

*Arceo, María Dolores*

## **Desarrollo del razonamiento clínico con scripts en la residencia de medicina familiar.**

Maestría en Educación para Profesionales de la Salud

*Tesis 2015*

**Cita sugerida:** Vancouver: Arceo MD. Desarrollo del razonamiento clínico con scripts en la residencia de Medicina Familiar. [tesis de maestría][Internet].[Buenos Aires]: Instituto Universitario del Hospital Italiano;2015.[Citado AAAA MM DD]. iii; 32 p. Disponible en: URL: <http://trovare.hospitalitaliano.org.ar/descargas/tesisytr/20150717113631/arceo-maria-dolores.pdf>





# Desarrollo del razonamiento clínico con Scripts en la residencia de Medicina Familiar

Por

María Dolores Arceo

Tutor: Dr. Eduardo Durante

Dirigida por Dr. Carlos Brailovsky

Presentado en cumplimiento de los requisitos para la obtención del Título de  
**Magister en Educación para los Profesionales de la Salud**

ante el

Instituto Universitario Escuela de Medicina Hospital Italiano de  
Buenos Aires

Buenos Aires, 6 de abril del 2015

## **Agradecimientos:**

Al Prof. Dr. Carlos Brailovsky, Co-Director de la Maestría en Educación para Profesionales de la Salud del Instituto Universitario Escuela de Medicina del HIBA, y al Prof. Dr. Eduardo Durante, subjefe Servicio de Medicina Familiar HIBA, Director de la Maestría en Educación para Profesionales de la Salud del Instituto Universitario Escuela de Medicina del HIBA, por sus aportes, paciencia y acompañamiento incondicional en cada uno de los pasos de esta tesis .

A mis compañeros del Servicio de Medicina Familiar y Comunitaria del HIBA, que colaboraron (como expertos) en la redacción de los casos y contestando el examen.

A los encargados de las residencias y residentes de los servicios de Medicina Familiar del Hospital Italiano de Buenos Aires, CEMIC, OSPeCon (sedes Buenos Aires, Rosario y Tucumán), Hospital Ramos Mejía, Hospital Piñero, Hospital Argerich, Hospital Municipal Bernardo Houssay de Vicente López y al Departamento de Medicina Familiar de la Carrera de especialización en Medicina Familiar del Hospital Nacional Clínicas (Facultad de Ciencias Médicas, Universidad Nacional de Córdoba) por haber colaborado contestando el examen.

A Micaela Gago, a quien extraño, por haber sido quien me introdujo en el mundo de la docencia y con quien comencé a pensar esta tesis.

## **Resumen**

El razonamiento clínico se lleva a cabo a partir de dos procesos: uno **analítico** y otro **no analítico** o basado en reconocimiento de patrones. Numerosos estudios demostraron que al hacer un diagnóstico, los expertos organizan su conocimiento en forma de redes conceptuales (scripts) que les permiten priorizar e interpretar de manera más eficiente los hallazgos clínicos (razonamiento no analítico). Por el contrario, los estudiantes razonan generando sucesivas hipótesis y contrastando las mismas con sus hallazgos, de manera individual y secuencial, siendo incapaces de organizar su conocimiento en estructuras más complejas (razonamiento analítico). Uno de los instrumentos de evaluación que fueron diseñados con el objetivo de medir el razonamiento clínico es el examen de comparación de scripts (para futuras referencias nos referiremos a él por su sigla en inglés, SCT: Script Concordance Test).

Hemos diseñado y validado un SCT para Medicina Familiar, y lo aplicamos en los primeros estadios de la formación de post grado (residencia) para ayudar a evaluar algunas dimensiones del proceso de desarrollo del razonamiento clínico en los residentes de medicina familiar.

## **Resultados:**

Se construyó y validó el SCT para Medicina familiar.

Se aplicó el examen a 57 residentes y evaluó la consistencia interna con el coeficiente Alpha de Cronbach, cuyo resultado fue de 0,614.

Se calculó, si existían diferencias entre los residentes de años inferiores, y los residentes de 4to año. Considerando las varianzas iguales, se obtuvo una diferencia estadísticamente significativa entre los R4 y los R123, de 4,18 (DS 1,97) y un p-valor de 0,038. Considerando las varianzas diferentes y utilizando el test de Welch, el valor p fue 0.0038.

## **Conclusiones**

Los residentes de medicina familiar de años superiores utilizan scripts en mayor medida que los de años inferiores.

**Palabras claves:** i- scripts, ii- scripts concordance test, iii- razonamiento clínico, iv- educación médica, v- evaluación

1- Introducción .....	2
2- Fundamentación .....	4
2.1- Razonamiento clínico (RC).....	4
2.2- Evaluación del RC .....	6
2.3- Test de Concordancia de Script .....	8
2.3.1- Puntuación y definición de puntaje de aprobación .....	11
2.3.2- Características de examen.....	12
3- Planteo del problema.....	15
4- Pregunta .....	16
5- Objetivos de la investigación .....	16
5.1- Objetivo general.....	16
5.2- Objetivos específicos .....	16
6- Materiales y métodos .....	16
6.1- Tipo y diseño general del estudio .....	16
6.2- Ámbito del estudio	16
7- Diseño de SCT .....	16
7.1- Desarrollo de la tabla de especificaciones .....	16
7.2- Selección y capacitación de expertos para el desarrollo de casos .....	17
7.3- Validación del examen.....	17
7.4- Análisis de la validación del examen.....	18
7.5- Selección de residencias .....	18
7.6- Aplicación del SCT a residentes .....	18
7.7- Análisis de los resultados del SCT en residentes:.....	19
8- Resultados .....	19
8.1- Resultados de la validación de examen .....	19
8.2- Resultados del SCT en residentes .....	20
9- Discusión.....	22
10- Conclusiones .....	24
11- Anexos .....	25
11.1- Viñetas .....	25
11.2- Instructivo para la realización del SCT.....	26
11.3- Carta enviada a residentes.....	29
12- Bibliografía .....	30

## 1- Introducción

Numerosos estudios demostraron que la forma de evaluar ejerce una influencia sustancial en la forma de aprendizaje (Durante, 2005), incluso Fredericksen (1984) afirmó que “el método más poderoso para cambiar la manera de estudiar de los estudiantes y la manera en que los profesores enseñan, es cambiar la manera de evaluar” (citado por Gormaz & Brailovsky, 2012).

Por lo tanto, si uno de los objetivos de la formación de un profesional competente es que desarrolle un correcto razonamiento clínico (RC), entendiendo al mismo como el conjunto de procesos mentales mediante los cuales el profesional plantea un diagnóstico y decide los planes de manejo y el pronóstico en un caso particular, es necesario incluirlo en la evaluación desde los inicios de la formación (Durante, 2005).

Por otro lado, Epstein (2007) ha definido tres razones principales que justifican la evaluación del RC del estudiante: 1) Estimular la motivación y optimizar el desarrollo de las capacidades del estudiante. 2) Brindar protección al paciente, identificando al estudiante incompetente y 3) Proveer una base confiable de selección de los candidatos más adecuados para profundizar su entrenamiento en un área determinada. (Durante, 2005; Epstein, 2007).

Si bien en el campo de la medicina se utilizan numerosos exámenes que, escalando la pirámide Miller, evalúan diferentes componentes de la competencia clínica (como por ejemplo el examen de opción múltiple, exámenes escritos de respuestas cortas, exámenes orales, exámenes orales estructurados, Examen Clínico Objetivo Estructurado (OSCE), Mini Cex estos exámenes evalúan conocimiento, el “sabe cómo” o “demuestra cómo” y en el mejor de los casos “el hacer” (Durante, 2006), pero ninguno de estos exámenes evalúa el razonamiento clínico excepto que estén específicamente diseñados para esto.

Hasta hace unos años, existían algunos instrumentos para evaluar el RC como el “Método Think aloud” y el “Método Stimulated recall”, que son exámenes orales muy difíciles de aplicar y tienen una validez y confiabilidad cuestionables (Gormaz & Brailovsky, 2012). A finales de la década del 90, Charlin, Tardiff y Boshuizen (2000) diseñan el Script Concordance Test (SCT), que es un instrumento que ha demostrado capacidad para medir, al

menos en parte, el RC. Este examen fue utilizado en diferentes especialidades dentro de la medicina como oncología, diagnóstico por imágenes, ginecología, urología, emergencias, medición del razonamiento intraoperatorio, entre otros. Sin embargo, pocas veces fue utilizado para medir el razonamiento en médicos de atención primaria (Lambert et al., 2009; Sibert et al., 2006; Boshuizen & Schmidt, 1990; Charlin, et al.,1998; Meterissian, Zabolotny & Gagnon, 2007; Monnier, Bédard & Gagnon, 2011).

Esta tesis tiene como objetivo desarrollar y validar el SCT para medicina familiar y aplicarlo en los primeros estadios de la formación de post grado (residencia) para evaluar el desarrollo del razonamiento clínico en los residentes de medicina familiar.



## 2- Fundamentación

La competencia profesional se define como “la capacidad de un profesional de utilizar su buen juicio así como también los conocimientos, habilidades y actitudes asociadas a la profesión para solucionar los problemas complejos que se presentan en el campo de su actividad profesional” (Brailovsky, 2001; Kane, 1992).

Esta definición enumera adecuadamente los diversos elementos que se deberían incluir a la hora de evaluar la competencia e implica que no alcanza solamente con que un profesional conozca las bases conceptuales acerca de cómo se debe realizar un procedimiento y lo realice con perfección, sino que también es necesario que sepa cuándo y cómo indicarlo, pueda comunicarse con los pacientes en forma efectiva, tome en cuenta las necesidades del paciente y el contexto, además interprete el resultado del estudio y decida cuales son las conductas necesarias, siempre dentro de un marco de incertidumbre. Por lo tanto, y según esta definición, ser un profesional de la salud competente incluye contar con diversas subcompetencias tales como habilidades reflexivas, organizacionales y técnicas. Sin embargo, la competencia central y más importante del médico consiste en el RC, que posibilita la toma de decisiones que conducen a la resolución de situaciones complejas en los diferentes contextos en los que deba actuar (Brailovsky, 2001).

### 2.1- Razonamiento clínico (RC)

Se entiende por RC al conjunto de procesos mentales mediante los cuales el médico plantea un diagnóstico y decide los planes de manejo y el pronóstico en un caso en **particular** (Brailovsky, 2001).

El RC se destaca por su función de guía de toda acción médica (el diagnóstico, el estudio de los casos clínicos, la elección de la conducta terapéutica y la estimación del pronóstico de un paciente) (Gormaz & Brailovsky, 2012).

El RC se lleva a cabo, básicamente, a partir de dos procesos: uno **analítico** y otro **no analítico** o basado en reconocimiento de patrones (Schmidt, Norman & Boshuizen, 1990).

Numerosos estudios demostraron que, al hacer un diagnóstico, los **estudiantes** razonan generando sucesivas hipótesis y contrastándolas con sus hallazgos, de manera individual y secuencial, sin poder organizar su conocimiento en estructuras más complejas. En general, los estudiantes cuentan con escaso conocimiento clínico, debido a que estuvieron con poca supervisión y expuestos a un número reducido de pacientes o no estuvieron expuestos a una práctica reflexiva. En general, su conocimiento es más del tipo biomédico y está organizado en largas listas de datos, sin conexiones entre ellos. Los estudiantes analizan cada una de estas “partes” por separado y las van confirmando o rechazando hasta arribar al diagnóstico (Bowen, 2006; Schmidt, Norman & Boshuizen, 1990; Norman, 2005).

Por el contrario, los expertos, tienen sus conocimientos organizados en forma de redes conceptuales que les permiten priorizar e interpretar de manera más eficiente los hallazgos clínicos, a través de un proceso no analítico o de reconocimiento de patrones (Schmidt, Norman & Boshuizen, 1990). Cuando un experto se enfrenta a un paciente, rápidamente relaciona el caso con patrones clínicos similares, incluso con casos particulares, a los que se enfrentó con anterioridad en su práctica clínica. El razonamiento del experto frente al enfermo pasa por una secuencia rápida de eventos dada por la percepción de los elementos claves pertinentes, la activación automática casi inconsciente de las redes de conocimiento que contienen estos elementos clínicos significativos ("*script triggering*"), el rápido acceso a diferentes hipótesis, la verificación de síntomas y signos que confirman o descartan el diagnóstico y el uso racional de sus conocimientos para guiar de manera apropiada las investigaciones y el tratamiento ("*script processing*") (Charlin & van der Vleuten, 2004).

Estas estructuras cognitivas que los expertos usan en su RC se denominan *illness scripts* (Charlin & van der Vleuten, 2004; Charlin, Tardif & Boshuizen, 2000; Schmidt & Boshuizen, 1990). De esta forma, los *illness scripts* funcionan como organizadores de la estructura cognitiva de cada profesional y permiten una eficiente aplicación de los conocimientos y un rápido arribo a soluciones de manera competente en la mayoría de las situaciones clínicas (Gormaz & Brailovsky, 2012; Bowen, 2006; Schmidt, Norman & Boshuizen, 1990; Brailovsky, 2001).

La importancia del contenido y del contexto (ambiente y circunstancias) así como de los factores personales (valores y actitudes) en la creación de estas redes o scripts hacen que cada una de ellas sea única para cada individuo. Por lo tanto, los scripts son únicos para cada

profesional y no son transferibles de docentes a alumnos, ya que no sólo incluyen información clínica relevante sino también información percibida inconscientemente de cada ambiente y circunstancias (Gormaz & Brailovsky, 2012).

Schmidt y colaboradores (Schmidt, Norman & Boshuizen, 1990), han definido cuatro estadios diferentes del desarrollo e integración de estos procesos cognitivos: los estudiantes pasan desde la creación de un mapa conceptual causal simple de signos y síntomas (estadio 1), a ser capaces de sintetizar el cuadro clínico del paciente en una o dos frases (estadio 2). En el estadio 3, ya aparecen los scripts de enfermedades que organizan el conocimiento en complejas redes conceptuales de patrones clínicos y tienen una estructura que incluye factores predisponentes, fisiopatología y consecuencias clínicas. Mientras que los estadios 1 y 2 se caracterizan por el razonamiento analítico, en el estadio 3 predomina el no analítico, y el estadio 4 representa el almacenamiento de estos scripts como scripts ejemplares (Schmidt, Norman & Boshuizen, 1990; Brailovsky, 2001). Los *scripts* comienzan a aparecer cuando los estudiantes se enfrentan con situaciones clínicas reales, y se desarrollan y perfeccionan durante toda la vida profesional (Charlin & van der Vleuten, 2004; Brailovsky, 2001). Los alumnos en cualquier estadio usan ambas formas de razonamiento en la práctica clínica: los más avanzados tienen la ventaja de contar con una colección de scripts más rica y descansan más sobre este mecanismo que sobre el analítico, excepto cuando se enfrentan a un caso complejo sobre el cual no tienen experiencia (Brailovsky, 2001).

## **2.2- Evaluación del RC**

Numerosas investigaciones indican que la forma de evaluar el conocimiento ejerce una influencia sustancial en el estilo de aprendizaje de los alumnos (Schmidt, Norman & Boshuizen, 1990). Por lo tanto, si consideramos que el razonamiento clínico es una competencia central para alcanzar la competencia profesional, su evaluación representa un enorme desafío y debe ser considerada como un elemento fundamental para orientar el proceso de aprendizaje, de forma tal que permita alcanzar los objetivos planteados.

Debido a la complejidad de los procesos implicados en el RC es muy difícil lograr que un instrumento de evaluación mida exclusivamente RC. Como se mencionó previamente, algunos de los métodos desarrollados para medir el RC son el “Método Think aloud” y el

“Método Stimulated recall”, que son exámenes orales, y el SCT, que se realiza en forma escrita (Gormaz & Brailovsky, 2012).

El “Método *Think aloud*” consiste en solicitar que el examinado relate el proceso mental que está siguiendo durante la resolución de un caso clínico (Norman, 2005; Boshuizen & Schmidt, 1992). Si bien éste es un examen oral estandarizado (lo que ayuda a mejorar su confiabilidad) y ha sido propuesto para medir el RC en estudiantes de medicina (Boshuizen & Schmidt, 1992; Banning, 2008), al ser un examen oral su implementación práctica presenta ciertos problemas (Memon, Joughin & Memon, 2010). En un examen oral habitualmente se interrogan pocos casos clínicos o temas (incluso a veces sólo uno, cuando se utilizan pacientes reales), hecho que implica un sesgo de muestreo que afecta la validez de contenido del mismo (Memon, Joughin & Memon, 2010).

Por otro lado, como todos los exámenes orales requieren para su realización un importante número de docentes y largas jornadas de evaluación, las variables afectivas, tanto de los examinadores como de los estudiantes, cobran una significativa importancia haciendo que las condiciones no sean equivalentes para todos los examinados. Además, existen factores no controlables como la interacción de los examinadores entre sí, el cansancio acumulado de los evaluadores, o las percepciones previas (o sesgos cognitivos) que tienen los examinadores sobre el desempeño del estudiante, que pueden tener un impacto sobre la manera de evaluar (llamado efecto halo) y que provocan sesgos o errores sistemáticos. Todos estos factores influyen sobre el puntaje que asignan los examinadores (Alves de Lima, 2012).

En el caso de los estudiantes, su desempeño también está influenciado por factores no controlables, como su estado emocional, el momento en que le toca dar su examen, la empatía que se produzca con el examinador, los prejuicios sobre su desempeño clínico, las características personales del estudiante (capacidad de enfrentar el estrés en el examen oral, habilidad de expresarse en público, etc.).

Aunque, en teoría, este tipo de examen oral (“Método *Think aloud*”) mide el RC en forma eficaz, los elementos conflictivos mencionados nos hacen pensar que su confiabilidad y su validez pueden ser cuestionables (Gormaz & Brailovsky, 2012).

El otro método de estudio del RC en forma oral es el examen de casos clínicos

mediante *Stimulated recall* (recordación estimulada), que consiste en que, una vez resuelto el caso clínico, se le solicita al examinado recordar y relatar cómo razonó mientras resolvía el caso (análisis de RC *post hoc*) (Norman, 2005; Boshuizen & Schmidt, 1992). Este tipo de examen es más complejo porque involucra la memoria de corto plazo del examinado en una situación de estrés, constituyéndose como una metodología de investigación, más que de evaluación de RC (Boshuizen & Schmidt, 1992; Gormaz & Brailovsky 2012).

### **2.3- Test de Concordancia de Script**

El Test de Concordancia de Script (SCT) es uno de los instrumentos más nuevos desarrollados para evaluar dimensiones del RC y particularmente aquellas que están en relación con la toma de decisiones clínicas (Gormaz & Brailovsky, 2012). Se trata de un examen escrito (Charlin, Tardif & Boshuizen, 2000; Charlin et al., 2000; Charlin et al., 1998) cuyo objetivo es medir la organización del conocimiento clínico en redes conceptuales que se revelan, básicamente, en la acción y en la toma de decisiones del médico, o sea, la organización de los conocimientos en scripts. El examen está compuesto por una serie de viñetas clínicas, que sitúan al examinado en un contexto clínico de incertidumbre, dentro del cual debe interpretar datos para tomar decisiones.

El examen consta de varios problemas, presentados en viñetas clínicas cortas, cada una de las cuales es seguida por una serie de ítems. Cada uno de estos ítems aporta nueva información clínica al caso, y contiene datos útiles que, a través del razonamiento, deben promover la reflexión y permitir la resolución del problema presentado, de manera que el examinado debe responder acerca del impacto que considera que puede tener esta nueva información sobre la hipótesis o la decisión propuesta. La respuesta va a depender de la experiencia de quien está respondiendo, el contexto en que esté planteada la situación clínica y el peso que le dé cada uno de los examinados a los distintos datos. En este tipo de examen no existe una única respuesta correcta, la puntuación que recibe cada ítem es asignada a partir de la frecuencia de las respuestas dadas por expertos y su valor depende del acuerdo entre ellos.

Para la construcción de un SCT es necesario seguir una serie de pasos que pueden resumirse de la siguiente manera:

- 1) Construcción de una tabla de especificaciones para seleccionar metodológicamente los contenidos a evaluar. Esta tabla se construye considerando

todas las áreas del campo que se quiere evaluar y el peso que se le dará a cada una en la evaluación final, con el fin de garantizar un muestreo y una validez de contenido adecuados para el examen (Fournier, Demeester & Charlin, 2008; Dory et al., 2012).

2) Una vez construida la tabla de especificaciones, se debe convocar a especialistas en el contenido o disciplina para construir los casos (las viñetas clínicas), previo entrenamiento en su elaboración. El examen de comparación de scripts debe contar con, al menos, veinte viñetas clínicas, cada una de las cuales debería incluir de tres a cinco ítems (introducción de nueva información) para alcanzar una confiabilidad adecuada (igual o mayor de 0.80) (van der Vleuten, et al., 2010; Charlin, Tardif. & Boshuizen, 2000). La necesidad de este número de viñetas y de ítems obedece al hecho de que es necesario asegurar un muestreo adecuado de las dimensiones a evaluar y se debe tener en cuenta la especificidad de caso (Durante, 2006), lo que implica que la resolución adecuada de un caso es específico para ese único caso y no predice o predice mal el desempeño en otros, aún muy similares. Esto obedece al hecho de que la adquisición de competencias es específica para cada contenido o contexto. Como consecuencia, es necesario aumentar el número de casos en las evaluaciones para asegurar una adecuada confiabilidad.

Para la construcción de los casos clínicos, se le solicita a un grupo de docentes, expertos en el área, que identifique distintas situaciones clínicas para las áreas definidas en la tabla de especificaciones, y que para cada una de ellas defina los siguientes aspectos:

- a) Las hipótesis relevantes, las estrategias de investigación o las opciones de tratamiento que se plantean.
- b) Las preguntas que se realizan ante dichas situaciones, los exámenes físicos y pruebas que piden para resolver el problema.
- c) La información clínica, positiva o negativa, que buscan en sus investigaciones.

#### Construcción de las viñetas clínicas

Una vez definidos estos puntos, se procede a la construcción de las viñetas.

Las viñetas clínicas deben describir un caso clínico o una situación clínica auténtica, representativa y problemática, de forma breve, aportando información básica, poco diferenciada, que permita varias posibilidades de diagnóstico o de toma de decisiones, pero que al mismo tiempo contenga casi todos los elementos necesarios para arribar a una respuesta congruente, puesto que a partir de esa viñeta se deben poder proponer diferentes hipótesis.

Es importante que, conservando la ambigüedad de las situaciones clínicas reales, la viñeta sea capaz de ubicar al médico en un contexto clínico similar al de su práctica real y que sea capaz de movilizar (“gatillar”) su razonamiento. Es decir, de activar sus redes cognitivas.

En aquellos casos en los que el paciente tiene un solo motivo de consulta sin ninguna otra manifestación clínica asociada, la información del enunciado puede completarse con información sobre antecedentes personales o familiares, medicación, etc. En todos los casos es importante dejar en claro, en el enunciado, en qué contexto está siendo atendido al paciente.

Las viñetas clínicas pueden ser diseñadas para evaluar el RC sobre diagnóstico, estudios complementarios o tratamiento (Fournier, Demeester, Charlin, 2008).

### Construcción de los ítems

Una vez construida la viñeta clínica, se procede a la construcción de los ítems. Cada **ítem** consta básicamente de tres partes:

- a) La hipótesis diagnóstica, la acción de investigación o la opción de tratamiento que es relevante para la situación.
- b) La presentación de nueva información (por ejemplo un signo, condición, estudio de imagen o resultado de un examen de laboratorio) que podría tener un efecto sobre la hipótesis diagnóstica, la decisión de investigación o la opción de tratamiento.

c) Una escala tipo Likert de 5 puntos que está diseñada para que el examinado marque el impacto que tiene la nueva información clínica sobre la hipótesis inicial o la decisión propuesta.

Para la construcción del examen, la elección de cada ítem debe centrarse en los elementos que son más útiles para resolver el problema clínico. Estos elementos, además deben tener un impacto positivo, neutro o negativo sobre las hipótesis diagnósticas, de investigaciones o terapéutica de modo que aseguren algún grado de dispersión entre las posibles respuestas.

Es importante que no exista relación de linealidad entre la hipótesis y el nuevo dato presentado, ya que debe ser necesario un proceso de reflexión y no la simple aplicación de conocimientos para la elección de la respuesta. Cada ítem dentro de las viñetas debe ser independiente de los otros. Es decir que se debe evitar que dos ítems sucesivos presenten opciones similares o que puedan estar relacionadas entre sí, ya que esto puede dar lugar a confusión y llevar a evaluarlas como si estuviesen encadenadas o tuviesen carácter acumulativo. Se debe tener presente que la finalidad de cada ítem es determinar los efectos o impacto de una información clínica aislada sobre la hipótesis propuesta y no efectos aditivos de una serie de elementos clínicos.

### **2.3.1- Puntuación y definición de puntaje de aprobación**

Luego de su construcción, cada viñeta clínica será sometida a la evaluación de un grupo de expertos (entre 10 y 20), para descartar términos confusos, validar el examen (verificando la relevancia de contenido, validez de construcción y confiabilidad) y construir la tabla de puntuación (Dory et al., 2012).

La tabla de puntuación se construye asumiendo que todas las respuestas dadas por los expertos son aceptadas como válidas. Así, a la respuesta marcada por el mayor número de expertos le corresponde la máxima puntuación y el resto de las respuestas tienen puntajes proporcionales. Por ejemplo, si en un determinado ítem 6 de 10 expertos han seleccionado la opción **C** como correcta, 3 la **B** y 1 la **D**, la opción **C** obtendrá la puntuación máxima de 1 punto ( $6/10 = 1$  punto), mientras que a la respuesta **B** se le asignarán 0,5 puntos ( $3/6$ ) y a la **D** 0,17 puntos ( $1/6$ ). Las opciones que no han sido elegidas por ningún experto tendrán una



puntuación igual a cero (Dory et al., 2012).

**Tabla 1:** Ejemplo de puntuación de SCT: Panel de 10 expertos que responden un ítem, considerando que seis expertos seleccionan la respuesta C, tres expertos la respuesta B y un experto de respuesta D.

Respuesta	A	B	C	D	E
Nº de expertos que seleccionan la opción	0	3	6	1	0
Puntuación	0	3/ 10	6/ 10	1/ 10	0
Transformación	0	3/ 6	6/ 6	1/ 6	
Puntuación final	0	0,5	1	0,17	0

Por lo tanto, la puntuación que recibe cada ítem es asignada a partir de la frecuencia de las respuestas dadas por los expertos y su valor depende del acuerdo entre ellos. El resultado total del examen está determinado por la suma de las puntuaciones obtenidas en cada ítem. Una puntuación elevada significa mayor coincidencia con los expertos en los diferentes ítems, indicando un mayor grado de desarrollo de su razonamiento clínico (Boshuizen & Schmidt, 1992).

Por el momento, el SCT no se usa para realizar evaluaciones sumativas a partir de las cuales se tomen decisiones de promoción, ya que todavía se requieren más estudios para poder establecer un punto de corte adecuado. Se trata de un problema psicométrico complejo que deberá ser resuelto antes de poder utilizar el SCT en instancias de evaluación sumativa (Gormaz & Brailovsky, 2012).

### 2.3.2- Características de examen

Para analizar la utilidad de este tipo de examen usaremos la ecuación de utilidad propuesta por van der Vleuten (Brailovsky, 2001; van der Vleuten, 1996) que relaciona las variables más importantes a tener en cuenta al momento de decidir usar un examen:

$$U = (C(p) + V(p) + IE(p) + (p) A) \times 1/C$$

Así la U es la utilidad de un instrumento y es igual a la suma de la proporción de

*confiabilidad* (C), más la proporción de *validez* (V), más proporción del *impacto educativo* (IE), más la proporción de *aceptabilidad* (A), todo multiplicado por la inversa del costo (Durante, 2005).

Cuando nos referimos a la confiabilidad hablamos de la precisión de la medición o la reproducibilidad del puntaje medido por el instrumento, así como de la medida en que los ítems individuales de un examen se comportan de manera similar (co-varían) en el interior del mismo. Esta determinación se realiza con métodos estadísticos, utilizando el coeficiente Alfa de Cronbach, para el caso de variables continuas como el SCT. El coeficiente varía entre 0 y 1 y 1 es el valor asignado a la máxima confiabilidad del instrumento (Durante, 2005).

En cuanto a la validez, existen diferentes dimensiones que se resumen con este término, por lo que podemos hablar de diferentes tipos de validez.: 1) la validez de constructo que tiene como propósito fundamental, de acuerdo a los criterios clásicos, representar cómo los puntajes de las evaluaciones reflejan y sintetizan el concepto bajo investigación. En este caso del SCT, cómo un puntaje más alto refleja mayor capacidad de razonamiento clínico. 2) La validez de criterio que muestra si los resultados de un examen son concordantes con los resultados de otras pruebas ya validadas, o si estos resultados predicen el desempeño futuro de los alumnos. 3) La validez de contenido que pone de manifiesto el grado de “muestreo” de todos los aspectos que se pretende medir (Durante, 2005).

Hasta el momento, el SCT ha demostrado confiabilidad y validez para diferenciar entre el razonamiento de un experto y un novato. Una serie de estudios ha mostrado la capacidad del SCT para discriminar entre diferentes niveles de experiencia y su aplicabilidad en dominios tan diversos como la Cirugía, el Diagnóstico por Imágenes, la Oncología y la Medicina de Urgencias. Para evidenciar la validez de constructo del SCT se realizaron diferentes estudios (Schmidt & Boshuizen, 1990; Boshuizen & Schmidt, 1992), que mostraron que los médicos más expertos obtenían mejores puntajes en el examen que aquellos menos expertos e incluso el grado de variabilidad en la respuesta de los expertos dependía de la complejidad del caso planteado y no de la evaluación de cada ítem en particular. Por el contrario, los alumnos presentaban mayor variabilidad en cada ítem, independientemente de la complejidad del caso presentado (Charlin et al., 1998) .

Además, se ha demostrado que el SCT no presenta el denominado “efecto

intermedio”, que consiste en que médicos experimentados obtienen un puntaje igual, o incluso peor, que los médicos menos experimentados o los estudiantes en los exámenes que evalúan únicamente conocimiento teórico. En el SCT, que está dirigido a medir RC (el cual aumenta con la experiencia clínica) este "efecto intermedio" desaparece, debido a que, para resolver este tipo de examen, es necesario interpretar datos clínicos para tomar decisiones en situaciones reales y no la simple aplicación de conocimientos teóricos o de memoria (validez de constructo) (Gormaz & Brailovsky 2012; Schmidt et al., 1990; Brailovsky et al., 2001; Charlin et al., 2009).

En cuanto a la confiabilidad, por lo general un examen se considera confiable cuando el coeficiente Alfa de Cronbach es de 0,80 o superior (Tornimbeni, Perez & Olaz, 2008). En diferentes series de estudios publicadas sobre el SCT los valores oscilaron desde 0,79 hasta 0,82 (Brailovsky et al., 2001; Fournier, Demeester, Charlin, 2008).

Por último, en un estudio realizado por Brailovsky y colaboradores (Brailovsky et al., 2001), el examen de concordancia de scripts también demostró que cuando un examinado ha presentado una buena organización del conocimiento clínico en un momento particular de su formación, muestra una buena organización en posteriores mediciones de este tipo de conocimiento (Brailovsky et al., 2001).

La aceptabilidad de un examen habla acerca de cuán adecuados los evaluados y los evaluadores, consideran los métodos de evaluación propuestos. En numerosos estudios (Charlin et al., 1998; Charlin et al., 2009; Sibert et al., 2006) en los cuáles se les pregunta a los estudiantes y evaluadores acerca de la aceptabilidad, la gran mayoría acuerda en que consiste en una buena experiencia.

Si bien el SCT es un examen difícil de construir ya que requiere la colaboración de expertos entrenados, su costo es aceptable si tenemos en cuenta que es un examen escrito, que se puede realizar en 90 minutos (tiempo estimado para un examen de 20 casos con 60 ítems en total) y con el cual se pueden evaluar un gran número de examinados de manera simultánea (Dory et al., 2012).

Poco se sabe aún acerca de si este tipo de examen produce algún tipo de impacto educativo, por lo que se requieren más estudios al respecto.

### **3- Planteo del problema**

El razonamiento clínico es un componente central en el desarrollo de la competencia clínica por lo que debe ser enseñado y evaluado. El SCT es uno de los instrumentos más nuevos desarrollados y validado para la evaluación del razonamiento clínico que permite diferenciar entre el razonamiento de un experto y un novato (Gormaz & Brailovsky, 2012).

Si bien existen numerosos estudios publicados utilizando el SCT que muestran la diferencia que existe en la estructura de razonamiento entre un experto y un novato, y la transformación en el tiempo de esta estructura, ningún estudio plantea o se focaliza en el momento de la formación de un profesional en que este cambio comienza. Por otro lado, si bien el SCT ha sido utilizado en diversos campos dentro de la medicina, hay hasta el momento muy pocos estudios publicados en el ámbito de la medicina familiar. Un ejemplo, es la investigación realizada por Charlin y colaboradores, con el objetivo de evaluar el efecto que tenía incluir docentes en el panel de expertos para confeccionar la tabla de puntuaciones del SCT. El resultado de la misma mostró que los docentes puntuaban de manera más parecida a los estudiantes que aquellos que no tenían funciones docentes. Por otro lado, Brailovsky y colaboradores realizaron una investigación, en la cual se correlacionaron los resultados obtenidos con distintos métodos de evaluación aplicados a una cohorte de estudiantes de medicina antes de empezar y después de terminar su formación en Medicina Familiar. Este estudio demostró, que cuando un examinado ha presentado una buena organización del razonamiento clínico en un momento particular de su formación, es de esperar que muestre buena organización en posteriores mediciones de este tipo de razonamiento en contextos diferentes y con contenidos diferentes (Charlin et al., 2007, Brailovsky et al., 2001). Sin embargo, ninguno de éstos estudios fue realizado en contextos de habla hispana. Por lo tanto, el objetivo de este trabajo es en primer lugar desarrollar y validar el SCT en el ámbito de medicina familiar e intentar estudiar el proceso de desarrollo del razonamiento clínico en los residentes de medicina familiar. Esto permitirá no sólo adaptar los planes de estudios sino que podría de ser utilidad para establecer estrategias que permitan promover el desarrollo de este tipo de razonamiento en forma más efectiva.

## **4- Pregunta**

¿Cómo se transforma el RC de los residentes de medicina familiar durante su formación?

## **5- Objetivos de la investigación**

### **5.1- Objetivo general**

- Determinar en qué etapa de su formación, un residente de medicina familiar adquiere el razonamiento mediante scripts.

### **5.2- Objetivos específicos**

- Diseñar el test de concordancia de Scripts para medicina familiar.
- Validar el SCT.

## **6- Materiales y métodos**

### **6.1- Tipo y diseño general del estudio**

Estudio de corte transversal.

### **6.2- Ámbito del estudio**

Residencias de medicina familiar con programas de formación similares de cuatro años de duración.

## **7- Diseño de SCT**

### **7.1- Desarrollo de la tabla de especificaciones**

Previo a la construcción del SCT se desarrolló una tabla de especificaciones para definir los contenidos de la especialidad. La construcción de la tabla se realizó utilizando como referencia las utilizadas para la confección de los exámenes de certificación por el

Colegio Médico de Canadá (College of Family Physicians of Canada, 2009) y por la Federación Argentina de Medicina Familiar y General. Luego de la confección de la tabla, ésta se validó con cinco expertos locales pertenecientes al Servicio de Medicina Familiar y Comunitaria del Hospital Italiano.

En este trabajo se han considerado expertos a médicos especialistas en Medicina Familiar que han completado su formación en una residencia y tienen más de cinco años de práctica clínica continua (Charlin et al., 2009).

### **7.2- Selección y capacitación de expertos para el desarrollo de casos**

Para la confección del SCT se convocó a diez expertos, los cuales fueron capacitados para la construcción de scripts (SCT) mediante un curso de 10 horas dictado por el Dr. Brailovsky, experto acreditado en el tema. El curso se dictó de acuerdo a la metodología empleada por Charlin (Charlin et al., 1998; Charlin et al., 2009) para la construcción de SCT.

Luego del curso se le pidió a cada experto que desarrolle al menos dos viñetas clínicas con cinco ítems cada una, con el objetivo de construir un SCT de veinte viñetas clínicas con cinco ítems cada una (Schmidt, Norman & Boshuizen, 1990; Brazeau-Lamontagne et al., 2004; Sibert et al., 2006). Cada ítem debía contar con:

- 1) Hipótesis diagnóstica, acción de investigación u opción de tratamiento, que sea relevante para la situación.
- 2) Presentación de nueva información (por ejemplo un signo, condición, estudio de imagen o resultado de test de laboratorio) que podría tener un efecto sobre la hipótesis diagnóstica, la acción de investigación o la opción de tratamiento.

### **7.3- Validación del examen**

Se obtuvieron 27 viñetas clínicas con las cuales se construyó el SCT inicial. Este examen fue sometido a la evaluación de un grupo de diez expertos, diferente al que intervino en la construcción del examen, para descartar términos confusos, validarlo (verificando la validez aparente y relevancia de contenido, validez de construcción y confiabilidad) y asignar los valores, según el grado de coincidencia en la escala de Likert.

#### **7.4- Análisis de la validación del examen**

Luego de que los 10 expertos completaran el SCT, se calculó su consistencia interna a través del coeficiente Alpha de Cronbach.

Se midió la correlación ítem/ ítem total con el fin de verificar el poder discriminante de cada ítem.y la presencia de ítems (casos) con correlaciones negativas.

Para los análisis estadísticos se utilizaron dos programas: SPPS IBM 20 y JMP 10 (SAS Corporation).

#### **7.5- Selección de residencias**

Se seleccionaron residencias de cuatro años de duración con un programa de formación similar.

Luego, se contactó a los responsables de la formación en cada residencia y una vez que ellos dieron su acuerdo para la participación en el trabajo, se le envió una carta a cada residente invitándolo a participar, explicando en qué consistía el trabajo, aclarando que su participación sería anónima y que los resultados del examen se compararían entre grupos y no de manera individual. Además, se aclaró que los resultados no serían vinculables a su desempeño en su lugar de residencia (anexo 3).

Una vez obtenido el acuerdo de los participantes, se acordó una reunión para explicar en qué consistía el examen y se realizó una prueba piloto para que se familiarizaran con el instrumento de evaluación. Al finalizar el encuentro, se les solicitó que resolvieran el SCT.

Para aquellos casos en que las residencias estuvieran localizadas en lugares poco accesibles para el grupo investigador, primero se les envió (vía e-Mail) un instructivo explicando el método de evaluación y el procedimiento a seguir (anexo 2), y luego se envió el examen a los responsables de la residencia para que lo aplicaran en su grupo de residentes.

#### **7.6- Aplicación del SCT a residentes**

Participaron 10 residencias de Medicina Familiar de diferentes regiones del país: Capital Federal, Tucumán, Córdoba y Rosario. El criterio de inclusión fue haber estado

cursando la residencia en el momento que se realizó esta investigación. Respondieron el examen 57 residentes, 20 de primer año, 17 de segundo, 13 de tercero y 7 de cuarto año.

### **7.7- Análisis de los resultados del SCT en residentes:**

Una vez que los residentes contestaron el examen, se calculó la consistencia interna del SCT con el coeficiente Alpha de Cronbach.

Se midió la correlación Item/ ítem total con el fin de verificar si había presencia de ítems (casos) con correlaciones negativas, es decir que no agregan elementos a la consistencia del examen, ni a la discriminación global del mismo.

En una primera etapa se efectuaron estadísticas descriptivas y se realizó un Análisis de varianza (ANOVA) a un factor para comparar los 4 grupos (por años) de residentes entre sí.

## **8- Resultados**

### **8.1- Resultados de la validación de examen**

Luego de que los 10 expertos contestaron el SCT, se construyó la tabla de puntuación. Los expertos no sugirieron eliminar ningún ítem, ni por razones de construcción, ni de contenido.

Se calculó el coeficiente Alpha de Cronbach global con el SCT de 27 viñetas cuyo resultado fue 0,79. Al calcular la correlación ítem/ ítem total para todos los casos, se encontraron índices negativos para los casos 13, 14, 17 y 21. En la tabla 2 se muestra el impacto sobre la confiabilidad global cuando cada uno de los casos o la totalidad de ellos eran eliminados. Finalmente se decidió eliminar estos 4 casos, obteniendo un examen con 23 casos, 111 ítems (dos de los casos tenían 3 ítems ) y un coeficiente de confiabilidad de 0,86.

**Tabla 2:** Correlación Ítem/ ítem total

<b>Ítem (caso) eliminado</b>	<b>Correlación Ítem/ ítem total</b>	<b>El coeficiente Alpha de Cronbach Con ítem (caso) eliminado</b>
13	-0,163	0,810
14	-0,269	0,816



17	-0,345	0,828
23	-0,333	0,811
23, 17, 14, 13		0,868

Eliminando estos 4 ítems el poder de discriminación del examen pasó de 0,31 a 0,42; diferencia que es significativa como se ha demostrado en otros estudios (Brailovsky, Allen, Laurence et al., 2014).

## 8.2- Resultados del SCT en residentes

Se evaluó la consistencia interna del examen con el coeficiente Alfa de Cronbach, cuyo resultado fue de 0,61.

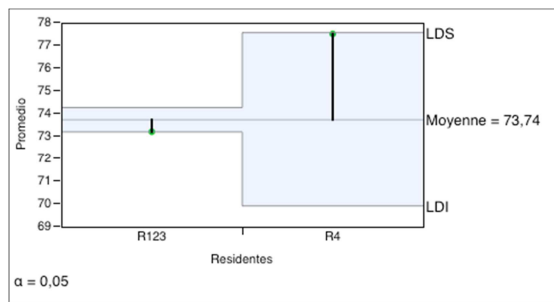
Al calcular la correlación ítem/ ítem total para todos los casos, se encontraron dos ítems (casos 11 y 12) con correlaciones negativas que fueron eliminados tal como está descrito en otros estudios (Charlin et al., 1998). Eliminando estos casos, el Alfa de Cronbach recalculado fue de 0,66.

El análisis con ANOVA no mostró diferencias estadísticamente significativas entre los 4 grupos de residentes entre sí. Por lo tanto, teniendo en cuenta la teoría del desarrollo de scripts, se decidió calcular si existían diferencias entre un grupo conformado por los residentes de 1ro., 2do. y 3er. año (junior), y otro conformado por los residentes de 4to año (senior).

Antes de proceder a los análisis estadísticos con estos dos grupos fue necesario verificar la normalidad de la distribución de los scores. Para esto se utilizó el test de Shapiro-Wilk ( p-value<sub>123</sub> 0,4554 y p-value<sub>4</sub>=0,2896 respectivamente) no significativo para ninguna de las distribuciones de los grupos, indicando que ambas eran normales.

El test de Levene (fig 1) mostró que las varianzas entre los dos grupos no eran iguales. Por lo cual, para analizar las diferencias existentes entre los dos grupos, residentes senior y residentes junior, en lugar de utilizar ANOVA a un factor (que utiliza un reagrupamiento de varianzas y que para los cálculos se sirve del promedio de las varianzas de los grupos), se utilizó el “Student T test” para varianzas desiguales.

(Figura 1) Test de Levene para el análisis de igualdad de varianzas



Las medias para cada uno de estos grupos fueron 73,21 (con un desvío estándar de 0,76) para el grupo R134, y 77,51 (con un desvío estándar de 2,0) para el grupo de R4.

Para testear la diferencia de las medias entre los grupos se usó un “Student T test” que, bajo la hipótesis nula, sigue una distribución t-student, lo que permitió testear la siguiente hipótesis nula

$$H_0: \mu_{R4} = \mu_{R123} \quad \text{vs.} \quad H_0: \mu_{R4} \neq \mu_{R123}$$

Considerando las varianzas desiguales, se obtuvo una diferencia estadísticamente significativa entre los R4 y los R123. La diferencia hallada fue de 4,29 con un error estándar de 1,29 y un valor de p de 0,0049 que permite rechazar la hipótesis nula de igualdad de medias entre los grupos con un 95% de confianza tal como se observa en la tabla 3.

**Tabla 3 :** Test de Student R4 vs. R123

Hipótesis de varianzas desiguales			
Diferencia	4,29918	t ratio	3,308483
Error estándar de la diferencia	1,29944	DL	14,76365
Límite de confianza superior (95 %) de la diferencia	7,07275	Prob. >  t	0,0049*
Límite de confianza inferior (95 %) de la diferencia	1,52562	Prob. > t	0,0024*

Finalmente, como siempre que se utiliza el “Student T test” se verificó el tamaño del efecto, para comprobar si la diferencia encontrada era pequeña o importante. Para esto, se utilizó la guía de Cohen (1992) que calcula el coeficiente “d” dividiendo la diferencia de medias entre los grupos por “t” ratio. Si  $d \geq 0,2$  muestra un pequeño efecto,  $d \geq 0,5$  el efecto es mediano y  $d \geq 0,8$  el efecto es importante. El tamaño del efecto en nuestra comparación de

grupos fue igual a 1,299 es decir que la diferencia entre los grupos es significativa y tiene un efecto importante.

## **9- Discusión**

El objetivo del trabajo fue validar el SCT en castellano para medicina familiar y evaluar algunas dimensiones del proceso de desarrollo del razonamiento clínico en los residentes de medicina familiar.

Como resultado de la investigación hemos constatado que el examen de SCT en castellano tuvo buena confiabilidad (Alfa de Cronbach: 0,84, durante la validación del examen con expertos) y buena validez de contenido para ser aplicado en residentes de medicina familiar. Al igual que Lambert et al. y Sibert et al. (entre otros) que habían validado y aplicado este tipo de examen en radio oncología y urología (Lambert et al., 2009, Sibert et al., 2006), hemos documentado que el examen es capaz de discriminar grupos diferentes en cuanto al razonamiento clínico entre los residentes de medicina familiar.

Como era de esperar, dado que la capacidad de RC aumenta con la experticia clínica, el grupo de residentes más avanzado en su formación (R4) mostró tener mayor grado de concordancia con los expertos en su RC que los de años inferiores.

Teniendo en cuenta que el desarrollo y la maduración del RC ocurre en forma gradual y dependiente del tiempo transcurrido ejerciendo la profesión, de la cantidad de pacientes a las que el profesional se ha visto expuesto y de una adecuada práctica reflexiva y supervisada, consideramos que no hemos contado con el poder estadístico suficiente para documentar este cambio gradual y que sólo hemos logrado constatar la diferencia en el RC de los residentes de cuarto año vs. los de los tres años iniciales en forma agrupada.

Si bien todas las residencias incluidas en el trabajo contaban con un curriculum similar, el hecho de que la implementación del mismo se realizara en diferentes sedes pudo haber tenido algún impacto en los resultados obtenidos en los primeros años, ya que pudieron existir diferencias en los contenidos de formación desarrollados en esos años y en las actividades de formación que cada sede implementó para alcanzarlos, así como al número y características de las situaciones clínicas a las que cada residente ha estado expuesto.

Estas diferencias tienden a desaparecer una vez completada la formación y luego de haber estado expuesto a un mayor número de pacientes, lo cual podría explicar el hecho de haber encontrado diferencias sólo entre el grupo de residentes inferiores y los residentes de cuarto año.

Si bien los resultados de nuestra investigación muestran, como otros estudios, que este tipo de examen logra medir adecuadamente el RC, como ya dijimos en este trabajo, se requieren más investigaciones para poder incluir este tipo de examen en las evaluaciones de tipo sumativo.

Por otro lado, la mayoría de los estudios consideran al SCT como un examen de costos aceptables, ya que pese a ser un examen difícil de construir es una evaluación escrita que se puede realizar en aproximadamente 90 minutos y permite evaluar a un gran número de examinados de manera simultánea. En nuestra experiencia, encontramos que para planificar este tipo de examen no sólo se debe considerar el tiempo que requiere el entrenamiento de los expertos para diseñar los casos, sino también el tiempo que se debe dedicar a entrenar a las personas que van a contestar este tipo de evaluación con el que no están familiarizados.

Nosotros tuvimos que diseñar un instructivo acerca de cómo contestar el examen para los casos en los que no era posible realizar un encuentro presencial para completar el entrenamiento previo. Esto no solo demoró los tiempos que habíamos planificado para realizar el estudio sino que nos hizo más dificultosa la tarea de sumar voluntarios al trabajo. Igualmente debemos destacar que todos los participantes del estudio consideraron su participación en el mismo como una experiencia positiva.

Es importante destacar que si bien el SCT ha demostrado poseer validez de contenido, validez de constructo, validez de predicción, confiabilidad y aceptabilidad, no es suficiente si lo que se quiere medir es la competencia clínica global, ya que no examina otras dimensiones importantes como: habilidades comunicacionales, habilidades actitudinales, manejo de situaciones de estrés, etc., que sí pueden ser evaluadas por otro tipo de exámenes como, por ejemplo, exámenes orales o el OSCE. Por lo cual es importante desarrollar programas de evaluación multidimensionales que incluyan diferentes pruebas que se complementen entre sí

e incluir en estos al SCT, ya que la calidad y utilidad de los instrumentos no se puede determinar de manera aislada sino, más bien, en relación con la de los demás.

Por último, también sería necesario desarrollar nuevas investigaciones que aporten información acerca del impacto educativo que podría tener el uso de este instrumento durante el desarrollo profesional.

En una segunda etapa de esta investigación nosotros planeamos incluir dos grupos más extremos: estudiantes del último año de la carrera de medicina a punto de comenzar su internado anual rotatorio y especialistas en medicina familiar con más de cinco años de experiencia clínica, con el objetivo de capturar una fracción mayor de la variabilidad del desarrollo de RC.

## **10- Conclusiones**

Nuestro trabajo muestra que los residentes de medicina familiar de años superiores utilizan el razonamiento no analítico (scripts) en mayor medida que los de años inferiores. Esto, a su vez, demuestra que el desarrollo del razonamiento clínico requiere de una práctica continua y reflexiva, que no es transferible de una persona a otra y que, además, inevitablemente requiere de un tiempo para su desarrollo.

Se requieren más estudios para contar con resultados más concluyentes y poder dilucidar los mecanismos subyacentes al desarrollo de este tipo de habilidades y su evaluación.

## 11- Anexos

### 11.1- Viñetas

a) Margarita Tulio de 25 años de edad concurre a la consulta con los resultados de estudios de laboratorio que usted le solicitó para evaluar su cansancio de tres meses de evolución. En el hemograma observa una Hb de 10,5

Si usted piensa en la siguiente hipótesis diagnóstica	Y encuentra	Tiene el siguiente efecto sobre su hipótesis:				
		A	B	C	D	E
Anemia ferropénica	Menstruaciones normales					
Talasemia	Ferritina baja					
Celiaquía	Volumen corpuscular medio de 75					
Anemia ferropénica	RDW de 13					
Hipotiroidismo	Mantuvo su peso el último año					

**A**-Excluye totalmente o casi totalmente la hipótesis

**B**-Hipótesis menos probable

**C**-No tiene impacto en la hipótesis

**D**-Hipótesis más probable

**E**-Confirma totalmente o casi con certeza la hipótesis

**Comentario:** \_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_

b) **Hernán de 50 años, tabaquista de 20 cigarrillos día desde los 18 años, consulta por cansancio. Durante la entrevista le cuenta que se separó de su mujer hace 6 meses.**

Si usted piensa en la siguiente hipótesis diagnóstica	Y encuentra	Tiene el siguiente efecto sobre su hipótesis:				
		A	B	C	D	E
Anemia	Sangre oculta en materia fecal negativa					
Hipotiroidismo	Hace un año estuvo medicado con Paroxetina					
Depresión	Duerme bien por la noche					
Cáncer de pulmón	Anemia					
Duelo por separación	Hemograma con hemoglobina de 12 g/dl					

**A**-Excluye totalmente o casi totalmente la hipótesis

**B**-Hipótesis menos probable

**C**-No tiene impacto en la hipótesis

**D**-Hipótesis más probable

**E**-Confirma totalmente o casi con certeza la hipótesis

**Comentario:** \_\_\_\_\_  
 \_\_\_\_\_  
 \_\_\_\_\_

## 11.2- Instructivo para la realización del SCT

El “*script concordance test*” (SCT) tiene como objetivo medir la organización del conocimiento clínico en redes conceptuales que se revelan básicamente en la acción y en la toma de decisiones del médico.

El test consiste en una serie de viñetas clínicas con cierto grado de incertidumbre, ante las cuales el médico debe tomar decisiones en base a la información que va encontrando tal como lo hace en la realidad.

Está organizado de la siguiente manera: primero se muestra una descripción del caso clínico, que contiene sólo una información básica, tal como la que se puede obtener en los primeros minutos de encuentro con el paciente. A continuación se muestran diferentes situaciones dilemáticas relacionadas con el mismo caso, cada una de las cuales está estructurada en tres partes:

- Una hipótesis de diagnóstico, una propuesta de estudios complementarios o una opción de tratamiento que podría haber surgido instantáneamente en la mente del médico, fruto de su experiencia con casos similares.
- Un nuevo hallazgo (previamente desconocido) que surge de la anamnesis, del examen físico o del resultado de algún estudio que se hubiera indicado, y que podría hacer cambiar la opinión del profesional respecto a la hipótesis o a la opción de manejo que estaba considerando.
- Un ejercicio reflexivo de toma de decisión, en el que se debe elegir una opción entre 5 soluciones alternativas,

a través de la valoración del efecto o impacto que tiene el nuevo dato encontrado sobre la hipótesis que se propone.

Por ejemplo:

Caso clínico:

Juana de 75 años de edad es diabética tipo II, de 25 años de evolución con buen control de la diabetes, hipertensa y tiene antecedente de isquemia miocárdica. Está medicada con Insulina NPH 20/10, enalapril 5mg cada 12 hs, atorvastatina 10 mg por día y aspirina 100 mg por día. Consulta por tos no productiva y regular estado general, disnea Clase Funcional I-II. El cuadro comenzó hace dos semanas



Hipótesis:

Si usted estuviera pensando en la siguiente hipótesis diagnóstica:

Neumonía



-Encuentra como nuevo dato:

Crepitantes en ambas bases pulmonares



-Pregunta:

¿Cómo influiría el nuevo hallazgo sobre la hipótesis que está considerando?



-Opciones de respuesta:



El nuevo dato encontrado:

- A. Excluye totalmente o casi totalmente la hipótesis.
- B. Hace menos probable la hipótesis.
- C. No tiene ningún impacto sobre la hipótesis.
- D. Hace más probable la hipótesis.
- E. Confirma casi con certeza o totalmente la hipótesis.

Viñeta completa:

**Juana de 75 años de edad es diabética tipo II, de 25 años de evolución con buen control de la diabetes, hipertensa y tiene antecedente de isquemia miocárdica. Está medicada con Insulina NPH 20/10, enalapril 5mg cada 12 hs, atorvastatina 10 mg por día y aspirina 100 mg por día. Consulta por tos no productiva y regular estado general, disnea Clase Funcional I-II. El cuadro comenzó hace dos semanas**

Si usted piensa en la siguiente hipótesis diagnóstica	Y encuentra	Tiene el siguiente efecto sobre su hipótesis:				
		A	B	C	D	E
Neumonía	Crepitantes en ambas bases pulmonares					
Angina de pecho	Ecostress sin cambios respecto a estudios previos					
Insuficiencia cardiaca	Soplo holosistólico 4/6 irradiado a axila y dorso					
Infarto	ECG con infradesnivel del ST-T en cara inferior, de 2mm					

El test consta de 23 viñetas clínicas con cinco hipótesis cada una, sólo se puede elegir una opción.

Es importante que cada hipótesis sea considerada por separado, sin tener en cuenta las anteriores y que todas sean contestadas teniendo en cuenta el contexto del paciente.

No hay una respuesta correcta, la puntuación del test es en base a la comparación con las respuestas obtenidas por los expertos.

Muchas gracias por su colaboración.

### 11.3- Carta enviada a residentes

Estimados:

Les escribo para solicitarle su participación en un trabajo de investigación desarrollado dentro del marco de la maestría en educación dictada en el Hospital Italiano: “Desarrollo de Script en la residencia de medicina familiar”

Numerosos estudios demostraron que al hacer un diagnóstico, los estudiantes razonan generando sucesivas hipótesis y contrastando las mismas con sus hallazgos, de manera individual y secuencial, siendo incapaces de organizar su conocimiento en estructuras más complejas. Por el contrario, los considerados expertos, logran organizar su conocimiento en forma de redes conceptuales que les permiten priorizar e interpretar de manera más eficiente los hallazgos clínicos. Estas redes conceptuales son denominadas *scripts*.

Partiendo de la base que, de acuerdo con la teoría, el razonamiento está compuesto por una serie de juicios cualitativos que pueden ser medidos, Charlin propone un método de evaluación llamado “*script concordance test*” (“SCT) cuyo objetivo es medir la organización del conocimiento clínico en redes conceptuales que se revelan básicamente en la acción y en la toma de decisiones del médico.

Si bien existen numerosos estudios publicados que muestran la diferencia que hay en la estructura de razonamiento entre un experto y un principiante, y la transformación en el tiempo de esta estructura, ningún estudio plantea o se focaliza en el momento en que ocurre este cambio durante la formación de un profesional dentro de un sistema de residencia. Por lo tanto, creemos que sería útil poder establecer en qué momento de su formación un residente adquiere el razonamiento mediante *scripts* y para esto le pedimos su colaboración.

Hemos desarrollado el *script concordance test* con los contenidos core de la Medicina Familiar. Su colaboración consistiría en resolver el test en forma anónima.

Aclaraciones:

- El test es anónimo.
- Los resultados se comparan entre los de expertos y los que arrojan los diferentes años de residencia
  - No se sacaran resultados individuales.
  - No es vinculante con su residencia.

Agradecemos su colaboración.

## 12- Bibliografía

- Alves de Lima, A. (2012). Variabilidad interobservador. Analizando algunas fuentes de error: heurísticas y categorizaciones. *Revista de Docencia Universitaria-REDU*, 10, 229-241.
- Banning, M. (2008). The think aloud approach as an educational tool to develop and assess clinical reasoning in undergraduate students. *Nurse Education Today*, 28, 814.
- Boshuizen, H.P.A., Schmidt, H.G.(1990). Biomedical knowledge and clinical expertise. Annual Conference of the American Educational Research Association, (1990) 2–16. Available at: [http://repub.eur.nl/res/pub/2782/eur\\_schmidt\\_197.pdf](http://repub.eur.nl/res/pub/2782/eur_schmidt_197.pdf).
- Boshuizen, H.P.A., Schmidt H. G., (1992). On the role of biomedical knowledge in clinical reasoning by experts, intermediates and novices. *Cognitive Science*, 16, 153-184.
- Bowen, J.L., (2006). Educational strategies to promote clinical diagnostic reasoning. *The New England Journal of Medicine*, 355, 2217–2225.
- Brailovsky C., (2001). Educación médica, evaluación de las competencias. Aportes para un cambio curricular en Argentina, 103 – 120. Available at: <http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Aportes+para+un+cambio+curricular+en+Argentina+2001#0>.
- Brailovsky, C., Allen, T., Laurence, K., Crichton, T., Laughlin, T., Van der Goes, T., Donoff, M., Schultz, K. & Bethune, C. (2014) Short answer questions based on Key Features have higher discrimination indices on a certification examination in family medicine. Ottawa Conference
- Brailovsky, C., Charlin, B., Beausoleil, S., Coté, S., van der Vleuten, C., (2001). Measurement of clinical reflective capacity early in training as a predictor of clinical reasoning performance at the end of residency: an experimental study on the script concordance test. *Medical Education*, 35 ,430–436.
- Brazeau-Lamontagne, L., Charlin, B., Gagnon, R., Samson, L., van der Vleuten, C., (2004). Measurement of perception and interpretation skills during radiology training: utility of the script concordance approach. *Medical teacher*, 26(4), 326–332.
- Charlin, B., Brailovsky, C., Brazeau-Lamontagne, L., Samson, L., Leduc, C., van der Vleuten, C., (1998). Script questionnaires: their use for assessment of diagnostic knowledge in radiology. *Medical Teacher*, 20(6), 567–571. Available at: <http://www.scopus.com/inward/record.url?eid=2-s2.0-0031672321&partnerID=40&md5=0cf926ddc7dbd8e1cddeb0519e4f2e28>.
- Charlin, B., Brailovsky, C., Leduc, C., Blouin, D., (1998). The Diagnosis Script Questionnaire: a new tool to assess a specific dimension of clinical competence. *Advances in Health Sciences Education : theory and practice*, 3(1990), 51–58. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/12386395>.
- Charlin, B., Tardif, J., Boshuizen, H.P., (2000). Scripts and medical diagnostic knowledge: theory and applications for clinical reasoning instruction and research. *Journal of the Association of American Medical Colleges*, 75(1), 182–190.
- Charlin, B., van der Vleuten, C., (2004). Standardized assessment of reasoning in contexts of uncertainty: the script concordance approach. *Evaluation & the Health Professions*, 27, 304–319.
- Charlin B., Gagnon R., Sauvé E., Coletti M., (2007). Composition of the panel of reference for concordance tests: do teaching functions have an impact on examinees' ranks and absolute scores?.\_Medical Teacher\_

- 29(1), 49-53.
- Charlin, B., Roy, L., Brailovsky, C., Goulet, F., Vleuten, C., (2009). The script concordance test: a tool to assess the reflective clinician. *Teaching and Learning in Medicine*, 12(4),189–195.
- Cohen, J. (1992) A Power Primer *Psychological Bulletin* 112, 155-159
- College of Family Physicians of Canada, (2009). Defining competence for the purposes of certification by the College of Family Physicians of Canada. The new evaluation objectives in family medicine. Available at: <http://www.cfpc.ca/uploadedFiles/Education/Defining%20Competence%20Complete%20Document%20bookmarkmarked.pdf>
- Dory, V., Gagnon, R., Vanpee, D., Charlin, B., (2012). How to construct and implement script concordance tests: insights from a systematic review. *Medical Education*, 46, 552–563.
- Durante, E., (2006). Algunos métodos de evaluación de las competencias: escalando la pirámide de Miller. *Revista Hospital Italiano de Buenos Aires*, 26(2), 55–61.
- Durante, E., (2005). La evaluación de los conocimientos: lo que parece ser, ¿es realmente lo que es? *Revista Hospital Italiano de Buenos Aires*, 25(1),18–23.
- Epstein, R., (2007). Assessment in medical education. *The New England Journal of Medicine*, 356, p.2109; author reply 2110.
- Fournier, J.P., Demeester, A., Charlin, B., (2008). Script concordance tests: guidelines for construction. *BMC Medical Informatics and Decision Making*, 8, 18–24.
- Frederiksen, N., (1984). The real test bias: influences of testing on teaching and learning. *American Psychologist*, 39, 193–202.
- Gormaz, C., Brailovsky, C., (2012). Desarrollo del razonamiento clínico en medicina. *Revista de Docencia Universitaria-REDU*, 10, 177–199.
- Kane MT. (1992). The assessment of professional competence. *Evaluation & The Health Professions*, 15(2),163-82.
- Lambert, C., Gagnon, R., Nguyen, D., Charlin, B., (2009). The script concordance test in radiation oncology: validation study of a new tool to assess clinical reasoning. *Radiation Oncology*, 4, 7.
- Memon, M.A., Joughin, G.R., Memon, B., (2010). Oral assessment and postgraduate medical examinations: establishing conditions for validity, reliability and fairness. *Advances in Health Sciences Education*, 15, 277–289.
- Meterissian, S., Zabolotny, B., Gagnon, R., Charlin, B., (2007). Is the script concordance test a valid instrument for assessment of intraoperative decision making skills? *American Journal of Surgery*, 193(2),. 248–51.
- Monnier, P., Be´dard, M., Gagnon, R., Charlin, B., (2011). The relationship between script concordance test scores in an obstetrics-gynecology rotation and global performance assessments in the curriculum. *International Journal of Medical Education*, 2, 3–6.
- Norman, G.R., (2005). Research in clinical reasoning: past history and current trends. *Medical Education*, 39, 418–427.
- Schmidt, H.G., Boshuizen, H.P.A., (1990). Effects of activation of prior knowledge on the recall of a clinical case. The Annual Conference of the American Educational Research Association. Available at: <http://www.eric.ed.gov/ERICWebPortal/contentdelivery/servlet/ERICServlet?accno=ED321713> \n<http://www.eric.ed.gov/ERICWebPortal/detail?accno=ED321713>.

- Schmidt, H.G., Norman, G.R., Boshuizen, H.P., (1990). A cognitive perspective on medical expertise: theory and implications. *Academic Medicine*, *65*, 611–21.
- Sibert, L., Darmoni, S., Dahamna, B., Hellot, M-F., Weber, J., Charlin, B., (2006). On line clinical reasoning assessment with script concordance test in urology: results of a French pilot study. *BMC Medical Education*, *6*, 45. Available at: <http://www.biomedcentral.com/1472-6920/6/45>.
- Tornimbeni, S., Pérez, E., Olaz, F., (2008). *Introducción a la Psicometría*.
- van der Vleuten, C. P. M., (1996). Assessment of professional competence: developments, research and practical implications. *Advances in Health Sciences Education*, *1*, 41–67.
- van der Vleuten, C. P. M., Schuwirth, L. W. T., Scheele, F., Driessen, E. W., Hodges, B., (2010). The assessment of professional competence: building blocks for theory development. *Best Practice and Research: Clinical Obstetrics and Gynaecology*, *24*, 703–719.
- Van de Wiel, M. W. J., Boshuizen, H. P. A., Schmidt, H. G., (1994). Differences in knowledge application by students and medical experts in clinical reasoning. *The Annual Meeting of the American Educational Research Association*.