

**AFACIMERA**

Asociación de Facultades de Ciencias Médicas  
de la República Argentina

---

**EVALUACIÓN  
EDUCACIONAL**

**Construcción y aplicación  
de exámenes**

---

**Módulo N° 2**

**- 2002 -**

## AFACIMERA

Dr. Acosta, Abraham	Decano de la Facultad de Ciencias de la Salud de la Universidad Adventista del Plata
Dr. Adamo, José	Decano de la Facultad de Ciencias de la Salud de la Universidad Nacional de Salta
Dr. Beveraggi, Enrique	Decano de la Escuela de Medicina del Hospital Italiano
Dr. Cerini, Carlos	Decano del Instituto Universitario de Ciencias Biomédicas de la Fundación Favaloro
Dr. Cherjovsky, Roberto	Decano de la Facultad de Medicina de la Universidad Abierta Interamericana
Dra. Chiara, Raquel Madis	Decano de la Facultad de Ciencias médicas de la Universidad Nacional de Rosario
Dr. De Michele, Daniel	Decano de la Facultad de Ciencias de la Salud de la Universidad Nacional de Entre Ríos
Dr. Deza, Horacio	Decano de la Facultad de Medicina de la Universidad Nacional de Tucumán
Dr. Escobar, José	Decano de la Facultad de Ciencias de la Salud de la Universidad Nacional de Formosa
Dr. Feldstein, Carlos	Decano de la Facultad de Medicina de la Fundación Barceló
Prof. Dr. Ferreira, Luis	Decano de la Facultad de Ciencias de la Salud de la Universidad de Ciencias empresariales y sociales -UCES-
Dr. Hernández, Rodolfo	Decano de la Facultad de Ciencias Médicas de la Universidad Nacional de La Plata
Dr. Liotta, Domingo	Decano de la Facultad de Medicina de la Universidad de Morón
Dr. Lizarraga, Adolfo	Decano de la Facultad de Medicina de la Universidad del Salvador
Dra. Magnelli, Norma	Decano de la Facultad de Ciencias Médicas de la Universidad Nacional de Cuyo
Dr. Mazzaro, Baltazar	Decano de la Facultad de Medicina de la Universidad Nacional del Nordeste
Dr. Panisse, Horacio	Decano de la Facultad de Medicina de la Universidad Nacional del Comahue
Dr. Saracho Cornet, Pedro	Decano de la Facultad de Medicina de la Universidad Nacional de Córdoba
Dr. Sonis, Abraham	Decano de la Facultad de Medicina de la Universidad Maimónides
Dr. Trakal, Esteban	Decano de la Facultad de Medicina de la Universidad Católica de Córdoba
Dr. Turin, Mario	Decano del Instituto Universitario Hospital CEMIC
Dr. Villar, Marcelo	Decano de la Facultad de Ciencias Biomédicas de la Universidad Austral

---

### Evaluación Educativa

Dirección: Lic. Amanda Galli

Secretaría Administrativa: Hugo Garavaglia

Diagramación: Martín Sesini

Impresión: Enrique Titakis

ISBN:

Todos los derechos reservados. Ninguna parte de esta publicación puede ser reproducida o transmitida en ninguna forma y por ningún medio electrónico, de fotocopia, grabación u otros, sin permiso previo por escrito de la dirección de este Curso. Las opiniones que se expresan en este libro son las de los autores y no necesariamente de AFACIMERA.

### Evaluación Educativa

Austria 2277 P.B. "B" (C1425EGK) - Buenos Aires, Argentina.

Tel/Fax: (54-11) 4805-5965

E-mail: grupoag@interserver.com.ar

# Índice

## Construcción y aplicación de exámenes

Introducción .....	7
Calidad de los instrumentos de evaluación. Confiabilidad, objetividad, validez. ....	12
Objetivos .....	21
Esquema de contenidos .....	22
¿Qué evaluar? .....	23
Exámenes orales .....	26
Examen oral individual .....	26
Examen oral en grupo: Coloquio .....	28
¿Cómo mejorar los exámenes orales tradicionales? .....	31
a) Antes del día del examen .....	31
b) Durante el examen .....	33
c) Después del examen .....	34
Exámenes escritos .....	36
Construcción de la prueba .....	38
Tabla de especificaciones .....	38
Redacción de preguntas .....	41
<i>Preguntas de respuesta abierta, de ensayo, de desarrollo</i> .....	42
<i>Preguntas de respuesta restringida</i> .....	44
• <i>Recordación simple:</i> .....	44
• <i>Completamiento:</i> .....	45
<i>Preguntas estructuradas</i> .....	47
• <i>Alternativas constantes:</i> .....	47
• <i>Apareamiento o Correspondencia:</i> .....	49
• <i>Ordenamiento:</i> .....	51
• <i>Selección múltiple:</i> .....	51
• <i>Preguntas de Menú largo</i> .....	59
Ejercicio 1 .....	61
Criterios de corrección .....	68
Diagramación de la prueba .....	71
Aplicación y corrección de la prueba. Tomar el examen .....	73
Criterios de aprobación. Nivel de exigencia .....	76
• <i>Criterio absoluto</i> .....	76
• <i>Criterio relativo</i> .....	77

Análisis del ítem .....	78
• <i>Índice de dificultad</i> .....	79
• <i>Índice de discriminación</i> .....	80
• <i>Distribución de las respuestas</i> .....	81
Ejercicio 2 .....	84
Adjudicación de notas. Sistema de calificaciones .....	87
Ejercicio 3 .....	91
Comunicación de los resultados. Brindar feedback al estudiante .....	93
Algunas sugerencias para volver al feedback más útil y eficiente .....	94
Devolución grupal: .....	95
Devolución individual: .....	96
Toma de decisiones .....	98
Clave de Respuestas .....	101
Anexo 1, Las apreciaciones personales del profesor .....	109
Anexo 2, Puntuación corregida. Calificación de la prueba objetiva .....	115
Anexo 3, Ejemplos de hojas de respuesta .....	117
Anexo 4, Ejemplo de instrucciones generales de la prueba .....	119
Anexo 5, Propuesta para la adjudicación de notas .....	121

# **Módulo N° 2**

## **Construcción y aplicación de exámenes**

**Lic. Claudia Castro\***  
**Lic. Amanda Galli\*\***

\* Licenciada en Ciencias de la Educación. UBA.  
Consultora en Educación Médica.

\*\* Licenciada en Ciencias de la Educación. UBA.  
Sociedad Argentina de Investigación y  
Desarrollo de la Educación Médica (SAIDEM)



## Introducción



*"Modificar un programa o unas técnicas de enseñanza sin cambiar el sistema de evaluación tiene todas las probabilidades de no conducir a nada.*

*Cambiar el sistema de evaluación sin modificar el programa de enseñanza ha tenido una mayor repercusión sobre la naturaleza y la calidad de aprendizaje que si se hubiera modificado el programa sin tocar los exámenes".*

G. E. Miller

Ud. ya sabe que el proceso de evaluación proporciona información, evidencias, que permiten tomar mejores decisiones pedagógicas. Los exámenes son los instrumentos que utilizan los docentes para recoger información acerca de los aprendizajes de sus alumnos y luego, en base a esa información, tomar decisiones: ofrecerles un recuperatorio, promoverlos o aplazarlos. Los estudiantes consideran con frecuencia que los exámenes son capciosos, confusos, muy subjetivos y arbitrarios.



*"-¡Pobre, así que estás en capilla! ¿Y por qué perdés el tiempo aquí?"*

*-Es mejor, ya no podemos estudiar más -dijo Juan-.*

*Entrenamiento liviano la víspera de la pelea. Clara va a aprobar, es seguro. Yo, no sé. A veces te preguntan cada cosa... De verdad dijo -Stella-... ("Quién descubrió el principio de la flotabilidad de los cuerpos?")*

*-Hay que apelar a los trucos -dijo Andrés-. A pregunta estúpida respuesta absurda. Los tres tipos de la mesa se quedan pensando si les estás tomando el pelo o si realmente tenés algo en el mate. Como pasa el tiempo, se aburren y te aprueban.*

*-A vos te parece fácil -dijo Juan- pero un examen final no es macana. Sobre todo para mí, que pago las culpas del autodidactismo más bien desorganizado, porque habría que ser idiota para creer que se aprende algo en las santas aulas argentinas.*

*-Clara debe saber -dijo Stella-. Seguro que estudió muchísimo.*

*-Todo el programa. -dijo Clara suspirando-. Pero es como un pozo: miro al fondo y no me veo más que a mí, con la cara lavada".*

Julio Cortazar <sup>1</sup>

<sup>1</sup> Cortazar Julio (1986), Fragmento del libro El examen. Buenos Aires, Sudamericana. p. 22.

Tomar exámenes no es sólo una práctica escolar. Muchas instituciones del sector salud diseñan y aplican exámenes y pruebas para seleccionar residentes, certificar especialistas, otorgar becas, etc.

En el área de las Ciencias de la Salud, los exámenes tanto de alumnos como de graduados implican una enorme responsabilidad social ya que a través de ellos se reconoce y se autoriza una determinada práctica profesional que se vincula directamente con la salud de las personas. Toda la sociedad confía en que este "control de calidad" se haya hecho con toda seriedad en las instituciones correspondientes.

**Este módulo tiene como propósito brindarle la información necesaria para que Ud. pueda diseñar instrumentos para la evaluación del aprendizaje en el dominio cognoscitivo.**

La construcción de exámenes que brinden resultados objetivos, válidos y confiables es todo un desafío.

.....  
***Antes de comenzar con el desarrollo de los temas de este módulo le proponemos que reflexione sobre las siguientes situaciones:***

▶ *"Yo no planteo una autoevaluación final a mis alumnos porque estoy seguro que todos se pondrían 8, 9 ó 10. Creo que no están maduros para estas cosas".*  
¿Qué concepto de autoevaluación tiene este profesor?

.....  
.....

▶ *"Llega el fin del cuatrimestre y aún no tengo notas. Tengo que tomar un examen".*  
¿Qué concepto de evaluación tiene este profesor?

.....  
.....

▶ *"Hoy tengo clase y no sé que dar. Voy a tomarles un parcialito".*  
¿Qué concepto de evaluación tiene este profesor?

.....  
.....



► "Los alumnos no están motivados para aprender, no leen, no estudian nada. Yo lo resuelvo fácilmente, cada dos por tres, les tomo una pequeña evaluación y les dijo que forma parte de la nota de concepto."

¿Qué concepto de evaluación tiene este profesor?

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

► "A Martínez le tomo yo porque fue alumno mío".

¿Qué concepto de evaluación tiene este profesor?

¿Qué podría pasar con Martínez si es examinado por otro docente?

Enumere algunas de las razones por las que un profesor quiere evaluar a sus propios estudiantes y estos, a su vez quieren ser evaluados por sus docentes.

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

► ¿Por qué cree que casi todos los estudiantes se oponen a los exámenes "centralizados" (examen único para todos los alumnos que han cursado una asignatura en diferentes sedes o unidades académicas)

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

► "Yo tengo pocos residentes, estoy con ellos muchas horas por día, los conozco "como si fueran mis hijos", no tengo necesidad de tomar pruebas o exámenes. Yo sé lo que saben, lo que cada uno puede dar ..."

¿Qué concepto de evaluación tiene este profesor?

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



## Recuerde

- La autoevaluación es un tema complejo. La evaluación externa y la autoevaluación no son excluyentes. Los docentes suelen argumentar falta de madurez y responsabilidad en sus alumnos para asumir una autoevaluación. Si este es un diagnóstico correcto, es imprescindible que esos mismos docentes ayuden a sus alumnos a superar estas limitaciones durante el transcurso del proceso de enseñanza y de aprendizaje. Si se tiene claro que evaluar no es sólo asignar una nota, la autoevaluación tampoco lo es, se trata de ofrecer a los estudiantes oportunidades que estimulen la reflexión sobre sus propios aprendizajes, sus logros y sus dificultades.
- En general, los docentes homologan los términos "evaluar" y "calificar". Se asocia estrechamente evaluación con aplicación de un examen para la promoción o el aplazo. Generalmente cuando se toman estos exámenes los alumnos ya están fuera del proceso. Se puede determinar cuánto saben y dónde están sus dificultades más importantes pero no se pueden modificar aquellas cosas (conocimientos, habilidades y actitudes) no aprendidas. La evaluación como retroalimentadora del proceso de enseñanza y de aprendizaje pretende mejorar la situación cuando todavía es posible y se pueden hacer modificaciones. Se convierte así en el instrumento que permite por un lado adecuar la enseñanza al aprendizaje del estudiante y por otro orientar el proceso de aprendizaje de cada alumno a través del feedback o la devolución de los resultados obtenidos en pruebas y exámenes.
- Desde la perspectiva científica didáctica<sup>2</sup> la evaluación debe tener dos funciones básicas: comprobar la validez de las estrategias metodológicas y ofrecer al alumno la información que le ayude a progresar hacia el autoaprendizaje. La evaluación no es un recurso accidental o aleatorio que el docente puede utilizar cuando le queda cómodo.
- Es cierto que aquellos docentes que han podido tener un curso a lo largo de todo el proceso han establecido un vínculo y una comunicación especial con sus alumnos. Si bien esta relación puede permitir conocerlos y comprenderlos mejor en el momento del examen, creemos que (por ejemplo en una cátedra) a partir de una buena definición de objetivos que explicita claramente cuáles son los aprendizajes que los alumnos deben demostrar al finalizar el proceso, es factible

---

<sup>2</sup> Gimeno Sacristán, J (1986) Teoría de la enseñanza y desarrollo del currículo, Buenos Aires, Ed. Rei.

que cualquier docente esté en iguales condiciones para hacerse cargo de la evaluación con objetividad. En la práctica se utilizan exámenes orales en los que los alumnos sienten que pueden "defenderse" mejor. En los exámenes orales tradicionales se advierten toda clase de sesgos que determinan que los resultados sean poco confiables. Los alumnos y los docentes, en general, defienden la posibilidad de manipular la situación.

- ▶ El conocimiento ordinario que nos rodea no es evaluación. Piense en el estudiante o en el residente que más conoce y revise las competencias profesionales enumeradas en el Anexo 1 del módulo anterior. ¿Hasta qué punto estaría en condiciones de decir para cada una de ellas en qué grado de desempeño se encuentra ese estudiante?



*"Hace unos años estábamos trabajando en una escala de desarrollo infantil y el profesor que daba el curso pidió a una profesora que pensara en el niño o la niña de la clase que conociera mejor. Empezaron a hacer las preguntas de la escala y no se pudo continuar, porque esa profesora, con mucha experiencia, que estaba hablando del niño que mejor conocía de su clase, no estuvo en condiciones de resolver si ese niño podía desarrollar bien, medianamente o no podía, ciertas conductas o habilidades.*

*Esto significa que sí tenemos un cierto conocimiento de los estudiantes con los que estamos trabajando, pero es un conocimiento muy superficial, muy genérico. Si alguien nos hace preguntas más en concreto, más específicas, esto nos falla. No es esta visión ordinaria lo que va a mejorar las clases"*

*Miguel Zabalza<sup>3</sup>*

En todos los casos en que se implementa una evaluación, una vez corregido el examen debe brindarse a los alumnos información sobre sus aciertos y errores para que la experiencia -haber rendido un examen- le permita a los sujetos tomar conciencia acerca de sus puntos fuertes y débiles. **Brindar feedback** es una de las mejores maneras de lograr que los resultados de la evaluación orienten el aprendizaje y contribuyan al mejoramiento de la calidad.

Por ejemplo, cuando el examen tiene como propósito seleccionar residentes o becarios y/o otorgar certificados de especialista, la información sobre los resultados de los exámenes puede ayudar a los sujetos a decidir sobre futuras actividades de educación continua.

---

<sup>3</sup> Zabalza M (Marzo 2002) ¿Qué puede aportar la evaluación para mejorar el proceso educativo? Revista Novedades Educativas Año 14. N° 135. p. 4



*"El criterio de utilidad implica que es condición básica que la evaluación pueda ser útil, que pueda ser utilizada en el propio proceso de mejora. Y esto tiene una serie de condiciones. Una evaluación en este sentido tiene muy en cuenta a los usuarios, está buscando información sobre aspectos que resultan relevantes para los agentes del proceso. Obliga a seleccionar aspectos significativos, que tengan algo que decir a las personas que se verán afectadas por la evaluación. Obliga, además, a buscar sistemas analíticos. Una evaluación es tanto más rica cuanto más analítica. Una visión general de que las cosas van bien o van mal no nos ayuda en nada. Tenemos que saber en qué sentido van bien o qué aspectos van mejor o van peor. Es decir, tiene que ser una evaluación lo más analítica posible."*

*Miguel Zabalza <sup>4</sup>*

Independientemente de para qué se evalúa o qué información se esté buscando, siempre habrá que utilizar instrumentos que permitan relevar datos. La calidad de los instrumentos que se utilicen incide en la calidad de información que se obtenga.

## **Calidad de los instrumentos de evaluación Confiabilidad, validez y objetividad.**

En educación los instrumentos que se utilizan para medir "cuánto saben" los estudiantes son los exámenes orales y/o escritos. En la práctica docente, en nuestro país, los profesores de una cátedra o departamento son los responsables de construir los "instrumentos" con los que se recogen los datos sobre el rendimiento de los alumnos.

Para ingresar a la Residencia también se rinden exámenes que son elaborados por los profesionales que colaboran en los Comité de Docencia e Investigación de los establecimientos asistenciales.

Los resultados de los exámenes, como los resultados obtenidos con otros instrumentos o procedimientos de medida (balanza, cinta métrica, tensiómetro, etc) deben ser confiables para que las decisiones, que se tomen en función de esos datos, sean acertadas.

Los resultados, las mediciones, son confiables cuando son "verdaderos", "correctos" y cuando los diferentes valores obtenidos no se deben a errores de medida o al azar. ¿Qué tan exacta es la medida de lo que medimos?

---

<sup>4</sup> Ibidem. p. 6

Cualquiera que sea el instrumento que se utilice, e independientemente del atributo que se mida, existe la posibilidad del error. Al medir la tensión arterial se repite el procedimiento dos y hasta tres veces para tener un valor más confiable ya que el registro puede estar muy influido por cierto grado de ansiedad del paciente y por la agudeza sensorial del evaluador. Hay que asumir la existencia de "errores de medida".

¿En qué medida es confiable la información obtenida con los instrumentos de medida construidos por los profesores? ¿Cuál es el grado de validez de dicha información?

¿En qué medida, los puntajes obtenidos por los alumnos en un examen revelan los conocimientos que tienen? ¿En qué medida los resultados de los exámenes expresan las diferencias entre los alumnos? ¿Los resultados de los exámenes permiten identificar a los alumnos que "saben" y a los que "no saben"?



*"Información insuficiente: el profesor no cuenta a menudo con suficientes datos acerca del trabajo del estudiante como para poder calificarlo con justeza. En la marcha cotidiana de la enseñanza el profesor se ve obligado a evaluar el rendimiento de gran número de alumnos en períodos muy cortos, por lo que sus juicios concluyen siendo frecuentemente improvisados. Algunos reglamentos escolares y disposiciones ministeriales reducen también la duración de los exámenes orales finales a períodos tan breves (15 minutos, por ejemplo) que la información recabada acerca del rendimiento del alumno resulta a todas luces incompleta. Los juicios estimativos emitidos sobre bases inseguras se ven, así, influidos más por circunstancias azarosas que por una cuidadosa ponderación de los logros y yerros de los alumnos."*

Alicia Camilloni <sup>5</sup>

Las pruebas construidas por los profesores tendrán errores de medición. Cuando se trata de pruebas estandarizadas o test psicométricos se pueden calcular los límites del error e intentar reducirlos a límites aceptables.

La **confiabilidad** de una prueba indica hasta qué punto pueden atribuirse a errores de medición las diferencias individuales en los puntajes obtenidos y hasta qué punto los puntajes revelan verdaderas diferencias en las características evaluadas. El trabajo estadístico para determinar el margen aceptable de error o la

---

<sup>5</sup> Camilloni A. (1989), Las apreciaciones personales del profesor. Centro de estudiantes de Filosofía y Letras (Cefyl) Universidad de Buenos Aires.

corrección del mismo generalmente es realizado por especialistas en medición. El docente o profesor deberá estar en condiciones de saber hasta qué punto los resultados de los exámenes que aplica respetan ciertos requisitos o criterios técnicos, especialmente cuando de esos resultados se desprenden decisiones importantes para el alumno o la sociedad.

### ¿De qué depende que los resultados de un examen sean confiables?

Varias condiciones y características de los exámenes contribuyen a la confiabilidad de los resultados. De manera práctica se pueden mencionar algunos factores que afectan a la confiabilidad de un instrumento y en los que puede intervenir el docente:

1. **Objetividad.** La confiabilidad será mayor cuando el otorgamiento de puntajes sea poco influido por la subjetividad del evaluador. Los resultados de un examen son objetivos cuando varios examinadores frente al mismo examen asignan la misma nota. O bien, una puntuación es objetiva cuando un examinador asigna la misma nota a un examen cuando lo vuelve a corregir y puntuar en otro momento.



*"Si a una prueba se le asigna un puntaje que sea independiente del juicio de quien debe juzgarla, se puede afirmar que se la ha calificado de modo objetivo"*

*Pedro Lafourcade <sup>6</sup>*



*"Toda técnica de medición de la competencia médica debe dar resultados objetivos; es decir, las observaciones independientes de distintos expertos deben concordar."*

*Charvat, McGuire, Parsons <sup>7</sup>*

2. **El número de preguntas o ítem incluidos en el examen.** A menor número de ítem mayor posibilidad de que los resultados no sean tan confiables. Existe una relación entre longitud del test y confiabilidad.

---

<sup>6</sup> Lafourcade P, (1969), Evaluación de los aprendizajes, Buenos Aires, Kapelusz. p. 187.

<sup>7</sup> Charvat J, McGuire C, Parsons V. (1969) Características y aplicaciones de los exámenes en la enseñanza de la medicina. Cuadernos de Salud Pública N° 36. OMS. Ginebra.



"Ebel (1965) ilustra convenientemente esta relación entre la longitud del test y su confiabilidad: partiendo de una longitud original de 5 ítem y 0.20 de confiabilidad, duplicando sucesivamente el número de preguntas se tendría:

Número de Ítem	Confiabilidad	Número de Ítem	Confiabilidad	Número de Ítem	Confiabilidad
5	0,20	40	0,67	320	0,94
10	0,33	80	0,80	640	0,97
20	0,50	160	0,89	infinito	1.00

Como se observa en esta tabla la relación entre el número de ítem y confiabilidad no es constante. El aumento de confiabilidad entre 80 y su duplo es bastante menor que entre 40 y 80. Sólo con un número infinito de ítem se logrará una confiabilidad perfecta (1.00)"<sup>8</sup>

En una prueba estructurada, 80 ítem, sería un número mínimo adecuado para alcanzar una confiabilidad aceptable. Esta extensión de un examen es especialmente importante en exámenes finales en los que se decide la promoción de un alumno, la admisión de un residente, la certificación de un especialista. En exámenes parciales, con un propósito formativo, podrían ser aceptables exámenes con un menor número de preguntas pero nunca inferior a 40 ó 50.

Es frecuente que los docentes tomen pruebas objetivas o "choice", de 20 ó 30 ítem lo que determina, como se observa en la tabla anterior, una confiabilidad bastante reducida. En otras palabras: estos exámenes no ofrecen garantías sobre el conocimiento que tienen los alumnos. Este asunto es particularmente importante si se tiene en cuenta que a través de los exámenes los profesores están "garantizando" a la sociedad que los futuros profesionales han adquirido los conocimientos necesarios para cuidar la salud de la población.

La confiabilidad de los resultados de un examen tiene que ver con la exactitud y precisión del procedimiento de medición.

3. El **grado de homogeneidad** de los aspectos evaluados. Es probable que los resultados sean más confiables cuando la prueba explora un universo de temas más acotado. Mientras mayor amplitud de temas en una prueba, menor cantidad de preguntas en cada tema, lo que hace que los resultados sean poco confiables. Con una o dos preguntas cerradas sobre un tópico no se puede concluir que el joven profesional o el estudiante domina el tema. El margen de error es muy amplio.
4. El **grado de dificultad** de la prueba. La confiabilidad disminuye en el caso de que una prueba sea muy difícil o muy fácil. Hay mayor probabilidad de que los ítem de dificultad media (respondidos correctamente por el 50% de los alumnos) otorguen mayor consistencia a la prueba.

<sup>8</sup> Citado por Lafourcade P, Ob. cit.

## ¿De qué forma podemos comprobar la exactitud-precisión de nuestra medición?

Existen varias formas de hacerlo:

- repetición de la misma prueba o examen.
- administración de una segunda prueba "equivalente"
- división de la prueba en dos fracciones equivalentes

### 1.- Repetir la prueba (test-retest).

De la misma forma en que se hace en la medición de la tensión arterial: el mismo instrumento aplicado al mismo sujeto.

Este procedimiento no es recomendable en el caso de los exámenes ya que se produce un "aprendizaje durante la prueba". La misma prueba aplicada por segunda vez no presenta el mismo grado de dificultad.

### 2.- Administración de pruebas paralelas.

Aplicación de dos instrumentos que midan exactamente lo mismo, como pesar una misma sustancia en dos o más balanzas igualmente calibradas.

La forma equivalente de un examen debe considerarse como una forma construida de acuerdo a las mismas especificaciones, pero con preguntas diferentes. Es decir, deben ser dos pruebas idénticas, con preguntas que exploren los mismos conocimientos pero con otras palabras. Pueden ser aplicadas, a los mismos sujetos, una a continuación de la otra con cierto lapso de tiempo. La correlación entre ambas proporciona un coeficiente de confiabilidad. El uso de dos pruebas paralelas proporciona una base muy firme para estimar la exactitud de una prueba educativa.

La correlación entre dos formas paralelas de prueba, normalmente administradas con un lapso de varios días, representa el procedimiento preferido para calcular la confiabilidad. Pero este procedimiento demanda mucho tiempo del profesor para la confección de dos pruebas y el doble de tiempo para administrarlas y tabular los resultados.

### 3.- División de una prueba en dos mitades iguales o equivalentes.

Para implementar este procedimiento la prueba debe ser razonablemente larga (60 ítem o más). En cada una de las mitades deben quedar incluidas preguntas sobre los mismos contenidos y del mismo nivel de dificultad.

Un procedimiento práctico es agrupar todas las preguntas pares en una mitad y todas las impares en la otra. Las dos mitades funcionan como pruebas equivalentes. Se administra como una sola prueba en una sola ocasión y luego los resultados se tabulan como dos pruebas: una de preguntas pares y otra de preguntas impares. La correlación entre estas dos puntuaciones proporciona una medida de la exactitud con la que la prueba mide el conocimiento que tiene el sujeto. Mientras más extensa sea la prueba más confiable será esta medida.





*"La confiabilidad se define como el grado de precisión con que un instrumento es capaz de suministrar información. La expresión de la confiabilidad de una prueba se realiza comparándola con ella misma o con otra de forma equivalente. En cada expresión de ella se obtiene una medida del atributo, o atributos, que posee cada miembro de un grupo de alumnos. Cada medida identifica la posición de cada uno de ellos con respecto a su grupo en función del atributo medido. La consistencia con que mantiene su posición de una medición a otra refleja la confiabilidad de la prueba"*

*Alicia Bertoni*<sup>9</sup>



En algunas Facultades se están empezando a generar espacios específicos para trabajar los temas de evaluación, tanto para la elaboración de los instrumentos -exámenes- como para el análisis de los resultados. Le sugerimos que averigüe si en la institución en la que Ud. se desempeña existe algún Comité o grupo de trabajo que se dedica a estos temas. Muchos de los cálculos e interpretaciones se realizan actualmente a través de programas de computación, pregunte en el área de informática la disponibilidad de los mismos y la posibilidad de realizar un trabajo en común.

### ¿Qué significa validez?

Además de la exactitud y precisión de las medidas existe otra preocupación en relación a los exámenes contruidos por los profesores:

¿Mide el examen lo que se desea medir? ¿mide sólo lo que se desea medir?

Generalmente se dice que una prueba es válida si mide aquello que se propone medir.



*"La validez es otro aspecto de la calidad de un instrumento de medida de la competencia.*

- 1. Muestra hasta qué punto un examen mide "realmente" lo que se espera que mida.*
- 2. El grado en que un instrumento de medida cubre o muestrea en todos los aspectos lo que se pretende medir. Está basada sobre la opinión de expertos.*

---


<sup>9</sup> Bertoni A y col. (1996) Evaluación nuevos significados para una práctica compleja. Kapelusz. Buenos Aires. p. 72.

3. Para la evaluación de la competencia clínica, la validez de contenido debe basarse sobre:
- a. Las habilidades que se desean evaluar;
  - b. El contenido clínico, que debe estar en relación estrecha con la esfera de peritaje esperada;
  - c. El instrumento utilizado, que debe ser congruente con la estrategia pedagógica utilizada y la habilidad que se desea medir".

Carlos Brailovsky <sup>10</sup>

Los docentes deben decidir cuál es el instrumento más adecuado para los objetivos que van a ser evaluados. No se puede utilizar el mismo instrumento para evaluar por ejemplo "recuerdo de información" que para evaluar el logro de "actitudes de respeto frente a los paciente". El instrumento debe ser adecuado a la variable que se desea medir.

Nunca se mide una cosa o una persona. Las mediciones que se hacen son sobre las cualidades o los atributos de las cosas o personas; por ejemplo se mide la longitud de una mesa, la temperatura de un horno, la inteligencia de un niño, la madurez emocional de un adolescente, la habilidad para resolver un problema, el razonamiento clínico, etc... Algunos atributos ya están suficientemente definidos - por ejemplo el significado de longitud- pero otros, como inteligencia y/o habilidad para resolver problemas y/o razonamiento clínico, deben ser definidos operacionalmente antes de intentar ser medidos. ¿Cómo se manifiesta ese atributo, cómo puede ser percibido? Hasta que no se establece un consenso sobre qué significa y cómo se manifiesta no puede acordarse la forma en la que podrá ser medido y finalmente evaluado.

 *"El problema de establecer acuerdos en lo que respecta a un concepto dado es aún más complejo cuando consideramos atributos con los que trata el psicólogo y el educador. Por ejemplo ¿qué se define como inteligencia? ¿qué clase de conducta se considerará como inteligente? ¿se definiría el concepto principalmente en términos de ideas y conceptos abstractos que maneja el individuo o tendríamos que incluir la forma en la que maneja eventos concretos?... Casi todos tenemos una idea general acerca de lo que queremos decir cuando caracterizamos una conducta como inteligente, pero existen muchos puntos específicos en los que podríamos estar en desacuerdo cuando tratamos de elaborar una definición más precisa. Esto pasa a menudo con casi todos los conceptos*

---

<sup>10</sup> Brailovsky C, (2001), Aportes para un cambio curricular en la Argentina 2001. Facultad de Medicina (UBA) - OPS. Buenos Aires, p. 110.

*psicológicos, aunque en distintos grados, y el primer problema que afronta el psicólogo o el educador conforme trata de medir los atributos en los que está interesado es lograr una definición precisa, clara y aceptada universalmente del atributo que se propone medir"*

R. Thorndike, E. Hagen <sup>11</sup>

## Validez de contenido o curricular

Es imprescindible una precisa especificación de qué se quiere evaluar y consensuar una definición operacional del atributo a medir. ¿Qué objetivos de aprendizaje se van a explorar? ¿En qué nivel taxonómico? ¿El alumno manifestará su "conocer" como simple recordación, comprensión o aplicación? La validez de contenido es muy importante en las medidas de aprovechamiento que son las que aplican los docentes para evaluar el aprendizaje de los alumnos.

Hay que relacionar el contenido del examen con el contenido del curso y con los objetivos instruccionales. En la medida que los objetivos, metas de un curso, estén representados en la prueba ésta será válida. Dado que este análisis es esencialmente racional y lleno de juicios, a veces se menciona como "validez racional o lógica".

El problema de establecer la validez de contenido de una prueba está estrechamente vinculado con la planificación de la prueba.

Una prueba tiene validez de contenido cuando una muestra significativa de los objetivos de aprendizaje establecidos y de los contenidos desarrollados en el curso o materia, está representada en la misma.



*"Ya hemos visto la dificultad que supone querer evaluar ciertos objetivos. La búsqueda de objetivos definidos en términos de componentes operacionalizados que permitan una evaluación precisa no es más que un esquema válido para ciertos objetivos o para determinadas parcelas de algunos de ellos. En el fondo toda operacionalización supone elección y selección de notas operativas, y, por lo tanto parcialidad y subjetividad. Querer apoyarse en ella como requisito de objetividad resulta un tanto pretencioso en ciencias humanas si se quiere llevar más allá de ciertos límites. (...)*

*Las taxonomías de objetivos, como instrumentos clasificadores y diferenciadores de objetivos o efectos potenciales, puede ayudar a una evaluación más comprensiva.*

*Nos parece lógico, pues, pensar en un modelo de evaluación amplio, pero que admita que muchos de esos efectos secundarios no van a ser evaluables, porque puede que no sea factible hacerlo o porque el evaluador no fija su atención en ellos.*

---

<sup>11</sup> Thorndike R, Hagen E. (1996) Medición y evaluación en Psicología y Educación. Trilla. Méjico.

*Finalmente, habría que recordar que todo aquello que implica creatividad, divergencia personal, variación individual, es difícil someterlo a evaluación, e incluso la evaluación no debe hacer sino constatar esa divergencia."*

*José Gimeno Sacristán*<sup>12</sup>

## Validez predictiva



*"...se refiere a la correlación existente entre los resultados obtenidos en una o varias pruebas combinadas y el desempeño posterior del estudiante en aquellos aspectos en los que ha sido evaluado"*

*Alicia Camilloni*<sup>13</sup>

Sería difícil afirmar que un estudiante será buen médico por haber obtenido durante sus años de formación altas calificaciones en los diferentes "choices" rendidos. Una buena nota en un examen escrito y/o en los exámenes orales ¿predice un buen desempeño frente al paciente? Este tema es motivo de numerosas investigaciones y hasta el momento no hay evidencias suficientes en relación a qué tipo de exámenes pueden "garantizar" que el estudiante será un "buen" profesional.

La evaluación de una prueba como predictora es principalmente una evaluación empírica y estadística por lo que este tipo de validez suele ser designada como "validez empírica o estadística". El procedimiento básico consiste en aplicar el examen a un grupo de personas que ingresan a un trabajo y seguirlas observando para obtener de cada una, una medida específica del éxito o "buen desempeño" en el trabajo. Luego se calcula la correlación entre la puntuación obtenida en el examen y la medida de éxito o buen desempeño. Cuanto más alta sea la correlación, más efectivo será el examen como predictor.

Se han realizado algunas investigaciones tratando de establecer la validez predictiva del examen de selección para ingreso a la residencia y el desempeño asistencial de los residentes; también entre exámenes de admisión a la carrera de medicina y rendimiento durante la carrera. En general, los exámenes de selección o admisión que se aplican en nuestro medio tienen bajo valor predictivo.

---

<sup>12</sup> Gimeno Sacristán J. (1986) Teoría de la enseñanza y desarrollo del currículo. Rei. Buenos Aires.

<sup>13</sup> Camilloni A (1998) La calidad de los programas de evaluación y de los instrumentos que la integran. En La evaluación de los aprendizajes en el debate didáctico contemporáneo. Ed. Paidós Buenos Aires. p. 80.

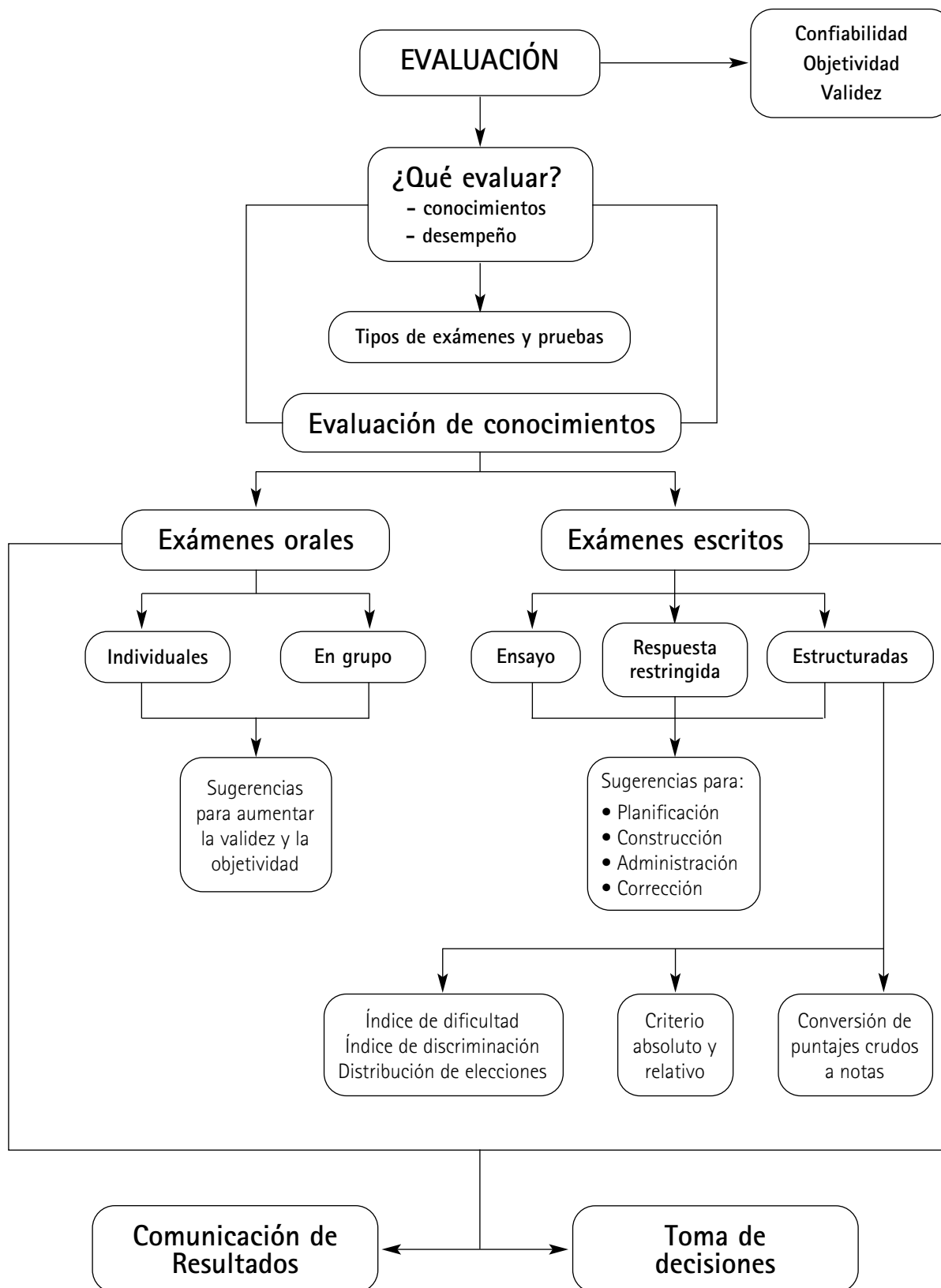
*El propósito de este Curso de Evaluación Educacional es capacitar a los docentes y profesionales de la salud para que puedan elaborar instrumentos de evaluación que brinden información válida y confiable. En última instancia, se trata de mejorar los sistemas de evaluación para "garantizar" a la sociedad que los profesionales formados, en la universidad y/o en los programas de Residencias y de Educación Continua, son confiables, que "saben" todo lo que necesitan saber para ser competentes en su ejercicio profesional.*

## Objetivos

Esperamos que al finalizar este Módulo Ud. sea capaz de:

- ✓ Reconocer las diferencias entre evaluar conocimientos y evaluar el desempeño.
- ✓ Reconocer las posibilidades y limitaciones de los distintos tipos de exámenes.
- ✓ Implementar estrategias para aumentar la validez y la objetividad de los exámenes orales y escritos.
  - elaborar una tabla de especificaciones.
  - seleccionar y redactar ítem.
- ✓ Revisar la calidad de una prueba estructurada aplicando Índices de Dificultad y de Discriminación.
- ✓ Emitir juicios de valor en función del análisis o interpretación de los datos.
- ✓ Convertir los puntajes obtenidos en un examen escrito al sistema de notas.
- ✓ Tomar decisiones en función de los propósitos de la evaluación y de los resultados obtenidos según criterio absoluto o relativo.

# Esquema de contenidos



## ¿Qué evaluar?

Hace algunos años han empezado a utilizarse términos como dominio, pericia, habilidad clínica, criterio médico, competencia, desempeño para referirse al comportamiento total de un trabajador y, en todos los casos, incluye el modo de organizar y emplear los conocimientos así como sus actitudes y sus interacciones con otras personas.

La competencia profesional se define como la capacidad del profesional para utilizar los conocimientos, las habilidades, las actitudes y la capacidad de juicio para solucionar problemas (complejos) que se presentan en el campo de su actividad profesional.

En el Módulo N° 1 desarrollamos el tema de las competencias profesionales como resultados esperados del aprendizaje. La mayoría de las cátedras expresan los resultados esperados en forma de objetivos. Si entendemos que los objetivos son los resultados esperados al finalizar el proceso de enseñanza y de aprendizaje, es lógico que se conviertan en el QUÉ de la evaluación.

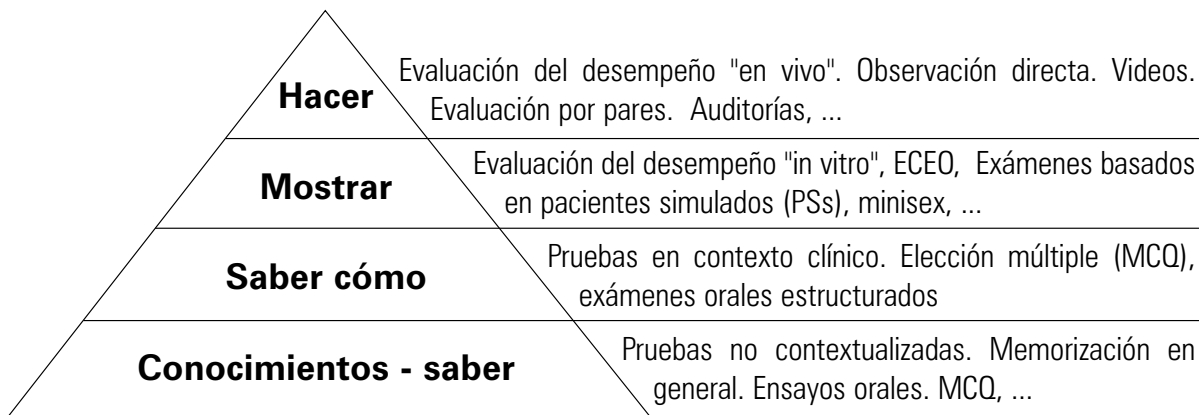
Según cómo se formulen los objetivos pueden servir o no como guía de la evaluación.


- ▶ Si se formulan los objetivos utilizando conductas observables como "describir, elegir, mencionar, ubicar en..., diferenciar, definir, etc." puede establecerse una relación entre ambos y los objetivos son una guía clara para la evaluación.
- ▶ Si se formulan utilizando conductas menos descriptivas y más amplias como "conocer, comprender, reflexionar, etc." es necesario realizar una especificación de estos términos para que, al formular los ítem de una prueba o las preguntas de un examen, no caigamos en el error de evaluar aquello que no nos hemos propuesto previamente como resultados de aprendizaje.

Así como el diseño de una planificación educacional comienza por establecer claramente los propósitos del docente y los objetivos que el estudiante debe alcanzar, también una buena evaluación debe comenzar por una revisión detallada de los objetivos educativos para identificar en forma precisa qué evaluar.

**La evaluación del desempeño significa esencialmente la evaluación del comportamiento del sujeto en una situación concreta de trabajo.**

**El docente debe evaluar aquello que se propuso como resultado del aprendizaje.**




 ... "Se puede decir que la validez de los instrumentos de medida está en relación con los estratos de la pirámide de Miller. No hay duda que la calidad profesional global se incrementa a medida que se trepa la pirámide. La complejidad taxonómica aumenta desde los simples conocimientos hasta la acción, la actividad en la vida real. Tener conocimientos (saber) no significa saber explicar cómo utilizarlos (decir lo que se debe hacer). Y ..., decir lo que se debe hacer no implica saber desempeñarse; y saber desempeñarse en una situación de evaluación no implica necesariamente actuar con sabiduría y profesionalismo en la vida real, es decir, de manera autónoma con verdaderos pacientes. Por lo tanto, si queremos categorizar estos niveles de la pirámide, los dos inferiores (en la base) representan los conocimientos, mientras que los dos superiores, mostrar y hacer, representan los comportamientos. El saber clásicamente se mide antes de medir la calidad de los desempeños. En general se utilizan exámenes de elección múltiple (MCQ), aunque también se pueden utilizar "ensayos" o redacciones cortas o aún exámenes orales, pero todas estas evaluaciones se practican de manera descontextualizada y miden principalmente la memorización de los hechos. El saber como se puede medir con modelos clínicos compuestos por preguntas contextualizadas, es decir a partir de una situación clínica simple, una pequeña descripción, pone en contexto "real" la situación. Los desafíos son mucho más complejos cuando se va trepando la pirámide; cuando queremos evaluar el mostrar (desempeño), enfrentamos situaciones taxonómicas de otro nivel, en las que no sólo se describe cómo sino que se muestra cómo. Se entra en una dimensión totalmente diferente de la evaluación. Para la evaluación del desempeño "in vitro" se utilizan situaciones que imitan lo que ocurre en la vida real, pero en un contexto homogéneo y controlado. Para la evaluación del desempeño se utilizan distintas técnicas y modelos. En este momento el más popular y probablemente el mejor estudiado y validado es el Examen Clínico Objetivo Estructurado (ECOE o ECEO). En la literatura de habla inglesa se representa con la sigla OSCE. Finalmente, se llega a la cúspide de la pirámide, es decir la evaluación de los profesionales en la práctica real, frente a verdaderos pacientes. Se puede realizar observación directa en consultorios. Las consultas se pueden grabar en video para analizarlas más tarde." ... <sup>14</sup>

<sup>14</sup> Brailovsky C, Ob. Cit. pp. 110-113



Es probable que antes del inicio de las clases el docente haya planificado una serie de objetivos que al final del período correspondiente no hayan sido cumplidos o también que muchos otros -no propuestos- hayan sido alcanzados. Siempre se producen aprendizajes no esperados. Los alumnos aprenden otras "cosas" no planificadas.

Una planificación flexible incluye una evaluación también flexible. En las decisiones acerca de qué incluir en la evaluación deben considerarse las vicisitudes del proceso de aprendizaje. Sobre todo en las instancias de evaluación formativa.

En las decisiones de evaluación sumativa final o de certificación, el docente debe asumir plenamente la responsabilidad de garantizar a la sociedad un profesional bien preparado. Desde esta perspectiva cobra una nueva dimensión la necesidad de establecer expresamente aquellas competencias que ineludiblemente todos los estudiantes, futuros profesionales, deben lograr.

Para valorar el aprendizaje de los alumnos, los docentes pueden tomar exámenes orales o escritos y pueden observar el comportamiento de los alumnos en determinadas situaciones. En los Servicios de Salud para evaluar la competencia de un trabajador, el supervisor utiliza sobre todo la observación del desempeño.

**La evaluación debe ser diseñada a partir de los objetivos, los exámenes y pruebas a utilizar deben ser pertinentes al tipo de conducta que se pretende evaluar.**

Qué evaluar	Tipos de exámenes y pruebas	
AREA COGNOSCITIVA (conocimientos)	Pruebas orales	Individuales Grupales (coloquio)
	Pruebas escritas	De desarrollo (ensayo) Semi - estructuradas Estructuradas Monografías, Informes, etc.
AREA PSICOMOTRIZ Y AFECTIVA (actitudes - destrezas)	Observación Directa	Listas de cotejo o de control Entrevistas Escala de evaluación ECEO / OSCE Otros

La evaluación de las competencias profesionales toma en cuenta las tres áreas.

## Recuerde

- ▶ La división de la conducta en tres áreas es un artificio, es útil para comprender la necesidad de distintos tipos de exámenes y pruebas.
- ▶ No existe un solo instrumento que brinde información suficiente y necesaria para la evaluación de todos los aprendizajes. Una evaluación del desempeño requiere necesariamente varios tipos de pruebas.




*En este módulo se presentan los exámenes para evaluar los dos primeros niveles de la pirámide de Miller. En el Módulo N° 3 se incluyen recursos para evaluar el desempeño.*

# Exámenes orales

## Examen oral individual

Se entiende por examen oral a la prueba rendida por un alumno ante un tribunal constituido por dos o tres docentes. Un alumno responde las preguntas que le son formuladas por los docentes. El alumno piensa y "trabaja" solo. Esta prueba tiene un duración muy variable (entre 5 y 30 minutos).

En la práctica, por la gran cantidad de alumnos, los exámenes orales, se realizan entre un estudiante con un docente. Frecuentemente, en los exámenes parciales, ese docente suele ser un auxiliar y en los finales un JTP y profesores. La mayoría de los docentes no planifican estas evaluaciones, en el momento del examen comienzan a formular preguntas y a partir de las respuestas del alumno exploran otros conocimientos que no siempre responden a los objetivos de aprendizaje. Suele suceder que repregunten y profundicen en aquellos conceptos que los alumnos menos dominan o que presentan dificultades en la exposición. De este modo el examen termina indagando aquello que el estudiante no sabe más que aquellos aspectos que conoce y que responden a los objetivos.

 *"-¿Tenés miedo del examen? -dijo Andrés.  
-No, más bien curiosidad. Por lo regular en la vida se sabe cómo van a ocurrir las cosas. Hasta podés imaginarte con bastante detalle lo que te va a hacer el dentista, lo que vas a comer en la casa de tu tía...  
Pero esto no: te repito que es un pozo, el enigma perfecto".*

*Julio Cortazar*<sup>15</sup>

La prueba oral ha sido tradicionalmente el método de examen más usado, especialmente en la enseñanza superior. Tanto las pruebas orales individuales como las grupales (coloquio) permiten establecer una relación personal directa entre evaluadores y evaluados.

.....

### *¿Qué piensa Ud. de los exámenes orales?*

▶ ¿Cuál ha sido su experiencia cuando era estudiante?  
¿Prefería los orales? ¿Por qué?

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

---

<sup>15</sup> Cortázar Julio. Ob. cit., p. 36

► Ahora como docente ¿Le gusta tomar exámenes orales?  
¿Por qué?

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

► ¿Para qué sirven los exámenes orales? ¿Qué tipo de información obtiene?

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

► ¿De qué depende que un examen oral sea "justo"?

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

► ¿Qué dicen los estudiantes sobre los exámenes que se toman en su unidad docente?

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



*En el Anexo 1 se incluye un interesante artículo sobre las dificultades en los exámenes orales.*

Hasta este momento se han analizado las características del examen oral individual.

Con frecuencia los docentes valoran como un aprendizaje deseable el desarrollo de la habilidad para trabajar cooperativamente en un grupo, para intercambiar información, para razonar frente a una situación compartida.

En el examen oral individual, difícilmente un estudiante pueda demostrar estas habilidades que son necesarias para la práctica profesional en el equipo de salud.

Para explorar el logro de este tipo de aprendizajes es útil organizar el examen como un coloquio.

## Examen oral en grupo: Coloquio<sup>16</sup>

El coloquio es una prueba oral que rinden, de manera simultánea, varios estudiantes con un docente.

La primera característica destacable y que diferencia marcadamente el coloquio de otras formas tradicionales de examen es que la situación en la que cada estudiante debe demostrar los aprendizajes realizados es una situación grupal. Lo totalmente nuevo es justamente el contexto en el que se desarrolla el examen: un grupo de alumnos con un docente como coordinador. La evidencia que se toma en cuenta para evaluar si se han logrado los objetivos propuestos es el desempeño de cada estudiante en esa situación de trabajo grupal. Lo que se juzga es el rendimiento de cada uno de los alumnos, de modo que la evaluación se expresa en una calificación individual.

¿Qué tipo de conductas pueden ser observadas en esta situación grupal y que no son visibles o perceptibles en un examen oral individual?

Se trata de comprobar si los alumnos han adquirido, además de los conocimientos específicos de la materia, la habilidad de trabajar cooperativamente en un grupo. Es decir: si el alumno es capaz de comunicar, discutir y compartir la información que posee.

Los grupos se constituyen con 6 o 7 estudiantes como máximo para que todos los integrantes del coloquio tengan oportunidad de participar activamente en la elaboración conjunta de los temas y, además, para permitir la tarea del coordinador ya que éste debe recordar y valorar las intervenciones de cada uno de los alumnos.

Habitualmente la situación de examen oral individual es estructurada con un tribunal constituido por 2 ó 3 profesores para interrogar a un alumno. En el coloquio se entrega la responsabilidad a un solo docente que debe asumir el papel de "animador" del grupo. Se aconseja que se integre al grupo coloquial un observador. Este observador registra las interacciones y puede completar la observación del coordinador y contribuir en la faz final de la evaluación.

Se establece una duración de 45 minutos como mínimo para el desarrollo del coloquio para brindar oportunidades a cada uno de los alumnos de intervenir en varias oportunidades y dando tiempo para la elaboración grupal de distintos temas.

Durante el examen coloquial los estudiantes deben abordar, por lo menos, 3 temas amplios, planteados en forma problemática, que promueva la discusión.

Uno de estos temas, el primero, puede ser propuesto por los alumnos para facilitar el comienzo del examen, aflojando la tensión inicial. Los otros temas, así como las preguntas individuales, son formulados por el docente coordinador.

Primera parte del coloquio: discusión en grupo del caso/problema

---

<sup>16</sup> Modificado de: Galli A, Hernández N C. de., Galperín J, "Una técnica no convencional para la evaluación del aprendizaje: el coloquio" Temas Universitarios. Revista de Planificación Universitaria. N° 5, junio 1978, Barquisimeto, Venezuela.

## **Ejemplo:**

### **Ejemplo de caso para el coloquio<sup>17</sup>:**

El objetivo de esta situación es observar la capacidad de los estudiantes para compartir información para elaborar posibles diagnósticos en un contexto específico y explorar la utilización de los conocimientos que poseen acerca de trastornos menstruales y de la alimentación.

El grupo tiene 15 minutos para leer y analizar la siguiente situación y responder las preguntas que figuran a continuación.

### **Una relación particular**

"En el bar de la esquina siempre hay mucho movimiento, sobre todo en las primeras horas de la mañana y las últimas de la tarde. El sanatorio de la ciudad está enfrente. En la pared principal se ve la obra de un artista local y en uno de los extremos un reloj de péndulo que marca las 18,30 hs. Se abre la puerta y entran dos mujeres que se acaban de encontrar. El mozo las mira y las invita a una mesa cerca de la ventana. Parecen madre e hija, es lo más probable. La señora aparenta entre 40 y 45 años, es de mediana estatura, elegante y prolija. La más joven, Laura, es una adolescente de aspecto muy agradable, ropa holgada que no permite adivinar su cuerpo. Lleva carpetas y un par de libros, como si quisiera aprovechar el tiempo para estudiar. Debe tener 14 años. En el medio del bullicio se escucha su charla, le piden al mozo una gaseosa y un yogur diet. La mamá dice: "Pedí algo para comer, tenemos más de media hora hasta la entrevista con la doctora".

- No, la verdad es que no tengo ganas, estoy bien así, ya piqué algo cuando pasé por casa al mediodía.

Se las nota ansiosas. La mamá parece cansada; tal vez los problemas en su oficina hoy fueron demasiados.

La hija protesta, -no sé qué necesidad hay de venir al médico. Me siento bien, aparte tengo que estudiar.

-Sí, pero hace bastante que no hacés un control. Además te veo más delgada.

-Siempre querés encontrarme algo. Estoy perfectamente bien, solamente un poco irregular. Seguro que es algo normal. Mejor vámonos, en otro momento volvemos. De paso podemos ver vestidos para mi fiesta de 15.

La mamá duda, tal vez tenga razón y no sea necesario consultar. Laura es muy juiciosa, estudia mucho, es excelente alumna. Tal vez exagero, el año pasado la vieron y estaba muy bien.

Llama al mozo para pedir la cuenta, la coca estaba todavía en el vaso y sólo unas cucharadas faltaban del yogur. Mira el reloj de la pared, mira a su hija que ya es otra. Piensa que su esposo va a preguntarle por la entrevista.

-Mejor vamos y hacemos la consulta. Es sólo un rato y nos quedamos tranquilas.

La hija mira y asiente. -Bueno pero después vamos a ver vestidos, quiero estar muy bien para la fiesta.

Salen abrazadas. Se las ve compinches. La mamá mira el reloj nuevamente. Ya es la hora (mientras calcula cuántos meses hace que su hija no menstrúa). Nuevamente algo ensombrece su mirada.

"Creo que tengo unas cuantas cosas que preguntar".

Toma a su hija de la cintura y cruza, llevándola casi corriendo".

1. ¿Qué tipo de relación tienen las dos mujeres?
2. ¿En qué se parecen y en qué se diferencian?
3. ¿Qué es lo que preocupa a la mamá? ¿Qué le preocupa a Laura?
4. ¿Hay algo en especial que les preocupa a ustedes?
5. ¿Qué características de Laura les llaman la atención?
6. ¿Qué les parece que le pasa a Laura? ¿Por qué?
7. Si ustedes fueran el médico que está por atender a Laura. ¿Qué información recabarían en primer término?

<sup>17</sup> Escrito por la Dra. Ángela Nakab para el Encuentro Presencial del Programa de Educación a Distancia PRONAP, 2001.

La segunda parte del coloquio -las preguntas individuales- le permiten al docente ratificar la apreciación del grado de conocimiento demostrado por el alumno durante la discusión de los temas. Esta instancia -interrogatorio- facilita al docente su tarea evaluativa dándole mayor seguridad al recurrir a una técnica ya conocida. Debe tener preparadas una serie de preguntas acorde en cantidad al número de participantes del coloquio de modo que todos tengan varias oportunidades para responder.

### **Ejemplo:**

#### **Preguntas individuales para la situación anterior:**

- *¿Cuáles son las causas más frecuentes que inhiben el ciclo menstrual?*
- *¿Cómo debe ser la relación LH/FSH para una normal maduración folicular?*
- *¿Qué índice es importante valorar en las adolescentes?*
- *¿Cómo se debe interpretar una amenorrea?*
- *¿Cuáles son los diagnósticos diferenciales ante una amenorrea?*
- *¿Cuáles son los estudios complementarios a realizar?*
- *¿Cómo se pueden clasificar los trastornos de la conducta alimentaria?*
- *¿Cuáles son los aspectos psicológicos y familiares de la anorexia?*
- *¿Cómo se aconseja el abordaje de este tipo de trastornos?*
- *¿Cuándo comienza a recuperarse la aparición de los ciclos menstruales?, etc.*

Una vez finalizadas las preguntas individuales se pasa a la última etapa: la comunicación o expresión de los juicios que merecen el desempeño de los alumnos, la coordinación del docente y la técnica de evaluación utilizada.

**Autoevaluación:** cada uno de los alumnos debe opinar respecto de su propio rendimiento y fundamentar su juicio frente al grupo, enriqueciendo o modificando el juicio de sus compañeros y del profesor. Se incorpora al coloquio esta instancia de autoevaluación considerando que el ejercicio profesional responsable reclama una constante autocrítica y que ésta es una actitud que se aprende ejercitándose en la valoración de los logros y limitaciones del propio rendimiento.

**Autoevaluación grupal:** el hecho de intervenir en la evaluación de los compañeros, analizando en grupo los aciertos y errores cometidos por cada uno, puede ser una forma de estimular la cooperación y reconocer que todos y cada uno de los integrantes de un equipo son responsables del producto final.

El grupo debe considerar además, el rol desempeñado por el coordinador. Estos comentarios deben ser registrados con vistas a contribuir a futuros ajustes o modificaciones de la técnica coloquial.

**Asignación de notas:** El docente responsable, en base a sus observaciones y tomando en cuenta la autoevaluación y las consideraciones hechas durante la evaluación grupal, asigna a cada uno de los estudiantes la calificación, fundamentando en todos los

casos cuáles son sus elementos de juicio para promover o no a un alumno. Esta circunstancia convierte el examen en una instancia de feedback muy valiosa para el estudiante que puede, con ayuda del profesor, percibir más nítidamente cuáles son sus falencias y cuáles son sus puntos fuertes.

Se introduce de esta manera una modificación en la rutina convencional de la asignación y comunicación de las notas. Generalmente, en los exámenes individuales la mesa examinadora asigna las notas una vez que se ha retirado el estudiante, de modo que éste no llega a conocer formalmente los fundamentos de la nota alcanzada. Además, el docente no comunica personalmente al alumno el resultado de su examen, delegando en un agente administrativo esta tarea, desentendiéndose de todo compromiso personal frente al estudiante.

**Escala de calificaciones:** En este tipo de exámenes suele utilizarse un sistema de dos categorías: Insuficiente y Aprobado. La nota de Aprobado significa el logro del 100% de los aprendizajes esperados. El Insuficiente expresa que el aprendizaje no alcanza el nivel mínimo deseado.

### **En síntesis:**

- ▶ En el coloquio varios estudiantes analizan y resuelven simultáneamente problemas o situaciones planteadas por el docente.
- ▶ Es importante que los estudiantes tengan experiencias previas en trabajos grupales durante la formación.
- ▶ Es conveniente que los docentes preparen con anterioridad las preguntas y problemas que presentarán a los alumnos en el momento el examen.
- ▶ La coordinación de un coloquio exige un entrenamiento específico en dinámica de grupo.

## **¿Cómo mejorar los exámenes orales tradicionales?**

Existen algunas estrategias para aumentar la objetividad y la validez de los exámenes orales de modo que los resultados de la evaluación sean más confiables.

### **a) Antes del día del examen**

Definir el propósito de la evaluación. ¿Para qué va a evaluar? Puede ser un examen parcial, un recuperatorio (evaluación formativa), un examen final (evaluación sumativa).

Jerarquizar los **objetivos de aprendizaje** que se van a explorar en la prueba.

Ajustar el examen a los objetivos de la materia, a los resultados de aprendizaje esperados. Es necesario que los docentes antes de formular preguntas para el examen oral revisen el programa de la materia y seleccionen, ponderen o jerarquicen los objetivos que un estudiante no puede dejar de lograr al término de la cursada. En esta revisión seguramente detectarán que no todos los objetivos tienen la misma importancia, que hay algunos esenciales en términos de la propia disciplina, del ejercicio profesional futuro y de las características de la sociedad en la que ese graduado se insertará (datos epidemiológicos, patología prevalente, condiciones de salud, etc). Esto es lo que no se puede dejar de preguntar. ¿Es igualmente pertinente preguntar a un alumno de pediatría, por ejemplo, Tratamiento de rehidratación oral en una diarrea, que preguntar ¿Qué test moleculares de ADN se deben solicitar si se sospecha atrofia espinal infantil progresiva? El tiempo que se destina a cada alumno en un examen oral es limitado, por lo que es absolutamente necesario explorar lo esencial e importante.

**La Tabla de especificaciones aumenta la validez del examen.**

Confeccionar la **Tabla de Especificaciones**. Durante la planificación del examen se puede hacer una lista de los temas más importantes y asignar un número de preguntas a cada tema.

**Elaborar las preguntas**, problemas, y/o situaciones con anticipación.

Una dificultad que frecuentemente se encuentra en los exámenes orales es que los docentes preguntan "grandes temas", comenzando generalmente por la frase "*Hábleme de ... ¿qué me puede decir de ... (TBC, hepatitis, etc.)*". En esta situación, con este formato de preguntas suelen aparecer diferentes problemas de comunicación. Cuando el docente formula esta frase, tiene en su mente una respuesta esperada, sabe lo que él diría, por donde comenzaría a hablar, que tópicos abordaría primero y cuáles después, tiene una expectativa de respuesta que difícilmente sea cumplida por el estudiante. Es improbable que el alumno empiece su respuesta por el mismo lugar que lo haría el docente. Generalmente los docentes comienzan, con cierta impaciencia, a reformular o especificar las preguntas interrumpiendo el razonamiento del estudiante. Esto disminuye la fluidez de la comunicación. Es conveniente formular preguntas más concretas y sucesivas, de respuesta más breve.

### **Ejemplo:**

#### **Preguntas en examen oral:**

- *Nombre tres causas frecuentes de ...*
- *¿Cuáles son los estudios complementarios que se deben pedir frente a la sospecha de... y por qué? ¿En qué orden los solicitaría?*
- *Explique tres recomendaciones que se le deben dar a un paciente que presenta...*
- *Realizaría una interconsulta con el especialista frente a ...? ¿Por qué? ¿Con qué especialista?*
- *Analice las diferentes alternativas terapéuticas y establezca una indicación para este paciente*
- *¿Qué variables tomó en cuenta para decidir la indicación de tal o cuál medicamento?*



Si el examen es en grupo, la situación o problema que se presente debe estar bien armado previamente. En este caso es particularmente importante no improvisar. Se debe elaborar una situación que genere controversia.

Otro problema de los exámenes orales es que se agota el repertorio de preguntas, especialmente luego de tomar durante varias horas a distintos alumnos. Es conveniente que todos los docentes de una cátedra, a lo largo del curso, vayan escribiendo preguntas de examen en función de los objetivos y de los textos que los alumnos deben leer. La elaboración de la **clave de respuesta esperada** implica definir la respuesta correcta para cada una de las preguntas o problemas. Todos los docentes deben coincidir o aceptar como correcta la misma respuesta. Es recomendable que en el momento en que se redacta la pregunta se elabore la clave. Anotar sintéticamente la respuesta esperada, detallando los puntos o subtemas que el alumno deberá desarrollar para que la respuesta sea considerada correcta.

**Elaborar previamente un repertorio de preguntas facilita la tarea.**

**Cada pregunta debe tener su clave de respuesta.**

Se llega así al momento del examen con una buena cantidad de preguntas. Cada pregunta puede escribirse en una pequeña tarjeta y clasificarse por objetivo o tema.

El hecho de preparar y disponer de un **banco de preguntas** con sus respectivas respuestas correctas en el momento de tomar examen oral tiene algunas importantes **ventajas**:

- amplía el repertorio y no se repiten las preguntas que "se nos ocurren en el momento".
- el docente no improvisa y todas las preguntas que se formulan al alumno corresponden a los objetivos de aprendizaje que deben ser evaluados.
- no se omite ningún tema importante desarrollado en el curso.
- es más fácil identificar los aciertos y los errores cometidos por el alumno, lo que permite brindarle una información más clara sobre el resultado de su desempeño.

Establecer el **nivel de exigencia para la aprobación del examen**.  
¿Cuántas preguntas deben ser respondidas correctamente para alcanzar el nivel de aprobación?

## **b) Durante el examen**

**Aplicar el instrumento de evaluación:** "tomar el examen". En el momento de la aplicación del examen oral es importante tener en cuenta algunos aspectos relacionados con la comunicación verbal y no verbal. Las preguntas que se formulan deben ser claras. Una vez formulada la pregunta o presentado el problema a resolver se debe esperar en silencio a que el estudiante responda. Muchos profesores, luego de formular la pregunta comienzan a agregar datos, a repreguntar lo mismo en otros términos y esta conducta puede distraer el razonamiento que está llevando a cabo el alumno para elaborar mentalmente la respuesta. Para promover una comunicación fluida será importante tomar conciencia de los elementos no

verbales que entran en juego en el momento del examen. Los estudiantes suelen estar muy pendientes de los gestos de aprobación o reprobación que manifiestan los profesores. Una manera de "sintonizar" o establecer empatía es acompañar el discurso o exposición del estudiante con gestos afirmativos de cabeza en aquellos momentos en los que está respondiendo correctamente. Cuando aparece un concepto equivocado, en el contexto del examen oral, muchos profesores suelen interrumpir al estudiante afirmando "No, esto no es así"; "Está mal lo que está diciendo" y a continuación comienzan a explicar ellos mismos el tema. En resumen el examen termina siendo una clase dada por el docente. Finalmente el docente siente, con razón, que fue él quien dio el examen y el alumno siente una enorme frustración ya que no ha tenido oportunidad de demostrar lo que había estudiado. La situación del error en un examen final debe quedar claramente explicitada tanto para el docente como para el alumno. Una buena estrategia es repreguntar sobre el tema con otras preguntas o situaciones que exploren el mismo concepto antes de "decidir" el conocimiento o desconocimiento.

**Registrar las respuestas de estudiante**, por ejemplo, con signos más (+) o menos (-), a cada una de las preguntas. Muchas veces ocurre que si el estudiante contesta muy bien, o muy mal, a la primera pregunta genera una imagen en el docente que tiñe la apreciación de las otras respuestas. Si el docente va registrando, ayudado por la explicitación de la clave de respuestas, que es lo que está contestando correctamente y que está incorrecto o incompleto, al finalizar el examen podrá rápidamente construir la calificación tomando en consideración la totalidad de las respuestas. La decisión de aprobar o no al estudiante será un poco más objetiva. Si son varios los examinadores puede asumir esta tarea alguno de los docentes que no está preguntando, de este modo podrá concentrarse solamente en lo que dice el alumno e ir comparándolo con las clave de respuestas esperadas. Una vez finalizado el examen, con el registro de las respuestas del alumno, será más fácil lograr el acuerdo entre los evaluadores acerca de la calificación a otorgar.

### c) Después del examen

**Informar los resultados a los estudiantes.** Habitualmente la devolución del examen oral la realiza algún administrativo que entrega las libretas al finalizar la mesa de examen o pone en la cartelera las listas con las notas. Muy pocas veces se planifica una instancia de devolución de resultados y análisis del rendimiento de cada estudiante en el examen. Se puede plantear una devolución personal al finalizar el examen. Si el profesor ha registrado de alguna manera las respuestas del estudiante podrá justificar con claridad la calificación obtenida por el estudiante. Podrá brindar retroalimentación o feedback específico, por ejemplo, "en el caso de diarrea pudiste resolver el diagnóstico e indicar el tratamiento correcto, pero no hubo claridad en las indicaciones al paciente en relación a

la dieta, signos de alarma", etc. En el coloquio el análisis del rendimiento grupal e individual está contemplado dentro de la estrategia, es un componente de la misma. La última etapa del coloquio consiste en el análisis del rendimiento por parte de cada estudiante, de los pares y del docente coordinador.

## Recuerde

- D La **validez** del examen aumentará significativamente si se incluye un muestreo representativo de los temas de aprendizaje. Revise los objetivos, elabore la tabla y prepare preguntas previamente.
  
- D La **objetividad** aumentará significativamente si se elabora una clave de respuestas y se controla la clave con otros colegas y si en el momento del examen participan por lo menos dos evaluadores simultáneamente. Registre las respuestas del estudiante.

**Validez de contenido** es el grado en que un instrumento de medida explora todos los aspectos de lo que pretende medir.

<b>Exámenes orales</b>	
<b>Posibilidades</b>	<b>Limitaciones</b>
<ul style="list-style-type: none"> <li>• Apreciar las habilidades de comunicación oral. Organización del discurso.</li>   <li>• Valorar la calidad de ciertos aprendizajes académicos, tales como: procesos de pensamientos seguidos en la solución de problemas, uso de terminología específica, etc.</li>   <li>• En un coloquio se puede apreciar la capacidad para intervenir en discusiones de grupos. Los objetivos más importantes oscilan desde los que se relacionan con el conocimiento y comprensión de la temática en discusión hasta los que toman en cuenta la habilidad para defender argumentos, rebatir hipótesis contrarias, saber escuchar, pedir aclaraciones oportunas, etc.</li> </ul>	<ul style="list-style-type: none"> <li>• <b>Escasa validez.</b> Generalmente se explora una parte limitada de la asignatura. En el desarrollo del examen intervienen variables no necesariamente relacionadas con los objetivos de aprendizaje que se quieren evaluar: influencia de rendimientos anteriores, disciplina, aspecto físico, conducta lingüística del examinado.</li> <li>• <b>Baja objetividad.</b> Uno de los principales problemas que expresan los alumnos en relación a estos exámenes se resume en "para aprobar depende de quién te tome", expresando que no todos los docentes preguntan sobre las mismas cosas, con igual nivel de profundidad y, lo que es peor, que no todos consideran correctas a las mismas respuestas.</li> <li>• <b>Baja confiabilidad.</b> La información obtenida en los exámenes orales además de su baja objetividad y escasa validez se caracteriza por ser poco confiable. Las decisiones que se toman -promover o aplazar al estudiante- generalmente tienen un amplio margen de error. Son muchos los "falsos positivos", es decir, los estudiantes que obtienen nota 4 y son aprobados cuando con otros instrumentos más confiables serían aplazados. Por este motivo muchos alumnos prefieren los orales, es más fácil llegar a "zafar".</li> <li>• <b>Otros inconvenientes:</b> <ul style="list-style-type: none"> <li>- En el docente pueden influir sus "simpatías" o "antipatías", estado anímico, cansancio, capacidad de atención a las respuestas, posibilidad de juicio crítico, etc.</li> <li>- También es muy común que se utilicen para evaluar más de un objetivo simultáneamente. Por ej.: ver cómo el alumno "expresa" la "información". Lo que conduce al error de confundir la capacidad de expresarse con el real conocimiento del tema.</li> <li>- La construcción de la calificación suele ser un problema de los exámenes orales, algunos profesores por lo general se manejan en un margen correspondido entre el 4 y el 7 o el 8 y otros docentes puntúan de 0 a 10. Algunos deciden la calificación de un alumno comparándola con los otros estudiantes que se han presentado al examen, otros docentes califican en comparación con los estándares, criterios u objetivos del aprendizaje.</li> </ul> </li> </ul>

## Exámenes escritos

Las pruebas escritas son los instrumentos de evaluación que con mayor frecuencia utilizan los docentes para tomar exámenes parciales, finales, recuperatorios. Requieren por parte del docente un trabajo meticuloso y ordenado para la confección del examen, la implementación y la posterior corrección del mismo.

Las pruebas escritas, según el **tipo de preguntas** que incluyan, pueden clasificarse del siguiente modo:

•ENSAYO	De ensayo o composición (de respuesta abierta). Es el típico examen escrito en el cual el alumno debe desarrollar un tema determinado. Demandan una respuesta elaborada por parte del evaluado.	
•SEMI ESTRUCTURADAS	Tradicional mejorada o de respuesta restringida. Recordación. Completamiento: llenar el o los espacios en blanco.	
•ESTRUCTURADAS	De reconocimiento	De doble alternativa (Si - No, verdadero-falso) Apareamiento o correspondencia Ordenamiento Selección múltiple

La implementación de un examen escrito requiere un esfuerzo importante por parte del docente.

Se pueden identificar varios pasos o etapas en la implementación de un examen escrito:

### A. Construcción de la prueba

Identificar los **objetivos de aprendizaje** que se van a evaluar en la prueba.

Si bien la definición de objetivos es uno de los primeros pasos dentro de la planificación educativa, puede ocurrir que sea necesario redefinirlos en términos más operacionales o seleccionar y jerarquizar algunos para evaluarlos. Recuerde que los exámenes escritos son instrumentos válidos para evaluar el dominio cognoscitivo. Recuerde que se puede evaluar para diagnosticar el nivel inicial de los alumnos, para comprobar como marcha el proceso de aprendizaje, como pre-final, para calificar, para orientar a los alumnos que tengan dificultades de aprendizaje, para seleccionar residentes, para recertificar especialistas, etc. El

propósito de la evaluación puede incidir en la extensión y profundidad con que se decidan explorar los contenidos en el examen.

Confeccionar la **Tabla de Especificaciones**: ¿qué temas se van a explorar? ¿cuántas preguntas corresponderán a cada tema? ¿qué nivel de profundidad tendrán las preguntas?. Esta herramienta ayuda a aumentar la validez de contenidos de la prueba.

**Redactar las preguntas**, ítem, problemas, situaciones y la clave de respuesta: definir la respuesta correcta para cada una de las preguntas o problemas. Todos los docentes deben coincidir o aceptar como correcta la misma respuesta. Esta práctica contribuye a mejorar el grado de objetividad de la prueba.

**Asignar puntaje** a cada uno de los ítem y el puntaje máximo posible de la prueba. Establecer el **nivel de exigencia** para la aprobación del examen.

**Diseño gráfico** de la prueba: ordenar las preguntas, agruparlas por temas, redactar las instrucciones generales (explicitar el tiempo disponible, el puntaje otorgado a cada pregunta, etc.)

Hacer las fotocopias necesarias.

## **B. Aplicación y corrección de la prueba**

Aplicar el instrumento de evaluación: "tomar el examen" el día y hora anunciado.

Corregir la prueba según la clave de respuesta elaborada. Adjudicar las calificaciones.

Analizar la calidad del instrumento. En las pruebas de selección múltiple se aplican los índices de Dificultad y de Discriminación.

## **C. Comunicación de resultados**

Brindar feedback. Informar los resultados a los alumnos y analizar con ellos los errores más frecuentes.

## **D. Toma de decisiones**

Tomar decisiones en relación con la promoción de los alumnos, (criterio absoluto y relativo), con el instrumento (modificación o supresión de ítem) y con la enseñanza (actividades remediales o compensatorias). Lógicamente las decisiones que tome el docente dependerán con los propósitos con los que el examen fue implementado.

## Construcción de la prueba

### Tabla de especificaciones

Se trata de planificar el examen de modo de garantizar la validez de contenido. Es una manera racional de planificar la prueba, ayuda a decidir cuáles son los aspectos útiles o importantes que se deben explorar.

Se puede confeccionar como un cuadro de doble entrada donde aparecen en una coordenada los distintos tipos de conducta esperados o niveles de profundidad de las preguntas (categorías taxonómicas) y, en la otra, los temas. Se pondera la importancia de cada tema y cada nivel taxonómico con mayor o menor número de preguntas.

La tabla de especificaciones:

- Define lo que se quiere medir antes de redactar las preguntas.
- Refleja el consenso por parte de los profesores acerca de lo que debe ser medido.
- Grafica la distribución y proporción de las preguntas según los contenidos a evaluar en el examen.
- Proporciona una "fotografía" exacta de lo que será el examen.

Veamos de qué modo se realiza a través de un ejemplo:

### Ejemplo:

*Tema a evaluar: Primer año de vida. Atención del niño sano*

<i>Subtemas a considerar</i>	<i>Conductas de</i>		
	<i>Información</i>	<i>Comprensión</i>	<i>Aplicación</i>
<i>1. Alimentación 1.a. Lactancia Materna 1.b. Alimentación complementaria</i>			
<i>2. Inmunizaciones 2.a. Indicaciones 2.b. Contraindicaciones</i>			
<i>3. Screening 3.a. Visión 3.b. Audición 3.c. Caderas</i>			
<i>4. Conductas anticipatorias 4.a. Evolución psicomotriz 4.b. Prevención de accidentes</i>			

• **Información:** comprende el recuerdo de elementos específicos y universales, el recuerdo de métodos y procesos, o el recuerdo de una configuración, estructura o contexto.

• **Comprensión:** se refiere a un tipo de entendimiento tal que el sujeto sabe lo que se comunica y puede utilizar el material, el conocimiento, la idea, etc. Se evidencia cuando se traduce un

conocimiento de una forma de comunicación a otra sin alterar el significado, cuando se lo interpreta, etc.

• **Aplicación:** es el uso de abstracciones en situaciones particulares y concretas. Las abstracciones pueden presentarse como ideas generales, reglas de procedimientos y métodos generalizados. También pueden ser principios técnicos, ideas y teorías que deben ser recordadas y aplicadas.

Que una pregunta corresponda a información, comprensión o aplicación, se relaciona con el contexto de aprendizaje y de la experiencia de la persona que responde.



*Si quiere profundizar este tema le recomendamos la lectura del Módulo N° 2 del curso Metodología Docente en Ciencias de la Salud. Páginas 24-25*

El segundo paso es decidir la distribución de ítem, es decir, cuántos ítem de información, comprensión y aplicación se elaborarán en relación a los objetivos.

**Ejemplo:**

*La tabla quedará del siguiente modo:*

<i>Subtemas</i>	<i>Ítem de información</i>	<i>Ítem de Comprensión</i>	<i>Ítem de Aplicación</i>	<i>Total de Ítem</i>
<i>1</i>	<i>10</i>	<i>9</i>	<i>3</i>	<i>22 27,5%</i>
<i>2</i>	<i>10</i>	<i>9</i>	<i>3</i>	<i>22 27,5%</i>
<i>3</i>	<i>8</i>	<i>6</i>	<i>2</i>	<i>16 20%</i>
<i>4</i>	<i>9</i>	<i>9</i>	<i>2</i>	<i>20 25%</i>
<i>TOTAL</i>	<i>37 46,25%</i>	<i>33 41,25%</i>	<i>10 12,5%</i>	<i>80 100%</i>

El total de preguntas que tendrá la prueba dependerá del tipo de ítem que incluya: Recuerde que en los exámenes "objetivos" (con ítem de selección múltiple, V-F, respuesta breve, etc.), la confiabilidad de la prueba depende, en gran medida, de la cantidad

de preguntas. Es recomendable entre 60 y 80 como mínimo, para alcanzar un grado de confiabilidad aceptable.

En caso de preguntas de respuesta breve, basadas en la presentación de un caso clínico, que implica varias preguntas a su interior, pueden ser en total 10 casos del tipo PMP (Patient-management-problem). La solución de estos casos corresponde a un nivel taxonómico de aplicación.

Con la confección de la tabla de especificaciones la prueba resultará equitativa, logrará una relación más ajustada entre los aprendizajes promovidos y los conocimientos explorados en el examen. En síntesis: el examen ganará en validez.

Es común que cuando se plantea en una cátedra elaborar un "choice" varios docentes redacten, en forma aislada algunos ítem o preguntas sobre su especialidad. Muchas veces los temas se repiten o hay un gran número de preguntas que exploran información, memorización de datos.

Elaborar la tabla de especificaciones como instrumento de trabajo permite orientar a los docentes sobre qué tipo de preguntas son las que hacen faltan para explorar todos los contenidos considerados en la prueba.



***Revise el último examen escrito que usted haya tomado e identifique:***

▷ ¿Cuántas preguntas tiene ese examen? .....

▷ ¿Cuáles fueron los contenidos o temas explorados en ese examen?

.....  
.....

▷ El examen incluye preguntas de distinto nivel taxonómico?  
Copie un ejemplo de pregunta de información:

.....  
.....

Copie un ejemplo de pregunta de comprensión:

.....  
.....

Copie un ejemplo de pregunta de aplicación:

.....  
.....



- ¿Algunos temas son más importantes que otros?  
¿Cómo se han ponderado? ¿Con mayor número de preguntas? ¿Con mayor puntaje asignado a esos temas?

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

- ¿Podría reconstruir la tabla de especificaciones correspondiente a ese examen?

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



### Redacción de preguntas

En la literatura se mencionan, a modo de sinónimos, los términos "preguntas", "ítem", "consignas" o "reactivos" para referirse al estímulo que se le presenta al estudiante para que éste de una respuesta. A veces en de pregunta se dice ítem porque, estrictamente, no siempre es una interrogación lo que se plantea.

El tipo de preguntas que se incluyen en una prueba dependerá de los objetivos que se deben evaluar, la habilidad de los docentes para elaborar preguntas cerradas-estructuradas, la cantidad de docentes disponibles para corregir las pruebas y la cantidad de alumnos a evaluar.

Las preguntas que se incluyen en los exámenes escritos pueden ser clasificadas en:

- De recordación.
- De reconocimiento.

Las preguntas de recordación, requieren que el alumno "piense" la respuesta, que busque, analice y seleccione entre toda la información que tiene, cuál es la respuesta posible. Luego además la tiene que "escribir", respetando los códigos y reglas del lenguaje escrito.

En cambio, en las preguntas de reconocimiento, al ofrecer las opciones de respuesta, requieren que el alumno "reconozca" la respuesta correcta entre las ya dadas. En todo caso piensa cuál es la respuesta analizando las alternativas que se le presentan. Además no tiene que escribir, redactando la respuesta sino sólo marcarla, señalarla.

Preguntas de Recordación

- De respuesta abierta = ensayo o composición
- De respuesta restringida =
  - respuesta breve
  - completamiento

Preguntas de Reconocimiento

De respuesta cerrada =

- alternativas constantes
- selección múltiple
- ordenamiento
- apareamiento o correspondencia



*Algunos ejemplos de preguntas han sido tomados de:*

- *Sociedad Argentina de Pediatría. PRONAP. Exámenes Finales.*
- *Sociedad Argentina de Cardiología: Curso bianual de cardiología.*
- *UBA. Facultad de Medicina. Tercera Cátedra de Farmacología. Exámenes finales y Guías de Trabajos Prácticos.*
- *Construir Salud. Programa de Educación Continua en Medicina Familiar.*
- *FASGO. FASGO XXI. Programa de actualización.*
- *Los chistes han sido tomados de Thyne JM. (1978) Principios y técnicas de examen.*

### ***Preguntas de respuesta abierta, de ensayo, de desarrollo***

Las preguntas de desarrollo son las que solicitan que "escriba" o que "desarrolle" un tema determinado. El alumno piensa y organiza el tema según su propio criterio y con total libertad.

Las respuestas escritas por el estudiante permiten evaluar habilidades y aptitudes tales como: redacción; organización del discurso; habilidad para estructurar el mensaje; lógica de razonamiento; capacidad de síntesis; dominio de la lengua escrita.

Se pueden utilizar prueba con preguntas de ensayo o desarrollo cuando:<sup>19</sup>

1. Se desee evaluar la capacidad de expresión escrita.
2. Se cuente con la posibilidad de que la prueba sea corregida por dos o más correctores.
3. El grupo de estudiantes es reducido.

<sup>19</sup> Modificado de Cox K y Ewan C (1990) La Docencia en Medicina. Ed. Doima Barcelona. p. 163

.....

***Analice la siguiente situación:  
Examen parcial de Pediatría***

**Tema "Inmunizaciones"**

- ▶ ¿Todos los examinados tienen obligación de responder en forma similar, desarrollando los mismos subtemas?  
Sí ( ) No ( )  
¿Por qué?

.....

.....

.....

- ▶ ¿Qué dificultades se plantean en la corrección de este tipo de pruebas?

.....

.....

.....

- ▶ ¿Cómo solucionarías estas dificultades?

.....

.....

.....



***Sugerencias***

Mediante las preguntas de ensayo se quiere evaluar la habilidad del estudiante para usar información por lo tanto al responder no se deben limitar a reproducir los textos leídos.

- Antes de empezar a elaborar la pregunta tenga en mente con toda claridad qué procesos mentales quiere que el estudiante use al contestar el examen.
- No empiece las preguntas de ensayo con palabras o frases como "qué", "quién", "cuándo", etc porque estas interrogaciones sólo requieren memoria.
- Empiece las preguntas de ensayo con palabras o frases como: "compare", "dé las razones para", "explique cómo", "haga una predicción de lo que sucedería si", "establezca la diferencia entre", "ilustre"...
- Una pregunta acerca de un punto controvertido debería evaluarse más en términos de la presentación de evidencias para tomar una posición que por la posición que se tomó. Ejemplo: "Se ha sugerido que el costo de todos los cuidados provistos por los médicos y el costo de los medicamentos deberían ser sufragados por los fondos provistos por el gobierno nacional. Escriba su grado de acuerdo o desacuerdo. Apoye su posición con argumentos lógicos"

El profesor deberá calificar no por la posición que adopte el estudiante sino por las razones y fundamentos en los que basa su posición.

Preguntas de ensayo o desarrollo	
Posibilidades	Limitaciones
<ul style="list-style-type: none"> <li>• No requieren elaboración previa</li> <li>• En la elaboración de la consigna o pregunta no se da ningún indicio de respuesta al estudiante</li> </ul>	<ul style="list-style-type: none"> <li>• Escasa Validez: no cubren todos los objetivos y contenidos propuestos ni una muestra representativa de los mismos. Generalmente se preguntan temas que exploran la memorización de información, el alumno "recita" el texto estudiado.</li> <li>• Baja objetividad: La corrección suele ser demasiado subjetiva al no existir una única respuesta correcta. Distintos correctores asignan diferente puntaje.</li> <li>• La corrección insume una enorme cantidad de tiempo.</li> </ul>

Algunas de las dificultades de las preguntas de desarrollo pueden ser superadas orientando la respuesta del alumno:

- a través de indicaciones más específicas.
- elaborando previamente la clave de corrección: las respuestas que son consideradas aceptables por un grupo de docentes.

### ***Preguntas de respuesta restringida***

Son las que restringen u orientan la respuesta del estudiante. Las preguntas son precisas y acotan la respuesta, no permitiendo que el alumno se explaye sobre aspectos que podrían ser irrelevantes, o sobre aspectos que el docente no decidió evaluar. Pueden solicitar el recuerdo de hechos o datos específicos, principios y leyes.

- ***Recordación simple:***

Es una pregunta o una instrucción que exige una respuesta muy breve; comúnmente una palabra, un porcentaje, una fecha. Generalmente exploran el nivel de memoria.

### **Ejemplo:**

1. Indicar por lo menos tres mecanismos por los cuales una droga puede cerrar un canal iónico (Debe responder: **mediante proteína G, por unión directa al canal, bloqueando un receptor con canal intrínseco**).

2. Mencione el efecto adverso característico de la beclometasona. (Debe responder: **moniliasis bucal por inmunosupresión**).

## **Recuerde**

- ▶ Evitar ambigüedades en la redacción de las preguntas.
- ▶ Puntualizar aspectos a tener en cuenta en la respuesta.
- ▶ Elaborar previamente una clave de respuestas correctas.

### • **Completamiento:**

Otro formato en el que suelen presentarse los ítem de respuesta restringida es el de completamiento.

Se presenta un párrafo con algunos espacios en blanco que el alumno debe completar con una palabra o cifra de modo que el enunciado sea correcto.

### **Ejemplo:**

*Las formas histopatológicas de cáncer mamario en las embarazadas no difieren de las halladas en las pacientes no gestantes. Predominan los carcinomas. ....*

(Debe responder: **ductales y lobulillares**)

## **Sugerencias**

- Las preguntas se redactan en forma tal que no den lugar a respuestas equívocas o incompletas: exigen univocidad.  
En este tipo de ítem conviene aclarar cuál es la especificidad que se espera en la respuesta en aquellas preguntas en las que se solicita "más frecuentes" (concepción estadística epidemiológica); "más importantes" (conviene aclarar en relación a qué es más importante, puede ser relacionado con la mortalidad, con la importancia para el diagnóstico diferencial, importancia funcional, etc.)
- Los espacios en blanco o punteados no deben ser precedidos por artículos ni seguidos de adjetivos sugerentes del concepto que se debe completar ya que estas palabras son indicadoras de género y número.
- Los espacios o líneas para contestar deben ser uniformes para que su tamaño no sugiera la respuesta.
- No deben tener tantos espacios en blanco que oscurezcan el sentido de la afirmación.
- **Para la puntuación:** Todo o nada, independientemente de la cantidad de términos o conceptos a completar.

Preguntas de respuesta restringida o corta	
Posibilidades	Limitaciones
<ul style="list-style-type: none"> <li>Las preguntas son de fácil diseño</li> <li>Aumenta el espectro de temas a explorar.</li> <li>En la elaboración de la consigna o pregunta no se da ningún indicio de respuesta al estudiante. El alumno debe buscar la respuesta ya que no se le ofrecen opciones.</li> <li>Los examinadores se pueden poner de acuerdo sobre la respuesta esperada con más facilidad que en las pruebas de desarrollo.</li> </ul>	<ul style="list-style-type: none"> <li>Escasa Validez: generalmente no cubren todos los objetivos y contenidos propuestos.</li> <li>Generalmente se preguntan temas que promueven la memorización de información. Las respuestas correctas dependen a menudo del recuerdo de palabras claves. Con frecuencia las preguntas suelen ser triviales.</li> </ul>

Con preguntas de respuesta restringida también se puede explorar la capacidad de análisis de la información en situaciones clínicas. En este caso se solicita la comprensión y la aplicación de los conocimientos.

### Ejemplo:

#### **Analice la siguiente situación clínica** <sup>20</sup>

Matías concurre al Centro de Salud para un control. Ha asistido desde el mes de edad en forma regular a cada citación. En la actualidad tiene 18 meses, aun continúa con lactancia materna. Su peso actual es de 11.450 g, su talla es de 84 cm y su perímetro cefálico es de 47 cm. Controles antropométricos efectuados a los 12 meses registraban un peso de 10.200 g y a los 15 meses de 10.800 g.

#### **Responda el siguiente cuestionario**

- Evalúe el crecimiento. Se adjuntan las gráficas de crecimiento
  - ¿En qué percentilo de peso para la edad se encuentra?
  - ¿Es adecuada la tendencia de la curva? ¿Por qué?
  - ¿En qué percentilo de longitud corporal se encuentra Matías?
  - ¿Cuándo lo cita para el próximo control? ¿Por qué?
- Al revisar la cavidad bucal de Matías, ¿Qué dientes espera encontrar?. Nómbralos
- Mencione 5 habilidades motoras que espera encontrar en su desarrollo neuromotriz.
- ¿Cuántas palabras, promedio, puede estar diciendo?
- Al revisar su carnet de vacunación, ¿qué vacunas espera encontrar?
- ¿Corresponde alguna indicación de vacunas en esta edad? ¿cuál?
- Mencione las indicaciones que le daría a la mamá de Matías en relación con los siguientes tópicos:
  - Control de esfínteres
  - Estimulación del desarrollo (mencione al menos 4)
  - Hábitos y Pautas de alimentación (mencione al menos 4)
  - Prevención de accidentes (mencione al menos 6 recomendaciones)

(Este ejemplo se retomará más adelante detallando la clave de respuesta)

<sup>20</sup> Modificado de: Cisaruk de Lanzotti E y Gorosito M, Pautas de control del niño de 1 a 5 años. PRONAP 2000, Módulo 2, Cap. 1

Este último ejemplo corresponde a "simulaciones escritas". Se trata de una situación clínica que explora el nivel de "aplicación". En la literatura suele identificarse este tipo de preguntas, a partir de un caso clínico, como pruebas PMP (patient management problem), manejo clínico y/o solución de problemas.

### **Preguntas estructuradas**

El alumno debe seleccionar la respuesta entre las opciones que se le ofrecen. Debe reconocer y luego ordenar, correlacionar o marcar la respuesta correcta.

- Cinco formatos:
- Alternativas constantes.
  - Apareamiento o correspondencia.
  - Ordenamiento.
  - Selección múltiple.
  - Menú largo.

#### **• Alternativas Constantes:**

Son aquellas en las que se ofrecen sólo dos opciones: que debe optarse por "Sí o No", "Verdadero o Falso", "Correcto o Incorrecto". También se las llama "de alternativas constantes".

La información se expresa en forma de proposiciones. Poder distinguir entre las proposiciones verdaderas o falsas indica generalmente dominio de la información verbal, aunque según cómo esté formulado el enunciado puede evaluarse comprensión de la información.

#### **Ejemplo:**

*Marque V (verdadero) o F (falso) según corresponda en relación a los estudios Milán I, II y III.*

<i>Todas las pacientes fueron irradiadas a las 4 semanas posteriores a la cirugía.</i>	V <input type="checkbox"/>	F <input checked="" type="checkbox"/>
<i>Las pacientes con ganglios axiliares positivos fueron sometidas a quimioterapia.</i>	V <input checked="" type="checkbox"/>	F <input type="checkbox"/>
<i>La masectomía de Halsted y el enfoque Q.U.A.D. mostraron los mismos resultados en pacientes mayores a 55 años con ganglios positivos y con negativos.</i>	V <input checked="" type="checkbox"/>	F <input type="checkbox"/>
<i>En pacientes portadoras de un componente intraductal extenso se observó una ventaja con masectomía y con Q.U.A.R.T.</i>	V <input checked="" type="checkbox"/>	F <input type="checkbox"/>
<i>La supervivencia global de las pacientes tratadas con masectomía es mayor.</i>	V <input type="checkbox"/>	F <input checked="" type="checkbox"/>
<i>Las curvas de recidivas locales fueron similares con Q.U.A.D o T.A.R.T:</i>	V <input type="checkbox"/>	F <input checked="" type="checkbox"/>

En los ítem de doble alternativa los estudiantes tienen una probabilidad del 50% de acertar por azar la respuesta correcta. Se puede elevar el nivel de dificultad si se solicita que corrijan los enunciados falsos. Además de marcar Falso debe reemplazar las palabras subrayadas.

**Ejemplo:**

1. El tratamiento quirúrgico, durante el primer trimestre de gestación, debe evitarse por sus graves efectos teratogénicos V ( ) F (X) (Clave). Quimioterápico .....
2. La insulina es segregada por la glándula pituitaria V ( ) F (X) (Clave). Páncreas .....

También se puede elevar el nivel de dificultad multiplicando la cantidad de elecciones que deben hacer en un mismo ítem.

**Ejemplo:**

En la siguiente lista de patologías identifique con una cruz (X) en C (correcto) aquellas que son causadas por virus y con una cruz (X) en I (incorrecto) aquellas que tienen otro agente etiológico.

- |              |             |             |             |
|--------------|-------------|-------------|-------------|
| a) Sarampión | C ( ) I ( ) | f) Varicela | C ( ) I ( ) |
| b) Difteria  | C ( ) I ( ) | g) TBC      | C ( ) I ( ) |
| c) Influenza | C ( ) I ( ) | h) Tifoidea | C ( ) I ( ) |
| d) Malaria   | C ( ) I ( ) | i) SIDA     | C ( ) I ( ) |
| e) Paperas   | C ( ) I ( ) |             |             |

(Debe colocar (X) en C en a), c) e), f) i) y debe colocar (X) en I en b), d), g), h).)

Generalmente los ítem de doble alternativa evalúan información puntual. Sin embargo, es posible construir ítem que exploren la comprensión de la información.

**Ejemplo:**

Una adolescente de 14 años de edad comió lo siguiente durante un periodo de 24 hs.

Desayuno	Almuerzo	Cena
1 vaso de leche (228 g)	1 vaso de gaseosa 1 hamburguesa con pan 1 porción de papas fritas	1 bife (250 g) 1 porción de puré 1 porción de pastel de manzana

Con base en la tabla anterior, identifique, para cada uno de los siguientes nutrientes, con una cruz (X) en S (suficiente) si la ingesta ha sido adecuada y con una cruz (X) en D si ha sido deficiente.

- |                  |             |               |             |
|------------------|-------------|---------------|-------------|
| a) Calcio        | S ( ) D ( ) | e) niacina    | S ( ) D ( ) |
| b) Calorías      | S ( ) D ( ) | f) proteína   | S ( ) D ( ) |
| c) carbohidratos | S ( ) D ( ) | g) vitamina A | S ( ) D ( ) |
| d) hierro        | S ( ) D ( ) | h) vitamina C | S ( ) D ( ) |

Debe señalar con (X) en S b), c), e), f) y con (X) en D: a); d), g) h)



- Es necesario construir más de un ítem para evaluar cada objetivo y para disminuir la posibilidad de azar.
- La verdad o falsedad no debe estar determinada por elementos menores, que no ocurra que una parte de la proposición sea verdadera y la otra falsa.
- Cuidar que las proposiciones no sean obvias.
- Cuidar que la proposición no sea copia textual del libro u otro material de estudio, ya que sólo medirá memoria.
- Deben evitarse enunciados negativos, ya que dificultan la comprensión y se eleva la dificultad no relacionada al conocimiento del tema. Esto además disminuye la validez.
- No usar enunciados con "nunca", "siempre", "todos", "imposible", etc., porque suelen ser falsos ya que en medicina (como en cualquier otra ciencia no exacta) es muy improbable que algo suceda siempre, en todos los casos, o que algo sea imposible.
- El enunciado debe expresar una sola idea en forma clara y concisa.
- No hay una regla fija en cuanto a la conveniencia de poner más enunciados verdaderos que falsos o a la inversa, nuestra experiencia es que los enunciados que son falsos discriminan mejor entre los alumnos de mejor y peor rendimiento en el examen (discriminación positiva).
- **Para la puntuación:** se puntúan todo o nada. Un punto por respuesta correcta y 0 si es incorrecta. Este criterio de corrección se mantiene en los casos en los que se pida la corrección de la respuesta falsa. Se asigna 1 punto en aquellos casos en que respondieron falso y además respondieron por qué era falso o corrigieron correctamente el dato y 0 si alguna de las dos partes del ítem está incorrecto o incompleto. En el caso de varias elecciones se asigna 1 punto si marcaron en el casillero correcto todas y cada una de las opciones y 0 si marcaron alguna incorrecta o dejaron alguna en blanco.

Preguntas de doble alternativa	
Posibilidades	Limitaciones
<ul style="list-style-type: none"> <li>• Fácil de elaborar.</li> <li>• La corrección insume poco tiempo.</li> <li>• Aumenta el espectro de temas a explorar.</li> <li>• Insume poco tiempo para responder.</li> <li>• Los examinadores se pueden poner de acuerdo sobre la respuesta esperada con más facilidad que en las preguntas de desarrollo.</li> </ul>	<ul style="list-style-type: none"> <li>• Alta probabilidad de azar.</li> <li>• Generalmente se preguntan temas que promueven la memorización de información específica. Las respuestas correctas dependen a menudo del recuerdo de palabras claves.</li> <li>• Exigen absoluta e incuestionable verdad o falsedad.</li> </ul>

**• Apareamiento o Correspondencia:**

Son aquellos ítem que solicitan establecer relaciones entre conceptos: hechos, fechas, causalidad, medidas, personajes, enfermedades, etc... Son aptos para evaluar una gran variedad de objetivos que representan manejo de conceptos. Implica pedir al sujeto que señale los casos que pertenecen a un concepto diferenciándolo de otro semejante. También se utilizan para

relacionar información visual (radiografías, ecografías, diagramas, cuadros, etc.). Sirven para cualquier contenido en que se espera evaluar "qué es lo que se relaciona con qué"; no sólo símbolos con términos, sucesos con fechas, sino también cuestiones más complejas como procesos con productos, hipótesis con deducciones, capacidades con tareas, instrumentos con trabajos, etc.

**Ejemplo:**

*Relacione las formas de TBC extrapulmonar que figuran en la columna de la izquierda con las afirmaciones que aparecen en la columna de la derecha. Coloque la letra correspondiente en el espacio punteado. Cada letra puede ser usada una o varias veces.*

<i>Formas de TBC extrapulmonar</i>	<i>Características</i>
1. <i>Pleuritis tuberculosa . . . . .</i>	a) <i>Es más frecuente en la mujer</i> b) <i>La PPD es negativa en el 50% de los casos</i> c) <i>Se puede presentar en forma brusca, similar a un cuadro de neumonía aguda</i>
2. <i>Meningitis tuberculosa . . . . .</i>	d) <i>El examen físico suele ser normal</i> e) <i>Se asocia a mortalidad elevada</i>
3. <i>Adenitis tuberculosa . . . . .</i>	f) <i>El cuadro clínico se inicia en forma incidiosa: cefalea, decaimiento, apatía, febrícula.</i> g) <i>Es la forma más grave de TBC extrapulmonar</i>
4. <i>TBC miliar . . . . .</i>	

*(Clave: 1. c; 2. b - f - g - e; 3. a; 4. b - d)*

**Sugerencias**

- En una de las columna se incluyen 3 o 4 conceptos y en la otra entre 8 y 10 enunciados, si son más de 10 habrá que preparar otro ítem.
- La correspondencia no debe ser 1 a 1, ya que al menos una de las relaciones se podrá contestar correctamente por descarte. En alguna de las dos columnas se pueden poner distractores, es decir, opciones que no correspondan y que deban ser consideradas por el estudiante y descartadas por incorrectas o no pertinentes.
- Es necesario indicar en el cuerpo del ítem si cada opción puede usarse una, varias o ninguna vez.
- Las opciones deben ser homogéneas entre si en cada una de las columnas.
- Use un encabezado, para cada columna, que describa con precisión su contenido. Un encabezado descriptivo ayuda al examinando a comprender la tarea. Si el constructor de la prueba no es capaz de identificar un encabezado que defina específicamente el contenido, entonces es muy probable que tenga delante de si una colección de reactivos heterogéneos.
- Organice las opciones de respuestas en un orden lógico. Ordenar nombres en orden alfabético o las fechas en orden cronológico reduce el tiempo que se necesita para contestar.
- **Para la puntuación:** Un punto si están todas las opciones correctamente emparejadas y 0 puntos si existe algún error en el apareamiento. Sin embargo, el docente puede decidir, en función del grado de dificultad que implique el ítem, si asignará un puntaje a cada uno de los enunciados correctamente emparejados.

En el ejemplo presentado de TBC, con el primer criterio se adjudica 1 punto si no existe ningún error. Con el segundo criterio de corrección, el ítem tendría un puntaje máximo de 4 puntos al adjudicar 1 punto a cada premisa de la columna de la izquierda correctamente respondida. En este caso el puntaje puede variar entre 0 a 4.

• **Ordenamiento:**

Son aquellos en los cuales se presentan elementos que conforman un proceso o una secuencia cronológica y debe ordenarse según un criterio establecido. También es posible presentar una secuencia o proceso en imágenes o diagramas. No siempre es fácil encontrar un tema en el que la secuencia de pasos determine la corrección o incorrección de la respuesta. Suelen ser útiles para explorar el conocimiento de algoritmos diagnósticos, o de procedimientos.

**Ejemplo:**

*Ordene la siguiente secuencia de eventos celulares (no observables clínicamente) que llevan a un cuadro de asfixia neonatal. Coloque al lado de cada opción el número que corresponda a la secuencia correcta.*

<i>Eventos</i>	<i>Debe contestar (Clave)</i>
<i>a) Falla secundaria de energía</i>	<i>4</i>
<i>b) Muerte neuronal</i>	<i>5</i>
<i>c) Falla primaria de energía</i>	<i>2</i>
<i>d) Alteración del intercambio gaseoso</i>	<i>1</i>
<i>e) Mecanismos de citotoxicidad</i>	<i>3</i>

**Sugerencias**

- El orden de los pasos de la secuencia, de las fases del proceso, de las etapas del procedimiento, etc. debe ser indiscutido, si hay más de un modo correcto de ordenar, se debe buscar otro tipo de ítem.
- **Para la puntuación:** Asignar el puntaje a todo o nada. Es decir, 1 punto si todos los pasos o elementos han sido ubicados en la secuencia correcta y 0 si existe algún error.

• **Selección múltiple:**

Son aquellos ítem en las que dada una pregunta o enunciado se ofrecen 4 ó 5 opciones entre las que el alumno debe elegir una como respuesta correcta. Es un tipo de ítem muy utilizado, generalmente se identifica como "choice". En las publicaciones internacionales se encuentran con la sigla MCQ (multiple choice question) y la sigla castellana es PEM (preguntas de elección múltiple).

Con frecuencia se escucha decir que las pruebas de selección múltiple sólo evalúan memoria. Esta observación es correcta. La mayoría de los exámenes estructurados elaborados por los docentes demandan pura memoria.

**Ejemplo:**

¿Cuál es el porcentaje de personas infectadas por el *M. Tuberculosis* que desarrollan enfermedad activa?

- a) 0 a 5%
- b) 6 a 10% (Clave)
- c) 11 a 15%
- d) 16 a 20%

Pero esto se debe a la falta de habilidad de los docentes para construir la prueba; no se trata de una limitación intrínseca al tipo de ítem.

Los docentes que construyen pruebas estructuradas que exploran sólo memoria frecuentemente en los exámenes orales también formulan preguntas sólo informativas.

En este tipo de ítem estructurados se pueden reconocer dos componentes:

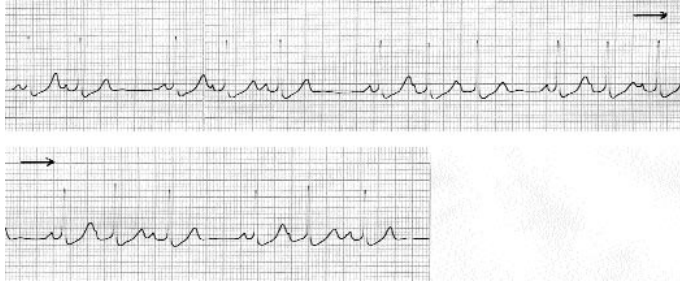
- la "**base**" o enunciado de ítem.
- las "**opciones**" de respuesta que se ofrecen.

La **base de un ítem** estructurado puede incluir diferente tipo de información. Las bases de los ítem pueden ser simples o complejos, con mayor o menor cantidad de datos. En general la interpretación de casos clínicos demandan en los alumnos procesos mentales de orden superior (análisis, síntesis, evaluación)

Entre las opciones se encuentran la respuesta correcta (clave) y los "**distractores**". Los distractores correctamente elaborados ayudan a graduar la dificultad de la pregunta.

**Ejemplos:**

	<b>Base del ítem</b>	<b>Opciones para seleccionar la respuesta</b>
<b>1.</b>	<p>Un varón de 3 meses de vida es llevado a su consultorio por dificultades en la alimentación y mal progreso de peso. Se trata de un niño nacido de término, con peso adecuado y sin complicaciones perinatológicas. La madre refiere que cuando lo prende al pecho succiona débilmente y que ha tenido varios episodios de tos y ahogo durante la alimentación. Solo aumentó 500 gramos desde el nacimiento, aún no tiene sonrisa social ni línea media, es francamente hipotónico y presenta criptorquidia bilateral. El resto del examen es normal.</p> <p>El estudio de este paciente se debe iniciar con:</p>	<ul style="list-style-type: none"> <li>a) Interconsulta con genética. (clave)</li> <li>b) Dosaje de enzimas musculares.</li> <li>c) Electromiograma y velocidad de conducción.</li> <li>d) Biopsia muscular y de nervio periférico.</li> </ul>

	<i>Base del ítem</i>	<i>Opciones para seleccionar la respuesta</i>
2.	<p>Aplique el principio de Fick a los datos siguientes y calcule la cantidad de glucosa (en mg/min) extraída por el bazo de un perro.</p> <ul style="list-style-type: none"> <li>- Gasto cardíaco - 2 litros/min.</li> <li>- Volumen diastólico - 200 ml/min.</li> <li>- Índice cardíaco - 2,8 litros/m<sup>2</sup></li> <li>- Glucosa arteria I - 100 mg/ml</li> <li>- Glucosa vena esplénica - 96 mg/ml</li> <li>- Flujo sanguíneo ( a nivel del bazo ) - 50 ml/min.</li> </ul> <p>Marque la opción que corresponde al resultado del cálculo:</p>	<p>a) 140</p> <p>b) 178</p> <p>c) 200</p> <p>d) 225</p> <p>e) 250</p>
3.	<p>La arritmia que se observa en el trazado (tiras continuas) corresponde a:</p> 	<p>a) taquicardia supraventricular con bloqueo AV de 2° grado.</p> <p>b) bloqueo sinoauricular de 2° grado.</p> <p>c) taquicardia sinusal con bloqueo AV de 2° grado (Wenckebach).</p> <p>d) extrasistolia supraventricular</p>

Especialmente interesante resulta la exploración del nivel de comprensión de relaciones.

Ya sabe que muchas veces el grado de dificultad de la pregunta se puede graduar con los distractores ofrecidos para la elección. Para elevar su grado de dificultad, el nivel taxonómico explorado en una pregunta se pueden presentar opciones que impliquen trabajar con relaciones entre hechos, eventos o acontecimientos.

Se pueden construir ítem que incluyan en el enunciado o base dos proposiciones vinculadas con un conector y en las opciones las posibles relaciones entre hechos, eventos o variables.

Se pueden redactar varios ítem cambiando el enunciado o la base y manteniendo las mismas opciones de respuesta o similares.

### **Ejemplo:**

	<i>Base del ítem</i>	<i>Opciones para seleccionar la respuesta</i>
1.	<p>Indique cuál de las opciones es correcta en relación a la siguiente afirmación:</p> <p>[<sup>1</sup>En condiciones normales, no hay paso de agua ni de electrolitos de los capilares hacia los alvéolos pulmonares] porque [la presión hidrostática de la sangre es inferior a la presión oncótica del plasma en todo el capilar pulmonar<sup>2</sup>].</p>	<p>a) las dos proposiciones son verdaderas y tienen una relación de causa - efecto</p> <p>b) las dos proposiciones son verdaderas pero no tienen relación de causa - efecto</p> <p>c) la primera proposición es verdadera pero la segunda falsa</p> <p>d) las dos proposiciones son falsas</p>

<i>Base del ítem</i>	<i>Opciones para seleccionar la respuesta</i>
<p>2. Indique cuál de las opciones es correcta en relación a la siguiente afirmación: [En un músculo en actividad, la diferencia de contenido de oxígeno entre la arteria aferente y la vena eferente está aumentada] porque [el aumento de las necesidades energéticas del músculo implica una vasodilatación importante]</p>	<p>a) las dos proposiciones son verdaderas y guardan una relación de causa - efecto b) las dos proposiciones son verdaderas pero no tienen una relación de causa-efecto c) la primera proposición es verdadera pero la segunda es falsa d) las dos proposiciones son falsas</p>
<p>3. Señale cuál de los siguientes opciones expresa la <b>relación correcta</b> entre peso y comienzo puberal.</p>	<p>a) La relación entre ambas variables es directamente proporcional (a más peso comienzo puberal más temprano) b) La relación es inversamente proporcional (a más peso comienzo puberal más tardío) c) No hay relación entre peso y comienzo puberal. d) La relación es variable según razas y latitudes.</p>
<p>4. Señale cuál de las siguientes opciones expresa la <b>relación correcta</b> entre la proposición "Signo de Babinski" y las variables: x: síndrome cerebeloso y: síndrome tabético</p>	<p>El signo de Babinski: a) se asocia solamente con X. b) se asocia solamente con Y. c) se asocia a la vez con X e Y. d) no está asociado ni con X ni con Y.</p>
<p>5. Señale cuál de las opciones es correcta respecto a la relación entre [ataxia], y x: [x síndrome cerebeloso] y: [y síndrome tabético]</p>	<p>La ataxia: a) se asocia solamente con X b) se asocia solamente con Y c) se asocia tanto con X como con Y d) no se relaciona ni con X ni con Y</p>

Existe un formato adecuado para explorar conocimientos con un mayor nivel de discriminación: todas las opciones pueden ser correctas pero una de ellas es mejor o más precisa que las demás. Este formato se denominan de "mejor razón" o "de mejor respuesta" y las palabras o conceptos usados habitualmente son: "la más frecuente", "la más probable", "lo primero que se debe sospechar", "la que describe mejor"; "la mejor conceptualización de", etc

### Ejemplo:

<p>1. Mariela, de 17 años, consulta porque en el último año presenta menstruaciones muy dolorosas que no responden a los analgésicos habituales, debiendo guardar reposo en cama. Es una adolescente sana y cuenta que desde su menarca, ocurrida a los 12 años, siempre presentó menstruaciones regulares con molestias leves. Su examen físico es normal. No ha tenido aun relaciones sexuales. Es muy responsable y estudiosa. ¿Cuál de los siguientes es el <b>estudio inicial más apropiado</b> en este caso?</p>	<p>a) Laparoscopia. b) Ecografía transvaginal. c) Ecografía pelviana. d) Evaluación psicológica</p>
--	---

<p>2. Un paciente, tabaquista, con diagnóstico de hipertensión estadio II, índice masa corporal de 31. Laboratorio: triglicéridos 350 mg/dl, HDL-C 35 mg/dl, colesterol total 235 mg/dl, ácido úrico 7,9 mg/dl, insulinemia basal 25 uU/ml y a las dos horas 125 uU/ml. Además de indicar dieta, desde un punto de vista fisiopatológico ¿Cuál es el esquema antihipertensivo que corresponde en <b>primera instancia</b>?</p>	<p>a) betabloqueantes no cardioselectivos b) antagonistas alfa 2 postsinápticos c) betabloqueantes cardioselectivos d) diuréticos</p>
<p>3. ¿Cuál es la secuela electrocardiográfica <b>más frecuente</b> en pacientes operados de Tetralogía de Fallot?</p>	<p>a) hemibloqueo anterior izquierdo b) bloqueo de rama derecha (Clave) c) bloqueo auriculoventricular completo d) bloqueo de rama izquierda</p>

Otro formato de ítem se caracteriza por ofrecer varias opciones correctas y una **incorrecta** o que no corresponde o que constituye la excepción.

### Ejemplo:

<p>1. ¿En cuál de las siguientes situaciones está <b>contraindicado</b> el desarrollo de un programa de ejercicios físicos?</p>	<p>a) Infarto de miocardio crónico b) Angor crónico estable c) Insuficiencia cardíaca compensada d) Estenosis aórtica sintomática</p>
<p>2. Las siguientes, <b>excepto una</b>, son indicaciones del Til Test. Marque la excepción.</p>	<p>a) Síncope en presencia de obstrucción severa del tracto de salida del VI b) Pacientes con lesiones graves durante las caídas (alto riesgo) c) Pacientes con síncope recurrentes d) Pacientes añosos con caídas recurrentes inexplicadas</p>
<p>3. Los siguientes, <b>excepto uno</b>, son signos de agrandamiento auricular derecho. Marque la excepción.</p>	<p>a) Desviación a la izquierda del eje de la onda P b) Ondas P con muescas y espastamientos en la derivación D2 c) Ondas P de tipo negativo positivo en la derivación VI d) Ondas P bimodales en las derivaciones V3 a V6</p>
<p>4.Cuál de las siguientes opciones es <b>incorrecta</b> respecto al síncope vasovagal de origen periférico?</p>	<p>a) Casi siempre ocurre en la posición de pie b) Estando de pie, el 70% del volumen sanguíneo se encuentra por debajo del nivel del corazón c) Los factores humorales, como el sistema renina-angiotensina y epinefrina-norepinefrina, no parecen jugar un rol importante d) En el corto plazo los mecanismos compensatorios dependen de la manipulación del cronotropismo, del inotropismo y de las resistencias periféricas</p>

Se las suele denominar "pruebas objetivas". Una prueba estructurada puede ser objetiva o no. Recuerde que la objetividad implica que el resultado de la corrección no varíe según el corrector. Va a depender de la discusión de los docentes sobre cuál es considerada la única respuesta correcta al elaborar la clave de respuesta.

Una prueba o examen puede incluir ítem de un solo tipo, de selección múltiple por ejemplo, o se pueden combinar dos o tres tipos.

Cualquiera sea el formato de este tipo de preguntas, la consigna debe estar clara. Deben preguntarse conocimientos o información. No deben preguntarse opiniones o creencias personales del estilo ¿Qué cree usted ... ?; ¿Qué piensa usted sobre .... ?, ¿Qué opinión le merece ... ?, ya que casi cualquier respuesta debería ser aceptada como válida porque el alumnos puede creer, pensar u opinar de manera muy distinta a la información científicamente aceptada como verdadera. Si no contesta como espera el docente no se podrá discriminar si es porque no sabe, no leyó, no estudió o porque no comparte la información y opina diferente.

Exigen univocidad: una única posibilidad de respuesta correcta.

## Sugerencias

### ► En relación a la base o el enunciado:

- Debe ser claro, evitar la ambigüedad y orientar sobre la cuestión que está preguntando. El estudiante debe saber qué es lo que se le está preguntando antes de leer las opciones.
- No se debe agregar nada que debilite o confunda la elección de la respuesta correcta.
- Se debe destacar cuál es la operación que debe hacer el alumno en esa pregunta: Marcar lo **INCORRECTO**, Marcar lo **CORRECTO**, Señalar la **EXCEPCIÓN**.

### ► En relación a las opciones:

- Las palabras que se repiten en todas las opciones deben aparecer en la base del ítem.
- Los distractores y la respuesta correcta deben ser homogéneos:
  - similar longitud: que la repuesta correcta no sea siempre la más larga y completa.
  - similar vocabulario: que la respuesta correcta no sea la única que utiliza términos técnicos.
  - similar tipo de contenidos: todos conceptos, todos fármacos, todos estudios diagnósticos
- Reducir la posibilidad de respuesta correcta por azar. Es conveniente usar 4 opciones como mínimo para que la posibilidad de azar sea de 25% como máximo. Ideal sería utilizar más distractores para bajar la probabilidad de azar, pero es muy difícil encontrar distractores plausibles. Estudios de Índices de Dificultad muestran que en las pruebas con 5 opciones generalmente aparece un distractor que no funciona, es decir, que no es elegido por nadie.
- Ubicar la clave en las distintas posiciones (a -b -c -d)
- Los distractores, más que inventarse, deben recolectarse de los errores comunes que van surgiendo en el aprendizaje para intentar que los alumnos discriminen entre lo correcto y lo incorrecto posible para ese grupo. Su función no es engañar, "pescar" o "poner trampas".



- Recordar que a través de los distractores se puede graduar la dificultad del ítem.
- Evitar como última opción "Todas las anteriores". Si el alumno sabe lo suficiente como para encontrar que al menos 2 opciones son correctas, puede deducir que la respuesta correcta es "Todas las anteriores"; al mismo tiempo, si reconoce a una opción como incorrecta elimina automáticamente esta opción. Evitar como última opción "Ninguna de las anteriores" o "Ninguna es correcta". La explicación lógica es semejante a la del apartado anterior.
- No hay una regla fija en cuanto a si siempre hay que preguntar sobre lo correcto, lo incorrecto o la excepción. Es más fácil construir ítem en los que haya una sola opción incorrecta, que construir 3 ó 4 alternativas incorrectas y una sola correcta.

► **Para la puntuación:**

Todo o nada. Un punto si la respuesta es correcta y cero si no lo es. Asignando el mismo puntaje a cada pregunta bien respondida. Por ejemplo: todas valen un punto. Si hay un tema más importante, construir más preguntas.

En caso de ponderar algunas preguntas, por ejemplo adjudicando 2 puntos, se corrige respetando el todo o nada. No se admiten 0,5 o 1 punto. Obtiene 2 puntos o 0.

**Recomendaciones generales para las pruebas estructuradas:**

- Comprobar que las preguntas estén relacionadas a objetivos importantes. Las preguntas deben ser apropiadas al nivel de conocimientos que se espera que tengan los estudiantes (esto es lo que ya fue especificado en los objetivos). Se debe respetar lo establecido en la Tabla de Especificaciones.
- Tratar que el instrumento en su totalidad sea válido, esto es que evalúe el logro de los objetivos de aprendizaje. Si hay un objetivo muy importante, se deberán construir tantos ítem como sea necesario para que el docente se asegure el aprendizaje del mismo y no construir un sólo ítem y asignarle una puntuación mayor. Este tipo de corrección puede distorsionar la realidad ya que un estudiante, por azar, puede contestar correctamente y al ser esa la única pregunta que indagaba sobre ese aspecto, el docente creer que lo sabe. A la inversa puede suceder lo mismo, el alumno puede saber ese tema y no haber comprendido la pregunta o desconocer ese detalle preguntado y el docente creer que el alumno desconoce totalmente el tema. Obviamente ninguna de estas dos inferencias es correcta.
- Es frecuente observar el uso de abreviaturas y siglas de síntomas, patologías, drogas, etc. Es aconsejable no abusar de las abreviaturas ya que puede ser un factor que confunda al alumno, es decir, que pueden responder mal no por carecer del conocimiento sino por desconocer o confundir las siglas. Conocer y manejar terminología específica de un campo del saber o disciplina es sin duda un objetivo valioso para

docentes y alumnos. Es por esto que debe ser específicamente enseñado y evaluado, con ítem elaborados para tal fin.

- Plantear el examen de tal forma que las preguntas sean completamente independientes de las respuestas a otras preguntas (por ejemplo, que el enunciado de una no induzca la respuesta de otra).
- Eliminar todos los detalles que no tengan relación con la pregunta.
- Evitar, de preferencia, las proposiciones negativas, pero si una locución negativa aparece en el cuerpo de la pregunta, subrayarla para atraer la atención del estudiante.
- Utilizar distractores (una opción que no es la correcta) plausibles o lógicos. Cada distractor debe, por su contenido o naturaleza, ser tal que parezca tener relación con la pregunta, si están demasiado alejados de la pregunta; parecerán ridículos a un estudiante inteligente. Con ello se reduce el número de respuestas posibles y, por lo tanto, la pregunta pierde su importancia y aumenta la posibilidad de azar.
- Cuando se trate de proposiciones de respuestas que comprenden cifras, disponer las respuestas en orden decreciente o creciente.
- Prever la ubicación de la respuesta correcta de tal modo que, para el conjunto de la prueba, la letra correspondiente a la respuesta correcta no aparezca más a menudo que otra. Con frecuencia se observa que la opción correcta se ubica en la opción c).

<b>Preguntas estructuradas</b>	
<b>Posibilidades</b>	<b>Limitaciones</b>
<ul style="list-style-type: none"> <li>• Mayor validez: Facilitan la evaluación de una gran cantidad de objetivos.</li> <li>• Mayor objetividad: Tienen una sola respuesta correcta y no permiten una interpretación personal del alumno al responder ni del docente al corregir.</li> <li>• Permiten evaluar a un gran número de estudiantes en poco tiempo con el mismo examen.</li> <li>• La corrección insume poco tiempo.</li> <li>• Permiten computar los resultados con facilidad lo que simplifica la realización de análisis estadísticos de los ítem.</li> <li>• Permiten la clara explicitación de los criterios de evaluación.</li> <li>• Permiten hacer un análisis de la calidad de la pregunta a través del índice de dificultad y de discriminación.</li> </ul>	<ul style="list-style-type: none"> <li>• Son difíciles de preparar, la mayor dificultad está relacionada con la falta de entrenamiento docente para la confección de las mismas.</li> <li>• Generalmente exploran niveles taxonómicos bajos.</li> <li>• Se puede producir en los estudiantes un "efecto de reconocimiento", lo que permite que aún sin dominar los contenidos pueden identificar la respuesta correcta al verla escrita dentro de la pregunta.</li> </ul>

En los ítem de selección múltiple -"choice" clásicos- siempre encontramos una pregunta y 4 ó 5 opciones de respuestas. Para responder el estudiante toma en consideración y analiza sólo las alternativas presentadas.

Desde hace unos años se está utilizando otro formato de preguntas de opción múltiple conocidas como de "menú largo".

- **Preguntas de Menú largo**

Estas preguntas de "menú largo" se caracterizan por tener un único listado de opciones de respuesta, palabras y/o frases ordenadas alfabéticamente, que incluye las clave y los distractores para todas las preguntas de la prueba. Para cada pregunta están vigentes las 100 ó más opciones del listado.

Si bien por el formato son preguntas de selección múltiple, por el proceso mental que realiza el alumno cuando elabora la respuesta este tipo de ítem constituye una variación de las preguntas, de respuesta breve y/o de completamiento del espacio en blanco (fill in the blank) ya que el estudiante primero "piensa" la respuesta y luego la confirma comprobando que su respuesta se encuentre entre las ofrecidas.

Lógicamente el estudiante no puede considerar en cada pregunta cada una de las 200 opciones. En este tipo de examen el estudiante después de leer la pregunta "piensa" una respuesta, luego busca en el listado único su respuesta pensada e identifica el código que le corresponde y lo copia en la hoja de respuesta -formulario- que puede ser tabulado manualmente o leído por un scanner.

**Ejemplo:**

**Preguntas menú largo.<sup>21</sup>**

*La prueba se aplicó a estudiantes de medicina que iniciaban el curso de patología en segundo año de la carrera. Las 30 preguntas exploran conocimientos de anatomía, bioquímica, fisiología y neurociencias desarrollados en los cursos del año anterior.*

- 1. Los lípidos son secretados por el hígado como . . . . .*
- 2. Los lípidos son transportados desde el sitio de absorción en el intestino hacia el hígado en forma de . . . . .*

<sup>21</sup> Ejemplo tomado de Damjanov I, Fenderson B, Veloski J, Rubin E. Testing of medical students with open-ended, uncued questions. Human pathology. Volume 26 N° 4 april 1995.

**Listado único de 200 opciones (se muestran sólo 82)**

A.1.- Acetilcolina	D.2.- Duodeno	I.2.- Insulina	Q.1.- Quilomicrones
A.2.- Ácido fenilacético	E.1.- Endometrio	I.3.- Islotes pancreáticos	R.1.- Reflejo de Babinski
A.3.- Aldosterona	E.2.- Esófago	L.1.- Laminina	R.2.- Riñón
A.4.- Apolipoproteína	E.3.- Espasticidad muscular	L.2.- LDL	R.3.- Rodopsina
A.5.- Astrocito	F.1.- Ferritina	L.3.- Lipoproteínas de Muy Alta Densidad	S.1.- Sistema Nervioso Central
B.1.- Biliverdina	Ganglios	L.4.- Lipoproteínas de Alta Densidad	S.2.- Sodio
C.1.- Calcio	G.1.- Linfáticos	L.5.- Lipoproteínas de Baja Densidad	T.1.- Testículo
C.2.- Célula beta	G.2.- Axilares	L.6.- Lipoproteínas de Muy Baja Densidad	T.2.- Tiroides
Células dell	G.3.- Inguinales	M.1.- Médula adrenal	T.3.- Tirosina
C.3.- la microglía	G.4.- Mamarios internos	M.2.- Metionina	T.4.- Transferrina
C.4.- la oligodendroglía	G.5.- Mediastinales	M.3.- Miometrio	T.5.- Trastorno autosómico recesivo
C.5.- túbulo proximal renal	G.6.- Periaórticos abdominales	M.4.- Mucus	T.6.- Trastorno poligénico
C.6.- ependimales	G.7.- Peribronquiales	M.5.- Muscarina estearasa	T.7.- Trastorno autosómico dominante
C.7.- epiteliales del glomérulo	G.8.- Perirectales	M.6.- Músculo esquelético	T.8.- Triptofano
C.8.- mesangiales	G.10.-Glucagon	N.1.- Noradrenalina	T.9.- Trompas de Falopio
C.9.-yuxtaglomerulares	H.1.- HDL	O.1.- Ovario	U.1.- Utero
Colágeno	H.2.- Hematina	P.1.- Páncreas	V.1.- VLDL
C.11.- Tipo I	H.3.- Hemoglobina	P.2.- Paratiroides	Y.1.- Yeyuno
C.12.- Tipo II	H.4.- Hígado	P.3.- Potasio	
C.13.- Tipo III	H.5.- Hipercalcemia	P.4.- Purinas	
C.14.- Tipo IV	H.6.- Hiperuricemia		
C.16.- Contracción muscular	H.7.- Hipoalbuminemia		
D.1.- Diabetes insípida	H.8.- Hipótesis		
	I.1.- Illo		

Administración: tiempo asignado para responder: 1 minuto por pregunta.

En el trabajo mencionado<sup>22</sup> se comparan los resultados de dos exámenes de 30 preguntas, uno con formato de selección múltiple clásico y el otro con formato menú largo. En dos grupos de estudiantes equivalentes, se encontró que el examen menú largo discrimina mejor entre los estudiantes que "saben" y los que "no saben". Además con este formato se eliminan los posibles aciertos por azar. Se menciona además que los profesores que elaboraron los exámenes opinaron que era mucho más fácil construir un examen con formato menú largo.

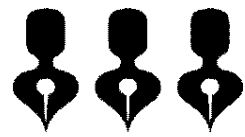
En el trabajo citado se concluye que es recomendable usar este tipo de pruebas menú largo porque son más fáciles de construir, se elimina el factor azar presente cuando ofrece un número limitado de opciones, son totalmente objetivas, presentan un índice de dificultad algo superior a las de selección múltiple clásico, discriminan mejor.

Y se parecen un poco más a la vida real: hay que "recuperar" la información entre los conocimientos adquiridos sin limitarse a considerar 4 ó 5 posibilidades sugeridas.



*Antes de continuar le proponemos un ejercicio para aplicar los conocimientos.*

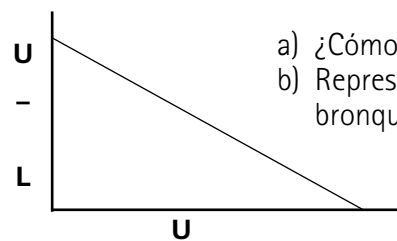
<sup>22</sup> Ibidem.



## EJERCICIO N° 1

**A. Analice los siguientes ítem e identifique a qué categoría taxonómica (información, comprensión, aplicación) corresponden.**

- 1.- Siempre se dice que "los estrógenos preparan a los órganos para que actúe la progesterona"  
¿Qué explicación a nivel molecular puede darse para este efecto de los estrógenos?  
Este ítem corresponde a: .....
- 2.- ¿Qué gas no anestésico, reciben en altas concentraciones casi todos los pacientes sometidos a anestesia general?  
Este ítem corresponde a: .....
- 3.- Lea la siguiente información y responda al cuestionario:  
La intoxicación aguda por paracetamol produce necrosis hepatocelular. Diversos estudios clínicos y experimentales han demostrado que:
  - El paracetamol con fenobarbital aumenta la potencia hepatotóxica del paracetamol
  - El pretratamiento con cimetidina disminuye la potencia hepatotóxica del paracetamol
  - La potencia tóxica del paracetamol es inversamente proporcional a la disponibilidad de glutatión.
  - a) ¿Qué conclusiones pueden extraerse de la interacción entre fenobarbital y paracetamol? ¿Por qué?
  - b) ¿Qué conclusiones pueden extraerse de la interacción entre cimetidina y paracetamol? ¿Por qué?
  - c) ¿Qué conclusiones pueden extraerse de la interacción entre disponibilidad de glutatión y potencia hepatotóxica del paracetamol?
  - d) En base a las tres respuestas anteriores, formule una hipótesis acerca del mecanismo de la hepatotoxicidad del paracetamol.  
Este ítem corresponde a: .....
- 4.- En un estudio de binding se estudiaron los receptores beta-2 adrenérgicos en el músculo liso bronquial en dos grupos de animales: controles y tratados crónicamente con salbutamol (agonista beta-2 adrenérgico)  
En el Gráfico ya está representado el resultado del control.



- a) ¿Cómo se llama este gráfico?
- b) Represente el resultado observado en los músculos lisos bronquiales de los animales tratados con salbutamol.

**Fundamente su dibujo**

Este ítem corresponde a: .....

*Compare sus respuestas con las que figuran en la CLAVE.*



**B. Analice los siguientes ítem, identifique los errores e intente mejorarlos.**

1.- La TBC no es una enfermedad no contagiosa V ( ) F ( )

Error: .....  
Ítem mejorado: .....

2.- La resistencia al sarampión obtenida mediante el uso de la vacuna contra el sarampión no se llama inmunidad activa" V ( ) F ( )

Error: .....  
Ítem mejorado: .....

3.- Mencione los síntomas más comunes de la Fiebre Hemorrágica Argentina.

Error: .....  
Ítem mejorado: .....

4.- Relacione la columna de la izquierda (países) con la de la derecha (capitales).  
Escriba al lado da cada país la letra correspondiente a su capital. Cada letra puede ser usada una vez."

Países	Capitales
1. Argentina	a) Dublin
2. Italia	b) Roma
3. Irlanda	c) Bamaco
4. Mali	d) Buenos Aires
5. Kuwait	e) Al Kuwait

Error: .....  
Ítem mejorado: .....

5.- ¿Cuál es el porcentaje de personas infectadas por el M Tuberculosis que desarrollan enfermedad activa?

- a) 0 a 5%
- b) 5 a 10%
- c) 10 a 15%
- d) 15 a 20%

Error: .....  
Ítem mejorado: .....

6.- "¿Quién cree usted que descubrió la penicilina?"

Error: .....  
Ítem mejorado: .....

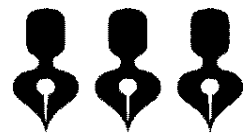
7.- La penicilina fue descubierta por ....."

Error: .....  
Ítem mejorado: .....

8.- "..... fue descubierta por ..... en ....."

Error: .....  
Ítem mejorado: .....





9.- "La presencia de ..... ventricular. .... en un hipertenso es. .... que aumenta. .... "  
Error: .....  
Ítem mejorado: .....

10.- "En pacientes operados de ..... la secuela ..... más frecuente es ..... rama derecha"  
Error: .....  
Ítem mejorado: .....

11.- "Penicilina ..... Fleming"  
Error: .....  
Ítem mejorado: .....

**C. Lea el siguiente texto y redacte preguntas de:**

1. alternativas constantes
2. correspondencia o apareamiento
3. selección múltiple

## Trastornos de la Conducta alimentaria. Bulimia Nerviosa.

Dra. Ana M. Armatta. EDISA. Módulo N° 9. 1997.

### Introducción

La bulimia nerviosa (BN) es un trastorno de la alimentación que implica un desorden afectivo con repercusión orgánica en diferentes órganos y sistemas.

Este trastorno de la alimentación está multideterminado, por factores etiológicos psicobiológicos, familiares y socioculturales.

La bulimia típica comienza en general entre los 17 y 20 años, a menudo, en pacientes con historia de sobrepeso que intentaron adelgazar a veces sometidas a dietas muy rígidas o planes de adelgazamiento rápido. Cuando estos se tornan difíciles y ya no pueden realizar la dieta con la tenacidad de los primeros tiempos, comienzan con alteraciones de la alimentación, pasando por períodos de ayuno muy largos, luego de los cuales se suceden los atracones de comida, episodios típicamente compulsivos, muy difíciles de dominar.

Aquí las ingestas son descontroladas con un alto valor calórico de 2000 o más calorías, pudiendo en algunos casos llegar a la ruptura gástrica.

La selección de alimentos es anárquica y desordenada a predominio de Hidrato de Carbono. El ataque se caracteriza no solamente por el valor calórico de las ingestas sino también por la velocidad con que se come. Comienzan aquí a provocarse el vómito o utilizan otros medios purgativos como laxantes o diuréticos como forma indiscriminada en forma de desembarazarse de lo ingerido.

Todo esto es mantenido en secreto por la paciente ya que se siente avergonzada por lo ocurrido, siente culpa, autodesprecio y entra en un estado de depresión intensa. Otras veces vomitan después de una comida normal, suponiéndose que no pueden tolerar la sensación de plenitud gástrica habitual después de una ingesta normal.

Toda esta problemática con la comida interfiere en los vínculos con las otras personas ya sean familiares o amigos, se tornan solitarias, muchas veces dejan su trabajo o estudio y quedan atrapadas en comer para vomitar y vomitar para después comer.



## Epidemiología

- 1- Relación hombre mujer 1/5.
- 2- Comienzo entre 17 y 20 años.
- 3- Mayor incidencia en las clases medias y altas aunque en la actualidad se extiende a las clases bajas.
- 4- Su aparición es más frecuente en países, occidentales.
- 5- Presencia en la familia de casos de depresión, alcoholismo y desórdenes en la alimentación.
- 6- Aumento de la frecuencia en ciertos grupos especiales.

Es un síndrome caracterizado por episodios de ataques de ingesta de comida (o "binges eating" = atracón) seguidos por vómitos autoinducidos, uso de laxantes o diuréticos, ayunos reparatorios, o actividad física extenuante con el propósito de bajar de peso.

## Etiopatogenia

Como las anorexias, las patologías bulímicas deben ser entendidas desde diversos aspectos como el biológico, el psicológico y el sociocultural, ya que se encuentran fuertemente intrincados:

### • Aspectos biológicos

La teoría neurofisiológica trata de explicar la etiopatogenia de los trastornos de la alimentación a través de la alteración de ciertas sustancias subcorticales, fundamentalmente hipotalámicas que constituyen sistemas relacionados con la regulación de la alimentación.

Hay coincidencia en que los niveles de serotonina cerebral se encuentran disminuidos en esta patología, lo que explicaría las ingestas copiosas que generalmente están constituidas por HC.

Con respecto al eje hipotálamo-hipófiso-gonadal, las pacientes bulímicas están sometidas a un cuadro de malnutrición, con períodos de restricción, atracones y vómitos que explicarían sus alteraciones menstruales (oligomenorrea o períodos de amenorrea).

La mayoría de los autores coinciden en que estas pacientes presentan niveles bajos de hormona folículo estimulante (FSH) y luteinizante (LH) mientras que los valores de prolactina y hormona de crecimiento se hallan normales.

Las pacientes bulímicas a diferencia de las anoréxicas tienen una vida sexual activa, por lo que se hace importante aconsejarlas sobre fertilidad ya que obviamente no conviene que se embaracen mientras dure el comportamiento patológico orgánico y/o psíquico.

El eje hipotálamo-hipófiso-adrenal no está alterado significativamente en la BN. No se describe aumento del cortisol.

El eje hipotálamo-hipófiso-tiroideo puede registrar alguna anormalidad, pero las hormonas T3, T4 y T3 reversa en general son normales.

Hay dos neuropéptidos a los que se les atribuye una gran influencia sobre la conducta alimentaria. Son el neuropéptido Y (NPY) y el neuropéptido YY (NPYY). A estos péptidos se los considera potentes estimulantes endógenos de la alimentación. Ambos son selectivos para la ingesta de carbohidratos.

La colecistoquinina (CCK) es una hormona con actividad generadora de saciedad. Se sugiere que la pérdida del sentimiento de saciedad en pacientes bulímicas estaría relacionado con la disminución de la secreción de la fracción 8-S de la CCK en respuesta a las comidas. El mecanismo principal de acción sería la activación del músculo pilórico, lo cual disminuye el vaciado gástrico.

### • Aspectos psicológicos y socioculturales

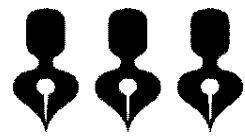
Las situaciones fundantes para la aparición de trastornos de la alimentación se ubican en las primeras épocas de la vida. El ser humano requiere desde su nacimiento de otros que puedan satisfacer sus necesidades. Se nace en un estado de dependencia absoluta, dependencia de un otro que pueda suprimir y calmar los estados de tensión endógena.

La posibilidad de reconocer el propio cuerpo como valioso, de sentirse bien con lo que se es y se tiene, depende primordialmente de estos primeros vínculos con la madre, de sus cuidados, de su ternura y de sus caricias.

Si la madre no satisface suficientemente las necesidades de su hijo, por ausencia o, aún presente y cumpliendo con aspectos básicos como la alimentación o la limpieza, la tarea







se realiza de manera desafectivizada, el niño encontrará en lugar de una madre satisfactoria, un vacío.

Se observa en estas pacientes, una precaria organización defensiva, ante la posibilidad de un derrumbe estructural por esos padecimientos ocurridos en la infancia, y que dan síntomas en la adolescencia.

En un mundo de pautas no suficientemente claras acerca del rol de la mujer; que de ama de casa, ha entrado hace un par de décadas, al mundo competitivo de los negocios y los éxitos profesionales, y en el que un cuerpo cercano al ideal social es indispensable; el profundo cambio en la identidad social no siempre es asimilado con facilidad. Y si no hay una adaptación suficiente a las nuevas exigencias, propulsado por un miedo mórbido a engordar se puede entrar en ese mundo psicopatológico nuevo que constituyen las alteraciones del impulso. Por eso encontramos asociados a la bulimia otros trastornos de la impulsividad como el abuso de drogas o el alcohol o la cleptomanía.

Respecto al tipo de personalidad que serviría de base a los trastornos bulímicos, si bien hay rasgos caracterológicos predominantes, no son suficientes para definir un perfil único. Los rasgos frecuentes son inestabilidad afectiva con predominio de depresión, impulsividad, baja tolerancia a la frustración y a la ansiedad y baja autoestima; pero las personas con trastornos bulímicos no tienen una estructura de personalidad definida.

En cuanto a las familias demuestran algunas pautas de funcionamiento particulares. En los distintos miembros se notan angustia y conflictos en las relaciones interpersonales con interacciones abiertamente hostiles y manipuladoras. A pesar de ello el contenido de lo discutido no está en relación con los conflictos reales que existen dentro del seno de la familia. En estas circunstancias, el bulímico ejerce una función de foco de distracción, permitiendo olvidar temporalmente las discusiones al provocar un sentimiento de unión familiar en pos de la ayuda al enfermo.

Si bien estas características de funcionamiento familiar son frecuentes de advertir, resta saber si este modelo conforma una precondition necesaria para la aparición de la patología o es una consecuencia de ella.

### Manifestaciones Clínicas

- Examen físico: *Puede ser normal o presentar algunos pocos signos y/o síntomas.*

La bulimia puede afectar a todos los órganos de la economía en diferentes grados, aunque el aspecto de la paciente no orienta acerca del diagnóstico.

Tampoco el peso corporal, ya que pueden tener sobrepeso, un peso adecuado o estar adelgazadas.

Son típicas la somnolencia, la sudoración y la debilidad después de los atracones.

Con el vómito autoinducido se pierden calorías cuya cantidad dependerá de la magnitud del vómito, del atracón previo y del tiempo en que los alimentos permanecieron en el estómago. Pero en general las calorías perdidas no son lo suficientemente importantes como para lograr el objetivo con el que las pacientes recurren a él, o sea bajar de peso.

En las pacientes bulímicas de bajo peso el mecanismo más utilizado es la restricción, aunque padezcan atracones y utilicen mecanismos compensatorios.

- Alteraciones hidroelectrolíticas

Las alteraciones metabólicas son muy frecuentes ya que dependen del desbalance hidroelectrolítico producido por los vómitos, y por otros mecanismos compensatorios, que generalmente deplecionan los líquidos del organismo. Por esto ante el cese de las purgas o en la realimentación puede haber repleción de líquidos con edema. Uno de los electrolitos más comprometido es el potasio (K), que puede disminuir por las siguientes causas:

- a- Pérdida por vómitos junto al ácido clorhídrico del estómago.
- b- Por el uso de diuréticos (furosemida-tiazidas).
- c- Pérdida intestinal por el abuso de laxantes que provocan diarreas crónicas.



La pérdida de K, altera la función celular normal que depende de este catión intracelular y repercute especialmente en la función neuromuscular.

La pérdida de ácido clorhídrico puede llevar a la alcalosis hipoclorémica. El riñón por su parte tratará de conservar el K perdiendo Na y complicando aún más el balance metabólico.

A nivel de los músculos esqueléticos puede haber disminución de la fuerza y resistencia muscular, produciendo astenia, calambres y/o parestesias.

El músculo cardíaco se ve afectado dando alteraciones del ritmo; ya que el corazón es especialmente sensible a la hipopotasemia, aumentando la irritabilidad, produciendo esas arritmias que en general son supraventriculares. Dentro de los cambios electrocardiográficos lo más común de observar es la prolongación del intervalo QT

- Aparato digestivo. El tubo digestivo se ve alterado en sus distintas porciones.

En la cavidad bucal puede haber disolución del esmalte dentario (perimilolisis) por erosión ácida debido al pasaje del contenido gástrico; con la aparición de manchas ocreas en los dientes y **fuerte predisposición a las caries**. La gran cantidad de hidratos de carbono ingeridos en los atracones colabora en la producción de las mismas.

El resto del aparato digestivo está sometido a daños por soportar períodos de sobrealimentación, seguido de ayuno, vómitos e ingesta de laxantes, todo lo cual obliga al organismo a adaptaciones constantes.

El esófago no está preparado para soportar el pasaje de ácidos ni las altas presiones generadas por los vómitos (más de 300 mm de Hg); dando lugar, a veces, a esofagitis secundarias, con dolor y ardor retroesternal, y otras veces várices esofágicas o ulceraciones.

El vaciado gástrico está influenciado por la cantidad y la composición de los alimentos, pudiendo presentar náuseas, dispepsias y/o dolor abdominal.

Las pacientes pueden referir vómitos de sangre que puede ser el motivo de consulta por la preocupación que despiertan.

El colon se ve afectado por el abuso de laxantes. Algunas pacientes se quejan de pasar por períodos de constipación a otros de diarrea debida a la alternancia de la ingestas.

Un signo frecuentemente visto como consecuencia de los vómitos autoprovocados es el signo de Russel, que consiste en la erosión y callosidades de la piel de los nudillos de los dedos que utilizan para provocarse el vómito.

Las glándulas salivales se encuentran agrandadas, probablemente por estimulación excesiva.

La secreción salival es responsable del aumento de la amilasemia (y no la pancreática).

- Nefropatías

Pueden aparecer alteraciones renales como consecuencia del abuso de diuréticos, por la pérdida de potasio o por un cuadro de deshidratación, pudiendo encontrarse aumento de la uremia y densidad urinaria baja.

- Aparato respiratorio

Está descripta la neumonitis por aspiración que puede llevar a la muerte. Esta complicación se ve con más frecuencia en pacientes que consumen drogas y/o alcohol, lo que disminuiría el nivel de conciencia durante los vómitos.

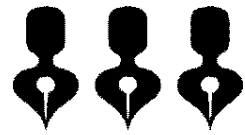
- Sistema Nervioso Central

La repercusión clínica es escasa. Están descriptos cambios en el EEG de sueño como acortamiento de los REM.

- Aparato Psíquico

La paciente bulímica suele llegar a la consulta por decisión propia cuando su estado físico o su estado emocional la llevó a considerar la necesidad de ayuda profesional. En estas condiciones su pedido puede ser genuino y estar dispuesta a comprometerse con un





tratamiento; o solamente desea un alivio sintomático o un esclarecimiento acerca de los riesgos potenciales que tiene su conducta alimentaria.

Estas dos muy distintas razones que motivan la consulta deben ser exploradas y definidas en la primera entrevista porque de ellas depende el destino final, éxito o fracaso del tratamiento.

Los comentarios de la paciente giran, exclusivamente alrededor de la comida, el peso, los valores calóricos de los alimentos, la figura o la belleza. Pareciera que toda ella estuviera en relación estricta con esta temática, fuera de la cuál sólo hay un vacío. Ningún otro tema o preocupación colorea su vida.

A partir de la idea base "terror mórbido a engordar", todos los pensamientos y las emociones se conectarán (Russell).

Respecto a la familia de la paciente bulímica, diversos estudios encuentran mayor incidencia de obesidad en las madres o de bulimia en los familiares cercanos, así como trastornos afectivos, alcoholismo y drogadependencia.

*Redacte preguntas de:*

1. Doble alternativa
2. Correspondencia o apareamiento
3. Selección múltiple

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

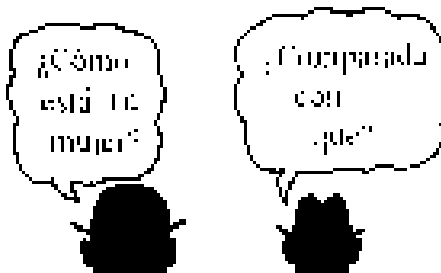
.....



*Se pueden redactar un gran número de preguntas... las que figuran en la CLAVE sólo son algunos ejemplos.*

## Criterios de corrección y de aprobación

Antes de aplicar la prueba, antes de tomar el examen a los alumnos, es necesario tener elaborada la Clave de Corrección.



- *Preguntas de Desarrollo o Ensayo:*

No existirá una patrón de respuestas esperadas; el evaluador tiene libertad para juzgar los resultados en función de lo que él mismo considere correcto o aceptable. Por este motivo es que estas pruebas no son objetivas.

De todos modos es recomendable establecer algunos criterios que se tomarán en cuenta en la corrección y advertirles a los alumnos. Por ejemplo: extensión de la respuesta (100 palabras como mínimo), errores de ortografía, subtemas a incluir: inclusión de información epidemiológica, explicación fisiopatológica, etc...

- *Preguntas de respuesta restringida:*

Para reducir la subjetividad del evaluador es aconsejable elaborar una respuesta modelo para cada pregunta de la prueba, explicitando lo que se espera del alumno y estableciendo un puntaje para cada una de las respuestas.

Recuerde el ejemplo de pregunta semi-estructurada de este módulo.

### **Ejemplo:**

#### *Analice y resuelva la siguiente situación clínica*

Matías concurre al Centro de Salud para un control. Ha asistido desde el mes de edad en forma regular a cada citación. En la actualidad tiene 18 meses, aun continúa con lactancia materna. Su peso actual es de 11.450g, su talla es de 84 cm y su perímetro cefálico es de 47 cm. Controles antropométricos efectuados a los 12 meses registraban un peso de 10.200 g y a los 15 meses de 10.800 g.

*Luego de la presentación del paciente se presentaba el siguiente cuestionario:*

1. *Evalúe el crecimiento. Se adjuntan las gráficas de crecimiento*
  - a. *¿En qué percentilo de peso para la edad se encuentra?*
  - b. *¿Es adecuada la tendencia de la curva? ¿Por qué?*
  - c. *¿En qué percentilo de longitud corporal se encuentra Matías?*
  - d. *¿A qué edad lo cita para el próximo control? ¿Por qué?*
2. *Al revisar la cavidad bucal de Matías, ¿Qué dientes espera encontrar?. Nómbralos*
3. *Mencione 5 habilidades motoras que espera encontrar en su desarrollo neuromotriz.*

4. ¿Cuántas palabras, promedio, puede estar diciendo?
5. Al revisar su carnet de vacunación, ¿qué vacunas espera encontrar?
6. ¿Corresponde alguna indicación en esta edad? ¿cuál?
7. Mencione las indicaciones que le daría a la mamá de Matías en relación con los siguientes tópicos:
  - a. Control de esfínteres
  - b. Estimulación del desarrollo (mencione al menos 4)
  - c. Hábitos y Pautas de alimentación (mencione al menos 4)
  - d. Prevención de accidentes (mencione al menos 6 recomendaciones)

**Clave de Corrección. Puntaje máximo posible: 46 puntos**

<i>A. Evaluación del crecimiento: 10 puntos como máximo</i>		
<i>Pregunta</i>	<i>Respuesta correcta</i>	<i>Puntaje</i>
1 a	Percentilo 50	2
1 b	Si. Aumento de peso constante y gradual	4
1 c	Percentilo 75	2
1 d	A los 24 meses. Crecimiento normal para edad y sexo	2
		10

<i>B. Evaluación del desarrollo: 10 puntos como máximo</i>		
<i>Pregunta</i>	<i>Respuesta correcta</i>	<i>Puntaje</i>
2.	16 dientes: incisivos centrales y laterales, inferiores y superiores, molares anteriores y caninos	2
3.	Corre con rigidez, sube escaleras de la mano, pateo pelotas, hace torres con 4 cubos, imita el trazo vertical	4
4.	10 palabras	4
		10

<i>C. Inmunizaciones: 10 puntos como máximo</i>		
<i>Pregunta</i>	<i>Respuesta correcta</i>	<i>Puntaje</i>
5	BCG (1ª dosis); Cuádruple (3ª dosis); Sabín (3ª dosis); Triple viral (1ª dosis)	5
6	4ª. dosis de Cuádruple y 4ª. dosis de Sabín	5
		10

<i>D. Consejería a la madre: 16 puntos como máximo</i>		
<i>Pregunta</i>	<i>Respuesta correcta</i>	<i>Puntaje</i>
7. a	Puede comenzar con el aprendizaje del control de esfínteres	2
7. b	Dejarlo comer solo, aunque derrame comida Dejarlo jugar con agua, con adecuada supervisión Permitirle que intente desvestirse Contarle cuentos, leerle libros con imágenes, cantarle, recitarle poesías cortas	4
7. c	Continuar con lactancia materna. Indicar que el niño puede comer los alimentos de la dieta familiar, insistiendo en la variedad en la incorporación de carnes de todo tipo, frutas y verduras de diferentes colores. Haciendo hincapié en métodos sanos de elaboración y cocción (evitar fritos y picantes, utilizar poca sal). Reforzar que el niño debe comer en familia, en un ambiente tranquilo y sin interferencias externas (TV). Debe comer con cuchara y solo, respetar sus preferencias y rechazos.	4
7. d	Debe explicar 6 de los siguientes: Quemaduras. Caídas. Intoxicaciones. Ahogamientos. Ahogamientos por inmersión. Heridas. Viaje en automóviles. Electrocutación.	6
		16

- *Preguntas estructuradas:*

Cada ítem o pregunta debe tener identificada la respuesta correcta.

Por ejemplo:

Pregunta 1= a

Pregunta 2= c etc.

Lo más práctico es completar una Hoja de Respuestas con las clave.

Así mismo debe estar establecido el valor asignado a cada respuesta correcta y el puntaje total máximo posible.

Por ejemplo: en una prueba de 80 preguntas, si cada respuesta correcta vale 1 punto, el puntaje máximo posible será 80 puntos. Si en esa misma prueba hubiese 20 preguntas ponderadas cuya respuesta correcta vale 2 puntos, el puntaje máximo posible sería 100 puntos.

Es importante que todas las preguntas de elección múltiple de una prueba tengan el mismo número de opciones. Existe una formula para corregir el factor azar que toma en cuenta el número de opciones.



*En el Anexo 2  
encontrará  
información sobre cómo  
usar puntuación  
corregida*

La corrección electrónica es altamente recomendable porque ahorra tiempo al docente y permite hacer una pronta comunicación de los resultados. Pero más importante aún es que la lectura con scanner carga los datos en la computadora y con un software específico se puede hacer una análisis de la calidad del examen. Esta "evaluación" de la calidad del instrumento de evaluación brinda información muy significativa para la toma de decisiones respecto a la promoción o no de los alumnos.

**Para controlar la subjetividad del evaluador:**

- ▶ Antes de aplicar la prueba asigne el puntaje a cada uno de los ítem.
- ▶ Escriba la respuesta esperada. Intente consensuar la respuesta entre los docentes.

## Diagramación de la prueba

Una vez redactadas las preguntas se debe armar la prueba tomando en cuenta algunas condiciones que influyen en el rendimiento de los estudiantes.

1. Controlar que las preguntas o ítem que se incluyen correspondan a la Tabla de Especificaciones. Es decir a lo que se había planificado en función de los objetivos que se querían evaluar.
2. Ordenar los ítem de manera de asegurar su legibilidad. Numerar las páginas del examen.  
El diseño gráfico de la prueba es importante.  
En las preguntas de elección múltiple:
  - es fundamental que la base y las opciones aparezcan en una misma página
  - son más fáciles de responder si cada una de las opciones está ubicada en una línea diferente (una debajo de la otra).
3. No abusar de abreviaturas y siglas que pueden confundir al examinando.
4. Agrupar los ítem de un mismo formato (verdadero-falso; apareamiento; selección múltiple). Escribir instrucciones específicas para cada formato de ítem. Por ejemplo:
  - Lea los siguientes enunciados y marque, en cada uno, V (verdadero) si es correcto y F (falso) si considera que no es cierto.
  - Establezca la correspondencia entre los conceptos que figuran en la columna de la izquierda y los enunciados que figuran en la columna de la derecha. Coloque las letras que correspondan en el espacio en blanco que figura a continuación de cada concepto. Cada letra puede ser utilizada una, varias o ninguna vez.

En los ítem de correspondencia o apareamiento **no incluir** instrucciones como "unir con flechas" porque dificulta y demora muchísimo la corrección.

5. Dentro del mismo formato de ítem agrupar aquellos que se relacionan con un mismo tema.
6. Ordenar los ítem de tal manera que el grado de dificultad vaya de lo fácil a lo difícil y numerar correlativamente las preguntas.
7. Controlar que una pregunta no proporcione pistas para contestar otra pregunta.
8. Verificar que la clave está ubicada en distintas posiciones. En las pruebas que tienen preguntas de selección múltiple se advierte una tendencia a ubicar la clave en el segundo y tercer lugar (clave en b) y c). Se debe revisar y asegurar que la clave se ubique en todas las posiciones por igual.
9. Es importante que todas las preguntas de elección múltiple de una prueba tengan el mismo número de opciones para aplicar

la formula para corregir el factor azar. Antes de aplicarse la prueba debe haberse decidido si en la corrección de la prueba se aplicará *puntuación corregida*<sup>24</sup> y debe anunciarse a los estudiantes.

10. Diseñar una hoja de respuestas. El alumno responde en el cuadernillo de preguntas y luego copia sus respuestas en una hoja aparte. La hoja de respuestas facilita y acelera el proceso de corrección. Para la corrección electrónica es necesario utilizar formularios inteligentes como "Hoja de Respuestas".
11. Reproducir tantos ejemplares como alumnos vayan a rendir la prueba.
12. En caso que tener que aplicar la prueba a un grupo numeroso de estudiantes y que las condiciones del aula o salón faciliten la copia, es recomendable preparar distintas "versiones" o "temas" del mismo examen. Se pueden construir pruebas paralelas manteniendo las mismas preguntas pero ubicándolas en distinto lugar y cambiando la clave de posición.
13. Escribir las Instrucciones Generales.
  - Debe explicitarse el **tiempo** disponible para responder. Se ha verificado que un alumno necesita entre 45 y 100 segundos para responder un ítem de selección múltiple dependiendo del grado de dificultad del ítem. Cada ítem de verdadero-falso demanda unos 30 segundos. En la práctica se calcula un minuto por pregunta y un tiempo adicional. Por ejemplo: una prueba de 100 ítem puede ser respondida en 100 minutos, mas un tiempo adicional de 20 minutos hace que el tiempo adecuado para completar el examen sea de 120 minutos (2 horas.)
  - Debe informarse el **puntaje asignado** a cada pregunta, si todas valen igual y destacar qué preguntas valen más, en caso de haber ponderado alguna.
  - Debe explicarse la forma de registrar las respuestas sobretodo cuando se utiliza Hoja de Respuestas.
  - Debe informarse sobre el nivel de exigencia establecido para la aprobación del examen.

La diagramación de una prueba influye en el rendimiento del estudiante. Con frecuencia una mala respuesta no corresponde a falta de conocimientos sino que se origina en instrucciones o indicaciones confusas.



*En el Anexo 3 figuran ejemplos de Hojas de Respuesta. y en el Anexo 4 puede ver un ejemplo de Instrucciones Generales.*



<sup>24</sup> Thorndike R Hagen E. (1996) Medición y evaluación en Psicología y Educación Trillas México. pp. 244-245



**Ud. ya sabe que la planificación de una prueba implica:**

- revisar los objetivos y seleccionar aquellos que va evaluar
- confeccionar la tabla de especificaciones
- redactar preguntas
- diagramar el examen



*¡Podemos continuar!*

## **Aplicación y corrección de la prueba**

### **Tomar el examen**

La mayoría de las pruebas escritas son administradas con modalidad presencial, es decir, los estudiantes concurren un día a una hora determinada a "rendir" el examen.

También se pueden tomar pruebas escritas no presenciales. En este caso los estudiantes tienen el examen en su poder y un tiempo asignado para responderlo y entregarlo al docente. Estos exámenes se utilizan con frecuencia en los cursos de educación a distancia. Los participantes reciben el examen por correo postal o electrónico y lo resuelven en tiempos que varían desde un par de semanas a dos o tres meses. De hecho, son pruebas a libro abierto cuyo propósito más importante es que se revise la bibliografía, se discuta la información con otros colegas, se consulte con especialistas, etc. El tipo de ítem que se incluye en los exámenes no presenciales puede ser variado, en general como los sistemas de Educación a Distancia tienen un gran número de participantes, se los confecciona con ítem estructurados de selección múltiple, con hoja de respuesta que es leída electrónicamente. También se pueden presentar problemas semiestructurados para que los estudiantes realicen búsquedas bibliográficas para responderlos y entregarlos, por ejemplo a las 48 horas.

En el momento de tomar la prueba, es necesario tomar algunas precauciones relacionadas con problemas frecuentes de la práctica.

.....  
***¿Cuál es su práctica en la "toma de exámenes"?***

► **¿Ha implementado exámenes no presenciales a libro abierto?  
¿Cuál es su opinión al respecto?**

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

**Si Ud. toma exámenes presenciales:**

Identifique en qué medida le son familiares las siguientes situaciones:	Casi siempre me sucede	A veces me sucede	Nunca me ha pasado
El día del examen escrito, como nunca antes, llegan todos temprano y me reciben sentados en las últimas filas, todos amontonados.			
Durante la primera media hora están en silencio y luego se comienzan a levantar a preguntarme cosas. Al rato la mitad está amontonada alrededor mío.			
En algún momento comienza el murmullo entre ellos y tengo que enojarme para que dejen de hablar.			
Noto que tratan de copiarse, algunos miran la hoja del compañero, otros directamente hablan.			
Generalmente no les alcanza el tiempo que tienen para terminar el examen.			
Cuando es la hora de finalización del examen retiro las hojas.			
Si no terminaron los espero, aunque sean solo 2 o 3 hasta que ellos quieran entregar.			
Generalmente utilizan poco tiempo para responder mis exámenes escritos.			
Me aburro cuando tomo exámenes escritos.			
Si tomo el examen con otro docente, se me hace más liviana la tarea ya que en algún momento nos ponemos a conversar.			
Cuando un alumno se levanta a preguntarme una duda no le contesto			
Cuando un alumno se levanta a preguntarme una duda, si ha sido un "buen alumno" o si me doy cuenta que tiene la "idea" de la respuesta lo ayudo, lo trato de orientar			

- ▶ Puede anotar otras situaciones que se le presentan cuando toma exámenes escritos

.....

.....

- ▶ Recuerde el último examen que rindió como alumno o participante de un curso. ¿Era escrito? ¿Cómo se comportó Ud.? Su conducta como alumno ¿se parecía a la de sus estudiantes? Anote en que se parecía y en que era diferente.

.....

.....



En general el momento de rendir un examen es para el alumno un momento estresante, casi independientemente de si ha estudiado mucho, poco o nada. Se los nota nerviosos, movedizos, inquietos.

Por otra parte, en los exámenes estructurados (de doble alternativa, completamiento, selección múltiple) es muy fácil "copiarse", espiar la respuesta que colocó el compañero de al lado, a veces por desconocimiento y otras sólo para "comprobar" qué puso y comparar.

Muchos de los problemas relacionados con la administración de la prueba pueden ser resueltos tomando algunas precauciones.

### **Sugerencias**

- Revise las fotocopias antes del día y hora del examen, revise si los juegos están correctamente armados, si no sobra o falta ninguna hoja, si están todas al derecho. Si las preguntas han quedado correctamente numeradas, etc. Todos estos detalles intervienen en el clima que se debe lograr para responder una evaluación.
- Cuando se toman pruebas estructuradas es conveniente preparar más de una versión del instrumento (por ejemplo Versión A, B). Se pueden redactar el doble de preguntas siempre y cuando sean equivalentes o, lo que es más fácil para el docente, modificar el orden de las preguntas y el orden de los distractores. En el momento de repartir los exámenes a los estudiantes intercalar las 2, 3 o 4 versiones.
- Calcular el tiempo que les llevará contestar el examen. En las pruebas estructuradas se calcula 1 minuto para leer y responder cada pregunta. Está comprobado que si el tiempo es excesivo y siguen con el examen en la mano, revisando y cambiando las respuestas se cometen más errores que si lo leen, contestan y revisan una sola vez.
- Explique, antes de repartir las fotocopias, cómo deben contestar: pueden marcar la opción elegida en la misma pregunta o deberán luego pasarla a la hoja de respuestas. Indique qué deben hacer si se equivocan en las marcas, cómo salvar los errores, etc.
- Establecer el encuadre la clase anterior o antes de repartir los exámenes.
  - Para evitar la tentación de revisar algún texto o apunte se puede solicitar que dejen todos los libros, carpetas, etc. en algún lugar alejado (en el frente, al costado de cada fila, etc.)
  - Explícite cómo quiere que se ubiquen los estudiantes, dependerá por supuesto del tipo de aula en la que esté evaluando. Comience a repartir el examen cuando crea que están ubicados correctamente, de lo contrario se producirá un clima de movimiento, ruido y confusión que no facilitará la concentración que necesitan los estudiantes para responder adecuadamente.
  - Es conveniente que los estudiantes tengan claro que no se aceptarán conversaciones en el momento del examen. Especificar claramente que el examen es individual y cómo se procederá en el caso que esta norma no se respete: se le sacará la hoja, se corregirá hasta donde ha llegado o estará aplazado sin corregir lo que ya fue respondido, etc.
  - Ud. podrá decidir si se podrán hacer preguntas o no, si contestará a las mismas, si podrán levantarse de los asientos para preguntar o deberán levantar la mano y Ud. se acercará hasta ellos. En general, se dice que la comprensión del ítem o la pregunta es parte de la evaluación del estudiante. Muchos alumnos llaman al profesor y le dicen "no entiendo lo que se quiere preguntar acá". Difícilmente el profesor, en ese momento, se podrá dar cuenta si no entiende, no estudio o qué cosa le está ocurriendo al alumno. También puede ocurrir que la pregunta esté confusa o mal formulada. Es importante aclarar que si en el momento de la corrección, se detecta algún problema en la formulación de un ítem, el

mismo será anulado y no contará en el puntaje global. Otros estudiantes se levantan y le preguntan al profesor "¿Está bien esta respuesta?". Asegúrese que entiendan que Ud. va a corregir el examen en otro momento. La mejor opción es no permitir ni contestar preguntas en el momento de la evaluación.

- Informar cuánto tiempo tienen para responder y avisar 15 minutos antes de finalizar el tiempo. Es conveniente respetar el tiempo asignado para el examen. La única excepción podrá ser que en el momento en que está tomando el examen se de cuenta que el mismo está resultando muy difícil y muchos estudiantes no llegaron al final en el tiempo establecido.
- Respete el encuadre con todos los alumnos. Trate a todos por igual, un alumno que ha sido flojo durante la cursada puede haberse esmerado mucho para preparar este examen y puede pasar que un alumno destacado haya pensado que el examen no le traería demasiadas dificultades y haber dedicado poco tiempo y esfuerzo al estudio.

---

## Criterios de aprobación. Nivel de exigencia.

La decisión que los profesores toman más frecuentemente en base a los resultados obtenidos en los exámenes es la promoción o no de los alumnos.

¿Cuántas preguntas deben ser respondidas correctamente para aprobar el examen?

Existen dos criterios para decidir el nivel de exigencia para la aprobación de un examen.

- *Criterio absoluto:*

El rendimiento de cada uno de los alumnos se compara con un estándar deseable preestablecido. Teóricamente el estándar lo constituyen los objetivos de aprendizaje establecidos como "resultados esperados del aprendizaje" y tomados como base para la planificación del examen.

¿Son todos los objetivos esenciales? ¿Qué porcentaje de "desconocimiento" o equivocación podría admitirse? ¿Cuál es el mínimo indispensable?

En respuesta a estas preguntas es que los profesores fijan el nivel de aprobación: por ejemplo 60% de las preguntas respondidas correctamente.

---

### **Ejemplo:**

*Si administramos una prueba de 80 preguntas con un nivel de exigencia de 60%, y cada pregunta vale un punto, el puntaje máximo posible es de 80 puntos y el puntaje mínimo para aprobar es 48 puntos. Todos los alumnos que obtengan 48 puntos y más estarán aprobados y todos aquellos que obtengan menos de 48 estarán aplazados en el examen. Este criterio se denomina "criterio absoluto".*

---

- **Criterio relativo:**

El puntaje necesario para aprobar se establece después de corregir todas las pruebas y conociendo el puntaje más alto obtenido. El nivel de exigencia para aprobar se relaciona con el puntaje obtenido por el "mejor" alumno. Por ejemplo 60% de respuestas correctas del máximo respondido por el mejor alumno.

**Ejemplo:**

*Si en la misma circunstancia anterior, se adopta un "criterio relativo" el puntaje necesario para aprobar dependerá de la cantidad de respuestas correctas que obtenga el "mejor estudiante". Si el mejor alumno de ese grupo obtiene, por ejemplo 70 puntos para aprobar el examen serán necesarios 42 puntos (60% de 70), si el mejor alumno responde bien sólo a 57 preguntas, el nivel de aprobación bajará a 34 respuestas correctas (60% de 57).*

Se trata de un criterio relativo porque depende de los resultados que se obtengan en cada oportunidad. Es decir se calcula el nivel de aprobación de acuerdo al mejor puntaje obtenido, en ese grupo, en esa oportunidad.

**Ejemplo:**

*El siguiente cuadro nos muestra cómo varía la cantidad de alumnos que resultan aplazados, según un criterio u otro.*

*Se trata de un examen final de un curso de post-grado.*

*Nº de profesionales que rinden el examen: 34*

*Tipo de prueba: Escrita, estructurada, 50 ítem de selección múltiple.*

*Puntaje máximo posible: 50. Puntaje máximo obtenido: 37. Puntaje mínimo obtenido: 19 puntos*

*Mediana: 28*

Puntajes obtenidos:	37	<b>Criterio absoluto:</b> 11 alumnos aprobados con el nivel de exigencia del 60% del puntaje máximo posible (30 ítem correctos)	
	35		
	34		
	34		
	34		
	32		
	32		
	31		
	31		
	31		
	30		NIVEL DE APROBACIÓN CON CRITERIO ABSOLUTO
	30		
	29		
	29		
	29		
	29		
	29		
	28		
	28		
	28		
	28		
	28		
	28		
	28		
	28		
	27	<b>Criterio Relativo:</b> 32 alumnos aprobados con el nivel de exigencia establecido 60% del máximo obtenido (23 ítem correctos)	
	27		
	26		
	26		
	26		
	26		
	25		
	24		NIVEL DE APROBACION CON CRITERIO RELATIVO
	23		
	22		
	20		
	19		

**El criterio absoluto siempre resulta más exigente que el relativo.**

En general los mejores estudiantes suelen responder bien entre un 85 y un 90 % de la prueba. En el ejemplo se advierte además, que el "mejor alumno" respondió correctamente sólo 37 de las 50 preguntas.

En ambos casos es importante determinar cuál es el puntaje máximo posible: después de corregida la prueba se debe hacer un análisis de la calidad de cada una de las preguntas. Si alguna/s pregunta/s resulta/n muy deficientes deben ser anuladas. Si un examen tiene 80 preguntas y dos de ellas resultan deficientes y son anuladas el puntaje máximo posible pasará a ser 78.

¿Cómo analizar la calidad de las preguntas?

### **Análisis de ítem**

Muchas veces cuando el docente observa los resultados de los exámenes suele pensar:

*"Tantos alumnos aplazados cuando sólo se exige un mínimo de conocimientos".*

Generalmente calma su preocupación con otro pensamiento:

*"¡Los alumnos cada vez estudian menos!"*

*"¡Con este tipo de exámenes se presentan a ver si aprueban por azar!"*

Pero antes de llegar a esta conclusión debería revisar el examen mismo para controlar la calidad de las preguntas.

- ¿Resultaron preguntas difíciles?
- ¿Realmente los estudiantes que más sabían contestaron bien las preguntas más difíciles?
- ¿Efectivamente este examen permite identificar a los que no saben?
- ¿Los puntajes obtenidos se deben al azar?

Para responder estas inquietudes el docente puede analizar cada una de las preguntas en función de las respuestas de sus alumnos.

¿Qué otra información se puede obtener analizando las respuestas y qué decisiones se pueden tomar?

El análisis de los resultados de un examen de rendimiento preparado por el profesor brinda información sobre la calidad del instrumento de evaluación. Estos análisis se pueden realizar a partir de un instrumento administrado a por lo menos 30 alumnos.

El análisis de las respuestas dadas por los alumnos frente a cada uno de los ítem suministra información sobre:

- ▶ Grado de dificultad del ítem. La dificultad o facilidad de una pregunta está dada por la proporción de alumnos que responden correctamente. Lo que permite identificar el logro de los objetivos indicados en la tabla de especificaciones. Se identifica qué objetivos y/o unidades temáticas no son dominados por los estudiantes (ítem que tienen muchas respuestas incorrectas).

- El poder de discriminación del ítem. Es decir, la posibilidad que tiene de distinguir entre los alumnos "que saben" y los que "no saben".
- La calidad de cada una de las opciones de respuestas ofrecidas.

• **Índice de dificultad de un ítem**

Se puede establecer si una pregunta es fácil o difícil en función de cómo la respondieron los alumnos.

Así establecemos que un ítem es<sup>25</sup>:

Ítem: x	Porcentaje de alumnos que lo respondió correctamente
Muy fácil:	100% a 85%
Relativamente fácil:	84% a 69%
Dificultad media:	68% a 32%
Relativamente difícil:	31% a 16%
Muy difícil:	15% a 0%

**El índice de dificultad expresa:** la proporción de sujetos que respondió correctamente un ítem. A mayor cantidad de sujetos que responden correctamente un ítem, menor dificultad del mismo.

La fórmula matemática que se utiliza para obtener el índice de dificultad es:

$$p: (\text{índice de dificultad}) = \frac{c}{n} \times 100$$

c: cantidad de respuestas correctas

n: número total de sujetos que rindieron la prueba.

**Ejemplo:**

*En un examen de Pediatría<sup>26</sup>, que fue contestado por 5.690 alumnos, se incluyó la siguiente pregunta:*

*Los siguientes, excepto uno, son hallazgos frecuentes en el RN con sepsis bacteriana. Señale la EXCEPCIÓN.*

- a) *Disminución del recuento total de plaquetas (<150.000/mm<sup>3</sup>)*
- b) *Disminución de la PCR (Clave)*
- c) *Aumento de la relación entre neutrófilos en cayado y neutrófilos totales*
- d) *Acidosis láctica*

*5.653 profesionales respondieron esta pregunta correctamente, es decir que marcaron la opción b). Tomando en cuenta que el 99,35% de los alumnos la respondió bien, puede afirmarse que se trata de una pregunta muy fácil.*

$$P = \frac{5.653}{5690} \times 100 = 99,35\%$$

<sup>25</sup> Modificado de Guilbert JJ (1981) Guía pedagógica para el personal de salud. OMS / OPS.

<sup>26</sup> Examen de PRONAP 2001

- **Índice de discriminación**

Si un determinado ítem es muy difícil, los pocos alumnos que lo respondieron correctamente ¿son los mejores alumnos?; ¿son los que tienen más conocimientos? Para este tipo de análisis se considera que los alumnos que tienen más conocimiento o que son "los mejores" de ese grupo, son los que obtuvieron en ese examen los puntajes más altos.

Para averiguar el Índice de discriminación de un ítem lo primero que se debe hacer es ordenar los alumnos según los puntajes obtenidos de mayor a menor y establecer dos grupos:

**Grupo superior:** es el 27% de alumnos que obtuvo en la prueba los puntajes más altos.

**Grupo inferior:** es el 27% de los alumnos que obtuvo en la misma prueba los puntajes más bajos.

Si un ítem ha sido contestado correctamente por la mayoría de los alumnos del grupo superior e incorrectamente por la mayoría de los alumnos del grupo inferior, se determina que este ítem tiene una **discriminación positiva (discrimina bien)**.

En el caso que un ítem sea respondido correctamente por un número similar de alumnos de rendimiento alto y bajo, se dice que el ítem tiene discriminación nula (no discrimina entre alumnos de alto y bajo rendimiento total).

Si un ítem es respondido correctamente por la mayoría de los alumnos con puntajes totales bajos y al mismo tiempo lo han respondido mal los alumnos con puntajes totales altos se dice que el ítem tiene **discriminación negativa (discrimina mal)**.

La fórmula para hallar el índice de discriminación es:

$$I.D. (\text{índice de discriminación}) = \frac{C_s - C_i}{\frac{n}{2}} \times 100$$

Cs: cantidad de respuestas correctas del grupo superior

Ci: cantidad de respuestas correctas del grupo inferior

n: (grupo superior + grupo inferior)

El índice de discriminación se interpreta de acuerdo a la siguiente tabla de uso convencional<sup>27</sup>

Valor	El ítem es:
40 y más	muy bueno
39 a 30	bueno a mejorar
29 a 20	regular, debe mejorarse
19 y menos	deficiente, debe descartarse o revisarse cuidadosamente

<sup>27</sup> Guilbert JJ, Ob. Cit.



### **Ejemplo:**

En un examen de Cardiología se incluyó la siguiente pregunta:

¿Cuál de las siguientes variables hemodinámicas expresa con mayor fidelidad la contractilidad miocárdica?

- a) volumen de fin de sístole
- b) índice cardíaco
- c) índice de trabajo sistólico
- d) fracción de eyección (Clave)

De los 34 profesionales que respondieron esta pregunta sólo 9 lo hicieron correctamente.

$$\text{El Índice de Dificultad. } P = \frac{9 \times 100}{34} = 26\%$$

Tomando en cuenta que sólo el 26% de los alumnos la respondió bien puede afirmarse que se trata de una pregunta relativamente difícil.

Para identificar la capacidad de discriminación de esta pregunta se ordenan los 34 puntajes de mayor a menor estableciendo el Grupo Superior y el Grupo Inferior con 9 alumnos cada uno.

Este ítem fue respondido de la siguiente manera:

Alumnos	Tipo de respuesta
Grupo Superior: 9 alumnos	8 respuestas correctas 1 respuesta incorrecta
Grupo inferior: 9 alumnos	2 respuestas correctas 7 respuesta incorrecta

$$\text{Aplicando la fórmula obtenemos: } \frac{8 - 2}{9} \times 100 = 66,66$$

Es decir que esta pregunta es muy buena para diferenciar a los que no saben de los que sí saben.

#### • **Distribución de las repuestas**

Existe otro análisis para profundizar el estudio de la calidad de cada ítem.

Es el **análisis cualitativo** que se hace en función de la cantidad de elecciones recibidas por cada una de las opciones propuestas. Se trata de analizar cómo se han distribuido las respuestas entre las diferentes opciones de cada pregunta. Si las opciones son homogéneas todas deben resultar atractivas y un cierto número de estudiantes debería elegir las.

## Ejemplo<sup>28</sup>:

Un varón de 3 meses de vida es llevado a su consultorio por dificultades en la alimentación y mal progreso de peso. Se trata de un niño nacido de término, con peso adecuado y sin complicaciones perinatológicas. La madre refiere que cuando lo prende al pecho succiona débilmente y que ha tenido varios episodios de tos y ahogo durante la alimentación. Sólo aumentó 500 gramos desde el nacimiento, aún no tiene sonrisa social ni línea media, es francamente hipotónico y presenta criptorquidia bilateral. El resto del examen es normal. Para iniciar el estudio de este paciente se debe solicitar:

- a) Interconsulta con genética. (Clave)
- b) Dosaje de enzimas musculares.
- c) Electromiograma y velocidad de conducción.
- d) Biopsia muscular y de nervio periférico.

Esta pregunta fue respondida por 5.690 personas con la siguiente distribución de respuestas:

opción	cantidad de personas	%
a)	3.232	56,80
b)	2.064	36,27
c)	336	5,91
d)	39	0,69
No contesta	19	0,33

El análisis de esta distribución entre las opciones informa que más de la mitad de los alumnos han identificado la respuesta correcta mientras que más del 40% han encontrado como respuesta posible otras opciones incorrectas.

La opción d) no ha funcionado como un buen distractor ya que ha sido elegido por muy pocos alumnos por lo que debe ser revisada. Y la opción b) ha sido elegida por un porcentaje significativo de participantes. ¿Podría ser también correcta?

En todos los casos que aparece este tipo de distribución entre las opciones es necesario revisar la pregunta. Si ambas opciones son correctas la pregunta debe ser anulada.

Muchas veces al hacer este análisis que expresa los errores frecuentes de los estudiantes, los docentes se dan cuenta que el ítem estaba mal formulado. Es generalmente, el caso en que la distribución se divide en forma pareja entre dos opciones. Tal vez averigüen que no tenía una respuesta unívoca, ya que las dos pueden ser correctas, en cuyo caso habrá que anular el ítem.

Tal vez el ítem era unívoco pero las respuestas equivocadas revelen, para ese grupo de alumnos un error conceptual que será necesario trabajar.

Si han elegido todas las opciones en forma pareja probablemente los alumnos desconocen el tema en profundidad.

<sup>28</sup> Este y los siguientes ejemplos han sido tomados de PRONAP. Examen final. 2001

### ¿Qué conclusiones puede sacar un profesor al obtener los índices de dificultad y discriminación?

- ▶ Los ítem muy difíciles están indicando que los alumnos no han logrado los objetivos evaluados.  
¿Fue bueno el ítem? ¿Estaba bien construido, tenía una única respuesta correcta, etc.?. Si está seguro de que el dato es confiable, entonces el docente deberá preguntarse:  
¿Por qué los alumnos no aprendieron?  
¿Por qué las experiencias de aprendizaje implementadas no fueron eficaces?  
¿Qué modificaciones se deben introducir en la planificación de la enseñanza?
- ▶ Los ítem que han sido respondidos correctamente por la mayoría de los alumnos: ¿son obvios, se sugiere la respuesta en la base del ítem, estaban bien construidos o dichos ítem exploran objetivos esenciales?
- ▶ ¿Cuáles son los alumnos que respondieron bien y cuáles son los que se equivocaron? ¿Cómo discrimina el ítem?
- ▶ ¿Qué se debe hacer ante un ítem deficiente?  
Se deberá revisar cuidadosamente el ítem ya que probablemente haya que anularlo. Si el ítem es anulado se modifica, entonces el puntaje máximo posible.

Este tipo de análisis es importante en aquellas materias o departamentos que tienen un banco de ítem para las evaluaciones de los estudiantes. Sería conveniente guardar en ese banco aquellos ítem que fueron probados y que han superado exitosamente estos análisis.

Generalmente, estos análisis resultan complicados para el docente, esta tarea se ve facilitada en gran medida con la utilización de software que se aplican luego de la corrección electrónica de la pruebas. Los software proveen una cantidad importante de información que luego se debe interpretar.

**Los ítem muy fáciles y los muy difíciles requieren un análisis particular para identificar las causas de su extrema facilidad o dificultad para decidir si deben ser descartados ya que, en la práctica no permiten discriminar entre los alumnos que saben y los que no saben.**



## EJERCICIO N° 2

### A. Responda las siguientes consignas.

- En el último examen que ha tomado el profesor Hacha se advierte que los ítem n° 1, 47 y 79 presentan un índice de dificultad del 11%. Estos ítem han resultado:
  - a) muy fáciles
  - b) relativamente fáciles
  - c) con dificultad media
  - d) relativamente difíciles
  - e) muy difíciles
- ¿Qué se debería hacer con estos ítem? ¿Se puede tomar alguna decisión con este dato o habría que buscar más información? ¿Cuál?

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

- En ese mismo examen se ha calculado el índice de discriminación para cada ítem. El ítem n° 47 presenta un índice de discriminación de 40. Este ítem es:
  - a) muy bueno
  - b) bueno
  - c) regular
  - d) deficiente
- ¿Qué interpretación, en función de los dos índices que presenta, puede realizar del ítem n° 47. ¿Qué se debería hacer con ese ítem?

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

### B. Analice las siguientes preguntas y calcule qué grado de dificultad presentan:

Todas las preguntas han sido respondidas por 5690 personas. El examen fue en domicilio, a libro abierto.

1. La madre de un bebé de 9 meses consulta preocupada porque el niño, desde el día anterior, no se sienta ni sostiene la cabeza. Ud. lo examina y lo encuentra llorón, irritable, con debilidad cerebral generalizada. Constata al intentar sentarlo que la cabeza cae hacia delante. La madre sostiene que se trata de un niño sano y el único antecedente es una diarrea de 3 días de evolución. El niño se encuentra normohidratado. ¿Qué estudio se debe solicitar, en primer término?.
  - a) Hemograma con recuento de plaquetas y una función renal.
  - b) Estado ácido-base y un ionograma. (Clave)
  - c) TAC cerebral con y sin contraste, de urgencia.
  - d) RMN de cerebro y médula en forma urgente.





Datos de la respuesta		%	De acuerdo a sus cálculos, ¿Cuál es el índice de dificultad de esta pregunta?
Opción	cantidad de personas		
a)	876		a) Muy fácil ( )
b)	2969 (Clave)		b) Relativamente fácil ( )
c)	740		c) Dificultad media ( )
d)	1076		d) Relativamente difícil ( )
No contesta	29		e) Muy difícil ( )

2. Gladys trae a la consulta a Carolina, que esta en 3° grado, porque ayer recibió un pelotazo en el pecho durante el recreo y presenta dolor e hinchazón sobre laaréola derecha. Ud. revisa a Carolina, tiene 8 años y 7 meses, y pesa 32,5 Kg y mide 137 cm. Al examen nota un nódulo palpable y móvil subareolar de 1 x 1,5 cm, sin cambio de color en la piel, sin flogosis que produce incomodidad más que dolor. ¿Que conducta se debe asumir?
- Interconsulta con traumatología por osteocondritis traumática
  - Interconsulta con cirugía por probable hematoma
  - Analgesia con ibuprofeno a 40 mg/kg/día por vía oral
  - Tranquilizar a la madres ya que se trata del inicio del desarrollo sexual (clave)

Datos de la respuesta		%	De acuerdo a sus cálculos, ¿Cuál es el índice de dificultad de esta pregunta?
Opción	cantidad de personas		
a)	8		a) Muy fácil ( )
b)	27		b) Relativamente fácil ( )
c)	228		c) Dificultad media ( )
d)	5423 (Clave)		d) Relativamente difícil ( )
No contesta	4		e) Muy difícil ( )

3. Identifique cuál de los siguientes signos y síntomas, es característico de intoxicación aguda con nicotina y suele aparecer cuando un individuo comienza a fumar.
- Sequedad de boca
  - Hipertensión arterial
  - Miosis. (Clave)
  - Retención urinaria

Datos de la respuesta		%	De acuerdo a sus cálculos, ¿Cuál es el índice de dificultad de esta pregunta?
Opción	cantidad de personas		
a)	155		a) Muy fácil ( )
b)	40		b) Relativamente fácil ( )
c)	5446 (Clave)		c) Dificultad media ( )
d)	46		d) Relativamente difícil ( )
No contesta	3		e) Muy difícil ( )

¿Qué análisis cualitativo podría realizar de cada una de las preguntas?

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



C. Analice la información presentada en el siguiente cuadro y determine el Índice de discriminación.<sup>29</sup>

Alumno	Puntaje total	Respuesta a la Pregunta
A	90	C
B	62	C
X	80	I
O	51	C
R	48	C
Z	27	I
Y	72	C
G	60	C
H	49	I
K	40	I
M	51	I
N	35	C
Ñ	57	C
C	78	C
J	56	C
P	30	I
D	33	I
E	94	C

C = Correcta

I = Incorrecta

I.D. =

Este ítem es: .....



*Controle sus respuestas con las que  
figuran en la Clave.*

<sup>29</sup> Tomado de: Caffaro de Hernández, N. Análisis Cuantitativo de Ítem. FEPAFEM. Documento interno.

*discriminación*

## Adjudicación de notas. Sistemas de calificaciones.

Generalmente para informar al estudiante, a los padres y a la sociedad toda sobre el rendimiento de los estudiantes se utilizan "las notas". Y las notas fundamentan decisiones importantes: el promedio de la carrera suele tomarse como uno de los criterios para la admisión al sistema de residencias, para otorgar una beca, etc.

Existen distintos Sistemas de Calificaciones que varían en la cantidad de "categorías o clases" en las que se pueden ubicar a los alumnos según su rendimiento.

- Sólo dos categorías: insuficiente - suficiente; aprobado - desaprobado
- Tres categorías: insuficiente - suficiente - distinguido
- Cinco categorías: A-B-C-D-E
- Diez categorías: escala decimal, de 0 a 10 .

Ejemplos:

Excelente	A
Muy bueno	B
Bueno- termino medio	C
Regular	D
Deficiente-malo	E

Con cierta frecuencia se combina la escala decimal con categorías cualitativas y resulta la siguiente combinación.

Sobresaliente	10
Distinguido	9
Muy bueno	7-8
Bueno	5-6
Regular	4
Aplazado	1-2-3
Reprobado	0



El sistema de calificaciones escolares es una convención, un código que está muy lejos de expresar con exactitud "una medida del conocimiento".

Chile califica con una escala del 0 al 7; Argentina con una escala del 0 al 10 y Venezuela utiliza del 0 al 21. Y la nota que significa "aprobado" varía en cada caso: Venezuela aprueba con el 7 que es el valor más alto de la escala chilena. Argentina aprueba con 4 que en Venezuela significa aplazado.

Los sistemas de calificaciones son convenciones y representan ciertas tradiciones.



*"Como cualquier otro aspecto profundamente implicado en la cultura, los procedimientos de calificación y las calificaciones a menudo se han considerado cosas ya hechas y sabidas, con un mínimo de análisis racional acerca de su naturaleza y sus funciones. A medida que las examinamos, nuestro enfoque debería ser, por lo menos en parte, el del antropólogo cultural que fija su interés en una serie de patrones de conducta que le parecen raras y presumiblemente significativas, y que trata de entender las funciones para las que sirven y la manera en la que se relacionan con la cultura de la que son parte. Debemos hacer a un lado nuestro compromiso personal y mirar estos fenómenos con los fríos ojos del científico social, con el fin de tratar de identificar las fuerzas que moldean y sustentan las prácticas de calificación actuales y las presiones dentro de la cultura educativa que hace de estas prácticas algo tan irracional y resistente al cambio.*

*Thorndike R. Hagen E.<sup>30</sup>*

En general, la tarea de medir o evaluar los rendimientos escolares es una cuestión sumamente difícil y compleja, aun para el experto en mediciones. El educador se ve enfrentado a justipreciar conductas refractarias a la evaluación con instrumentos de medida poco confiables. Hace más de medio siglo Thorndike ya había señalado que la dificultad para medir la actuación de los alumnos se debe a

- Ausencia o imperfección de las unidades de medida
- Inconsistencia de los hechos a ser medidos
- Extrema complejidad de las mediciones

---

<sup>30</sup> Thorndike R. Hagen E. Ob. cit, p. 570





*"Roberto, en una exposición oral sobre un tema de historia fue calificado con un 10*

*¿qué significa esta nota?*

- *¿Que Roberto, según el criterio personal del profesor conoce extraordinariamente bien el punto expuesto?*
- *¿Que Roberto, que siempre se porta bien, nunca molesta a su profesor, es solícito con él y cumple el papel de "preferido" y recibe su premio? (aunque tanto el profesor como sus compañeros saben que no es tan excelente en historia)*
- *¿Que la exposición de Roberto, comparándola con la de sus compañeros, resultó la mejor? (aunque no se sepa qué calificación habría recibido si todos hubieran estudiado el punto y fueran tan capaces e interesados en historia como él)*
- *¿Que Roberto, comparando su actuación con normas obtenidas en grupos similares de estudiantes, realmente conoce el tema de modo muy satisfactorio?*
- *¿Que Roberto se ha esforzado denodadamente en aprender el tema expuesto, y aunque si bien la exposición no era de lo mejor, el profesor premia su esfuerzo con la nota más alta?*
- *¿Que Roberto es un brillante expositor y que, aunque sus informaciones no son tan exactas ni tan profundas, impresiona a su profesor con un lenguaje vivaz e ininterrumpido, siendo recompensado, la mayoría de las veces, con la nota más alta?*
- *¿Que Roberto, juntamente con el resto del curso, sin saber demasiado el tema en cuestión, recibe notas altas porque la modalidad de su profesor es, precisamente, utilizar casi indiscriminadamente las dos o tres categorías superiores de la escala?*

*Según se ve, en una única nota existen multiplicidad de significados a los que habría que agregar los derivados del propio rendimiento: ¿el 10 se refiere a la capacidad de recordar hechos específicos o a la habilidad para resolver con rapidez y precisión situaciones problemáticas?*

*Lafourcade P.<sup>31</sup>*

Múltiples factores intervienen en la adjudicación de las notas por parte de los docentes. Algunas investigaciones<sup>32</sup> han demostrado que:

- A igualdad de rendimiento y capacidad las niñas reciben mejor notas que los varones.
- Los alumnos más aceptados por sus maestros, a igualdad de rendimiento y capacidad, obtienen calificaciones más altas que los alumnos menos aceptados.

<sup>31</sup> Lafourcade P. Ob cit. p. 226

<sup>32</sup> Citadas en Lafourcade P. Evaluación de los aprendizajes. Kapelusz. Buenos Aires 1969 p. 227

- Las maestras tienden a asignar mejores notas a las niñas que a los varones y a ser menos generosas que los maestros. De modo que una nota alta o baja puede estar relacionada hasta con el sexo de quien califica.

Cuando los maestros y profesores adjudican notas (y aquí con independencia de los otros factores) el valor de las mismas depende, quiérase o no, de lo que ellos suponen debe ser aprendido en su curso.

Una de las críticas más serias al sistema de calificaciones en vigencia es que en un solo símbolo se incluye una gran variedad de objetivos cumplidos. Una única nota jamás podrá ser un buen indicador de logros tan diversos en calidad y jerarquía.

**Las informaciones deben ser forzosamente más analíticas y descriptivas.**

A pesar de todas estas limitaciones un sistema de calificaciones es necesario para comunicar los resultados. Cualquiera sea el sistema que se utilice (cualitativo, alfabético, decimal, etc.) las "notas" informan:

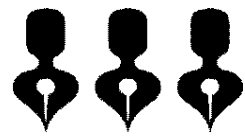
- al alumno sobre su propio rendimiento
- a los padres sobre el progreso de sus hijos
- a los maestros sobre la calidad de su enseñanza
- al futuro empleador sobre la capacidad de los postulantes

Uno de los problemas que enfrentan los docentes frecuentemente es la corrección de los exámenes escritos estructurados y la adjudicación de las notas de 0 a 10. En la práctica se utilizan distintos procedimientos con las más diversas fundamentaciones.



¿Cómo adjudica Ud. las notas de 0 a 10 en un examen final, escrito, estructurado, de 80 preguntas? En nuestro país está bastante difundido el criterio de exigir el 60% de respuestas correctas para aprobar el examen. ¿Al alumno que alcanza ese nivel de exigencia le corresponde la nota 4? ¿Cómo asigna las notas del 5 al 10?

*A continuación le proponemos un ejercicio sobre adjudicación de notas.*



## Ejercicio N° 3

*A. Busque el último examen escrito que ha tomado.*

1. ¿Qué tipo de preguntas ha utilizado?

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

2. Analice los puntajes que ha otorgado. Si no son muchos estudiantes puede confeccionar un listado ordenando a los estudiantes desde el de mayor puntaje hasta el de menor puntaje. ¿Cuál es la nota promedio?

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

3. ¿Cuántos alumnos obtuvieron nota 10?

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

4. ¿Cuántos alumnos aplazados hay? ¿Qué nota tienen: 0, 1, 2 o 3?

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

5. ¿Ud. cree que en todos los grupos de alumnos debe haber algunos aplazados?  
¿Por qué?

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



**B. Resuelva el siguiente problema**

En el cuadro siguiente figuran los resultados de un examen de 80 preguntas rendido por 32 alumnos. ¿Qué notas asignaría a cada uno de los alumnos? Escriba la nota correspondiente a cada uno de los alumnos utilizando una escala de 1 a 10.

Alumno	Puntaje Obtenido	Nota Adjudicada
AP	78	
BZ	77	
CV	74	
DF	72	
HJ	72	
PL	71	
BM	70	
VP	70	
FG	70	
TR	69	
LB	68	
EG	68	
PC	67	
RC	66	
TM	65	
PZ	64	
AV	64	
ST	64	
BU	63	
LC	63	
VM	62	
BK	61	
FC	59	
GH	57	
JR	55	
LG	53	
AN	52	
MB	50	
SS	48	
MA	48	
XZ	45	
WK	42	



*Compare sus notas con las que figuran en la CLAVE. En el anexo 5 se presenta una propuesta para la conversión de puntajes "crudos" a notas.*



## Comunicación de los resultados Brindar feedback al estudiante



*"Los estudiantes necesitan información acerca de ellos mismos, para guiarlos en su aprendizaje real y para sus planes futuros; necesitan retroalimentación inmediata que les diga lo que ya saben y dónde están sus errores y sus deficiencias".*

*Thorndike R. Hagen E.<sup>33</sup>*

Este momento tiene una importancia fundamental en las etapas de la evaluación que no siempre es reconocida por los docentes. Se ha expresado reiteradamente que la evaluación no es un mero trámite que el docente debe cumplir sino que tiene una especial significación para el aprendizaje de los alumnos.

La devolución de los resultados de la evaluación, especialmente si se trata de una evaluación formativa, es una oportunidad para brindar feedback.

La palabra "feedback", generalmente se traduce por "retroalimentación", "reprovisión", "observación interpretativa" o "mecanismo de revisión". Se usa principalmente en los sistemas electrónicos, y también en los viajes espaciales para controlar el mecanismo que regula las naves en el vuelo espacial. La estación terrestre observa y obtiene informaciones de la nave que, a su vez, envía informaciones a la tierra, para proseguir su viaje o para corregir desviaciones de la ruta marcada.

¿Puede un alumno tener una retroalimentación adecuada de su aprendizaje si busca su nombre en una lista y lee una calificación?. ¿Qué información le brinda al estudiante su nota? ¿Qué significa, además de saber que aprobó, un 6? ¿Es igual un 6 en un estudiante que en otro?


El momento de la devolución de los resultados es, por un lado, la instancia en la que el docente puede revisar y profundizar los temas incluidos en el examen. Por otro lado es la oportunidad de que cada estudiante sepa qué contenidos "domina" y cuáles son aquellos que desconoce o ha comprendido en una forma equivocada.

**El análisis de los resultados de un examen reorienta el aprendizaje.**

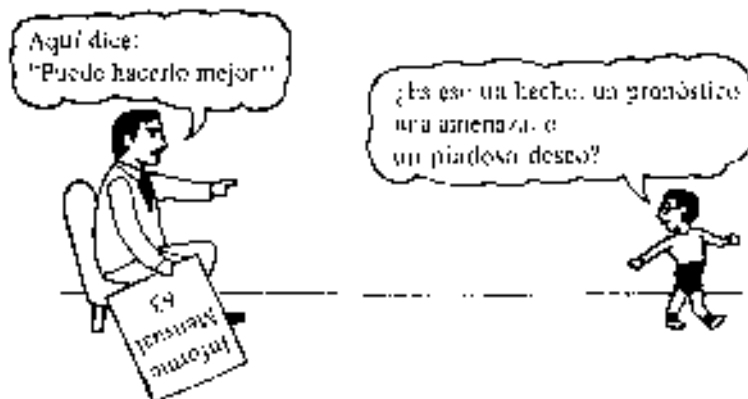
---

<sup>33</sup> Thorndike R. Hagen E. Ob cit.

## Algunas sugerencias para volver el feedback más útil y eficiente

 "El mensaje debe proveer información clara sobre los estándares de desempeño y el rendimiento del estudiante, para dilucidar cualquier diferencia"

Sachdeva AK<sup>34</sup>



**El feedback debe ser específico**

1. El feedback debe brindarse en un lenguaje más descriptivo que evaluativo. Evitando el lenguaje evaluador, se reduce la posibilidad de que la otra persona reaccione de una manera defensiva. No debe referirse a las características personales de los estudiantes. No es correcto decirle a un estudiante: "Sos un confuso"; "Sos un desprolijo", etc. Sí se puede señalar la confusión demostrada en la pregunta X entre tal y tal concepto. También será correcto señalar que ha sido difícil corregir y calificar la prueba debido a la mala caligrafía, la cantidad de palabras o frases tachadas, las llamadas fuera del renglón, la falta de respeto por las reglas de puntuación, etc. Será más productivo el feedback si estos comentarios se realizan con el examen a la vista del estudiante.
2. Debe ser específico y no general. En vez de decir "siempre te equivocás", sería mejor especificar cuál es el error que comete y en oportunidades ha demostrado la falla.
3. El *feedback* puede ser destructivo cuando sólo responde a las necesidades de aquel que lo ofrece, sin tomar en cuenta las de la persona que lo recibe. Cuando un docente brinda feedback a un estudiante el objetivo es la toma de conciencia por parte del estudiante de sus errores para poder corregirlo.
4. Se debe dirigir a los comportamientos que pueden ser modificados; caso contrario, aumentamos la frustración.

<sup>34</sup> Sachdeva A K (1996) Use of effective feedback to facilitate adult learning. J. Cancer Educ. Summer; 11 (2): 106-18

5. Es necesario ofrecerlo en el momento oportuno. Muchas veces, es más eficiente si es ofrecido en seguida de la ocurrencia de conducta, dependiendo, naturalmente, de la preparación de la persona o del grupo.



*El ambiente en el que se brinda el feedback debe ser de orientación y debe fomentar el establecimiento del diálogo entre el profesor y el estudiante."*<sup>35</sup>

6. Debe ser verificado para asegurarse que se ha establecido una buena comunicación. Se puede pedir al estudiante exprese con sus palabras cuáles fueron los principales problemas en su examen.

A continuación se proponen algunos ejemplos de cómo brindar feedback. Seguramente deberán adaptarse a las situaciones particulares de cada docente: cantidad de alumnos, propósito de la evaluación, tipo de examen implementado, etc.

### **Devolución grupal:**

Ya hemos dicho que cuando el docente corrige una prueba puede darse cuenta de los errores más frecuentes en ese grupo de estudiantes. Si se detecta algún error conceptual importante, compartido por muchos estudiantes, es necesario corregirlo y esta es la oportunidad que tiene el docente para hacerlo.

Se puede iniciar la clase de devolución de los resultados del examen con comentarios generales acerca de lo que la mayoría del grupo contestó bien y lo que han contestado mal. El docente puede realizar comentarios generales acerca de cómo ha funcionado la prueba, si ha resultados fácil o no, si las preguntas han sido buenas o alguna ha tenido dificultades, etc.

Luego de esto se puede entregar a cada estudiante su examen corregido para que lo revisen y luego de unos minutos comenzar, en grupo total, a analizar pregunta por pregunta y justificar entre todos la respuesta correcta.

Es conveniente que sea el docente el que coordine la discusión preguntando por ejemplo ¿Qué contestaron en la primera pregunta? ¿Por qué? ¿Alguien marcó la opción a)? ¿Por qué no era correcta la opción b)? etc. pero dejando que sean los estudiantes quienes respondan y discutan entre sí.

Permitir que se expresen defendiendo sus argumentos o justificando una elección de respuesta es un excelente ejercicio para ellos. En los casos en los que no haya acuerdo, el docente puede explicitar la clave de respuesta que se consideró correcta y fundamentar desde el punto de vista teórico.

---

<sup>35</sup> Ibidem

Este tipo de análisis ayuda a revisar nuevamente los contenidos, corregir errores, fijar el aprendizaje. El requisito indispensable es la planificación del tiempo para esta tarea. Deberá preverse en el cronograma una clase para la "devolución de los resultados de la evaluación"

Si el docente tiene un grupo numeroso de estudiantes puede realizar una clase en la que cada estudiante, con su examen en la mano, pueda observar las claves de respuesta correcta presentadas en una transparencia mientras se va explicando por qué esa y no otra era la respuesta que se esperaba. La participación en grupos muy numerosos, deberá ser menos espontánea y más pautada por el profesor. (levantar la mano si hay alguna duda, preguntar cuando se terminen de explicar las preguntas del subtema X, etc.)

Cuando se realiza un análisis grupal de los resultados de la evaluación, generalmente se aclaran la mayoría de las dudas individuales, pero siempre se debe dejar un tiempo para aquellos alumnos que necesiten alguna explicación adicional o deseen formular una observación a la corrección, a la puntuación, etc.

### **Devolución individual:**

Siempre es necesario para el estudiante tener una devolución individual de sus aciertos y errores. La mejor manera que tiene una persona de mejorar, modificando lo que está mal, es tomando conciencia del error.

Es importante, aún con un número de alumnos elevado, realizar una devolución de su rendimiento.

Esto es posible aún en Programas de Educación a Distancia con más de 5000 inscriptos. Se puede generar una hoja de devolución individual, que se manda por correo en la que se especifica qué puntaje obtuvo el participante, cuál era la opción correcta para cada pregunta (clave) y cuál fue la opción elegida por él. También se pueden explicitar los valores de las medidas de tendencia central del grupo total de individuos para que cada persona se ubique en relación al resto.

Si el docente trabaja con un pequeño grupo de estudiantes, es probable que pueda realizar una devolución personalizada de lo que ha resultado fácil o difícil para cada estudiante. Será un buen momento para ayudarlos a reflexionar sobre sus aprendizajes o sobre la metodología de estudio que tienen, esto es principalmente importante en los docentes de las material básicas, de los primeros años de la carrera. Muchos trabajos demuestran que los estudiantes universitarios ingresan a las carreras con escasas herramientas para



la lectura comprensiva y no dominan estrategias simples de estudio (identificación de ideas principales, resumen, etc.).

Se pueden explorar, en una entrevista o por escrito en una hoja fotocopiada, las siguientes cuestiones:

¿Suponías que este iba a ser el resultado/ la nota de tu examen?

¿Por qué crees que te fue así?

¿Por qué este tema te resultó tan difícil?

¿Cuál crees que es tu principal problema cuando rendís un examen?

¿Considerás que el tiempo que dedicaste a estudiar fue suficiente?

¿Cómo estudiaste, de a poco o dejaste todo para los últimos días?

¿Estudias solo o en grupo de compañeros?

¿Utilizaste las oportunidades de consulta con el profesor / tutor a lo largo del cuatrimestre?

¿Entendiste lo que se te estaba preguntando?

¿Leíste todo el examen antes de comenzar a responder?

¿Releíste o revisaste el examen antes de entregar?

¿Cuando entregaste el examen, estabas satisfecho con tu evaluación?

¿Qué crees que te faltó (leer, hacer, preguntar, discutir) para lograr el máximo posible?

¿Qué modificarías para la próxima evaluación?

Este tipo de preguntas, que están orientadas a la reflexión sobre el rendimiento del estudiante pueden facilitar la identificación de los problemas que muchos estudiantes tienen con las evaluaciones. Además son una buena oportunidad de mejorar la comunicación docente - alumno y generar un clima más propicio para el aprendizaje. Si los estudiantes perciben que el docente se interesa por lo que les pasa e intenta ayudarlos seguramente estarán más comprometidos con la tarea y aumentará la motivación.

Es importante que cuando el docente brinde feedback o explore las dificultades en los resultados de algunos estudiantes, mantenga una distancia afectiva óptima y no pierda de vista que el objetivo es ayudar a que el otro se re-orienta, retome el curso, mejore su rendimiento. De ninguna manera es una cuestión personal. Si algún docente se toma el trabajo de citar a un alumno a una entrevista personal para ayudarlo a reflexionar sobre su rendimiento debe estar dispuesto a que algún alumno le exprese que lo que pasó es que no estudió, o que la materia no le importa, o que no entiende para qué tiene que estudiar todo esto, etc. En este caso será responsabilidad del docente demostrar la utilidad de esta materia en el Plan de la Carrera, e intentar convencer al estudiante del valor de los objetivos propuestos.

## Toma de decisiones

Las decisiones más claramente relacionadas con la evaluación son aquellas en las los profesores deciden, con base en la mejor evidencia disponible, si un estudiante esta aprobado o no. La correcta confección de un examen, la administración en condiciones adecuadas y la posterior corrección y análisis de resultados favorecen las decisiones racionales acerca de la promoción del estudiante.

Por otra parte una correcta comunicación de resultados orienta al estudiante en el esfuerzo que será necesario para alcanzar un desempeño óptimo. Se puede afirmar que el propósito más importante de la evaluación es reorientar el aprendizaje del estudiante.

Sin embargo, la evaluación, también brinda muchas evidencias acerca de la calidad de la enseñanza ofrecida. Los docentes deberían revisar esta información para tomar decisiones que modifiquen u optimicen la enseñanza. Los resultados de la evaluación de los estudiantes ponen al docente en el foco de análisis. ¿Qué capacidad han tenido como planificadores de la asignatura? ¿Qué tipo de experiencias y materiales de aprendizaje han ofrecido al estudiante? ¿Qué habilidades han demostrado para la confección de exámenes? ¿Qué calidad han tenido los ítem o preguntas que han hecho? ¿Cómo han calificado a los estudiantes? ¿Han logrado aumentar la validez y la objetividad de los instrumentos?

Se podría reiterar que la evaluación es útil para la sociedad, ya que permite asegurar la calidad de los graduados que se insertan en ella, para el estudiante, ya que propone un acompañamiento en el esfuerzo de aprender y finalmente para el docente, ya que le proporciona evidencias para tomar decisiones racionales, asumir su responsabilidad social con fundamentos y mejorar permanentemente la calidad de la enseñanza que propone.



*Para finalizar este  
módulo le proponemos  
un ejercicio de  
Autoevaluación.*

¿En qué medida Ud. siente que...	En gran medida	Modera- damente	Muy poco	Nada
¿sus expectativas en relación a este tema se han visto satisfechas?				
¿es capaz de construir exámenes escritos cuyos resultados sean objetivos y válidos?				
¿va a mejorar los exámenes orales?				
¿es capaz de realizar un análisis de la calidad de las preguntas de un examen estructurado?				
¿podrá mejorar su práctica docente?				
¿será capaz de influir en sus colegas para mejorar el sistema de evaluación?				

## Recuerde

Ud. ha comenzado a evaluar el desempeño de sus alumnos. Ya ha indagado el área cognoscitiva.

Sabe qué y cuánto saben sus alumnos. Como ha hecho la devolución de los resultados, ellos también lo saben.

- ▶ Para evaluar el desempeño es necesario, además de indagar el área cognoscitiva, explorar los logros en el área psicomotora y afectiva (destrezas y actitudes).

*En el módulo 3 Ud. tendrá la información necesaria para realizar esta exploración*

*Si desea profundizar sus conocimientos, le sugerimos la lectura de la **bibliografía** citada a pie de página.*



# Clave de Respuestas

*¿Un "4"? ¿Me saqué un "4" por mi escultura de alambre?*



1

*¿Puedo preguntar algo?*



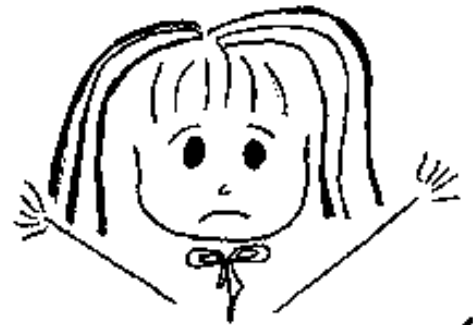
2

*¿Me calificaron por mi talento? En tal caso, es verdad que se nos juzga por una parte de nuestra vida que no podemos controlar.*



3

*Si me calificaron por mi esfuerzo, ¡esto es injusto!, porque lo hice lo mejor que pude*



4

*Quizás tuvieron en cuenta lo que aprendí en esta unidad de trabajo. Pero en ese caso la maestra debió juzgarse a sí misma por su capacidad para enseñarme ...*



5

*¿No querrá compartir mi "4"?*



6



## EJERCICIO N° 1

### A.

- 1.- Comprensión.
- 2.- Memoria. Recuerdo de información.
- 3.- a), b) y c) Comprensión de la información. D) aplicación de los conocimientos.
- 4.- a) Recuerdo de información y b) Aplicación de conocimientos.

### B.

- 1.- El error de este ítem es la doble negación.  
Ítem mejorado: "La TBC es una enfermedad contagiosa" V ( X ) F ( )
- 2.- El error de este ítem es la doble negación.  
Ítem mejorado: "La resistencia al sarampión obtenida mediante el uso de la vacuna contra el sarampión se llama inmunidad pasiva" V ( ) F ( X )
- 3.- Error no está especificado qué cantidad de síntomas debe colocar el estudiante para que su respuesta sea considerada correcta. El término "comunes" como adjetivo de síntomas es inespecífico. Se puede referir a frecuencia, a que son compartidos por otra enfermedad, etc.  
Ítem mejorado: "Mencione los 5 (cinco) síntomas más frecuentes Fiebre Hemorrágica Argentina."
- 4.- Este ejemplo tiene varios errores: Es una correspondencia 1 a 1; Kuwait - Al Kuwait es obvio. Si es necesario preguntar este contenido se debe indagar, por ejemplo, con un ítem de completamiento: "La capital de Kuwait se llama .....".
- 5.- El error de este ejemplo es que las categorías presentadas en las opciones de respuesta no son mutuamente excluyentes. ¿Dónde debería marcar un estudiante que sepa o crea que la respuesta es el 5%? En la opción a) o en la b)?
- 6.- Este ítem puede ser contestado de muchas maneras, ya que el alumno puede creer cualquier cosa y se debe respetar su creencia.

Ítem mejorado: "¿Cómo se llamó el hombre que descubrió la penicilina?"

- 7.- En este ítem, las respuestas posibles y correctas son múltiples: un científico, un hombre, la casualidad, un aventurero e incluso Fleming.  
Ítem mejorado: "La penicilina fue descubierta por un hombre llamado ....."
8. 9 y 10.- En estas afirmaciones se puede contestar casi cualquier contenido ya que no queda claro qué es lo que se está preguntando. Es casi imposible que los estudiante "adivinen" las siguientes respuestas:  
"La penicilina fue descubierta por Fleming en 1929"  
"La presencia de hipertrofia ventricular izquierda en un hipertenso es un factor que aumenta la mortalidad cardiovascular".  
"En pacientes operados de Tetralogía de Fallot la secuela electrocardiográfica más frecuente es el bloqueo de rama derecha".
- 11.- Falta el verbo. Las relaciones que unen estos dos términos son extremadamente variables y múltiples para alumnos con imaginación.

Seguramente en estos últimos 4 ítem no habrá logrado reformularlos o mejorarlos ya que no quedaba claro qué era lo que se intentaba explorar.

### C.

- 1.- Alternativas constantes:
  - 1.1.- La secreción salival es responsable del aumento de la amilasaemia. V  F
  - 1.2.- La pérdida de potasio altera la función celular normal y repercute especialmente en la función neuromuscular. V  F

Clave de respuestas 1.1. verdadero - 1.2 verdadero

- 2.- Apareamiento:  
Establezca la correspondencia entre las patologías de la columna de la izquierda y los enunciados que figuran en la columna de la derecha. Escriba la letra que corresponda a continuación de cada patología. Cada letra puede ser usada una, varias o ninguna vez.

1. bulimia nerviosa ..a.c...f.i..	a) uso de laxantes y/o diuréticos b) sedentarismo c) actividad física extenuante d) está demostrado un fuerte componente genético e) sangrado intermenstrual f) ciertas pautas familiares favorecen el desarrollo de la enfermedad
2. obesidad ..b.d..f.g..	g) se observa una mayor tendencia entre hispanos y mediterráneos h) obstrucción intestinal i) vómitos autoinducidos j) infantilismo sexual

Nerviosa. Marque la EXCEPCIÓN.

- a) Es mucho más frecuente en las mujeres
- b) Tiene mayor incidencia en las clases medias y altas
- c) Hay antecedentes de depresión y/o adicciones en el grupo familiar
- d) Existe mayor incidencia en los países orientales. (Clave)



Recuerde que las preguntas siempre pueden ser mejoradas si son analizadas entre varios docentes!!!

### 3.- Selección múltiple

- 3.1.- La bulimia típica generalmente comienza:
- a) antes de los 10 años
  - b) entre los 11 y 15 años
  - c) entre los 17 y los 20 años (clave)
  - d) después de los 21 años
- 3.2.- ¿Cuál de las siguientes afirmaciones en relación a los Trastornos Alimentarios es CORRECTA?
- a) Las situaciones fundantes para la aparición de los trastornos de la alimentación se instalan en la vida intrauterina
  - b) El aumento de la secreción de la fracción 8-S de la colecistoquinina se produce en respuesta a las comidas
  - c) Existe generalmente en este tipo de pacientes, un alto grado de tolerancia a la frustración
  - d) El abuso de drogas, de alcohol y/o cleptomanía suelen estar asociados a estos cuadros (clave)
- 3.3.- Las siguientes, excepto una, son manifestaciones clínicas de la Bulimia Nerviosa. Marque la EXCEPCIÓN.
- a) Significativas alteraciones en el EEG- insomnio (Clave)
  - b) Constipación - diarrea
  - c) Disminución de la fuerza y resistencia muscular
  - d) Signo de Russel
- 3.4.- ¿Cuál de las siguientes afirmaciones en relación a la etiopatogenia de la Bulimia Nerviosa es INCORRECTA?
- a) Los niveles de serotonina están disminuidos
  - b) Los valores de prolactina están aumentados (Clave)
  - c) No se registra aumento del cortisol
  - d) Las hormonas T3- T4 y T3 reversa son normales
- 3.5.- Las siguientes afirmaciones, excepto una, son ciertas en relación a la epidemiología de la Bulimia

## EJERCICIO N° 2

### A.

1. Respuesta correcta e) Los ítem han resultado muy difíciles. Lo primero que debe hacerse es revisar la calidad de los mismos. ¿Había una sola respuesta correcta? ¿Cómo se han distribuido las respuestas de los estudiantes en el resto de las opciones? ¿Hubo alguna opción que resultó atractiva para la mayoría? ¿Podría ser también correcta? ¿Qué índice de discriminación presentaron estas preguntas? Seguramente estos ítem deban revisarse muy cuidadosamente. Si la formulación de los mismos es inobjetable, se debe reflexionar sobre la enseñanza de los temas explorados y modificar aspectos relacionados con la planificación de las estrategias de aprendizaje propuestas.
2. Respuesta correcta a) El ítem discrimina muy bien entre los estudiantes que saben y los que no. El ítem se debe mantener sin modificaciones.

### B.

- 1.
- |             |             |              |                              |
|-------------|-------------|--------------|------------------------------|
| a)          | 876         | 15,40        | a) Muy fácil                 |
| <b>b)</b>   | <b>2969</b> | <b>52,18</b> | b) Relativamente fácil       |
| c)          | 740         | 13,01        | <b>c) Dificultad media X</b> |
| d)          | 1076        | 18,91        | d) Relativamente difícil     |
| No contesta | 29          | 0,5          | e) Muy difícil               |



2.				
a)	8	0,14	a) Muy fácil	X
b)	27	0,47	b) Relativamente fácil	
c)	228	4,01	c) Dificultad media	
d)	<b>5423</b>	<b>95,31</b>	d) Relativamente difícil	
No contesta	4	0,07	e) Muy difícil	

3.				
a)	155	2,72	a) Muy fácil	X
b)	40	0,70	b) Relativamente fácil	
c)	<b>5446</b>	<b>95,71</b>	c) Dificultad media	
d)	46	0,81	d) Relativamente difícil	
No contesta	3	0,05	e) Muy difícil	

Se debe analizar la distribución de las respuestas de los alumnos, especialmente revisar aquellas que no han funcionado como plausibles para casi nadie. Probablemente habría que reemplazar estas opciones por otras más pertinentes.

### C.

ID: 28%. Es un ítem regular, debe revisarse.

## EJERCICIO N° 3

Alumno	Puntaje obtenido	Nota
AP	78	10
BZ	77	
CV	74	
DF	72	9
HJ	72	
PL	71	
BM	70	
VP	70	8
FG	70	
TR	69	
LB	68	
EG	68	
PC	67	
RC	66	
TM	65	
PZ	<b>64</b>	7
AV	<b>64</b>	
ST	<b>64</b>	
BU	63	
LC	63	6
VM	62	
BK	61	
FC	59	5
GH	57	
JR	55	4
LG	53	
AN	52	
MB	50	
SS	48	
MA	48	2
XZ	45	
WK	42	

¿Le ha quedado igual o diferente?



Consulte el Anexo 5



# ANEXOS

## **Anexo 1**

### **Las apreciaciones personales del profesor**

Alicia Camilloni. Cefyl-UBA 1989

## **Anexo 2**

### **Puntuación corregida. Calificación de la prueba objetiva**

Thorndike R, Hagen E. *Medición y evaluación en Psicología y Educación*. Trillas. México. 1996

## **Anexo 3**

### **Ejemplos de hojas de respuesta**

## **Anexo 4**

### **Ejemplo de instrucciones generales de la prueba**

## **Anexo 5**

### **Propuesta para la adjudicación de notas**

Lic Amanda Galli



---

# Anexo 1

## Las apreciaciones personales del profesor

Alicia Camilloni

---

Cuando el profesor debe calificar a un alumno, sea en la calificación de un examen oral o de un examen escrito no objetivo, sea en la elaboración del llamado "concepto" del escolar, es frecuente que deba recurrir a apreciaciones personales. Aunque muchas veces estos juicios estimativos presenten mayor valor predictivo que las calificaciones obtenidas por medio de instrumentos objetivos, también pueden encontrarse viciados por un cúmulo de factores psicológicos que actúan como causas de error.

Según McClelland (1), las estimaciones de los profesores tienen las siguientes características:

- a) Son generalmente demasiado altas.
- b) Tienden a sobreestimar los logros de los alumnos débiles y a subestimar los de los más capaces.
- c) Los patrones de medida utilizados varían considerablemente de escuela en escuela.

Estas características restan exactitud a las calificaciones personales. Pueden, sin embargo, ser controladas, si el profesor procura depurar los juicios valorativos, de los elementos que deben permanecer ajenos a su elaboración.

La arbitrariedad de las estimaciones ha sido observada desde hace mucho tiempo por las autoridades y por los alumnos. Los Estatutos de las Universidades de Bolonia y de Padua, que datan del siglo XV prescriben, para evitar la odiosa parcialidad que se manifestaba anteriormente en los exámenes, que los examinados sean tratados con afecto paternal, y que cualquier acto de hostilidad del profesor sea penado hasta con un año de suspensión.(2)

Todavía hoy los estudiantes reaccionan contra la misma parcialidad, contra el "humor y juicio imprevisible que son elementos indispensables para la buena marcha del examen oral" (3), o que al menos, parecen indispensables por la frecuencia con que lo acompañan.

Para juzgar con acierto se requiere poseer, primero, la capacidad de percibir adecuadamente a la otra

persona y tener, además, la capacidad de tomar decisiones acordes con lo percibido. Es una ardua tarea porque el profesor es a la vez juez y parte en ella y porque debe realizarla ejerciendo un poder que el alumno mismo no le ha conferido.

Esa facultad de juzgar es el símbolo de la autoridad encarnada por el profesor e implica tantos riesgos que conviene analizar las posibles causas de error operantes en la elaboración de juicios estimativos sobre la actuación de un alumno. "El arte de examinar -dice Piobetta- importa cualidades muy diferentes de aquéllas que requiere el arte de enseñar. Se puede ser un excelente profesor sin ser por eso un buen examinador."(4) Veamos, pues, algunas de las causas corrientes de error:

**Información insuficiente:** el profesor no cuenta a menudo con suficientes datos acerca del trabajo del estudiante como para poder calificarlo con justeza. En la marcha cotidiana de la enseñanza el profesor se ve obligado a evaluar el rendimiento de gran número de alumnos en períodos muy cortos, por lo que sus juicios concluyen siendo frecuentemente improvisados. Algunos reglamentos escolares y disposiciones ministeriales reducen también la duración de los exámenes orales finales a períodos tan breves (15 minutos, por ejemplo) que la información recabada acerca del rendimiento del alumno resulta a todas luces incompleta. Los juicios estimativos emitidos sobre bases inseguras se ven, así, influidos más por circunstancias azarosas que por una cuidadosa ponderación de los logros y yerros de los alumnos.

**Efecto de halo:** "Es la tendencia a formar una impresión o juicio sobre una característica particular de un individuo basándose sobre la impresión general que se tiene de él. Si la impresión general es favorable será hipercalificado en los rasgos positivos y viceversa." (5) El efecto de halo es una manifestación

de la tendencia a evaluar una parte en función del todo. Aparece tanto en la predisposición a calificar positivamente las nuevas realizaciones de los buenos alumnos, o a la inversa con los malos, como en la tendencia a calificar positivamente a un alumno cuya presencia física y modales son agradables, o lo que es más absurdo, porque sus hermanos han sido buenos alumnos. Actúa, pues, como un prejuicio que impide la consideración objetiva del rendimiento. Se origina en la necesidad de lograr una imagen homogénea de cada uno de los escolares, reduciendo el margen de error derivado de las incoherencias que corrientemente se presentan entre las distintas observaciones. Su efecto, sin embargo, es el opuesto, pues sólo consigue distorsionar las imágenes.

**Hipótesis de la personalidad implícita:** la forma en que un profesor percibe a un alumno se halla influida por sus creencias acerca de cómo se organiza la personalidad, y particularmente por la forma en que considera que los distintos rasgos de la personalidad se asocian.<sup>(6)</sup> Si piensa, por ejemplo, que el ser indisciplinado implica el ser haragán, al observar en un alumno un comportamiento carente de disciplina tenderá a pensar que ese alumno es haragán, y su actitud hacia él le hará producir calificaciones desfavorables. Las variedades en las hipótesis acerca del apareamiento de los rasgos son muchas. Cambian según los profesores, que podrán unir rasgos como linda e inteligente, o separarlos.

La hipótesis de la personalidad implícita actúa también como un prejuicio, pero a la inversa del efecto de halo que consistía en derivar la cualidad de una parte de lo que se sabía o barruntaba del todo, esta hipótesis opera derivando la cualidad del todo, de lo conocido acerca de una parte.

**Tendencia a la categorización:** la percepción de un objeto entraña su colocación en una categoría de objetos que es ya conocida por el sujeto. No se percibe simplemente un ruido sino el tañir de una campana o un cañonazo. No se percibe una forma amarilla, sino una flor o un libro. Lo mismo ocurre con las personas. Se las percibe como pertenecientes a un tipo determinado y se les atribuyen rasgos supuestamente característicos del grupo al que pertenecen.

Las categorizaciones ayudan al profesor que tiene que calificar a doscientos, trescientos o más alumnos cada año. No configuran necesariamente causas de error, pero sí tienden a serlo, y con mucha intensidad cuando no se manejan de manera flexible.

La base de las categorizaciones está constituida por estereotipos: "todos los hijos únicos son malcriados", "todos los japoneses son inteligentes y aplicados", "todos los hijos de médicos son buenos alumnos", etc. Son, como las llama Hoftätter (7), "soluciones patentadas" que liberan de la penosísima inseguridad que acomete al profesor que debe calificar a centenares de alumnos en un tiempo limitado. Abrevian, aparentemente, la tarea, pero la exactitud y sensibilidad de la evaluación se pierden al perderse la riqueza que se seguiría de una captación original de cada una de las conductas individuales.

Los efectos distorsionadores de la categorización se manifiestan, en forma particular, cuando las categorías empleadas son pocas, esto es, cuando la categorización va acompañada de una tendencia a la sobresimplificación.

Cuanto mayor es la distancia geográfica, histórica o social entre los que juzgan y los que son juzgados, mayor es la disposición a trabajar con pocas categorías, y a definir pobremente cada una de ellas. Así como a los bantúes pueden parecerles iguales todos los esquimales, un niño puede encontrar parecidos a todos los ancianos. "Es muy posible que un estudiante universitario sepa apreciar distinciones individuales en sus camaradas y los instructores más jóvenes, pero mira a todos los profesores de más edad con el mismo rasero, viéndolos como "vejstorios" sin individualidad. Recíprocamente, hay profesores que ven a todos los estudiantes como si fueran iguales y como tales los tratan." (8)

La tendencia a la categorización puede ser, en consecuencia, la fuente de tres tipos de error: ubicación equivocada del individuo en una categoría que no le corresponde por causa de una defectuosa interpretación de los datos; deformación de las percepciones y los juicios por el manejo de una clasificación organizada sobre un número insuficiente de categorías y, por último, empobrecimiento de la observación por el empleo de categorías definidas sobre la base de muy pocos rasgos. Pensemos en "el

abanderado de la escuela", en "el alumno aplazado", en "el alumno de siete", en "el faltador", etc.

**Primacia de la primera impresión:** las primeras impresiones parecen tener una influencia fundamental en la forma en que una persona percibe y juzga a otra. Sus juicios están teñidos por ese efecto inicial, y las impresiones posteriores se distorsionan buscando ajustarse a él.

De manera semejante al efecto de halo o a la hipótesis de la personalidad implícita, una impresión, en este caso la primera, irradia sobre las que le siguen. Por esta razón encontramos con frecuencia entre los estudiantes la afirmación de que es difícil levantar una clasificación inicial baja, o que el alumno que ha comenzado bien su curso "puede dormirse sobre sus laureles" sin grave riesgo para sus posibilidades de promoción.

La intensidad del efecto de primacia de las primeras impresiones disminuye grandemente si advertimos a los profesores sobre el carácter falaz que presentan.

"El orden que observamos ahora en las acciones de una persona no sigue un plan establecido. Si de todas maneras nos ingeniamos para llegar a una opinión razonable, es porque pudimos sobreponernos al orden accidental en que se nos aparecen las diversas características de la persona, las cuales pudieron relacionarse con las experiencias anteriores y posteriores".(9) Una actitud abierta, que permita desde un comienzo el empleo de criterios amplios y flexibles, hará que el profesor se mantenga alerta contra la producción del efecto de primacia de la primera impresión, origen muy común de errores. "Al parecer, los jueces cuyas primeras impresiones son numerosas, ricas y variadas, se muestran luego flexibles al recibir información adicional, modificando y ampliando fácilmente el juicio original. En cambio los que ven poco de la personalidad al principio, es probable que conserven su reducida y estereotipada impresión inicial a pesar de obtener una información mayor." (10)

**Primacia de la última impresión:** así como en una serie que se ha memorizado, los elementos que mejor se recuerdan son los primeros y los últimos, también en una serie de impresiones, la impresión de data más

reciente puede conducir a una categorización de todo el conjunto.

Piénsese, por ejemplo, en la importancia que adquiere la última pregunta en un examen oral, que, bien contestada, hace olvidar al profesor muchas de las respuestas erróneas anteriores correspondientes, sin embargo, a temas de igual valor que el tratado al final del interrogatorio.

Este fenómeno, al igual que el anterior, no origina exclusivamente errores. Puede darse el caso de que la primera o última impresión concuerden no sólo con la primera o la última respuesta real del alumno sino también con las respuestas intermedias. Si fuera así, la resonancia de una de esas impresiones permitiría una inferencia acertada acerca de las cualidades personales o del trabajo futuro del alumno. Pero debe acordarse que la exactitud de la inferencia no depende del simple peso afectivo ejercido casi siempre por la primera impresión, y a veces por la última, sino de las circunstancias que hemos descrito y a las que la realidad no se ajusta sino excepcionalmente.

La aceptación de la resonancia afectiva de una impresión constituye una vía ilegítima para la elaboración de juicios estimativos, y como tal debe ser desechada. Resultará sencillo contraponerle un análisis cuidadoso de todas las respuestas de los alumnos y de todas las impresiones que ellas han producido.

#### **Influencia del aspecto físico y en especial de los rasgos faciales:**

cuando no se cuenta con informaciones suficientes acerca del rendimiento de un individuo suele buscarse apoyo en la observación de su aspecto físico. Se denomina fisiognómica al arte de descubrir las características de la personalidad sobre la base del estudio del cuerpo, y sobre todo, de la conformación y expresión del rostro. Desde el primer tratado que se conoce sobre el tema, Physiognomónica, atribuido a Aristóteles, muchos ensayos se han realizado para convertir a la fisiognómica en una ciencia. Estos intentos han sido hasta ahora infructuosos, por lo que no se puede conceder gran importancia a las evaluaciones fundadas sobre una fisiognómica improvisada. En este campo todas las reglas presentan gran número

de excepciones. Rostros "muy inteligentes" pueden esconder grandes incapacidades, una mímica facial que aparentemente denota indiferencia corresponde, en ocasiones, a una realidad personal de muy otra naturaleza.

Conviene, pues, tratar de marginar las consideraciones referentes a la configuración física, los rasgos faciales y los gestos de los estudiantes, cuando se elabora una calificación referida a su trabajo e incluso a sus características personales.

**Proyección:** la proyección es un mecanismo de defensa al que algunas veces recurre el profesor. Engañándose a sí mismo, atribuye faltas a los alumnos para no ver sus propios defectos como enseñante. Intenta evadirse de esta manera el sentimiento de culpa que lo aqueja.

"Los alumnos de este curso -dice- muestran una gran pedertería; no aceptan nada de lo que se les enseña sin discutirlo previamente". Podríamos preguntarnos ¿quién es el pedante? También oímos decir a veces: "Este curso es indolente, no tiene ningún entusiasmo por aprender lo que les enseñó". ¿De quién es la responsabilidad? ¿De quién es la indolencia? "Este alumno hace un gran esfuerzo. Sé lo que significa tratar de estudiar cuando hay enfermos en la familia". ¿Mensuramos su rendimiento o la intensidad de los problemas personales del estudiante? ¿A quién tratamos de disculpar, al alumno que está frente a nosotros o al estudiante que fuimos?

Los ejemplos se multiplican en nuestra memoria. Proyectamos sobre los alumnos, sobre los colegas, sobre las autoridades de la escuela, sobre el clima. Las calificaciones resultantes pierden objetividad porque están preñadas de nuestros propios conflictos.

**Marco de referencia estrecho:** el marco de referencia es "el marco dentro del cual juzgamos las propiedades de un determinado objeto".(11) La variabilidad presentada por las calificaciones de un profesor puede atribuirse frecuentemente a la variabilidad de los marcos de referencia empleados en su elaboración.

Para cada ocasión se adopta, como término de comparación, una tanda diferente de alumnos, de exámenes o de trabajos. Sabiendo esto, los alumnos

de rendimiento mediano evitan rendir examen a continuación de un alumno excelente. Los profesores repiten: "X merecía 9, pero si lo comparamos con Y, que ha estado tan bien, debemos calificarlo con 8". O a la inversa: "Z merecía 8, pero comparando su examen con los demás, que han sido tan flojos, convendrá calificarlo con 9 ó 10".

Estas consideraciones son ilegítimas, ya que hacen depender la calificación de cada alumno, de circunstancias azarosas que nada tienen que ver con su rendimiento real. La mejor forma de evitar la variabilidad consiste en fijar previamente y de manera concreta, los estándares de rendimiento deseados, debiéndose usar sólo estos estándares como marco de referencia.

**Error por generosidad:** es común que el profesor se sienta inseguro al fijar una calificación para el alumno. Esta inseguridad se origina en la indeterminación del significado atribuido a cada uno de los grados de la escala que utiliza, en la indefinición de los objetivos cuyo logro debe evaluar, o en la falta de exactitud de las medidas obtenidas por medio de técnicas de evaluación en las que no confía totalmente. Con todas estas notas in mente, y sabiéndose juez y parte al mismo tiempo, no es raro que tienda a calificar generosamente a los estudiantes.

Es el profesor quien concede significado a cada grado de la escala, quien ha definido los objetivos y quien ha construido las pruebas de evaluación. Lo ha hecho sobre la base de un cierto número de opciones, algunas de las cuales reconoce defendibles, aunque no las haya elegido.

Cuando una calificación desfavorable puede afectar seriamente al alumno, el profesor no puede evitar el interpretarla, al modo en que lo hace el estudiante a menudo, como un ataque personal a otro individuo. "El resultado es claro, las calificaciones tienden a acumularse en el extremo superior de cualquier escala". "Se descubre muy frecuentemente la paradoja de que una gran mayoría del grupo está clasificado por encima del término medio".(12)

Si todos los profesores evaluaran con la misma generosidad a sus alumnos, la rectificación sería sencilla, pero esto no ocurre, pues hay diferencias, no



sólo entre distintos profesores, sino para el mismo profesor en diferentes días.

Aunque el principio que rige para los fallos judiciales, *in dubio pro reo*, debe aceptarse también en la evaluación escolar, es menester precaverse contra el error por generosidad. La única forma de hacerlo será limitar el sentimiento de inseguridad que aqueja al profesor, definiendo con precisión cada uno de los grados de la escala y los objetivos cuyo logro ha de evaluar, y sometiendo a control constante las técnicas de evaluación empleadas con el objeto de darles la mayor objetividad y exactitud posibles. Debemos reconocer, sin embargo, que en algunos casos, a este error por generosidad se contrapone un error por mezquindad el que tiende a disminuir, sin razón aparente, las calificaciones.

Hemos mencionado hasta ahora las siguientes causas posibles de error en las apreciaciones personales del profesor:

- 1) información insuficiente.
- 2) efecto de halo.
- 3) hipótesis de la personalidad implícita.
- 4) tendencia a la categorización y la sobresimplificación.
- 5) primacía de la primera impresión.
- 6) primacía de la última impresión.
- 7) influencia del aspecto físico y de los rasgos faciales.
- 8) proyección.
- 9) marco de referencia estrecho.
- 10) error por generosidad (o por mezquindad).

Entendemos que todos estos errores pueden ser eludidos mediante una cuidadosa crítica de los factores que intervienen en el proceso de elaboración de las calificaciones.

Existe, sin embargo, otra causa de error, cuya influencia es más difícil de evitar. Está dada por la falta de capacidad para juzgar a otros que presentan ciertos individuos. "Hay personas que son mejores jueces que otras".(13)

Al examinar las calificaciones que diferentes profesores atribuyen a sus alumnos, se observa que existe en algunos una disposición a ser severos y, en otros, a ser indulgentes.

Es lo que suele denominarse ecuación personal de los

maestros: se manifiesta en cada uno, una actitud valorativa general que lo impulsa a ser siempre riguroso, siempre benévolo, o inestable, en sus juicios acerca de los alumnos o del trabajo que ellos realizan. Las calificaciones de un profesor tienden, de esta manera, a agruparse en un cierto sector de la escala. Por ejemplo: 5-10; 4-9; 2-8, etc. Para poder interpretar entonces cada calificación, es menester someterla a un proceso de corrección ajustado a las características de la ecuación personal del profesor que la ha elaborado.

Si analizamos ahora los rasgos peculiares de las estimaciones de los profesores, tales como los había descrito McClelland, comprobamos que ellos son el resultado de algunas de las causas de error que hemos indicado.

La primera característica, el que las calificaciones sean generalmente demasiado altas, puede atribuirse a un efecto del error por generosidad.

La segunda, sobreestimación de los logros de los alumnos débiles y subestimación de los éxitos de los alumnos buenos, tiene como consecuencia la disminución de la amplitud de variación de las calificaciones y es un efecto de la ecuación personal de cada profesor.

La tercera característica, que consiste en la variabilidad de los patrones de medida empleados, tiene varias causas, por ejemplo la limitación de los marcos de referencia.

Luego de examinar las posibles causas de error intervinientes en la elaboración de juicios estimativos, veamos cuáles son las cualidades que debe tener un buen evaluador escolar.

Debe reunir, en primer lugar, experiencia docente y conocimientos. Estos últimos, correspondientes a dos áreas. Por un lado, es necesario que conozca a fondo el campo de instrucción en el cual se ha de evaluar el rendimiento de los estudiantes. Es probable que se requiera poseer más conocimientos para evaluar con eficacia que para enseñar con eficacia. Pero además de conocer su materia, es indispensable que el evaluador conozca perfectamente las técnicas de evaluación que ha de emplear, las ventajas y desventajas que tiene cada una de ellas, sus condiciones de practicidad y sus alcances en relación

con la apreciación del logro de los objetivos o de los propósitos de la educación.

Un buen evaluador no se improvisa. La experiencia docente le ayuda a afinar sus juicios y a distinguir lo que hay de original en las expresiones de cada estudiante.

En cuanto a las cualidades personales, las exigencias son muchas. En el aspecto moral, el evaluador debe reunir honestidad, equidad, independencia de criterio, sentido de la responsabilidad y altruismo. En el aspecto afectivo es fundamental que tenga estabilidad emocional. Esta se manifestará en su igualdad de humor y paciencia, sin las cuales, el estudiante estaría sometido a la arbitrariedad y el capricho del momento. En el aspecto funcional, deberá tener la posibilidad de mantener la atención durante largo tiempo, poseer poder imaginativo y capacidad mnemónica, agudeza perceptiva, capacidad para tomar decisiones y, fundamentalmente, gran capacidad intelectual, representada por flexibilidad de espíritu, poder de captar los matices y rapidez mental, esta última indispensable en los exámenes orales.

Agreguemos finalmente otro rasgo indispensable en un evaluador educacional: la intracepción. Esta consiste en la capacidad para comprender intuitivamente los sentimientos, los deseos, y las intenciones de otra persona. Ella sensibiliza al profesor con respecto al estado de ánimo del estudiante y le permite, en consecuencia, mejorar las condiciones de la situación en que se evalúa su rendimiento. Debe tenerse presente, sin embargo, que la intracepción no debe conducir nunca a la identificación con el estudiante, pues para juzgar con equidad hay que mantener cierto desapego ante la persona que es juzgada. La excesiva cercanía puede distorsionar también los juicios. (14)

Como se ve, las exigencias son muchas. Las posibilidades de error también. Como dice Allport: "Aún cuando el proceso de conocimiento del otro es muy complejo, aun cuando hay muchas fuentes de error que alteran el juicio, una de las tareas que más esforzadamente realiza el hombre en la vida es la de

salvar mediante la comprensión el abismo que separa el tú del yo". (15) Esta tarea no es opcional para el profesor. Forma parte de su labor cotidiana y es uno de los pilares de la acción educadora. En ella se pone en juego su ética profesional y se manifiestan todas sus cualidades personales.

- 1) William McClelland, *Selection for Secondary Education*, Univ. of London Press, Londres, 1942, pág. 34.
- 2) José Antonio Encinas, *Historia de las Universidades de Bolonia y de Padua*, Ed. Ercilla, Santiago de Chile, 1935, pág. 85.
- 3) "Los estudiantes de Ca'Foscari, Venecia", en *Las luchas estudiantiles en el mundo*, Ed. Galerna, Buenos Aires, Argentina, 1969, pág. 166.
- 4) *Exámenes y concursos*, Kapelusz, Buenos Aires, Argentina, 1952, pág. 77.
- 5) D. Krech, R. S. Crutchfield y E. L. Ballachey, *Psicología social*, Biblioteca Nueva, Madrid, España, 1965, pág. 81.
- 6) Krech, Crutchfield y E. L. Ballachey, op. cit., pág. 60-61.
- 7) Peter R. Hofstätter, *Introducción a la psicología social*, Ed. Luis Miracle, Barcelona, España, 1966, pág. 405.
- 8) Gordon W. Allport, *La personalidad. Su configuración y desarrollo*, Ed. Herder, Barcelona, España, 1966, pág. 605.
- 9) Solomon E. Asch, *Psicología social*, Eudeba, Buenos Aires, Argentina, 1962, pág. 217.
- 10) Gordon W. Allport, op. cit., pág. 586.
- 11) Krech, Crutchfield y Ballachey, op. cit., pág. 81.
- 12) Robert L. Thorndike, Elizabeth Hagen, *Measurement and Evaluation in Psychology and Education*, John Wiley & Sons Inc., 1967, pág. 361.
- 13) Gordon W. Allport, op. cit., pág. 589.
- 14) Sobre las cualidades que deben tener un buen juez y un buen examinador escolar: François Gorphe, "La psicología de los jueces y de los modos de juzgar", en H. Piéron, *Tratado de Psicología aplicada*, Ed. Kapelusz, Buenos Aires, Argentina, 1961, pág. 309-332; J. B. Piobetta, op. cit., pág. 83-84; Gordon W. Allport, op. cit., págs. 590-5.
- 15) Gordon W. Allport, op. cit., pág. 606.

---

## Anexo 2

### Puntuación corregida. Calificación de la prueba objetiva

Thorndike R, Hagen E.

---

Una decisión básica que debe tomarse antes de aplicarse la prueba es determinar si se va a hacer una corrección por adivinación o no. Si así es, debe decirse a los estudiantes, porque esto puede influir en su estrategia para contestar la prueba. Los expertos en pruebas no están en completo acuerdo acerca de la utilidad de la corrección por adivinación.

Desde hace mucho tiempo se sabe que hay marcadas diferencias entre los estudiantes en su tendencia a adivinar en una prueba objetiva cuando no están seguros de una respuesta. Estas diferencias en el patrón de adivinación incluyen algunas variaciones en las calificaciones que no están relacionadas con las diferencias reales en el aprovechamiento entre los estudiantes. Uno de los propósitos que se intentan al instruir a los examinados a no adivinar y al imponer una sanción por las respuestas incorrectas, es tratar de que la conducta de adivinar sea cada vez menos posible para todos los examinados.

Waters (1967) estudió los efectos de diferentes fórmulas de calificación sobre el número de reactivos omitidos de una prueba. Las fórmulas de calificación empleadas eran: sólo las respuestas correctas; las correctas menos las incorrectas; las correctas menos dos por las incorrectas, y las correctas menos cuatro por las incorrectas. Waters encontró que el número de omisiones aumentaba a medida que la sanción por las respuestas incorrectas se incrementaba, pero que los examinados que recibían sanciones de calificación más severas no omitían tantos reactivos como deberían si es que iban a lograr la calificación más alta. También encontró que el número promedio de reactivos correctos era más o menos el mismo para todos los patrones de calificación. Parece que ni la instrucción para no adivinar, con una amenaza de multa o de castigo por las respuestas incorrectas, ni la instrucción para contestar todas las preguntas, ya sea que se esté seguro de la respuesta o no, persuade a las personas a comportarse de la misma manera en una

prueba objetiva. La probabilidad de correr el riesgo parece depender mucho de la personalidad del individuo, así como de las instrucciones de la prueba.

La fórmula para la corrección por adivinación que se usa generalmente es

$$\text{Puntuación corregida} = R - \frac{W}{n - 1}$$

donde:

R es el número de preguntas contestadas correctamente,

W es el número de preguntas contestadas incorrectamente, y

n es el número de opciones de respuesta para cada reactivo.

Nótese que las preguntas para las que los examinados no dan la respuesta no cuentan en esta fórmula. En una prueba de falso - verdadero en la que sólo hay dos respuestas posibles,  $n - 1$  en la fórmula se convierte en un  $2 - 1$  (o sea, 1) y la corrección por adivinación es el número de respuestas correctas menos el número de respuestas incorrectas.

Los dos ejemplos que aparecen a continuación, uno basado en una prueba de falso - verdadero y el otro basado en una prueba de opción múltiple, muestran cómo funciona la fórmula.

#### Ejemplo 1

*Desempeño del estudiante*

<i>Tipo de prueba</i>	<i>n</i>	<i>R</i>	<i>W</i>	<i>Omisiones</i>
Falso - Verdadero	2	48	20	7

$$\text{Puntuación corregida} = 48 - \frac{20}{2 - 1} = 48 - \frac{20}{1} = 28$$

## Ejemplo 2

### *Desempeño del estudiante*

<i>Tipo de prueba</i>	<i>n</i>	<i>R</i>	<i>W</i>	<i>Omisiones</i>
Opción múltiple	5	48	20	7

$$\text{Puntuación corregida} = 48 - \frac{20}{5 - 1} = 48 - \frac{20}{4} = 43$$

En la fórmula se supone que una persona conoce la respuesta correcta o la adivina, es decir, que nunca responde sobre la base de estar erróneamente informado, y todas las opciones son igualmente atractivas para la persona que no conoce la respuesta. Estas suposiciones rara vez están justificadas. Por un lado, un buen constructor de pruebas formula opciones erróneas con el fin de que sean contestadas por los estudiantes mal informados; por el otro, la persona con una información parcial o con alguna experiencia en contestar reactivos objetivos normalmente puede eliminar una o dos opciones de su consideración. Por tanto, la corrección de la fórmula es cuando mucho un ajuste aproximado e inexacto que puede compensar las diferencias significativas en la capacidad para marcar los reactivos.

La adivinación de los reactivos presenta uno de los problemas más serios en las pruebas con restricción de tiempo que tienen reactivos de sólo dos opciones. Las calificaciones que se obtienen de este tipo de pruebas probablemente serían más confiables si se hiciera la corrección por adivinación. En las pruebas de opción múltiple con cuatro o cinco opciones y con tiempos suficientes para permitir que todos o casi todos los examinados intentaran contestar todos los reactivos, una puntuación que es simplemente el número de respuestas correctas resulta bastante satisfactoria y no hay ventaja, o sólo muy poca, si se aplica la corrección por adivinación.

Ha habido múltiples sugerencias para patrones de calificaciones más complejos para las pruebas objetivas, particularmente para los reactivos de opción múltiple. La mayoría de ellas incrementan el tiempo de calificación y, en cierta medida, también el tiempo de aplicación de la prueba. Algunas de ellas parecen producir ciertas ventajas en la precisión de la medición por reactivo, pero no queda claro que proporcionen una ganancia por minuto de tiempo de aplicación. Los lectores interesados en estas variaciones pueden consultar los tratamientos especiales que se han dado a este problema (Ebel, 1965; Davis, 1959; Combs y colaboradores, 1955).

---

# Anexo 3

## Ejemplos de hojas de respuesta

---

### HOJA DE RESPUESTAS

APELLIDO Y NOMBRE: .....

	A	B	C	D
1				
2				
3				
4				
5				
6				
7				
8				
9				
10				
11				
12				
13				
14				
15				
16				
17				
18				
19				
20				
21				
22				
23				
24				
25				
26				
27				
28				
29				
30				
31				
32				
33				
34				
35				
36				
37				
38				
39				
40				

	A	B	C	D
41				
42				
43				
44				
45				
46				
47				
48				
49				
50				
51				
52				
53				
54				
55				
56				
57				
58				
59				
60				
61				
62				
63				
64				
65				
66				
67				
68				
69				
70				
71				
72				
73				
74				
75				
76				
77				
78				
79				
80				

FIRMA: ..... FECHA: .....

AFACIMERA

Asociación de Facultades de Ciencias Médicas de la República Argentina



POR FAVOR LEA ATENTAMENTE ESTE INSTRUC TIVO DEL EXAMEN



- \* Use sólo lápiz de grafito blando
- \* No use bolígrafos, ni pluma, ni pizarra
- \* No falsifique ni reproduzca
- \* En caso de error, borrar cuidadosamente y rellene de nuevo

Marcas Incorrectas Marca Correcta



METODOLOGIA DOCENTE

EXAMEN 2000

Fecha de Entrega: 10 de Marzo de 2001

Nombre y Apellido

Two rows of empty boxes for writing name and surname.

Tipo y Nº de Documento

- D.N.I.
- C.L.
- L.C.
- L.P.

Seven empty boxes for document number.

Two rows of circles for marking document number, each row containing 10 circles.

Grid of 80 multiple-choice questions (numbered 1-80) arranged in two rows of 40. Each question has four options (A, B, C, D) represented by circles.

Developed by PDC MathWorks® PROCEMI® 01/01

S.C.B. S.A. ARGENTINA 10800

Verdadero - Falso: Coloque una cruz en el casillero correspondiente en cada una de las preguntas.

	1	2	3	4	5	6	7	8	9	10
Verdadero										
Falso										

Correspondencia

Pregunta N°	Escriba las opciones que correspondan
1	
2	
3	

---

## Anexo 4

### Ejemplo de instrucciones generales de la prueba

---

**FACULTAD DE .....**  
**EXAMEN FINAL**  
**AÑO .....**

VERSIÓN A

NOMBRE Y APELLIDO..... DNI:.....

DOCENTE:.....FECHA:.....

#### Instrucciones generales

- La prueba tiene 80 ítem.
- Cada ítem tiene una sola respuesta correcta.
- Usted dispone de 120 minutos para responder.
- Lea y responda las preguntas en el cuadernillo y luego pase las respuestas a la grilla de HOJA DE RESPUESTAS que figura en la última página.
- Tenga en cuenta que **sólo** se corregirá la grilla.

Si la corrección de los exámenes se hará electrónicamente se debe avisar que deberán completar la grilla u hoja de respuestas en el Formulario correspondiente con LAPIZ NEGRO N° 2.





# Anexo 5

## Propuesta para la adjudicación de notas

Lic Amanda Galli

A continuación se presenta una propuesta para convertir, al sistema de notas, los puntajes "brutos" o "crudos" obtenidos en un examen escrito estructurado.

La propuesta que se presenta es una combinación

- se basa en un criterio absoluto para decidir la aprobación y/o el aplazo.
- se basa en un criterio relativo para distribuir las notas de 4 a 10.

1.- Identificación de los alumnos aplazados según Criterio Absoluto.

El examen tiene 80 preguntas, cada una de ellas vale 1 punto, por lo tanto el puntaje máximo posible es 80 puntos. Si el nivel de exigencia es el 60% de respuestas correctas del máximo posible, todos los alumnos que obtienen 48 y más puntos han aprobado el examen.

En el ejemplo, los alumnos XZ y WK están aplazados porque no alcanzaron el 60% de respuestas correctas. El profesor puede decidir qué nota inferior a 4 adjudica. Se sugiere calificar con nota 2. Se puede informar a los estudiantes que todos los alumnos que no alcancen el nivel de aprobación recibirán la nota 2.

Los puntajes de los dos alumnos aplazados se sacan de la serie para la adjudicación de las notas de 4 a 10.

2.- Para asignar las notas de 4 a 10 se puede seguir el siguiente procedimiento:

- a) Ordenar de mayor a menor los puntajes de los alumnos que aprobaron. Deben ser excluidos los alumnos aplazados
- b) Calcular la mediana que es el valor que divide al grupo por la mitad, de manera tal que una mitad de los casos queda por encima de ella y la otra mitad por debajo. En el ejemplo la mediana es el puntaje 64.

### Examen A:

Alumno	Puntaje obtenido	Nota	
AP	78	10	
BZ	77		
CV	74	9	
DF	72		
HJ	72		
PL	71		
BM	70	8	
VP	70		
FG	70		
TR	69		
LB	68		
EG	68		
PC	67		
RC	66		
TM	65		7
<b>PZ</b>	<b>64</b>		
<b>AV</b>	<b>64</b>		
<b>ST</b>	<b>64</b>		
BU	63		
LC	63		
VM	62	6	
BK	61		
FC	59		
GH	57	5	
JR	55		
LG	53	4	
AN	52		
MB	50		
SS	48		
MA	48		
XZ	45	2	
WK	42		

- c) Adjudicar nota 7 al puntaje correspondiente a la mediana  $\pm 1$  punto.

En el ejemplo, los alumnos con puntajes de 63-64 y 65 reciben nota 7

d) Adjudicar las notas 8 - 9 y 10 .

Contar cuantos sujetos quedan en el grupo superior y dividir por 3 (son tres las notas a adjudicar). En el ejemplo: 14 alumnos se ubican por encima de la mediana  $\pm 1$  punto.  $14 \div 3 = 4,66$ . Siempre que obtenga una cifra con decimal redondee hacia arriba = 5

Cada 5 puntos más se adjudica la siguiente nota más alta: entre 66 y 70 puntos corresponde la nota 8; entre 71 y 75 corresponde nota 9 y entre 76 y 80 corresponde nota 10.

e) Adjudicar las notas 4 -5 y 6.

Contar cuantos sujetos quedan en el grupo inferior y dividir por 3 (son tres las notas a adjudicar). En el ejemplo: quedan 10 alumnos por debajo de la mediana  $\pm 1$ .

$10 \div 3 = 3,33$ . Siempre que obtenga una cifra con decimal redondee hacia arriba = 4

Cada 4 puntos menos se adjudica la siguiente nota más baja: entre 62 y 59 corresponde nota 6; entre 58 y 55 corresponde nota 5; todos los alumnos que queden por debajo de 5 les corresponde nota 4.

#### *Veamos otro ejemplo:*

Examen final, de 80 preguntas de selección múltiple, rendido por 25 alumnos. El nivel de exigencia para la aprobación también es 60% de respuestas correctas.

Siete alumnos resultan aplazados porque logran puntajes por debajo de los 48 puntos que marca el nivel de exigencia.

Aplicando el procedimiento ya descrito se pusieron las notas. El valor de la mediana es 58.

ATENCION: Es importante advertir que en ambos ejemplos algunos alumnos obtienen nota 10 independientemente de la cantidad de respuestas correctas.

En el Examen A los estudiantes que consiguen nota 10 habían respondido más del 95% del examen mientras que en el Examen B el alumno que obtiene nota 10 había respondido correctamente sólo el 85% de la prueba.

**Con el procedimiento descrito y ejemplificado siempre habrá alumnos con nota 10.**

#### **Examen B:**

Alumno	Puntaje obtenido	Nota
AP	68	10
BZ	67	9
CV	65	
DF	62	8
HJ	61	
PL	61	
BM	61	
VP	61	
<b>FG</b>	<b>59</b>	7
<b>TR</b>	<b>57</b>	
<b>LB</b>	<b>57</b>	
EG	56	6
PC	55	
RC	54	
TM	54	
PZ	54	
AV	52	5
ST	48	4
BU	47	2
LC	47	
VM	44	
BK	42	
FC	41	
GH	40	
JR	39	

Si el profesor considera que no es justo adjudicar nota 10 independientemente del porcentaje de respuestas correctas se puede modificar el procedimiento estableciendo un limite superior: sólo reciben nota 10 los alumnos que respondan correctamente el 95% y más del total de la prueba.

Esta segunda propuesta es una combinación

- se basa en un criterio absoluto para decidir el aplazo y para adjudicar la nota 10.
- se basa en un criterio relativo para distribuir las notas de 4 a 9.

*El procedimiento es el siguiente:*

c) Ordenar de mayor a menor los puntajes de los alumnos que aprobaron. Deben ser excluidos los alumnos aplazados.

- d) Calcular la mediana que es el valor que divide al grupo por la mitad, de manera tal que una mitad de los casos queda por encima de ella y la otra mitad por debajo. En el ejemplo A la mediana es 64. En el ejemplo B la mediana es 58.
- f) Adjudicar nota 7 al puntaje correspondiente a la mediana  $\pm 1$  punto.  
Examen A: los alumnos con puntajes de 63 - 64 y 65 reciben nota 7.  
Examen B: los alumnos con puntajes 57 - 58 y 59 reciben la nota 7.
- g) Adjudicar la nota 10 a los alumnos que hayan respondido el 95% y más de la prueba. En el examen A los alumnos AP y BZ han respondido correctamente, respectivamente, el 97% y el 96% de la prueba y les corresponde la nota 10.  
En el examen B ningún alumno alcanza la nota 10.
- g) Adjudicar las notas 8 y 9 dividiendo por dos la cantidad de alumnos que han quedado entre la nota 10 y la nota 7.
- h) Adjudicar las notas 4 -5 y 6.  
Contar cuantos sujetos quedan en el grupo inferior entre la nota 7 y el limite del aplazo y dividir por 3 (son tres las notas a adjudicar).



*Si Ud. quiere ampliar y profundizar este tema le sugerimos la siguiente Bibliografía.*

- Thorndike R. Hagen E. Medición y Evaluación en Psicología y Educación. Trillas Méjico. 1996 Cap. 15: Calificaciones y Forma de Calificar.

- Lafourcade P. Evaluación de los aprendizajes. Kapelusz. Buenos Aires 1969. Cap 11 El sistema de calificaciones.

- Van Dalen D. Meyer W. Manual de técnica de la investigación educacional. Paidós. Barcelona. 1981. cap 13 La Estadística Descriptiva.

Los resultados en uno y otro ejemplo quedarían modificados como se muestra a continuación.

**Examen B:**

Con limite superior para obtener nota 10: 95% de respuestas correctas

Alumno	Puntaje obtenido	Nota
AP	68	9
BZ	67	
CV	65	8
DF	62	
HJ	61	
PL	61	
BM	61	
VP	61	
FG	59	7
TR	57	
LB	57	
EG	56	6
PC	55	
RC	54	
TM	54	
PZ	54	
AV	52	5
ST	48	4
BU	47	2
LC	47	
VM	44	
BK	42	
FC	41	
GH	40	
JR	39	

De este modo la nota 10 queda reservada para los alumnos que tienen un excelente rendimiento en la prueba: 95% de respuestas correctas.

