

Los simuladores clínicos escritos para evaluación

Psic. Adrián Cañedo Dorantes
Investigador del Centro Latinoamericano de Tecnología Educativa para la Salud, México.

Dr. Fernando Suárez Sánchez
Jefe de la Sección de Evaluación de la Secretaría de Educación Médica.
Facultad de Medicina. UNAM.

Antes de 1970, la simulación ofrecía la tecnología tanto para la instrucción como para la determinación de la competencia para la solución de problemas. Sin embargo, a pesar de la gran difusión que se había dado a este procedimiento, su empleo se circunscribía a esferas de acción muy reducidas, conociéndose poco en las escuelas, donde no se le utilizaba en la enseñanza de las estrategias para la solución de problemas. Una de las principales razones para la limitación de su uso seguramente fue la sospecha del costo muy elevado que la sola palabra "simulador" suponía. Fue necesario entonces, que se demostraran las grandes posibilidades que la simulación ofrecía en cuanto a la obtención de magníficos resultados en la enseñanza-aprendizaje de cualquier situación de la vida cotidiana. Lo anterior no sólo implicaba reducción de costos, sino facilidades para su utilización en cualquier lugar y momento donde su aplicación fuera de utilidad.

De acuerdo con McGuire y cols., la simulación consiste en colocar al individuo en una situación realista para confrontarlo con un problema que deberá resolver mediante una secuencia de selección de posibilidades, decisiones y acciones. Cada una de estas posibilidades le dará acceso a re-

troalimentación apropiada que podrá o no podrá modificar la situación, pero que deberá utilizarse para tomar decisiones subsiguientes en relación con las acciones que hayan quedado pendientes, y que a su vez podrán modificar el problema. La situación inicial evoluciona a través de varias etapas hasta que termina cuando el individuo alcanza una resolución aceptable o se enfrenta a alternativas fatales por la sola configuración de sus propias selecciones y acciones.

Teniendo en cuenta la esencia de la simulación delineada por McGuire, y en vista de las necesidades existentes para una utilización más amplia y económica de este instrumento, el Centro para el Desarrollo de la Educación del Colegio de Medicina de la Universidad de Illinois produjo tres modalidades de simulación útiles como instrumento de enseñanza y que también se emplearon para evaluación en grupos pequeños y mayores. Una de esas modalidades fue la simulación escrita, que utiliza la imagen latente o cualquier otro tipo de técnica que evite la exposición de los sistemas de retroalimentación no solicitados; puede utilizarse en forma individual o de grupo sin perder individualidad, y que al mismo tiempo es susceptible de revisión y análisis por computación. La simulación escrita tiene además un costo notoriamente bajo y se puede reproducir casi en cualquier lugar.

La simulación escrita tiene ciertas características como son la necesidad de iniciarse de manera realista; contener series de decisiones secuenciales e interdependientes que representen las diferentes etapas de la definición, análisis y resolución del proble-

ma; proporcionar en forma realista la información relacionada con la pregunta haciéndolo como base de acciones subsecuentes; que una vez obtenidos los datos sea imposible retractarse de la decisión tomada; y estar construida de tal manera que proporcione diferentes posibilidades para cada acción decidida.

En estas condiciones, la simulación escrita constituyó un instrumento útil para enseñar al individuo a tomar decisiones, puesto que el resultado de su acción era rápidamente conocido, y en base a él se podía continuar con la resolución del problema presentado y agregar un nuevo problema al ya existente, el que a su vez tendría que ser solucionado a menos que la decisión fuera tan grave que no pudiera ser reparada y entonces se acabaría todo el problema, aunque el desenlace no fuera el deseado.

Tal como estaba diseñada, la simulación escrita empezó a utilizarse para la enseñanza en medicina, pero pronto se vieron posibilidades de ampliar su campo de acción al terreno de la evaluación, pues se pensaba que el manejo de un problema clínico escrito permitiría conocer el comportamiento del médico ante un problema clínico real.

Para objetivizar ese conocimiento, se diseñaron varios sistemas de puntuación como base para el establecimiento de cifras que pudieran catalogar al individuo en su desempeño "clínico". El sistema más empleado fue el de asignar puntuación positiva o negativa a cada una de las opciones ofrecidas, tanto en las secciones de información como en las relacionadas con la toma de decisiones, y ponderarlas de acuerdo al valor que se atribuyera a cada sec-

ción como parte integrante del todo. Como resultado de esto, se obtuvieron varios índices que señalan los errores de comisión (aquello que se hizo y no debería haberse hecho), los errores de omisión (aquello que debería haberse hecho y no se hizo); la habilidad (relación entre las cosas útiles y el total de las opciones escogidas) y la competencia del que resuelve el simulador (manejo integral del caso).

Estos índices se consideraron lo suficientemente útiles como para obtener resultados que se puedan relacionar estrechamente con la situación real del evaluado; por ello, los simuladores clínicos escritos se aplicaron con fines de evaluación. Sin embargo, en nuestro medio, los resultados obtenidos con la aplicación de cerca de 90,000 simuladores a un total de casi 30,000 alumnos durante dos años, nos ha permitido concluir que, tal como esta diseñado en la actualidad, este instrumento es poco confiable para utilizarse en evaluación.

El análisis de la situación hizo posible la detección de ciertos aspectos que contribuyen a que disminuya la confiabilidad de los simuladores clínicos escritos como instrumentos de evaluación.

Los principales problemas encontrados son:

PRIMERO. El problema de la estrategia. Para la resolución de un simulador escrito, se presentan varios posibles caminos; pero algunos de ellos eliminan secciones enteras, lo que da lugar a que la puntuación obtenida por un alumno que resuelva bien el caso saltando secciones, pueda ser menor que la del que resuelva mal el caso pero que pase por todas las secciones. Esto sucedió con un gran grupo de alumnos que, trabajando

Opción	1	2	3	4	5	6	7	8	9	10
Valor	4	0	0	1	2	-4	-2	-1	2	4

Figura 1. Valores de cada opción en una sección dada.

con el mismo caso clínico simulado, obtuvieron menor número de puntos como egresados del curso que como aspirantes al mismo. Vemos que lo anterior es factible si tomamos en cuenta que el especialista, por tener mayor familiaridad con el problema que se le presenta, puede manejar su caso clínico simulado en forma más directa, dejando de revelar opciones que por su implicación obvia da por supuestas, a diferencia del estudiante, que requiere mayor información para descartar o aceptar una hipótesis. Pero también este resultado puede conducir a la peligrosa conclusión de que durante el tiempo que dura dicho curso, los alumnos "desaprenden", lo que pondría en tela de juicio el sistema para preparar especialistas.

SEGUNDO. Las condiciones de aplicación del caso clínico simulado constituyen otro problema, ya que la existencia de retroalimentación sobre los errores de manejo propicia las trampas. Haciendo un símil con el método de opción múltiple, la retroalimentación sobre los errores sería lo mismo que poner una nota que dijera "esta opción es falsa" lo que daría lugar a que nadie, por honesto que fuera, se quedara con las respuestas inadecuadas, sobre todo cuando la evaluación constituyera un factor determinante en su carrera. Este problema trae como consecuencia la dificultad en el manejo de la tarjeta perforada y con ello el incremento en tiempo y costo que implica utilizar otros procedimientos.

TERCERO. Un tercer problema es la calificación atomística, es decir, la asignación de puntos (positivo, negativo o cero) a cada opción del simulador. Hemos encontrado que, con este procedimiento, pueden surgir situaciones contradictorias

como es el hecho de que alguien que maneje bien el caso clínico simulado obtenga menor número de puntos que aquél que cometa errores. Para aclarar esta situación veamos algunos ejemplos. Supongamos que en una sección existen diez opciones, y que a cada opción se le asignó un valor, como se muestra en la Figura 1. Un estudiante que hubiera revelado las opciones 4, 5 y 10 obtendría 7 puntos, y no habría cometido errores de comisión, aunque si dos errores de omisión, el 1 y el 9. En cambio, el estudiante que hubiera revelado las opciones 1, 5, 6, 9 y 10, obtendría 8 puntos, a pesar de haber cometido un error de omisión, el 4, y un error grave de comisión, el 6. Como lo que cuenta para el resultado es la eficiencia global y no las dimensiones de los errores, el número de puntos para la sección es mayor para el segundo alumno, que comete errores, sin tomar en cuenta que el primero no reveló la opción 10, posiblemente porque ya tenía toda la información que suponía necesaria para el manejo de su caso.

CUARTO. Un cuarto problema, que consideramos más grave, es la diversidad de atribución de puntos que existen entre las secciones de información y las de estrategia. Por ejemplo, un estudiante puede acumular 20 puntos durante el interrogatorio, y lo más que puede perder en la sección de toma de decisiones correspondiente son tres puntos, pues sólo se le permite que revele una opción. Si la decisión es inadecuada, parecería que la información proporcionada por el interrogatorio fue inútil, ya que con ella tomó una decisión inadecuada. En cambio, otro estudiante que haya obtenido 12 puntos en el interrogatorio y haya tomado una decisión adecuada tendría menor puntaje que

el anterior, no obstante haber utilizado más adecuadamente la información proporcionada.

Conociendo estos problemas, decidimos elaborar un tipo de simulador clínico escrito que, manteniendo el mismo realismo, pudiera utilizarse con mayor confiabilidad como instrumento de evaluación. Para ello, volvimos a diseñar un caso clínico simulado, que se había construido con la técnica convencional, haciéndole las siguientes modificaciones:

1. Se añadieron secciones de diagnóstico de presunción, y se colocaron antes de cada sección de toma de decisiones.

2. Se quitó toda la retroalimentación (negativa y positiva) de las secciones de información.

3. Se retiraron algunas opciones de las secciones de toma de decisiones (sólo con el fin de facilitar el trabajo por ahora).

4. Se añadió una sección de control para la detección de trampas cuando se utilice la tarjeta perforada.

5. Se cambió la filosofía de la calificación, tratando de evitar las contradicciones que se producen con la calificación atomística.

Lo anterior hizo que el simulador fuera más sencillo en su construcción. Empieza con una escena inicial de características similares a las del convencional, pero inmediatamente después se le pide al que lo resuelve que, con la información proporcionada, seleccione una sospecha diagnóstica de entre las posibilidades que se le ofrecen; con ello, podemos juzgar qué tan útil fue para él tal información, y a partir de que pensamiento diagnóstico empezará el estudio clínico. Pasará después a la sección con la que él mismo decida continuar, y en ella seleccionará la información que requiera, con lo que podrá confirmar o descartar su sospecha inicial; es obvio que la selección de la información en cada sección deberá estar fundamentada sobre la sospecha diagnóstica, y que podrá ampliarse cuando tal sospecha pierda fundamento con los resultados obtenidos; de esta manera queda justificada la

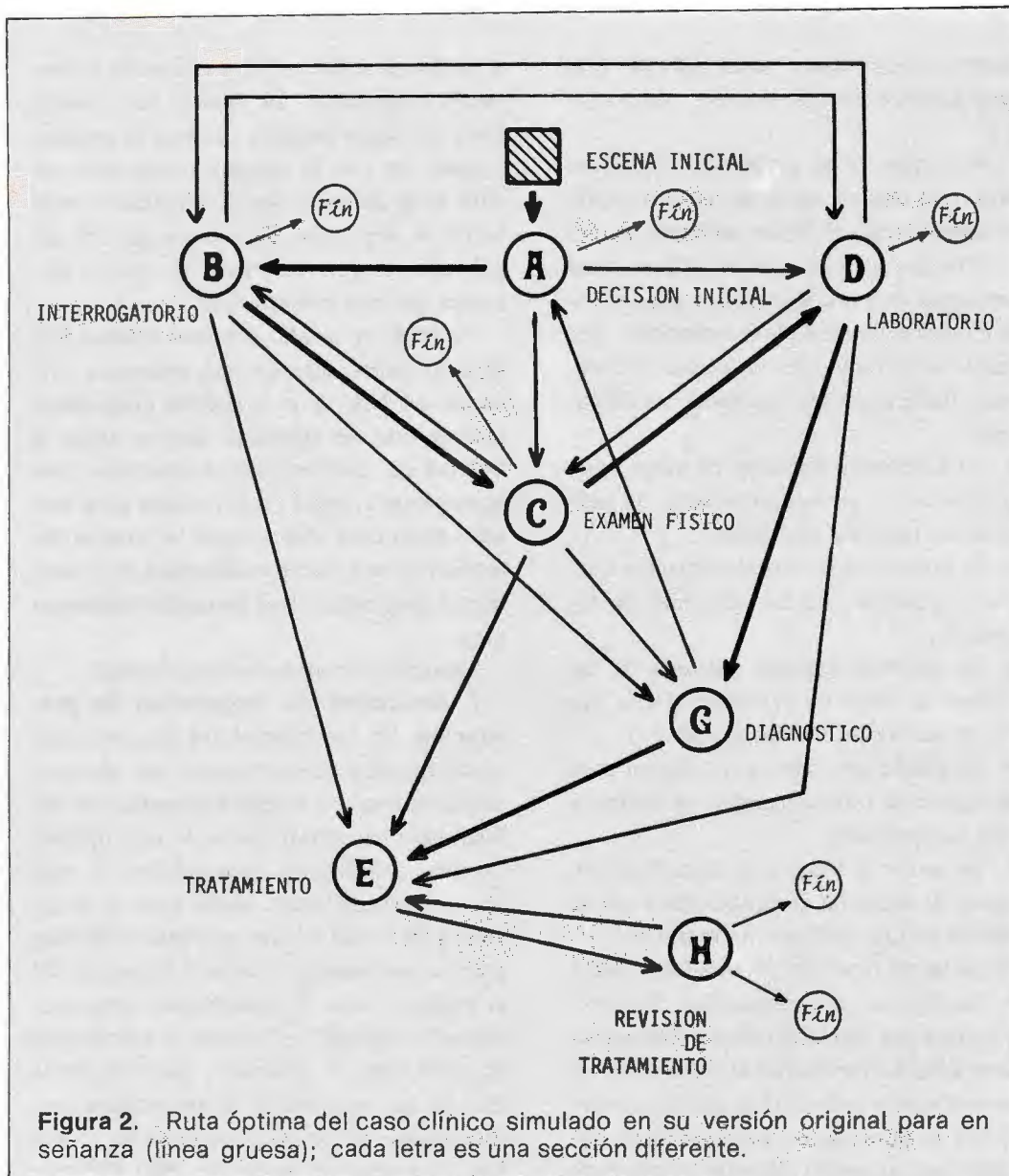
selección de información y el cambio de sospecha diagnóstica. El manejo será similar cada vez que se requiera cambiar de sección; y cada vez que se agregue nueva información a la anterior, podrá afinarse o cambiarse la impresión diagnóstica que irá adquiriendo mayor valor cada vez que se disponga de más información.

Al final, el estudio del caso clínico presentado habrá seguido una secuencia planeada en base a la sospecha diagnóstica, aunque esto no signifique que se limita la libertad del que resuelve el simulador para equivocarse cuantas veces lo desee en el manejo de su caso clínico; pero si para la obtención de una buena calificación, se evitará que el diagnóstico final no quede fundamentado.

Veamos lo anterior con más detalle.

1. Secciones de diagnóstico de presunción. En los simuladores de enseñanza, la información proporcionada por una sección es si revela o no una respuesta, y se califica bajo un patrón único: la ruta óptima. Cuando aplicábamos simuladores de este tipo e interrogábamos sobre porqué se revelaba tal o cual opción, nos encontrábamos ante un razonamiento como el siguiente: "Si el paciente tiene X enfermedad debe presentar Y síntoma", o "si tiene X enfermedad no debe tener Y síntoma", que nos dio la idea de que una sección de información podría resolverse, ya no a través de un patrón fijo, sino desde el punto de vista de comprobar o desechar un diagnóstico; o mejor dicho, de buscar e interpretar la información a partir de una hipótesis. Esto era posible si se conocía previamente la hipótesis, o sea el diagnóstico de presunción, independientemente de que la hipótesis fuera correcta o incorrecta.

La situación inicial proporciona al alumno información suficiente para suponer un diagnóstico; este diagnóstico A, le orientará en el tipo de exploración o interrogatorio que deberá realizar, y sería obligado preguntar antecedentes familiares en caso de que la enfermedad supuesta tuviera componentes



etiológicos de tipo genético; pero ese tipo de interrogatorio podría no ser tan útil en caso de que la enfermedad supuesta fuera de índole infecciosa, la que deberá tener sus opciones relevantes propias. Si el estudiante supone el diagnóstico A, y la información que le permite desechar o comprobar dicho diagnóstico se encuentra en las opciones x, y, z, estará utilizando una estrategia adecuada si revela tales opciones; pero si no las revela, su método de obtener información será inadecuado, independientemente de que el diag-

nóstico A sea correcto o incorrecto. Supongamos ahora que las opciones x, y, z, proporcionaron tal información que, con ella, se descartaría el diagnóstico supuesto; si a pesar de ello el estudiante persiste en el diagnóstico A, la interpretación de la información obtenida no será correcta; pero si además persiste con el mismo diagnóstico equivocado en varias secciones, se podrá deducir la rigidez de pensamiento para descartar una hipótesis que no tiene elementos para ser sostenida. Cuando analicemos el proce-

dimiento de calificación trataremos sobre otras características que se obtienen al añadir la sección de diagnóstico de presunción.

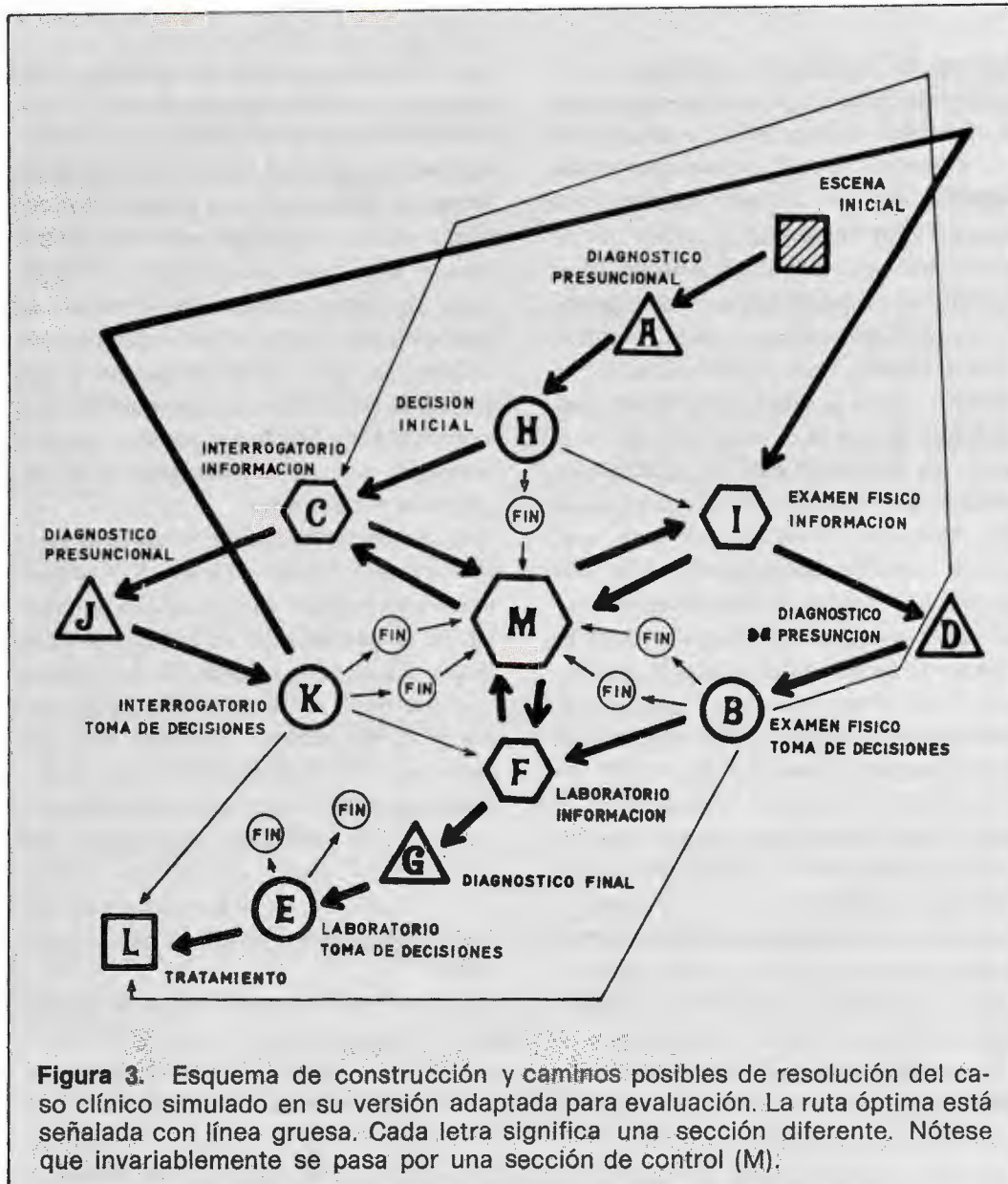
2. Eliminación de la retroalimentación negativa. En el simulador convencional (anexo 1), se encuentran opciones que al ser reveladas proporcionan la siguiente información: “El paciente regresa al día siguiente y lo ve el jefe del curso quien considera que el tratamiento estuvo mal indicado, por lo que lo llama a usted a su oficina para notificarle que no se le puede aceptar en el curso”. Es muy difícil para un estudiante en situación de evaluación, permanecer ecuánime ante una información de este tipo, pues la sensación que le queda es de que “no aprobó” pues no es consciente del procedimiento completo de calificación. Por ello la hemos eliminado del simulador de evaluación que diseñamos (anexo 2), aunque permanece en el programa que lo califica y el alumno la conoce unos días después de haber manejado su caso clínico simulado. Con ello, se pierde la retroalimentación inmediata; pero, tratándose de una situación de evaluación y no de enseñanza, no resulta grave y permite un control más confiable, ya que el alumno desconoce si ha manejado adecuada o inadecuadamente el caso clínico simulado, hasta que se le entregan los resultados.

3. Simplificación de la toma de decisiones. En el simulador clínico escrito convencional se encuentran las opciones 4 y 5 después de la escena inicial. En la primera de ellas, se le indica al alumno que escoja otra opción en esa misma sección. Por ser cíclica, esa opción puede permanecer puesto que también ayuda a conservar el realismo del simulador. En cambio, la opción 5 le indica: “Fin del problema” lo que puede dar lugar a que se hagan trampas ya que el estudiante se da cuenta inmediata de que no anda por el camino correcto y, por lo tanto, consideramos necesario eliminarla. Analizando las frecuencias en los simuladores que se han aplicado con anterioridad con fines de evaluación, nos dimos cuenta de que ningún estudiante o solo uno en mil, escoge este

tipo de opciones, por lo que también consideramos conveniente su eliminación, simplificando así la toma de decisiones al obligar al alumno a meterse necesariamente en las secciones importantes. En el simulador clínico escrito convencional, existe una estructura tal (Figura 2) que permite al estudiante dejar sin contestar hasta dos secciones, lo que complica mucho el procedimiento de calificación que hemos propuesto y que pone más atención en la evaluación del procedimiento que siguen los alumnos en cada actividad, como el interrogatorio y la exploración por ejemplo.

Lo anterior nos decidió a tomar una estructura más simple (Figura 3), aunque todavía no estamos conformes con la eliminación de posibilidades en la toma de decisiones. Estamos conscientes de que pueden quedarse todas las opciones analizadas menos la 5. Sin embargo, pensamos que esto podrá lograrse en un siguiente paso, ya que confiamos en que el simulador de evaluación propuesto, de estructura más simple, nos proporcionará mayor claridad para encontrar el método que permita trabajar con este tipo de opciones sin complicar tanto el problema.

4. Sección de control. Existe la posibilidad de que un estudiante maneje el simulador de alguna manera y perfore la tarjeta de otra; para fines de evaluación esto es un error que podría tirar por tierra todo el proyecto. Para evitar este error se añadió la sección de control, cuya utilización obliga al alumno a perforar dos opciones en lugar de una sola en determinado momento del manejo de su caso clínico simulado. Si dichas opciones son cruciales dentro de las secciones, el alumno que revele alguna de ellas necesariamente revelará la otra y cualquier incongruencia podrá ser fácilmente descubierta al observar la última página del simulador, donde se coloca la sección de control. Por ejemplo, en la sección de laboratorio y gabinete del simulador de evaluación que se presenta en el anexo 2, existe una opción que permite enterarse de los resultados ob-



tenidos con una serie gastroduodenal y que proporciona la clave del diagnóstico diferencial al demostrarse hernia del hiato diafragmático. El alumno que revele esta opción se remitirá a la sección de control para que revele otra donde se le señala: "Esta opción es de control para el uso adecuado de la tarjeta, perforala y vuelva a la sección..." El procedimiento permite la clasificación de los estudiantes en cuatro grupos (Figura 4).

La máquina computadora envía una lista de los folios que caen en cada una de las

categorías, y con ello se pueden seleccionar los simuladores en las mismas categorías y determinar qué estudiantes trataron de hacer trampa observando sólo la última página y descubriendo allí cualquier incongruencia, como es el hecho de que la opción haya sido revelada pero que la tarjeta no haya sido perforada.

Hicimos un ensayo de lo anterior y tomamos tiempo para determinar que tan factible sería ver la última hoja, clasificar el simulador, cotejar si su número de folio estaba o no en la categoría apropiada, etc., y encon-

tramos que se puede realizar en 8 horas si se revisan 3,000 simuladores entre cuatro personas, lo que consideramos adecuado si tenemos en cuenta la calidad de la información que se obtiene y que básicamente es la detección de las trampas y verificación del empleo adecuado de la tarjeta de respuestas. Esto no puede obtenerse con el sistema convencional, ya que la revisión de la totalidad del cuaderno de respuestas, sin que se tome en cuenta la tarjeta, es un procedimiento muy laborioso que implica perforación de tarjetas suplementarias, lo que redundaría en elevación de costos y retrasos considerables en la obtención de los resultados.

5. Calificación heurística. Si una persona sabe que otra está enferma y esto es lo único que conoce, difícilmente podrá hacer algún diagnóstico ya que el término "enfermo", implica cualquier enfermedad. Pero si, además de ello, sabe que la persona está enferma del aparato digestivo, podrá encaminar su interrogatorio o exploración, ya que las enfermedades del aparato digestivo son un subconjunto de todas las enfermedades; si a lo anterior se agrega información acerca de que hay fiebre y ésta ha ido aumentando cada día, el pulso es lento, hay diarrea y deshidratación, la enfermedad empezó como gripa, y el paciente tomó "agua fresca" unos días antes de iniciarse la enfermedad, entonces el diagnóstico podrá estar mejor fundamentado. Este esquema nos muestra que la información que el sustentante tiene al principio le permite sospechar un cierto número de posibilidades diagnósticas y que la estrategia que deberá seguir estará orientada a captar la información que le permita reducir las posibilidades diagnósticas hasta llegar a una sola, el diagnóstico correcto. En este procedimiento existen tres elementos importantes: 1) La información que permite al estudiante hacer una hipótesis, es decir, la que enfoca el campo de posibilidades; 2) una estrategia que permita obtener información adicional; y 3) una interpretación adecuada de la información tanto inicial como adicional

A	0	0
B	0	1
C	1	0
D	1	1

Figura 4. Clasificación de los grupos de estudiantes.

para poder aceptar o descartar la hipótesis.

En los simuladores de evaluación, la hipótesis está dada por el diagnóstico de presunción. Cada diagnóstico que se presenta al alumno tiene cierta jerarquía que permite conocer el valor que la información previa le ha dado. En el simulador de evaluación que se presenta, el diagnóstico "pancreatitis" tiene menos posibilidades de certeza que el diagnóstico "úlceras gástricas" por la información que se obtiene, desde la situación inicial hasta la indicación terapéutica. Dicho diagnóstico está asociado a determinada estrategia; por ejemplo: el sustentante que, habiendo pensado en pancreatitis después de la situación inicial, entre a la sección de interrogatorio y tenga una estrategia adecuada, deberá revelar las opciones 13, 14, 15, 17, 18, 21, 22, 23, 26, 27 y 29 y cometerá errores en su estrategia si revela las opciones 1, 2, 3, 4, 5. Si al terminar de interrogar persiste en su diagnóstico de pancreatitis, sabiendo que ninguna de las opciones reveladas le ha proporcionado elementos para confirmar su hipótesis, estará interpretando la información de manera inadecuada. Existe la posibilidad de que un alumno entre a la sección de interrogatorio con la hipótesis de pancreatitis, revele las opciones 15, 17 y 21, y cambie su hipótesis por la de úlcera gástrica porque la información no le proporcionó ninguna evidencia. Esto hará que cambie la estrategia aún dentro del interrogatorio y lo sabremos porque el diagnóstico de presunción que escoja al terminar la sección será el de úlcera gástrica. Lo anterior obliga

A	2
A y B	1
B	9
No A	6
No B	8
No A y B	3
A y no B	4
No incluidos	5 y 7

Figura 5. Estrategia resultante de las posibilidades diagnósticas A y B.

a calificar al sustentante no sólo por un patrón de estrategia sino por el conjunto de ambos. Veamos un ejemplo. Supongamos que para el diagnóstico A, la estrategia adecuada para interrogar al paciente sean las opciones 1 y 2, y la inadecuada (No A), sean las opciones 3 y 6; y que para el diagnóstico B, la estrategia adecuada sean las opciones 1 y 9, y la inadecuada (No B), la 4 y la 8. La estrategia resultante de ambas sería como se muestra en la figura 5.

Si juntamos ambas estrategias, las opciones que el sustentante deberá revelar son las que incluyen A, AB y B; las opciones que no deberá revelar son No A, No B, y No A y No B; y las opciones que podrá revelar sin que le afecten la calificación son No A y B, A y No B, y las no incluidas. En este ejemplo, el sustentante que siga una estrategia adecuada deberá revelar las opciones 1, 2 y 9, pero también podrá revelar la 3, la 4, la 5 y la 7, sin modificar por ello lo adecuado de su estrategia; pero cometerá errores si revela las opciones 3, 6 y 8 (Figura 5).

Para simplificar el problema, por el momento, hemos hecho la restricción de conjuntar en una estrategia dos diagnósticos solamente, pero esto se puede hacer con tres o más. Con el objeto de asegurar la cohe-

rencia de posibilidades, se diseñó un programa que conjunta de dos en dos todos los diagnósticos propuestos para cada sección de información en el simulador que envía información como la que aparece en la figura 6.

Con la información anterior se garantiza que el constructor de simuladores ponga las opciones de tal forma que, para cada sección el apareamiento de dos diagnósticos produzca opciones que deban revelarse y opciones que no deban revelarse. Además, existe otro camino para determinar las opciones que deben y las que no deben revelarse; nos referimos a las opciones lógicas, que son cadenas de opciones unidas por conectores del tipo (.) lo que indica que una opción A y otra B deben revelarse ambas para decidir si alguna de ellas o las dos pertenecen a la categoría de "deben revelarse" o a la de "no deben revelarse". El conector (V) indica que se puede revelar A o B o ambas. El conector (o), el OR excluyente, indica que solamente puede revelarse A o B pero no ambas, para que se cumpla la condición.

Este tipo de operaciones lógicas enriquece las posibilidades para calificar más realísticamente el simulador, pues permite desechar una contradicción típica en la calificación atomística: la acumulación de puntos por el error de revelar dos opciones contradictorias entre sí pero buenas por sí solas. Por ejemplo, en la sección de tratamiento de un simulador pueden presentarse las opciones "tetraciclina" o "ampicilina"; si se seleccionan ambas, el tratamiento será incorrecto, pero cualquiera de ellas por separado puede hacer un tratamiento adecuado. En el simulador convencional, si un estudiante revela ambas opciones obtiene mayor puntaje que el que no comete el error.

El siguiente paso por calificar es el diagnóstico. Con la calificación atomística, el diagnóstico tiene una escala, y entre más se acerque el sustentante al diagnóstico correcto mayor será el puntaje que se obtenga, aunque al fin y al cabo no se incluya

en la eficiencia global. Nuestra proposición de incluir secciones de diagnóstico de presunción en el simulador proporciona mayor calidad a la calificación del diagnóstico, ya que no únicamente se conoce el diagnóstico definitivo, que el alumno selecciona, sino que también se sabe el proceso mediante el cual llegó a él. No sólo podremos conocer la información que requirió para llegar al diagnóstico definitivo, sino también la que le llevó al diagnóstico diferencial, el grado de rigidez para cambiar de diagnóstico y la "intuición diagnóstica". Para ello contamos con la forma de calificación que se muestra en la figura 7.

A cada sección de la ruta se le asocia un diagnóstico y se le pesa o no dependiendo de si la pregunta ¿estaría en un grave error? se contesta en forma afirmativa o negativa, suponiendo una información previa adecuada. Con este procedimiento tenemos una plantilla para cada sección de diagnóstico de presunción que nos indica cual diagnóstico esta pesado y cual no. Como cada sección se califica por dos diagnósticos, el antecedente y el consecuente, las posibilidades que podremos encontrar son las que mostramos en la figura 8.

Como el número de posibilidades es alto, hemos pensado que cuando el sustentante no pefore la sección de diagnóstico (SD) o la conteste erróneamente perforando dos o más diagnósticos (DD), le podremos asignar la estrategia Diagnóstico Correcto (DC), añadiéndole un peso; y este Diagnóstico Pesado (DP), además de que nos permite asignarle una estrategia dentro del programa, nos reduce la tabla de posibilidades a las que se muestran en la figura 9.

Las posibilidades 2 y 4 no afectan a ninguna sección, pero las posibilidades 1 y 3 le quitan un cierto peso a la sección pues la información proporcionada no ha sido de utilidad, lo que hace que persista el diagnóstico incorrecto, lo cual para esos momentos es un error grave. La excepción a la regla la constituye la posibilidad 2, siempre y cuando DP se produzca por per-

Número de reactivos de la sección de interrogatorio (xx)	
Errores:	
A. No A	No hay
A. No A.B	No hay
A. No A No B	No hay
A. No AB No B	No hay
A. B No B	No hay
B. No B	29****
No A.B. No B	No hay
Diagnóstico: <i>debe revelar</i>	
A	xx, xx, xx, xx
A.B	No hay
B	xx, xx, xx, xx
Porcentaje de la sección	%
<i>no debe revelar</i>	
No A	xx, xx, xx, xx
No B	No hay
No A. No B	xx, xx, xx, xx
Porcentaje de la Sección	%
Incongruencias: (items eliminados)	
No A. No B	x
No A.B	x
Porcentaje de la Sección	%
Items que no se toman en cuenta:	
x, x, x, x, x, x, x.	
Porcentaje de la Sección	%

Figura 6. Diagnóstico 1 comparado con el diagnóstico 2.

forar diagnósticos múltiples o por no haber contestado. En ambos casos, se deberá aplicar un peso de castigo a la sección.

Nuestra proposición para calificar la toma de decisiones o, en otras palabras, la ruta del simulador, es que se proporcione a cada ruta un cierto valor global. Como se había señalado para un simulador tipo, las posibles rutas no deben exceder de 15. En los casos en que se haya descubierto alguna trampa, se le asignará un valor CERO a la toma de decisiones.

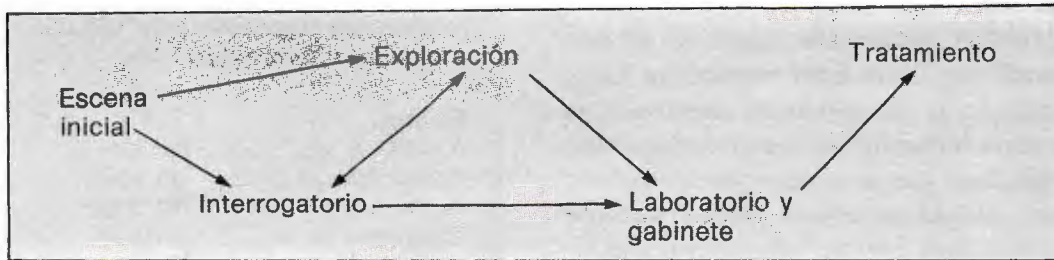


Figura 7. Rutas posibles en el simulador de evaluación anexo.

Nos queda por discutir la eficiencia global. El método que sugerimos para su calificación es el siguiente:

Todo simulador de evaluación está constituido por los siguientes tipos de secciones: A, secciones de información; B, secciones de diagnóstico; C, secciones de toma de decisiones; y D, secciones de control. Cada sección, dependiendo de los objetivos del simulador, se pondera en una gráfica circular que familiarmente llamamos "pastel" (Figura 10). Puede tomarse un ejemplo del simulador de evaluación que presentamos. A cada trozo del pastel le asociamos un porcentaje; cada trozo alcanza también, internamente, un cierto porcentaje que se transforma según el porcentaje relativo que le haya tocado a esa sección. La eficiencia global no es más que el promedio de cada porcentaje pesado dentro del pastel. Veamos un ejemplo para aclarar la situación.

Supongamos que las secciones del simulador están ponderadas de la siguiente forma:

Interrogatorio	25%
Examen físico	25%
Laboratorio y gabinete	25%
Tratamiento	10%
Diagnóstico	10%
Estrategia	5%

Supongamos ahora que un sustentante obtiene los siguientes puntajes para cada sección:

Interrogatorio	50%
Examen físico	100%
Laboratorio y gabinete	50%
Tratamiento	100%
Diagnóstico	50%
Estrategia	100%

Entonces, su eficiencia global será como sigue:

$$50 \times .25 + 100 \times .25 + 50 \times .25 + 100 \times .10 + 50 \times .10 + 100 \times .05$$

lo que da un total de 70%.

6. **Puntajes.** En los aspectos tratados se evitó tocar en lo posible los puntajes concretos que se le asignan a cada sección, ya que interesaba explicar el método con más generalidad; pero, una vez hecho esto, podremos tocar dicho punto y, para ello, comenzaremos con la asignación de puntajes a las secciones de información. Tenemos dos grandes grupos de opciones, las que se deben revelar y las que no se deben revelar. Hemos dividido a cada grupo en tres subgrupos que, en el caso de las opciones que se deben revelar, son: *Forzosas*: aquellas opciones que le dan la información más importante en una estrategia; por ejemplo, si el sustentante tiene como diagnóstico de presunción "diabetes mellitus", es forzoso mandar hacer un análisis de sangre para determinar la concentración de glucosa en ella, ya que la información que le proporcione le será de mucha utilidad. *Importantes*: aquellas opciones que le dan al sustentante el marco general para orientarlo sobre el grado de avance del padecimiento, o que le dan información sobre el tipo de complicaciones más frecuentes. En el caso del simulador de la hernia hiatal, será importante conocer el tamaño de la hernia, saber si existe sangrado, etc. *Rutinarias*: opciones que, sin estar directamente relacionadas con el padecimiento, son necesarias para conocer el estado general del paciente, como sucede en situación de exploración con la medición de las variables

Tabla de posibilidades	
Antecedente	Consecuente
SD	SD
SD	DD
SD	DP
SD	DC
DD	SD
DD	DD
DD	DP
DD	DC
DP	SD
DP	DD
DP	DP
DP	DC
DC	SD
DC	DD
DC	DP
DC	DC

Figura 8. Posibilidades para la calificación de cada sección por dos diagnósticos, el antecedente y el consecuente. SD = Sin Diagnóstico; DD = Doble Diagnóstico; DP = Diagnóstico Pesado; DC = Diagnóstico Correcto.

vitales.

En el caso de las opciones que el sustentante no debe revelar, los subgrupos están formados por: *Errores leves*: aquéllos que al cometerse no causan daño significativo al paciente, como pudiera ser el preguntar antecedentes hereditarios cuando el paciente se ha roto una pierna. *Errores graves*: aquéllos que al cometerse resultan costosos, le causan molestias al paciente o inclusive, pueden poner en peligro su vida. Un error de este tipo sería, por ejemplo, realizar una intervención quirúrgica sencilla, pero innecesaria. *Errores fatales*: aquéllos que, al cometerse, causan la muerte del paciente, directa o indirectamente.

El tratamiento que damos a las opciones que el sustentante debe revelar es distinto al que damos a las opciones que no debe revelar. A estas últimas, les asignamos pesos directamente, procedimiento muy similar al de asignación de puntajes en la calificación atomística, pues nos hemos dado cuenta en las sesiones de validación de un simu-

Nueva tabla de posibilidades	
Antecedente	Consecuente
DP	DP
DP	DC
DC	DP
DC	DC

Figura 9. Reducción de la tabla de posibilidades por aplicación de la estrategia de DP (Diagnóstico Pesado) y DC (Diagnóstico Correcto).

lador clínico escrito, que los médicos generalmente están de acuerdo al señalar los errores y que éstos se penalizan disminuyendo en un porcentaje determinado la eficiencia de la sección. Por ejemplo, si un sustentante comete un error leve, probablemente se le quite 5-10% al 100% del valor total de la sección; si comete un error importante, quizá se le reduzca en un 25-30%; y si comete un error fatal puede perder un 50% o la sección entera. Si cometiera tantos errores que la suma de ellos sobrepasara el 100% del valor de la sección, se le anula la sección pero no se le dan puntajes negativos. Se ha pensado además, que el programa de calificación no sólo señale los errores cometidos en la sección sino que, de ser posible, también señale las razones por las que se consideran errores.

Por lo que se refiere al tratamiento de las opciones que el sustentante debe revelar, nuestro punto de vista es el siguiente: existe cierto número de opciones necesarias para determinada estrategia que sirven para confirmar o desechar una hipótesis; cuando éstas se han utilizado, la búsqueda de más opciones se torna redundante. Por ejemplo, si se supone una infección viral tipo gripe y se revela la opción "cuerpo cortado" y el paciente no lo tiene, la opción "calentura" y tampoco la tiene, la opción "tos" y tampoco la tiene, la hipótesis inicial puede ser desechada volviéndose inútil la búsqueda de más síntomas y signos al respecto como pudiera ser "si fluye la nariz", u otros problemas nasales y respiratorios enfocados a

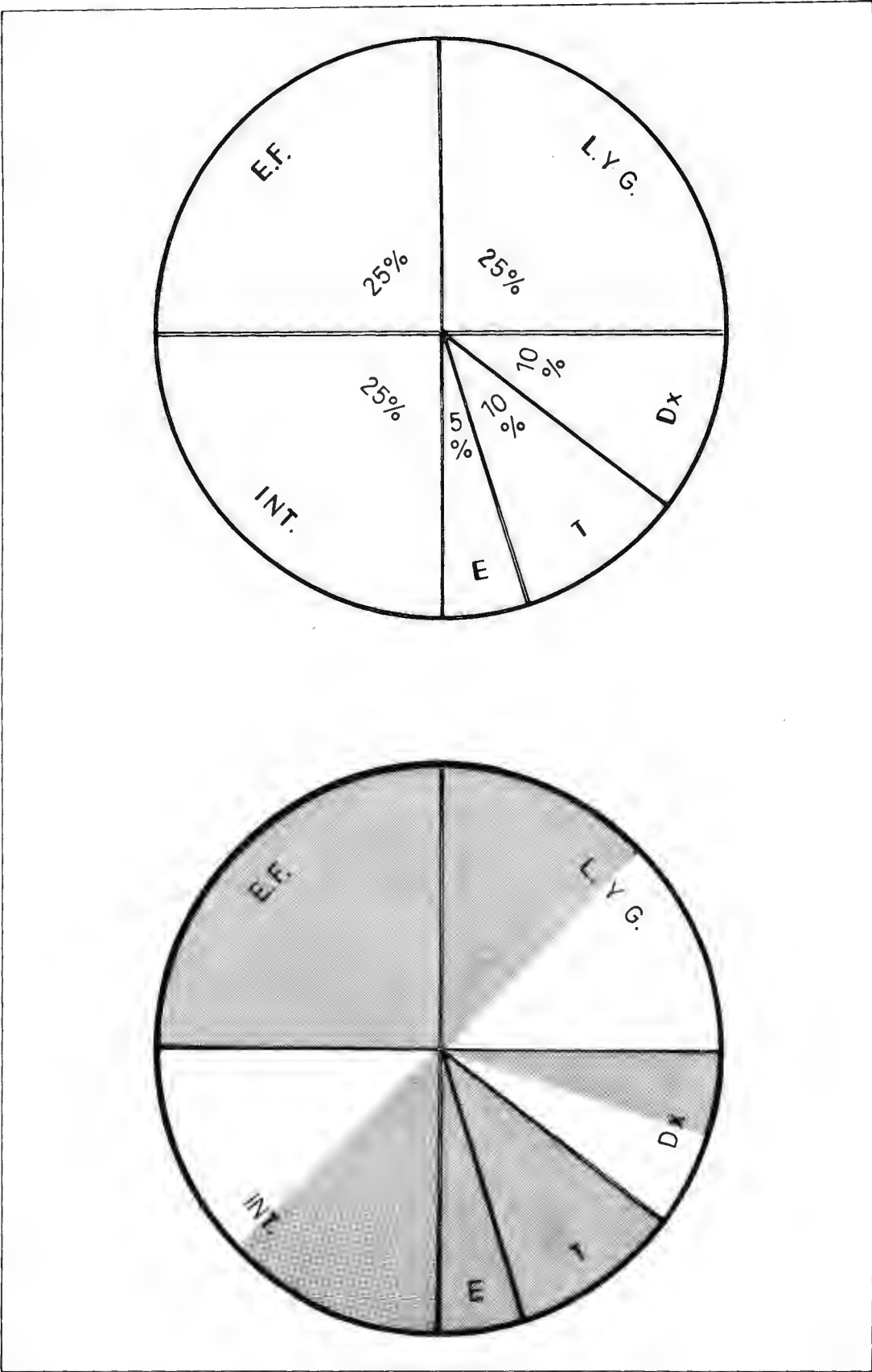


Figura 10. Eficiencia global

comprobar si el paciente tiene gripe. Si para úlcera gástrica, las “agruras”, el “dolor en la boca del estómago”, y “el horario” en el que aparecen esos síntomas y signos es para un especialista suficiente información, y decide pasar a explorarlo o a solicitar exámenes de laboratorio, sería injusto pedirle que solicite mayor información que le sería redundante o inútil.

El asunto se complica si suponemos que para el diagnóstico X pueden existir, en la sección de interrogatorio, las opciones a, b, c, d, y e, que el sustentante deberá revelar, y éste sólo revela tres (a, b, c), (a, b, d), (a, b, e), (a, c, d), (a, c, e), (b, c, d), (b, c, e), (b, d, e), (c, d, e). Cualquiera de estas combinaciones podrá ser suficiente para darle información que le permita aceptar o desechar el diagnóstico, de tal forma que las dos opciones sobrantes para cada tríada representan información redundante. En estos casos se podrá decidir si el sustentante buscó la estrategia adecuada en el interrogatorio, independientemente de la combinación, con un 60% de las opciones que debe revelar. En otros casos puede ser que la tríada (b, c, d) o (c, d, e), no sea suficiente por sí misma, y entonces se pondrán como calificaciones condicionales, manteniendo el mismo porcentaje de aceptación.

Para fines prácticos, se han escogido cinco niveles de estrategias: la óptima, con un valor de 100%; la buena, 90%; la adecuada, 70%; la regular, 50% y la mala, 30%. Sobre ellos, se define el puntaje asignado a la sección, el cual puede reducirse hasta llegar a 0% en la medida que se cometan errores. De esta manera, estrategia y errores, están correlacionados.

7. Análisis del simulador. Para terminar, expondremos un resumen de nuestro trabajo con respecto a lo que consideramos un capítulo importante en la construcción de los simuladores: el análisis estadístico del simulador. Lo consideramos importante porque constituye una manera de retroalimentar al constructor de simuladores, señá-

lándole las partes del simulador que, para una población dada, funcionaron de manera distinta a la que se suponía. El análisis que presentamos a continuación no está completo, pero sí orientado hacia lo que creemos será el análisis definitivo. En principio, lo hemos dividido en dos partes: Análisis de frecuencias y análisis de correlaciones.

A. Análisis de frecuencias. El análisis de frecuencias es ilustrativo con respecto a las secciones de diagnóstico, ya que en ellas existen diagnósticos “probables” y diagnósticos “distractores” o “ruidosos”. El constructor de simuladores pone uno o dos diagnósticos muy cercanos al diagnóstico correcto y espera, como consecuencia, frecuencias cercanas al diagnóstico correcto después de la situación inicial; por ejemplo, en el simulador de evaluación que anexamos se supone que el diagnóstico “úlcera gástrica” debe elegirse con la misma frecuencia que el de “hernia hiatal” después de la situación inicial, ya que es la probabilidad más compatible con el cuadro clínico; pero si para la población evaluada la frecuencia es muy dispareja, supondremos que la situación inicial es tendenciosa. Por otra parte, la frecuencia entre hernia hiatal y úlcera gástrica debe ser muy dispareja después de la sección de laboratorio y gabinete, pues suponemos que en ella debe confirmarse el diagnóstico. Sin embargo, la situación no deberá suceder antes de esta sección, pues la haría redundante. Existe una prueba estadística no paramétrica que ante una banda de confianza arbitraria nos puede indicar si las diferencias de frecuencias son significativas o no y, de esta forma, ilustrarnos sobre cuándo se dio el diagnóstico diferencial y, por lo tanto, cómo están funcionando las secciones según las suposiciones del constructor. Por lo que respecta a los diagnósticos “ruidosos” o “distractores”, se esperan ciertas frecuencias de algunos de ellos, o ninguna o muy escasas de otros. El análisis también puede hacerse relacionando las diferencias entre lo esperado y lo observado con el funcionamiento de las

secciones de información.

B. Análisis de correlaciones. Si correlacionamos cada opción, o varias opciones en caso de estar condicionadas, con la eficiencia de la sección y la eficiencia global, esperamos que ésta tenga valores positivos para las opciones forzosas, importantes y de rutina, y que los valores sean cero o negativos para las opciones o conjunto de opciones que consideramos como errores, sean leves, graves o fatales, para aquellos sustentantes que han seguido una determinada estrategia. Si esto no sucede, es posible que las opciones confundan a los sustentantes o los sustentantes estén confundidos. Lo anterior permitirá que el constructor de simuladores analice dichas opciones para averiguar el por qué están confundiendo a los sustentantes. Si correlacionamos la eficiencia obtenida en cada sección con la eficiencia global, esperamos que el valor de correlación sea positivo; pero si el valor de la correlación es cero, podremos saber que algo está pasando en relación con el grado de dificultad o confusión de una sección completa, puesto que si ésta es muy fácil, podrán resolverla adecuadamente tanto los alumnos malos como los buenos. Del análisis detallado de las correlaciones podremos obtener información del funcionamiento de cada una de las opciones y de las secciones completas. Esto aún no está completamente desarrollado, nos hace falta analizar la consistencia de la escala y la confiabilidad interna del simulador, para lo cual tenemos algunas hipótesis que esperamos sistematizar pronto.

Resumen

1. Se consideran varios problemas que resultan de aplicar los simuladores de enseñanza a las condiciones de evaluación.
2. Se describen las modificaciones que hay que hacer a los simuladores clínicos escritos de enseñanza para que resulten instrumento útil en situación de evaluación.
3. Se justifica la adición de una nueva sección que se ha llamado de diagnóstico.

4. Se discute la forma de calificar los casos clínicos simulados para evaluación.

5. Se hace una breve descripción de lo que se considera sería el análisis del simulador, algo análogo a lo que es el análisis de los exámenes de opción múltiple.

6. La estructura de los simuladores clínicos escritos que actualmente existen no es inadecuada, pero su confiabilidad como instrumentos de evaluación es pobre.

Bibliografía

McGuire, C. H.; Solomon, L.M.; Bashook, P.G.: Handbook of Written Simulations: Their Construction and Analysis. Center of educational Development. University of Illinois. College of Medicine. Chicago, Illinois. 1972. (Pre-publication Copy).

Instrucciones para el manejo del caso clínico simulado que se presenta en el Anexo 1

Esta es la simulación escrita de un problema clínico, y está construida de manera que semeje lo más estrechamente posible la situación real. Consta de varias secciones que corresponden a diferentes posibles soluciones. Todas las secciones incluidas le proporcionan acceso a diferentes vías que podrá seguir para la resolución del problema. En cada sección, existe una lista de opciones de las que podrá seleccionar la(s) que considere significativa(s). En algunas secciones tendrá que escoger SOLO UNA opción, pero en otras podrá escoger TANTAS OPCIONES COMO CONSIDERE NECESARIAS. Sin embargo, el hecho de que se ofrezca una opción o se imprima una sección completa no implica que necesariamente se deba incluir en la solución óptima del problema, PUEDE HABER MAS SECCIONES QUE LAS UTILES. Puesto que las secciones están colocadas al azar, es necesario que proceda de sección en sección PRECISAMENTE COMO SE LE INDIQUE en las instrucciones que se le den en el desarrollo del problema mismo, y que ponga mucha atención para que sólo seleccione el número de opciones permitidas.