	22
	24	18	18	18	24	12
	18	28	28	10	24	24
	22	22	28	18	26	18
	24	20	18	10	20	14
	18	28	20	14	18	24
	17	10	16	28	28	26
	26	18	22	22	28	
	18	28	26	16	29	
	12	28	26	12	18	
	16	16	21	24	20	
N	68					
\bar{X}	20.63					
S	5.52					

$$N = 68$$

$$T O T A L E S : \bar{X} = 20.63$$

$$S = 5.52$$

Prueba No. 7

Instituto: G

Valor de la Prueba: 100 puntos

SECCION	G	H
	P U N T E O	P U N T E O
	41 10 53 32 55 20 55 27 50 48 35	80 70 25 40 15 10 60 15 35 30 43
	14 23 33 23 57 22 24 60 28 20 5	20 24 48 17 18 33 95 20 63 20 45
	28 5 18 19 30 25 10 63 10 30 35	46 48 80 74 37 58 45 45 20 35 10
	30 38 60 100 25 32 47 14 25 30	48 52 20 32 30 43 45 70 10 60
	9 34 75 30 88 0 18 35 5 90	33 58 30 25 33 19 46 51 15 85
N	53	53
\bar{X}	33.8	40.17
S	22.0	20.89

N = 106

T O T A L E S : \bar{X} = 37

S = 21.68

Prueba No. 8

Instituto: H

Valor de la Prueba: 50 puntos

SECCION	A	B	C
	P U N T E O	P U N T E O	P U N T E O
	27 38 20 4 45 41 25 22	24 50 43 26 38 28 16	27 37 18 34 12 26 48 27
	43 26 35 47 45 35 36	28 47 26 33 32 21 24	39 9 24 28 24 40 10
	45 35 34 25 44 40 36	22 32 42 28 36 24 28	22 24 12 21 21 33 20
	32 37 41 25 29 34 22	26 45 49 39 24 33 36	25 48 42 26 28 11 46
	29 23 35 45 30 39 37	28 32 39 32 35 42 28	32 45 46 27 31 16 14
N	36	35	36
\bar{X}	33.50	32.46	27.44
S	9.07	8.34	11.44

N = 107

T O T A L E S :

 $\bar{X} = 31.12$

S = 10.04

Prueba No. 9

Instituto: I

Valor de la Prueba: 100 puntos

SECCION	A	B	C
	83 80 74 72	85 74 63 51	93 56 30 13
	69 74 74	85 60 62	68 62 72
	84 60 68	73 66 71	99 72 92
	86 80 62	68 60 88	60 84 95
	85 68 79	52 57 60	96 85 69
	72 75 55	85 91 70	95 55 56
	78 99 22	70 73 91	92 61 70
	57 58 72	68 60 37	76 92 86
	71 43 52	99 59 79	91 84 52
	79 50 70	91 64 64	43 59 75
	69 91 69	76 71 76	62 85 86
	74 69 81	77 85 20	91 50 47
	70 39 50	49 82 92	96 72 84
N	40	40	40
\bar{X}	69.08	70.10	72.65
S	14.64	15.71	19.70

N = 120

T O T A L E S : \bar{X} = 70.61

S = 16.89

Prueba No. 10

Instituto: J

Valor de la Prueba: 100 puntos

SECCION	A	B
	45 40 35 45 55	65 45 70 70 75
	60 55 80 45 45	30 40 75 65 60
	80 75 45 65 75	80 65 50 45 70
	63 40 53 55 55	73 50 55 60 60
	55 55 45 60 63	55 70 55 35 30
	55 55 30 35 45	55 80 60 40 47
	50 40 70 45 80	30 65 68 55
	98 55 35 53 42	45 90 20 50
	80 70 45 40 35	63 60 55 35
	78 73 70 83	60 50 30 55
N	49	46
\bar{X}	56.14	55.67
S	15.59	15.25

$$N = 95$$

$$T O T A L E S : \bar{X} = 55.92$$

$$S = 15.43$$

