

**UNIVERSIDAD NACIONAL DE CÓRDOBA**  
**FACULTAD DE CIENCIAS MÉDICAS**

**IMPLEMENTACIÓN DE INSTRUMENTOS DE EVALUACIÓN  
OBJETIVA DE COMPETENCIAS TÉCNICAS EN CIRUGÍA**

Trabajo de Tesis para optar al  
Título de Doctor en Medicina y Cirugía

**Andrés Carlos Kasparian**

CÓRDOBA  
REPÚBLICA ARGENTINA  
2015

## **COMISIÓN DE SEGUIMIENTO DE TESIS**

**DIRECTOR:** Profesor Doctor Ricardo A. Chércoles

**INTEGRANTES:** Profesor Doctor Rolando B. Montenegro  
Profesor Doctor Hugo O. Juri

Artículo 30º del Reglamento de Carrera de Doctorado en Medicina y Cirugía  
“LA FACULTAD DE CIENCIAS MÉDICAS NO SE HACE SOLIDARIA CON LAS  
OPINIONES DE ESTA TESIS”

El presente trabajo de investigación ha sido realizado con el aporte de una Beca de Finalización de Doctorado (2010 y 2011) otorgada por la Secretaría de Ciencia y Tecnología de la Universidad Nacional de Córdoba. RR 2239/10 y 2263/11.

## DEDICATORIAS

A mis abuelos, los sobrevivientes.

A mis padres, Jorge y Alicia.

A mi compañera, Ana Luisa.

A mi hijo Juani.

## AGRADECIMIENTOS

- A los Dres. Rafael Jover Clos y Alfredo C. Martínez, docentes de la Unidad Hospitalaria de Cirugía N° I, y médicos de planta del Servicio de Cirugía N° I del Hospital Nacional de Clínicas (UNC) por el tiempo y el esfuerzo dedicado a la evaluación de las filmaciones.
- Al personal médico del Servicio de Cirugía N° I (Unidad Hospitalaria de Cirugía N° I) del Hospital Nacional de Clínicas, que participó en los procedimientos quirúrgicos que fueron filmados.
- Al personal del Quirófano Central del Hospital Nacional de Clínicas (servicio de anestesiología, instrumentadores quirúrgicos y auxiliares) por facilitar la tarea de la filmación en el ámbito de trabajo.
- A los Profesores Dra. Marta Fiol y Dr. Rubén D. Ruiz, de la Cátedra de Fisiología Humana de la Facultad de Ciencias Médicas (UNC) por el asesoramiento estratégico en el planteamiento y ejecución de la tesis.
- Al Lic. en Cine y Técnicas Audiovisuales Juan Avaltroni, a cargo de la captura de imágenes y filmación de las cirugías en quirófano.
- A la Dra. Marina Ponzio por su colaboración en la confección de los gráficos del capítulo Resultados.
- A mis maestros, mis compañeros y mi familia, particularmente mi padre, quienes promovieron y estimularon en mí la ejecución del proyecto de la tesis doctoral.

## ÍNDICE

Capítulo 1: INTRODUCCIÓN	Pág. 10
Capítulo 2: MATERIALES Y MÉTODOS	Pág. 28
Capítulo 3: RESULTADOS	Pág. 35
Capítulo 4: DISCUSIÓN	Pág. 43
Capítulo 5: BIBLIOGRAFÍA	Pág. 77

## RESUMEN

**Introducción:** La adquisición de habilidades quirúrgicas constituye un factor central en la formación de todo cirujano. Sin embargo, la evaluación de las habilidades técnicas es uno de los factores más débiles y menos desarrollados. En la actualidad los recursos para evaluar las competencias técnicas adolecen de subjetividad, falta de confiabilidad y validez. La observación directa, método de evaluación más frecuentemente utilizado en nuestro medio, presenta inconvenientes instrumentales y está fuertemente influenciada por las relaciones intersubjetivas y los rasgos de personalidad. El objetivo de esta investigación es proponer la creación y el uso de un instrumento objetivo para evaluar el desempeño técnico, y determinar su confiabilidad y validez.

**Materiales y métodos:** se seleccionaron dos procedimientos: la colecistectomía laparoscópica y la hernioplastia inguinal (técnica de Lichtenstein). Se constituyeron tres grupos de comparación según la experiencia quirúrgica: inicial, intermedio, y expertos. Se filmaron las cirugías en tiempo real, sin identificación del paciente ni del cirujano. Las filmaciones sin edición fueron asignadas a dos cirujanos expertos en una secuencia aleatoria por sorteo y con sobres sin identificación. Para la evaluación se propuso el uso de un instrumento objetivo (explicitación de pasos a evaluar y cuantificación mediante escala de Likert) y específico para cada procedimiento. Así mismo se utilizó la escala global OSATS (R. Reznick). Se aplicó un análisis de varianza no paramétrico para determinar la validez. Valores de  $p$  menores a 0.05 fueron considerados estadísticamente significativos. Valores superiores a 0,80 del Coeficiente alfa de Cronbach aseguraron confiabilidad.

**Resultados:** Desde Abril del 2010 hasta Diciembre del 2012 se filmaron 36 colecistectomías videolaparoscópicas y 31 hernioplastias inguinales. Se encontraron diferencias significativas entre los grupos en todos los ítems evaluados  $p < 0.05$ . El coeficiente  $\alpha$  de Cronbach fue mayor a 0,80 para ambas técnicas. No hubo diferencias significativas entre las calificaciones de ambos evaluadores. No hubo diferencias entre nuestro instrumento específico y la escala global OSATS.

**Discusión:** Es posible y útil aplicar un instrumento objetivo de evaluación del desempeño técnico en cirugía. La herramienta presentó validez de constructo y confiabilidad aceptables. La filmación confiere perdurabilidad a un evento efímero: la cirugía. La objetividad se basa en la enunciación y cuantificación explícita de cada paso, y en la aleatorización y anonimato de la muestra. La uniformidad de criterios entre los evaluadores es fundamental para obtener resultados satisfactorios. Evaluar siempre implicará un recorte de la realidad.

## SUMMARY

**Introduction:** technical skills acquisition is considered to be of paramount importance in surgical training. Yet, formal assessment of technical skills is the weakest and less developed area. Currently available resources to evaluate technical skills are largely subjective, and lack of validity and reliability. Direct observation, one of the most frequently used methods, is largely biased by interpersonal subjectivity and personality traits. We propose the creation and use of a new procedure-specific tool for objective assessment of technical skills in surgery to evaluate validity and reliability.

**Materials and Method:** laparoscopic cholecystectomy and Lichtenstein's inguinal hernia repair were the chosen procedures. Three groups of comparison were defined according to surgical expertise: initial, intermediate, and experts. Surgeries were videorecorded in real time without identification of the patient or the surgeon. Tapes without any posterior edition were assigned to two expert surgeons in a blind and randomized sequence. A newly proposed procedure-specific rating scale was used for evaluation, as well as Reznick's OSATS global scale. Kruskal-Wallis non-parametric test was used to assess validity.  $p < 0,05$  was considered statistically significant. Chronbach's  $\alpha$  values  $> 0.8$  granted reliability.

**Results:** from April 2010 to December 2012 36 laparoscopic cholecystectomies and 31 inguinal hernia repairs were recorded. Significant difference was found among groups of comparison for every item ( $p < 0.05$ ). Chronbach's  $\alpha$  was largely over 0,80 for both procedures. There was no significant difference between the assessments of the two raters. There was no significant difference between our newly proposed procedure-specific instrument and Reznick's OSTATS.

**Discussion:** the objective assessment of technical skills in surgery is feasible and useful. The tool we proposed showed construct validity and reliability. Video recording of surgical procedures grants durability over time to an ephemeral phenomenon. The objectivity is based on the explicit statements and quantification of every step to be evaluated, and the blind randomization and anonymous treatment of the sample. Sharing the same quality criteria between evaluators is of paramount importance to reach satisfactory results. The process of evaluation always implies a shortened view of the reality.

## Capítulo 1

# INTRODUCCIÓN

En su sentido más amplio y desde la perspectiva de la antropología cultural la educación puede considerarse como aquella parte de la experiencia endocultural que, a través del proceso de aprendizaje, equipa a un individuo para que ocupe su lugar como miembro adulto de su sociedad. Este proceso forma parte de otro fenómeno más general conocido como “enculturación primitiva”, que le permite al individuo apropiarse y manejar su cultura. Un significado mucho más acotado del concepto “educación” lo restringe a los procesos de transmisión de conocimientos y su aprendizaje, que tienen lugar en tiempos específicos, en sitios particulares fuera del hogar, por períodos definidos, y por personas especialmente preparadas para esta tarea. Esta perspectiva más limitada y específica le confiere al concepto de educación el significado de **enseñanza dentro de un marco institucional**. (1)

La **educación** es una práctica tan antigua como las sociedades mismas. A través de ella se producen, reproducen y transforman los bienes simbólicos y materiales que posibilitan la vida en sociedad. En comunidades pequeñas y escasamente organizadas la educación se desarrolla espontáneamente en el seno de la vida familiar y comunitaria, y transmite saberes prácticos y por lo general poco sofisticados, que son imprescindibles para dominar la naturaleza inmediata al ser humano y permitir la subsistencia.

A medida que las comunidades se expanden numéricamente y se complejizan sus modos de funcionamiento, las prácticas educativas comienzan a diferenciarse entre sí y a alejarse espacialmente de las actividades domésticas. Surge la necesidad de la creación de espacios sociales específicos para su desarrollo, y la existencia de individuos que asuman la responsabilidad de realizar esta transmisión de saberes. Se produce entonces la **institucionalización de la educación**.

Es justamente en el ámbito de la aparición de los procesos educativos institucionalizados donde **surge la evaluación**, incluso mucho antes de la aparición de la pedagogía moderna, adoptando probablemente la figura del “examen”. Tanto la educación institucionalizada como la evaluación constituyen herramientas de mantenimiento y consagración del orden social existente y se requieren recíprocamente. (2)

El diccionario de la Real Academia Española define el concepto evaluar como un verbo transitivo cuyo significado es “señalar el valor a algo”, “estimar,

apreciar, calcular el valor a algo” y finalmente, “estimar los conocimientos, aptitudes y rendimiento de los alumnos” (3)

El uso del lenguaje cotidiano le otorga al verbo evaluar un amplio rango de significados que incluye conceptos tan diversos como estimar, interpretar, medir, comprender, conocer, comparar, valorar, juzgar, aprehender, cifrar, apreciar, calcular, justipreciar, valorar y atribuir valor a algo. La operación de evaluar algo o alguien consiste en estimar su valor no material.

En la práctica diaria del ámbito de la enseñanza el significado de evaluar es menos plurívoco: consiste en calificar a los estudiantes aplicando distintas clases de instrumentos para obtener la información a partir de la cual se emitirá un juicio.

Resulta interesante destacar cómo se pueden distinguir dos perspectivas sobre el significado de evaluación derivadas de las acepciones anteriores. Una de ellas, que agrupa algunos términos como medir, valorar, cifrar, que involucran mesurar con precisión; en este sentido expresan una cantidad precisa, atendiendo a un aspecto cuantificado. La segunda perspectiva sobre evaluación, implica ciertas operaciones como estimar, aprehender, comprender, en otras palabras, emitir un juicio cualitativo y, eventualmente, aproximativo sobre una realidad. “Evaluar es en consecuencia un término bien singular que puede expresar una cosa y su contrario: lo preciso y lo aproximado, lo cuantitativo y lo cualitativo” (4, 5)

La distinción entre lo cualitativo y lo cuantitativo ha sido motivo de un largo debate en el ámbito de la teoría e historia de la ciencia. Las fuentes consultadas en relación a estos aspectos de la investigación revelan que a medida que la modernidad se fue consolidando se sentaron las bases filosóficas, sociales y políticas que dieron origen a la ciencia. Hamilton señala la importancia de los textos de Descartes (1596-1650) quien proclamó “la importancia de las matemáticas y la objetividad en la búsqueda de la verdad” que fueron considerados los cimientos fundamentales de la **investigación cuantitativa**. Por otro lado se le atribuye a Kant (1724-1804), específicamente con su obra *Crítica a la Razón Pura*, cuestionar los ideales del “objetivismo cartesiano” y proponer un modelo de conocimiento basado en la interpretación y la comprensión, aspectos asociados a la **investigación cualitativa**.

La distinción entre “explicación” (refiriendo a lo cuantitativo) y “comprensión” (refiriendo a lo cualitativo) será luego profundizada por Dilthey (1833-1911) quien remarcó las diferencias entre las ciencias de la naturaleza y las ciencias del espíritu o humanas.

Para Conde F, con las figuras de Platón y Aristóteles en la Grecia clásica, ya se pueden reconocer posturas filosóficas pro-cuantitativas y pro-cualitativas, respectivamente.

El debate entre cantidad y cualidad continúa abierto en la actualidad y se vincula con la incidencia que han tenido los textos de Kuhn y el concepto clave de *paradigma*. Valles señala que los paradigmas deben entenderse como un sistema de creencias básicas (principio, supuesto) sobre: “a) la naturaleza de la realidad investigada (supuesto ontológico) b) sobre el modelo de la relación entre el investigador y lo investigado (supuesto epistemológico) y c) sobre el modo en que podemos obtener conocimiento sobre dicha realidad (supuesto metodológico). Es decir que el concepto no se refiere solo al método, si no que sirve de guía para orientar al investigador también en los aspectos ontológicos y epistemológicos. (6)

El concepto de evaluación fue aplicado inicialmente para referirse a cuestiones vinculadas exclusivamente con el proceso y los resultados del aprendizaje. Sin embargo, es durante la década del 60 que muda su significado asociado al **rendimiento escolar** hacia un concepto diferente y mucho más amplio: la **evaluación educativa**. Esta nueva significación abarca numerosos otros aspectos como son la evaluación del docente, del programa, de la institución educativa, del plan de estudios, etc. e hizo perder a la evaluación su precisión respecto a su objeto de estudio, poniendo en evidencia una problemática subyacente en la actualidad: **la dificultad para dar rigor al campo conceptual y a la actividad evaluatoria (7)**

Desde la perspectiva de su origen y evolución histórica, el examen como tal, aparece en el año 1200 a.C. dando respuesta a un problema de la burocracia china en la selección de personas para el servicio de una casta a otra. (Max Weber en “*Economía y Sociedad*”). Stuffelbeam y Shinkfield remontan el mismo hecho al año 2000 ac “cuando algunos oficiales chinos dirigieron unas investigaciones de los servicios civiles” constituyendo un sistema de examinación primitivo de selección de candidatos a ocupar cargos

públicos que subsiste hasta la actualidad. Los mismos autores afirman que en el Siglo V a.C. “Sócrates y otros maestros griegos utilizaron cuestionarios evaluativos como parte de su metodología didáctica” y reconocen que hacia el 1800, en Inglaterra, “se utilizaron comisiones reales para evaluar los servicios públicos” (8). Este dato demuestra que el examen no nace en la relación de transmisión de conocimiento, sino como respuesta a una exigencia social, y que este surgimiento le imprime características constitutivas que son incorporadas en su tránsito a la educación.

Según Durkheim, en su *Historia de la educación y de las doctrinas pedagógicas*, es en la Edad Media cuando aparece el examen vinculado a la práctica escolar con algunas características particulares. Sólo hay tres exámenes: el de bachiller, el de la licenciatura y el de doctor. Sólo se permitía presentar examen a aquellos candidatos de los que se estaba seguro que podían ser aprobados. Las funciones del examen eran demostrar en público el grado de madurez intelectual y el dominio de habilidades cognitivas desarrollado por el estudiante (9). La adquisición de los títulos (bachiller, licenciado y doctor) cumplía dos funciones: una de ellas es la más explícita y consistía en otorgar crédito de validez internacional a los saberes de los aspirantes a los cargos docentes. Y la otra, no revelada ni explicitada, que tenía que ver con el mantenimiento del equilibrio de poder dentro de la institución universitaria. De forma tal que los maestros ya consagrados ejercieran en forma monopólica el derecho de seleccionar quiénes ingresarían en la corporación educativa, y continuaran siendo ellos mismos sus reguladores.

En el Siglo XVII durante el conflicto de la Reforma y la Contrarreforma aparece la didáctica y surgen dos procesos divergentes de institucionalización del examen: en uno se considera un problema metodológico (Comenio, en su *Didáctica Magna*, 1657) y en otro, como una supervisión permanente (La Salle, en la *Guía de las escuelas cristianas*, 1720) Según Foucault (10), La Salle establece el examen como un instrumento de vigilancia continua e ininterrumpida. Los resultados de los exámenes periódicos son entregados por la autoridad, y se estimula a quienes tienen mejores resultados y se reprende en público a quienes se atrasan.

Para **Comenio** el examen forma parte del bagaje metodológico del docente, es decir constituye un instrumento profesional para enfrentar cualquier problema de aprendizaje. Desde esta perspectiva, los resultados no se califican ni permiten una promoción, colocando al método en sí mismo en el centro del trabajo docente.

En los colegios jesuíticos se presenta el fenómeno de asociación de la evaluación con el control y el disciplinamiento, que la transforma en un ritual vacío de significado pedagógico, y ligado a un hecho externo al proceso mismo del aprendizaje: el premio o estímulo. El maestro de los colegios jesuíticos actúa como un vigilante, más parecido a un preceptor, que entabla con el alumno una relación asimétrica, en la que éste último aparece puerilizado, concebido intrínsecamente débil, maleable y potencialmente corregible.

Hacia fines del Siglo XIX y principios del XX comienza el período de industrialización en los Estados Unidos de América. Aparece en este momento el **Taylorismo y la organización científica del trabajo**, que propulsaron el uso de **tests** de aptitudes como herramientas de reclutamiento y promoción del personal en las empresas. De este modo, se pretendía hacer más eficiente la producción industrial y menos conflictiva la relación entre obreros y empresarios. Del mismo modo, el proceso de industrialización influyó notablemente para que la institución escolar se adaptara a los requerimientos del aparato productivo. Una de las características de este momento histórico fue la **uniformidad**. A fin de preparar a los jóvenes para el mercado de trabajo, los educadores crearon cursos uniformados. Aparecieron los primeros tests de inteligencia (Binet y Terman) sistemas de admisión escolar, reglas de acreditación, y promoción escolar. Surge en este período el test de opción múltiple. Entendida como homogeneidad de criterios, la uniformidad tendrá obviamente efectos en la evaluación escolar.

En el Siglo XX la **Teoría del test** incorpora el supuesto que es posible medir el aprendizaje, entendiéndose éste como un cambio de conducta, susceptible de ser medible y verificable. El aprendizaje deja de ser una cualidad y se transforma en un objeto cuantificable que puede ser medible. La incorporación de la teoría del test en la práctica escolar implica la aceptación de una visión experimental de los programas de estudio: muestrear el contenido, validar instrumentos, determinar atributos y operaciones, aplicar

estadística descriptiva para valorar el instrumento y asignar la calificación. Esta teoría permitió legitimar como naturales, es decir, del orden de lo biológico, desigualdades sociales. De esta manera la teoría del test es una teoría política, que permitió, según Bowles-Gintis “establecer sutiles vinculaciones entre delincuentes, pobres, y poco inteligentes”

Con la publicación de la obra de Henry Fayol “Administración general e industrial (1916) en la que se establecen las fases del planeamiento organizacional (planear, ejecutar, y evaluar), se gesta el moderno discurso científico de la educación.

Con posterioridad al uso del test, a partir de la década del 40 se incorpora la noción de **evaluación**, que aparece como más científica y neutra, pero recoge el sentido implícito de control de la administración científica del trabajo y la medición de la teoría del test.

En suma, es hacia los años 20 cuando la sociedad estadounidense comienza a manifestar una confianza ciega en este nuevo dispositivo que era el test. En este contexto surge el discurso científico de la evaluación, signado por el sesgo tecnocrático del control y la homogeneización y por el sustrato del eficientismo. Estas características quedan muy marcadas dentro del campo educativo, y perviven hasta nuestros días generando no pocas controversias (11,12).

Por otro lado, retomando cuestiones vinculadas más directamente con el complejo aprendizaje de la cirugía podemos afirmar que, la adquisición de habilidades quirúrgicas constituye un factor central en la formación de todo cirujano. No hay duda que la destreza técnica es considerada por el público en general y por los propios médicos en formación como uno de los factores más importantes de la lista de cualidades que debería poseer todo cirujano (13, 14).

Una encuesta publicada en el año 2001 realizada con cirujanos del Reino Unido que alcanzaron el grado de “maestro”, reveló que más del 80% de los encuestados identificaron la excelencia en las habilidades psicomotoras como el factor más determinante de una carrera quirúrgica exitosa (15, 16). Otro reporte publicado en el mismo año, basado en encuestas y reuniones de consenso, realizado con cirujanos de diferentes especialidades con grado de “maestro” provenientes de centros europeos y norteamericanos, consideró a la

destreza operatoria como el atributo de mayor importancia luego del juicio clínico (17, 18).

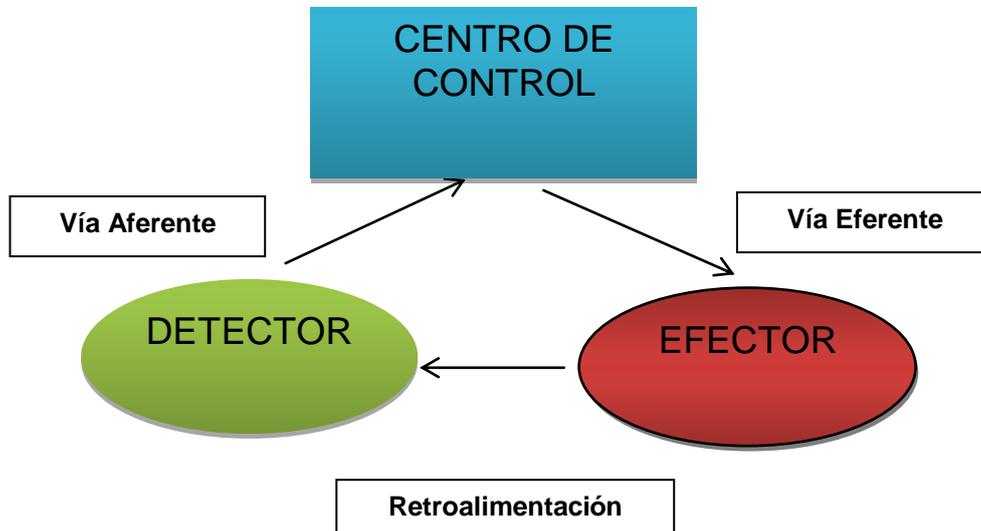
Sin embargo, dentro de todos los requisitos incluidos en la formación de un cirujano la evaluación de las habilidades técnicas constituye el factor más débil y menos desarrollado. En la actualidad las herramientas de evaluación de los programas de residencias quirúrgicas hacen fuertemente hincapié en la adquisición de conocimientos y en la aplicación de los mismos en forma criteriosa, en tanto que los recursos para evaluar las competencias técnicas adolecen de subjetividad, falta de confiabilidad y validez (19).

En este punto resulta importante explicar que cuando mencionamos el término “habilidades técnicas” estamos haciendo referencia a la destreza o desempeño operatorio, es decir a un tipo particular de motilidad somática dentro del grupo de los movimientos voluntarios. La habilidad operatoria se alcanza luego de un entrenamiento sistemático y específico que habitualmente lleva varios años, responde a una serie de principios básicos que, una vez incorporados, permiten la ejecución satisfactoria y exitosa de un procedimiento quirúrgico completo y más complejo.

Desde el punto de vista de la fisiología del sistema nervioso, la ejecución de cualquier tipo de movimiento, al igual que de todas las funciones orgánicas que permiten la existencia de un sujeto, implica la existencia de un mecanismo de control homeostático. Estos mecanismos son los encargados de controlar las variables vitales y asegurar el correcto funcionamiento del sistema, ya que garantizan el mantenimiento de la constancia del medio interno, concepto conocido como “**homeostasis**”. En el año 1926 el fisiólogo estadounidense Walter B. Cannon introdujo el concepto de homeostasis para describir “las condiciones uniformes que prevalecen en el organismo a pesar de todas las influencias del medio externo que tienden a cambiarlo”. Se trata de un equilibrio dinámico, en constante intercambio con el medio circundante, que permite la independencia del organismo de las fuentes de agua y de sustratos energéticos necesarios para la subsistencia (20, 21, 22).

Los mecanismos de control homeostáticos (figura 1) están formados por los siguientes elementos constituyentes: un **detector**, encargado de medir el parámetro controlado y enviar la información a un **centro de control** a través de una vía aferente. Éste compara el parámetro fijado por la naturaleza con las

variaciones que eventualmente se producen. Y un **efector**, comunicado con el centro de control a través de una vía eferente, que corrige la variación producida.



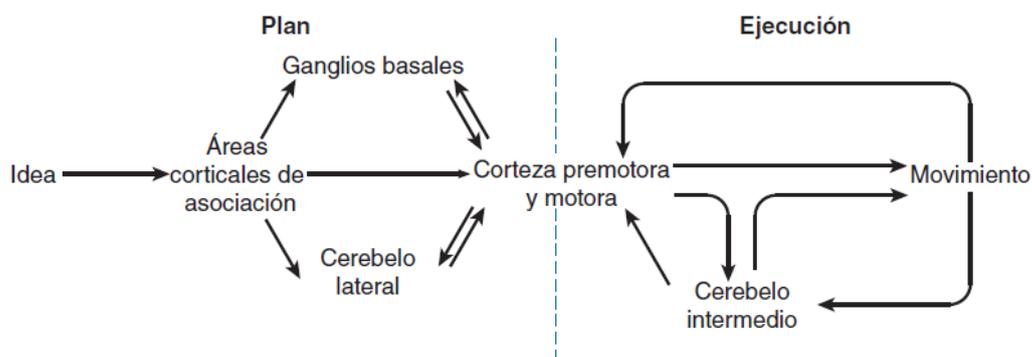
**Figura 1:** Componentes de un mecanismo de control homeostático tal como fue descrito originalmente por Cannon WB en 1926.

La actividad motora somática depende en última instancia del patrón de descarga de la neurona motora espinal  $\alpha$  (alfa) o **motoneurona inferior**, localizada en el asta anterior de la médula espinal, y de las neuronas homólogas de los núcleos motores de los pares craneales del tallo encefálico. Estas neuronas constituyen lo que se conoce como la **vía final común** del sistema nervioso, ya que sobre ellas confluyen todas las señales provenientes de las estructuras superiores, y es justamente a través de ellas que emergen todo tipo de órdenes hacia **los efectores**, en este caso el **músculo esquelético**.

Las señales que convergen en las neuronas motoras efectoras tienen tres funciones: 1) permitir la actividad llamada voluntaria, 2) ajustar la postura corporal para establecer un escenario de estabilidad para la ejecución de los movimientos y 3) coordinar la acción de los diversos músculos para que los movimientos sean suaves y precisos.

Los comandos para el movimiento voluntario (figura 2) se originan en las áreas de asociación cortical. Los movimientos se planean en esa parte de la

corteza, los ganglios basales, y en las porciones laterales de los hemisferios cerebelosos. Estas estructuras activan a la corteza premotora y motora principal, previo relevo en el tálamo. La eferencias de la corteza motora llegan directamente a la motoneurona inferior a través de la vía córtico-espinal o su análoga, la córtico-bulbar (vía piramidal). La ejecución del movimiento produce alteraciones en la sensibilidad propioceptiva (músculos, tendones y articulaciones) y la exteroceptiva (piel). Esta información ascendente generada por el propio movimiento ejecutado hace relevo en el espinocerebelo (neocerebelo) y finalmente alcanza a la corteza motora. Constituye una forma de retroalimentación que permite ajustar y suavizar el movimiento, confiriéndole las características de normalidad.



**Figura 2:** Control del movimiento voluntario. Las órdenes se originan en las áreas de asociación cortical. La corteza, los ganglios basales, y el cerebelo participan en la planificación del movimiento. El movimiento ejecutado por la corteza motora primaria se transmite por la vía piramidal a la motoneurona inferior. El cerebelo brinda retroalimentación para afinar y ajustar el movimiento y le confiere las características de normalidad. Tomado de Ganong, *Fisiología Médica*, 23ª edición.

Es difícil definir exactamente qué es un movimiento voluntario y separar de él lo que tiene de automático. Esto es así porque un movimiento voluntario tiene componentes automáticos que, muchas veces, son controlados voluntariamente. En el repertorio de los movimientos ya aprendidos lo único voluntario es la iniciación, y los demás componentes del mismo son automáticos inconscientes. Por lo tanto se podría sostener que lo único voluntario del movimiento “voluntario” es su propósito.

Otro problema que se plantea al abordar el estudio de los movimientos voluntarios tiene que ver con la notable capacidad de ejecutar acciones tan complejas como las requeridas en ciertas profesiones, entre las que se

encuentra el ejercicio de la práctica quirúrgica, o el desarrollo de ejercicios acrobáticos, o el uso profesional de instrumentos musicales. Por otro lado, también es sorprendente la complejidad y precisión de los movimientos musculares requeridos para hablar, cantar o gesticular. Toda esta perfección del acto motor se debe a la existencia de una vía piramidal correctamente desarrollada e indemne. La terminación directa de fibras piramidales en las motoneuronas alfa permite un control preciso y bien discriminado de las unidades motoras, que se traduce en un movimiento preciso. Para ello el sistema piramidal está vinculado con ajustes espaciales que permiten adaptar de manera exacta el movimiento a las características del estímulo (23).

Los centros de control y las vías de conducción de los músculos esqueléticos del tronco y las partes proximales de las extremidades están localizados en la región ventral y medial. Estos grupos musculares están encargados de los ajustes posturales y los movimientos más gruesos. Mientras que los centros y las vías encargadas del control de los músculos distales de las extremidades, responsables de los movimientos finos y diestros, se ubican en las regiones laterales.

**El sistema motor “aprende con la práctica” y el desempeño mejora con la repetición. Este fenómeno implica plasticidad neuronal (24).**

En el año 1967 Fitts y Posner proponen una teoría de adquisición de habilidades motoras basada en tres etapas o estadios, que ha sido ampliamente aceptada en diversos ámbitos, incluida la literatura científica especializada en cirugía (25) (Tabla 1).

La primera etapa denominada **cognitiva**, implica la intelectualización de la tarea por parte del estudiante, es decir la comprensión de cada uno de los pasos constituyentes que pueden desprenderse de una maniobra compleja. Se hace necesario un proceso de deconstrucción del movimiento para que pueda ser enseñado a otros (26). La ejecución característica en este momento es esencialmente errática y discontinua ya que implica para el ejecutante una reconstrucción de la maniobra (modo y secuencia) que requiere una estrecha supervisión consciente, y muchas veces no se puede realizar con éxito. En la segunda etapa, llamada **integrativa**, el entendimiento consciente de cada paso

constitutivo de una maniobra o gesto comienza a traducirse en un desempeño motor apropiado incipiente. La ejecución puede ser más fluida y tener menos interrupciones, pero no está libre de repeticiones que implican recomenzar con la secuencia de movimientos desde el inicio por falta de efectividad. La tercera etapa es llamada **autonómica**. En esta instancia la ejecución va adquiriendo gradualmente fluidez, y la acción se transforma en una **unidad** compleja de movimientos precisos. No se hace necesario pensar en cómo ejecutar cada uno de los pasos y se puede prestar atención a otros aspectos más complejos, que podrían incluso no estar relacionados directamente con la técnica.

**Tabla 1:** Teoría de la adquisición de habilidades motoras en tres etapas, de Fitts y Posner. Tomada y adaptada de Fitts PM, Posner MI. *Human Performance*. Belmont, CA: Brooks/Cole, 1967.

Etapa	Objetivo	Actividad	Desempeño
<b>Cognitiva</b>	Conocimiento y entendimiento de la tarea y de sus pasos	Explicación y demostración	Errático y discontinuo
<b>Integrativa</b>	Inicio de ejecución	Práctica intencionada y feedback del instructor	Menos interrupciones. Con repeticiones
<b>Automatización</b>	Ejecución con velocidad, efectividad y precisión	Escasa supervisión cognitiva. Foco en el refinamiento de la ejecución	Continuo, fluido y adaptativo

En nuestro ámbito, la evaluación de las competencias técnicas se hace habitualmente presente en dos instancias: una de ellas es el final de cada período lectivo del programa de formación de postgrado (residencia de cirugía) y constituye un requisito formal para la promoción al siguiente año, y que se complementa con otros exámenes orientados a evaluar conocimientos específicos. La otra, guarda relación con los requisitos solicitados por las instituciones formadoras o de certificación profesional para la obtención del título de especialista. En ambos casos, hay un equipo de evaluadores constituido por cirujanos con experiencia y vinculados a la tarea formativa, que

se hace presente en el quirófano con el propósito de presenciar una cirugía asignada al candidato y emitir luego un veredicto calificador.

Esta forma de evaluación conocida como observación directa presenta algunos inconvenientes que no pueden ser soslayados y que se relacionan con los siguientes hechos inherentes al método en sí mismo:

- El procedimiento quirúrgico reviste siempre la calidad de efímero, o sea es transitorio con un inicio y un final, y unidireccional, o sea no permite la vuelta al principio ni el retroceso.
- Las posibilidades físicas de observación directa en el quirófano no son óptimas: el campo quirúrgico suele ser estrecho o demasiado profundo, la iluminación puede no ser la adecuada para el observador, el riesgo de contaminación hace difícil el acercarse en proximidad, y la visualización es intermitente.
- El proceso de evaluación en sí mismo se lleva a cabo en una instancia posterior a la cirugía, fuera del ámbito del quirófano, y exige al evaluador un ejercicio de memoria.
- El evaluador no tiene definido con claridad qué criterios debe evaluar, ni cómo debe calificar, y por lo tanto termina apelando a la intuición.

Bajo estas circunstancias la evaluación que resulta de la aplicación de este método queda fuertemente influenciada por las relaciones intersubjetivas y hace que el resultado esté sesgado por los rasgos de personalidad y los vínculos de afectividad más que basado en criterios objetivos, y por lo tanto es muy poco confiable.

Este hecho guarda relación con la vigencia hasta la actualidad de un modelo clásico de formación en cirugía introducido por Sir William Halsted en el año 1898 en el John Hopkins Hospital de los Estados Unidos de Norteamérica. Dicho modelo, diseñado a semejanza del sistema alemán, estaba basado en una relación maestro-aprendiz y en la adquisición de responsabilidades graduadas y progresivas. El entrenamiento estaba regido por el concepto de “ver y hacer”, es decir, el aprendiz trabajaba a la par de su maestro para asegurar la progresión de sus habilidades por imitación. La solvencia del aprendiz estaba determinada por el maestro de acuerdo a su

desempeño. Esta evaluación estaba basada muchas veces en rasgos de la personalidad más que en criterios objetivos (27, 28).

Sin embargo, los avances en las teorías de la educación, y las presiones crecientes en la práctica profesional han llevado a cuestionar seriamente la confianza en este sistema para la adquisición de habilidades técnicas. La creciente complejidad de los casos hospitalarios, y el mayor énfasis en mitigar el error médico por los problemas médico legales, hace disminuir la posibilidad de los instructores para formar a los residentes. La limitación de la carga horaria para los residentes, legislada en algunos centros formadores del hemisferio norte, y el interés en optimizar el uso del quirófano han determinado una menor exposición de los médicos en formación a los procedimientos con pacientes en la sala de operaciones.

Surge por estos motivos la necesidad de un nuevo modelo de educación en cirugía basado en un programa estructurado de formación, en el cual el entrenamiento de las técnicas operatorias comienza fuera del quirófano, en un laboratorio, y la adquisición y evaluación de competencias se realiza con instrumentos que pretenden ser lo más objetivos posibles (27, 28, 29).

Kane (30) definió la **competencia** como: *“... el grado de utilización de los conocimientos, las habilidades y el buen juicio asociados a la profesión, en todas las situaciones que se pueden confrontar en el ejercicio de la práctica profesional”* es decir que la competencia representa la capacidad de un profesional de utilizar su buen juicio así como también sus conocimientos, habilidades, y actitudes para solucionar problemas complejos que se presentan en el ámbito de su actividad.

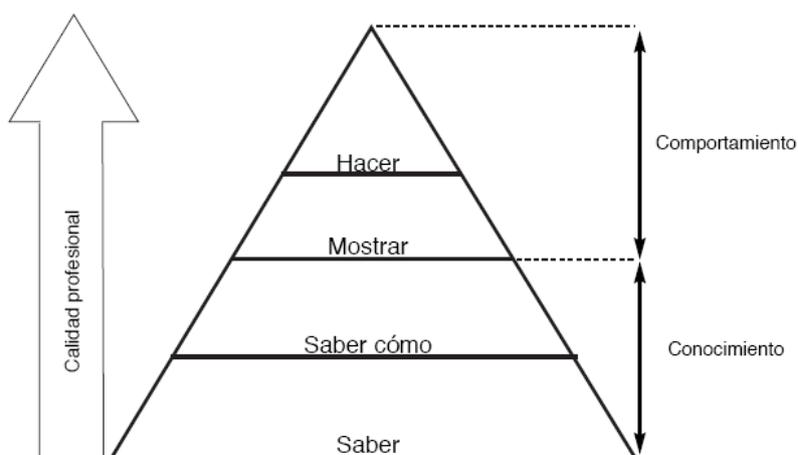
El concepto de competencia es una construcción conceptual compleja que se puede abordar de manera parcial a través de los diferentes aspectos mensurables que la constituyen: conocimientos, habilidades técnicas, actitudes, juicio, capacidad resolutiva frente a nuevos problemas, capacidad de comunicación, etc.

Hay muchos **motivos** por los cuales hay que **evaluar las competencias** tanto en estudiantes como en profesionales. Se pueden mencionar: 1) la evaluación como parte integrante del proceso educativo 2) para certificar la competencia (evaluación sumativa) 3) para ejercer retroalimentación sobre el proceso (evaluación formativa) 4) para evaluar los programas de formación 5)

para confirmar los valores y los estándares de la profesión y 6) por la responsabilidad e imputabilidad frente a la sociedad.

En forma análoga Beard (31) enumera diversas razones para evaluar las competencias, y afirma que el método de evaluación depende del motivo: 1) como refuerzo del aprendizaje a través de una retroalimentación apropiada 2) para determinar el nivel alcanzado por el cirujano en formación 3) para controlar el progreso 4) para garantizar la seguridad de los pacientes antes de comenzar una práctica no supervisada 5) para certificación.

Miller (32) ha desarrollado un modelo de competencia profesional representado por una pirámide compuesta de varios niveles (figura 3). En la base se encuentran los conocimientos (saber) sobre los que se apoya la competencia (saber como). A un nivel superior se encuentra el desempeño (mostrar como) y finalmente la acción en una práctica real (el hacer).



**Figura 3:** Pirámide de Miller. Modificado de Miller G.E. *The assessment of clinical skills/competence/performance. Academic Medicine (Supl.) 1990, 65: S63-S67.*

En 1996 Van Der Vleuten (33) propone un modelo teórico conceptual para definir la utilidad de una herramienta de evaluación de competencias profesionales que correlaciona distintas propiedades y que puede expresarse con la siguiente ecuación:

$$U = \left\{ \binom{(p)}{F} + \binom{(p)}{V} + \binom{(p)}{E} + \binom{(p)}{A} \right\} \times \binom{(p)}{C} \frac{1}{C}$$

Donde U: es la utilidad del instrumento; F: la fidelidad o confiabilidad; V: la validez; E: el impacto educativo; A: la aceptabilidad; C: el costo; y (p) un índice de contribución parcial de cada variable.

La **confiabilidad** puede definirse como la expresión cuantitativa de la **reproductibilidad** con la que un instrumento mide el mismo atributo. Hace referencia a la precisión del test. Es decir, en que medida el test genera los mismos resultados si se repite en dos sucesivas oportunidades bajo las mismas circunstancias, en este caso, sin la modificación propia del aprendizaje entre las tomas (34, 35). Se representa con el coeficiente alfa de Cronbach. Cuando las evaluaciones son llevadas a cabo por más de un observador se aplica también el concepto de **confiabilidad inter-evaluadores**, que mide el grado de concordancia entre dos o más evaluadores que utilizan el mismo instrumento (36).

La **validez** es un concepto complejo que hace referencia a la capacidad del instrumento de evaluación de medir o no lo que se propone como objetivo. Puede desglosarse en distintos subconceptos: validez **de constructo** hace referencia a la medida en la cual el examen puede diferenciar distintos niveles de experiencia. Validez **de contenido** refiere si el objeto o campo de conocimiento que pretende ser evaluado es medido por el instrumento. Por ejemplo: si se pretende evaluar aspectos técnicos y se terminan evaluando conocimientos, el instrumento de evaluación carece de validez de contenido. Validez **concurrente** es el grado en el cual los resultados de la evaluación se correlacionan con el gold Standard para ese campo del conocimiento. Validez **de superficie** es el grado en el cual el instrumento tiene el aspecto de situaciones reales. Validez **predictiva** la habilidad del examen de predecir el desempeño a futuro (13, 32, 37, 38, 39).

En febrero del año 1999 una comisión asesora del **ACMGE** (Accreditation Council for Medical Graduate Education) en Estados Unidos de Norteamérica definió 6 competencias generales que forman parte de todos los programas de educación de postgrado. Estas competencias generales son: 1)

cuidado del paciente 2) conocimiento médico 3) profesionalismo 4) práctica profesional acorde al sistema 5) aprendizaje y progreso basado en la práctica 6) habilidades interpersonales y de comunicación. En la competencia 1 (cuidado del paciente) se incluye la habilidad técnica (40).

El **Colegio Real de Cirujanos del Reino Unido** (UK) introdujo un programa de formación en cirugía basado en competencias: ISCP (Intercollegiate Surgical Curriculum Programme) (41) que define como competencias esenciales en cirugía: 1) conocimientos 2) destrezas 3) criterio de juzgamiento y 4) comunicación (27, 31).

En resumen, el conocimiento actual sobre el tema pone en evidencia falencias de peso en los métodos de evaluación de las habilidades técnicas vigentes en la actualidad. Los métodos tradicionales vigentes en la mayoría de los centros formadores de nuestro medio carecen de validez y confiabilidad en relación a la destreza quirúrgica.

La formación quirúrgica dentro del marco de un programa formal estructurado de residencias médicas requiere una retroalimentación objetiva constante de la adquisición de habilidades técnicas para corrección y reorientación, y para ello precisa de instrumentos de evaluación objetivos.

Existen métodos validados de evaluación objetiva que podrían comenzar a utilizarse para evaluar las habilidades quirúrgicas con fines formativos y de certificación. Aunque se requieren futuras investigaciones antes de ser utilizados para evaluaciones sumativas y de revalidación (42, 43, 44, 45, 46, 47, 48).

De lo expuesto anteriormente se ponen en evidencia diversas situaciones problemáticas. Por un lado, el escaso desarrollo del campo de la evaluación del desempeño técnico, particularmente en los ámbitos académicos donde se forman los futuros cirujanos. Por otro lado, la carencia de herramientas válidas y confiables para la evaluación objetiva de la técnica quirúrgica y, finalmente, la importancia asignada por médicos y público general a la adquisición de las habilidades técnicas.

En función de lo presentado se plantean los siguientes **OBJETIVOS** en esta investigación:

- Transformar el carácter efímero del acto quirúrgico en otro perdurable e inalterable en el tiempo, mediante la implementación de tecnología de registro y almacenamiento de imágenes de las cirugías, de forma sistemática.
- Elaborar un instrumento de evaluación basado en criterios objetivos y cuantificables, que sea aplicable a la ejecución específica de dos procedimientos quirúrgicos: la colecistectomía videolaparoscópica y la hernioplastia inguinal con técnica de Lichtenstein.
- Evaluar confiabilidad y validez del instrumento.

La **HIPÓTESIS** que plantea este proyecto de trabajo es que la explicitación y la enunciación de criterios objetivos de calificación, como así también la cuantificación en la ejecución del procedimiento quirúrgico permitirá la elaboración de un instrumento con confiabilidad y validez aceptables, para lograr una evaluación objetiva de los aspectos técnicos operatorios.

## Capítulo 2

# MATERIALES Y MÉTODOS

## **2. 1. Selección de los procedimientos quirúrgicos**

Para la ejecución de esta investigación se eligieron dos procedimientos quirúrgicos: la colecistectomía videolaparoscópica según posicionamiento norteamericano (50, 51, 52) y la hernioplastia inguinal con técnica de Lichtenstein (53, 54, 55, 56) (técnica de reparación de la pared posterior del conducto inguinal “libre de tensión”, en la cual se coloca una prótesis de refuerzo de polipropileno, luego de la reducción del saco herniario).

La selección de estos procedimientos está fundamentada en la alta prevalencia que tienen en nuestro medio las patologías que motivan las intervenciones quirúrgicas, y al hecho de que los cirujanos incluidos en los programas de formación de nuestro servicio participan en ellos de manera habitual desde el comienzo mismo de su experiencia y cumpliendo distintos roles: inicialmente como ayudantes, y luego como cirujanos. Esto permite poner a consideración la aplicación del instrumento en todo el espectro de la curva de aprendizaje.

## **2. 2. Grupos de comparación. Criterios de inclusión.**

Para el análisis se constituyeron 3 grupos de comparación teniendo en cuenta las tres etapas de la teoría de adquisición de habilidades motoras propuesta por Fitts y Posner. El criterio de inclusión fue el nivel de entrenamiento al momento de la ejecución de cada uno de los procedimientos. El **Grupo 1**, incluyó a los cirujanos con experiencia inicial, definida como la ejecución de los primeros diez procedimientos. El **Grupo 2**, comprendió a los cirujanos con experiencia intermedia, es decir aquellos que cuentan en su haber una cifra superior a diez procedimientos pero que se encuentran aún en etapa de formación, y el **Grupo 3** integrado por los cirujanos expertos, que incluye a aquellos médicos con grado de especialista, pertenecientes al staff permanente del servicio de cirugía. En ningún caso fue posible la identificación del paciente operado ni del cirujano evaluado, garantizando de este modo en todo momento el anonimato.

### **2. 3. Registro y almacenamiento de imágenes**

Las imágenes de las cirugías se registraron en tiempo real y sin captura de audio, desde el inicio hasta el final del procedimiento. Para las colecistectomías laparoscópicas se utilizó una placa de captura de video marca Pinnacle® Studio MovieBox HD versión 14, que fue conectada al procesador de imágenes de la endocámara del equipo de videolaparoscopia. Para las hernioplastias inguinales, una videocámara marca SONY MiniDV/DVcam, modelo DSR-PD 170P, con norma PAL y chip de 3CCD montada sobre un trípode ubicado detrás del cirujano, con angulación picada y con un plano detalle sobre el campo de trabajo. Las imágenes capturadas fueron bajadas a un soporte físico tipo DVD (.m2p) sin ningún tipo de edición posterior.

### **2. 4. Elaboración del instrumento de evaluación**

Con el propósito de conferir la máxima objetividad a la evaluación se elaboró un instrumento específico para cada procedimiento, que permitió una calificación numérica en función del desempeño técnico del cirujano. La calificación es aplicable en cada uno de los cinco pasos de la técnica quirúrgica que fueron seleccionados por ser considerados críticos o de mayor importancia.

Cada paso de la técnica seleccionado para ser evaluado fue claramente definido. Su ejecución fue cuantificada mediante una escala de Likert con un puntaje de 1 a 5, y con descriptores explícitos para los extremos y el medio, donde la calificación “ 1 “ corresponde a un desempeño inadecuado, “ 3 “ al desempeño intermedio y la “ 5 ” a la ejecución sobresaliente.

Para la confección del instrumento de evaluación específico se analizó minuciosamente la técnica elegida y se enumeró en orden la secuencia de pasos a ejecutar en el procedimiento. Se da respuesta de este modo a “QUÉ” es lo que se debe hacer en el procedimiento (Tabla 2).

**Tabla 2:** Enumeración de la secuencia ordenada de pasos a ejecutar en una colecistectomía laparoscópica. “el QUÉ”

1. Colocación de trocar óptico mediante minilaparotomía umbilical (técnica abierta).
2. Insuflación del abdomen: creación de neumoperitoneo.
3. Colocación de trócares operadores bajo visión directa.
4. Posicionamiento del paciente: Trendelenburg invertida y lateralización a izquierda.
5. Exposición de vesícula biliar:
6. Identificación de vía biliar principal, y de la unión bacinete cístico.
7. Apertura de serosa que recubre el triángulo de Calot y disección de estructuras que se encuentran en el mismo.
8. Disección y jalonamiento del conducto cístico y de la arteria cística.
9. Colocación de clips proximales y distales sobre conducto y arteria cística, y sección de las estructuras.
10. Disección del lecho vesicular de cuello a fondo.
11. Control de hemostasia y extracción de trócares bajo visión directa.

Posteriormente se seleccionaron de la lista 5 pasos que fueron considerados fundamentales o críticos en la ejecución del procedimiento, y se explicitó el modo en que cada uno de ellos debería ser ejecutado correctamente, teniendo en cuenta las referencias de seguridad, el direccionamiento de las maniobras, la ubicación de instrumental y otras recomendaciones, etc. Se dio respuesta al “CÓMO” de la ejecución (Tabla 3).

**Tabla 3:** Recomendaciones para la ejecución de la maniobra “Exposición de vesícula biliar”. “el CÓMO”

#### **Exposición** de vesícula biliar:

Maniobras de **tracción** que permiten ejecutar la colecistectomía con una correcta visualización de estructuras y con seguridad. Reviste particular importancia la **tracción en sentido cefálico del fondo vesicular y lateral del bacinete** ya que permite desplazar el eje del conducto cístico por fuera del eje del conducto colédoco y de esta forma previene las lesiones inadvertidas de la vía biliar principal.

Y finalmente se describió con claridad cuál sería, o más precisamente cómo se observaría, el desempeño insuficiente, el regular y el sobresaliente para cada una de las maniobras seleccionadas en el instrumento. Es decir se cuantificó numéricamente la ejecución de cada paso: “el CUÁNTO” (Tabla 4).

**Tabla 4:** Calificación numérica del desempeño en la maniobra "Exposición de vesícula biliar. Uso de mano no hábil". "el CUÁNTO"

1 Tracción Incorrecta	2	3 Exposición Insuficiente	4	5 Exposición Apropriada
Tracción cefálica en vez de lateral del bacinete, o falta de tracción.  Pinza izquierda manejada por el instructor.		Pobre tracción del fondo, insuficiente lateralización del bacinete  Poca utilidad de la pinza de mano izquierda: vesícula tomada pero "aplastada" contra el lecho		Tracción del fondo vesicular en sentido cefálico, y del bacinete en sentido lateral  Uso fluido de pinza de mano izquierda para exponer lateral derecho e izquierdo.

El resultado final de este proceso fue la creación de un instrumento de evaluación específico para colecistectomía videolaparoscópica y otro para hernioplastia inguinal (Anexos 3 y 4)

Con la finalidad de conferirle mayor robustez a la validación del instrumento propuesto, se utilizó la escala de evaluación global de Reznick, conocida como OSATS (Objective Structured Assessment of Technical Skills) a cada uno de los procedimientos evaluados. Esta escala ha tenido una amplia aceptación y está correctamente validada por numerosos estudios (Tabla 5).

## **2. 5. Evaluación del material**

Para cumplir el rol de evaluadores se convocaron a dos cirujanos con vasta experiencia en la docencia y formación de médicos residentes del ámbito de la universidad, pertenecientes al Servicio de Cirugía N° I del Hospital Nacional de Clínicas de la Universidad Nacional de Córdoba. Se concretaron tres reuniones de información y entrenamiento con cada uno de ellos por separado, donde se procedió a explicar con detalle su rol y la tarea, y se brindaron ejemplos de las calificaciones numéricas propuestas en función del desempeño, mediante la visualización de fragmentos de filmaciones en cada caso.

Se entregó a cada uno de ellos la información necesaria para la calificación, contenida en un póster con los siguientes datos: características generales de los grupos a evaluar definidos en función de la teoría de Fitts y Posner, tabla de evaluación global de R. Reznick (OSATS), tabla de evaluación

específica para colecistectomía laparoscópica y tabla de evaluación específica para hernioplastia inguinal. Los evaluadores no tuvieron límite de tiempo para evaluar. Las filmaciones se pudieron observar todas las veces que fueran consideradas necesarias, teniendo la posibilidad de ser adelantadas o atrasadas a voluntad.

El cegamiento de la muestra fue realizado colocando los DVDs de las filmaciones en sobres de papel "ad hoc" sin ventana ni identificación junto con una copia de la grilla para completar con los datos de la evaluación. Estos elementos fueron incluidos en un sobre de mayor tamaño con un número correspondiente al orden de la secuencia. La muestra se randomizó por sorteo consiguiendo de este modo una secuencia aleatoria de los procedimientos para calificar.

## **2. 6. Análisis estadístico**

Los resultados obtenidos en el presente trabajo fueron expresados como Término Medio  $\pm$  Error Estándar de la Media (TM  $\pm$  ESM). Para el tratamiento estadístico de los mismos se realizaron análisis de la varianza no paramétricos (pruebas de Kruskal-Wallis) para comparar los tres grupos en función de los puntajes adjudicados por los evaluadores. Cuando el valor  $p$  fue menor al nivel de significación ( $\alpha$ ) de 0.05 se rechazó la hipótesis de igualdad de los grupos, y se concluyó que hay diferencias estadísticamente significativas entre los mismos.

Para la determinación de la confiabilidad del test se aplicó el Coeficiente alfa de Cronbach. Valores superiores a 0,80 aseguran una correcta consistencia interna de la escala y confieren la característica de confiable al instrumento de evaluación.

Se utilizó el programa estadístico Infostat (Infostat 1,1 Versión 2000, Grupo Infostat, Facultad de Agronomía, Universidad Nacional de Córdoba, Córdoba, Argentina).

**Tabla 5:** Escala de evaluación global OSATS de Reznick. Objective Structured Assessment of Technical Skills.

1. Respeto por los tejidos				
1	2	3	4	5
Aplica frecuentemente una fuerza innecesaria sobre los tejidos o causa daño por el uso inapropiado del instrumental		Tiene manejo cuidadoso de los tejidos pero ocasionalmente causa daño inadvertido	Tiene manejo adecuado y consistente de los tejidos con un mínimo daño de los mismos	
2. Tiempo y Movimiento				
1	2	3	4	5
Muchos movimientos inadecuados		Hay una relación eficiente tiempo/movimiento pero hace movimientos innecesarios	Hay clara economía de movimientos Eficiencia máxima	
3. Manejo y Conocimiento del instrumental				
1	2	3	4	5
Movimientos dudosos y torpes Uso inapropiado		Uso competente del instrumental con torpezas ocasionales	Movimientos fluidos con el instrumental. No hay movimientos torpes	
4. Flujo de la cirugía				
1	2	3	4	5
Detenimiento frecuente, impresiona con inseguridades del próximo paso		Hay algo de planificación hacia delante, progresión razonable de la cirugía	Curso claramente planeado, flujo sin esfuerzo	
5. Conocimiento del procedimiento específico				
1	2	3	4	5
Conocimiento deficiente. Requiere instrucciones específicas en la mayoría de los pasos		Conocimiento de los pasos importantes del procedimiento	Demuestra familiaridad con todos los pasos del procedimiento	
6. Utilización de los asistentes (si aplicable)				
1	2	3	4	5
Ubica inadecuadamente a los asistentes o no los utiliza		Uso apropiado de los asistentes durante la mayoría del tiempo	Uso estratégico de los ayudantes para la máxima ventaja todo el procedimiento	

## **Capítulo 3**

# **RESULTADOS**

Desde Abril del año 2010 hasta Diciembre del 2012 se filmaron 67 procedimientos quirúrgicos de carácter electivo en el ámbito del quirófano central del Hospital Nacional de Clínicas (Facultad de Ciencias Médicas, Universidad Nacional de Córdoba) Se contabilizaron 36 colecistectomías videolaparoscópicas y 31 hernioplastias inguinales, realizados en su totalidad por el personal perteneciente al Servicio de Cirugía Nº I de dicha institución. Como cada procedimiento fue calificado por dos evaluadores de manera independiente, el número total de observaciones fue: 72 colecistectomías videolaparoscópicas y 62 hernioplastias inguinales. (*Tabla 6*)

**Tabla 6:** Distribución de procedimientos según grupos de ejecución y número total de evaluaciones.

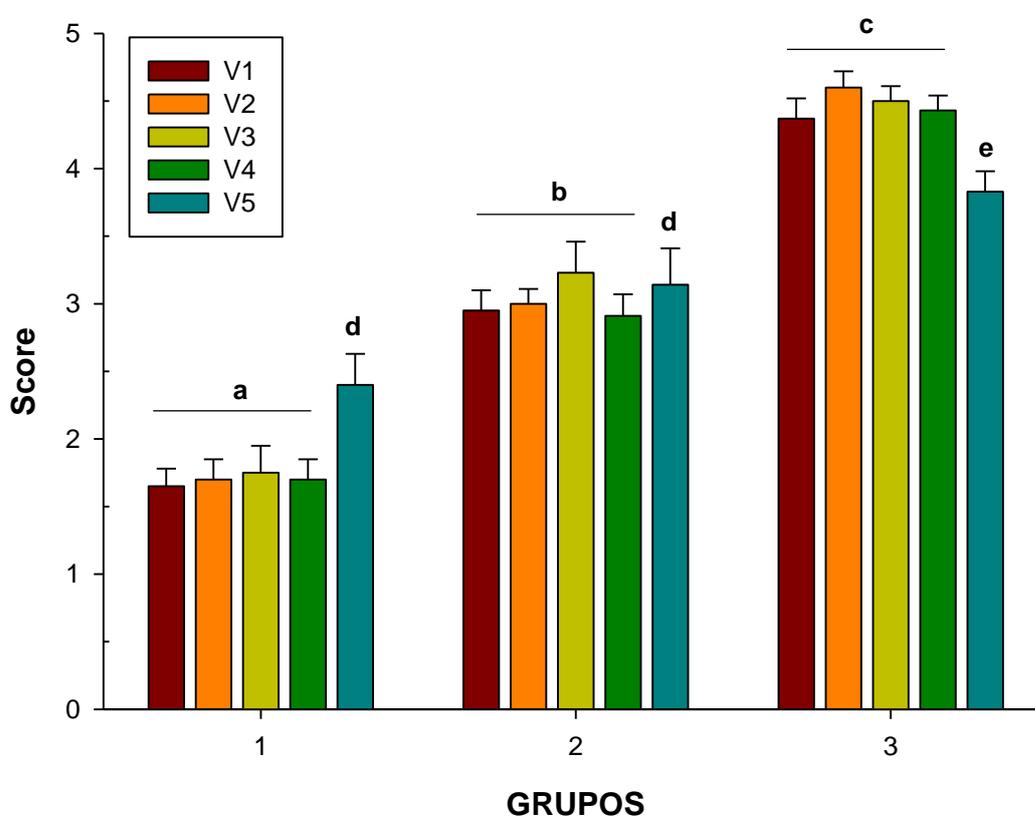
<b>Grupo</b>	<b>Colecistectomía videolaparoscópica</b>	<b>Hernioplastia inguinal</b>
I	10	10
II	11	10
III	15	11
<b>Total de casos</b>	n = 36	n = 31
<b>Total de evaluaciones</b>	n = 72	n = 62

El tratamiento estadístico de los resultados permitió analizar los siguientes aspectos:

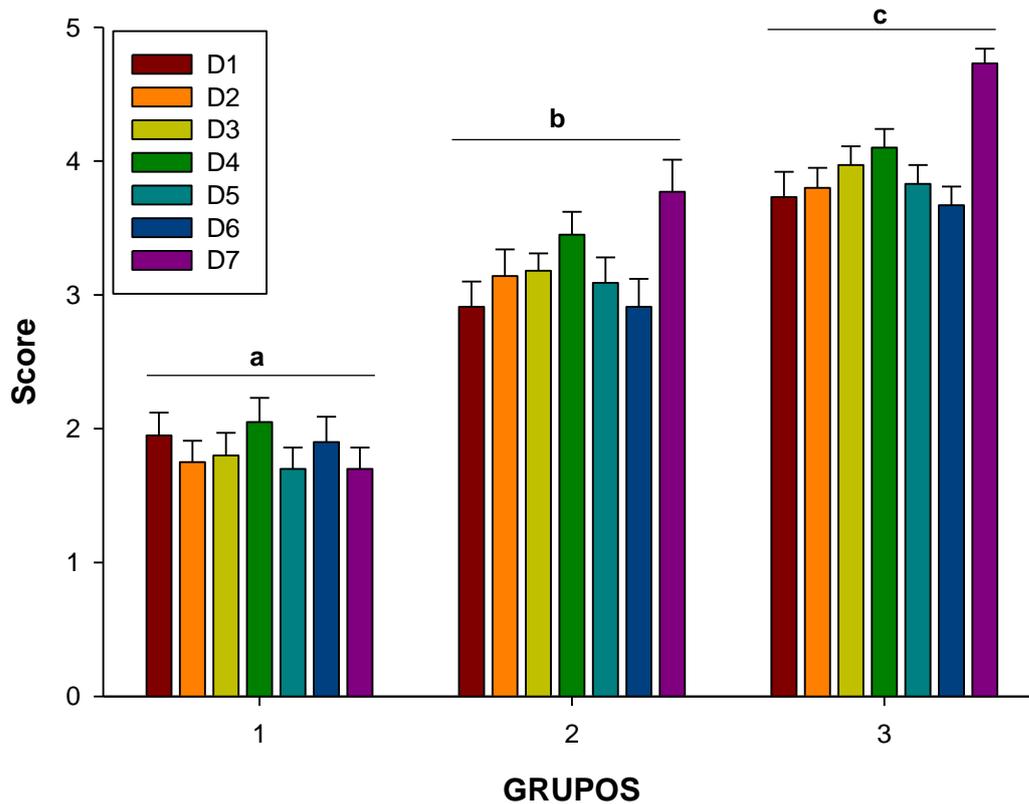
1. Comparación entre los grupos (validez de constructo)
2. Análisis de los evaluadores (confiabilidad inter evaluador)
3. Comparación entre los instrumentos específicos y el goldstandar (validez concurrente)
4. Determinación de la confiabilidad.

## 1. COMPARACIÓN ENTRE GRUPOS

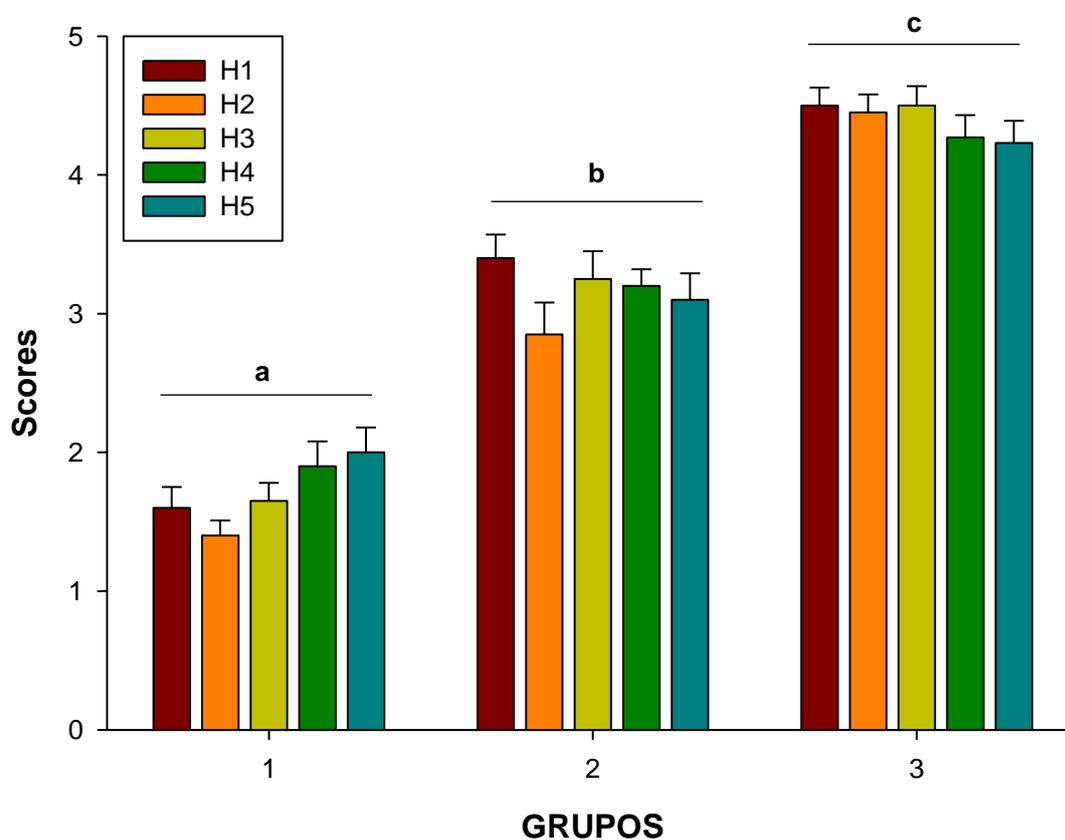
Con la finalidad de comparar los grupos entre sí se realizó un análisis de varianza no paramétrico (prueba de Kruskal-Wallis) a las calificaciones numéricas adjudicadas por los dos evaluadores, y luego la prueba “a posteriori” que hace referencia a la diferencia entre los grupos. Valores de  $p$  menores de 0,05 fueron considerados significativos.



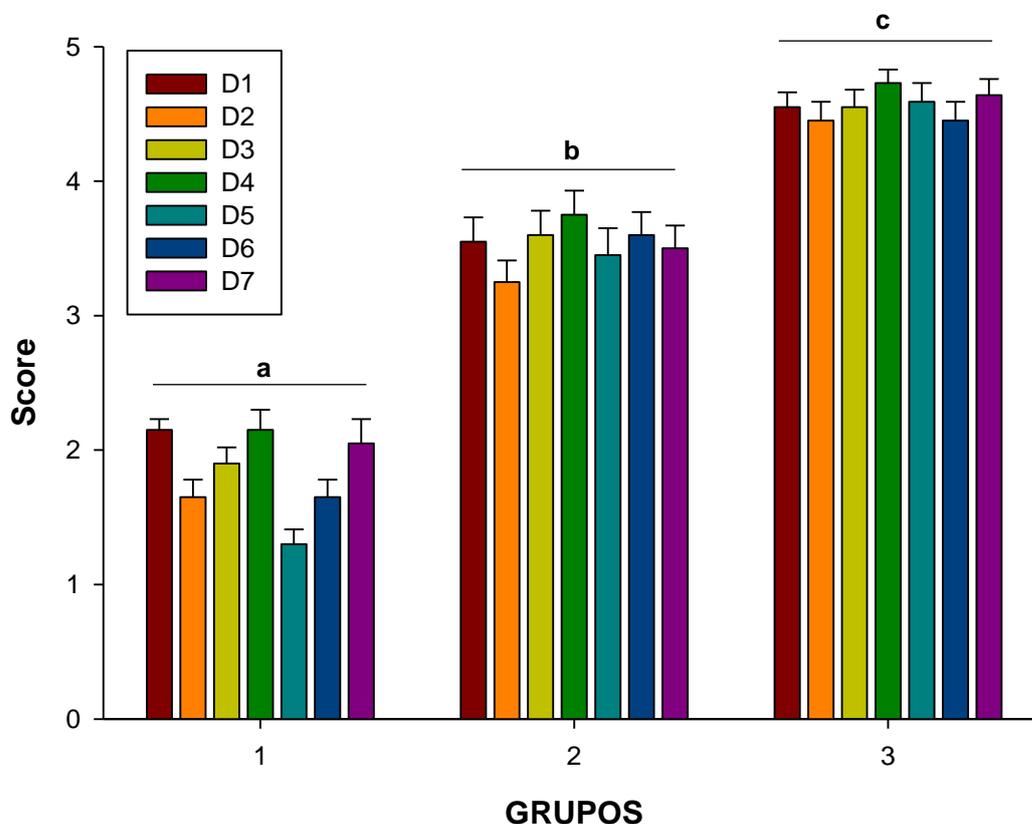
**Figura 4:** Evaluación objetiva de competencias técnicas en **Colecistectomía Laparoscópica** mediante la aplicación de un **nuevo instrumento específico** propuesto para el procedimiento, presentada por grupos de experiencia. Grupo 1 (principiantes):  $n=20$ ; Grupo 2 (experiencia intermedia):  $n=22$ ; Grupo 3 (expertos):  $n=30$ . **V1:** exposición de vesícula biliar. Uso de mano no hábil; **V2:** Disección del Triángulo de Callot **V3:** Disección, ligadura y sección del conducto cístico **V4:** Identificación de la arteria cística, ligadura y sección **V5:** Disección del lecho vesicular. **a vs. b vs. c=  $p < 0,0001$ ; d vs. e=  $p < 0,0002$ .**



**Figura 5:** Evaluación objetiva de competencias técnicas en **Colecistectomía Laparoscópica** mediante la aplicación del **instrumento de evaluación global** propuesta por R. Reznick (OSATS), presentada por grupos de experiencia. Grupo 1 (principiantes): n=20; Grupo 2 (experiencia intermedia): n=22; Grupo 3 (expertos): n=30. **D1:** Respeto por los tejidos **D2:** Tiempo y movimientos **D3:** Manejo del instrumental **D4:** Conocimiento del instrumental **D5:** Flujo y planificación de la cirugía **D6:** Rol de los ayudantes **D7:** Conocimiento del procedimiento específico. **a vs. b vs. c= p <0,0001.**



**Figura 6:** Evaluación objetiva de competencias técnicas en *Hernioplastia Inguinal* mediante la aplicación de un **nuevo instrumento específico** propuesto para el procedimiento, presentada por grupos de experiencia. Grupo 1 (principiantes):  $n=20$ ; Grupo 2 (experiencia intermedia):  $n=20$ ; Grupo 3 (expertos):  $n=22$ . **H1:** Acceso al conducto inguinal y jalonamiento del cordón; **H2:** Exploración y disección de los elementos del conducto inguinal **H3:** Tratamiento del saco herniario **H4:** Ajuste del orificio inguinal profundo **H5:** Inserción y fijación de prótesis. **a vs. b vs. c=  $p < 0,000$ .**



**Figura 7:** Resultados de la evaluación objetiva de competencias técnicas en **Hernioplastia Inguinal** mediante la aplicación del **instrumento de evaluación global** propuesto por R. Reznick (OSATS), presentada por grupos de experiencia. **a vs. b vs. c = p < 0,0001**. Grupo 1 (principiantes): n=20; Grupo 2 (experiencia intermedia): n=22; Grupo 3 (expertos): n=30. **D1:** Respeto por los tejidos **D2:** Tiempo y movimientos **D3:** Manejo del instrumental **D4:** Conocimiento del instrumental **D5:** Flujo y planificación de la cirugía **D6:** Rol de los ayudantes **D7:** Conocimiento del procedimiento específico

## **2. ANÁLISIS DEL COMPORTAMIENTO DE LOS EVALUADORES**

Se realizó un análisis de la varianza no paramétrico (pruebas de Kruskal-Wallis) para comparar entre sí las calificaciones adjudicadas por los dos evaluadores en cada una de las variables consideradas. Cuando el valor  $p$  fue menor al nivel de significación ( $\alpha$ ) de 0.05 se rechazó la hipótesis de igualdad de los evaluadores y se concluyó que hay diferencias estadísticamente significativas entre ellos. Cuando el valor  $p$  fue mayor a 0.05, se concluyó que los evaluadores tuvieron igual comportamiento para la adjudicación de los puntajes.

No se obtuvieron diferencias significativas en ninguna de las variables analizadas, tanto de los instrumentos específicos para los procedimientos, como en el instrumento global, con excepción de la variable D7 (conocimiento del procedimiento específico) del instrumento de evaluación global, aplicado a las colecistectomías laparoscópicas en donde el valor de  $p = 0,037$ . Es decir que esta fue la única variable donde no hubo coincidencia en las calificaciones asignadas por los evaluadores.

## **3. COMPARACIÓN ENTRE INSTRUMENTO ESPECÍFICO Y GLOBAL**

### **3. 1. Hernioplastia Inguinal**

No se encontraron diferencias significativas entre las evaluaciones provenientes del instrumento específico y del global ( $n= 62$ ;  $p = 0.23$ )

### **3. 2. Colecistectomía Laparoscópica**

No se encontraron diferencias significativas entre las evaluaciones provenientes del instrumento específico y del global ( $n= 72$ ;  $p = 0.51$ )

## **4. ANÁLISIS DE LA CONFIABILIDAD ( $\alpha$ de Cronbach)**

Para la evaluación de la confiabilidad se utilizó el coeficiente  $\alpha$  de Cronbach. Valores por encima de 0,80 revelan datos a favor de la consistencia interna de la tabla.

El coeficiente  $\alpha$  de *Cronbach* calculado para el instrumento que evalúa de manera específica la colecistectomía laparoscópica arrojó un valor de 0.93, y en el caso del instrumento diseñado para la hernioplastia inguinal fue de 0.95. Así mismo, el instrumento de evaluación global obtuvo un coeficiente de 0.94 y 0.97 cuando fue aplicado a los mismos procedimientos, respectivamente.

## Capítulo 4

# DISCUSIÓN

Tal como se mencionó en el Capítulo 1 Introducción, La ecuación de Van Der Vleuten identifica y relaciona entre sí los factores que definen la utilidad de una herramienta de evaluación de competencias. Estos son: confiabilidad, validez, impacto educativo, aceptabilidad, y costo (33) y son los parámetros que hemos utilizado de guía para el análisis de y la calificación de nuestra propuesta.

La lectura analítica de nuestros resultados luego del tratamiento estadístico nos permite afirmar que el instrumento de evaluación específico que ha sido propuesto para ser aplicado en colecistectomías laparoscópicas y en hernioplastias inguinales presenta **validez y confiabilidad**.

La diferencia estadísticamente significativa que se obtuvo entre las calificaciones respecto al desempeño técnico de los integrantes de los tres grupos comparados hace referencia al concepto de **validez de constructo**. Esto significa que el instrumento tiene probada capacidad para discernir entre diversos niveles de entrenamiento. Creemos que este hecho constituye uno de los hallazgos de mayor importancia en esta investigación ya que implica la superación de la puesta a prueba del valor calificador y clasificador de la herramienta que hemos propuesto. La posibilidad de diferenciar niveles de desempeño le confiere trascendencia formativa, educativa, y de certificación al uso de este tipo de instrumentos para la evaluación. Se debe destacar que en el ítem V5 (disección del lecho vesicular) del instrumento específico para colecistectomía laparoscópica no se encontraron diferencias entre el grupo 1 (experiencia inicial) y el 2 (experiencia intermedia). Sin embargo estos dos grupos, considerados en conjunto, sí se diferenciaron del grupo 3 (expertos). Creemos que este resultado tiene que ver con el hecho de que la disección del lecho vesicular tal como se enseña en nuestro servicio, mediante la utilización de una pinza llamada "hook" que es un gancho para electrocauterio, apela a habilidades que pueden ser rápidamente aprendidas con solvencia y que por lo tanto tienen una curva de aprendizaje con un ascenso más rápido que los otros pasos de una colecistectomía laparoscópica, haciendo quizás más difícil discernir los límites entre el grupo de experiencia inicial y el intermedio. Desde nuestro punto de vista, la capacidad discriminante del instrumento propuesto para este paso del procedimiento queda puesta en evidencia, aunque merece nuevas evaluaciones para determinar su verdadero potencial. Entendemos que

lo que requiere es una “mirada” más atenta de los evaluadores en este punto específico y un trabajo más orientado a discernir entre los grupos para este ítem.

Análogamente se puede afirmar que existe **validez de contenido**, ya que el objetivo propuesto fue la evaluación de las competencias técnicas, y la metodología descrita y aplicada para la creación del instrumento de evaluación propuesto está basada en un estudio exhaustivo de la técnica quirúrgica exclusivamente, o sea de los aspectos puramente procedimentales, excluyendo otros como el razonamiento clínico, la evolución postoperatoria, la morbimortalidad, la estadía hospitalaria, etc. por no estar considerados dentro de la categoría “técnica” y quedar por lo tanto fuera del propósito de este trabajo.

Por otro lado no se encontraron diferencias significativas entre los resultados devueltos por el instrumento específico de procedimiento propuesto por nosotros y el instrumento de evaluación global diseñado por Reznik, que está ampliamente validado y aceptado, y que hemos elegido como referencia. Este hecho remite a otro de los subconceptos de validez denominado **validez concurrente**.

Los valores obtenidos del Coeficiente alfa de Crohnbach por encima de 0.80 demuestran que nuestro instrumento es confiable. La **confiabilidad** es un atributo por el cual una herramienta de evaluación presenta la suficiente estabilidad y precisión como para devolver el mismo resultado frente a las mismas circunstancias. Esto quiere decir que los resultados que se obtendrán serán los mismos cuando se evalúe a diferentes cirujanos con mismo grado de entrenamiento, o bien frente al mismo cirujano sin mediar mejoras en su desempeño técnico.

En relación a la **aceptabilidad**, la predisposición y colaboración del personal médico interviniente en las cirugías, en especial la de los médicos residentes, fue muy buena. Del mismo modo respondió el resto del equipo quirúrgico ya que solo bastó una breve explicación inicial del motivo de las grabaciones y las necesidades técnicas en relación a la ubicación de la cámara para adaptarse a la modalidad de trabajo necesaria. La propuesta de trabajo fue instrumentada con mínimos recaudos que permitieron asegurar la adecuada implementación de los dispositivos de captura de imagen en el

quirófano, sin que esto haya significado un obstáculo o haya aparejado demoras en el desarrollo habitual de la actividad. Medidas tan simples como recubrir el trípode que sostuvo la cámara con una bata estéril, un pequeño desplazamiento lateral de la mesa de instrumental, y rectificaciones puntuales hechas a demanda de la posición de la cabeza del cirujano cuando obstruía el campo de filmación, fueron suficientes. Es decir, todas adaptaciones menores que no demandaron un esfuerzo significativo ni implicaron un desvío mayor del curso habitual. Se pudo trabajar en un clima de cooperación y buena voluntad.

El **costo** de la implementación de un instrumento objetivo de evaluación de competencias técnicas como el que se presenta en este trabajo de investigación dependerá del tipo de tecnología y equipamiento que disponga la institución médica. En la actualidad los equipos más nuevos de videolaparoscopia permiten la captura directa de las imágenes provenientes de la videocámara utilizando un dispositivo de almacenamiento portátil de datos que es conectado a un puerto de tipo USB que suele estar en el procesador de imágenes. Esta situación que podría considerarse la óptima, implica un costo elevado en relación al equipo de cirugía laparoscópica, pero una erogación mínima en dispositivos para almacenamiento de imágenes (pendrive, tarjetas SD, etc) En caso de no acceder a esta opción existe la posibilidad, como en nuestro caso, de emplear una placa capturadora de video y una computadora portátil que permitiría la digitalización y el almacenamiento de las imágenes de la cámara. Los costos en esta opción implican una erogación que cualquier institución médica e incluso un profesional médico interesado por cuenta propia, podría afrontar sin problemas. En el caso de las cirugías “abiertas” se plantea una situación análoga. Algunos quirófanos cuentan en sus instalaciones con un sistema de filmación que forma parte de la infraestructura hospitalaria, que puede tener una cámara incorporada a los dispositivos de iluminación, con orientación vertical sobre la camilla quirúrgica. En ese caso el costo quedaría reducido al dispositivo de almacenamiento de imágenes. Si el quirófano no contara con este tipo de dispositivos se debería adquirir una videocámara portátil y un sistema de soporte que permita la filmación de los procedimientos quirúrgicos. El costo en este caso será variable según las características de la cámara, sin embargo en función de nuestra experiencia sugerimos cualquiera que cuente con un chip 3 CCD, pero que sea de

pequeño tamaño y liviana, ya que facilita su ubicación para la filmación en posición “de picada” o cenital, y el eventual desplazamiento de la misma durante la cirugía.

Todas las consideraciones presentadas hasta aquí hacen referencia a los costos en términos económicos, sin embargo existe otro tipo de costo que no puede ser soslayado: el **recurso humano**. La implementación de un sistema de evaluación objetiva de competencias técnicas como el que se presenta en esta tesis implica designar y destinar personal médico calificado a dicha tarea, lo que presupone agregar un conjunto de nuevas actividades que demandarán tiempo y algún grado de esfuerzo complementarios, o quizás, menos probablemente, distraer parte del tiempo destinado a otras actividades habituales para este proyecto.

El equipo de profesionales a cargo de la propuesta deberá trabajar en una primera etapa en la **elaboración consensuada** del instrumento de evaluación mediante el análisis exhaustivo de la técnica, para definir cuáles son los aspectos claves del procedimiento que deberían ser puntualmente observados, y cómo se cuantificará luego el desempeño del cirujano. La siguiente etapa será destinada a la **puesta a prueba y el entrenamiento** de los evaluadores, que implica la observación y calificación de cirugías filmadas aplicando el instrumento creado para “aprender a mirar” y calificar. Nuestra sugerencia es que los profesionales que se involucren en esta tarea deben ser cirujanos con experiencia en la práctica quirúrgica, con cabal conocimiento de la técnica quirúrgica y que, fundamentalmente, compartan los mismos criterios de calidad en la ejecución de los procedimientos. A nuestro entender, la situación óptima sería que hayan sido formados en el mismo centro educativo. Finalmente y luego de la etapa de entrenamiento los evaluadores estarán capacitados para observar las filmaciones y calificar en función de los instrumentos elaborados.

Es importante presentar en este capítulo algunas consideraciones y reflexiones a cerca de dos conceptos a los cuales se ha hecho referencia permanentemente en nuestro trabajo y constituyen un eje central para el debate: evaluación, y objetividad.

En primer lugar, evaluar supone efectuar una lectura orientada sobre el objeto que se evalúa, en función de la cual el evaluador emite un juicio sobre la realidad. Dicho de otra manera, no existe una lectura directa de la experiencia. Hay siempre un proceso de interacción entre el evaluador y la realidad a evaluar. En ese proceso, en función de las propiedades de esa realidad, el evaluador construye el referente, es decir, aquello con relación a lo cual se va a efectuar la evaluación, aquello que le permitirá “pronunciarse sobre la realidad” que evalúa.

En la **construcción del referente** intervienen diferentes factores entre los que se pueden mencionar: las expectativas que conciernen al propio objeto evaluado, la concepción de evaluación que se sustenta, los propósitos que se derivan de ésta y el contexto decisional en que se inserta el proceso evaluativo.

**Tanto el objeto que se evalúa como el proceso de valoración son contruidos por el sujeto que evalúa.**

En la evaluación “no se capta nunca directa y totalmente la realidad examinada. Ésta es siempre objeto de un proceso previo de definición o de elección de los caracteres que permitirán aprehenderla. “no basta decir que la evaluación lleva a la realidad; todavía falta construir o adoptar los indicadores de la realidad considerada (57).

El proceso de construcción del objeto supone, entonces, proceder por **recortes sucesivos de la realidad, realizados conforme a criterios que pueden ser explicitados o bien tácitos**. Por esta razón los indicadores que se construyen **no cubren nunca la totalidad** de la realidad que se evalúa. Al evaluar siempre se introduce un esquema de lectura de la realidad que circunscribe, recorta, reduce lo observable.

“Con respecto al producto de la evaluación, también concierne al evaluador construir los resultados de la evaluación, construir los datos, ‘lo referido’ (con relación al referente) Ello significa relevar en la realidad los signos indicadores que den cuenta de la presencia de los criterios previamente establecidos” (58).

En otros términos, el **referente** puede definirse como el conjunto de normas o de criterios que sirven de grilla de lectura del objeto a evaluar. Constituye una especie de “ideal” o de “deber ser” y hace que la evaluación implique, siempre, una lectura orientada por el marco, por la óptica, por la

perspectiva de análisis privilegiada por parte del evaluador (ya sea éste un sujeto o un equipo). Por el contrario, **lo referido** está constituido por el conjunto de atributos observables que se recortan de un objeto de evaluación y que se consideran representativos de éste. Cabe aclarar también aquí que el evaluador, durante el proceso de evaluación, toma decisiones y construye tanto el referente como lo referido. Tomar decisiones significa, entre otras muchas cuestiones, recortar, seleccionar y plantearse desde qué marcos teóricos de referencia se organizará la lectura de la realidad a evaluar.

Lo anteriormente expuesto trae a consideración la existencia de referentes universales por los cuales el evaluador, de manera intencional o no, termina optando por alguna perspectiva o recorte teórico, o alguna toma de posición en particular en relación a la evaluación. Y por otro lado desnuda el hecho de que un programa o proyecto difícilmente pueda ser evaluado en toda su complejidad, ya que en realidad se puede evaluar mucho menos de lo que se pretende, y sin embargo se supone erróneamente que se ha abarcado al objeto de la evaluación casi en su totalidad.

Por otro lado creemos importante hacer algunas consideraciones sobre la **idea de objetividad** aplicada en el ámbito de la evaluación. La evaluación “objetiva” implica la existencia de una situación cercana a lo ideal con un sentido de valor francamente positivo, en donde el examen queda liberado de la influencia distorsiva de la subjetividad del evaluador, que sería la responsable de una valoración sesgada, imprecisa, injusta, e incluso muchas veces malintencionada. Siendo estos atributos de valor que habitualmente quedan vinculados a “lo subjetivo”. Esta sobrevaloración de lo objetivo contrapuesta a la carga de sentido negativo impuesta a lo subjetivo, queda en gran medida cuestionada por las consideraciones presentadas anteriormente, que hacen referencia a las generalidades de la evaluación.

El concepto de construcción del objeto que se evalúa y del proceso de evaluación, implica necesariamente una **participación protagónica de la subjetividad**, ya que supone un posicionamiento frente a la realidad por parte de un sujeto (entiéndase autoridad competente, un idóneo o persona encargada) dentro de un marco institucional, donde se ha definido de un perfil de egresado o de especialista en función de una necesidad o demanda. Este

perfil servirá de guía para la enunciación de los patrones de desempeño y las expectativas de rendimiento, los requerimientos mínimos para la aprobación, las condiciones para la certificación, el número, el tipo y el momento de evaluaciones, que quedará plasmado en un curriculum formal.

Un programa de formación de postgrado en cirugía general tendiente a la obtención del título de especialista en el área no es una mera enumeración aséptica y universalmente aplicable de contenidos teóricos y competencias. Muy por el contrario, está cargado de intencionalidad. Es el resultado final de una sucesión de definiciones y recortes de la realidad, que en última instancia son elecciones que implican priorizar determinados contenidos sobre otros, y que están condicionados por una demanda o necesidad en un contexto histórico particular. Todo esto dará como resultado una forma adecuada y deseable de práctica profesional. Es así como se modela un profesional desde una perspectiva determinada. Se prioriza un patrón de práctica y se eligen los medios para conseguirlo.

De un modo análogo, la investigación que motiva esta tesis es el resultado de una sucesión de recortes y definiciones dentro del área de conocimiento de la cirugía. La primera definición queda puesta en evidencia en el enunciado mismo del título. Se elige evaluar específicamente las competencias técnicas, es decir se delimita el estudio a los aspectos puramente procedimentales de la práctica quirúrgica, y se decide, intencionalmente, dejar por fuera a otras competencias como el conocimiento, el razonamiento clínico, la comunicación interpersonal, el profesionalismo, etc. Por otro lado, se seleccionan dos procedimientos para ser evaluados, se definen tres grupos de comparación en función de un sustento teórico, se elabora un instrumento que es específico para cada procedimiento en el que se seleccionan cinco puntos críticos de la técnica para ser observados, y se cuantifica el desempeño de un modo particular, etc. Las elecciones no son fortuitas ni responden al azar, sino son dirigidas y están impregnadas de una clara intencionalidad y orientación. Todos los fundamentos y los detalles forman parte de la información de los capítulos introducción y material y método.

Ahora bien, hechas estas consideraciones ¿dónde radicaría entonces la objetividad en la modalidad de evaluación propuesta? Los aspectos más

objetivos que quedan claramente identificados son: la enunciación explícita de los criterios de evaluación y la cuantificación del desempeño técnico. También podría mencionarse el anonimato, el cegamiento y la aleatorización de las muestras, y el hecho de contar con un registro de procedimientos que es perdurable en el tiempo. Es decir se debe entender el concepto como una **objetividad instrumental** vinculada con cuestiones del método de obtención de la información y las medidas (59).

Sin embargo, todos los criterios definidos como objetivos fueron contruidos por un sujeto. La interpretación y el uso de la medida no pueden situarse fuera de un contexto interpersonal, institucional, histórico y social. Por lo que el “pre”-supuesto valor agregado de equidad y justicia con que se vincula a la idea de objetividad termina resultando una entelequia. Aparecen en juego de este modo los aspectos cualitativos de la evaluación.

El origen de la carga de sentido positivo que “lo objetivo” le ha conferido a la evaluación como valor agregado se pone en evidencia a lo largo de su evolución histórica. Un breve repaso permitirá echar luz sobre el surgimiento de esta asociación y sus probables causas.

El planteo dilemático que ha generado la polarización entre **evaluación cuantitativa y evaluación cualitativa**, y las controversias derivadas, no deberían existir como tales ya que la calidad y la cantidad aluden a dimensiones diferentes de lo real. Las discusiones remiten tanto a cuestiones metodológicas como a las concepciones epistemológicas que las sustentan: la relación entre paradigma y método, la oposición entre subjetividad/objetividad como patrimonio de uno u otro, la relación entre validez y confiabilidad en cada uno de los métodos, el énfasis en el proceso o en los resultados, el carácter exploratorio e inductivo en un caso y confirmatorio y deductivo en otro, la posibilidad de establecer generalizaciones, la relación con el contexto de descubrimiento o de verificación, etc.

C. Gutiérrez hace referencia a dos paradigmas complementarios en el ámbito de la metodología de las ciencias sociales:

“Uno de los paradigmas correspondería al método heurístico; el otro, al método positivo. Los dos paradigmas, complementarios en el sentido de la disyunción, no en el de la conjunción, serían dos vías alternativas hacia el

dominio de la realidad. Los dos enfoques serían importantes y necesarios. Habría que usar los dos, especialmente por su tendencia a contrarrestar los posibles excesos de cada uno: el peligro de la inflación verbal o especulación sin garantía, por un lado; el peligro de depresión intelectual, anulación del impulso heurístico, por el otro. Sería bueno que todos tratáramos de cultivar ambos enfoques simultáneamente...” (60).

Como señalan Cook y Reichart en su obra *Métodos cuantitativos y cualitativos en investigación evaluativa*: “...para una comprensión completa de una evaluación, ésta tendría que realizar al menos tres tareas: comprobación, valoración del impacto y explicación causal. Se trata de una gama muy amplia de tareas que, para ser eficazmente atendidas, requerirán quizás el empleo de métodos tanto cualitativos como cuantitativos. Aunque no de un modo inevitable, puede suceder a menudo que la comprobación será realizada con mayor eficacia conforme a un método cualitativo, que la valoración del impacto se realice con mayor precisión mediante métodos cuantitativos, y que la explicación causal se obtenga de mejor manera a través del empleo conjunto de métodos cuantitativos y cualitativos” (61).

Es decir que la combinación flexible de ambos enfoques contribuye a corregir los sesgos que, inevitablemente, presentan los métodos cuantitativos y cualitativos cuando se los usa en forma separada, permite atender los múltiples propósitos de la investigación evaluativa con una variedad de alternativas, y finalmente, la potenciación mutua derivada de la complementariedad (62).

En este trabajo de investigación se pone claramente de manifiesto esa condición ambivalente, cuantitativa y cualitativa, que hace referencia a la doble vertiente de la que se nutre la evaluación. Lo cuantitativo es lo vinculado con la métrica, lo numérico, la obtención de una cifra y una medida calificatoria. Pretende con argumentos razonables ser objetivo. Se contrapone a la “interferencia” que supuestamente genera un vínculo intersubjetivo en la obtención o la lectura de las cifras. Y trata de llenar un vacío relacionado con la calificación de las habilidades técnico-procedimentales. Podría interpretarse como un modo de obtención de información, o una forma de abordar la lectura de la realidad. Lo cualitativo tiene que ver con la interpretación de la medida. Qué significado y valor se le otorga a la cifra, qué utilidad tiene como indicador, que destino se le dará a esa información, y qué consecuencias traerá. Si será

tenida en cuenta como parámetro para la certificación o como indicador del aprendizaje, si corresponde rectificar o reforzar algún aspecto particular detectado por el instrumento, si desde lo formal el candidato estará promovido al nivel inmediatamente superior o no. Todas son determinaciones que si bien están basadas en una cifra obtenida gracias a un método que pretende ser lo más objetivo posible, responden a necesidades y directivas institucionales, en marcos históricos claramente definidos, y a situaciones personales diversas.

Nuestro instrumento ha sido validado. Ha superado la prueba del poder calificador y discriminante entre los grupos con diferentes grados de entrenamiento. Tal como fue expuesto previamente presenta validez de constructo y confiabilidad. Lo cuantitativo tiene relación, precisamente, con la validez interna del instrumento. Mientras que los aspectos cualitativos de nuestro trabajo no quedan explicitados, sino más bien insinuados. Se abre un abanico de probables aplicaciones donde podría resultar de utilidad, que van desde lo estrictamente formativo, como feedback del proceso de aprendizaje para el estudiante o médico residente, lo sumativo para promoción de grado, y la certificación eventualmente, para la obtención del grado de especialidad normatizado por alguna institución profesional o educativa. Lo cualitativo hace referencia a la validez externa o de generalización.

Reznick (63) en el año 1993 reportaba los métodos disponibles para la evaluación de las competencias técnicas, con grados variables de confiabilidad y validez:

Los **exámenes de certificación** administrados por las instituciones formadoras o reguladoras al final de un período lectivo o luego de cumplir con el programa de formación, basan su evaluación en los conocimientos y las habilidades clínicas, y no evalúan la capacidad técnica del candidato (35). Constituyen el paradigma de la evaluación externa.

Los **registros de procedimientos** efectuados a lo largo del programa de formación (cuadernos de bitácora) son obligatorios en los programas de residencia de algunos lugares del mundo como el Reino Unido, y son requeridos al momento de las evaluaciones de fin de año o final de residencia, entrevistas de trabajo, etc. Sin embargo ha sido demostrado que es un registro

meramente indicativo de la ejecución de los procedimientos y no un reflejo de la habilidad operatoria, es decir carecen de validez (de constructo)

El **registro del tiempo operatorio** empleado para la ejecución de la cirugía no evalúa la calidad de la técnica, y es un parámetro muy poco confiable en evaluaciones de procedimientos en tiempo real debido a la influencia modificadora de numerosas eventualidades durante el curso del acto quirúrgico que no pueden ser previstas.

La **observación directa** es uno de los métodos que con mayor frecuencia se utiliza en nuestro medio. Tal como se realiza actualmente en el quirófano es subjetiva, presenta dificultades instrumentales relacionadas con la dificultad para la observación y la visión intermitente, e implica un ejercicio de memoria ya que la evaluación es realizada en una instancia posterior, fuera del quirófano, y luego que el procedimiento finalizó. Como la evaluación es global y no está basada en criterios específicos no resulta confiable. Como está influenciada por la subjetividad del observador tiene una escasa confiabilidad de reevaluación e interobservador.

Los **datos de morbilidad** con frecuencia son utilizados como marcadores sucedáneos del desempeño operatorio, sin embargo estos datos están fuertemente influenciados por factores dependientes de los pacientes, como la gran variabilidad biológica y los diferentes estadios de la enfermedad, y por otros factores no relacionados con el paciente: soporte institucional, acceso y disponibilidad de recursos tecnológicos, etc. Por lo tanto, frente a tal grado de heterogeneidad, la información carece de confiabilidad, y no puede reflejar verdaderamente el desempeño técnico.

Como consecuencia de los problemas de confiabilidad y subjetividad de los métodos mencionados previamente, se fueron desarrollado diversos **métodos objetivos** de evaluación de habilidades técnicas, entre los que se pueden mencionar:

Las **listas de cotejo** (“checklist” en idioma inglés) consisten en una enumeración de aspectos parciales constitutivos de un proceso, enunciados explícitamente y destinados a ser verificados. Pueden ser tenidas en consideración cuestiones estrictamente técnicas y procedimentales, como así también de otro orden: cognitivas, actitudinales, vinculadas a la comunicación y al liderazgo, etc. El tipo de respuesta que se espera responde a una lógica

binaria “presente/ausente”, “sí/no”, “logrado/no-logrado” en la mayoría de los casos (Tablas 7 y 8).

Como se puede observar en los ejemplos que se presentan a continuación, existe una enumeración explícita y exhaustiva de lo que se propone verificar. La claridad en los enunciados es justamente una de las fortalezas de este tipo de instrumentos ya que el evaluador no tendrá duda alguna a cerca de qué es lo que deberá constatar, y su mirada estará correctamente dirigida hacia un aspecto puntual y definido. Sin embargo la respuesta de tipo todo o nada y la ausencia de descriptores para lo correcto/incorrecto, presuponen que el evaluador es un individuo implícitamente idóneo y con la capacidad innata de desarrollar su tarea con solvencia, sin contemplar la posibilidad de la existencia de disparidad de criterios en la interpretación de las maniobras o los hechos.

**Tabla 7:** Lista de cotejo para la evaluación de una maniobra quirúrgica específica: ejecución de un punto y nudo laparoscópicos en un modelo animal. Tomado de Revista Argentina de Cirugía. Número Extraordinario. Relatos 2013. p.71.

Acción	Correcto = 1	Incorrecto = 0
1 Selección de instrumentos		
2 Utiliza la sutura adecuada		
3 Coloca la aguja bien en el porta y la introduce adecuadamente en el tejido		
4 Pasa del primer intento el tejido > 80%		
5 Recupera la aguja del tejido correctamente > 90%		
6 No daña los tejidos con las pinzas		
7 Hay una adecuada separación de puntos en la sutura		
8 Simétricos		
9 La aposición de los bordes evita su inversión		
10 Sutura sero-muscular > 80%		
11 Falla intento de nudo en < 20% intentos		
12 Nudo plano		
13 Nudo lateralizado		
14 Al menos tres lazos por nudo		
15 Corte de hilos a distancia adecuada		
16 Intento de extracción de aguja por trocar de 10 > 90%		
17 Tensión correcta en la sutura		
18 No hematomas peri-sutura		
19 No deja restos de hilos en el campo		
20 No pierde aguja		

**Tabla 8:** Modelo de lista de cotejo (checklist) para lobectomía tiroidea utilizado en el Hospital Alemán de Buenos Aires. Tomado de Revista Argentina de Cirugía. Número Extraordinario. Relatos 2013. p.72.

	Si	NO
<ol style="list-style-type: none"> <li>1. Paso a paso de la técnica quirúrgica</li> <li>2. Posiciona al paciente en la camilla de operaciones en decúbito dorsal</li> <li>3. Luego de la intubación del paciente por parte del anesthesiólogo realiza una cofia con un campo de tela chico para mantener el cabello fuera del campo operatorio</li> <li>4. Verifica que los electrodos estén por debajo de las tetillas para mantenerlos fuera del campo operatorio</li> <li>5. Coloca un realce a la altura de los hombros para hiper-extender el cuello</li> <li>6. Verifica que los brazos del paciente estén fijados al costado del cuerpo</li> <li>7. Realiza la antisepsia con Iodopovidona</li> <li>8. Coloca adhesivo biológico para fijar los campos de tela por debajo de las clavículas hasta los hombros y desde allí ascender sobre el trapecio de cada lado hasta el mentón</li> <li>9. Coloca adecuada de campos de tela y fijación con puntos a la piel del cuello</li> <li>10. Realiza una incisión cervical anterior arciforme de 4-5 cm, a dos traveses de dedo de la horquilla esternal con bisturí frío</li> <li>11. Incide el TCS y platisma con bisturí eléctrico</li> <li>12. Coloca tres pinzas de Allis en el borde superior del platisma seccionado y realiza el colgajo superior hasta el cartilago tiroides;</li> <li>13. Coloca tres pinzas de Allis en el borde inferior del platisma seccionado y realiza el colgajo inferior hasta las clavículas</li> <li>14. Coloca dos separadores angostos, uno superior y otro inferior</li> <li>15. Tracciona, junto con el primer ayudante, de los músculos pretiroideos con pinzas de De Bakey y realiza la apertura del rafe medio con bisturí eléctrico</li> <li>16. Coloca un separador angosto en el extremo superior y otro lateral del lado que inicia la disección</li> <li>17. Realiza la disección subcapsular del lóbulo tiroideo seccionando la riendas avasculares con bisturí eléctrico monopolar o bipolar</li> <li>18. Diseca el pedículo superior tiroideo. Liga y secciona o secciona con bisturí eléctrico bipolar</li> <li>19. Diseca el pedículo superior tiroideo. Liga y secciona o secciona con bisturí eléctrico bipolar</li> <li>20. Diseca la cara lateral del lóbulo tiroideo, e identifica las dos glándulas paratiroides y del nervio laríngeo recurrente al cual diseca hasta su entrada en la laringe</li> <li>21. Separa el lóbulo tiroideo de la tráquea controlando minuciosamente el nivel de la entrada del nervio en la laringe</li> <li>22. Incluye la pirámide de Lalouette en la resección, la cual separa con bisturí eléctrico monopolar</li> <li>23. Secciona la glándula tiroides a nivel de la vertiente entre el istmo y el lóbulo contralateral con bisturí eléctrico monopolar</li> <li>24. Envía la pieza de resección a anatomía patológica por congelación y espera el resultado</li> <li>25. Lava con solución fisiológica estéril y aspira sobre una gasa chica el líquido de lavado</li> <li>26. Controla la hemostasia con buena visión del lecho quirúrgico –utilizan separadores los ayudantes</li> <li>27. Cierra el rafe medio con un surget de Vicryl 3/0 comenzando desde el extremo superior dejando los últimos 2 cm sin cerrar</li> <li>28. Cierra el platisma con puntos separados o surget de Vicryl 3/0</li> <li>29. Cierra la piel con sutura intradérmica reabsorbible –Monocryl o Vicryl Rapide- 4/0 o 5/0</li> <li>30. Realiza una cura plana</li> </ol>		

Es destacable mencionar un ejemplo de checklist implementado por la OMS (Organización Mundial de la Salud) en el marco del programa “la cirugía segura salva vidas”, por su amplia difusión mundial y por la utilidad que ha tenido.

En octubre de 2004 la Organización Mundial de la Salud (OMS) creó la Alianza Mundial para la Seguridad del Paciente en respuesta a la Resolución 55.18 de la Asamblea Mundial de la Salud, en la que se instaba a la OMS y a los Estados Miembros a prestar la mayor atención posible al problema de la seguridad de los pacientes. La Alianza organiza programas que abarcan aspectos sistémicos y técnicos, y formula cada dos años diversos retos mundiales por la seguridad del paciente en todo el mundo. Cada desafío está destinado a impulsar el compromiso y la acción internacionales en un aspecto que suponga un área de riesgo importante para todos los estados miembros de la OMS. El primer desafío se centró en las infecciones relacionadas con la atención sanitaria. El tema elegido para el segundo reto mundial por la seguridad del paciente es la seguridad de las prácticas quirúrgicas.

Datos provenientes de 56 países demostraron que durante el año 2004 el número total de cirugías mayores fue estimado entre 187 y 281 millones de operaciones, o dicho en otros términos, una operación anual cada 25 personas vivas (64). En países industrializados la tasa de complicaciones mayores oscila entre un 3 a 22% de los procedimientos que requieren hospitalización, y la tasa de mortalidad es del 0,4 al 0,8% (65, 66) En estos estudios se determinó que casi la mitad de los eventos adversos eran prevenibles (67). En países en desarrollo la tasa de mortalidad asociada a cirugía mayor ha sido reportada entre un 5% y hasta un 10% (68, 69) La seguridad de los pacientes sometidos a cirugía se ha transformado en un problema de salud a nivel mundial.

El segundo Reto Mundial por la Seguridad del Paciente, llamado “La cirugía segura salva vidas” aborda la seguridad de la atención quirúrgica. La Alianza Mundial para la Seguridad del Paciente empezó a trabajar en este reto en enero de 2007.

El objetivo es mejorar la seguridad de la cirugía en todo el mundo, definiendo para ello un conjunto básico de normas que puedan aplicarse en todos los Estados Miembros de la OMS. Con este fin, se reunieron grupos de trabajo compuestos por expertos internacionales que revisaron la bibliografía y las experiencias del personal vinculado a la tarea asistencial en todo el mundo, y coincidieron en señalar cuatro áreas en las que se podrían realizar grandes progresos en materia de seguridad de la atención quirúrgica: 1) Prevención de

las infecciones de la herida quirúrgica, 2) Seguridad de la anestesia, 3) Seguridad de los equipos quirúrgicos y 4) Medición de los servicios quirúrgicos, creándose de este modo cuatro grupos de trabajo.

Los cuatro grupos de trabajo definieron diez objetivos esenciales que todo equipo quirúrgico debería alcanzar durante la atención quirúrgica. Esos objetivos se condensaron en una lista de verificación (checklist) de una página que los profesionales de la salud podrán utilizar para cerciorarse de que se cumplen las normas de seguridad.

**Tabla 9:** *Objetivos planteados por la Organización Mundial de la Salud en el Desafío Mundial "La Cirugía Segura Salva Vidas". OMS. 2007.*

Objetivo 1.	El equipo operará al paciente correcto en el sitio anatómico correcto.
Objetivo 2.	El equipo utilizará métodos que se sabe que previenen los daños derivados de la administración de la anestesia, al tiempo que protegen al paciente del dolor.
Objetivo 3.	El equipo se preparará eficazmente para el caso de que se produzca una pérdida de la función respiratoria o del acceso a la vía aérea, y reconocerá esas situaciones.
Objetivo 4.	El equipo se preparará eficazmente para el caso de que se produzca una pérdida considerable de sangre, y reconocerá esas situaciones.
Objetivo 5.	El equipo evitará provocar reacciones alérgicas o reacciones adversas a fármacos que se sabe que suponen un riesgo importante para el paciente.
Objetivo 6.	El equipo utilizará sistemáticamente métodos reconocidos para minimizar el riesgo de infección de la herida quirúrgica.
Objetivo 7.	El equipo evitará dejar accidentalmente gases o instrumentos en la herida quirúrgica.
Objetivo 8.	El equipo guardará e identificará con precisión todas las muestras quirúrgicas.
Objetivo 9.	El equipo se comunicará eficazmente e intercambiará información sobre el paciente fundamental para que la operación se desarrolle de forma segura.
Objetivo 10.	Los hospitales y sistemas de salud públicos establecerán una vigilancia sistemática de la capacidad, el volumen y los resultados quirúrgicos.

**Figura 8:** Lista de verificación de seguridad de la cirugía propuesta por la OMS en el marco del 2° Reto Mundial “La Cirugía Segura Salva Vidas” OMS. 2009.

Antes de la Inducción de la anestesia	Antes de la Incisión cutánea	Antes de que el paciente salga del quirófano
(Con el enfermero y el anestesista, como mínimo)	(Con el enfermero, el anestesista y el cirujano)	(Con el enfermero, el anestesista y el cirujano)
<p><b>¿Ha confirmado el paciente su identidad, el sitio quirúrgico, el procedimiento y su consentimiento?</b></p> <input type="checkbox"/> Sí	<p><input type="checkbox"/> Confirmar que todos los miembros del equipo se hayan presentado por su nombre y función</p> <p><input type="checkbox"/> Confirmar la identidad del paciente, el sitio quirúrgico y el procedimiento</p> <p><b>¿Se ha administrado profilaxis antibiótica en los últimos 60 minutos?</b></p> <input type="checkbox"/> Sí <input type="checkbox"/> No procede	<p><b>El enfermero confirma verbalmente:</b></p> <input type="checkbox"/> El nombre del procedimiento <input type="checkbox"/> El recuento de instrumentos, gases y agujas <input type="checkbox"/> El etiquetado de las muestras (lectura de la etiqueta en voz alta, incluido el nombre del paciente) <input type="checkbox"/> Si hay problemas que resolver relacionados con el instrumental y los equipos
<p><b>¿Se ha marcado el sitio quirúrgico?</b></p> <input type="checkbox"/> Sí <input type="checkbox"/> No procede	<p><b>Previsión de eventos críticos</b></p> <p>Cirujano:</p> <input type="checkbox"/> ¿Cuáles serán los pasos críticos o no sistematizados? <input type="checkbox"/> ¿Cuánto durará la operación? <input type="checkbox"/> ¿Cuál es la pérdida de sangre prevista?	<p><b>Cirujano, anestesista y enfermero:</b></p> <input type="checkbox"/> ¿Cuáles son los aspectos críticos de la recuperación y el tratamiento del paciente?
<p><b>¿Se ha completado la comprobación de los aparatos de anestesia y la medicación anestésica?</b></p> <input type="checkbox"/> Sí	<p>Anestesista:</p> <input type="checkbox"/> ¿Presenta el paciente algún problema específico?	<p><b>Equipo de enfermería:</b></p> <input type="checkbox"/> ¿Se ha confirmado la esterilidad (con resultados de los indicadores)? <input type="checkbox"/> ¿Hay dudas o problemas relacionados con el instrumental y los equipos?
<p><b>¿Se ha colocado el pulsioxímetro al paciente y funciona?</b></p> <input type="checkbox"/> Sí	<p><b>¿Pueden visualizarse las imágenes diagnósticas esenciales?</b></p> <input type="checkbox"/> Sí <input type="checkbox"/> No procede	
<p><b>¿Tiene el paciente...</b></p> <p>... Alergias conocidas?</p> <input type="checkbox"/> No <input type="checkbox"/> Sí		
<p>... Via aérea difícil / riesgo de aspiración?</p> <input type="checkbox"/> No <input type="checkbox"/> Sí, y hay materiales y equipos / ayuda disponible		
<p>... Riesgo de hemorragia &gt; 500 ml (7 ml/kg en niños)?</p> <input type="checkbox"/> No <input type="checkbox"/> Sí, y se ha previsto la disponibilidad de líquidos y dos vías IV o centrales		

La presente lista no pretende ser exhaustiva. Se recomienda completarla o modificarla para adaptarla a la práctica local.

Revisado 1 / 2009

© OMS, 2009

La valoración de la implementación del uso de la lista propuesta por la OMS, con algunas modificaciones según el caso, ha sido reportada como inicialmente favorable en relación a la disminución de las complicaciones operatorias y de la tasa de mortalidad asociada a la cirugía, aunque no hay resultados definitivos aún. Como se puede observar, la lista propone verificaciones en distintos momentos del acto quirúrgico (antes de la inducción anestésica, antes de la incisión cutánea, y antes de que el paciente salga del quirófano). No se trata de una verificación puntual de la técnica operatoria, sea global o específica de un procedimiento, si no que pretende asegurar una serie de medidas preventivas básicas para evitar la aparición de complicaciones o eventualmente la muerte (70, 71, 72).

Las escalas de **evaluación global** son instrumentos que no son específicos para un procedimiento, esto significa que son aplicables a cualquier tipo de acto quirúrgico, tanto en seres humanos como en modelos animales e inanimados. También pueden utilizarse para evaluación de aspectos no estrictamente técnicos, como por ejemplo la seguridad del paciente en el quirófano. Diversos estudios han demostrado que tienen buena validez y confiabilidad.

A continuación se presenta una escala denominada “Índice Global de Evaluación del Desempeño Técnico” o **GRITS** (Tabla 10) (46) (del inglés, Global Rating Index for Technical Skills), que a su vez está basada en otras escalas de evaluación global previamente validadas en otros ámbitos como la “Evaluación Objetiva Estructurada del Desempeño Técnico” u **OSATS** (Tabla 5) (13, 29, 37, 45, 48, 49,) (del inglés, Objective Structured Assessment of Technical Skills) y la “Evaluación Global de la Destreza Quirúrgica Laparoscópica” o **GOALS** (Tabla 11) (73) (del inglés, Global Operative Assessment of Laparoscopic Skills).

La escala **GRITS** ha demostrado tener una buena validez de constructo y confiabilidad, cuando fue utilizada en la sala de operaciones aplicada a diversos procedimientos quirúrgicos. Nótese la diferencia con nuestro trabajo en los siguientes aspectos: la escala es de aplicación universal y no específica de procedimiento, la evaluación se realiza en el mismo quirófano mediante observación directa, y no a través de un registro filmográfico en condiciones de anonimato, los procedimientos evaluados son diversos. La observación directa en quirófano podría ser interpretada como un factor de simplificación de la tarea, pero hace perder la ventaja metodológica del cegamiento de la muestra, y pone en juego nuevamente la influencia de la intersubjetividad en la toma de la medida calificadora. Ya ha sido demostrado que la observación directa es un método poco confiable para la evaluación (63). Por otro lado uno de los aspectos incluidos en la escala hace referencia a las habilidades de comunicación, que no debe ser considerada como una habilidad técnica propiamente dicha, y también incluye dos puntos que solo son aplicables en el caso de cirugía laparoscópica, como son la percepción profunda, y la destreza bi-manual.

**Tabla 10: GRITS (Global Rating Index for Technical Skills) Índice Global de Evaluación del Desempeño Técnico. Nótese que el ítem 7 evalúa “habilidades de comunicación” que no es una habilidad técnica, y los ítems 8 “Percepción profunda” y 9 “Destrezas bi-manuales” son aplicables en procedimientos laparoscópicos exclusivamente.**

1. Respeto por los tejidos				
1	2	3	4	5
Aplica frecuentemente una fuerza innecesaria sobre los tejidos o causa daño por el uso inapropiado del instrumental		Tiene manejo cuidadoso de los tejidos pero ocasionalmente causa daño inadvertido	Tiene manejo adecuado y consistente de los tejidos con un mínimo daño de los mismos	
2. Tiempo y movimiento				
1	2	3	4	5
Muchos movimientos inadecuados		Hay una relación eficiente tiempo/movimiento pero hace movimientos innecesarios	Hay clara economía de movimientos Eficiencia máxima	
3. Manejo y conocimiento del instrumental				
1	2	3	4	5
Movimientos dudosos y torpes Uso inapropiado		Uso competente del instrumental con torpezas ocasionales	Movimientos fluidos con el instrumental. No hay movimientos torpes	
4. Flujo de la cirugía				
1	2	3	4	5
Detenimiento frecuente, impresiona con inseguridades del próximo paso		Hay algo de planificación hacia delante, progresión razonable de la cirugía	Curso claramente planeado, flujo sin esfuerzo	
5. Conocimiento del procedimiento específico				
1	2	3	4	5
Conocimiento deficiente. Requiere instrucciones específicas en la mayoría de los pasos		Conocimiento de los pasos importantes del procedimiento	Demuestra familiaridad con todos los pasos del procedimiento	
6. Utilización de los asistentes (si aplicable)				
1	2	3	4	5
Ubica inadecuadamente a los asistentes o no los utiliza		Uso apropiado de los asistentes durante la mayoría del tiempo	Uso estratégico de los ayudantes para la máxima ventaja todo el procedimiento	
7. Habilidades de comunicación				
1	2	3	4	5
Problemas frecuente para el trabajo en equipo o falla en la comunicación		Comunicación apropiada con el equipo durante la mayoría del procedimiento	Coordina el equipo quirúrgico de una manera superadora	
8. Percepción profunda (solo para cirugía laparoscópica)				
1	2	3	4	5
Excesivos movimientos laterales, constantes movimientos en eje de profundidad, corrección lenta		Algunos movimientos en eje de profundidad pero con rápida corrección	Dirige los instrumentos con precisión en el eje correcto	
9. Destrezas bi-manuales (solo para cirugía laparoscópica)				
1	2	3	4	5
Utilización exclusiva de una mano, coordinación entre manos pobre		Usa ambas manos pero no optimiza su interacción	Esa ambas manos adecuadamente para ofrecer una correcta exposición	

**Tabla 11:** GOALS (*Global Operative Assessment of Laparoscopic Skills*). Evaluación Global de la Destreza Quirúrgica Laparoscópica. Nótese que los ítems a evaluar son “no-específicos” de un procedimiento, y que se contemplan aspectos elementales de la cirugía laparoscópica exclusivamente.

1. Percepción profunda				
1	2	3	4	5
Excesivos movimientos laterales, constantes movimientos en eje de profundidad, corrección lenta		Algunos movimientos en eje de profundidad pero con rápida corrección		Dirige los instrumentos con precisión en el eje correcto
2. Destrezas bi-manuales				
1	2	3	4	5
Utilización exclusiva de una mano, coordinación entre manos pobre		Usa ambas manos pero no optimiza su interacción		Usa ambas manos adecuadamente para ofrecer una correcta exposición
3. Eficiencia				
1	2	3	4	5
Incierta, mucho esfuerzo malgastado, muchos movimientos innecesarios, cambio constante del foco de la cirugía, o falta de progreso en la cirugía		Lento pero con un plan razonable y organizado		Confiado, eficiente y conducción segura de la cirugía, mantiene el foco en los pasos de la cirugía que los realiza adecuadamente
4. Manejo de los tejidos				
1	2	3	4	5
Rudeza, rompe los tejidos por excesiva fuerza, daño estructuras adyacentes, pobre control de la coagulación, los tejidos no suelen ser firmemente tomados por los “graspers”		Manejo cuidadoso de los tejidos pero con daño menor de tejidos vecinos, (ej: coagula el hígado, induce sangrado hepático, los tejidos se resbalan del “grasper”)		Tiene manejo adecuado de los tejidos con una tracción correcta. Mínima injuria de tejidos vecinos. Óptima utilización de la energía y sin exceso
5. Autonomía				
1	2	3	4	5
No puede completar todo el procedimiento, aun en casos fáciles y con una guía externa adecuada		Puede completar el procedimiento pero con ayuda externa moderada		Realiza el procedimiento solo sin ayuda

La escala GOALS se crea en el año 2003 en la Universidad de McGill (Montreal, Canada), como respuesta a la falta de una herramienta válida, confiable, objetiva y práctica para la evaluación de las habilidades quirúrgicas en cirugía laparoscópica. Si bien ya existían herramientas de evaluación validadas para la cirugía abierta (OSATS), no se contaba con una escala equivalente para la cirugía laparoscópica. Frente a esta carencia no había un método estandarizado para evaluar la competencia ni devolver al residente en formación información objetiva y específica en relación a su desempeño en cirugía laparoscópica.

La escala GOALS es una escala global que contempla la evaluación de los aspectos básicos de la cirugía laparoscópica, es decir que no es específica de un procedimiento en particular. Los aspectos evaluados son: percepción de profundidad, habilidad bimanual, eficiencia, manejo de tejidos, y autonomía. Estos ítems surgen del análisis y la deconstrucción al nivel más elemental de la motricidad que se requiere durante el uso de esta tecnología. Esta escala ha demostrado ser válida y confiable, y permite proveer al cirujano en etapa de entrenamiento con información específica en relación a su desempeño, para orientarlo en su formación (73).

Se debe aclarar que la evaluación mediante el uso de esta escala fue realizada mediante la observación directa y en tiempo real en el quirófano, y esto marca una clara diferencia con nuestro trabajo de investigación. Nuevamente cuestionamos este punto. La evaluación fue realizada por varios observadores en simultáneo, y existen enunciados explícitos sobre qué evaluar y cómo calificar, como resguardo de la objetividad. Sin embargo la situación presencial (coincidencia física evaluador-evaluado en quirófano) implica la pérdida de la ventaja metodológica del cegamiento de la muestra, y la influencia de la intersubjetividad en la toma de la medida calificadora.

La amplia aceptación del Examen Clínico Objetivo Estructurado (**ECO**E) (74, 75, 76, 77, 78, 79, 80, 81, 82, 83) llevó a un grupo de médicos de la Universidad de Toronto liderado por Richard Reznick (84) a desarrollar un concepto similar aplicado a la evaluación de habilidades técnicas, que fue denominado “Evaluación objetiva estructurada del desempeño técnico” u **OSATS** (del inglés Objective Structured Assessment of Technical Skills) (Tabla 5), que consiste en seis estaciones donde los candidatos desarrollan procedimientos quirúrgicos estandarizados en modelos animales o inanimados, durante un período limitado y previamente definido de tiempo, en un laboratorio de entrenamiento y simulación quirúrgica. El desempeño es evaluado con una lista de cotejo específica de la consigna a desarrollar, y una evaluación del desempeño global que consiste en siete componentes genéricos de destreza operatoria que son cuantificados con una escala de cinco puntos de Likert, con descriptores explícitos para los puntos extremos y el medio (13, 49). Por su parte, Regehr (85) demostró que las listas de cotejo no le agregan ningún valor

adicional a la evaluación y que la confiabilidad es menor que la de las escalas de evaluación global, haciendo prescindible su utilización. El inconveniente de este tipo de evaluaciones es la inversión en recursos humanos y el tiempo que debe dedicarse en el análisis de las grabaciones u observación directa.

Vale la pena destacar que el **OSATS** es una escala no específica de procedimiento, ampliamente validada, que puede ser aplicada tanto en modelos inanimados como en animales de laboratorio y en seres humanos, sea mediante la observación presencial o bien por el registro de imágenes. Nuestro trabajo de investigación tomó este método de evaluación como disparador pero propuso trascender los aspectos generales o globales de la práctica quirúrgica, y plantear un instrumento específico para cada procedimiento seleccionado. OSATS ha servido como referencia y fuente de inspiración para un importante número de publicaciones relacionadas con el área de la evaluación objetiva de los aspectos procedimentales en diversas áreas de la práctica profesional, tanto en cirugía como en otras especialidades médicas, sea en la carrera de medicina como en otras relacionadas con el cuidado de la salud. (86, 87, 88, 89, 90, 91, 92, 93, 94)

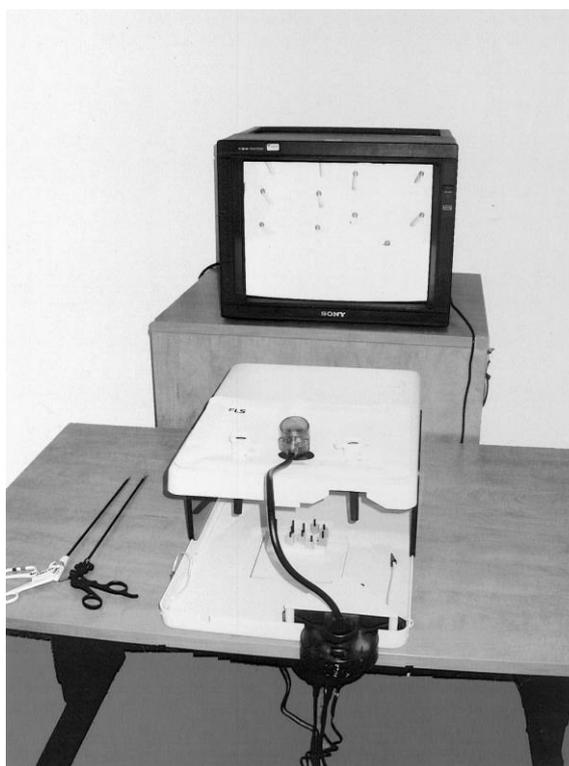
En las listas de cotejo y en las escalas globales se explicitan los aspectos parciales a verificar en la ejecución de la secuencia de pasos que constituyen un procedimiento. Definir con claridad y enunciar con palabras (explicitar) qué es lo que se deberá observar hace posicionar la mirada calificadora del evaluador sobre la ejecución del acto y la corre de la persona que lo está realizando. Las escalas globales cuentan con otro aspecto destacable: la explicitación de la medida calificadoria en función de la realidad observada, confiriéndole al proceso de evaluación mayor objetividad, validez y confiabilidad. Creemos que esta característica constituye una de las mayores fortalezas de la herramienta. Transforma a los evaluadores en observadores más que en interpretadores de la conducta, tendiendo a eliminar de esta forma la influencia de la intersubjetividad en el proceso de la calificación (85). Por otro lado, a diferencia de la clásica evaluación por observación directa, brindan información relacionada con el desempeño real en la ejecución que sirve de retroalimentación al médico en formación para la identificar fortalezas, detectar errores y redireccionar el entrenamiento quirúrgico fuera del quirófano.

Y finalmente mencionamos un sistema desarrollado en la Universidad de McGill (Montreal, Québec, Canadá) a partir de 1996, a cargo de un equipo liderado por Gerald M. Fried, denominado **MISTELS** (en inglés, McGill Inanimate System for Training and Evaluation of Laparoscopic Skills) que propone cinco ejercicios fundamentales que no son específicos para un procedimiento determinado, a desarrollar en un dispositivo de entrenamiento que es un **simulador** de cirugía laparoscópica (en inglés, “endotrainer”)

Un grupo de cinco expertos revisó una serie de grabaciones de diversos procedimientos realizados por vía laparoscópica: colecistectomía, hernioplastia inguinal, apendicectomía, y funduplicatura de Nissen. Del resultado de este análisis surge la selección de un conjunto de habilidades psicomotoras que son vinculadas de manera específica con la tecnología laparoscópica, es decir le son propias, y que constituyen el objeto de evaluación: 1) percepción de la profundidad con un sistema de visión monocular 2) percepción visual y espacial 3) el uso de ambas manos de manera complementaria 4) ligadura utilizando un sutura pre-armada (endoloop) 5) corte preciso con la mano dominante y contracción de la mano no dominante 6) la sutura mediante nudos intracorpóreos y 7) sutura con nudos extracorpóreos. Estas destrezas específicas fueron finalmente plasmadas en cinco ejercicios: 1) transferencia 2) corte 3) ligadura con nudo preformado 4) nudos intracorpóreos y 5) nudos extracorpóreos. El cirujano trabaja con dos trócares ubicados en posiciones fijas, que atraviesan una cubierta opaca que tapa la zona de trabajo e impide la visualización directa de los objetos destinados al desarrollo del ejercicio. El sistema óptico era inicialmente un laparoscopio de 10 mm y 0 grados ubicado en una posición fija, que luego fue reemplazado por una cámara que permitía la conexión con cualquier monitor de computadora o televisor con la entrada de video adecuada (figura 9). Cada ejercicio evalúa dos parámetros: la **eficiencia**, concepto relacionado con el tiempo de ejecución, y la **precisión**, entendido como la correcta ejecución de la consigna. Se establece un tiempo máximo para cada ejercicio y se aplica una penalidad por la falta de exactitud y por cometer errores.

Las ventajas de este sistema basado en un simulador inanimado son varias: es poco costoso y portátil, es reproducible, y tiene flexibilidad. Una ventaja sobre los sistemas de realidad virtual es que, a diferencia de estos,

permite el empleo de cualquier instrumental laparoscópico de uso habitual en el quirófano. Además, provee información de retroalimentación específica del desempeño de la ejecución, que está ausente en la mayoría de los sistemas de realidad virtual. Finalmente un aspecto fundamental: es mucho menos costoso. Es flexible en el sentido que permite la práctica de otras maniobras adicionales, por ejemplo suturas continuas, anastomosis, etc. Hacer esto en un sistema de realidad virtual implicaría agregar un desarrollo muy complejo, y los modelos hasta la actualidad no están apropiadamente desarrollados



**Figura 9:** Dispositivo de entrenamiento MISTELS. Obsérvese la cámara adosada que proyecta la imagen al monitor. La altura de la mesa puede modificarse para ajustarse a una altura de trabajo confortable.

El sistema MISTELS ha pasado por un riguroso proceso de validación y de determinación de su confiabilidad, por lo tanto provee datos útiles para la evaluación e información significativa para el estudiante, constituyendo de este modo una herramienta educacional efectiva (95, 96). En la actualidad se encuentra disponible comercialmente, y ha sido incorporado como una parte del programa Fundamentals of Laparoscopic Surgery (FLS™) ([www.flsprogram.org](http://www.flsprogram.org)) (97), desarrollado por la Sociedad Americana de Cirujanos Gastroenterológicos y Endoscopistas, **SAGES** (en inglés Society of

American Gastrointestinal and Endoscopic Surgeons) El FLS incluye la evaluación de conocimientos mediante un examen de opciones múltiples en el que se hace hincapié en el razonamiento clínico y en la toma de decisiones. Y un examen de destreza manual: el MISTELS. El FLS es en la actualidad un requisito para acceder a ser miembro de la Asociación Americana de Cirugía.

Desde nuestro punto de vista el MISTELS ha demostrado ser una herramienta sólida y efectiva. Su mayor fortaleza radica en su validez y confiabilidad, que son atributos internos del propio sistema. Debería considerarse como el paradigma de los sistemas “físicos o reales” de entrenamiento y evaluación. Hay muchos aspectos teóricos y prácticos que le confieren claras ventajas con relación a las propuestas de entrenamiento y evaluación basados en la realidad virtual, expuestos previamente y a los cuales adherimos. También presenta alguna ventaja en comparación con análisis de la destreza manual, relacionada con el tipo de información que se obtiene mediante el uso de estos sistemas, la posibilidad de interpretación de las medidas calificadoras y utilidad de las mismas para el médico en entrenamiento. Conceptos como los de desviación angular devuelven una medida válida de destreza manual, pero deja dudas si es útil como herramientas educacional, ya que la interpretación de la misma es compleja y carece de una correlación directa clara con la ejecución de un ejercicio.

Sin embargo hay algunas cuestiones para tener en cuenta. Este sistema puede ser utilizado en forma autónoma, tanto en el laboratorio de cirugía de cualquier institución médica como en el domicilio particular de cualquier profesional interesado en el entrenamiento, pero requiere de la presencia física de un supervisor para la evaluación formal de los ejercicios. Por otro lado plantea la calificación en función del tiempo, que desde nuestro punto de vista y en concordancia con la opinión de R. Reznick es un dato que no se relaciona de manera tan directa con la calidad de la ejecución. Si bien se trata de ejercicios en modelo de simulación, el cuestionamiento tiene valor.

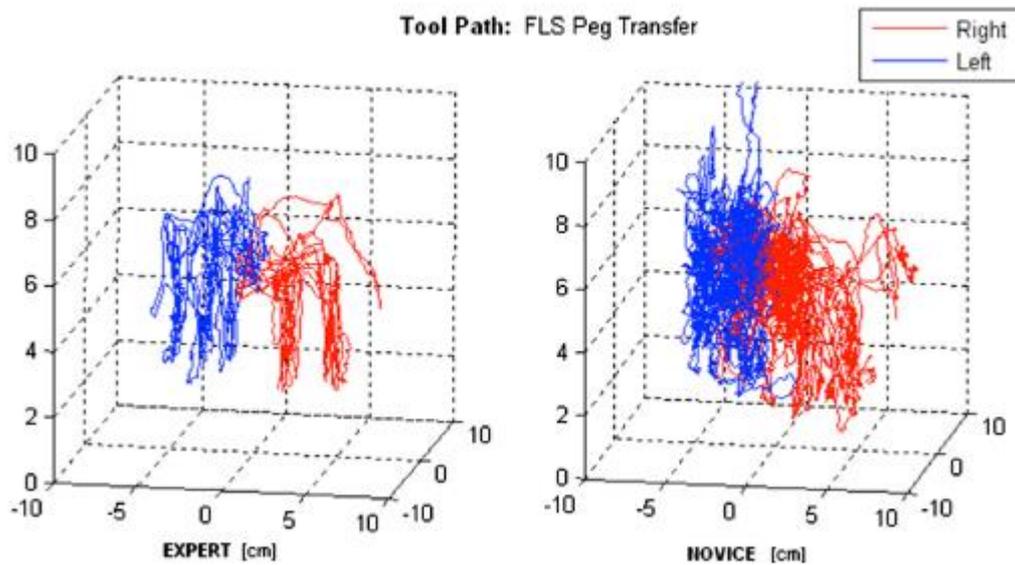
Nuestro trabajo presenta diferencias marcadas con el MISTELS. Por comenzar nosotros propusimos la aplicación de un instrumento de evaluación en el ámbito de un quirófano con un paciente real que está siendo sometido a un procedimiento quirúrgico específico. No trabajamos sobre ejercicios de habilidades básicas exclusivamente laparoscópicas en el marco de un

programa de entrenamiento con un simulador. Dejamos enunciado con claridad qué evaluar, cómo hacerlo, y qué calificación numérica corresponde en una escala de 1 a 5 en función de cómo se ejecutan los pasos. La forma de la calificación del parámetro “precisión” no ha sido claramente publicada. Se menciona un descuento de puntaje por “falta de precisión o errores” pero no existe explicitación comunicada.

Tal como afirma Fried GM, creador y promotor del MISTELS, este sistema no está diseñado para medir competencias en el quirófano. Él sostiene que se han identificado un conjunto de habilidades básicas que son propias de la cirugía laparoscópica y diferentes de la cirugía abierta, y que estas destrezas básicas deben ser aprendidas fuera del ámbito del quirófano, y en una instancia previa a actuar sobre un paciente real. Es justamente en esta instancia previa donde el uso del MISTELS tiene mejor lugar. (95, 98)

**Sistemas de análisis de la destreza manual o rastreo mecánico del movimiento.** Se puede mencionar el **ICSAD** (Imperial Collage Surgical Assessment Device) que es un sistema de rastreo, disponible comercialmente bajo el nombre Isotrak II® de Polhemus, que consiste en un generador de campo electromagnético y dos sensores que son colocados en las manos del cirujano en una posición estandarizada. Hay un procesador de datos que utiliza un software que convierte la información proveniente de los sensores en medidas de destreza manual, como el número y velocidad de movimientos, la distancia recorrida por las manos y el tiempo necesario para ejecutar la tarea (Figura 10).

Diversos estudios han demostrado que este sistema presenta validez de constructo cuando es aplicado a una variedad de procedimientos, tanto laparoscópicos como abiertos (99) Existe una fuerte correlación entre los resultados del análisis de la destreza manual y la experiencia previa en cirugía laparoscópica, provenientes tanto de la ejecución de una tarea simple en un simulador para entrenamiento, como en la ejecución de una consigna más compleja como una colecistectomía laparoscópica en un modelo animal porcino (100). Un cirujano entrenado y habilidoso es más económico con referencia al número de movimientos que necesita para completar una tarea, y más preciso



**Fig 10:** Resultado del análisis de destreza manual para el parámetro “trayectoria del instrumental” durante la ejecución del ejercicio de transferencia de objetos que forma parte del programa *Fundamentals in Laparoscopic Surgery*. En rojo mano derecha, en azul la mano izquierda. El gráfico de la izquierda representa el resultado de un experto, y el de la derecha el de un novato.

a la hora de localizar y alcanzar sus objetivos. Lo que significa que usa caminos más cortos para completar y conseguir su meta quirúrgica (101).

Otro sistema que corresponde mencionar es el **ADEPT** (Advanced Dundee endoscopic psychomotor trainer) que surge originalmente diseñado como una herramienta para la selección de médicos en formación en el área de la cirugía endoscópica, y está basado en la capacidad de algunos tests psicomotores de predecir la habilidad innata para el desarrollo de tareas relevantes. Hay estudios que han demostrado la validez y confiabilidad de este sistema de entrenamiento (102). Las pinzas están equipadas con sensores que miden las desviaciones angulares y pueden medir errores como el contacto excesivo (entre ellas y con los tejidos) Las mediciones son hechas con un software específico y se usa con una caja de entrenamiento que tiene un laparoscopio. (103)

Los **sistemas de realidad virtual** pueden definirse como un conjunto de tecnologías que permiten a un individuo interactuar eficientemente con bases de datos computarizadas, en tres dimensiones y en tiempo real, usando sus órganos de los sentidos naturales y habilidades (104). Se trata de equipos de alta complejidad tecnológica que cuentan con un software que emula con gran fidelidad y alta resolución un entorno virtual en el cual es posible la ejecución de un amplio rango de actividades que van desde ejercicios básicos hasta procedimientos complejos (Figuras 11 y 12). Entre los ejercicios básicos disponibles se pueden mencionar aquellos originalmente diseñados para las cajas de entrenamiento con sistema óptico (endotrainer) que consisten en el movimiento de instrumentos, la transferencia de objetos, y el corte de estructuras (Figura 13). También es posible la navegación de un endoscopio y la manipulación de órganos intra-abdominales, sin transmisión táctil. Entre las acciones más complejas se incluyen la simulación de intervenciones quirúrgicas con aplicación de clips, corte de estructuras, uso de electrocoagulación y aplicación de nudos y suturas (Figura 14). Algunos equipos ya cuentan con un sistema que integra la transmisión de la sensación táctil, llamada transmisión “háptica”.

Los simuladores de realidad virtual son sistemas informáticos que generan datos de “salida”. En un encuentro internacional de expertos donde se revisaron todos los métodos de evaluación disponibles se sugirieron qué parámetros deberían formar parte de las medidas de calificación de las habilidades técnicas. Se propusieron: economía de movimientos, largo de los desplazamientos, y uso inadecuado de los instrumentos entre otros (105).

Los datos son almacenados en la memoria del equipo para ser analizados y cada usuario cuenta con su propio registro. Se elaboran gráficos del desempeño conforme se avanza a los ejercicios más complejos y se lleva una estadística del tipo bitácora de la agenda de prácticas.

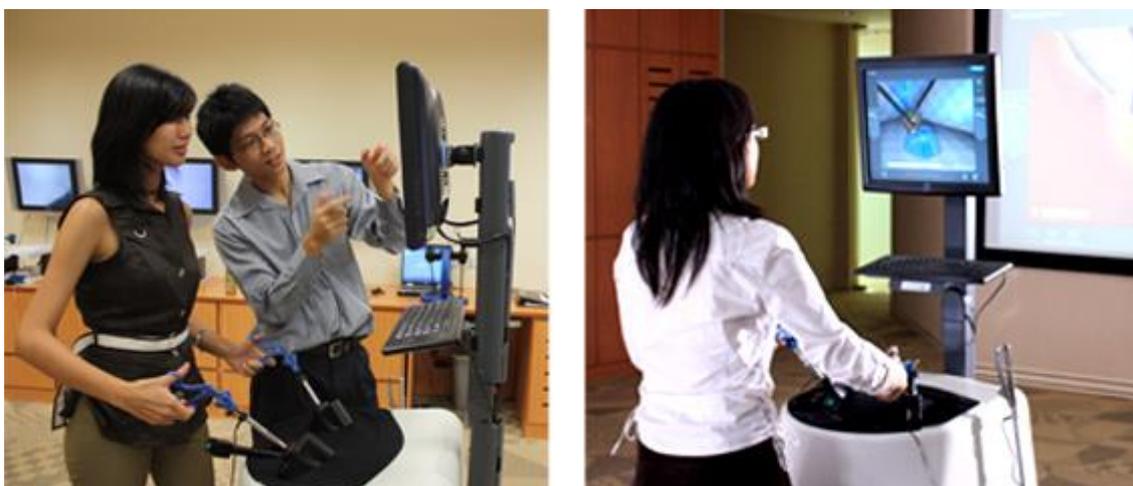
Existen por lo menos 5 equipos comerciales, todos ellos similares en su sistema de evaluación:

- 1) Lapsim® (Surgical Science, Gothemberg, Suecia)
- 2) MIST-VR® (Mentice, San Diego, California, USA)
- 3) Xitact LS500® (Xitact, Morges, Suiza)

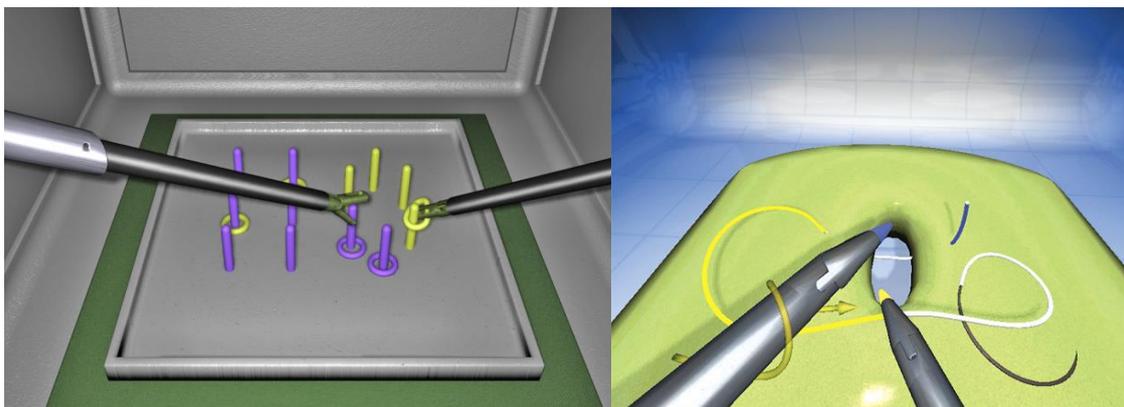
- 4) Xitact LS500® (Xitact, Morges, Suiza)
- 5) LapMentor® (Symbionix, Cleveland, Ohio, USA)



**Figura 11:** Simulador laparoscópico LapSim® de SurgicalScience. A la izquierda el modelo con transmisión háptica. A la derecha la versión sin transmisión háptica.



**Figura 12:** Simulador laparoscópico LapMentor® de Symbionix. El sistema ofrece 13 módulos de entrenamiento, más de 65 tareas y casos, incluyendo cirugía general, ginecología, urología, cirugía bariátrica y cirugía colorrectal.

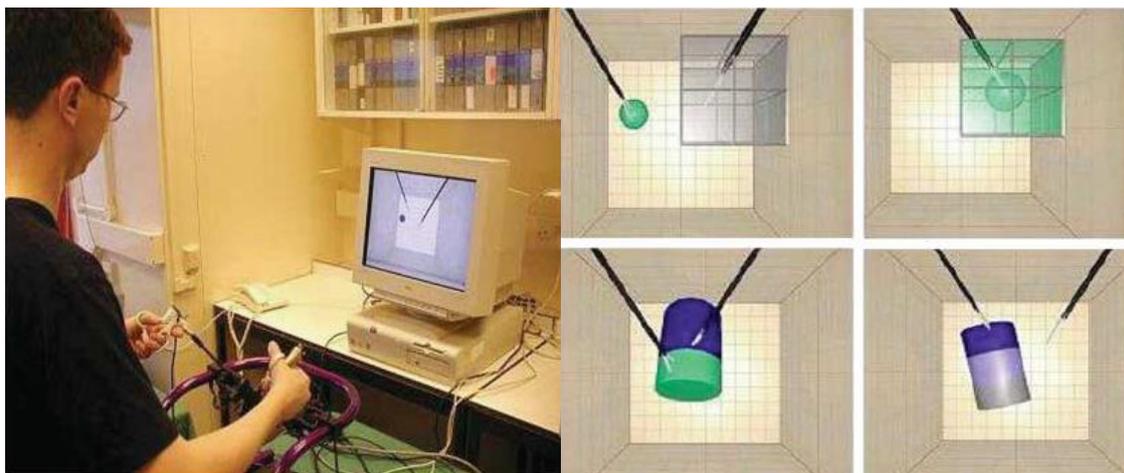


**Figura 13:** Sistema de Realidad Virtual aplicada a la ejecución de ejercicios básicos laparoscópicos. A la izquierda LapSim® y a la derecha LapMentor®.



**Figura 14:** Sistema de Realidad Virtual LapMentor® aplicada a dos procedimientos quirúrgicos laparoscópicos: colecistectomía (izquierda) y apendicectomía (derecha).

Se puede mencionar el **MIST-VR** (Minimally Invasive Surgical Trainer – Virtual Reality) (figura 15) que reproduce una serie de imágenes tridimensionales que pueden ser manipuladas. Surge de la colaboración entre cirujanos y psicólogos quienes trabajando en conjunto hicieron un análisis de la colecistectomía laparoscópica y determinaron un grupo de habilidades motoras básicas necesarias para ejecutar este procedimiento con éxito. Los datos que surgen del uso de este sistema se pueden medir y evaluar, y MIST-VR ha sido validado ampliamente para la evaluación de habilidades básicas en cirugía laparoscópica.



**Figura 15:** *MIST-VR propone la simulación de habilidades laparoscópicas básicas en un entorno virtual. Tomado de BMJ. Vol. 327. p. 1035. Noviembre del 2003.*

Desde nuestra perspectiva los sistemas de realidad virtual constituyen un avance tecnológico sorprendente que han revolucionado el campo del entrenamiento mediante simulación. A medida que surgen nuevas líneas de desarrollo en el área aparecen nuevas aplicaciones posibles que deben ser puestas a prueba. En la actualidad se ofrecen procedimientos quirúrgicos complejos, como el by-pass gástrico en Y de Roux, o una colectomía, en un entorno de realidad virtual. Sin embargo no existe clara evidencia de ser superiores a otros sistemas de simulación y entrenamiento menos complejos y costosos, y más accesibles. Por lo tanto, es en el área de las habilidades laparoscópicas básicas donde la realidad virtual ha demostrado hasta el presente su mayor fortaleza, compitiendo con el sistema MISTELS del programa Fundamentals of Laparoscopic Surgery del American Board of Surgeons. Los parámetros de evaluación son los mismos que los sistemas de rastreo de destreza manual, y hacen claramente referencia a un tipo particular de habilidad motora característica de los procedimientos laparoscópicos. No está aún validado para procedimientos específicos y complejos. Resulta particularmente interesante el entorno virtual presentado por su semejanza con la realidad anatómica, pero solo recientemente se ha incorporado la tecnología “háptica” que no es otra cosa que la simulación virtual del sentido del tacto en el instrumental de los simuladores (106, 107).

**El análisis de los productos finales:** los resultados finales de la cirugía y la evolución posterior del paciente no pueden ser atribuidos exclusivamente a cuestiones técnicas. De hecho las consecuencias de una mala técnica pueden ponerse en evidencia muchos años después. Szalay y col. (108) propusieron la evaluación del producto final de los modelos inanimados que constituyen las diferentes estaciones del OSATS. Sus estudios demostraron validez de constructo y una muy buena correlación con la evaluación del OSATS en vivo, superando de esta forma los inconvenientes derivados de las observaciones en vivo y las grabaciones. Datta et al. analizaron la tasa de filtrado y el área de sección transversal de modelos de anastomosis vasculares (109). Hanna et al estudiaron la calidad de los nudos hechos por vía laparoscópica usando un tensiómetro y desarrollaron una escala de calidad y confiabilidad (110)

La principal ventaja de este tipo de determinaciones es que resuelve alguna de las limitaciones concernientes a la disponibilidad de recurso humano que son propias de las evaluaciones basadas en observaciones en vivo y en grabaciones. Sin embargo se debe resaltar que el producto final es precisamente el resultado último de un proceso. Por lo tanto la evaluación se hará sobre la consecuencia última de una serie de pasos que dan origen a un producto. El resultado es el reflejo indirecto de un proceso que queda fuera del análisis, ya que se excluye la observación del accionar del cirujano. Es decir solo se podrá hacer una inferencia sobre el procedimiento.

En nuestra opinión la información proveniente del análisis de los productos finales, es sin dudas valiosa, pero debería ser complementada con datos relevantes del procedimiento mismo, sean de destreza manual para obtener valores de efectividad, o de grabaciones o de observaciones directas del procedimiento quirúrgico, para lograr que la evaluación sea más objetiva y válida.

En resumen, el conocimiento actual sobre el tema pone en evidencia falencias de peso en los métodos de evaluación de las habilidades técnicas vigentes en la actualidad. Los métodos tradicionales vigentes en la mayoría de los centros formadores de nuestro medio carecen de validez y confiabilidad en relación a la destreza quirúrgica.

La formación quirúrgica dentro del marco de un programa formal estructurado de residencias médicas requiere una retroalimentación objetiva

constante de la adquisición de habilidades técnicas para corrección y reorientación, y para ello precisa de instrumentos de evaluación objetivos. Todas las profesiones en general, relacionadas o no con la salud, implican la ejecución de procedimientos que requieren la adquisición de destrezas técnicas de variada complejidad. La formación de profesionales de ámbitos tan disímiles como son un piloto de avión y un cirujano, en apariencia sin ningún punto de contacto, tienen en común el hecho que ambas requieren el desarrollo de habilidades técnicas y que ambas deben enfrentar condiciones en las cuales, eventualmente, se pone en riesgo la vida. Muchos aspectos relacionados con el entrenamiento y la evaluación de los pilotos de avión podrían ser aplicables a la formación de los cirujanos (28)

## CONCLUSIONES

1. La implementación de instrumentos para la evaluación objetiva de competencias técnicas constituye una herramienta de enorme potencial para todos los actores en el complejo proceso de formación de un cirujano.
2. Existen instrumentos objetivos, válidos y confiables, que constituyen herramientas útiles para la evaluación del desempeño técnico, y que pueden ser utilizados con distintas finalidades: formativa, sumativa, y de certificación.
3. La objetividad debe entenderse como instrumental, es decir relacionada con el método de obtención de la calificación. El uso que se le dará a los datos (la medida) responde a definiciones institucionales, históricas, y a un modelo deseado y procurado de práctica profesional.
4. La captura de la imagen es un factor fundamental, ya que le confiere el carácter de permanente a un procedimiento efímero (la operación). Constituye un testimonio de lo actuado, permite comparar, detectar errores para corregirlos, y redireccionar la práctica.
5. La tarea de los evaluadores es de capital importancia. Dos aspectos deben tenerse en cuenta: la necesidad de compartir los mismos estándares de calidad, y el entrenamiento en evaluación.
6. La evaluación debe ser entendida como un proceso de construcción de conocimiento sobre el objeto de evaluación.

7. La evaluación implica un acto de interrogación sobre aquello que se evalúa, que se puede formular tanto sobre el objeto de evaluación como sobre las propias estrategias y prácticas del evaluador.
8. Evaluar siempre supone algún tipo de recorte de la realidad, tanto desde los referentes que se construyen, como desde los aspectos referidos que se seleccionan del objeto a evaluar y que se consideran representativos de éste.

## **Capítulo 5**

# **BIBLIOGRAFÍA**

1. Herskovits MJ. *El Hombre y sus Obras. La ciencia de la Antropología Cultural*. Primera reimpresión Argentina. Buenos Aires. Fondo de Cultura Económica. p. 341. 1981.
2. Carlino F, Augustowsky GP, Di Alessio MB, Singer L. “La genesis de las prácticas de evaluación: la prehistoria del campo” En: *La Evaluación Educacional. Historia, problemas y propuestas*. Primera edición. Buenos Aires. ISBN 950-701-573-6. Aique Grupo Editor S.A. p. 30-34. 1999.
3. Real Academia Española. *Diccionario de la Lengua Española*. [en línea] 22ª edición [consultado en marzo del 2014] Disponible en la URL [www.rae.es](http://www.rae.es)
4. Barlow M. *L'évaluation scolaire. Décoder son langage*. ISBN: 2-85008-164-7. Lyon. Chronique Social. p. 60-63. 1992.
5. Bertoni A, Poggi M, Teobaldo M. “Los significados de la evaluación educativa: alternativas teóricas” En: *Evaluación. Nuevos significados para una práctica compleja*. Primera edición. Buenos Aires. ISBN 950-13-6184-5. Editorial Kapelusz. p. 24-25. 1995.
6. Valles MS. *Técnicas cualitativas de Investigación Social. Reflexión metodológica y práctica profesional*. 3ª edición. Madrid. Editorial Síntesis. p. 23. 2003.
7. Díaz Barriga A. *Docente y Programa. Lo institucional y lo didáctico*. Primera edición. Buenos Aires. Aique Grupo Editor S.A. Cap 4. p.125. 1994
8. Stufflebeam D, Shinkfield A. *Evaluación Sistemática. Guía teórica y práctica*. 1ª Edición. Barcelona. Paidós. p.33. 1987.
9. Durkheim E, Ortega F, Delgado ML. *Historia de la educación y de las doctrinas pedagógicas: la evolución pedagógica en Francia*. 1ª Edición. Madrid. La Piqueta. p. 119. 1982.
10. Foucault M. *Vigilar y Castigar. Nacimiento de la Prisión*. 23ª edición en español. México. Editorial Siglo XXI. p. 151-154. 1995.
11. Díaz Barriga A. *Docente y Programa. Lo institucional y lo didáctico*. Primera edición. Buenos Aires. Aique Grupo Editor S.A. Cap 4. p.128-133. 1994
12. Carlino F, Augustowsky GP, Di Alessio MB, Singer L. *La Evaluación Educacional. Historia, problemas y propuestas*. Primera edición. ISBN

- 950-701-573-6. Buenos Aires. Aique Grupo Editor S.A. Cap 1 y 2. p. 34-47. 1999
13. Reznick R, Regehr G, MacRae H, Martin J, McCulloch W. Testing Technical Skill via an Innovative “Bench Station” Examination. *Am J Surg*. 173: 226–230. 1997.
  14. Kasparian AC, Chércoles RA. Nuevos instrumentos de evaluación objetiva de competencias técnicas en cirugía. *Revista Facultad de Ciencias Médicas*. 68(1): 20-24. 2011.
  15. Cosman P, Hemli JM, Ellis AM, Hugh TJ. Learning the Surgical Craft: A Review of Skills Training Options. *ANZ J. Surg*. 77: 838-845. 2007.
  16. Cuschieri A. 2001. Master surgeons and surgical proficiency. En *Metrics: Objective Assessment of Surgical Skills – Developing Quantitative Measurements through Surgical Simulation*. Scottsdale, Arizona: US Army Medical Research Command, Telemedicine and Advances Technology Research Center (TATRC)
  17. Cuschieri A, Francis N, Crosby J, Hanna GB. What do Master Surgeons Think of Surgical Competence and Revalidation. *Am J Surg* 182:110-116. 2001.
  18. Yule S, Flin R, Paterson-Brown S, Maran N. Non-technical Skills for surgeons in the operating room: A review of the literature. *Surgery*. 139: 140-149. 2006.
  19. Darzi A, Datta V, Mackay S. The Challenge of Objective Assessment of Surgical Skill. *Am J Surg* 181: 484-486. 2001
  20. Cannon W.B. Physiological regulation of normal states: Some tentative postulates concerning biological homeostatics. En A. Pettit (Ed.), *A Charles Richet: ses amis, ses collègues, ses élèves*, 22 Mai 1926, Les Editions Medicales, Paris. p. 91–93. 1926.
  21. Cannon W.B. Organization for physiological homeostasis. *Physiological Reviews*, 9: 399–431. 1929.
  22. Cannon W.B. The sympathetic division of the autonomic system in relation to homeostasis. *Archives of Neurology & Psychiatry*. 22: 282–294. 1929.
  23. Loyber I. *Introducción a la fisiología del sistema nervioso*. 2ª edición. Córdoba. El Galeno. p. 213. 2001.

24. Barrett KE, Barman SM, Boitano S, Brooks H. *Ganong. Fisiología Médica*. 23<sup>a</sup> edición en castellano. México. McGrawHill. p.241-243. 2010.
25. Fitts PM, Posner MI. *Human Performance*. 1<sup>a</sup> edición. Belmont, California. Brooks/Cole Pub. Co. 1967.
26. Kirk RM. Teaching de Craft of Operative Surgery. *ANN R Coll Surg Engl* 1996; 78 (Supl 1): 25-8.
27. Hamford JM, Hall JC. Acquiring Surgical Skills. *BJS* 87: 28 – 37. 2000.
28. Moen MD, Moen RO, Moen O.L. Teaching and Evaluating Technical Skills: Applying the Aviation Model to Gynecologic Surgery. *J Pelvic Med Surg*. 13: 107-112. 2007.
29. Reznick R, MacRae H. Teaching Surgical Skills – Changes in the Wind. *N Engl J Med* 355: 2664-2669. 2006.
30. Kane MT.. The assessment of clinical competence. *Eval Health Prof*. 15(2): 163-182. 1992
31. Beard JD. Assessment of Surgical Competence. *BJS* 94: 1315-1316. 2007.
32. Miller G. The assessment of clinical skills/competence/performance. *Academic Medicine*. 65: S63-S67. 1990.
33. Van Der Vleuten C. The assessment of professional competence: developments, research and practical implications. *Adv Health Sci Educ Theory Pract*. 1: 41-67. 1996
34. Greenberg JA, Irani JL, Greenberg CC, Blanco MA, Lipsitz S, Ashley SW, Breen EM, Hafler JP. The ACGME competencies in the operating room. *Surgery*. 142: 180-184. 2007
35. Scott DJ, Valentine RJ, Bergen PC, Rege RV, Laycock R, Tesfay ST. Evaluating surgical competency with the American Board of Surgery in-training examination, skill testing, and intraoperative assessment. *Surgery*. 128: 613-622. 2000.
36. Moorthy K, Munz Y, Sarker SK, Darzi A. Objective assessment of technical skills in surgery. *BMJ*. 327: 1032-1037. 2003.
37. Goff BA, Lentz GM, Lee D, Fenner D, Morris J, Mandel LS. Devolpement of a Bench Station Objective Structured Assessment of Technical Skills. *Obstet Gynecol*. 98: 412-416. 2001.

38. Smee S. ABC of learning and teaching in medicine. Skill based assessment. *BMJ*. 326: 703-706. 2003.
39. Van Der Vleuten C, Schuwirth L. Assessing professional competence: from methods to programmes. *Medical Education*. 39: 309-317. 2005.
40. Swing SR. The ACGME outcome project: retrospective and prospective. *Medical Teacher*. 29: 648-654. 2007.
41. The Royal College of Surgeons of England. Intercollegiate Surgical Curriculum Programme. *Intercollegiate Surgical Curriculum Overview*. [en línea] Octubre 2013. [Consultado 23 de Noviembre del 2014] Disponible en la URL [https://www.iscp.ac.uk/surgical/curr\\_intro.aspx](https://www.iscp.ac.uk/surgical/curr_intro.aspx)
42. Anderson CI, Jentz AB, Harkema JM, Rao Kareti L, Apelgren KN, Slomski CA. Assessing the competencies in general surgery residency training. *Am J Surg* 189: 288-292. 2005.
43. Bann S, Khan M, Datta V, Darzi A. Surgical skill is predicted by the ability to detect errors. *Am J Surg*. 189: 412-415. 2005
44. Boehler ML, Schwind CJ, Rogers DA, Ketchum J, O'Sullivan E, Mayforth R, Quin J, Wohltman C, Johnson C, Williams RG, Dunnington G. A Theory-Based Curriculum for Enhancing Surgical Skillfulness. *J Am Coll Surg*. 205: 492-497. 2007.
45. Dath D, Regehr G, Birch D, Schlachta C, Poulin E, Mamazza J, Reznick R, MacRae H.. Toward Reliable Operative Assessment. *Surg Endosc* 18: 1800-1804. 2004.
46. Doyle JD, Webber EM, Sidhu RS. A universal global rating scale for the evaluation of technical skills in the operating room. *Am J Surg*. 193: 551-555. 2007.
47. Larson JL, Williams RG, Ketchum J, Boehler M, Dunnington GL.. Feasibility, reliability and validity of an operative performance rating system for evaluating surgery residents. *Surgery*. 138: 640-649. 2005.
48. MacRae H, Regehr G, Leadbetter W, Reznick R.. A Comprehensive Examination for Senior Surgical Residents. *Am J Surg*. 179: 190-193. 2000.
49. Martin JA, Regehr G, Reznick R, MacRae H, Murnaghan J, Hutchinson G. Objective structured assessment of technical skill (OSATS) for surgical residents. *Br J Surg*. 84: 273-278. 1997.

50. Dubois F., Berthelots G., Levard H. Cholecystectomy par coelioscopie, *Presse Med.* 18: 980-982, 1989
51. Pellegrini C.A.: Cirugía Videoscópica. Relato Oficial del LXV Congreso Argentino de Cirugía. *Rev Arg Cirug.* N° Extraordinario, 1994.
52. Meinero M, Melotti G, Mouret Ph. *Cirugía Laparoscópica.* 1ª edición. Buenos Aires. Editorial Médica Panamericana S.A. p. 147. 1996.
53. Lichtenstein IL, Shulman AG, Amid PK, Monitor MM. The tension-free hernioplasty. *Am J Surg.* 57: 188-193. 1989.
54. Usher FC. Hernia repair with Marlex mesh. *Arch Surg.* 84: 325-328. 1962.
55. Shulman AG, Amid PK, Lichtenstein IL. The safety of mesh repair for primary inguinal hernias: results of 3019 from five diverse surgical sources. *Am Surg.* 58: 256-261. 1998.
56. Shulman AG, Amid PK, Lichtenstein IL. A survey of non-expert surgeons using the open tension-free mesh patch repair from primary inguinal hernias. *Int Surg.* 80: 35-36. 1995.
57. Barbier JM. *La evaluación en los procesos de formación.* 1ª edición. Barcelona. Ediciones Paidós. p. 63. 1993.
58. Poggi, M. Evaluación Educativa. Sobre Sentidos y Práctica. *Revista Iberoamericana de Evaluación Educativa,* 1 (1), p. 36-44. 2008.
59. Kasparian AC, Martínez AC, Jover Clos RJ, Chércoles RA. Evaluación Objetiva de Técnicas en Cirugía. *Revista Facultad de Ciencias Médicas.* 71 (3) 2014; in press.
60. Gutierrez C. Teoría de los métodos en las Ciencias Sociales. 1ª edición. Costa Rica. AULA-EDUCA. p. 209. 1971.
61. Cook TD, Reichardt ChS. Métodos cualitativos y cuantitativos en investigación evaluativa. 5ª edición. Madrid. Ediciones Morata SL. P. 44. 2005.
62. Bertoni A, Poggi M, Teobaldo M. *Evaluación. Nuevos significados para una práctica compleja.* Primera edición. Buenos Aires. Editorial Kapelusz. p. 38-40. 1995.
63. Reznick R. Teaching and testing surgical skills. *Am J surg.* 165: 358-36. 1993.

64. Weiser TG, et al. An estimation of the global volume of surgery. *Lancet*. 372:139-144. 2008.
65. Gawande AA, et al. The incidence and nature of surgical adverse events in Colorado and Utah in 1992. *Surgery*. 126:66–75. 1999.
66. Kable AK, Gibberd RW, Spigelman AD. Adverse events in surgical patients in Australia. *International Journal of Quality in Health Care*., 14:269–765. 2002.
67. Brennan TA, et al. Incidence of adverse events and negligence in hospitalized patients. Results of the Harvard Medical Practice Study I. *New England Journal of Medicine*. 324:370–611. 1991.
68. Bickler SW, Sanno-Duanda B. Epidemiology of paediatric surgical admissions to a government referral hospital in the Gambia. *Bulletin of the World Health Organization*. 78:1330–6. 2000.
69. McConkey SJ. Case series of acute abdominal surgery in rural Sierra Leone. *World Journal of Surgery*. 26:509–13. 2002.
70. Haynes A, Weiser T, Berry W, et al. A surgical safety checklist to reduce morbidity and mortality in a global population. *N Engl J Med*. 360: 491–9. 2009
71. Haugen AS, Søfteland E, Almeland SK, Sevdalis N, Vonen B, Eide GE, Nortvedt MW, Harthug S. Impact of the World Health Organization's Surgical Safety Checklist on safety culture in the operating theatre: a controlled intervention study. *British Journal of Anaesthesia* 110 (5): 807–15. 2013.
72. van Klei WA, Hoff RG, van Aarnhem EE, Simmermacher RK, Regli LP, Kappen TH, van Wolfswinkel L, Kalkman CJ, Buhre WF, Peelen LM. Effect of the World Health Organization Checklist on Patient Outcomes: A Stepped Wedge Cluster Randomized Controlled Trial. *Ann Surg*. 255:44–49. 2012.
73. Vassiliou MC, Feldman LS, Andrew CD, et al. A global assessment tool for evaluation of intraoperative laparoscopic skills. *Am J Surg*. 190: 107-113. 2005.
74. Harden RM, Stevenson M, Downie WW, Wilson GM. Assessment of clinical competence using objective structured examination. *Br Med J*. 1(5955):447-51. 1975.

75. Cuschieri A, Gleeson FA, Harden RM, Wood RA. A new approach to a final examination in surgery. Use of the objective structured clinical examination. *Ann R Coll Surg Engl.* Sep;61(5):400-5. 1979.
76. Harden RM. What is an OSCE? *Med Teach.* 10(1):19-22. 1988.
77. Harden RM. Twelve tips for organizing an Objective Structured Clinical Examination (OSCE). *Med Teach.* 12(3-4):259-64. 1990.
78. Yang YY, Lee FY, Hsu HC, Huang CC, Chen JW, Lee WS, et al. A core competence-based objective structured clinical examination (OSCE) in evaluation of clinical performance of postgraduate year-1 (PGY(1)) residents. *J Chin Med Assoc.* 74(5): 198-204. 2011
79. Wagner DP, Hoppe RB, Lee CP. The patient safety OSCE for PGY-1 residents: a centralized response to the challenge of culture change. *Teach Learn Med.* 21(1): 8-14. 2009.
80. Zyromski NJ, Staren ED, Merrick HW. Surgery residents' perception of the Objective Structured Clinical Examination (OSCE). *Curr Surg.* 60(5): 533-7. 2003.
81. Altshuler L, Kachur E. A culture OSCE: teaching residents to bridge different worlds. *Acad Med.* 76(5): 514. 2001.
82. Griesser MJ, Beran MC, Flanigan DC, Quackenbush M, Van Hoff C, Bishop JY. Implementation of an objective structured clinical exam (OSCE) into orthopedic surgery residency training. *J Surg Educ.* 69(2): 180-9. 2012.
83. Schoenmakers B, Wens J. The objective structured clinical examination revisited for postgraduate trainees in general practice. *Int J Med Educ.* 5:45-50. 2014.
84. Webster PCh. Richard Reznick: leading innovator of surgical education. *The Lancet.* 9785 (378): 21. 2011.
85. Regehr G, MacRae H, Reznick R, Szalay D. Comparing the psychometric properties of checklists and global rating scales for assessing performance on an OSCE format examination. *Acad Med.* 73: 993-997. 1998.
86. Zevin B, Bonrath EM, Aggarwal R, Dedy NJ, Ahmed N, Grantcharov TP, ATLAS group. Development, feasibility, validity, and reliability of a scale

- for objective assessment of operative performance in laparoscopic gastric bypass surgery. *J Am Coll Surg.* 216(5):955-965. 2013.
87. Alici F, Buerkle B, Tempfer CB. Objective Structured Assessment of Technical Skills (OSATS) evaluation of hysteroscopy training: a prospective study. *Eur J Obstet Gynecol Reprod Biol.* 178:1-5. 2014.
88. Niitsu H, Hirabayashi N, Yoshimitsu M, Mimura T, Taomoto J, Sugiyama Y, Murakami S, Saeki S, Mukaida H, Takiyama W. Using the Objective Structured Assessment of Technical Skills (OSATS) global rating scale to evaluate the skills of surgical trainees in the operating room. *Surg Today.* 43(3):271-5. 2013.
89. Jaffer U, Singh P, Pandey VA, Aslam M, Standfield NJ. Validation of a novel duplex ultrasound objective structured assessment of technical skills (DUOSATS) for arterial stenosis detection. *Heart Lung Vessel.* 6(2):92-104. 2014.
90. Marglani O, Alherabi A, Al-Andejani T, Javer A, Al-Zalabani A, Chalmers A. Development of a tool for Global Rating of Endoscopic Surgical Skills (GRESS) for assessment of otolaryngology residents. *B-ENT.* 8(3):191-5. 2012.
91. Khaliq T. Reliability of results produced through objectively structured assessment of technical skills (OSATS) for endotracheal intubation (ETI). *J Coll Physicians Surg Pak.* 23(1):51-5. 2013.
92. Egi H, Tokunaga M, Hattori M, Ohdan H. Evaluating the correlation between the HUESAD and OSATS scores: concurrent validity study. *Minim Invasive Ther Allied Technol.* 22(3):144-9. 2013.
93. Nimmons GL1, Chang KE, Funk GF, Shonka DC, Pagedar NA. Validation of a task-specific scoring system for a microvascular surgery simulation model. *Laryngoscope.* 122(10):2164-8. 2012.
94. Fransson BA, Ragle CA, Bryan ME. Effects of two training curricula on basic laparoscopic skills and surgical performance among veterinarians. *J Am Vet Med Assoc.* 15;241(4):451-60. 2012.
95. Fried GM, Feldman LS, Vassiliou MC, Fraser SA, Stanbridge D, Ghitulescu G, Andrew CG. Proving the value of simulation in laparoscopic surgery. *Ann Surg.* 240(3):518-25; discussion 525-8. 2004.

96. Vassiliou MC, Ghitulescu GA, Feldman LS, Stanbridge D, Leffondré K, Sigman HH, Fried GM. The MISTELS program to measure technical skill in laparoscopic surgery: evidence for reliability. *Surg Endosc.* May;20(5):744-7. 2006.
97. Fraser SA, Klassen DR, Feldman LS, Ghitulescu GA, Stanbridge D, Fried GM. Evaluating laparoscopic skills: setting the pass/fail score for the MISTELS system. *Surg Endosc.* 17(6):964-7. 2003.
98. Fried GM. FLS assessment of competency using simulated laparoscopic tasks. *J Gastrointest Surg.* 12(2):210-2. 2008.
99. Datta V, Chang A, Mackay S, Darzi A. The use of electromagnetic motion tracking analysis to objectively measure open surgical skills in the laboratory based model. *J Am Coll Surg.* 193: 479-485. 2001.
100. Taffinder N, Smith S, Mair J, Russel R, Darzi A. Can a computer measure surgical precision? Reliability, validity and feasibility of the ICSAD. *Surg Endosc* 13 (supl 1) : 81. 1999.
101. Smith SG, Torkington J, Brown TJ, Taffinder NJ, Darzi A. Motion analysis. *Surg Endosc* 16:640-5. 2002.
102. Francis NK, Hanna GB, Cuschieri A. The performance of master surgeons on the advanced Dundee endoscopic psychomotor tester: contrast validity study. *Arch Surg.* 137:841-4. 2002.
103. Francis NK, Hanna GB, Cuschieri A. Reliability of the Advanced Dundee Endoscopic Psychomotor Tester for bimanual tasks. *Arch Surg.* 136(1):40-43. 2001
104. McCloy R, Stone R. Science, medicine, and the future. Virtual reality in surgery. *BMJ.* 323:912-5. 2001.
105. Satava RM, Cuschieri A, Hamdorf J. Metrics for objective assessment. *Surg Endosc.* 17:220-6. 2003
106. Moorthy K, Jiwanji M, Shah J, Bello F, Munz Y, Darzi A. Validation of a web-based training tool for lumbar puncture. *Stud Health Technol Inform.* 94: 219-25. 2003.
107. Moorthy K, Munz Y, Jiwanji M, Bann S, Chang A, Darzi A. Validity and reliability of a virtual reality upper gastrointestinal simulator and cross validation using structured assessment of individual performance with video playback. *Surg Endosc.* 18(2): 328-33. 2004.

108. Szalay D, MacRae H, Regehr G, Reznick R. Using Operative Outcome to Assess Technical Skill. *Am J Surg* 180: 234-237. 2000.
109. Datta V, Mandalia M, Mackay S, Chang A, Chesire N, Darzi A. Relationship between skill and outcome in the laboratory based model. *Surg.* 131: 318-323. 2002.
110. Hanna GB, Frank TG, Cuschieri A. Objective assessment of endoscopic knot quality. *Am J Surg.* 174: 410-413. 1997.

## ANEXO 1

**CRITERIOS GENERALES DE DESEMPEÑO POR GRUPO**

Basado en la Teoría de Adquisición de Habilidades Motrices  
De Fitts y Posner.

<b>GRUPO 1</b>	<b>GRUPO 2</b>	<b>GRUPO 3</b>
<p>No puede Hace muy mal Muy poco eficiente</p> <p>Repite maniobras muchas veces (&gt;3) sin obtener resultados favorables</p> <p>Comete errores graves</p> <p>Transgresiones mayores</p>	<p>Puede, pero es poco efectivo</p> <p>Debe repetir algunas veces (&lt;3) para conseguir un resultado favorable</p> <p>Comete errores leves</p> <p>Transgresiones menores</p>	<p>Puede, y es eficiente</p> <p>Rara vez necesita repetir maniobras</p> <p>No comete errores</p> <p>Sin transgresiones</p>

## ANEXO 2

## ESCALA DE EVALUACIÓN GLOBAL DEL DESEMPEÑO TÉCNICO OSATS de R. Reznick

<b>1. Respeto por los tejidos.</b>				
<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
Daño tisular evidente  Aplicación frecuente de fuerza innecesaria Uso inapropiado del instrumental		Daño menor o inadvertido  Manipulación cuidadosa de los tejidos pero ocasionalmente aparece daño inadvertido menor		Ausencia de daño tisular  Manipulación apropiada del tejido con mínimo daño o ausencia completa de daño
<b>2. Tiempo y movimientos.</b>				
<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
Numerosos movimientos innecesarios Fuera de tiempo		Tiempos y movimientos eficientes, pero Algunos movimientos innecesarios		Clara economía de movimientos con máxima eficiencia
<b>3. Manejo del instrumental</b>				
<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
Movimientos vacilantes e inapropiados que realiza con el instrumental de manera repetitiva		Uso apropiado del instrumental, pero A veces parece tosco o inapropiado		Movimientos fluidos y apropiados
<b>4. Conocimiento de los instrumentos</b>				
<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
Pide o usa con frecuencia el instrumental incorrecto		Conoce los nombres de la mayoría del instrumental y lo usa apropiadamente para la tarea		Obviamente familiarizado con el nombre y el uso del instrumental
<b>5. Flujo y planificación de la cirugía</b>				
<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
Detiene con frecuencia la cirugía y tiene necesidad de pensar el siguiente paso		Habilidad demostrada de planificación con progresión continua del procedimiento		Curso quirúrgico obviamente planificado, sin esfuerzos evidentes de un paso al otro.
<b>6. Rol de los ayudantes</b>				
<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
Permanente error en la ubicación de los ayudantes o falla en el uso de los mismos		Correcto uso de los ayudantes la mayor parte del tiempo		Uso estratégico de los ayudantes con máxima ventaja en todo momento
<b>7. Conocimiento del procedimiento específico</b>				
<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
Conocimiento deficiente, necesita instrucciones específicas en la mayoría de los pasos quirúrgicos.		Conoce todos los aspectos importantes de la operación		Demuestra familiaridad con todos los pasos de la operación

## ANEXO 3

## ESCALA DE EVALUACIÓN PARA COLECISTECTOMÍA LAPAROSCÓPICA

1. Exposición de vesícula biliar. Uso de mano no hábil				
1 Tracción Incorrecta	2	3 Exposición Insuficiente	4	5 Exposición Apropiaada
Tracción cefálica en vez de lateral del bacinete, o falta de tracción.  Pinza izquierda manejada por el instructor.		Pobre tracción del fondo, insuficiente lateralización del bacinete  Poca utilidad de la pinza de mano izquierda: vesícula tomada pero "aplastada" contra el lecho		Tracción del fondo vesicular en sentido cefálico, y del bacinete en sentido lateral  Uso fluido de pinza de mano izquierda para exponer lateral derecho e izquierdo.
2. Disección del triángulo de Calot				
1 Disección a ciegas Disección ineficiente y/o temeraria	2	3 Disección poco eficiente o mal direccionada ( hasta 3 repeticiones)	4	5 Disección precisa y eficiente
Sin referencias anatómicas de seguridad.  Sin reconocer unión cístico bacinete y/o vía biliar principal.  3 o más intentos ineficaces Maniobras temerarias		Diseca la serosa en sentido de la vía biliar principal.  No logra abrir la serosa del triángulo de Calot  No abre la serosa de la cara derecha		Abre la serosa en la unión bacinete cístico y diseca hacia el lecho hepático sobre el triángulo de Calot.  Diseca la serosa que recubre la cara derecha. Maniobras precisas y eficientes (no requiere repetición, con margen de seguridad)
3. Disección, ligadura y sección del conducto cístico				
1 Disección, ligadura y sección fallidas (múltiples intentos fallidos)	2	3 Disección insuficiente y/o Clips mal localizados o aplicados (repite maniobras hasta 3 veces)	4	5 Disección, ligadura y sección efectivas
No puede liberar el conducto cístico por cuenta propia.  Los clips quedan mal colocados (oblicuos, parcialmente oclusivos o sin espacio entre ambos) y/o se sueltan espontáneamente La sección del conducto cístico suelta los clips		Libera parcialmente el conducto de estructuras adyacentes, y labra un pedículo demasiado corto.  Reubicación, o reaplicación de clips. Pedículo entre clips insuficiente que dificulta el corte.		Libera completamente el conducto cístico de estructuras adyacentes y labra un pedículo adecuado.  Correcta aplicación de clips, buena distancia entre ellos. Sección neta y apropiada del conducto.

<b>4. Identificación de la arteria cística, ligadura y sección</b>				
<b>1</b> <b>Disección, ligadura y/o coagulación fallidas</b>	<b>2</b>	<b>3</b> <b>Disección insuficiente, ligadura y/o coagulación poco eficaces que obliga a repetir gestos</b>	<b>4</b>	<b>5</b> <b>Disección, ligadura y/o coagulación efectivas</b>
<p>No logra discar la arteria (a pesar de ser factible).</p> <p>La arteria sangra y <u>no puede hacer hemostasia</u>, no logra identificar los cabos, o hace maniobras peligrosas (aplica puntos de sutura o clips a ciegas, coagula en proximidades de la vía biliar principal, y sin éxito en la maniobra.</p>		<p>Identifica las referencias anatómicas pero no consigue un pedículo suficiente para seccionar entre clips, o</p> <p>Coagula de manera ineficaz, uno de los cabos sangra. Debe retomar el cabo sangrante y volver a clipar.</p>		<p>Identifica las referencias anatómicas, disecciona la arteria logrando un pedículo suficiente para aplicar clips en posición proximal y distal, y seccionar entre ambos, o coagular. Los cabos quedan correctamente ligados.</p> <p>No liga pero coagula correctamente sin hemorragia residual de ningún cabo</p>
<b>5. Disección del lecho vesicular</b>				
<b>1</b> <b>Disección a ciegas, maniobras ineficientes y/o temerarias</b>	<b>2</b>	<b>3</b> <b>Disección poco eficiente e imprecisa</b>	<b>4</b>	<b>5</b> <b>Disección eficiente y precisa</b>
<p>Disección <u>fuera de plano</u>, con invasión del lecho hepático o ruptura grosera de la vesícula biliar.</p> <p>Eventual lesión vascular o biliar del pedículo hepático.</p>		<p>Disección <u>pierde el plano</u>: abre vesícula o invade lecho en zonas puntuales. Debe volver al plano de disección reiteradamente</p> <p>Diseca demasiado cerca de vía biliar principal sin tomar precauciones, o empleando el electrobisturí de manera temeraria en zonas de peligro.</p>		<p>Disección se mantiene <u>en el plano</u> de adosamiento.</p> <p>a distancia de la vía biliar principal, evitando el uso de electrocoagulación en zonas más próximas a vía biliar principal.</p>

## ANEXO 4

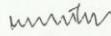
## ESCALA DE EVALUACIÓN PARA HERNIOPLASTIA INGUINAL

<b>1. Acceso al conducto inguinal y jalamiento del cordón.</b>				
<b>1</b> Disección ineficaz y/o temeraria Visualización pobre e insuficiente	<b>2</b>	<b>3</b> Disección poco eficaz y/o mal direccionada. Visualización intermitente.	<b>4</b>	<b>5</b> Disección precisa y efectiva. Correcta visualización
No reconoce ni logra identificar las estructuras en el paciente. No tiene en cuenta referencias anatómicas en la disección. Precisa directivas y correcciones constantes		Reconoce con dificultad las estructuras y referencias anatómicas. Requiere alguna indicación y/o corrección para ejecutar la maniobra apropiadamente.		Identifica correctamente las estructuras y referencias anatómicas que permiten acceder y jalar el cordón.
<b>2. Exploración y disección de los elementos del cordón (vasos, conducto deferente, y saco)</b>				
<b>1</b> Disección ineficaz y/o temeraria Visualización pobre e insuficiente	<b>2</b>	<b>3</b> Disección poco eficaz y/o mal direccionada. Visualización intermitente	<b>4</b>	<b>5</b> Disección precisa y efectiva. Correcta visualización
Sangrado significativo que interfiere la visualización de elementos. Disección a ciegas Falta de reconocimiento del conducto deferente, y/o daño parcial o total. No logra identificar saco herniario No logra liberar el saco de estructuras vecinas		Liberación parcial del saco de las estructuras vecinas. Requiere repetición?  Maniobras dubitativas.  Sangrado menor.		Libera correctamente (hasta el cuello) el saco, de los elementos vasculares, sin producir sangrado.  Reconoce conducto deferente y lo disecciona con cautela haciendo hincapié en conservar su indemnidad.
<b>3. Tratamiento del saco herniario (hernias indirectas)</b>				
<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
La disección, sutura y/o invaginación no permite la reducción del saco herniario		Reducción incompleta luego de jareta o sutura Requiere repetir disección y/o reducción		Saco ligado y seccionado en posición proximal (cuello) El muñón peritoneal tiende a retraerse hacia el interior de la cavidad.
<b>3'. Tratamiento del saco herniario (hernias directas)</b>				
<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
La disección, sutura y/o invaginación no permite la reducción del saco herniario		Reducción incompleta luego de jareta o sutura Requiere repetir disección y/o reducción		Reducción completa con jareta invaginante o sutura

<b>4. Ajuste del orificio inguinal profundo</b>				
<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
<p>No tiene en cuenta el tamaño del orificio inguinal profundo.</p> <p>No evalúa la necesidad de cerrarlo</p>		<p>Tiene en cuenta el tamaño del orificio pero no lo cierra efectivamente, o no calibra el cierre.</p>		<p>Evalúa la necesidad de achicamiento del orificio inguinal profundo</p> <p>Cierra y calibra su tamaño luego de la sutura.</p>
<b>5. Inserción y fijación de la prótesis</b>				
<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
<p>La prótesis no tiene el tamaño suficiente para cubrir la zona disecada.</p> <p>Deja al descubierto el triángulo de Hessert.</p> <p>La fijación es incorrecta y no se hace sobre los tejidos adecuados.</p>		<p>La prótesis tiene el tamaño adecuado pero su fijación requiere rehacer algunos puntos de sutura o volver sobre los pasos previos.</p>		<p>La prótesis cubre extensamente la zona inguinal disecada.</p> <p>Es fijada apropiadamente (arcada crural al tendón conjunto) con puntos separados o sutura continua.</p>

## ANEXO 5

**Certificados de presentación de los resultados preliminares de la investigación  
en la Conferencia Argentina de Educación Médica CAEM 2011  
Organizado por AFACIMERA – BUENOS AIRES**

	<b>XI Conferencia Argentina de Educación Médica CAEM 2011</b>		
<p><b>“La Enseñanza y Aprendizaje de Competencias: Moda o Necesidad?”</b> 14, 15 y 16 de septiembre de 2011. Buenos Aires, Argentina.</p> <p>Por el presente certificado se acredita que</p> <p align="center"><i>Andrés Carlos Kasparian</i></p> <p>ha participado en calidad de Asistente en la XI Conferencia Argentina de Educación Médica “CAEM’2011” Buenos Aires, 16 de Septiembre de 2011.</p>			
 Prof. Dr. Abraham Acosta Bustillo Presidente CAEM’2011	 Prof. Dr. Carlos Alvarez Vice-Presidente CAEM’2011	 Prof. Dr. Osvaldo Blanco Presidente Comité Organizador CAEM’2011	 Prof. Dr. Angel Centeno Presidente Comité Científico CAEM’2011
0621			

	<b>XI Conferencia Argentina de Educación Médica CAEM 2011</b>		
<p><b>“La Enseñanza y Aprendizaje de Competencias: Moda o Necesidad?”</b> 14, 15 y 16 de septiembre de 2011. Buenos Aires, Argentina.</p> <p>Por el presente certificado se acredita que</p> <p align="center"><b>Kasparian A, Chércoles R</b></p> <p>han presentado el trabajo científico: <b>“Evaluación objetiva de competencias técnicas en Cirugía.”</b></p> <p>en la XI Conferencia Argentina de Educación Médica “CAEM’2011” Buenos Aires, 16 de Septiembre de 2011.</p>			
 Prof. Dr. Abraham Acosta Bustillo Presidente CAEM’2011	 Prof. Dr. Carlos Alvarez Vice-Presidente CAEM’2011	 Prof. Dr. Osvaldo Blanco Presidente Comité Organizador CAEM’2011	 Prof. Dr. Angel Centeno Presidente Comité Científico CAEM’2011

Poster de presentación con resultados preliminares de la investigación en la  
**Conferencia Argentina de Educación Médica CAEM 2011**  
**Organizado por AFACIMERA – BUENOS AIRES**

**087 EVALUACION OBJETIVA DE COMPETENCIAS  
 TÉCNICAS EN CIRUGÍA**

**A Kasparian, R Chércoles**

Servicio de Cirugía N° I - Hospital Nacional de Clínicas - Facultad de Ciencias Médicas  
 Universidad Nacional de Córdoba, Argentina.



**OBJETIVOS**

- Proponer y aplicar un instrumento de evaluación objetiva de competencias técnicas en cirugía.

**MATERIAL Y MÉTODO**

- Modelo a evaluar: hernioplastia inguinal con técnica de Lichtenstein.

- Cirujanos expertos vs. cirujanos novatos.

- Filmación anónima de los procedimientos y evaluación posterior del registro en formato DVD, según tablas 1 y 2.

**ESCORE DE EVALUACIÓN GLOBAL**

**1. Campo operatorio**

1	2	3	4	5
El espacio de trabajo no pudo ser creado o mantenido adecuadamente a lo largo del procedimiento. Pobre visualización de estructuras.		No se puede mantener el espacio de trabajo de manera constante. Visualización intermitente de estructuras.		Creación y mantenimiento de un espacio de trabajo adecuado. Correcta visualización de las estructuras sobre las cuales se realiza el procedimiento.

**3. Exploración y disección de los elementos cordonaes:**

1	2	3	4	5
Disección dificultosa, poco efectiva, y/o con sangrado de elementos vasculares. Falta de reconocimiento del conducto deferente, daño parcial o total del mismo.		No logra aislar el saco herniario de manera completa hasta el cuello, y/o produce sangrado, y/o no tiene en cuenta al conducto deferente mientras disecciona.		Libera correctamente el saco herniario de los elementos cordonaes sin producir sangrado. Reconoce el conducto deferente y lo disecciona con cautela haciendo hincapié en preservar su indemnidad.

**RESULTADOS**

- Se evaluaron 12 cirugías.
- Cirujanos novatos 7 (siete)
- Cirujanos expertos 5 (cinco)
- Confiabilidad 0,83 para el instrumento global y 0,81 para la específica.
- Las diferencias entre los grupos fueron significativas ( $p < 0,05$ )

**CONCLUSIONES**

- Definición de categorías de evaluación
- Calificación numérica según desempeño
- Posibilidad de volver sobre la ejecución del procedimiento
- Evaluación a ciegas
- Validez de constructo entre grupos extremos
- Confiabilidad alta
- Pendiente: discriminar grupos de entrenamientos intermedios

## ANEXO 6

**Revisión publicada en la Revista de la Facultad de Ciencias Médicas  
Año 2011**

## TRABAJO DE REVISIÓN

**NUEVOS INSTRUMENTOS DE EVALUACIÓN OBJETIVA DE COMPETENCIAS TÉCNICAS EN CIRUGÍA**

NEW TOOLS FOR OBJECTIVE ASSESSMENT OF TECHNICAL SKILLS IN SURGERY.

Andrés C. Kasparian, Ricardo Chércoles

**INTRODUCCIÓN**

La adquisición de habilidades quirúrgicas constituye un factor central en la formación de todo cirujano. No hay duda que la destreza técnica es considerada por el público en general y por los propios médicos en formación como uno de los factores más importantes de la lista de cualidades que debería poseer todo cirujano.

Una encuesta publicada en el año 2001 realizada con cirujanos del Reino Unido que alcanzaron el grado de “maestro”, reveló que más del 80% de los encuestados identificaron la excelencia en las habilidades psicomotoras como el factor más determinante de una carrera quirúrgica exitosa (6, 7). Otro reporte publicado en el mismo año proveniente de centros europeos y norteamericanos, basado en encuestas y reuniones de consenso con cirujanos “maestros” de diferentes especialidades, consideró a la destreza operatoria como el atributo de mayor importancia luego del juicio clínico (8, 35).

Sin embargo, dentro de todos los requisitos incluidos en la formación de un cirujano, la evaluación de las habilidades técnicas constituye el factor más débil y menos desarrollado. En la actualidad las herramientas de evaluación de los programas de residencias quirúrgicas hacen fuertemente hincapié en la adquisición de conocimientos y en la aplicación de los mismos en forma criteriosa, en tanto que los recursos para evaluar las competencias técnicas se caracterizan por presentar subjetividad, y falta de confiabilidad y validez (11, 29).

Este hecho guarda relación con la vigencia hasta la actualidad de un modelo clásico de formación en cirugía introducido por Sir William Halsted en el año 1898 en el John Hopkins Hospital de los Estados Unidos de Norteamérica. Dicho modelo, diseñado a semejanza del sistema alemán, estaba basado en una relación maestro-aprendiz y en

la adquisición de responsabilidades graduadas y progresivas. El entrenamiento estaba regido por el concepto de “ver y hacer”, es decir, el aprendiz trabajaba a la par de su maestro para asegurar la progresión de sus habilidades por imitación. La solvencia del aprendiz estaba determinada por el maestro de acuerdo a su desempeño. Esta evaluación estaba basada muchas veces en rasgos de la personalidad más que en criterios objetivos (16,24).

Sin embargo, los avances en las teorías de la educación, y las presiones crecientes en la práctica profesional han llevado a cuestionar seriamente la confianza en este sistema para la adquisición de habilidades técnicas. La creciente complejidad de los casos hospitalarios, y el mayor énfasis en mitigar el error médico por los problemas médico legales, ha modificado de manera significativa la posibilidad de los instructores para formar a los residentes. La limitación reglamentaria de la carga horaria para los residentes, y el interés en optimizar el uso del quirófano han determinado una menor exposición de los médicos en formación a los procedimientos con pacientes en la sala de operaciones.

Surge por estos motivos la necesidad de un nuevo modelo de educación en cirugía basado en un programa estructurado de formación, en el cual el entrenamiento de las técnicas operatorias comienza fuera del quirófano, en un laboratorio, y la adquisición y evaluación de competencias se realiza con instrumentos que pretenden ser lo más objetivos posibles (1,16, 24, 28).

**COMPETENCIAS Y SU EVALUACIÓN**

Kane (19) definió la competencia como: “... el grado de utilización de los conocimientos, las habilidades y el buen juicio asociados a la profesión, en todas las situaciones que se pueden confrontar en el ejercicio de la práctica profesio-

nal" es decir que la competencia representa la capacidad de un profesional de utilizar su buen juicio así como también sus conocimientos, habilidades, y actitudes para solucionar problemas complejos que se presentan en el ámbito de su actividad.

El concepto de competencia es una construcción conceptual compleja que se puede abordar de manera parcial a través de los diferentes aspectos mesurables que la constituyen: conocimientos, habilidades técnicas, actitudes, juicio, capacidad resolutoria frente a nuevos problemas, capacidad de comunicación, etc.

Hay muchos motivos por los cuales hay que evaluar las competencias tanto en estudiantes como en profesionales. Se pueden mencionar: 1) la evaluación como requisito integrante del proceso educativo 2) para certificar la competencia (evaluación sumativa) 3) para retroalimentar el proceso educativo (evaluación formativa) 4) para evaluar los propios programas de formación 5) para confirmar los valores y los estándares de la profesión y 6) por la responsabilidad e imputabilidad frente a la sociedad.

Miller (23) ha desarrollado un modelo de competencia profesional representado por una pirámide compuesta de varios niveles. En la base se encuentran los conocimientos (saber) sobre los que se apoya la competencia (saber como). A un nivel superior se encuentra el desempeño (mostrar como) y finalmente la acción en una práctica real (el hacer).

En 1996 Van Der Vleuten (34) propone un modelo teórico conceptual para definir la utilidad de una herramienta de evaluación de competencias profesionales que correlaciona distintas propiedades y que puede expresarse con la siguiente ecuación:

U: utilidad del instrumento; F: fidelidad o confiabilidad; V: validez; E: impacto educativo; A: aceptabilidad; C: costo; (p) contribución parcial. Modificado de Van Der Vleuten C. 1996. The assessment of professional competence: developments, research and practical implications. *Adv Health Sci Educ Theory Pract.* 1: 41-67

La confiabilidad puede definirse como la expresión cuantitativa de la reproducibilidad con la que un instrumento mide el mismo atributo. Hace re-

ferencia a la precisión del test. Es decir, en que medida el test genera los mismos resultados si se repite en dos sucesivas oportunidades bajo las mismas circunstancias, o sea, sin la modificación propia del aprendizaje entre las tomas (15, 30). Se representa con el coeficiente alfa de Cronbach. Cuando las evaluaciones son llevadas a cabo por más de un observador se aplica también el concepto de confiabilidad inter-evaluadores, que mide el grado de concordancia entre dos o más evaluadores que utilizan el mismo instrumento (25).

La validez es un concepto complejo que hace referencia a la capacidad del instrumento de evaluación de medir o no lo que se propone como objetivo. Puede desglosarse en distintos subconceptos: validez de constructo hace referencia a la medida en la cual el examen puede diferenciar distintos niveles de experiencia. Validez de contenido refiere si el objeto o campo de conocimiento que pretende ser evaluado es medido por el instrumento. Por ejemplo: si se pretende evaluar aspectos técnicos y se terminan evaluando conocimientos, el instrumento de evaluación carece de validez de contenido. Validez concurrente es el grado en el cual los resultados de la evaluación se correlacionan con el gold Standard para ese campo del conocimiento. Validez de superficie es el grado en el cual el instrumento tiene el aspecto de situaciones reales. Validez predictiva la habilidad del examen de predecir el desempeño a futuro (14, 23, 29, 31, 33).

Reznick (27) en el año 1993 reportaba cinco métodos disponibles para la evaluación de las competencias técnicas, con grados variables de confiabilidad y validez:

Los exámenes de certificación administrados por las instituciones formadoras o reguladoras al final de un período lectivo o luego de cumplir con el programa de formación, basan su evaluación en los conocimientos y las habilidades clínicas, y no evalúan la capacidad técnica del candidato (30).

Los registros de procedimientos (cuadernos de bitácora) efectuados a lo largo del proceso de formación son obligatorios en los programas de residencia de algunos lugares del mundo como el Reino Unido, y son requeridos al momento de las evaluaciones de fin de año o final de residencia, entrevistas de trabajo, etc. Sin embargo ha

## TRABAJO DE REVISIÓN

sido demostrado que es un registro meramente indicativo de la ejecución de los procedimientos y no un reflejo de la habilidad operatoria, es decir carecen de validez de constructo.

El registro del tiempo empleado para la ejecución de los procedimientos tampoco evalúa la calidad de la técnica, y es un parámetro no confiable durante procedimientos reales debido a la influencia de numerosas circunstancias que pueden no ser previstas, y que tendrían poca relación con el desempeño del cirujano.

La observación directa tal como se realiza actualmente en el quirófano en casi todas las residencias de cirugía, es global y no está basada en criterios específicos, por lo tanto no resulta confiable. Como está influenciada por la subjetividad del observador tiene una pobre confiabilidad de reevaluación e interobservador. A demás los procedimientos quirúrgicos son únicos y efímeros, por lo tanto no hay posibilidad de revisar el desempeño del candidato en caso de dudas, y esto aumenta aún más la subjetividad y la falta de confiabilidad.

Los datos de morbilidad y mortalidad con frecuencia son utilizados como marcadores del desempeño operatorio, sin embargo esta información está influenciada en gran medida por factores dependientes del paciente: variabilidad biológica, disparidad en la presentación de las enfermedades, etc. y por otros factores no relacionados con los pacientes: soporte institucional, acceso a los recursos, etc. Por lo tanto estas cifras carecen de confiabilidad, y no reflejan verdaderamente la competencia técnica.

Como consecuencia de los problemas de confiabilidad y subjetividad de los métodos mencionados previamente, se han propuesto y desarrollado métodos más objetivos de evaluación de habilidades técnicas, entre los que se pueden mencionar:

Las listas de cotejo y evaluaciones globales: la enumeración de criterios definidos para cotejar la ejecución o no de algún procedimiento hace que el proceso de evaluación sea más objetivo, válido y confiable. Se transformaría a los evaluadores en observadores más que en interpretadores de la conducta, tendiendo a minimizar de esta forma la subjetividad en el proceso de evaluación (26).

La amplia aceptación del examen clínico objetivo estructurado (ECOE) llevó a un gru-

po de Toronto a desarrollar un concepto similar aplicado a la evaluación de habilidades técnicas (OSATS), que consiste en 6 estaciones donde los candidatos desarrollan procedimientos quirúrgicos estandarizados en modelos animales o inanimados, durante un período limitado de tiempo. El desempeño es evaluado con una lista de cotejo específica de la consigna a desarrollar, y una evaluación del desempeño global que consiste en 7 componentes genéricos de destreza operatoria que son cuantificados con una escala de 5 puntos de Likert, con descriptores explícitos para los puntos extremos y el medio (22, 29). El inconveniente de este tipo de evaluaciones es la inversión en recursos humanos y el tiempo que debe dedicarse en el análisis de las grabaciones u observación directa (26).

Los sistemas de análisis de destreza manual como el ICSAD (Imperial Collage Surgical Assessment Device) es un sistema de rastreo, disponible comercialmente, que consiste en un generador de campo electromagnético y dos sensores que son colocados en las manos del cirujano. Hay un procesador de datos que utiliza un software para medir el número y velocidad de movimientos, la distancia recorrida por las manos y el tiempo para ejecutar la tarea. Diversos estudios han demostrado la validez de constructo de este sistema (9).

Los sistemas de realidad virtual pueden definirse como un conjunto de tecnologías que permiten a un individuo interactuar eficientemente con bases de datos computarizadas en tres dimensiones y en tiempo real, usando sus órganos de los sentidos naturales y habilidades. Se puede mencionar el MIST-VR (Minimally Invasive Surgical Trainer – Virtual Reality) reproduce una serie de imágenes tridimensionales que pueden ser manipuladas, fue ideado como entrenamiento para las colecistectomías laparoscópicas. Los datos que surgen del uso de este sistema se pueden medir y evaluar, y MIST-VR ha sido validado ampliamente para la evaluación de habilidades básicas en cirugía laparoscópica (25).

El análisis de los productos finales: los resultados de la cirugía y la posterior evolución del paciente no se pueden atribuir a cuestiones técnicas exclusivamente. De hecho las consecuencias de una mala técnica pueden ponerse en evidencia muchos años después. Szalay y col. (32) propu-

sieron la evaluación del producto final de los modelos inanimados que constituyen las diferentes estaciones del OSATS. Sus estudios demostraron validez de constructo y una muy buena correlación con la evaluación del OSATS en vivo, superando de esta forma los inconvenientes derivados de las observaciones en vivo y las grabaciones. Datta et al. analizaron la tasa de filtrado y el área de sección transversal de modelos de anastomosis vasculares (10). Hanna et al. estudiaron la calidad de los nudos hechos por vía laparoscópica usando un tensiómetro y desarrollaron una escala de calidad y confiabilidad (17). En resumen, el conocimiento actual sobre el tema pone en evidencia falencias de peso en los métodos de evaluación de las habilidades técnicas vigentes en la actualidad. Los métodos tradicionales vigentes en la mayoría de los centros formadores de nuestro medio carecen de validez y confiabilidad en relación a la destreza quirúrgica. La formación quirúrgica dentro del marco de un programa formal estructurado de residencias médicas requiere una retroalimentación objetiva constante de la adquisición de habilidades técnicas para corrección y reorientación, y para ello precisa de instrumentos de evaluación objetivos. Todas las profesiones en general, relacionadas o no con la salud, implican la ejecución de procedimientos que requieren la adquisición de destrezas técnicas de variada complejidad. La formación de profesionales de ámbitos tan disímiles como son un piloto de avión y un cirujano, en apariencia sin ningún punto de contacto, tienen en común el hecho que ambas requieren el desarrollo de habilidades técnicas y que ambas deben enfrentar condiciones en las cuales, eventualmente, se pone en riesgo la vida. Muchos aspectos relacionados con el entrenamiento y la evaluación de los pilotos de avión podrían ser aplicables a la formación de los cirujanos. (24) Existen métodos validados de evaluación objetiva que podrían comenzar a utilizarse para evaluar las habilidades quirúrgicas con fines formativos y de certificación. Aunque se requieren futuras investigaciones antes de ser utilizados para evaluaciones sumativas y de revalidación (2, 3, 5, 12, 13, 20, 21).

## BIBLIOGRAFÍA

1. ACGME. *Outcome Project*. <http://www.acgme.org/Outcome/>
2. Anderson CI, Jentz AB, Harkema JM, Rao Kareti L, Apelgren KN, Slomski CA. 2005. *Assessing the competencies in general surgery residency training*. *Am J Surg* 189: 288-292
3. Bann S, Khan M, Datta V, Darzi A. 2005. *Surgical skill is predicted by the ability to detect errors*. *Am J Surg*. 189: 412-415.
4. Beard JD. 2007. *Assessment of Surgical Competence*. *BJS* 94: 1315-1316
5. Boehler ML, Schwind CJ, Rogers DA, Ketchum J, O'Sullivan E, Mayforth R, Quin J, Wohltman C, Johnson C, Williams RG, Dunnington G. 2007. *A Theory-Based Curriculum for Enhancing Surgical Skillfulness*. *J Am Coll Surg*. 205: 492-497.
6. Cosman P, Hemli JM, Ellis AM, Hugh TJ. 2007. *Learning the Surgical Craft: A Review of Skills Training Options*. *ANZ J. Surg*. 77: 838-845.
7. Cuschieri A. 2001. *Master surgeons and surgical proficiency*. En *Metrics: Objective Assessment of Surgical Skills – Developing Quantitative Measurements through Surgical Simulation*. Scottsdale, Arizona: US Army Medical Research Command, Telemedicine and Advances Technology Research Center (TATRC)
8. Cuschieri A, Francis N, Crosby J, Hanna GB. 2001. *What do Master Surgeons Think of Surgical Competence and Revalidation*. *Am J Surg* 182:110-116.
9. Datta V, Chang A, Mackay S, Darzi A. 2001. *The use of electromagnetic motion tracking analysis to objectively measure open surgical skills in the laboratory based model*. *J Am Coll Surg*. 193: 479-485.
10. Datta V, Mandalia M, Mackay S, Chang A, Chesire N, Darzi A. 2002. *Relationship between skill and outcome in the laboratory based model*. *Surg*. 131: 318-323.
11. Darzi A, Datta V, Mackay S. 2001. *The Challenge of Objective Assessment of Surgical Skill*. *Am J Surg* 181: 484-486.
12. Dath D, Regehr G, Birch D, Schlachta C, Poulin E, Mamazza J, Reznick R, MacRae H. 2004. *Toward Reliable Operative Assessment*. *Surg Endosc* 18: 1800-1804.
13. Doyle JD, Webber EM, Sidhu RS. 2007. *A universal global rating scale for the evaluation of technical skills in the operating room*. *Am J Surg*. 193: 551-555.

## TRABAJO DE REVISIÓN

14. Goff BA, Lentz GM, Lee D, Fenner D, Morris J, Mandel LS. 2001. Development of a Bench Station Objective Structured Assessment of Technical Skills. *Obstet Gynecol.* 98: 412-416.
15. Greenberg JA, Irani JL, Greenberg CC, Blanco MA, Lipsitz S, Ashley SW, Breen EM, Hafner JP. 2007. The ACGME competencies in the operating room. *Surgery.* 142: 180-184.
16. Hamford JM, Hall JC. 2000. Acquiring Surgical Skills. *BJS* 87: 28 – 37.
17. Hanna GB, Frank TG, Cuschieri A. 1997. Objective assessment of endoscopic knot quality. *Am J Surg.* 174: 410-413.
18. Intercollegiate Surgical Curriculum Programme. <http://www.iscp.ac.uk>.
19. Kane MT. 1992. The assessment of clinical competence. *Eval Health Prof.* 15(2): 163-182.
20. Larson JL, Williams RG, Ketchum J, Boehler M, Dunnington GL. 2005. Feasibility, reliability and validity of an operative performance rating system for evaluating surgery residents. *Surgery.* 138: 640-649.
21. MacRae H, Regehr G, Leadbetter W, Reznick R. 2000. A Comprehensive Examination for Senior Surgical Residents. *Am J Surg.* 179: 190-193.
22. Martin JA, Regehr G, Reznick R, MacRae H, Murnaghan J, Hutchinson G. 1997. Objective structured assessment of technical skill (OSATS) for surgical residents. *Br J Surg.* 84: 273-278.
23. Miller G. 1990. The assessment of clinical skills/competence/performance. *Academic Medicine.* 65: S63-S67.
24. Moen MD, Moen RO, Moen O.L. 2007. Teaching and Evaluating Technical Skills: Applying the Aviation Model to Gynecologic Surgery. *J Pelvic Med Surg.* 13: 107-112
25. Moorthy K, Munz Y, Sarker SK, Darzi A. 2003. Objective assessment of technical skills in surgery. *BMJ,* 327: 1032-1037
26. Regehr G, MacRae H, Reznick R, Szalay D. 1998. Comparing the psychometric properties of checklists and global rating scales for assessing performance on an OSCE format examination. *Acad Med.* 73: 993-997.
27. Reznick R. 1993. Teaching and testing surgical skills. *Am J surg.* 165: 358-361.
28. Reznick R, MacRae H. 2006. Teaching Surgical Skills – Changes in the Wind. *N Engl J Med* 355: 2664-2669.
29. Reznick R, Regehr G, MacRae H, Martin J, McCulloch W. 1997. Testing Technical Skill via an Innovative “Bench Station” Examination. *Am J Surg.* 173: 226–230.
30. Scott DJ, Valentine RJ, Bergen PC, Regehr RV, Laycock R, Tesfay ST. 2000. Evaluating surgical competency with the American Board of Surgery in-training examination, skill testing, and intraoperative assessment. *Surgery.* 128: 613-622.
31. Smee S. 2003. ABC of learning and teaching in medicine. Skill based assessment. *BMJ.* 326: 703-706.
32. Szalay D, MacRae H, Regehr G, Reznick R. 2000. Using Operative Outcome to Assess Technical Skill. *Am J Surg* 180: 234-237.
33. Van Der Vleuten C, Schuwirth L. 2005. Assessing professional competence: from methods to programmes. *Medical Education.* 39: 309-317.
34. Van Der Vleuten C. 1996. The assessment of professional competence: developments, research and practical implications. *Adv Health Sci Educ Theory Pract.* 1: 41-67.
35. Yule S, Flin R, Paterson-Brown S, Maran N, 2006. Non-technical Skills for surgeons in the operating room: A review of the literature. *Surgery.* 139: 140-149.

## ANEXO 7

**Comunicación Oral de resultados finales, ganadora del 3er Premio en la  
Conferencia Argentina de Educación Médica CAEM 2014  
Organizado por AFACIMERA – BUENOS AIRES**



acredita que

**Kasparian AC, Martínez AC, Jover Clos RJ, Chércoles RA**

han presentado la Comunicación Oral  
*"Evaluación objetiva de competencias técnicas en cirugía"*

en la **XIV CONFERENCIA ARGENTINA DE EDUCACIÓN MÉDICA**  
en el marco de **"La calidad en la formación de recursos humanos en salud"**

Ciudad Autónoma de Buenos Aires, 27, 28, 29 de Octubre de 2014.

  
Dr. Roberto Cherjovsky  
Presidente Comité Científico CAEM 2014

  
Dr. Gerardo Omar Larroza  
Presidente AFACIMERA

  
Dr. Miguel Ángel Schiavone  
Presidente CAEM 2014

2014



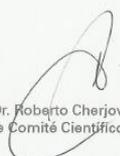
A los 29 días del mes de Octubre de 2014, certificamos que el trabajo  
**"Evaluación objetiva de competencias técnicas en Cirugía"**

presentado por

**Dres. Kasparian A.C; Martínez A.C.; Jover Clos R.J., Chércoles R.A.**

en el marco del **XIV CONFERENCIA ARGENTINA DE EDUCACIÓN MÉDICA**  
**"La calidad en la formación de recursos humanos en salud"** se ha hecho acreedor  
al **TERCER PREMIO** en el área de "Educación Médica".

Ciudad Autónoma de Buenos Aires.

  
Dr. Roberto Cherjovsky  
Presidente Comité Científico CAEM 2014

  
Dr. Gerardo Omar Larroza  
Presidente AFACIMERA

  
Dr. Miguel Ángel Schiavone  
Presidente CAEM 2014

2014

**Título:** EVALUACIÓN OBJETIVA DE COMPETENCIAS TÉCNICAS EN CIRUGÍA

**Autores:** Kasparian AC, Martínez AC, Jover Clos RJ, Chércoles RA.

**Correo electrónico de contacto:** kaspariandres@gmail.com

**Institución:** Hospital Nacional de Clínicas de Córdoba – Facultad de Cs. Médicas – Universidad Nacional de Córdoba

## RESUMEN

**Introducción:** La adquisición de habilidades quirúrgicas constituye un factor central en la formación de todo cirujano. Sin embargo, la evaluación de las habilidades técnicas es uno de los factores más débiles y menos desarrollados. En la actualidad los recursos para evaluar las competencias técnicas adolecen de subjetividad, falta de confiabilidad y validez. La observación directa, método de evaluación más frecuentemente utilizado en nuestro medio, presenta inconvenientes instrumentales y está fuertemente influenciada por las relaciones intersubjetivas y los rasgos de personalidad. El objetivo de esta investigación es proponer creación y el uso de un instrumento objetivo para evaluar el desempeño técnico y determinar su confiabilidad y validez.

**Material y método:** se seleccionaron dos procedimientos: colecistectomía laparoscópica y hernioplastia inguinal (técnica de Lichtenstein). Se constituyeron tres grupos de comparación según la experiencia quirúrgica: inicial, intermedio, y expertos. Se filmaron las cirugías en tiempo real, sin identificación del paciente ni del cirujano. Las filmaciones sin edición fueron asignadas a dos cirujanos expertos en forma aleatoria por sorteo y con sobres sin identificación. Para la evaluación se propuso el uso de un instrumento objetivo (explicitación de pasos a evaluar y cuantificación mediante escala de Likert) y específico para cada procedimiento. Así mismo se utilizó la escala global OSATS (R. Reznick). Se aplicó análisis de varianza no paramétrico para determinar validez. Valores de  $p$  menores a 0.05 fueron considerados estadísticamente significativos. Valores superiores a 0,80 del Coeficiente alfa de Cronbach aseguraron confiabilidad.

**Resultados:** Desde Abril del 2010 hasta Diciembre del 2012 se filmaron 36 colecistectomías videolaparoscópicas y 31 hernioplastias inguinales. Se encontraron diferencias significativas entre los grupos en todos los ítems evaluados  $p < 0.05$ . El coeficiente  $\alpha$  de Cronbach fue mayor a 0,80 para ambas técnicas. No hubo diferencias significativas entre las calificaciones de ambos evaluadores. No hubo diferencias entre nuestro instrumento específico y la escala global OSATS.

**Discusión:** Es posible y útil aplicar un instrumento objetivo de evaluación del desempeño técnico en cirugía. La herramienta presentó validez de constructo y confiabilidad aceptables. La filmación confiere perdurabilidad a un evento efímero: la cirugía. La objetividad se basa en la enunciación y cuantificación explícita de cada paso, y en la aleatorización y anonimato de la muestra. La uniformidad de criterios entre los evaluadores es fundamental para obtener resultados satisfactorios. Evaluar siempre implicará un recorte de la realidad.

## ANEXO 8

**Artículo Original publicado en la Revista de la Facultad de Ciencias Médicas  
Año 2014**

## EVALUACIÓN OBJETIVA DE COMPETENCIAS TÉCNICAS EN CIRUGÍA

Autores: Kasparian AC, Martínez AC, JoverClos RJ, Chércoles RA.

**Resumen**

**Introducción:** La adquisición de habilidades quirúrgicas constituye un factor central en la formación de todo cirujano. Sin embargo, la evaluación de las habilidades técnicas es uno de los factores más débiles y menos desarrollados. En la actualidad los recursos para evaluar las competencias técnicas adolecen de subjetividad, falta de confiabilidad y validez. La observación directa, método de evaluación más frecuentemente utilizado en nuestro medio, presenta inconvenientes instrumentales y está fuertemente influenciada por las relaciones intersubjetivas y los rasgos de personalidad. El objetivo de esta investigación es proponer creación y el uso de un instrumento objetivo para evaluar el desempeño técnico y determinar su confiabilidad y validez.

**Material y método:** se seleccionaron dos procedimientos: colecistectomía laparoscópica y hernioplastia inguinal (técnica de Lichtenstein). Se constituyeron tres grupos de comparación según la experiencia quirúrgica: inicial, intermedio, y expertos. Se filmaron las cirugías en tiempo real, sin identificación del paciente ni del cirujano. Las filmaciones sin edición fueron asignadas a dos cirujanos expertos en forma aleatoria por sorteo y con sobres sin identificación. Para la evaluación se propuso el uso de un instrumento objetivo (explicitación de pasos a evaluar y cuantificación mediante escala de Likert) y específico para cada procedimiento. Así mismo se utilizó la escala global OSATS (R. Reznick). Se aplicó análisis de varianza no paramétrico para determinar validez. Valores de p menores a 0.05 fueron considerados estadísticamente significativos. Valores superiores a 0,80 del Coeficiente alfa de Cronbach aseguraron confiabilidad.

**Resultados:** Desde Abril del 2010 hasta Diciembre del 2012 se filmaron 36 colecistectomías videolaparoscópicas y 31 hernioplastias inguinales. Se encontraron diferencias significativas entre los grupos en todos los ítems evaluados  $p < 0.05$ . El coeficiente  $\alpha$  de Cronbach fue mayor a 0,80 para ambas técnicas. No hubo diferencias significativas entre las calificaciones de ambos evaluadores. No hubo diferencias entre nuestro instrumento específico y la escala global OSATS.

**Discusión:** Es posible y útil aplicar un instrumento objetivo de evaluación del desempeño técnico en cirugía. La herramienta presentó validez de constructo y confiabilidad aceptables. La filmación confiere perdurabilidad a un evento efímero: la cirugía. La objetividad se basa en la enunciación y cuantificación explícita de cada paso, y en la aleatorización y anonimato de la muestra. La uniformidad de criterios entre los evaluadores es fundamental para obtener resultados satisfactorios. Evaluarsiempreimplicaráunrecorte de la realidad.

**Summary**

**Introduction:** technical skills acquisition is considered to be of paramount importance in surgical training. Yet, formal assessment of technical skills is the weakest and less developed area. Currently available resources to evaluate technical skills are largely subjective, and lack of validity and reliability. Direct observation, one of the most frequently used methods, is largely biased by interpersonal subjectivity and personality traits. We propose the creation and use of a new procedure-specific tool for objective assessment of technical skills in surgery to evaluate validity and reliability.

**Materials and Methods:** laparoscopic cholecystectomy and Lichstenstein's inguinal hernia repair were the chosen procedures. Three groups of comparison were defined according to surgical expertise: initial, intermediate, and experts. Surgeries were videorecorded in real time without identification of the patient or the surgeon. Tapes without any posterior edition were assigned to two expert surgeons in a blind and randomized sequence. A newly proposed procedure-specific rating scale was used for evaluation, as well as Reznick's OSATS global scale. Kruskal-Wallis non-parametric test was used to assess validity.  $p < 0,05$  was considered statistically significant. Chronbach's  $\alpha$  values  $> 0.8$  granted reliability.

**Results:** from April 2010 to December 2012 36 laparoscopic cholecystectomies and 31 inguinal hernia repairs were recorded. Significant difference was found among groups of comparison for every item ( $p < 0.05$ ). Chronbach's  $\alpha$  was largely over 0,80 for both procedures. There was no significant difference between the assessments of the two raters. There was no significant difference between our newly proposed procedure-specific instrument and Reznick's OSTATS.

**Discussion:** the objective assessment of technical skills in surgery is feasible and useful. The tool we proposed showed construct validity and reliability. Video recording of surgical procedures grants durability over time to an ephemeral phenomenon. The objectivity is based on the explicit statements and quantification of every step to be evaluated, and the blind randomization and anonymous treatment of the sample. Sharing the same quality criteria between evaluators is of paramount importance to reach satisfactory results. The process of evaluation always implies a shortened view of the reality.

## Introducción

La adquisición de habilidades quirúrgicas constituye un factor central en la formación de todo cirujano. No hay duda que la destreza técnica es considerada por el público en general y por los propios médicos en formación como uno de los factores más importantes de la lista de cualidades que debería poseer todo cirujano.

Una encuesta publicada en el año 2001 realizada con cirujanos del Reino Unido que alcanzaron el grado de "maestro", reveló que más del 80% de los encuestados identificaron la excelencia en las habilidades psicomotoras como el factor más determinante de una carrera quirúrgica exitosa<sup>(1,2)</sup>. Otro reporte publicado en el mismo año, basado en encuestas y reuniones de consenso, realizado con cirujanos de diferentes especialidades con grado de "maestro" provenientes de centros europeos y norteamericanos, consideró a la destreza operatoria como el atributo de mayor importancia luego del juicio clínico<sup>(3,4)</sup>.

Sin embargo, dentro de todos los requisitos in-

cluidos en la formación de un cirujano la evaluación de las habilidades técnicas constituye, parece ser, el factor más débil y menos desarrollado. En la actualidad las herramientas de evaluación de los programas de residencias quirúrgicas hacen fuertemente hincapié en la adquisición de conocimientos y en la aplicación de los mismos en forma criteriosa, en tanto que los recursos para evaluar las competencias técnicas adolecen de subjetividad, falta de confiabilidad y validez<sup>(5,6)</sup>.

En nuestro ámbito, la evaluación de las competencias técnicas se hace presente habitualmente en dos instancias: al final de cada período lectivo del programa de formación de postgrado (residencia de cirugía) que constituye un requisito formal para la promoción anual, sumada a otros exámenes orientados a evaluar conocimientos específicos. La otra, guarda relación con los requisitos de certificación profesional para la obtención del título de especialista. En ambos casos, hay un equipo de evaluadores constituido por cirujanos con experiencia y vinculados a la tarea formativa,

que se hace presente en el quirófano con el propósito de presenciar una cirugía asignada al candidato y emitir luego un veredicto calificador.

Esta forma de evaluación conocida como observación directa presenta algunos inconvenientes que se relacionan con los siguientes hechos inherentes al método en sí mismo:

- 1.- El procedimiento quirúrgico reviste siempre la calidad de efímero, o sea es transitorio con un inicio y un final, y unidireccional, o sea no permite la vuelta al principio ni el retroceso.
- 2.- Las posibilidades físicas de observación directa en el quirófano no son óptimas: el campo quirúrgico suele ser estrecho o demasiado profundo, la iluminación puede no ser la adecuada para el observador, el riesgo de contaminación hace difícil el acercarse en proximidad, y la visualización es intermitente.
- 3.- El proceso de evaluación en sí mismo se lleva a cabo en una instancia posterior a la cirugía, fuera del ámbito del quirófano, y exige al evaluador un ejercicio de memoria.
- 4.- El evaluador no tiene definido con claridad qué criterios debe evaluar, ni cómo debe calificar, y por lo tanto termina apelando a la intuición.

La modalidad instrumental aplicada para evaluar bajo estas condiciones queda fuertemente influenciada por las relaciones intersubjetivas y hace que el resultado esté sesgado por los rasgos de personalidad y los vínculos de afectividad en lugar de estar basado en criterios objetivos, resultando por lo tanto poco confiable.

Este hecho guarda relación con la vigencia hasta la actualidad de un modelo clásico de formación en cirugía introducido por Sir William Halsted en el año 1898 en el John Hopkins Hospital de los Estados Unidos de Norteamérica. Dicho modelo, diseñado a semejanza del sistema alemán, estaba basado en una relación maestro-aprendiz y en la adquisición de responsabilidades graduadas y progresivas. El entrenamiento estaba regido por el concepto de "ver y hacer", es decir, el aprendiz trabajaba a la par de su maestro para asegurar la progresión de sus habilidades por imitación. La solvencia del aprendiz estaba determinada por el maestro de acuerdo a su desempeño. Esta evaluación estaba basada

muchas veces en rasgos de la personalidad más que en criterios objetivos<sup>(7, 8)</sup>.

En la actualidad, el entrenamiento en las técnicas operatorias comienza fuera del quirófano, en un laboratorio, y la adquisición y evaluación de competencias se realiza con instrumentos que pretenden ser lo más objetivos posibles<sup>(7, 8, 9)</sup>.

Kane<sup>(10)</sup> definió la competencia como: "... el grado de utilización de los conocimientos, las habilidades y el buen juicio asociados a la profesión, en todas las situaciones que se pueden confrontar en el ejercicio de la práctica profesional" es decir que la competencia representa la capacidad de un profesional de utilizar su buen juicio así como también sus conocimientos, habilidades, y actitudes para solucionar problemas complejos que se presentan en el ámbito de su actividad.

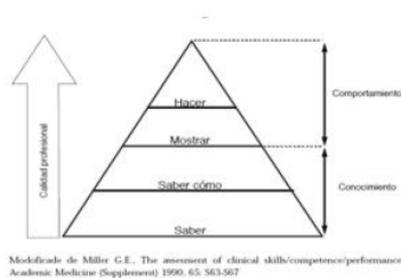
Hay muchos motivos por los cuales hay que evaluar las competencias tanto en estudiantes como en profesionales. Se pueden mencionar:

- 1) la evaluación como parte integrante del proceso educativo
  - 2) para certificar la competencia (evaluación sumativa)
  - 3) para ejercer retroalimentación sobre el proceso (evaluación formativa)
  - 4) para evaluar los programas de formación
  - 5) para confirmar los valores y los estándares de la profesión y
  - 6) por la responsabilidad e imputabilidad frente a la sociedad.
- En forma análoga Beard<sup>(11)</sup> enumera diversas razones para evaluar las competencias, y afirma que el método de evaluación depende del motivo:

- 1) como refuerzo del aprendizaje a través de una retroalimentación apropiada
- 2) para determinar el nivel alcanzado por el cirujano en formación
- 3) para controlar el progreso
- 4) para garantizar la seguridad de los pacientes antes de comenzar una práctica no supervisada
- 5) para certificación.

Miller<sup>(12)</sup> ha desarrollado un modelo de competencia profesional representado por una pirámide compuesta de varios niveles. En la base se encuentran los conocimientos (saber) sobre los que se apoya la competencia (saber cómo). A un nivel superior se encuentra el desempeño (mostrar cómo) y finalmente la acción en una práctica real (el hacer).

## Evaluación objetiva de competencias técnicas en cirugía



En 1996 Van Der Vleuten<sup>(13)</sup> propone un modelo teórico conceptual para definir la utilidad de una herramienta de evaluación de competencias profesionales que correlaciona distintas propiedades y que puede expresarse con la siguiente ecuación:

$$U = \left\{ \frac{F}{(p)} + \frac{V}{(p)} + \frac{E}{(p)} + \frac{A}{(p)} \right\} \times \frac{1}{C}$$

U: utilidad del instrumento; F: fidelidad o confiabilidad; V: validez; E: impacto educativo; A: aceptabilidad; C: costo; (p) contribución parcial. Modificado de Van Der Vleuten C. 1996. The assessment of professional competence: developments, research and practical implications. AdvHealthSciEducTheory-Pract. 1: 41-67

En resumen, el conocimiento actual sobre el tema pone en evidencia falencias de consideración en los métodos de evaluación de las habilidades técnicas vigentes en la actualidad. Los métodos tradicionales carecen de validez y confiabilidad en relación a las competencias técnicas. La formación quirúrgica dentro del marco de un programa formal estructurado de residencias médicas requiere una retroalimentación objetiva constante de la adquisición de habilidades técnicas para corrección y reorientación, y para ello precisa de instrumentos de evaluación objetivos. Existen métodos validados de evaluación objetiva que podrían comenzar a utilizarse para evaluar las habilidades quirúrgicas con fines formativos y de certificación. Aunque se requieren futuras investigaciones antes de ser utilizados para evaluaciones sumativas y de revalidación<sup>(14, 15, 16, 17, 18, 19, 20)</sup>.

### Material y Método

Se seleccionaron dos procedimientos quirúr-

gicos: la colecistectomía videolaparoscópica y la hernioplastia inguinal con técnica de Lichtenstein, debido a la alta prevalencia de la patología que motiva la intervención quirúrgica, y al hecho de que los cirujanos en formación están expuestos a dichos procedimientos desde el inicio mismo de su experiencia, permitiendo la aplicación del instrumento en todo el espectro de la curva de aprendizaje.

Para el análisis se constituyeron 3 grupos de comparación teniendo en cuenta las tres etapas de la teoría de adquisición de habilidades motoras propuesta por Fitts y Posner. El criterio de inclusión fue el nivel de entrenamiento al momento de la ejecución de cada uno de los procedimientos. El Grupo 1, incluyó a los cirujanos con experiencia inicial, que se corresponde con la ejecución de los primeros diez procedimientos. El Grupo 2, cirujanos con experiencia intermedia que cuentan en su haber una cifra superior a diez procedimientos pero aún en etapa de formación, y el Grupo 3 integrado por los cirujanos expertos con grado de especialistas. En ningún caso fue posible la identificación del paciente operado ni del cirujano evaluado, garantizando de este modo en todo momento el anonimato.

Las imágenes de las cirugías se registraron en tiempo real y sin captura de audio, desde el inicio hasta el final del procedimiento. Para las colecistectomías laparoscópicas se utilizó una placa de captura de video marca Pinnacle® Studio MovieBox HD versión 14, que fue conectada al procesador de imágenes de la endocámara del equipo de videolaparoscopia. Para las hernioplastias inguinales, una videocámara marca SONY MiniDV/DVcam, modelo DSR-PD 170P, con norma PAL y chip de 3CCD montada sobre un trípode ubicado detrás del cirujano, con angulación picada y con un plano detalle sobre el campo de trabajo. Las imágenes capturadas fueron bajadas a un soporte físico tipo DVD (.m2p) sin ningún tipo de edición.

Con el propósito de conferir la máxima objetividad se confeccionó un instrumento de evaluación específico para cada procedimiento, que permitió una calificación numérica en función del desempeño técnico del cirujano en cada uno de los pasos de la técnica quirúrgica considerados de mayor importancia.

Cada ítem seleccionado para ser evaluado

fue claramente definido, y su ejecución cuantificada mediante una escala de Likert con un puntaje de 1 a 5, y con descriptores explícitos para los extremos y el medio, donde la calificación <sup>(1)</sup> se corresponde a un desempeño inadecuado y <sup>(5)</sup> a la ejecución sobresaliente. Se convocó para cumplir el rol de evaluador a dos cirujanos de experiencia dedicados a la docencia y formación de médicos residentes del ámbito de la universidad, pertenecientes al Servicio de Cirugía N° I del Hospital Nacional de Clínicas de la Universidad Nacional de Córdoba. Se entregó a cada uno de ellos la información detallada necesaria para la calificación. Los evaluadores no tuvieron límite de tiempo para su tarea. Las filmaciones se pudieron observar todas las veces consideradas necesarias, teniendo la posibilidad de ser adelantadas o atrasadas a voluntad.

La muestra fue cegada colocando los DVDs de las filmaciones sin identificación en sobres de papel ad hoc sin ventana junto con una copia de la grilla para completar con los datos de la evaluación. Estos elementos fueron incluidos en un sobre de mayor tamaño con un número correspondiente al orden de la secuencia. La muestra se randomizó por sorteo consiguiendo de este modo una secuencia aleatoria de los procedimientos para calificar. Se realizaron análisis de la varianza no paramétricos (pruebas de Kruskal-Wallis) para comparar los tres grupos en función de los puntajes adjudicados por los evaluadores. Cuando el valor p fue menor al nivel de significación ( $\alpha$ ) de 0.05 se rechazó la hipótesis de igualdad de los grupos, y se concluyó que hay diferencias estadísticamente significativas entre los mismos.

Para la determinación de la confiabilidad del test se aplicó el Coeficiente alfa de Cronbach. Valores superiores a 0,80 aseguran una correcta consistencia interna de la escala y confieren la característica de confiable al instrumento de evaluación.

## Resultados

Desde Abril del año 2010 hasta Diciembre del 2012 se filmaron 67 procedimientos quirúrgicos de carácter electivo en el ámbito del quirófano central del Hospital Nacional de Clínicas (Facultad de Ciencias Médicas, Universidad Nacional de Córdoba) Se contabiliza-

ron 36 colecistectomías videolaparoscópicas y 31 hernioplastias inguinales, realizados en su totalidad por el personal perteneciente al Servicio de Cirugía N° I de dicha institución. Como cada procedimiento fue calificado por dos evaluadores de manera independiente, el número total de observaciones fue: 72 colecistectomías videolaparoscópicas y 62 hernioplastias inguinales. (Ver tabla 1)

Grupo	Colecistectomía videolaparoscópica	Hernioplastia inguinal
I	10	10
II	11	10
III	15	11
Total de casos	n = 36	n = 31
Total de evaluaciones	n=72	n=62

Tabla 1: Distribución de procedimientos según grupos de ejecución y número total de evaluaciones.

### 1. Comparación entre grupos

Con la finalidad de comparar los grupos entre sí se realizó un análisis de varianza no paramétrico (prueba de Kruskal-Wallis) a las calificaciones numéricas adjudicadas por los dos evaluadores, y luego la prueba "a posteriori" que hace referencia a la diferencia entre los grupos. Valores de p menores de 0,05 fueron considerados significativos.

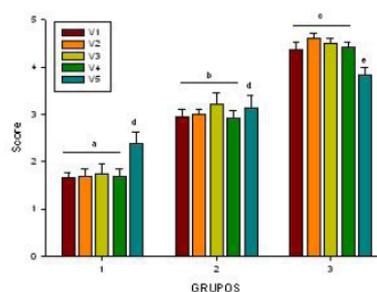


Figura XX: Resultados de la evaluación objetiva de competencias técnicas en Colecistectomía Laparoscópica mediante la aplicación de un nuevo instrumento específico propuesto para el procedimiento, presentada por grupos de experiencia. a vs. b vs. c:  $p < 0,0001$ ; d vs. e:  $p < 0,0002$ . Grupo 1 (principiantes): n=20; Grupo 2 (experiencia intermedia): n=22; Grupo 3 (expertos): n=30. V1: exposición de vesícula biliar. Uso de mano no hábil; V2: Disección del Triángulo de Callot; V3: Disección, ligadura y sección del conducto cístico; V4: Identificación de la arteria cística, ligadura y sección; V5: Disección del lecho vesicular

## Evaluación objetiva de competencias técnicas en cirugía

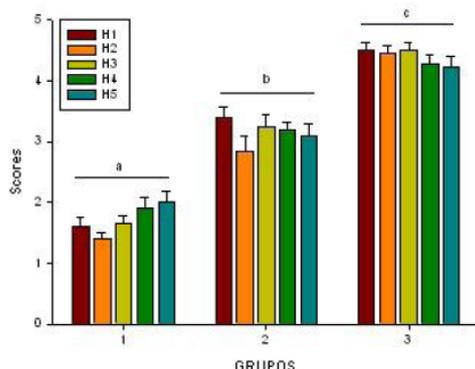


Figura XX: Resultados de la evaluación objetiva de competencias técnicas en Hernioplastia Inguinal mediante la aplicación de un nuevo instrumento específico propuesto para el procedimiento, presentada por grupos de experiencia. a vs. b vs. c=  $p < 0,0001$ . Grupo 1 (principiantes):  $n=20$ ; Grupo 2 (experiencia intermedia):  $n=20$ ; Grupo 3 (expertos):  $n=22$ . H1: Acceso al conducto inguinal y jalónamiento del cordón; H2: Exploración y disección de los elementos del conducto inguinal H3: Tratamiento del saco herniario H4: Ajuste del orificio inguinal profundo H5: Inserción y fijación de prótesis

## 2. Análisis del comportamiento de los evaluadores

Se realizó un análisis de la varianza no paramétrico (pruebas de Kruskal-Wallis) para comparar entre sí las calificaciones adjudicadas por los dos evaluadores en cada una de las variables consideradas. Cuando el valor  $p$  fue menor al nivel de significación ( $\alpha$ ) de 0.05 se rechazó la hipótesis de igualdad de los evaluadores y se concluyó que hay diferencias estadísticamente significativas entre ellos. Cuando el valor  $p$  fue mayor a 0.05, se concluyó que los evaluadores tuvieron igual comportamiento para la adjudicación de los puntajes. No se obtuvieron diferencias significativas en ninguna de las variables analizadas.

## 3. Análisis de la confiabilidad ( $\alpha$ de Cronbach)

Para la evaluación de la confiabilidad se utilizó el coeficiente  $\alpha$  de Cronbach. Valores por encima de 0,80 revelan datos a favor de la consistencia interna de la tabla. El coeficiente  $\alpha$  de Cronbach calculado para el instrumento que evalúa de manera específica la colecistectomía laparoscópica arrojó un valor de 0.93, y en el caso del instrumento diseñado para la hernioplastia inguinal fue de 0.95. Así

mismo, el instrumento de evaluación global obtuvo un coeficiente de 0.94 y 0.97 cuando fue aplicado a los mismos procedimientos respectivamente.

## Discusión

Evaluar supone efectuar una lectura orientada sobre el objeto que se evalúa, en función de la cual el evaluador se "pronuncia sobre la realidad". No existe una lectura directa de la experiencia si no que siempre hay un proceso de interacción entre el evaluador y la realidad a evaluar. En ese proceso, en función de las propiedades de esa realidad, el evaluador construye el referente, es decir, aquello con relación a lo cual se va a efectuar la evaluación.

En la construcción del referente intervienen diferentes factores: las expectativas en torno del objeto evaluado, la concepción de evaluación que se sustenta, los propósitos que se derivan de ésta y el contexto decisional en que se inserta el proceso evaluativo. Tanto el objeto que se evalúa como el proceso de valoración son construidos por el sujeto que evalúa.<sup>(21)</sup>

El proceso de construcción del objeto implica recortes sucesivos de la realidad, realizados conforme a criterios que pueden ser explicitados o tácitos. Por esta razón los indicadores que se construyen no cubren nunca la totalidad de la realidad que se evalúa. Al evaluar siempre se introduce un esquema de lectura de la realidad que circunscribe, recorta, y reduce lo observable.

En otros términos, el referente puede definirse como el conjunto de normas o de criterios que sirven de grilla de lectura del objeto a evaluar. Constituye una especie de "ideal" o de "deber ser" y hace que la evaluación implique, siempre, una lectura orientada por el marco, por la óptica, por la perspectiva de análisis privilegiada por parte del evaluador. Por el contrario, lo referido está constituido por el conjunto de atributos observables que se recortan de un objeto de evaluación y que se consideran representativos de éste.<sup>(22)</sup>

La noción de objetividad en la evaluación suele ser entendida como una situación cercana a lo ideal en donde el examen queda liberado de la influencia distorsiva de la subjetividad

del evaluador, que genera una valoración sesgada, es imprecisa, injusta, y muchas veces hasta malintencionada. Esta sobrevaloración de lo objetivo en contraposición a la carga de sentido negativo de lo subjetivo, queda en gran medida cuestionada por las consideraciones presentadas previamente.

El concepto de construcción del referente y del proceso de evaluación, implica necesariamente una participación protagónica de la subjetividad, ya que supone un posicionamiento frente a la realidad en un marco institucional, que determina la definición de un perfil de egresado o de especialista en función de una necesidad o demanda. Este perfil servirá de guía para la enunciación de los patrones de desempeño y rendimientos esperados, los requerimientos mínimos para la aprobación, las condiciones para la certificación, el número, el tipo, y el momento de evaluaciones, etc. que quedará plasmado en un programa de formación.

De un modo análogo, la investigación presentada es el resultado de una sucesión de recortes y definiciones dentro del área de conocimiento de la cirugía. Se toma la decisión, revelada en el mismo título, de evaluar específicamente las competencias técnicas. Es decir se delimita el estudio a los aspectos procedimentales de la práctica quirúrgica, y se decide, intencionalmente, dejar por fuera a otras competencias, como los conocimientos, el razonamiento clínico, la comunicación interpersonal, el profesionalismo, etc. Por otro lado, se seleccionan dos procedimientos para ser evaluados, se definen tres grupos de comparación, se elabora un instrumento específico para cada procedimiento en el que se seleccionan cinco puntos de la técnica para ser observados, se cuantifica el desempeño de un modo particular, etc. Las elecciones no son fortuitas ni responden al azar, si no son dirigidas y están impregnadas de una clara intencionalidad.

El análisis de los resultados luego del tratamiento estadístico permite afirmar que el instrumento de evaluación propuesto presenta validez y confiabilidad.

La diferencia estadísticamente significativa que se obtuvo en las calificaciones del desempeño técnico de los tres grupos compara-

dos hace referencia a la validez de constructo, esto es la propiedad para discernir entre los diversos niveles de entrenamiento: inicial, intermedio, y experto. Este hecho constituye uno de los hallazgos de mayor importancia que significa la puesta a prueba del valor calificador y clasificador de la herramienta, y por lo tanto le confiere trascendencia formativa, educativa, y probablemente de certificación al uso de este tipo de instrumentos para la evaluación.

Los valores obtenidos del Coeficiente alfa de Cronbach por encima de 0.80 demuestran que nuestro instrumento es confiable. La confiabilidad es un atributo por el cual una herramienta de evaluación presenta la suficiente estabilidad y precisión como para devolver el mismo resultado frente a las mismas circunstancias.

Ahora bien, hechas estas consideraciones ¿dónde radicaría entonces la objetividad en la modalidad de evaluación que se ha propuesto? Los aspectos más objetivos que podrían identificarse son: la enunciación explícita de los criterios de evaluación y la cuantificación del desempeño técnico. También podría mencionarse el anonimato, el cegamiento y la aleatorización de las muestras, y el hecho de contar con un registro de procedimientos que es perdurable en el tiempo. Es decir se debe entender el concepto como una objetividad instrumental.

Sin embargo, todos los criterios definidos como objetivos fueron construidos por un sujeto. La interpretación y el uso de la medida no pueden situarse fuera de un contexto interpersonal, e institucional, e histórico. Por lo que la idea de equidad y justicia con que se vincula a la objetividad en la evaluación termina resultando una entelequia. Aparecen en juego los aspectos cualitativos de la evaluación.

### Conclusiones

- 1 La evaluación debe ser entendida como un proceso de construcción de conocimiento sobre el objeto de evaluación.
2. Implica, un acto de interrogación sobre aquello que se evalúa, tanto sobre el objeto de evaluación como sobre las propias estrategias y prácticas del evaluador.
3. Siempre supone algún tipo de recorte, tanto

desde los referentes que se construyen, como desde lo referido.

4. Es factible y resulta útil evaluar las competencias técnicas en cirugía mediante la utilización de instrumentos objetivos.

5. La objetividad en la evaluación mediante el método presentado debe entenderse como instrumental. El uso de los datos (la medida) responde a definiciones institucionales, históricas, y a un modelo deseado de práctica profesional acorde a las necesidades.

### Bibliografía

1. Cosman P, Hemli JM, Ellis AM, Hugh TJ. 2007. *Learning the Surgical Craft: A Review of Skills Training Options*. ANZ J. Surg. 77: 838-845.
2. Cuschieri A. 2001. *Master surgeons and surgical proficiency*. En *Metrics: Objective Assessment of Surgical Skills – Developing Quantitative Measurements through Surgical Simulation*. Scottsdale, Arizona: US Army Medical Research Command, Telemedicine and Advances Technology Research Center (TA-TRC)
3. Cuschieri A, Francis N, Crosby J, Hanna GB. 2001. *What do Master Surgeons Think of Surgical Competence and Revalidation*. Am J Surg 182:110-116.
4. Yule S, Flin R, Paterson-Brown S, Maran N. 2006. *Non-technical Skills for surgeons in the operating room: A review of the literature*. Surgery. 139: 140-149
5. Darzi A, Datta V, Mackay S. 2001. *The Challenge of Objective Assessment of Surgical Skill*. Am J Surg 181: 484-486.
6. Reznick R, Regehr G, MacRae H, Martin J, McCulloch W. 1997. *Testing Technical Skill via an Innovative "Bench Station" Examination*. Am J Surg. 173: 226-230
7. Hamford JM, Hall JC. 2000. *Acquiring Surgical Skills*. BJS 87: 28 – 37.
8. Moen MD, Moen RO, Moen O.L. 2007. *Teaching and Evaluating Technical Skills: Applying the Aviation Model to Gynecologic Surgery*. J Pelvic Med Surg. 13: 107-112
9. Reznick R, MacRae H. 2006. *Teaching Surgical Skills – Changes in the Wind*. N Engl J Med 355: 2664-2669.
10. Kane MT. 1992. *The assessment of clinical competence*. Eval Health Prof. 15(2): 163-182
11. Beard JD. 2007. *Assessment of Surgical Competence*. BJS 94: 1315-1316
12. Miller G. 1990. *The assessment of clinical skills/competence/performance*. Academic Medicine. 65: S63-S67
13. Van Der Vleuten C. 1996. *The assessment of professional competence: developments, research and practical implications*. Adv Health SciEduc Theory Pract. 1: 41-67
14. Anderson CI, Jentz AB, Harkema JM, RaoKareti L, Apelgren KN, Slomski CA. 2005. *Assessing the competencies in general surgery residency training*. Am J Surg 189: 288-292
15. Bann S, Khan M, Datta V, Darzi A. 2005. *Surgical skill is predicted by the ability to detect errors*. Am J Surg. 189: 412-415.
16. Boehler ML, Schwind CJ, Rogers DA, Ketchum J, O'Sullivan E, Mayforth R, Quin J, Wohltman C, Johnson C, Williams RG, Dunnington G. 2007. *A Theory-Based Curriculum for Enhancing Surgical Skillfulness*. J Am Coll Surg. 205: 492-497.
17. Dath D, Regehr G, Birch D, Schlachta C, Poulin E, Mamazza J, Reznick R, MacRae H. 2004. *Toward Reliable Operative Assessment*. SurgEndosc 18: 1800-1804.
18. Doyle JD, Webber EM, Sidhu RS. 2007. *A universal global rating scale for the evaluation of technical skills in the operating room*. Am J Surg. 193: 551-555.
19. Larson JL, Williams RG, Ketchum J, Boehler M, Dunnington GL. 2005. *Feasibility, reliability and validity of an operative performance rating system for evaluating surgery residents*. Surgery. 138: 640-649.
20. MacRae H, Regehr G, Leadbetter W, Reznick R. 2000. *A Comprehensive Examination for Senior Surgical Residents*. Am J Surg. 179: 190-193
21. Bertoni A, Poggi M, Teobaldo M. En "Evaluación. Nuevos significados para una práctica compleja". Los significados de la evaluación educativa: alternativas teóricas" Buenos Aires, Editorial Kapelusz. p 20-29. 1995
22. Poggi M. 2008. *Evaluación educativa. Sobre sentidos y práctica*. Revista Iberoamericana de Evaluación Educativa. 1(1): 36-44.