



La evaluación

Concepto y sentido de la evaluación

La evaluación es un proceso de valoración de conocimientos y actitudes ampliamente estudiado, pues se encuentra estrechamente ligada al aprendizaje. En esencia, se trata de averiguar en qué medida se han logrado unos objetivos formulados inicialmente, a fin de mejorar la eficacia en la enseñanza.

En el caso de las oposiciones a los cuerpos de bomberos, quienes evalúan no ejercen previamente una labor docente, solamente han de seleccionar a los más capacitados, por lo que igualmente habrán de averiguar en qué medida se cumplen unos requisitos previamente formulados.

Las características del oficio hacen que las distintas administraciones públicas tiendan a evaluar a sus candidatos en los tres ámbitos del aprendizaje: psicomotor, cognitivo y emocional.

Las pruebas psicomotoras tratan de cuantificar la potencia y resistencia, así como la coordinación y el equilibrio. Son las denominadas «pruebas físicas».

El «examen teórico», cognitivo, cuantifica el grado de conocimientos y aptitudes intelectivas, tales como la memoria, la comprensión de conceptos o la resolución de problemas.

La «prueba psicológica», ámbito emocional, intenta prever respuestas de comportamiento en el ejercicio de sus funciones.

Objetividad y subjetividad

La objetividad resulta imprescindible para conseguir un resultado satisfactorio. Evaluamos objetivamente cuando no influye en nuestro juicio ningún factor ajeno al elemento que se persigue. Como factores que contribuyen a la objetividad se pueden citar:

- El juicio atento, que supone no descuidar ni perder ningún detalle.
- La capacidad de análisis, que consiste en ir preguntándose por las causas y las consecuencias.

- El ánimo sereno o estabilidad emocional.

La subjetividad va referida a las emociones que trastornan el análisis de los jueces, pudiendo ser debida a diferentes causas, como por ejemplo:

- El impresionismo, que consiste en confundir experiencias transitorias con verdades comprobadas.
- El dogmatismo, que es erigir fórmulas o procedimientos en verdades indiscutibles al margen del estudio, la crítica y la discusión.
- El estereotipo, que es una idea preconcebida que nos hemos formado, caracterizada por una aceptación rutinaria y desprovista de toda crítica.
- El ser juez y parte interesada, que consiste en la tendencia, incluso inconsciente, a autofavorecerse.

La evaluación ha de empezar por definir los objetivos, estableciendo de manera concreta los requisitos y características que tendrá que reunir el sujeto. La realización de un tipo de prueba u otro, con un nivel de dificultad mayor o menor, habrá de cimentarse en el análisis de pruebas anteriores, pues es importante conocer en qué medida se cumplieron los objetivos predeterminados.

Planteamiento general de la prueba

Una prueba de evaluación no se puede improvisar, sino que debe ser fruto de un trabajo organizado, riguroso y responsable, máxime cuando se trata de oposiciones a los cuerpos de bomberos, dada su complejidad y trascendencia, pues no estamos hablando de un trabajo temporal, sino permanente hasta la jubilación, lo que sin duda condicionará por largo tiempo la buena marcha del servicio.

El primer paso consiste en definir las características del puesto de trabajo para, seguidamente, construir el perfil ideal de los candidatos que buscamos, lo que a su vez va a implicar el establecimiento de unos objetivos. Las herramientas que tenemos para alcanzar tales objetivos son las diferentes pruebas.

El diseño de todo el proceso de selección será por tanto una tarea muy laboriosa, que requerirá altas dosis de reflexión y debate. Las denominadas «bases de la oposición» son el documento público que describe el proceso y fija las reglas.

En cuanto a la prueba de conocimientos, el área en el que se mueve este libro, diré que se parte de un temario previamente establecido, es decir, de una relación de contenidos destinados a cubrir los objetivos de ámbito gnoseológico. Por tanto, su redacción ha de ser escrupulosa, pues solamente podrá ser objeto de examen aquello que figure en las bases. Un temario deficiente dará lugar a un examen parcial y, consecuentemente, poco fiable.



Distintas categorías de ítems

Se denomina «ítem» a cada una de las partes o unidades que componen una prueba, ya se trate de ítems de evocación (respuesta breve), de verdadero o falso, de correspondencia (asociación por pares), de selección múltiple (tipo test), interpretativos o de ensayo breve.

Los ítems, ya sean de un tipo u otro, se organizan en distintas categorías. Puesto que se trata de saber en qué medida los individuos son capaces de aplicar los conocimientos y habilidades intelectivas, los ítems se pueden clasificar según «niveles taxonómicos». El esquema de taxonomía más generalizado es el de B. Bloom quien nos categoriza los ítems de la siguiente manera:

- **Conocimiento:** se refiere a conductas bastante elementales, basadas en el recuerdo de datos, hechos o procesos.
- **Comprensión:** es un nivel elemental en el que el individuo interioriza lo que aprende y lo interpreta.
- **Aplicación:** supone la utilización del conocimiento y la comprensión de situaciones particulares más concretas.
- **Análisis:** es la descomposición de un contenido o mensaje en las partes o elementos que lo constituyen con el fin de organizarlo, jerarquizando dichos elementos y explicando las relaciones que haya entre ellos.
- **Síntesis:** es el proceso de combinar y organizar partes, elementos significativos de un todo para constituir una nueva estructura que contiene en sí los aspectos fundamentales del todo inicial pero que es, a la vez, un todo con entidad en sí mismo, distinto del primero.
- **Evaluación:** este nivel se refiere al pensamiento crítico. Son juicios cuantitativos y cualitativos sobre diferentes aspectos del conocimiento. Supone el nivel de mayor autonomía intelectual.

Como se ve, están ordenados desde la categoría o nivel taxonómico más simple al más complejo. Por ello es que los ítems más sencillos de elaborar y, en consecuencia, los más abundantes son los que única y exclusivamente miden conocimientos. Resulta un poco más complejo establecer preguntas que incidan en la comprensión. Más difícil aún es establecer problemas de aplicación, y mucho más los que vayan a requerir un análisis. Los niveles de síntesis y de evaluación no son posibles mediante ítems de selección múltiple.

Los ítems de selección múltiple

En los ítems de selección múltiple, la parte correspondiente al planteamiento del problema recibe el nombre de base o, sencillamente, pregunta. Las posibilidades que se ofrecen para la respuesta se llaman opciones o alternativas, y las alternativas incorrectas son denominadas distractores.

Existen dos tipos básicos de ítems de selección múltiple:

- Selección de la respuesta correcta, o sea, elegir entre tres o más opciones la verdadera.
- Selección de la mejor respuesta, es decir, elegir entre una serie de opciones verdaderas la mejor de ellas.

Ventajas y limitaciones:

La principal ventaja es que se puede aplicar tanto a sujetos de bajo nivel verbal, pues no tienen necesidad de escribir o redactar una respuesta, como a individuos con un alto nivel de conocimientos, a lo que hay que añadir la facilidad para su tratamiento estadístico.

Sus limitaciones radican en la dificultad para encontrar suficientes distractores que puedan parecer aceptables y, sobre todo, que se mide si el individuo sabe o comprende lo que hay que hacer ante un problema determinado, pero no se puede precisar de qué manera actuará en una situación real.

Fórmula de corrección:

En la corrección de estos ítems se debe contemplar la incidencia del azar en la elección de la respuesta, siendo este inversamente proporcional al número de alternativas, es decir, cuantas más respuestas menor posibilidad hay de acertar por suerte.

Si no introducimos una puntuación negativa para aquellas respuestas que han sido incorrectas, provocamos que los sujetos respondan a todas, pues el acierto por azar de alguna que no se sabe siempre hará subir la nota. Por ello es que se deben calificar los errores mediante una fórmula que divide la puntuación entre el número de distractores, es decir, si al acierto de una pregunta le otorgamos 1 punto, ofreciendo tres alternativas, una válida acompañada de dos distractores, el error habrá de calificarse como $-1/2$. Si se ofrecen cuatro alternativas, una válida y tres distractores, el error habrá de tener un valor de $-1/3$. Y si son cinco las alternativas, la pregunta mal contestada deberá calificarse con $-1/4$ de punto. De esta manera conseguimos que, estadísticamente, los aciertos por azar se vean justamente compensados.

El emplear una puntuación negativa superior a la de la fórmula supone estar valorando no solo los aciertos sino también los errores, siendo preceptivo comunicar esta circunstancia y demás condiciones de la prueba antes de su comienzo, pues los opositores habrán de saber a qué atenerse. He visto más de un caso en el que cada error se valoró con -1 punto, lo que no deja de ser una desmesura, pues trae como consecuencia notas muy bajas y un altísimo número de suspensos. Si la prueba se compone de 50 preguntas, contestamos a todas y cometemos 13 errores, suspenderemos, pues obtendremos: $37 \text{ puntos} - 13 \text{ puntos} = 24 \text{ puntos}$, o sea, un 4,8 en una escala de 10.

Normas y consejos para la elaboración:

Todos los distractores deben parecer razonables pues, de ser incoherentes o absurdos, darían lugar al acierto por eliminación.



Generalmente la respuesta correcta suele ser la más larga debido a la necesidad de hacerla inequívocamente correcta, por lo que se debe acompañar de distractores suficientemente extensos para evitar claves.

Deben evitarse aquellos ítemes que plantean trampas o juegos de palabras que puedan despistar. Las formulaciones negativas también pueden confundir al examinado. El objetivo no debe ir encaminado a eliminar candidatos induciéndoles a algún tipo de equívoco, sino a cuantificar con la mayor fiabilidad su conocimiento y capacidad para resolver problemas.

Se debe emplear con cautela la alternativa «todas las anteriores» o «ninguna de las anteriores», máxime si surge por el agotamiento del diseñador ante la materia, pues suelen ser de un nivel de dificultad inferior al resto. También es preferible evitar la inclusión de términos absolutos en los distractores, tales como «siempre», «nunca» o «solamente», pues suelen ir asociados a respuestas falsas.

Composición de la prueba

El denominado «examen teórico» suele elaborarse mediante ítemes de selección múltiple, dado que es fácil de responder por parte de los opositores y su corrección no lleva asociada ninguna clase de subjetividad. Sin embargo, su construcción requiere bastante esfuerzo, pues habrá de redactarse un número de preguntas superior al que se vaya a utilizar, realizar diferentes combinaciones y, finalmente, elegir la más coherente y equilibrada. Así es que cada ítem debe entrar a un banco base debidamente etiquetado, indicándose el objetivo a cubrir, el área de contenido o tema al que va referido, su categoría taxonómica (conocimiento, comprensión, aplicación) y su nivel de dificultad estimado.

Lo más recomendable para no liarse es utilizar una tabla de especificaciones. Así, cuando nos enfrentemos a la tarea de componer el cuestionario de examen, podremos ir anotando fácilmente y viendo en qué medida vamos cumpliendo los objetivos. También podremos saber qué partes de la materia están siendo empleadas, cuáles son los contenidos que más veces aparecen y los que menos. Nos daremos cuenta igualmente del nivel al que estamos evaluando, de si por ejemplo hay demasiadas preguntas sobre conocimientos y ninguna sobre aplicación. Se trata, en definitiva, de establecer una relación metódica entre los objetivos, los contenidos y los ítemes que se vayan a utilizar en la medición.