

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 α 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 α 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 α 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^(1,2). 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^(3, 4).

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^(5, 6).

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^(7, 8).

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^(7, 8, 9).

Kane⁽¹⁰⁾ 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⁽¹¹⁾ 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⁽¹²⁾ 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).



Modificado de Miller G.E. The assessment of clinical skills/competence/performance. Academic Medicine (Supplement) 1990. 65: 563-567

En 1996 Van Der Vleuten⁽¹³⁾ 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:

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

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^(14, 15, 16, 17, 18, 19, 20).

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 colectistectomías 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 ⁽¹⁾ se corresponde a un desempeño inadecuado y ⁽⁵⁾ 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 (α) 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.

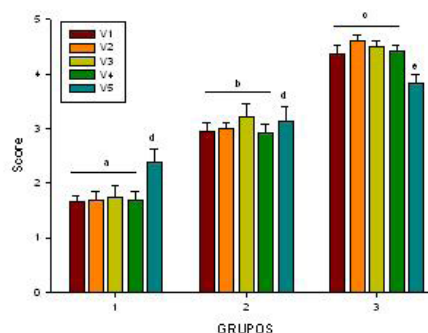


Figura XX: Resultados de la evaluación objetiva de competencias técnicas en Colecistectomía Laparoscópica mediante la aplicación de un nuevo instrumento específico propuesto para el procedimiento, presentada por grupos de experiencia. a vs. b vs. c= p <0,0001; d vs. e= p <0,0002. Grupo 1 (principiantes): 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

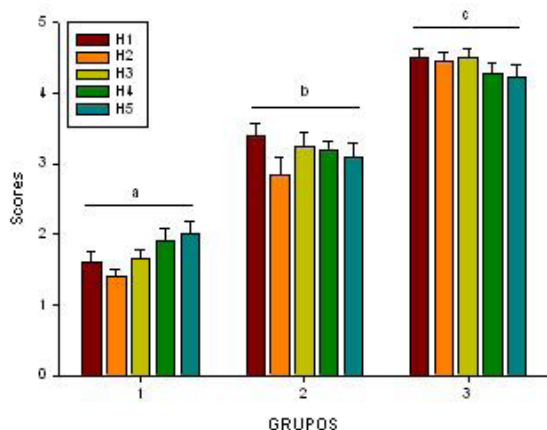


Figura XX: Resultados de la evaluación objetiva de competencias técnicas en Hernioplastia Inguinal mediante la aplicación de un nuevo instrumento específico propuesto para el procedimiento, presentada por grupos de experiencia. a vs. b vs. c= $p < 0,0001$. Grupo 1 (principiantes): $n=20$; Grupo 2 (experiencia intermedia): $n=20$; Grupo 3 (expertos): $n=22$. H1: Acceso al conducto inguinal y 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

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 (α) 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 (α de Cronbach)

Para la evaluación de la confiabilidad se utilizó el coeficiente α de Cronbach. Valores por encima de 0,80 revelan datos a favor de la consistencia interna de la tabla. El coeficiente α 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.⁽²¹⁾

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.⁽²²⁾

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.