

## CAPÍTULO DOS

---

# DISEÑO Y ELABORACIÓN DE TESTS

La cantidad de esfuerzo invertido en la elaboración de un test psicológico o educativo varía con el tipo de prueba y con los propósitos para los cuales se crea. Es probable que la mayoría de los maestros dedique relativamente poco tiempo a preparar pruebas de ensayo o de respuesta corta para evaluar el progreso de sus alumnos en una unidad de enseñanza. Por otro lado, las pruebas de habilidad y de personalidad diseñadas por especialistas en evaluación psicológica por lo general requieren del esfuerzo de muchos individuos que trabajan por períodos prolongados.

Los procedimientos empleados en la elaboración de una prueba también varían con el tipo de ésta y los propósitos de los usuarios. Preparar un inventario de lápiz y papel, de intereses o de características de la personalidad, implica problemas diferentes a los de construir una prueba de aprovechamiento o de aptitud. De igual modo, los complejos procedimientos seguidos por los diseñadores profesionales de pruebas son poco familiares para la mayoría de los maestros. Cualquiera que sea el tipo de prueba o las metas de los usuarios, se necesita cierto grado de planeación del contenido antes de escribir los reactivos que contendrá el instrumento. La planeación de la prueba deberá incluir definiciones claras de las variables o constructos que van a medirse, descripciones de las personas que van a ser examinadas, las condiciones bajo las cuales se administrará la prueba, e información concerniente a la calificación, interpretación de las puntuaciones y uso que se dará a los resultados.

### PLANEACIÓN DE UN TEST

La elaboración de un test requiere la consideración cuidadosa de sus propósitos específicos. Las pruebas cumplen muchas funciones diferentes, y su proceso de elaboración varía en cierto grado de acuerdo con el propósito que se pretenda lograr. Por ejemplo, se siguen procedimientos diferentes al elaborar pruebas de aprovechamiento, de inteligencia, de aptitud especial o un inventario de personalidad. Sin embargo, de manera ideal, la elaboración de una prueba u otro instrumento psicométrico empieza con la definición de las variables o constructos que van a medirse y con el esbozo del contenido propuesto.

#### Pruebas de observación

La elaboración de una prueba de aptitud para observar a solicitantes de un trabajo particular comienza con un análisis detallado de las actividades que componen ese trabajo. Un análisis de tareas, o *análisis de trabajo*, consiste en especificar los componentes del trabajo de modo que puedan construirse las situaciones de prueba o reactivos más adecuados para predecir el desempeño del empleado. Esas especificaciones pueden incluir *incidentes cruciales*, conductas que son decisivas para el desempeño exitoso o fallido, así como otra información que describa las actividades del trabajo. Dado que la descripción de un trabajo particular por lo general es larga y requiere de-

dicación, la prueba final no medirá todos los aspectos del desempeño del empleado. Tratará sólo con una muestra de los comportamientos más importantes relacionados con el trabajo en cuestión, muestra que en el mejor de los casos debe ser representativa de todas las tareas a realizar.

## Pruebas de inteligencia

En el capítulo 7 se describen con detalle los procedimientos empleados por los diseñadores de pruebas de inteligencia, por lo que aquí sólo se hará una breve descripción. Como en la elaboración de cualquier otra prueba, se reúne un conjunto de reactivos que supuestamente miden algún aspecto del constructo "inteligencia". Esos reactivos pueden ser elaborados de acuerdo con una teoría específica de la conducta inteligente o haciendo referencia sólo a los tipos de tareas que la gente muy inteligente puede realizar de manera más efectiva que las personas menos inteligentes. La selección de los reactivos a incluir en la prueba final puede hacerse con base en las relaciones de las respuestas dadas a reactivos con criterios tales como la edad cronológica, así como con las relaciones entre los reactivos de la prueba.

## Inventarios y escalas de personalidad

Al elaborar inventarios de personalidad y escalas de calificación se han empleado varios enfoques, algunos basados en el sentido común, otros en las teorías de personalidad y otros más en procedimientos estadísticos. Como se describe en los capítulos 16 y 17, muchos de los instrumentos de evaluación de la personalidad publicados recientemente han sido elaborados combinando enfoques teóricos, racionales y empíricos. Uno o más de estos enfoques pueden emplearse en diferentes etapas del desarrollo del instrumento.

## Pruebas de rendimiento

Se ha dedicado más atención a los procedimientos usados para elaborar pruebas de rendimiento académico que a los de otras clases de pruebas. Esto es comprensible cuando nos percatamos de que se aplican más pruebas de rendimiento que todos los otros tipos de pruebas combinados. A pesar del uso generalizado de las pruebas de rendimiento, la mayoría de los profesores, quienes supuestamente están familiarizados con su materia de estudio, no dedica tiempo suficiente a la evaluación del progreso de los estudiantes. Con mucha frecuencia los maestros consideran que los exámenes son algo desagradable adjunto a la enseñanza, en lugar de verlos como parte integral y formativa del proceso educativo. Sin embargo, cuando se usan de manera efectiva, los resultados de los exámenes no se limitan a la sola evaluación y motivación de los estudiantes. También proporcionan información a los maestros, al personal administrativo y a los padres, concerniente a la medida en que se han alcanzado los objetivos educativos específicos. Al proporcionar datos sobre la efectividad del currículo escolar y los procedimientos de enseñanza, las puntuaciones de los exámenes pueden contribuir significativamente en la planificación educativa para estudiantes individuales o grupos, incluso para distritos escolares enteros.

## Preguntas para las personas que planifican los instrumentos

Quienes planifican las pruebas de rendimiento de un salón de clases deben empezar por atender con cuidado las siguientes preguntas:

1. ¿Cuáles son los temas y materiales sobre los que se examinará a los estudiantes?
2. ¿Qué tipos de preguntas deben elaborarse?

3. ¿Qué formatos o esquemas de reactivos y pruebas deben utilizarse?
4. ¿Cuándo, dónde y cómo debe administrarse la prueba?
5. ¿Cómo debe calificarse y evaluarse la prueba resuelta?

Las preguntas 1, 2 y 3 se analizan en este capítulo, las preguntas 4 y 5 en el capítulo 3.

### Taxonomías de objetivos cognoscitivos

Así como elaborar una prueba de observación para usar en la selección de personal requiere un análisis preliminar del trabajo a desempeñar, la preparación de una prueba para medir objetivos instrucionales específicos es más efectiva cuando las conductas que van a evaluarse se definen claramente al inicio. Desde mediados de la década de 1950 se ha prestado mucha atención a los sistemas formales y estándar de clasificación de los objetivos cognoscitivos, afectivos y psicomotrices de la instrucción. En la tabla 2.1 se presentan las principales categorías de cuatro de esas taxonomías de objetivos cognoscitivos. Las seis principales categorías de la primera taxonomía, la *Taxonomía de objetivos educativos: el dominio cognoscitivo* (Bloom y Krathwohl, 1956), se presentan en orden de la más simple a la más compleja. Esas categorías no son exclusivas, sino más bien progresivamente in-

**TABLA 2.1 Compendio ilustrativo de los objetivos cognoscitivos**

---

#### Bloom y Krathwohl (1956)

Conocimiento  
Comprendión  
Aplicación  
Análisis  
Síntesis  
Evaluación

#### Educational Testing Service (1965)

Memoria  
Comprendión  
Pensamiento

#### Ebel (1979)

Comprendión de la terminología (o vocabulario)  
Comprendión del hecho y del principio (o generalización)  
Habilidad para explicar o ilustrar (comprendión de relaciones)  
Habilidad para calcular (problemas numéricos)  
Habilidad para predecir (qué es probable que suceda bajo condiciones especificadas)  
Habilidad para recomendar la acción apropiada (o algunas situaciones de problemas prácticos específicos)  
Habilidad para formular un juicio evaluativo

#### Gerlach y Sullivan (1967)

Identificación  
Nominación  
Descripción  
Elaboración  
Ordenamiento  
Demostración

---

clusivas. Por ejemplo, tanto el Conocimiento (categoría I) como la Comprensión (categoría II) son esenciales para la Aplicación (categoría III) y por ende están incluidas en la tercera categoría. En la tabla 2.2 se presenta una descripción de las categorías registradas en esta taxonomía.

Otra taxonomía presentada en la tabla 2.1, la propuesta por Gerlach y Sullivan (1967), enfatiza la conducta del examinado en la identificación, nominación, descripción, elaboración, ordenamiento o demostración de algo. La *identificación* consiste en indicar qué miembro de un conjunto pertenece a una categoría particular. En la *nominación* debe proporcionarse la etiqueta verbal correcta para un referente o conjunto de referentes. La *descripción* consiste en reportar categorías relevantes de objetos, eventos, propiedades o relaciones. En la *elaboración* se crea un producto de acuerdo con ciertas especificaciones. El *ordenamiento* requiere arreglar en un orden específico dos o más referentes, y la *demonstración* consiste en realizar ciertas acciones para cumplir una tarea especificada.

La aplicación de cualquiera de las taxonomías presentadas en la tabla 2.1 debe alentar a la persona que diseña la prueba a ir más allá de los reactivos que miden el reconocimiento simple o la memoria, y a preparar reactivos que midan objetivos educativos de orden superior y demanden reflexión. Los siguientes reactivos, que pueden presentarse en un formato de ensayo o de prueba objetiva, ejemplifican lo anterior:

¿Cuál es la fórmula para calcular el error estándar de medición? (*Conocimiento*)

Examine la siguiente gráfica y determine cuántos reactivos deben agregarse a una prueba de 50 reactivos para aumentar su confiabilidad de .60 a .80. (*Comprensión*)

**TABLA 2.2 Categorías de la Taxonomía de objetivos educativos: el dominio cognoscitivo**

- I. *Conocimiento* implica el recuerdo de hechos específicos. Los verbos de muestra en los reactivos de conocimiento son *definir, identificar, mencionar y nombrar*. Ejemplo de un reactivo de conocimiento es: “Mencione las seis categorías principales de *La taxonomía de objetivos educativos: el dominio cognoscitivo*”.
- II. *Comprensión* significa entender el significado o propósito de algo. Los verbos de muestra en los reactivos de comprensión son *convertir, explicar y resumir*. Ejemplo de un reactivo de comprensión es: “Explique lo que quiere decir el revisor de la prueba cuando dice que ésta no es confiable”.
- III. *Aplicación* implica utilizar la información y las ideas en situaciones nuevas. Los verbos de muestra en los reactivos de aplicación son *calcular, determinar y resolver*. Ejemplo de un reactivo de aplicación es: “Calcule la media y la desviación estándar del siguiente grupo de calificaciones”.
- IV. *Análisis* es descomponer algo para revelar su estructura y las interrelaciones que hay entre sus partes. Los verbos de muestra en los reactivos de análisis son *analizar, diferenciar y relacionar*. Ejemplo de un reactivo de análisis es: “Analice esta unidad instruccional en varias categorías conductuales y de contenido”.
- V. *Síntesis* es combinar varios elementos o partes en un todo estructural. Los verbos de muestra en los reactivos de síntesis son *diseñar, crear, formular y planificar*. Ejemplo de un reactivo de síntesis es: “Diseñe una tabla de especificaciones para una prueba de estadística elemental”.
- VI. *Evaluación* es formular un juicio basado en el razonamiento. Los verbos de muestra en los reactivos de evaluación son *comparar, criticar, evaluar y juzgar*. Ejemplo de un reactivo de evaluación es: “Evalúe el procedimiento usado en la estandarización de esta prueba”.

Calcule el error estándar de estimación para una prueba que tiene una correlación de .70 con un criterio que tiene desviación estándar de 10. (Aplicación)

Distinga entre una prueba de rendimiento para el salón de clase y una prueba estandarizada de rendimiento en términos de lo que mide cada una y la manera en que se utilizan. (Análisis)

Formule una teoría que relacione los intereses con la personalidad y mencione la evidencia de investigación apropiada que la apoya. (Síntesis)

Evalúe las críticas concernientes al contenido y los usos del SAT. (Evaluación)

## Objetivos afectivos y psicomotrices

Una función importante de la educación es inculcar en los estudiantes ciertas actitudes, valores y otros estados afectivos. No existe un método que sea completamente satisfactorio para clasificar los objetivos afectivos de la instrucción, pero se ha propuesto todavía una serie de sistemas de clasificación. Un ejemplo es la *Taxonomía de objetivos educativos: dominio afectivo* (Krathwohl, Bloom y Masia, 1964). Las principales categorías de esta taxonomía son: I. Recibir o atender. II. Responder o participar. III. Valorar o creer en el valor de algo. IV. Organizar los valores en un sistema. V. Caracterización mediante un valor o valor complejo. En comparación con su contraparte en el dominio cognoscitivo, esta taxonomía no se ha aplicado con mucha frecuencia.

También se han propuesto taxonomías de objetivos educativos en el dominio psicomotriz (por ejemplo, Harrow, 1972; Nixon y Jewett, 1980; Simpson, 1966). Las seis categorías en la *Taxonomía del dominio psicomotriz* de Harrow, por ejemplo, son: movimientos reflejos, movimientos básicos fundamentales, habilidades perceptuales, habilidades físicas, movimientos hábiles y comunicación no discursiva. Los niveles inferiores de la taxonomía de Nixon y Jewett se interesan en la percepción de los componentes de un movimiento y en los esfuerzos de los examinados por repetirlo o recordarlo después de que se ha demostrado. Los niveles superiores ponen de relieve la creación de movimiento para una situación específica en los deportes, la danza u otras actividades físicas. Al aplicar dicha taxonomía, deben tomarse decisiones con respecto a los pesos numéricos que se asignarán a cada componente del desempeño y a si se harán deducciones por errores, torpezas y falta de pulcritud.

## Tabla de especificaciones

La mayoría de los diseñadores de pruebas no se adhiere rígidamente a una taxonomía formal al especificar los objetivos que van a medirse. No obstante, al planificar una prueba es útil construir una tabla de especificaciones de dos vías. En dicha tabla, los objetivos conductuales que van a evaluarse se presentan en los encabezados de renglón y los objetivos de contenido (temáticos) como encabezados de columna. Luego se escriben en el cuerpo (celdas) de la tabla las descripciones de los reactivos específicos que caen bajo los encabezados apropiados de renglón y columna.

Una tabla de especificaciones debe ser razonablemente detallada en términos del conocimiento y las habilidades que se espera demuestren los examinados, pero es importante no enfatizar en exceso un objetivo particular. Por ejemplo, puede ser más sencillo elaborar reactivos que evalúen el conocimiento de términos y hechos que reactivos donde se mida la habilidad de analizar y evaluar, pero en la prueba deben incluirse reactivos de las dos últimas categorías.

La tabla 2.3 es una tabla de especificaciones para una unidad sobre preparación, aplicación y análisis de reactivos de pruebas. Advierta que el número total de reactivos que se dedica

a cada tema aparece entre paréntesis debajo del mismo. Una vez que se ha determinado un conjunto de objetivos para un curso o unidad de instrucción y que se ha preparado el bosquejo temático, pueden elaborarse los reactivos de la prueba para medir el grado en que los estudiantes han alcanzado los objetivos mencionados para cada tema.

Ciertos tipos de reactivos de prueba son más apropiados que otros para medir la obtención de objetivos específicos. Los reactivos de respuesta corta y de completamiento son adecuados para evaluar el conocimiento de la terminología, pero inadecuados para evaluar habilidades cognoscitivas de orden superior. Por esta razón, la tabla de especificaciones para una prueba debe ser inspeccionada con cuidado antes de decidir qué tipo de reactivos y cuántos de cada uno son apropiados. Al planifi-

**TABLA 2.3 Especificaciones para una prueba sobre preparación y administración de pruebas**

OBJETIVO CONDUCTUAL	CONTENIDO (TEMA)				
	Preparación	Elaboración	Aplicación	Calificación	Análisis de reactivos
Conocimiento de la terminología	Análisis de trabajo; incidentes críticos; muestra representativa (3 reactivos)	Reactivo de aparajamiento; colectivo en espiral; grupo de respuestas (5 reactivos)	Rapport; efecto de halo (2 reactivos)	Clave de lista; calificación compuesta; calificación con máquina (3 reactivos)	Criterio; consistencia interna; homogeneidad de la prueba (3 reactivos)
Conocimiento de hechos específicos	Categorías en la <i>Taxonomía de objetivos educativos</i> (2 reactivos)	Ventajas y desventajas de los reactivos de ensayo y de los reactivos objetivos (4 reactivos)	Factores que afectan el desempeño en la prueba (3 reactivos)	Reglas para calificar pruebas de ensayo y pruebas objetivas (3 reactivos)	Métodos para determinar la validez de los reactivos; propósitos del análisis de reactivos (3 reactivos)
Comprensión	Explicación de los propósitos de hacer plan de prueba (2 reactivos)	0 reactivos	0 reactivos	Efectos de la ponderación de los reactivos sobre la calificación total (1 reactivo)	Explicación de la relación entre <i>p</i> y <i>D</i> (1 reactivo)
Aplicación	Especificaciones para una unidad sobre examinación (1 reactivo)	Ejemplos de reactivos de opción múltiple para medir comprensión, aplicación, análisis, síntesis y evaluación (4 reactivos)	Instrucciones para una prueba (2 reactivos)	Corrección para la adivinación; ponderación de confianza; uso del nomograma para calificar los reactivos de reordenamiento (4 reactivos)	Cálculo de los índices de dificultad y discriminación; distribución de las respuestas a los distractores (4 reactivos)
Total	8 reactivos	13 reactivos	7 reactivos	11 reactivos	11 reactivos

car una prueba también es necesario considerar cuestiones prácticas como el costo, el tiempo disponible para la administración, la disposición de los reactivos y las condiciones de la prueba.

## PREPARACIÓN DE LOS REACTIVOS DEL TEST

La meta principal de la planificación de la prueba es la preparación de un bosquejo detallado, como una tabla de especificaciones, que sirva como guía al elaborar los reactivos con los que se van a evaluar o predecir ciertos objetivos. Una vez preparada dicha tabla o el bosquejo detallado del contenido de la prueba, el siguiente paso es elaborar los reactivos. Por lo general, es recomendable que, en las pruebas objetivas, inicialmente se prepare alrededor de 20% más de reactivos de los que en realidad se necesiten, de modo que se disponga de una cantidad adecuada de buenos reactivos para la versión final de la prueba. Las organizaciones comerciales de tests, como el Educational Testing Service, emplean como elaboradores de reactivos a personas que poseen un conocimiento profundo de la materia de la prueba y destreza suficiente para la creación de reactivos. Cualquier persona que desee aprender cómo elaborar buenos reactivos puede beneficiarse al inspeccionar una muestra de reactivos de las pruebas publicadas, ya que éstos se encuentran entre los mejores disponibles.

Todos los reactivos representan procedimientos para obtener información acerca de los individuos, pero la cantidad y los tipos de información varían con la naturaleza de las tareas planteadas por diferentes tipos de reactivos. Pedir a los examinados que comparan la Batalla de Bulge con la Batalla de Hastings requiere un tipo de respuesta diferente a la que se obtiene cuando se les pide señalar, de entre una serie de acontecimientos, los que ocurrieron en cada batalla. En el primer reactivo se requieren habilidades de organización e integración complejas, mientras que sólo se necesita memoria de reconocimiento para responder al segundo.

Se han sugerido varios métodos para clasificar los reactivos de acuerdo con el formato o la forma de la respuesta requerida. *Completamiento o llenado contra selección, recuerdo contra reconocimiento, y construcción de respuesta contra identificación* son formas de diferenciar entre los reactivos donde se pide a los examinados que escriban o construyan una respuesta y aquellos en que se les pide señalar cuál de varias alternativas es correcta. Otro método popular de clasificación de reactivos es *ensayo contra objetivo*, de los cuales se presentan ejemplos en la tabla 2.4. Todos los reactivos de ensayo son del tipo de completamiento o llenado, donde la respuesta del examinado se construye en lugar de ser meramente identificada.

Un reactivivo objetivo puede ser del tipo de completamiento o llenado, o de selección, dependiendo de si se pide que los examinados construyan una respuesta o seleccionen la mejor respuesta de entre una lista de alternativas. El rasgo crucial de los reactivos objetivos no es la forma de la respuesta, sino la objetividad con la que pueden calificarse. Dos o más calificadores de un reactivivo de ensayo pueden estar en desacuerdo en si una respuesta dada es correcta y en cuántos puntos debería recibir. Sin embargo, salvo que ocurran errores administrativos, los diferentes calificadores de una prueba objetiva asignarán la misma calificación a una determinada prueba.

### Reactivos de ensayo

La ventaja principal de los reactivos de ensayo es que pueden medir la habilidad personal para organizar, relacionar y comunicar, conductas que no son fáciles de evaluar con los reactivos objetivos. Las pruebas de ensayo tienen las ventajas de que requieren menos tiempo para su elaboración y reducen la probabilidad de que los examinados respondan en forma correcta a los reactivos por

**TABLA 2.4 Ejemplos de varios tipos de reactivos de test****I. Reactivos de ensayo**

*Instrucciones:* Escriba una respuesta de media página para cada uno de los siguientes reactivos.

1. Compare las ventajas y desventajas de los reactivos de ensayo y los reactivos objetivos.
2. Explique las razones para realizar un análisis de reactivos en una prueba para el salón de clases.

**II. Reactivos objetivos****A. Respuesta corta**

*Instrucciones:* Escriba la(s) palabra(s) apropiada(s) en cada espacio.

1. La única cosa que es objetiva acerca de una prueba objetiva es \_\_\_\_\_.
2. ¿Cuál es el primer paso formal en la elaboración de una prueba para predecir el desempeño laboral? \_\_\_\_\_.

**B. Verdadero-falso**

*Instrucciones:* Encierre V en un círculo si la afirmación es verdadera; encierre F en un círculo si la afirmación es falsa.

- |   |   |  |
|---|---|--|
| V | F | 1. El sistema de clasificación de pruebas más global es el de <i>The Mental Measurements Yearbooks</i> .   |
| V | F | 2. El grupo de respuesta de deseabilidad social es la tendencia a dar una calificación alta a un examinado en un rasgo sólo porque obtuvo una calificación alta en otro rasgo. |

**C. Aparejamiento**

*Instrucciones:* Escriba la letra correspondiente al nombre correcto, de la lista que aparece en la segunda columna, en el espacio apropiado de la línea al margen de la primera columna.

- |       |                                      |              |
|-------|--------------------------------------|--------------|
| _____ | 1. prueba colectiva de inteligencia  | A. Binet     |
| _____ | 2. prueba individual de inteligencia | B. Darwin    |
| _____ | 3. inventario de intereses           | C. Galton    |
| _____ | 4. inventario de personalidad        | D. Otis      |
| _____ | 5. correlación producto-momento      | E. Pearson   |
| _____ | 6. pruebas sensoriomotrices          | F. Rorschach |
|       |                                      | G. Spearman  |
|       |                                      | H. Strong    |
|       |                                      | I. Woodworth |

**D. Opción múltiple**

*Instrucciones:* Escriba la letra de la opción correcta en el espacio del margen al lado del reactivio.

- \_\_\_\_\_ 1. Los adverbios como *nunca*, *en ocasiones* y *siempre*, que revelan la respuesta a un examinado que no tiene información sobre la materia del reactivio, se llaman
  - a. generalidades brillantes
  - b. adverbios de enlace
  - c. grupos de respuesta
  - d. determinantes específicos
- \_\_\_\_\_ 2. Jimmy, quien tiene 8 años 4 meses de edad, obtiene una calificación de edad mental de 9 años 5 meses. ¿Cuál es su razón CI en la prueba?
  - a. 88
  - b. 90
  - c. 113
  - d. 120

simple adivinación. Sin embargo, las preguntas de ensayo pueden ser tan generales que se interpretan de manera muy diferente por distintas personas. Además, el número de preguntas de ensayo que pueden ser respondidas con respuestas de media página en una clase típica de 50 minutos (aproximadamente seis) puede ser insuficiente para determinar el conocimiento que tiene una persona de la materia de la prueba. No deberíamos esperar que las respuestas fueran tan inclusivas o detalladas como las requeridas por los reactivos del cuadro 2.1, pero podrían alcanzar cierta comprensión relativamente profunda del material. Otras desventajas de las pruebas de ensayo es que son susceptibles al engaño por parte de individuos con facilidad de palabra que no cuentan con información; además, su calificación es subjetiva y toma mucho tiempo.

Un profesor de historia informó haber aplicado una prueba de ensayo que incluía la pregunta: “¿Cuáles fueron las causas y las consecuencias de la Batalla de Hastings?” Y un estudiante apático, cuya preparación sobre la historia de Inglaterra no incluía los acontecimientos anteriores al siglo XIV, empezó a responder la pregunta con la afirmación: “No puedo comentar sobre la Batalla de Hastings, pero dirijamos nuestra atención hacia la Guerra de los Cien Años”. Éste es un ejemplo bastante ostensible de la tendencia que los examinados no informados tienen a responder una pregunta ligeramente diferente de la que fue planteada para enfatizar lo que saben, en lugar de lo que no saben. Una manera de enfrentar este problema, aunque laborioso para los que presentan la prueba y para los que la califican, es el famoso procedimiento chino que consiste en hacer que los estudiantes escriban todo lo que saben del tema. Es posible que lo medido por esa prueba sea la susceptibilidad a la fatiga más que el conocimiento general.

Como regla, no deberán usarse reactivos de ensayo cuando el mismo conocimiento o habilidad puedan ser evaluados con reactivos objetivos. Si se plantean preguntas de ensayo, la persona que redacta los reactivos debe tratar de hacer las preguntas de manera tan objetiva como sea

## CUADRO 2.1

### ¿UN EXAMEN FINAL GLOBAL PARA LOS ALUMNOS UNIVERSITARIOS DEL ÚLTIMO AÑO?

1. Describa la historia del pontificado desde sus orígenes hasta el presente, concentrándose especialmente, pero no de manera exclusiva, en su impacto social, político, económico, religioso y filosófico en Europa, Asia, África y América.
2. Con base en el conocimiento que tenga usted de sus obras, evalúe la estabilidad emocional, el grado de ajuste y las frustraciones reprimidas de Alejandro de Afrodísias, Ramsés II, Gregorio de Nícea y Hammurabi. Apoye sus respuestas con citas del trabajo de cada uno de estos hombres, mencionando las referencias apropiadas.
3. Desarrolle un plan realista para reducir la deuda nacional. Identifique los efectos de su plan sobre el cubismo, la controversia donatista y la teoría de las ondas de la luz. Bosqueje un método para prevenir esos efectos. Critique este método desde todos los puntos de vista posibles. Señale las deficiencias en su punto de vista, según lo demuestra su respuesta a la pregunta anterior.
4. Bosqueje el desarrollo del pensamiento humano. Estime su relevancia y compárela con el desarrollo de cualquier otra clase de pensamiento —animal o alienígena.
5. Suponga que 2 500 aborígenes amotinados y enloquecidos están asaltando el aula. ¿Cómo los calmaría usted? Puede usar cualquier idioma antiguo excepto el latín o el griego y cualquier técnica no verbal diferente a la violencia.
6. Tome una posición a favor o en contra de la lógica y la verdad. ¿Cómo probaría la validez de su posición sin involucrar a nadie más o sin poner en peligro su propia salud?

posible. Esto puede lograrse al (1) definir la tarea y redactar los reactivos de manera clara, por ejemplo, pedir a los examinados que comparen y expliquen en lugar de que discutan; (2) usar un número pequeño de reactivos que deberán responder todos los examinados; (3) estructurar los reactivos de forma que los expertos en la materia estén de acuerdo en que puede demostrarse que una respuesta es mejor que otra, y (4) hacer que los examinados respondan a cada reactivo en una hoja por separado.

### **Reactivos de respuesta corta, de verdadero y falso y de aparejamiento**

Los reactivos objetivos no se limitan a los cuatro tradicionales (respuesta corta o completamiento, verdadero y falso, aparejamiento y opción múltiple), pero éstos son los más populares. Entre las ventajas atribuidas a las pruebas objetivas está el que pueden calificarse de manera fácil e imparcial y que, como se requiere menos tiempo para responder a cada reactivo, puede hacerse un muestreo más amplio del contenido que en las pruebas de ensayo. Al preparar las pruebas objetivas debe tenerse cuidado de lograr que los reactivos resulten claros, precisos y gramaticalmente correctos. Deben escribirse en un lenguaje adecuado para el nivel de lectura de las personas a las que se dirigen. Debe incluirse en el reactivo toda la información y los requerimientos necesarios para seleccionar una respuesta razonable, omitiendo las palabras y frases no funcionales o estereotipadas.

Resulta tentador elaborar reactivos objetivos mediante la copia literal de afirmaciones de un texto o de otras fuentes, pero esta práctica sólo enfatiza la memoria. Las personas que redactan reactivos también deben tener cuidado de no incluir claves para las respuestas correctas y evitar los reactivos interrelacionados o entrelazados. Dos reactivos están *interrelacionados* cuando el planteamiento de uno proporciona una señal para la respuesta del otro. Dos reactivos están *entrelazados* cuando es necesario conocer la respuesta a uno de ellos para llegar a la respuesta correcta del otro.

**Reactivos de respuesta corta.** Un reactivo de respuesta corta o de completamiento plantea una tarea tipo fuente, en la cual se requiere que los examinados completen o llenen uno o más espacios en blanco de una afirmación incompleta con las palabras o frases correctas, o que den una respuesta breve a una pregunta. En términos de la longitud de la respuesta elaborada, los reactivos de respuesta corta caen entre los reactivos de ensayo y los de reconocimiento. Los reactivos de respuesta corta se encuentran entre los más sencillos de elaborar, y requieren que los examinados proporcionen la respuesta correcta en lugar de simplemente reconocerla. Aunque son especialmente útiles para evaluar el conocimiento de la terminología, los reactivos de respuesta corta tienen serias limitaciones: son inapropiados para medir objetivos instruccionales complejos y, debido a que puede haber más de una respuesta correcta, la calificación no siempre es por completo objetiva.<sup>1</sup>

Al elaborar reactivos de respuesta corta deberán seguirse las siguientes directrices:

1. Las preguntas directas son preferibles a las afirmaciones incompletas.
2. Plantee los reactivos de forma que las respuestas sean breves y no ambiguas.
3. Si se utiliza una afirmación incompleta, coloque el espacio en blanco al final de la afirmación.

<sup>1</sup>Un tipo de reactivo de completamiento diseñado para evaluar la habilidad de lectura es la *técnica cloze*. En este procedimiento se instruye a los individuos para que reemplacen las palabras faltantes que han sido borradas al azar en determinados párrafos. Una medida de la habilidad de lectura del individuo es el grado en que puede llenar correctamente los espacios en blanco y dar así sentido a los pasajes.

4. Haga que todos los espacios en blanco sean de la misma extensión.
5. Evite usar múltiples espacios en blanco en el mismo reactivó, en especial si tornan poco claro el significado de la tarea.
6. Indique las unidades en que deben expresarse las respuestas numéricas.

**Reactivos de verdadero y falso.** Uno de los tipos de reactivos que es más sencillo de elaborar, pero probablemente el más criticado por los examinadores profesionales, es el de verdadero y falso. Los reactivos de verdadero y falso pueden escribirse y leerse con rapidez y, por ende, permiten un muestreo amplio del contenido de la materia. Un defecto notorio de los reactivos de verdadero y falso es que a menudo se interesan en información trivial o se elaboran copiando afirmaciones literales de un texto. En consecuencia, se dice que alientan la memorización y así encaminan mal los esfuerzos por aprender. Otra crítica a estos reactivos es que a menudo son ambiguos y no pueden usarse para medir objetivos instrucionales más complejos. Además, debido a que la calificación total en una prueba de este tipo puede ser afectada por la tendencia del examinado a adivinar cuando tiene dudas o a estar de acuerdo (o en desacuerdo), la precisión de la calificación puede ser cuestionable.<sup>2</sup>

En promedio, los examinados obtendrán un 50% de aciertos en los reactivos de verdadero y falso simplemente por adivinar. Las calificaciones pueden ser infladas todavía más cuando los reactivos contienen *determinantes específicos* —palabras como *todos*, *siempre*, *nunca* y *sólo*—, los cuales indican que la afirmación probablemente es falsa, o palabras como *a menudo*, *en ocasiones* y *usualmente*, sugerentes de que la afirmación es verdadera.

A pesar de esos defectos, los reactivos de verdadero y falso no tienen que ser triviales o ambiguos o encaminar mal el aprendizaje. En defensa de los reactivos de verdadero y falso, Ebel (1979) afirma que el grado de dominio que tienen los estudiantes en un área particular del conocimiento es indicado por su éxito al juzgar la veracidad o falsedad de proposiciones relacionadas con él (p. 112). Él ha considerado que tales proposiciones son expresiones del conocimiento verbal, que es la esencia del logro educativo.

La defensa que hizo Ebel de los reactivos de verdadero y falso puede ser cuestionada, pero no se cuestiona el hecho de que estos reactivos, bien diseñados, pueden medir más que la simple memoria. Por ejemplo, al incluir dos conceptos, condiciones o eventos en un reactivó de verdadero y falso, el examinador puede preguntar si es verdad que tienen una relación de moderada a fuerte (Diekhoff, 1984). Otras posibilidades son preguntar si (1) un concepto, condición o evento implica o es una consecuencia de otro evento; (2) un concepto, condición o evento es un subconjunto, ejemplo o categoría de otro evento, (3) ambos conceptos, condiciones o eventos son verdaderos. Dichos reactivos pueden medir la comprensión así como el conocimiento significativo de conceptos y eventos.

Cualesquiera que sean los objetivos de una prueba de verdadero y falso, al elaborar reactivos de este tipo es recomendable atender las siguientes sugerencias:

1. Asegúrese de que las afirmaciones planteen asuntos importantes (no triviales).
2. Establezca afirmaciones relativamente cortas, y verdaderas o falsas sin lugar a dudas.
3. Evite los reactivos planteados de manera negativa, especialmente los que contienen doble negación.
4. Evite los reactivos ambiguos y capciosos.

<sup>2</sup>La tendencia a estar de acuerdo cuando se tiene duda (o conformidad) es un grupo de respuesta. Los *grupos de respuestas* son las tendencias por parte de los examinados a responder a los reactivos de una prueba de acuerdo con su forma, es decir, a la manera en que están planteados, en lugar de hacerlo con base en su contenido.

5. Como regla, evite los determinantes específicos. Si se usan determinantes específicos para hacer que se equivoquen las personas sin conocimientos, pero hábiles para presentar pruebas, deben incluirse en las afirmaciones verdaderas tan a menudo como en las falsas.
6. En las afirmaciones de opinión, cite la fuente.
7. Haga que las afirmaciones verdaderas y las falsas sean aproximadamente de la misma longitud, y que el número de afirmaciones verdaderas sea aproximadamente igual al de las falsas. Puede argumentarse que, dado que los reactivos falsos tienden a discriminar más que los reactivos verdaderos, el número de afirmaciones falsas debería ser mayor que el de afirmaciones verdaderas. Sin embargo, si el maestro sigue esta práctica en pruebas sucesivas, los estudiantes pueden darse cuenta de ello y comenzar a responder "falso" cuando tengan duda acerca de la respuesta.
8. Asegúrese de que las respuestas erróneas sean más atractivas planteando los reactivos de tal manera que la lógica superficial, los errores populares o los determinantes específicos sugieran que las respuestas erróneas son correctas. Las afirmaciones falsas que parecen verdaderas también pueden hacer que se equivoquen los examinados sin conocimientos.

**Reactivos de aparejamiento.** Tanto los reactivos de verdadero y falso como los de opción múltiple son, en cierto sentido, variedades de los reactivos de aparejamiento. En estos tres tipos de reactivos, un conjunto de opciones de respuesta se equipara con un conjunto de opciones de estímulo (premisas). La distinción es que los reactivos de verdadero y falso y los de opción múltiple tienen sólo una premisa (el *tronco* del reactivo) y dos o más opciones de respuesta, mientras que los reactivos de aparejamiento tienen múltiples premisas y múltiples opciones de respuesta. La tarea del examinado en un reactivo de aparejamiento es acoplar las opciones de respuesta con la premisa correcta. El aparejamiento usualmente es de uno a uno (una respuesta por premisa), pero también puede ser de una respuesta a varias premisas, de varias respuestas a una premisa, o de varias respuestas a varias premisas. Por supuesto, debe informarse a los examinados cuál de esos procedimientos se aplica en un reactivo particular.

Los reactivos de aparejamiento son más sencillos de elaborar y cubren el material de manera más eficiente que muchos otros tipos de reactivos; por desgracia, usualmente sólo miden la memorización de acontecimientos.<sup>3</sup> Además, la necesidad de hacer que las opciones sean homogéneas (que todas las opciones de respuesta sean del mismo tipo, como fechas, lugares o nombres) limita el tipo de material que puede adaptarse a un marco de aparejamiento. A continuación se presentan algunos lineamientos para elaborar reactivos de aparejamiento:

1. Ordene la premisa y las opciones de respuesta en un formato claro y lógico de columnas, con las premisas en la columna izquierda y las opciones de respuesta en la columna derecha.
2. Use entre seis y quince premisas, con dos o tres opciones de respuesta más que premisas.
3. Numere las premisas de manera sucesiva, y coloque letras (a, b, c, etc.) antes de las opciones de respuesta.
4. Especifique con claridad las bases para realizar el aparejamiento.
5. Coloque todo el reactivo en una sola página.

Un tipo especial de reactivo de aparejamiento es el *reactivo de reordenamiento*, en el cual se requiere que los examinados clasifiquen un número fijo de categorías predeterminadas. En un

<sup>3</sup>Al menos un estudio encontró que los reactivos de aparejamiento pueden diseñarse para ser iguales o incluso superiores a los de opción múltiple como medidas tanto del dominio del contenido de interés como de las actitudes de las personas que presentan la prueba (Shaha, 1984).

tipo particular de reordenamiento conocido como *reactivo de rango*, los individuos reordenan un conjunto de opciones en orden de la primera a la última (o de la más alta a la más baja).

## Reactivos de opción múltiple

No se sabe quién elaboró el primer reactivo de opción múltiple para una prueba, pero desde el punto de vista de la evaluación psicológica fue algo fortuito.<sup>4</sup> Los reactivos de opción múltiple son los más versátiles de todos los reactivos objetivos, ya que pueden usarse para medir logros de aprendizaje simples y complejos en todos los niveles y en todas las áreas temáticas. Aunque los reactivos de respuesta de ensayo demandan mayor habilidad de organización que la selección de respuestas a los reactivos de opción múltiple, responder de manera correcta a un reactivo de opción múltiple bien preparado requiere buena habilidad para discriminar y no sólo capacidad para reconocer o recordar la respuesta correcta. Las calificaciones en los reactivos de opción múltiple también son menos afectadas por la adivinación y por otros grupos de respuesta que las calificaciones en otros reactivos objetivos. Además, puede obtenerse información de diagnóstico útil a partir de un análisis de las opciones incorrectas (*distractores*) seleccionadas por los examinados.

Entre los defectos de los reactivos de opción múltiple están que (1) los buenos son difíciles de elaborar, en especial aquellos en los que todas las opciones resulten igualmente atractivas para los examinados que no conocen la respuesta correcta; (2) enfatizan el reconocimiento más que el recuerdo y la organización de la información, y (3) requieren más tiempo para la respuesta y pueden muestrear el dominio temático de manera menos adecuada que los reactivos de verdadero y falso. También se ha argumentado, pero no demostrado, que las pruebas de opción múltiple favorecen a los lectores sagaces, hábiles y rápidos, y penalizan a los personas más reflexivas y que piensan con más profundidad (Hoffman, 1962).

En el cuadro 2.2 se presentan lineamientos para facilitar la elaboración de reactivos de opción múltiple de alta calidad. Tales lineamientos son sobre todo producto de la lógica y de la experiencia, más que de la investigación, y su seguimiento no garantiza la elaboración de buenas pruebas de opción múltiple. La elaboración de buenos reactivos depende mucho o más que del conocimiento de la materia de la prueba, de la comprensión de lo que los estudiantes deberían saber y de lo que es poco probable que sepan acerca de la materia, y del arte o habilidad de plantear preguntas. Incluso cuando los lineamientos no se siguen con precisión, los reactivos de opción múltiple tienden a ser bastante sólidos en su capacidad para medir el conocimiento y la comprensión.

**Elaboración de distractores.** Un factor crucial en la determinación de la efectividad de los reactivos de opción múltiple es la selección o elaboración de los elementos distractores (las opciones incorrectas). Para la selección de reactivos puede emplearse una aproximación racional o una empírica. El enfoque *racional* demanda a la persona que elabora la prueba formular juicios personales concernientes a qué distractores son apropiados. En contraste, el enfoque *empírico* consiste en seleccionar distractores de entre las respuestas incorrectas más populares a los troncos de los reactivos aplicados en afirmaciones abiertas-cerradas. No hay consenso acerca de qué método da lugar a los mejores distractores, pero el juicio del examinador parece ser al menos tan efectivo como la aproximación empírica (Hanna y Johnson, 1978; Owens, Hanna y Coppedge, 1970).

<sup>4</sup>Se acredita a Arthur Otis haber sido pionero en el uso del formato de reactivo de opción múltiple en las pruebas colectivas de inteligencia. Los primeros instrumentos publicados que emplearon este formato fueron las Pruebas autoaplicables de Otis de habilidad mental (1916-1917).

**CUADRO 2.2****LINEAMIENTOS PARA ELABORAR REACTIVOS DE OPCIÓN MÚLTIPLE**

1. Debe utilizarse como tronco una pregunta o una afirmación incompleta, pero se prefiere el formato de pregunta. Si el tronco es una afirmación incompleta, coloque el espacio en blanco al final de la afirmación.
2. Establezca claramente el problema específico de la pregunta o afirmación incompleta en el tronco y a un nivel de lectura apropiado para los examinados, pero evite tomar preguntas o afirmaciones literales de los textos.
3. Coloque la mayor parte del reactivio en el tronco. Es ineficiente repetir las mismas palabras en cada opción y a los examinados les resulta menos difícil revisar las opciones más cortas.
4. Emplee preguntas de opinión con moderación; cuando las utilice, cite la fuente de la opinión.
5. Cuatro o cinco opciones son típicas, pero también pueden escribirse buenos reactivos que tengan sólo dos o tres opciones. Con los estudiantes de los primeros grados, tres opciones son preferibles a cuatro o cinco. Haladyna y Downing (1993) concluyeron que tres opciones pueden ser adecuadas para la mayor parte de las pruebas de habilidad y rendimiento.
6. Si las opciones tienen un orden natural, como fechas o edades, es aconsejable disponerlas en ese orden. De otro modo, ordénelas aleatoria o alfabéticamente (siempre que la alfabetización no proporcione señales para la respuesta correcta).
7. Haga que todas las opciones sean aproximadamente de la misma extensión, que sean gramaticalmente correctas y apropiadas en relación con el tronco. Sin embargo, no deje que el tronco revele la opción correcta por medio de asociaciones verbales u otras señales.
8. Haga que todas las opciones sean plausibles para los examinados que no conocen la respuesta correcta, pero haga que sólo una opción sea la correcta o "la mejor". Los errores populares o las afirmaciones que sólo son parcialmente correctas son buenos distractores.
9. Al elaborar cada distractor, plantee una razón por la cual los examinados que no conocen la respuesta correcta podrían seleccionarlo.
10. Evite, o al menos minimice, el uso de expresiones negativas como "no" en el tronco o las opciones.
11. Aunque cierta cantidad de novedad e incluso de humor es apropiada y puede servir para interesar y motivar a los examinados, no deben usarse reactivos y opciones ambiguos o capciosos.
12. Use con moderación las expresiones: ninguno de los anteriores, todos los anteriores, o más de uno de los anteriores. Además, evite el uso de determinantes específicos como: siempre o nunca.
13. Coloque las opciones en un formato apilado (párrafo) en lugar de hacerlo en tandem (una tras otra); use números para designar los reactivos y letras para las opciones.
14. Prepare el número correcto de reactivos para el grado o nivel de edad que se pondrá a prueba, haciendo que cada reactivio sea independiente de otros reactivos (que no se entrelacen o se interrelacionen).
15. Haga que los niveles de dificultad sean tales que el porcentaje de examinados que responden a un reactivio de manera correcta esté aproximadamente a la mitad entre el porcentaje de azar (adivinación aleatoria) y el 100 por ciento: % correcto =  $50(k + 1)/k$ , donde k es el número de distractores por reactivio.

**Elaboración de reactivos complejos.** Los diseñadores de pruebas por lo general tienen más dificultad para elaborar reactivos que midan la comprensión y el pensamiento que los que miden el conocimiento directo de la materia. Se han propuesto varias formas de redactar reactivos objetivos que evalúen objetivos instrucionales más complejos. Opciones como: todas las anteriores, ninguna de las anteriores, dos de las anteriores y todas salvo una de las anteriores, pueden tornar más difícil la elección de un examinado. Dicha elección también puede complicarse haciendo que todas las opciones sean correctas (o incorrectas) y pidiendo a los examinados que seleccionen la mejor o la más apropiada para cada reactivio. Otras maneras de hacer más difícil la decisión de un

examinado son: (1) incluir reactivos de respuesta múltiple en los cuales números variables de opciones sean correctos y el examinado deba indicar qué opciones (si las hay) son correctas o incorrectas; (2) hacer que los examinados seleccionen una respuesta y la mejoren o escriban una breve justificación de la misma, y (3) pedir a los examinados que identifiquen el planteamiento correcto (como una ecuación o método de solución) en tareas de resolución de problemas.

En el cuadro 2.3 se ilustran otros procedimientos para incrementar la complejidad de los reactivos de opción múltiple. Todas esas técnicas están diseñadas para hacer que la selección de la opción correcta sea un proceso reflexivo y analítico, en el cual se pongan en práctica varias capacidades cognoscitivas en lugar de sólo la memoria. Por último, el uso de un formato de conjunto de problemas, en el cual dos o más reactivos de opción múltiple se relacionan con la misma ilustración, gráfica, pasaje o escenario, se ha vuelto popular en los exámenes de acreditación o certificación (Hambleton, 1996).

### **Uso de computadoras en la elaboración de pruebas**

Las aplicaciones más comunes de las computadoras en la elaboración de pruebas consisten en programas de procesamiento de textos para ayudar en la mecanografía de los reactivos, la formación, la revisión de errores de ortografía y de sintaxis, etc. La elaboración de pruebas es facilitada aún más por una combinación del procesador de textos y programas de gráficos que apoyan la preparación de pruebas compuestas por palabras e ilustraciones. Esos programas contienen bancos de reactivos a los cuales se puede tener acceso ingresando ciertas palabras clave que indican el contenido y las características psicométricas deseadas en la prueba. Los bancos de reactivos, de los que pueden seleccionarse y recuperarse los reactivos al diseñar las pruebas, están disponibles con los editores de libros de texto como complementos para determinadas obras.

Los redactores de reactivos de prueba basados en la computadora, algoritmos de especificación de dominio para generar reactivos de prueba, y enfoques basados en la lingüística o el aprendizaje de conceptos para la redacción de reactivos, pueden proporcionar procedimientos más eficientes y precisos para la elaboración de reactivos de prueba (Herman, 1994). En la actualidad la preparación de buenos reactivos de prueba es tanto un arte como una ciencia.

## **FORMACIÓN Y REPRODUCCIÓN DE UN TEST**

Una vez que se han preparado los reactivos para una prueba, es aconsejable hacer que los revisen y editen personas conocedoras. Incluso los esfuerzos más concienzudos no necesariamente producen una buena prueba, y un amigo o asociado con frecuencia puede detectar errores y hacer sugerencias valiosas para mejorar los reactivos.

Suponiendo que el diseñador de la prueba ha elaborado un número suficiente de reactivos satisfactorios, antes de formar una prueba deben tomarse decisiones finales concernientes a varios asuntos:

1. ¿Es la longitud de la prueba apropiada para los límites de tiempo?
2. ¿Cómo deberán agruparse u ordenarse los reactivos en las páginas del cuadernillo de la prueba?
3. ¿Deben marcarse las respuestas en el cuadernillo de la prueba o se utilizará una hoja especial de respuestas?
4. ¿Cómo se reproducirán el cuadernillo de la prueba y la hoja de respuestas?
5. ¿Qué información debe incluirse en las instrucciones de la prueba?

### **CUADRO 2.3**

## ALGUNAS FORMAS COMPLEJAS DE REACTIVOS DE OPCIÓN MÚLTIPLE

## Extensión de la prueba

La decisión de cuántos reactivos incluir en una prueba depende de los límites de tiempo, del grado y nivel de lectura de los examinados, y de la extensión y dificultad de los reactivos. Los reactivos cortos y/o los que sólo requieren memorización de acontecimientos pueden responderse en menos tiempo que los más largos, donde son necesarios cálculos laboriosos y/o razonamiento abstracto. La experiencia previa con reactivos del mismo tipo general que los incluidos en una prueba ayudará a determinar si los límites de tiempo son apropiados. En las pruebas de dificultad moderada aplicadas a partir del nivel de las escuelas secundarias, una buena regla empírica es conceder un minuto por cada reactivo de opción múltiple o de respuesta corta y un minuto por cada dos reactivos de verdadero y falso. De este modo, una prueba de 50 reactivos de opción múltiple o de respuesta corta y una de 100 reactivos de verdadero y falso suelen ser apropiadas para un periodo de clase típico de 50 minutos en el nivel de secundaria. Cinco o seis preguntas de ensayo que requieren respuestas de media página pueden ser respondidas en este mismo periodo. A menos que los reactivos sean muy largos o sumamente difíciles, al menos 90% de los estudiantes en un grupo típico de secundaria podrán terminar la prueba en el tiempo asignado. La extensión de la prueba y los límites de tiempo necesitarán ajustarse hacia abajo o hacia arriba cuando se examine a alumnos de escuela primaria o a estudiantes de universidad.

Existen, por supuesto, diferencias entre los estudiantes en cuanto al tiempo que requieren para terminar una prueba. Puede esperarse que aquellos con más conocimientos o habilidades en la materia de la prueba terminen primero, pero no siempre sucede así. Los estudiantes menos informados pueden simplemente adivinar o “rendirse” y entregar la prueba antes del tiempo límite cuando se permite hacerlo. Además, los hábitos de presentación de pruebas de los examinados con altas calificaciones pueden llevarlos a revisar los reactivos de la prueba en varias ocasiones para estar seguros de que no pasaron algo por alto o lo interpretaron mal. Ciertos estudiantes, con altas y bajas calificaciones, también habrán escuchado que es más probable que sus respuestas iniciales sean las correctas, y por lo tanto no es buena idea perder tiempo reconsiderando la primera elección. Todos esos factores hacen difícil predecir cuánto tiempo le tomará a un alumno determinado terminar una prueba. Todo depende de una interacción compleja entre la preparación, la personalidad y el estado emocional y físico del estudiante, de la naturaleza y dificultad del material de la prueba, y del ambiente del examen (ruido y otras distracciones, conducta del examinador o supervisor, etcétera). Es probable que quien administre la prueba pueda hacer que el tiempo real dedicado a resolverla sea más uniforme al pedir que los examinados permanezcan en sus asientos después de terminarla, pero aún así puede haber diferencias sustanciales en el tiempo que necesitan los examinados para completar la prueba.

## Ordenamiento de los reactivos

En lo que respecta al ordenamiento de las opciones en los reactivos de opción múltiple, se ha dicho que los examinados muestran preferencias por la posición de las opciones, y cuando no están seguros de la respuesta es más probable que elijan ciertas opciones (digamos *b* y *c*) que otras (*a* y *d*). Aunque la investigación no ha logrado demostrar que estas preferencias tengan un efecto significativo en las calificaciones de una prueba (Jessell y Sullins, 1975; Wilbur, 1970), es aconsejable ordenar los reactivos de opción múltiple y los de verdadero y falso de forma que las respuestas no sigan un patrón. Ordenar las opciones para los reactivos de opción múltiple en orden alfabético puede ser satisfactorio, pero una mejor estrategia es aleatorizar el orden de las opciones dentro de los reactivos. Esto asegurará que al menos la persona que elabora la prueba no tenga ninguna inclinación al ordenar las opciones correctas. Por supuesto, cuando se usan, op-

ciones como: todas las anteriores y ninguna de las anteriores, éstas deben colocarse en la última posición.

En los reactivos de aparejamiento o reordenamiento, a los examinados les resulta más conveniente y la calificación se facilita si todas las premisas y opciones de respuesta se colocan en la misma página. Colocar los reactivos de respuesta corta en grupos de cinco o algo así también puede reducir los errores al presentar y calificar una prueba. Por último, debe proporcionarse espacio suficiente para responder los reactivos de respuesta corta y los de ensayo, sea que las respuestas se escriban en el cuadernillo de la prueba o en una hoja por separado.

Con relación al esquema de la prueba como un todo, puede esperarse que la tarea de los examinados se haga más sencilla si se agrupan juntos los reactivos del mismo tipo (opción múltiple, verdadero y falso, etc.) y los que tratan del mismo tema. Es cierto que ordenar los reactivos en grupos de acuerdo con el tipo o tema puede simplificar la preparación, aplicación y calificación de la prueba, pero no hay evidencia de que esta práctica mejore las calificaciones del instrumento. En las pruebas que contienen reactivos objetivos y de ensayo, estos últimos suelen colocarse al final, ya que suelen requerir más tiempo y diferentes procesos de pensamiento que los primeros.

Otra suposición razonable es que las calificaciones de la prueba serán más altas si se ordenan subconjuntos de reactivos del más fácil al más difícil. Se supone que el éxito al responder los reactivos más sencillos crea expectativas favorables de éxito, y que ello anima a los examinados a poner más empeño en los reactivos más difíciles. Sin embargo, los hallazgos de la investigación no siempre han confirmado esta suposición (Allison, 1984; Gerow, 1980; Klimko, 1984). Un reactivo sencillo ocasional puede mejorar el desempeño en los reactivos subsecuentes, pero, en general, ordenar los reactivos en orden de dificultad parece tener poco efecto sobre las calificaciones globales. Las excepciones a esta conclusión son las pruebas de velocidad (Plake, Ansorge, Parker y Lowry, 1982) o las muy difíciles (Green, 1984; Savitz, 1985). En una prueba de velocidad o en una que es muy difícil, colocar los reactivos más difíciles al final de la prueba parece mejorar un tanto las calificaciones.

Una conclusión lógica de los hallazgos de la investigación sobre los efectos del ordenamiento en los reactivos de acuerdo con el nivel de dificultad es que, al elaborar pruebas que no son de velocidad desde fáciles hasta de dificultad moderada, los diseñadores harían bien en preocuparse menos por el ordenamiento de los reactivos e interesarse más en asegurarse de que estén bien escritos y midan lo que se supone deben medir. Cuando una prueba es muy difícil o de velocidad, colocar los reactivos en orden de los más fáciles a los más difíciles puede asegurar el uso más eficiente del tiempo del examinado, así como mejorar la motivación y, por consiguiente, dar por resultado calificaciones más altas.

## Hojas de respuestas

Para la mayoría de las pruebas que se administran en un aula, en especial en los primeros grados, es aconsejable hacer que los estudiantes marquen o escriban sus respuestas en el cuadernillo de la prueba (Airasian y Terrasi, 1994). Esto genera menos errores al indicar las respuestas. En los reactivos objetivos, también facilita la calificación si se requiere que los examinados escriban las letras o respuestas apropiadas en los espacios marginales situados a la izquierda de las preguntas.

Las hojas de respuestas por separado, que son más fáciles de calificar, pueden usarse a partir de los últimos años de la escuela elemental. Si la prueba se va a calificar con una máquina deberán usarse hojas de respuestas distribuidas comercialmente. En dichas hojas, los examinados responden colocando en un círculo o en un espacio al lado del número del reactivo el número o letra correspondiente. Si la prueba va a calificarse de manera manual, el profesor puede

preparar fácilmente una hoja de respuestas y duplicarla. Una hoja de respuestas para una prueba de 75 reactivos de opción múltiple puede tener el siguiente formato:

- |               |               |               |
|---------------|---------------|---------------|
| 1. a b c d e  | 26. a b c d e | 51. a b c d e |
| 2. a b c d e  | 27. a b c d e | 52. a b c d e |
| ...           | ...           | ...           |
| 25. a b c d e | 50. a b c d e | 75. a b c d e |

Se indica a los examinados que marquen la letra correspondiente a la respuesta correcta para cada reactivo. También se dispone de hojas de respuestas SCANTRON que pueden ser calificadas por una máquina o a mano.

Toda institución educativa tiene recursos que facilitan la reproducción de materiales escritos o impresos para su uso en el aula. Las máquinas fotocopiadoras pueden utilizarse para reproducir los cuadernillos de prueba en un formato de impresión por uno o ambos lados, en ocasiones a color. Si se va a usar el mismo tipo de hoja de respuestas<sup>5</sup> para diferentes pruebas, puede imprimirse una gran cantidad en una sola operación de la máquina y almacenarse para otras aplicaciones de pruebas.

### Instrucciones en los tests

Las instrucciones generales para una prueba de ensayo u objetiva que se aplica de manera simultánea a un grupo de personas se colocan al frente de la prueba, y las instrucciones específicas para cada parte de una prueba múltiple se colocan antes de la parte respectiva. Por lo general, resulta sensato mecanografiar las instrucciones en negritas de forma que sea menos probable que los examinados las salten o las pasen por alto. Como su planteamiento puede llegar a tener cierto efecto sobre las calificaciones obtenidas, las instrucciones deben ser precisas más que generales (Joncas y Standig, 1998). También es aconsejable que el examinador lea en voz alta las instrucciones globales si son inusuales o poco familiares para los examinados. En una prueba individual en la cual el examinador presenta cada tarea e interactúa de manera continua con el examinado, las instrucciones se dan en forma oral. Sea que se den de manera oral, impresa o en ambas formas, las instrucciones deben informar a los examinados acerca del propósito de la prueba (o reactivo), cómo deben indicarse las respuestas,<sup>5</sup> qué tipo de ayuda pueden esperar si no entienden algo, cuánto tiempo tienen para terminar la prueba, cómo se calificarán las respuestas, si es recomendable adivinar cuando se tenga duda, y cómo corregir una respuesta si cometieron un error. Las siguientes instrucciones generales para una prueba de rendimiento aplicada a un grupo son representativas:

Escriba su nombre en la esquina superior derecha de la hoja de respuestas, pero no escriba en el cuadernillo de la prueba. Esta prueba está diseñada para evaluar su conocimiento y comprensión de estadística elemental. Son 50 reactivos y usted tendrá exactamente 50 minutos para completar la prueba. Indique su respuesta a cada reactivo llenando el espacio apropiado en la hoja de respuestas debajo de la letra que corresponde a la respuesta correcta. Su calificación en la prueba será igual al número de reactivos que haya respondido correctamente. Aunque la adivinación al azar no aumentará su calificación, si puede eliminar al menos una opción en un reactivo, es sensato hacer una con-

<sup>5</sup>Como el método de respuesta en las pruebas aplicadas por computadora puede no ser familiar para algunos examinados, debe asignarse tiempo suficiente para dar las instrucciones y mostrar cómo funciona el equipo. Además, los examinados deben ser supervisados durante la prueba para asegurarse de que están usando el equipo de manera apropiada.

jetura informada a partir de las opciones restantes. Debe tener tiempo suficiente para responder todos los reactivos y revisar sus respuestas. Si termina antes de tiempo, por favor permanezca sentado en silencio hasta que todos hayan terminado.

Cuando las instrucciones de una prueba se den de manera oral, deben leerse de forma lenta, clara y exactamente como aparecen impresas. Después de haber leído las instrucciones, debe permitirse a los examinados hacer preguntas, e independientemente de su trivialidad o redundancia aparente, deben responderse de manera paciente e informativa.

En las pruebas múltiples que constan de cierta variedad de temas y/o tipos de reactivos, puede ser necesario dar instrucciones específicas para cada parte. Las instrucciones que atañen a muchos de los mismos asuntos (cómo marcar las respuestas, cómo corregir los errores, si se pueden omitir respuestas o adivinar cuando se tenga duda) pueden variar con el tipo de reactivos objetivos. Las instrucciones para responder los reactivos de ensayo pueden incluir sugerencias acerca de cómo estructurar las respuestas (bosquejo, formato y cosas similares); cómo deben ser las respuestas largas; qué tanto peso de calificación se dará al contenido, forma, gramática, caligrafía y otros rasgos de las respuestas, y si debe intentarse responder a todas las preguntas, a un número selecto de éstas, o si algunas son obligatorias y otras opcionales.

## PRUEBAS ORALES

Las pruebas orales se definen como una situación de evaluación en la cual los examinados responden de manera oral a las preguntas planteadas. Las preguntas pueden presentarse de manera oral, por escrito o de ambas formas. Las pruebas orales de rendimiento son más comunes en las instituciones educativas europeas que en Estados Unidos, donde la práctica de las pruebas orales declinó durante el siglo XX y es menos común en los grados superiores que en los inferiores.

A muchos estudiantes no les gustan las pruebas orales y sienten que son medidas injustas del conocimiento y la comprensión. Sin embargo, los maestros de expresión oral, arte dramático, inglés e idiomas extranjeros, a menudo deploran la falta de atención a la evaluación de las habilidades del lenguaje hablado y sienten que la consecuencia de semejante descuido es una ciudadanía que no puede hablar de manera correcta, comprensible y cómoda. Aunque muchos maestros de idiomas y de otras materias en las cuales es importante el desarrollo de las habilidades del habla admiten lo deseable de los ejercicios y evaluaciones orales, también se dan cuenta de que las pruebas orales no sólo son muy subjetivas sino que a menudo resultan ineficientes (Crowl y McGinitie, 1974; Platt, 1961).

### Ventajas de las pruebas orales

Desde los primeros años del siglo XX, las pruebas orales de rendimiento se han venido percibiendo como carentes de eficiencia y rigor psicométrico. También se les ha criticado por requerir demasiado tiempo, proporcionar una muestra limitada de respuestas y por estar mal planeadas en la mayoría de los casos.

Sin embargo, a pesar de sus limitaciones, incluso los críticos de las pruebas orales admiten que éstas poseen algunas ventajas sobre las pruebas escritas. Una ventaja es la situación social interactiva que proporcionan, lo que permite evaluar cualidades personales como apariencia, estilo y manera de hablar. La situación cara a cara también hace poco probable la copia y quizás los engaños. Otras ventajas de las pruebas orales es que con frecuencia requieren respuestas a un

nivel intelectual más alto que las escritas, y proporcionan práctica en comunicación oral e interacción social. También alientan una revisión más cuidadosa del material de prueba y pueden ser terminadas en menos tiempo que exámenes escritos comparables. Los individuos que aplican pruebas orales pueden seguir los procesos de pensamiento de los examinados y localizar con más facilidad los límites de su conocimiento y comprensión de la materia. Esos límites pueden ser determinados pidiendo a los examinados que expliquen, defiendan o se esmeren en sus respuestas. Por último, el tiempo que se necesita para preparar y evaluar las respuestas orales puede ser menor que para una prueba escrita comparable (Glovrozov, 1974; Platt, 1961).

Las pruebas orales son especialmente apropiadas para los alumnos de primaria y para otros que experimentan dificultades en la lectura o escritura. Incluso en los niveles superiores puede estar justificada la aplicación de una prueba oral ocasional cuando el tiempo y/o los recursos para reproducir los materiales de prueba son escasos (Green, 1975). Los exámenes orales son cruciales en materias como expresión oral, idiomas y arte dramático.

Las entrevistas estructuradas que constan de preguntas y respuestas orales a menudo se realizan con solicitantes de puestos en organizaciones gubernamentales e industriales. Es frecuente que tales entrevistas se efectúen por teléfono cuando los solicitantes no pueden viajar al sitio del examen. En exámenes de este tipo es posible introducir cierta cantidad de estandarización y control planteando a todos los examinados las mismas preguntas, limitando el tiempo del que disponen para responder y registrando electrónicamente sus respuestas para reproducirlas y evaluarlas más tarde.

### Pruebas orales contra pruebas escritas

El hecho de que las calificaciones en las pruebas orales de rendimiento sólo tengan correlaciones moderadas con las calificaciones en pruebas escritas comparables, sugiere que miden aspectos diferentes del rendimiento. En general, el conocimiento de hechos específicos puede ser determinado con mayor rapidez por las pruebas objetivas escritas, por lo que los exámenes orales no deben contener grandes cantidades de esos tipos de preguntas. Como sucede con las pruebas de ensayo, las pruebas orales son más apropiadas cuando las preguntas requieren de respuestas extensas.

Dado que los logros o las conductas evaluadas mediante pruebas orales son tan importantes como las mediciones de pruebas escritas, debería prestarse más atención a la principal fuente de error en las pruebas orales: los examinadores o evaluadores. Las personas que aplican pruebas orales deben poseer un conocimiento profundo de la materia y una conciencia muy aguda de las respuestas apropiadas. Además, las categorías usadas por los examinadores al describir o calificar las respuestas de los examinados deberían citar conductas observables específicas en lugar de conceptos vagos como *potencial creativo, carácter, habilidad general o efectividad interpersonal*. Estos conceptos indefinidos, y que quizás no puedan definirse, no son medidos más fácilmente por las pruebas orales que por las escritas.

## PRUEBAS DE DESEMPEÑO

Las pruebas de lápiz y papel son las más eficientes y objetivas de todos los tipos de pruebas, pero por lo regular sólo proporcionan información indirecta acerca de la habilidad de una persona para hacer o fabricar algo. El conocimiento de la materia puede demostrarse de manera bastante minuciosa en un periodo corto por medio de una prueba de ensayo, una de opción múltiple u otra prueba escrita. Sin embargo, poseer un bagaje de información acerca de un tema o ser capaz de explicar cómo hacer algo no es lo mismo que usar la información o destreza en situacio-

nes prácticas. En alguna ocasión el autor condujo un taller de relaciones humanas con un grupo de supervisores de una línea de ensamblaje. Aunque todos los supervisores salieron bien en las pruebas escritas del material presentado en el taller y coincidieron en que un enfoque democrático hacia la supervisión era superior a uno autoritario, la mayoría reanudó su conducta autoritaria en la supervisión al regresar a la línea de ensamblaje.

Hay muchos otros ejemplos de conductas específicas a la situación, en las cuales los estudiantes aprenden a dar la respuesta correcta en clase o en una prueba de lápiz y papel, pero la abandonan cuando enfrentan una situación de la vida real en la que podría ser aplicable. Buena parte del aprendizaje que se da en el salón de clase se relaciona con conductas en contextos no académicos, pero la relación está lejos de ser perfecta. La generalización del conocimiento y las habilidades del salón de clases a las situaciones de la vida real es particularmente endeble en el caso del conocimiento verbal. Los maestros se dan cuenta de que si la escuela debe preparar a los estudiantes para la vida, las habilidades y el conocimiento deben enseñarse de tal manera que se transfieran a situaciones laborales y otros contextos no académicos. Los maestros de ciencia, atletismo, arte dramático, música, artes industriales, expresión oral, lenguas extranjeras, caligrafía, agricultura, y muchas otras áreas temáticas, reconocen la necesidad de que los estudiantes practiquen repetidamente y tengan experiencia directa para que las habilidades sean bien aprendidas y transferibles. Los laboratorios y proyectos de ciencia, las habilidades psicomotrices aprendidas en juegos y deportes, tocar instrumentos musicales y cantar, actuar en obras, construir o aplicar objetos útiles en un taller, practicar el hablar en público y la conversación en español y en otros idiomas, todo lo anterior, proporciona oportunidades para aprender y practicar habilidades que son potencialmente útiles fuera de la clase y servirán como cimientos para el aprendizaje práctico experiencial posterior. Debido a un mayor realismo que las pruebas escritas, a las pruebas de ejecución en ocasiones se les conoce como *evaluación auténtica* o, para enfatizar que son una opción a las pruebas escritas, *evaluación alternativa*.

Aunque puede no ser necesario seguir una taxonomía de objetivos psicomotrices al planificar una prueba para medir qué tan bien ha aprendido una persona una habilidad particular, es útil elaborar una lista detallada de las conductas que son indicadoras de un rango de competencia en esa habilidad. Deben tomarse de antemano decisiones como qué tanto peso (numérico) se dará a cada aspecto del desempeño y qué deducciones (si las hay) se harán por errores, lentitud o descuido.

Una prueba de ejecución debe concentrarse, sobre todo, en el producto o resultado final de ejecutar una habilidad, pero también es importante observar la forma en que se realiza (el proceso). Por ejemplo, lo que cuenta más al jugar golf es el número de golpes requeridos para meter la bola en el hoyo, pero todos los instructores de golf se dan cuenta de que la forma, o estilo, también es importante. En las pruebas de ejecución que involucran un producto terminado tangible no sólo debe advertirse la cantidad y calidad del producto, sino también la eficiencia con la que fue hecho.

Tanto los productos como los procesos del desempeño suelen evaluarse de manera subjetiva, principalmente por observación combinada con un registro escrito o electrónico y una lista de verificación o escala de calificación. Es posible examinar y evaluar *portafolios* enteros, o colecciones de los desempeños y productos de los estudiantes a lo largo de un periodo. Para la evaluación precisa del desempeño es crucial una observación cuidadosa que esté tan libre de sesgos como sea posible. Las pruebas de ejecución estructuradas, en las cuales se prueba a cada examinado bajo las mismas condiciones, suelen ser más objetivas que las no estructuradas, donde se observa y evalúa a los estudiantes de manera subrepticia durante la clase, en los pasillos o en otras áreas de la escuela. Pero incluso aunque se tenga sumo cuidado, por su misma naturaleza, las pruebas de ejecución son menos objetivas y, en consecuencia, menos confiables que las pruebas escritas. Además, las pruebas de ejecución requieren más tiempo que las escritas y a menu-

do también equipo costoso y otras condiciones que consumen tiempo. Por esas razones, siempre que el costo y la ineficiencia de una prueba de ejecución no sean compensados por su carácter realista, es preferible una prueba escrita.

## RESUMEN

Este capítulo trata principalmente de procedimientos para diseñar y elaborar pruebas de rendimiento educativo, pero los principios analizados también pueden aplicarse a otros tipos de instrumentos de evaluación psicológicos y educativos.

El primer paso en la elaboración de una prueba de rendimiento es preparar una lista de los objetivos conductuales que van a evaluarse. Luego debe construirse una tabla de especificaciones que presente el número de reactivos necesarios en cada categoría de contenido (temático) para cada objetivo conductual. Se han propuesto varias taxonomías o métodos de clasificación de objetivos conductuales en los dominios cognoscitivo, afectivo y psicomotriz. La taxonomía de objetivos educativos más popular es la *Taxonomía de objetivos educativos: el dominio cognoscitivo*, de Bloom y Krathwohl.

Tanto las pruebas de ensayo como las objetivas poseen ventajas y desventajas. Los reactivos de ensayo son más fáciles de elaborar, pero los reactivos objetivos pueden calificarse de manera más rápida y precisa. Las pruebas objetivas también proporcionan una muestra más representativa del contenido de la materia. Las preguntas de respuesta corta, de verdadero y falso, de opción múltiple y de aparejamiento son variedades de los reactivos objetivos. De éstos, los reactivos de opción múltiple son los más versátiles y populares.

Al formar una prueba debe prestarse atención a factores como la longitud y el formato, el método para registrar las respuestas, las facilidades para la reproducción de la prueba, y las instrucciones para la aplicación. Las instrucciones de aplicación de una prueba incluyen el(s) propósito(s), los límites de tiempo, el procedimiento de calificación y lo aconsejable de adivinar cuando se tenga duda.

Las pruebas orales no se usan tan a menudo como las pruebas escritas, pero cuando se planean, aplican y evalúan con cuidado pueden proporcionar información que por lo regular no se obtiene con otros métodos de evaluación. En cierto sentido, tanto las pruebas escritas como las orales son medidas de ejecución, pero el concepto de pruebas de ejecución por lo general se ha concentrado en conducta no verbal. Dado que las pruebas de ejecución son más realistas que las verbales, en ocasiones se les conoce como evaluación auténtica. En lugar de limitarse a describir cómo hacer algo o qué se hizo, las pruebas de ejecución requieren que los examinados demuestren un proceso. Dichas pruebas se emplean de manera extensa para evaluar habilidades aprendidas en el laboratorio y en situaciones de campo, las cuales abarcan desde el laboratorio de ciencia hasta la arena deportiva y otros contextos aplicados. A menudo los maestros conservan y evalúan portafolios del desempeño y los productos de los estudiantes.

---

## PREGUNTAS Y ACTIVIDADES

1. Elija un tema para desarrollar una prueba en un área que le interese, plantee sus objetivos conductuales y de contenido, elabore una tabla de especificaciones y diseñe una prueba objetiva de una hora sobre el tema elegido usando varios tipos de reactivos.

2. Diseñe un sistema de objetivos educativos del dominio cognoscitivo para su salón de clases. ¿En qué difiere de los sistemas que se describieron en el texto? ¿Qué ventajas y desventajas particulares posee?
3. Elabore una taxonomía de objetivos para las humanidades o el currículo básico en la universidad. Incluya al menos cinco objetivos de su taxonomía, con dos o tres subobjetivos bajo los cinco objetivos principales. Defina cada una de las principales categorías y subcategorías de su taxonomía de manera tan clara y objetiva como sea posible.
4. Diseñe una tabla de especificaciones para una prueba global de humanidades que vaya a aplicarse a todos los estudiantes al final de su segundo año en la universidad. Base los objetivos conductuales y de contenido de su tabla de especificaciones en la taxonomía que elaboró en la actividad 3.
5. Describa las fortalezas y debilidades relativas de las pruebas de ensayo, orales y de ejecución. ¿Para qué propósitos y bajo qué condiciones es más apropiado cada tipo de prueba?
6. ¿Por qué suele considerarse que los reactivos de opción múltiple son superiores a los otros tipos de reactivos objetivos? ¿Puede pensar en una situación donde los reactivos de verdadero y falso, completamiento o aparejamiento sean preferibles a los de opción múltiple?
7. Escriba cinco reactivos de respuesta corta (completamiento), cinco de verdadero y falso, y cinco de opción múltiple basándose en la siguiente selección adaptada de Aiken (1980):

Una razón para la escasez de datos psicométricos sobre los adultos mayores es que en este grupo de edad la gente, cuya conducta es menos susceptible de ser controlada por psicólogos y educadores, a menudo se muestra renuente a ser examinada. Hay muchas razones para explicar la poca cooperación de los adultos mayores en las situaciones de prueba, incluyendo la falta de tiempo, la percepción de las tareas de la prueba como triviales y sin sentido, y el temor de salir mal y parecer tontos. A los adultos mayores, en mayor medida que los adultos más jóvenes que están más conscientes de la prueba, no les entusiasma realizar tareas que los hagan ver ridículos o que son percibidas como irrelevantes en sus vidas.

Debido a que los adultos mayores tienen poca motivación para ser examinados, se requiere sensibilidad y tacto de parte de los examinadores psicológicos para obtener respuestas válidas. Por desgracia, a menudo se cuestiona si los examinadores técnicamente competentes pero jóvenes pueden establecer suficiente *rappor* con los examinados mayores como para comunicarles adecuadamente las instrucciones de la prueba y estimularlos para hacer lo mejor que puedan. Relativamente pocos psicómetras parecen tener el entrenamiento y la experiencia suficientes en la examinación psicológica de los adultos mayores como para hacer un trabajo creíble. Sin embargo, la mayoría de los examinadores encuentra que una vez que las personas mayores aceptan ser probadas, tienen una motivación tan alta como la de los examinados jóvenes para hacer las cosas bien.

Incluso cuando los adultos mayores se muestran cooperativos y motivados, los límites de tiempo de muchas pruebas, la presencia de defectos sensoriales, la tendencia a la distracción y la facilidad con que se fatigan les dificulta desempeñarse de manera satisfactoria. Una de las cosas más características acerca de ser mayor es que los reflejos y los movimientos físicos tienden a ser más lentos. Por esta razón, las explicaciones de la declinación relacionada con la edad en las calificaciones de pruebas en áreas como el aprendizaje y la memoria deben considerar el hecho de que los adultos mayores por lo general no reaccionan con tanta rapidez como los adultos jóvenes.

Aunque la gente mayor suele estar en desventaja en las pruebas cronometradas, su desempeño mejora de modo significativo cuando se le da tiempo suficiente para responder. En las pruebas que no están cronometradas los adultos mayores muestran poca o ninguna inferioridad en comparación con los adultos más jóvenes.

Los defectos sensoriales, en especial en las modalidades visual y auditiva, también pueden interferir con el desempeño en la vejez. Puede ser útil contar con materiales especiales de prueba, como caracteres grandes, y examinadores entrenados que estén alerta en cuanto a la presencia de defectos sensoriales. Sin embargo, en ocasiones un supuesto defecto sensorial en realidad puede ser una máscara para ocultar un problema de lectura o comprensión auditiva. El autor tuvo la experiencia de prepararse para probar a un hombre anciano que, avergonzado por su poca habilidad para la lectura, convenientemente olvidó sus lentes y, en consecuencia, no pudo leer los materiales de la prueba.

8. ¿Cuáles son las ventajas y desventajas de las pruebas orales en comparación con las pruebas escritas? ¿En qué circunstancias son apropiadas las pruebas orales? ¿Cómo deberían diseñarse, aplicarse y calificarse?
9. ¿Qué miden las pruebas de ejecución que no pueda ser medido por medio de pruebas de lápiz y papel (escritas) o por pruebas orales? Describa dos o tres pruebas de ejecución que haya presentado.