

Cómo citar la publicación:

García Aretio, L. (14/05/2020). Las tablas de especificación y las rúbricas. *Contextos universitarios mediados*. (ISSN: 2340-552X). Recuperado de <https://aretio.hypotheses.org/4115>.

Las tablas de especificación y las rúbricas

Lorenzo García Aretio
UNED

Las tablas de especificación

Para no perderse, sobre todo cuando se trate de redactar pruebas de carácter objetivo con multitud de preguntas, y no sesgar el resultado de la evaluación, sugerimos elaborar un esquema (**tabla de especificación**) mediante el que determinemos el número total de *items* de la prueba y los que exactamente van a corresponder a cada bloque de competencias/destrezas y objetivos y a los contenidos (García Aretio, 1994).

Consideremos, por ejemplo, que hemos optado por redactar una **prueba objetiva** con un total de **50 items** referida a las tres primeras unidades del curso o a los tres bloques temáticos del curso completo. Por supuesto, esta tabla se puede referir también a otro tipo de contenidos, por ejemplo:

- las partes de una sola Unidad,
- la primera mitad del curso,
- el curso entero,
- contenidos referidos a:
 - hechos, ideas, conceptos, principios,
 - procedimientos técnicos,
 - normas de acción y valores y actitudes,
 - epígrafes amplios de los contenidos de una unidad...

En este ejemplo vamos a contemplar sólo estas categorías de la [taxonomía de Bloom](#): *recordar*, *comprender* y *aplicar*, distribuyendo de esta manera la carga del número *items* en cada categoría. Esta distribución o ponderación, dependerá del tipo de saberes que corresponde evaluar de las citadas tres unidades. Pretendemos que los 50 *items*, se distribuyan en función de la importancia y énfasis que hayamos propuesto en el diseño de la asignatura o curso con respecto a categorías cognitivas de los objetivos y destrezas. En este ejemplo, proponemos este grado o proporción de importancia con respecto al total de la prueba:

- *Recordar*, reconocer, identificar, definir...: 10%. Corresponderían 5 elementos.
- *Comprender*, asociar, interpretar, argumentar, ejemplificar...: 50%. Corresponden 25 elementos.
- *Aplicar*, resolver, ejecutar, modificar...: 40%. Corresponden 20 elementos.

Según la relevancia de los contenidos concretos que nos proponemos evaluar, en este caso de cada Unidad Didáctica o Bloque temático, hemos decidido que de la número 1 extraemos un 30% de *items*, o lo que es lo mismo, 15 preguntas; de la nº 2, un 20% (10 preguntas), y la nº 3 la entendemos como la que contiene contenidos más relevantes y le vamos a asignar un 50% de *items*, es decir 25 preguntas..

Tomada esta primera decisión de asignación de importancia relativa a objetivos y contenidos, nos corresponde aplicar los porcentajes correspondientes de los pertinentes objetivos que han de seleccionarse de cada Unidad. Así, si de objetivos de «*aplicar*» han de proponerse un total de 20 *items* y de la primera Unidad corresponden un 30%, de esta Unidad habrán de formularse 6 *items* (30% de 20). Y así sucesivamente. Obviamente al obtener porcentajes, los inevitables decimales de los números de *items* de cada Unidad u objetivo habrán de redondearse.

En consecuencia el número de preguntas, elementos o *items* de cada parte, unidad o bloque, deberá ser siempre proporcional a la relevancia y extensión de los contenidos a evaluar y estará también en función de la importancia asignada a las diferentes competencias y objetivos. A propósito, esta tabla que he preparado es sencilla, con pocas columnas y escasas filas.

TABLA DE ESPECIFICACIÓN (*Prueba objetiva de 50 cuestiones*)

CONTENIDOS	OBJETIVOS			TOTALES
	RECORDAR... (10%)	COMPRENDER... (50%)	APLICAR... (40%)	
UNIDAD I (30%)	2	7	6	15
UNIDAD II (20%)	1	5	4	10
UNIDAD III (50%)	2	13	10	25
TOTALES	5	25	20	50

Sin duda, preparando pruebas de evaluación que contemplen una ponderación de la importancia, del peso, de competencias/objetivos y contenidos de una determinada parte del curso, existe mayor garantía de que lo que se evalúa representa a aquella parte de los contenidos, procesos mentales o habilidades y destrezas que se pretendían valorar.

Obviamente, la extensión de la prueba, el número de elementos o preguntas que la conforman, dependerá del *carácter* de la misma (a distancia, presencial, formativa, para acreditar saberes, etc.) y del *tiempo* asignado para resolverla, cuestión fundamental, sobre todo cuando se trata de pruebas en línea. Y, obviamente, mayor consistencia tendrá una prueba si contempla mayor número de *items*.

El índice de dificultad

Conforme un profesor o equipo docente lleva más años impartiendo enseñanzas en la misma asignatura o curso, resulta muy recomendable que se vaya construyendo un base de datos o banco de *items*, clasificados por contenidos de la materia (epígrafes, unidades, bloques temáticos...) y por objetivos. Una vez conformada esa base de datos, resulta de suma utilidad asignar a cada *item* un *índice de dificultad* que se puede ir modulando conforme se aplica a diferentes grupos de estudiantes.

El índice de dificultad de un *item* (ID) muestra la relación, en una determinada prueba, entre el número de aciertos a ese *item* y el número de alumnos que realizan la prueba. Así, un ID = 0 asignado a un *item* supondría que ningún alumno respondió a ese elemento adecuadamente, por su dificultad o por la deficiente redacción del elemento. Un ID = 1, supondría suma facilidad del *item*, dado que todos respondieron correctamente. En este ejemplo, serían dos *items* que habrían de ser eliminados de la base o banco de *items* y nunca más reutilizados, dado que no suponen discriminación alguna. Los restantes items de ese bloque de contenidos, podríamos ordenarlos o clasificarlos según su índice de dificultad, *muy difíciles*, *difíciles*, *normales*, *fáciles* o *muy fáciles*. Ello facilitaría la elaboración de pruebas más equilibradas.

Las rúbricas

Muy en línea con estas tablas de especificación ([García Aretio, 1994](#)) aparecieron en los contextos educativos las ***rúbricas*** como instrumentos apropiados para el aprendizaje y la evaluación de competencias. Su mayor expansión en educación se reflejó con motivo de la implantación del [proceso de Bolonia](#) y el aprendizaje por competencias. Su plasmación gráfica es similar a la de la matriz de valoración o tabla de especificación antes descrita, tabla en fin, de doble entrada, lista de control o escala, que supone una guía de *registro*, en la que se contempla *en una de las dos dimensiones los niveles de dominio, de logro o de calidad (escala cualitativa y/o*

cuantitativa) y, asociados a ellos, en la otra dimensión los elementos, criterios o indicadores preestablecidos objeto de evaluación (relacionados con los objetivos y competencias), propios de la actividad o tarea que ha de realizar el estudiante.

Las rúbricas suponen un catálogo de indicadores altamente útiles para el estudiante, con pautas y criterios claros de valoración de las diferentes etapas o aspectos del trabajo, actividad o prueba motivo de valoración. En suma, se estructuran en criterios y niveles graduales de logro. Estas rúbricas y las tablas antes referidas, se presentan como una especie de chequeo (*check-list*).

Existen rúbricas de elaboración más sencilla, serían las denominadas *globales u holísticas*. Se trataría de una valoración global o de conjunto, sin separar las partes. Es decir, ante una actividad, tarea, prueba..., se trataría de especificar qué queremos decir (describir) cuando a un estudiante le adjudicamos un suspenso, un aprobado, un notable, un sobresaliente... O lo que es lo mismo, para nosotros docentes que diseñamos, desarrollamos y evaluamos ¿qué significamos al considerar este trabajo como *muy malo, malo, aceptable, bueno o muy bueno*?

Las *rúbricas analíticas* son bastante más complejas. Se elabora una tabla en la que cada casilla debería estar rellena de texto descriptor del tipo de desempeño o realización de la tarea o actividad, en función del indicador y del nivel de logro correspondiente. Habrá de cuidarse, por tanto:

- Seleccionar el número de niveles de logro, numéricos o cualitativos (podría ser la primera fila de la tabla).
- En los aspectos o criterios de valoración, tomar decisiones respecto a lo que se quiere, principalmente (primera columna de la tabla):
 - objetivos y procesos mentales requeridos;
 - contenidos de la unidad de estudio, de la parte del curso a que hace referencia la rúbrica, de la actividad...;
 - según procedimientos requeridos.
- En los cruces correspondientes, explicitar, describir, resultados para ubicar una prueba o tarea determinada en esa casilla, con respecto a ese criterio y nivel de logro.

Por ejemplo, en las filas podríamos ubicar los niveles de logro: notas del 0 al 5, de insuficiente a sobresaliente, muy bueno-muy malo, escaso-excelente, etc. Y en las columnas, criterios, aspectos o indicadores que pretendemos evaluar o sugerir para la mejora.

Un ejemplo. En el primer aspecto de la actividad o prueba que deseamos evaluar, se trataría de *definir* el concepto o los conceptos manejados. En cada casilla se debería insertar una descripción del criterio referida a cada nivel de logro: el estudiante ha definido mal o no ha definido ninguno (0), o alguno deficientemente

(1), o varios de forma desigual (2), o todos y de forma excelente (3). Naturalmente, los niveles podrían ser más, y también los indicadores o criterios.

En la actividad, tarea o prueba diseñada, desarrollada y, ahora dispuestos a que sea evaluada, el segundo criterio sería el de *interpretar*, por ejemplo, ¿ha asociado, ha contrastado, ha comprendido...?, ¿cómo?, de forma totalmente equivocada (0), ha interpretado una parte pequeña de lo exigido (1), asoció, contrastó o comprendió partes relevantes (2), o interpretó de forma excelente..., y así sucesivamente en los otros criterios de clasificación y de elaboración personal. Podríamos representarlo en una tabla similar a ésta, completando texto en cada una de las casillas, en la que, de izquierda a derecha juzgamos, valoramos, los criterios establecidos, concretando los niveles propios de dominio en cada casilla.

	Nivel 0	Nivel 1	Nivel 2	Nivel 3
Definir...	Ninguna/deficiente	(texto)	(texto)	Todas/perfectas
Interpretar...	Equivocación	(texto)	(texto)	Excelente
Clasificar...	Nula	(texto)	(texto)	Completa
Elaborar...	Mínima	(texto)	(texto)	Máxima

Podrían mirar [algunos ejemplos, mejores y peores, de rúbricas](#), con la idea de enriquecer la información. En todo caso, hoy existen plataformas virtuales y multitud de herramientas en la red que facilitan la elaboración de este instrumento. Cuando estas herramientas se elaboran desde la red y se complimentan también en entornos digitales, se les vienen denominando, ¡cómo no!, *e-rúbricas*.

En fin, cuando los docentes diseñan y los estudiantes conocen estas *tablas de especificación* así como las *rúbricas*, desde el inicio del curso, les aportan:

- una clarificación y *motivación* para el aprendizaje y su *autorregulación*;
- a los docentes les exige inicialmente una previsión, *diseño*, planificación meditada sobre las expectativas relacionadas con los aprendizajes deseados o pretendidos (competencias y objetivos) y, posteriormente registrar esas evidencias de forma más objetiva, reduciendo la subjetividad en la evaluación;
- a ambos, docentes y estudiantes, una *hoja de ruta* o *guía* durante todo el proceso de aprendizaje respecto a las expectativas descritas;
- una buena herramienta del estudiante para aprender a *autoevaluarse* y a *coevaluar*;
- un instrumento que ayuda a la *realimentación* de los estudiantes, ofreciendo información clara, *feedback*, sobre la superación o no de los diferentes

criterios e indicadores (*evaluación formativa* y también *sumativa*), ayuda a la autorregulación;

- una herramienta que puede representar una garantía en los procesos de *certificación* de competencias y de *acreditación*, dada su mayor objetividad.

Podrían elaborarse rúbricas referidas a una actividad, a una tarea, una unidad de estudio..., pero también pueden existir rúbricas referidas a toda una asignatura e, incluso, a una titulación.

Por supuesto, no puede ocultarse que trabajar con rúbricas requiere una mayor dedicación por parte del profesorado ¿*rentabilidad?*, tanto en el diseño, seguimiento de la tarea y evaluación de la misma, y esa circunstancia ha de tenerse muy presente. En otro sentido, personalmente pienso que, si la rúbrica no está bien diseñada y redactada, mejor emplear otros sistema de evaluación: listas de control, escalas, portafolios..., o en todo caso, aquellas *rúbricas globales* referidas con anterioridad, de más sencilla elaboración y aplicación.

En la entrada siguiente al blog, me referiré a las pruebas de evaluación de tipo abierto. Ahí encontrarán más pistas para la fijación de criterios, aspectos, indicadores de evaluación. Los niveles o grados de logro, ya es algo más sencillo de optar porque estamos más habituados a ello. Les animo a leer la próxima entrada.

Finalizo con una cautela. Al igual que señalé en alguna entrada anterior al referirme a las competencias y, sobre todo, a los objetivos, a las taxonomías, respecto al peligro de estandarizar y encorsetar en exceso los aprendizajes de nuestros estudiantes: todas estas estrategias deberán equilibrarse y hacer contrapesos con lo que debe suponer la conformación desde la universidad de *espíritus abiertos, críticos, autónomos y flexibles*. Ese es el reto. No olvidemos aquella categoría de la [taxonomía actualizada de Bloom](#): la **creación**.

Referencia: García Aretio, L. (1994). *Educación a distancia hoy*. UNED.