

CONSTRUCCION Y ANALISIS DE PRUEBAS DE RENDIMIENTO

Dr. Eduardo ABALDE PAZ

Profesor titular de Universidad. Facultad de Humanidades La Coruña

INTRODUCCION

No hay ninguna duda, creo, en admitir, al menos a nivel teórico, que la evaluación es un eslabón más de la cadena de aprendizaje y que, por tanto, no hay razón alguna para no entenderla como formativa, admitiendo su función de retroalimentación, de motivación generadora de nueva actividad, de potencial para supervisión y mejora del propio diseño curricular.

En la práctica, la evaluación desempeña un papel sancionador; en el mejor de los casos, como señala Gimeno, pretende ser una toma de conciencia de la distancia que separa al alumno de unos objetivos pretendidos, y que en las prácticas al uso son incluso no explícitos.

Acepto la idea respecto a la necesidad de recurrir a las más variadas fuentes de información a la hora de realizar la evaluación. Me parece laudable pretender cotas cada vez más altas en la búsqueda de la objetividad, sin llegar a caer en el espejismo de lo objetivo y en el fetichismo numérico.

Son tres los aspectos que están condicionando el cambio en el profesor al realizar la evaluación:

a) La falta de preparación de tipo técnico del profesorado, por carecer de información respecto a temas tales como:

1. Definición operacional del rendimiento escolar.
2. Condiciones de la validez de las pruebas o exámenes que aplica.
3. Condiciones de fiabilidad aplicada a la medición de aspectos de la conducta humana.
4. La complejidad de los fenómenos que abarca la más insignificante manifestación de una conducta de rendimiento académico.

b) La estructura objetiva del sistema de autoridad.

c) La resistencia al cambio en los procedimientos de evaluación.

Asumiendo que la utilización de pruebas de evaluación del rendimiento no es, en modo alguno, la única posibilidad de obtener información para realizar la eva-

luación; eso sí, creemos que es una forma muy extendida de realizar la evaluación del rendimiento y que supone, en principio, un deseo de objetivar el proceso de medición que necesita de planteamientos más rigurosos a la hora de interpretar los resultados.

PRINCIPIOS PARA LA CONSTRUCCION DE INSTRUMENTOS DE MEDIDA

1. *Los contenidos del aprendizaje.* Conocer la estructura de los contenidos nos ayudará a planificar un instrumento que dé la relevancia conveniente a cada núcleo del programa así como la importancia relativa en el contexto global del mismo. Es por ello conveniente desmenuzar el campo que va a ser sometido a medición de forma que todo lo importante pueda ser incluido en el instrumento, lo que lleva aparejado que puede prescindirse de lo anecdótico, incidental, o poco relevante que, en caso contrario, podría incluirse con función de relleno; igualmente el conocimiento de tal estructura llevará a un equilibrio entre las diversas partes a la vez que a una coherencia que dé unidad a la prueba.

Se hace necesario realizar un cierto *desmenuzamiento* de los temas, lo que nos dará una visión global del mismo que nos permitirá:

- No olvidar de incluir ningún aspecto importante en el instrumento.
- Incluir más elementos de aquellos apartados más amplios.
- Ponderar cada apartado en función de su importancia intrínseca y relativa.

2. *Los objetivos.* No basta con conocer el campo sobre el que se va a actuar. Será necesario, además, determinar de antemano cuáles son los objetivos que se pretenden alcanzar con aquellos contenidos tanto porque es bien claro y palpable que con un mismo contenido pueden perseguirse objetivos diferenciados, y viceversa, cuanto porque el instrumento de medida debe reflejar la situación que pretende medir.

Los objetivos a incluir en un instrumento de medida pueden ser muy variados y, lógicamente, estarán en función, fundamentalmente, de tres variables:

- a) El sujeto al que se aplicará el instrumento.
- b) Las exigencias del contenido que se ha de medir.
- c) Las pretensiones de la investigación.

3. *Tabla de especificaciones.* Existe una íntima relación entre los contenidos y los objetivos. Esa relación conviene ponerla de manifiesto en una tabla de doble entrada, en que se ponen en relación ambos aspectos. Lo que nos permite conocer qué objetivos deben cubrir los distintos contenidos y, al contrario, con qué contenidos se piensa lograr los objetivos propuestos.

PRUEBAS OBJETIVAS

De los dos grandes tipos de pruebas existentes, objetivas y de ensayo, nos limitaremos a exponer las primeras.

1.^a Pedir brevemente y con claridad que se realice un trabajo (contestar a una pregunta, hacer un dibujo, interpretar un gráfico...).

2.^a Prever las realizaciones o respuestas correctas.

3.^a Establecer de antemano la puntuación que debe atribuirse a cada respuesta o realización pedida.

4.^a Hacer posible la elaboración de datos o resultados de la prueba.

TIPOS DE ITEMS DE PRUEBAS OBJETIVAS

Se suelen dividir en dos amplias categorías, según que el sujeto dé la respuesta o la reconozca.

A. *Evocación de un recuerdo.*

1. Respuestas simples. P.e.: El Quijote fue escrito por.....

2. Textos mutilados. P.e.: Indicativo, subjuntivo son.....del verbo.

B. *Reconocimiento de recuerdos.*

1. Respuestas sugeridas:

1.1. Elección única. P.e.: La palabra que designa acción es.....

a) Pronombre.

b) Artículo.

c) Verbo.

d) Interrogación.

1.2. Elección múltiple. P.e.: Subraya los nombres de obras de la literatura española: Romeo y Julieta, La Dorotea, Los Nibelungos, La Vida es sueño, Las Moradas, La Iliada.

2. Apreciar la mejor respuesta. P.e.: La circunferencia es.....

a) Una superficie redonda.

b) Una línea curva.

c) Un círculo.

d) Una línea curva cerrada y convexa.

3. Distinguir lo verdadero de lo falso. P.e.: Escribir V o F según la frase sea verdadera o falsa.

El adverbio modifica al nombre.

El adverbio es invariable.

4. Asociar conocimientos. P.e.: Coloque, frente a las oraciones de la derecha, la letra que corresponda según la clasificación de la lista de la izquierda.

a) Aseverativas.....¡Qué triste será no verla más!

b) Interrogativas.....Los ovnis aparecieron otra vez.

c) Admirativas.....¿Qué día es mañana?

5. Algunos autores añaden los elementos de ordenación: contexto verbal, histórico, espacial...

ELECCION DEL TIPO DE ITEMS

Determinada la extensión de la materia y con una idea del número aproximado de preguntas, se establecerá el tipo de items que se ha de utilizar. Dentro de lo posible conviene emplear varios tipos. Se da más variedad a la prueba y se evita la rutina en el modo de reaccionar los alumnos.

Las cuestiones de verdadero-falso aparecen como las más fáciles de proponer; conviene tener presente que hace falta un número elevado de ítems, porque este tipo de preguntas es el más influido por el azar.

NORMAS A OBSERVAR EN LA FORMULACION DE ÍTEMS OBJETIVOS

1. Cada ítem debe medir un solo resultado del aprendizaje.
2. El contenido del ítem debe ser siempre relevante y estar de acuerdo con los objetivos operativos que se persiguen con la prueba.
3. Las preguntas deben estar formuladas con claridad y ser muy concretas. Sólo debe haber una respuesta correcta. El pie del ítem debe tener un sentido completo, independientemente de las alternativas.
4. En la pregunta no debe haber ni elementos irrelevantes ni pistas para la respuesta. Ni dar pistas para responder otras preguntas.
5. Se deben evitar los ítems negativos, pues dan lugar a confusión.
6. Todas las opciones deben ser gramaticalmente consistentes con la pregunta.
7. Todos los distractores (opciones incorrectas), deben tener las mismas posibilidades.
8. Evitar el uso de la opción «todas las anteriores» pues facilita la respuesta. Utilizar con precaución la expresión «ninguna de las anteriores».
9. La posición de la alternativa correcta se debe distribuir aleatoriamente.
10. Se debe evitar que la respuesta correcta sea más larga o esté mejor redactada que las falsas.
11. Las normas para cada tipo de ítems deben expresarse claramente y por escrito.
12. Los ítems deben ir agrupados según los diferentes modelos (V-F, opción múltiple, etc.) y dentro de cada tipo deben agruparse los de contenido homogéneo.
13. La ordenación de los ítems en una prueba debe ir de más fáciles a más difíciles.
14. Se debe evitar siempre que una pregunta sea la clave para otras respuestas.

CALIFICACION DE LOS RESULTADOS

Para la calificación de las pruebas objetivas, se da un punto por cada pregunta bien contestada; cuando se trate de apreciar la mejor respuesta se concederán dos puntos a la mejor, y uno si elige una buena, pero no la mejor.

La calificación total de la prueba se hace sumando los puntos obtenidos, teniendo en cuenta las dos observaciones siguientes:

1.^a En el caso de distinguir lo verdadero de lo falso, el resultado viene dado por la siguiente fórmula:

$$R = A - E$$

R = Puntuación final

A = Aciertos

E = Errores.

Las preguntas sin contestar no se tienen en cuenta.

2.^a Cuando se trate de elección entre varias respuestas, el cómputo de puntos obedece a la siguiente fórmula:

$$R = A \frac{E}{N - 1}$$

R = Puntuación final

A = Aciertos

E = Errores

N = N.º de respuestas posibles en cada caso.

Se utilizan las dos fórmulas para evitar la influencia del azar. En el caso de V - F, es ley general que al azar lo mismo influye en sentido negativo que en positivo.

En el caso de la elección de varias respuestas, el azar no influye tanto, porque se encuentra el sujeto ante varias posibilidades de elegir respuesta. Por eso se utiliza N - 1, justamente las posibilidades de errar.

ANÁLISIS DE LAS PRUEBAS Y ELEMENTOS

Según Lafourcade, «El análisis de los items que integran una prueba de rendimiento preparada por el maestro o profesor representa una tarea que indefectiblemente debe ser emprendida si se desea conocer la real efectividad de la enseñanza y de los instrumentos utilizados en la evaluación.

I. Dificultad de los elementos.

I. D. = a/n .

a: es el número de sujetos que han acertado el elemento que se considera.

n: el número de sujetos que han intentado resolver ese elemento; esto es, los que acertaron, los que lo fallaron y los que lo omitieron.

Los valores de la fórmula varían entre 0 y 1, indicando la proporción de sujetos que resolvió correctamente el elemento. Si multiplicamos por 100 tenemos el %.

Cuanto mayor sea el porcentaje menor será su dificultad.

I. D = 1. Todos los sujetos resolvieron satisfactoriamente el elemento.

I. D = 0. Ningún sujeto acertó el elemento.

Si el fin de la prueba es poner de manifiesto las diferencias individuales. Cada uno de los elementos que la componen ha de contribuir a esa discriminación.

Los elementos realizan el mayor número de discriminaciones entre pares de individuos cuando su nivel de dificultad es tal que son resueltos satisfactoriamente por el 50% de los individuos.

El estudio de la dificultad de los elementos se basa en la propiedad matemática de «máximos y mínimos» que dice: «El producto de dos números cuya suma es constante es máximo cuando estos números son iguales y mínimo cuando uno de ellos es la unidad». P.e.: $50 * 50 = 2.500$ pares de discriminaciones, $99 * 1 = 99$, $10 * 90 = 900$.

Cuando la prueba tiene esta finalidad, es conveniente que la mayor parte de los temas tengan una dificultad media.

La prueba puede tener una doble intención:

a) «Objetivos mínimos de clase». En este caso caben los criterios anteriores.

b) «Objetivos mínimos de alumnos». Para analizar la consecución de estos objetivos serían aconsejables criterios didácticos. Lo que supone elegir items con índices de dificultad más bajos (en torno al 80 %, es decir, que puedan superar al menos el 80 % de los alumnos), a sabiendas que se sacrifican aspectos tales como la fiabilidad de la prueba y la validez.

Utilidad del índice de dificultad.

Es necesario para poder ordenar la prueba, como su nombre indica, por la dificultad. Al construir una prueba suele colocarse al principio un pequeño porcentaje de items con poca dificultad, a fin de que los sujetos más torpes no se frustren. Después se aumenta progresivamente la dificultad.

HOMOGENEIDAD DE LOS ELEMENTOS

Todos los items que forman la prueba deben contribuir a medir lo mismo que el total de la prueba. La correlación que existe entre cada elemento y el total se denomina «Índice de Homogeneidad».

Cálculo del Índice de Homogeneidad.

1.º Correlación biserial entre cada elemento y el total.

$$r_b = I. H. = \frac{\bar{X}_p - \bar{X}_t}{S_t} * p/y$$

2.ª Utilización de los 27 % superiores e inferiores.

Pasos: a) Se hacen dos grupos, uno con el 27 % superior y otro con el 27 % inferior del total de los aciertos.

b) Se hallan las proporciones que en cada grupo aciertan el elemento (Ps y Pi).

c) Consultando el ábaco para estimar la correlación biserial entre un elemento y la puntuación total.

Ejemplo: Aplicado un test a un grupo de 300 sujetos. De los 300 tomamos el 27% con mayor número de aciertos (81) y el 27% con menor número de aciertos (81).

Hallamos las proporciones que en cada grupo acierta el elemento:

—De los 81 superiores 74 han acertado el elemento n.º 10, es decir, $P_s = . 92$.

—De los 81 inferiores 58 han acertado el mismo elemento, es decir, $P_i = . 72$.

Consultamos el ábaco correspondiente y obtenemos $I. H. = . 32$ (aprox.).

VALIDEZ DE LOS ELEMENTOS

La validez se obtiene mediante una cuidadosa planificación y delimitación del universo conductual que va a ser objeto de medición en la prueba. Exige un estudio de los programas, libros de texto e información que provenga de expertos. Delimitación de contenidos y objetivos y, elaboración de la tabla de especificaciones.

ANÁLISIS DE LAS RESPUESTAS INCORRECTAS

En las pruebas de elección múltiple interesa conocer, además de la dificultad, homogeneidad y validez de los elementos, la eficacia de las respuestas incorrectas.

Cuando se incluyen respuestas incorrectas hay que procurar elegir las de tal

forma que presenten el mismo atractivo que la correcta a los sujetos con menor aptitud o conocimientos, en la materia. Si esto se consigue los sujetos más aptos tenderán a elegir la respuesta correcta, mientras que los de menor aptitud se distribuirán homogéneamente entre todas las alternativas del elemento.

BIBLIOGRAFIA

- CABRERA, F. y ESPIN, J. V. (1986): *Medición y evaluación educativa*. Barcelona, PPU.
- FERNANDEZ PEREZ, M. (1974): *Evaluación escolar y cambio educativo*. Madrid, Cincel.
- GIMENO, J. (1981): *Teoría de la enseñanza y desarrollo del currículo*. Madrid, Anaya
- LAFOURCADE, P. (1977): *Evaluación de los aprendizajes*. Madrid, Cincel.
- SAWIN, E. I. (1970): *Técnicas básicas de evaluación*. Madrid, Magisterio Español.