

La ECOE

(Extractos del Informe presentado por la SEMG al Ministerio de Sanidad y Consumo)

Se supone que la última edición prevista de la prueba ECOE es una Evaluación Clínica Objetiva y Estructurada que tiene como función conceder el título de especialista en Medicina Familiar a los licenciados en medicina Pre-95 admitidos en este proceso que la superen. En la pasada edición, la Sociedad Española de Medicina General aceptó tener un representante en el comité técnico de la prueba, exclusivamente en calidad de observador. El fruto fue un completo informe que analiza el desarrollo de la última edición realizada y la conclusión es clara: vulnera la equidad de los aspirantes, además de seguir una metodología inadecuada para la situación que pretende analizar que la hace poco eficiente, con lo que se puede afirmar que se trata de una prueba no apropiada.

Por todos esos motivos, la SEMG presentó al Ministerio de Sanidad y Consumo y al Comité Técnico de la Prueba este informe para que lo tomasen en consideración e incluso ha solicitado posteriormente la suspensión del desarrollo de la prueba que ha comenzado recientemente. En este artículo se resumen las ideas básicas de ese informe.

EL INFORME

El 28 de Agosto de 1998 fue publicado el Real Decreto 1753/1998, que establece un sistema excepcional y transitorio de acceso al título de Médico Especialista en Medicina Familiar y Comunitaria y sobre el ejercicio de la Medicina de Familia en el Sistema Nacional de Salud. Así, también en el BOE 295 del 10 de diciembre de 2005, el Ministerio de Sanidad y Consumo, mediante resolución de la Dirección General de Recursos Humanos y Servicios Económico-Presupuestarios, convocó la licitación del

concurso público de consultoría y asistencia de la prueba objetiva modelo ECOE para el 2005, encaminada al diseño, producción y realización de la prueba objetiva por un importe total de presupuesto base de licitación de 430.000 euros. El concurso fue adjudicado a SEMFyC y desarrollado por una Unión Temporal de Empresas establecida por SEMFyC y SEMERGEN mediante Convenio de Colaboración firmado el 21 de diciembre de 2005. Las distintas pruebas se celebraron en las sedes de Sevilla, Barcelona y Madrid durante cinco fines de semana sucesivos de los meses de febrero y marzo de 2006.

La prueba utilizada para conocer el nivel de competencia profesional de los candidatos corresponde a un tipo concreto de valoración de la competencia por métodos indirectos "tipo multiestación". Su desarrollo es diseñado por el grupo evaluador que recibe la denominación de "Comité de Prueba". En el examen se evalúan de manera ponderada las competencias profesionales de los candidatos, mediante la selección de una serie de situaciones priorizadas entre las más prevalentes y de mayor importancia de las habituales en el ejercicio profesional propio de la atención primaria. Los casos clínicos presentados se estructuran como "estaciones" y cada participante se inicia en una estación distinta. Estas estaciones suelen ser de igual duración y combinan situaciones en las que se examina la capacidad para realizar una anamnesis adecuada, una exploración física correcta, un manejo de las situaciones clínicas competente con un diagnóstico diferencial bien dirigido, un plan de acción o terapéutico acertado, actividades preventivas relacionadas con el caso propuesto y unas habilidades de comunicación que permitan una adecuada relación médico-enfermo o médico-familiar. Las estaciones suelen precisar de la presencia de "actores que

interpretan el guión” diseñado por el Comité de Prueba, que pretende reproducir las circunstancias habituales de una consulta. De manera habitual, las estaciones cuentan con la presencia más o menos explícita de un observador médico que registra, en un formulario especialmente preparado, la realización o ejecución por parte del médico evaluado de una serie de preguntas, exploraciones, actitudes y condiciones preestablecidas y desconocidas para él y que constituyen los distintos ítems evaluativos de la estación indicada.

No obstante, el desarrollo de una prueba evaluativa que utilice el modelo ECOE necesita, de manera adicional, una serie de condiciones de diseño, estructura, desarrollo y análisis exigentes. La celebración de una prueba modelo ECOE para un número de candidatos tan elevado como el que nos ocupa (1.600 aspirantes en esta convocatoria 2005), supone un motivo adicional de preocupación. Mantener las necesarias condiciones de equidad, objetividad, imparcialidad, confidencialidad y eficacia, imprescindibles en una prueba de evaluación para un colectivo con una experiencia profesional y formación continuada de calidad acreditada, que opta a la obtención de una titulación académica especializada, parece *a priori* una tarea que precise de un gran esfuerzo de diseño, garantías profesionales y científicas, infraestructuras, logística y rigor de elaboración, desarrollo y evaluación.

Por ello es necesario disponer de un alto número de espacios para intentar reproducir las condiciones habituales del entorno de trabajo. La selección de posibles sedes es difícil y la posibilidad de descentralizarlas acercando la prueba al entorno geográfico de los candidatos está dificultada, lo que obliga a realizar las pruebas en un número muy reducido de sedes y, por ello, a realizar largos y costosos desplazamientos a los candidatos de Comunidades Autónomas periféricas (Galicia, Canarias, Baleares, Ceuta, Melilla, Asturias...). En este contexto, la necesaria evaluación de distintas áreas competenciales en el seno de situaciones clínicas dramatizadas, con selección de múltiples ítems y la necesidad de cumplimentar por

medio de preguntas o acciones dichos ítems, produce un abigarramiento de las condiciones científicas que condiciona de manera grave el desarrollo de cada estación. La fragmentación en muchos ítems facilita la orientación del manejo global del paciente pero va introduciendo una creciente complejidad en su evaluación de manera que la múltiple sectorización de ítems y su necesario cumplimiento para obtener la máxima nota revierten en un probable olvido (condicionado por la premura de tiempo) de varios de ellos, que repercuten por lo tanto en una disminución de la puntuación global obtenida. Esto es evidente incluso aunque el diseño científico de la estación estuviese bien realizado. Los participantes son plenamente conscientes de la teatralidad del ambiente al que se enfrentan, lo que redundará inevitablemente en una disminución sensible de su rendimiento. Es importante considerar que, incluso dentro del propio examen, el candidato no tiene tiempo para habituarse al modelo, porque no existen estaciones iniciales blancas (sin incidencia en la calificación global) que podrían facilitar el rodaje previo de los candidatos. Además, de las encuestas que se realizan a los candidatos una vez finalizada la prueba, se desprende que la mayoría confirma que la presencia de observadores dentro de las estaciones ha afectado negativamente a su rendimiento. A la vista de las respuestas expresadas de manera anónima por los candidatos tras la prueba, existe un considerable rechazo a este formato, probablemente por infrecuente en el ámbito de población aplicado. La presencia de evaluadores ocultos o camuflados en el interior de la sala produce también un rechazo manifiesto en la mayoría de los candidatos, quienes consideran que menoscaba su rendimiento de manera grave.

La presencia de actores que representan un determinado papel, simulando padecer unos síntomas determinados previamente ensayados con ayuda de personal médico de entrenamiento, resulta en la práctica interpretado con altas dosis de improvisación que introduce un factor de distorsión muy evidente en la prueba. Entre los motivos: los candidatos no están



acostumbrados a este tipo de simulación. Por otra parte, los actores tienen la importante misión de calificar a los candidatos en el componente competencial de habilidades en la comunicación; la cumplimentación de este apartado, compuesto de 10 ítems, debe realizarse en el breve período de tiempo entre actuaciones, que es el espacio teóricamente reservado para descanso.

Con estos preparativos, llegamos a la prueba ECOE-convocatoria 2005 con 1.171 candidatos, para la mayoría de ellos (950) la primera vez que la hacían y para el resto la segunda. Así, en esta convocatoria se ha obtenido un número de asistentes superior a cualquiera de las convocatorias anteriores. Las pruebas se realizaron en 3 sedes durante 5 fines de semana sucesivos. La elección de las sedes se hizo en función de la procedencia mayoritaria de los candidatos (Madrid, las dos Castillas, Cataluña y Andalucía concentraban un 50% del total); por ello, se eligieron las ciudades de Sevilla, Barcelona y Madrid. La presencia de candidatos en relación con el total de convocados en cada una de las sedes fue: Sevilla 69,8%; Barcelona 69,2% y Madrid 78%.

Las cifras obtenidas como calificación en las estaciones que evalúan Habilidades Técnicas y Comunicación fueron las más altas. En concreto la estación "Sutura" obtuvo un resultado global de 64,42 puntos y "Comunicación" de 66,73. Los resultados analizados por ítems de manera individual han presentado una importante asimetría de distribución dentro de cada estación, que es imposible atribuir exclusivamente a cualquiera de los múltiples factores de diseño, de ejecución (actores) o de evaluación (acción de los evaluadores).

Así es que nuestra hipótesis inicial, y nuestro principal motivo de preocupación (pero no único), se centraba en la posible situación de inequidad que podría presentarse, y se refería a la utilización de los mismos materiales y contenidos evaluativos de las estaciones para todas las convocatorias de la prueba realizadas en fines de semana sucesivos; es decir, la utilización de las mismas estaciones con idénticos criterios

evaluativos. El probable efecto copiado debería producirse tras la celebración de la primera convocatoria y por ello mejorar el rendimiento de las sucesivas celebraciones de la prueba. El conocimiento previo de las áreas temáticas que se abordan por parte de los evaluados, así como las competencias que se solicita que sean desarrolladas por parte del examinado, facilitan de manera importante la posibilidad de repasar y preparar específicamente las habilidades que el médico posea. La hipótesis desarrollada se centra en si los candidatos examinados en la primera sede o, más aún, los examinados el primer día, podrían soportar por ello una condición negativa que afectara a la necesaria e imprescindible equidad de la prueba para con todos los candidatos. Así pues, el efecto sede, si se comprobara esta hipótesis, debería producir el tan temido efecto copiado.

A tenor de esta particularidad, se pudieron detectar diferencias significativas relativas al factor sede (peores notas en la sede de Sevilla, sobre todo el sábado). Una vez efectuado el ajuste de factor de confusión Sede-Sevilla-Día, los resultados obtenidos por los candidatos se reflejan en las siguientes puntuaciones promedio: global de todas las sedes: 59,798; global de todas las sedes excluyendo Sevilla: 60,652; global de la sede de Sevilla: 55,982; promedio de la sede de Sevilla el sábado: 54,536.

La simple observación de los datos no deja lugar a dudas en cuanto a la comprobación de la hipótesis. En efecto, las cifras obtenidas como promedio por los candidatos examinados en Sevilla, y sobre todo por los evaluados el sábado (día de inicio de la prueba) y que, por tanto, no tenían posibilidad de ningún tipo de información previa, son significativamente menores que las del resto de candidatos que se examinaron en fines de semana posteriores, que ya conocían las áreas temáticas y los aspectos que se iban a evaluar. La falta de equidad que de ello se deriva, y que per se es motivo suficiente a nuestro criterio como para confirmar la existencia de efecto copiado, se agrava al establecerse la nota de corte que clasifica a los candidatos en *aptos* y *no aptos*. El informe elaborado

por el Comité de Prueba reconoce que para elegir una nota umbral definitoria de los dos grupos que considera necesario diferenciar, y dado que la distribución estadística de las calificaciones obtenidas es muy parecida a una distribución normal, sin la aparición de grupos no posee ningún criterio que le permita establecer un umbral de corte natural; utiliza simplemente la distribución de notas globales. Por todo ello, y en función de un criterio forzado, designado previamente, que pretende de manera obligada dividir de manera artificial a los candidatos, no en función exclusiva de sus resultados individuales sino de sus resultados comparados con los del global de candidatos, ese informe recurre a un criterio estadístico discriminatorio y determina el umbral de corte en los 55 puntos. La nota media obtenida por los candidatos examinados en primera convocatoria del sábado en la sede de Sevilla fue de 54,536 (situada claramente por debajo del punto de corte) y el promedio de calificaciones global de todos los evaluados en la sede de Sevilla se situó en 55,982, muy cerca del punto considerado como discriminatorio para el establecimiento de los grupos de *aptos* y *no aptos*. Por todo ello es esperable que la proporción de *no aptos* en la convocatoria de Sevilla (especialmente en su primer día de celebración) haya sido muy elevada en relación con el resto de sedes.

Por todo lo expuesto, la SEMG considera que está demostrado el *efecto copiado*, que se agrava por la elección de un sistema de selección de *nota de corte* con criterios erróneos. Todo ello afecta de manera manifiesta a la necesaria equidad de la prueba, y condiciona la calificación final como *apto* o *no apto* de un número importante de candidatos que pueden haber sido perjudicados de manera injusta en su pertinente intención de acceder a una titulación que podrían haber conseguido si la prueba se hubiera efectuado en otras condiciones. Este *efecto copiado*, pues, perjudica de manera innegable a los médicos que se examinaron el primer día, lo que para esta sociedad científica supone una clara discriminación negativa que, por sí sola, debería ser suficiente para invalidar la prueba.