



Evaluar a los médicos: Quién y cómo

Escuela de salud Pública de Menorca, 2019

Primera Mesa.

Evaluar a los estudiantes en el grado.

Cees van der Vleuten, Maastricht University.

OMC §

ORGANIZACIÓN
MÉDICA COLEGIAL
DE ESPAÑA

CONSEJO GENERAL
DE COLEGIOS OFICIALES
DE MÉDICOS



Fundación
Educación Médica

Programmatic assessment and its application to the medical students' assessment

FEM, SEDEM y CGCOM Meeting, Menorca, Spain

Cees van der Vleuten, Maastricht University
The Netherlands, www.ceesvandervleuten.com

Seminario 1: Evaluar en el Grado
Cees van der Vleuten, Maastricht University
18 September, 2019



Evaluar a los médicos:
Quién y cómo
Escuela de salud Pública
de Menorca, 2019

Introducción: Mila García Barbero

Buenas tardes. Intentar coordinar a gente tan complicada y con agendas tan complejas como la mayoría de los que estamos aquí, la verdad es que no resulta nada fácil y no me voy a extender. Esta sesión dentro del programa general está dedicada a la “*Evaluación de los estudiantes de grado, quién y cómo*”, que es el título general de toda la mesa de hoy.

Primero voy a presentar al ponente y luego presentar en dos palabras a los discusores y explicaré como va a funcionar el resto de la sesión. En primer lugar, vamos a presentar al Prof. Cees van der Vleuten, realmente una de las eminencias en evaluación y en educación en general. Lleva trabajando desde principio de los años 80 en la Facultad de Medicina de Maastricht que para todo el mundo que ha estado metido en educación médica ha sido una de las facultades punteras sobre todo en los programas basados en solución de problemas (*Problem-Based Learning*).

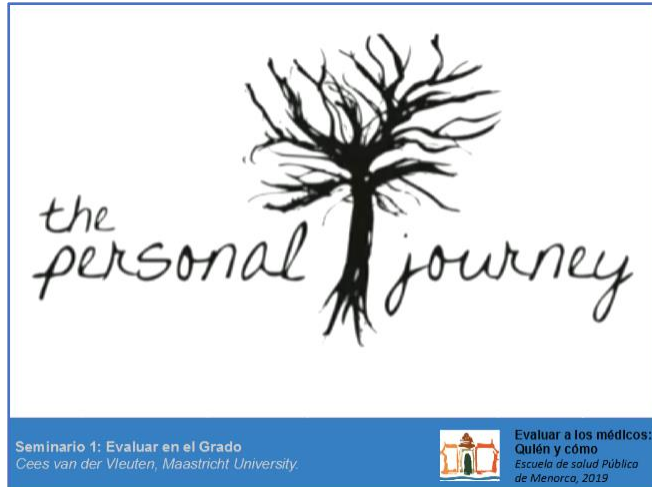
La Facultad de Medicina de Alicante cuando empezó en España al principio se inspiró bastante en Maastrich y en toda una serie de universidades que trabajaba en ese campo. Por supuesto en aquella época no pudimos llegar a lo que llegaron ellos ni lo intentamos y seguimos sin haber llegado. Vamos casi con 20 años de retraso en casi todas las facultades con respecto a las de otras, pero vuelvo a Cees. El es en este momento director del

Departamento de Desarrollo Educativo investigación de la Facultad de Medicina de Ciencias de la Vida, que si nos fijamos tiene una connotación solamente en el nombre muy diferente del nuestro. Es un departamento de desarrollo educativo e investigación en educación en medicina y en ciencias de la vida que no es como nuestros departamentos de educación médica (donde existen) que se dedican nada más a la gestión. Además, es director científico que también tiene otra connotación especial, de la Facultad de Educación de las profesiones de salud que se dedica fundamentalmente a formar profesionales en posgrado en educación médica. Es asesor de muchas facultades y del *National Board of Medical Examiners* en Estados Unidos, ha trabajado para las facultades más innovadoras de todo el mundo, Australia, Estados Unidos por supuesto en Europa y por lo tanto realmente, yo creo que es de las personas que más ha trabajado en educación médica en investigación en todo el mundo. En la sesión de hoy hablará de evaluación programática de los estudiantes luego ya seguiremos con las preguntas y la discusión.

CEES VAN DER VLEUTEN

Gracias por invitarme. Estoy muy honrado de estar aquí en este sitio histórico, maravilloso, para hablar de educación médica. He sido presentado muy amablemente, pero les voy a hablar de evaluación programática, y luego tendremos tiempo suficiente para preguntas. Si necesitan aclaramientos, o voy muy rápido, por favor díganmelo.

Básicamente, voy a llevarlos a través del viaje de mi carrera profesional. Empecé en la Universidad de Maastrich, donde existía un enfoque basado en PBL desde 1974. Poníamos un gran énfasis en el aprendizaje auto dirigido y teníamos un sistema de evaluación muy tradicional y fue muy complejo, y llevó mucho tiempo resolver este conflicto. Y les voy a hablar de esto precisamente.



Overview

- From practice to research
- From research to theory
- From theory to practice
- From practice to research
- Conclusions



Seminario 1: Evaluar en el Grado
Cees van der Vleuten, Maastricht University.

Evaluar a los médicos:
Quién y cómo
Escuela de salud Pública
de Menorca, 2019


Este es el resumen de lo que vamos a ver. Me voy a referir a la investigación que se ha hecho en evaluación de la práctica educativa y de los mensajes que se derivan de ello.

Con estos mensajes, podemos pensar más teóricamente sobre evaluación. Si tenemos un modelo de evaluación y lo implementamos en la práctica, una vez implementado podemos realizar investigación otra vez sobre como y cuando funciona. Esto es lo que


haremos esta tarde con ustedes.

The Toolbox

- MCQ, MEQ, OEQ, SIMP, Write-ins, Key Feature, Progress test, PMP, SCT, Viva, Long case, Short case, OSCE, OSPE, DOCEE, SP-based test, Video assessment, MSF, Mini-CEX, DOPS, assessment center, self-assessment, peer assessment, incognito SPs, portfolio.....

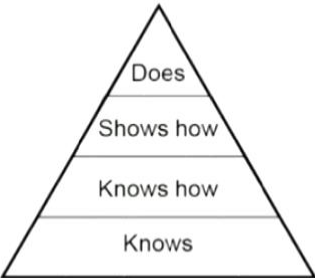


Seminario 1: Evaluar en el Grado
Cees van der Vleuten, Maastricht University



**Evaluar a los médicos:
Quién y cómo**
Escuela de salud Pública de Menorca, 2019

En evaluación en educación médica, como saben disponemos de gran número de instrumentos.

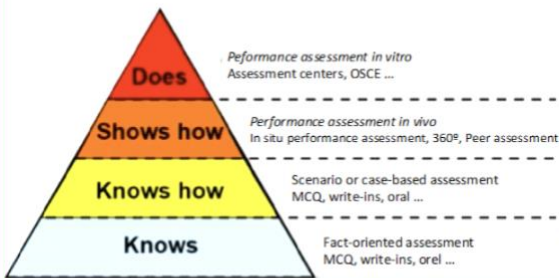
The way we climbed.....



Seminario 1: Evaluar en el Grado
Cees van der Vleuten, Maastricht University


**Evaluar a los médicos:
Quién y cómo**
Escuela de salud Pública de Menorca, 2019

The way we climbed.....




Performance assessment in vitro
Assessment centers, OSCE ...

Performance assessment in vivo
In situ performance assessment, 360°, Peer assessment

Scenario or case-based assessment
MCQ, write-ins, oral ...

Fact-oriented assessment
MCQ, write-ins, oral ...

Seminario 1: Evaluar en el Grado
Cees van der Vleuten, Maastricht University


**Evaluar a los médicos:
Quién y cómo**
Escuela de salud Pública de Menorca, 2019

Ustedes, probablemente estarán familiarizados con esa pirámide de la competencia, un modelo muy simple de competencia, y básicamente hemos ido ascendiendo por ella en los últimos 50-60 años, y estamos ya desde hace 10 años en el nivel más alto, nivel en el que queremos evaluar en contexto real (en la vida real).

Characteristics of instruments

Seminario 1: Evaluar en el Grado
 Cees van der Vleuten, Maastricht University

Evaluar a los médicos:
 Quién y cómo
 Escuela de Salud Pública
 de Menorca, 2019

Y se ha desarrollado para ello un gran número de instrumentos. De todos estos instrumentos, ustedes pueden estudiar sus características. Típicamente, esto se hace mediante diferentes estudios sobre fiabilidad y validez o estudiando otros aspectos del instrumento. Yo les resumiré algo de la investigación sobre los tres inferiores.

Si me refiero a la **validez**, estoy hablando de lo qué estoy midiendo. Si me refiero a la **fiabilidad**, estoy hablando de cuán precisa es la medida. Si me refiero al **impacto de la evaluación**, me refiero al efecto de la evaluación sobre el aprendizaje.

En términos de validez, de aquello que estoy evaluando, pienso que el mundo se está moviendo desde una orientación basada en el input, [desde un currículo en el que lo importante es cuantas horas de fisiología o anatomía se imparten, o de cuantas horas se dedican a la Medicina Interna], a otra orientación o modelo basado en output, es decir en los resultados.

Validity: what are we assessing?




- Curricula have changed from an input orientation to an output orientation
- We went from haphazard learning to integrated learning objectives, to end objectives, and now to (generic) competencies
- We went from teacher oriented programs to learning oriented, self-directed programs

Seminario 1: Evaluar en el Grado
 Cees van der Vleuten, Maastricht University

Evaluar a los médicos:
 Quién y cómo
 Escuela de Salud Pública
 de Menorca, 2019

Esto es, definamos en primer lugar los resultados, y después nos fijaremos en el currículo y en cómo el currículo permite llegar mejor a estos resultados. Estos son fundamentalmente los dos enfoques de la educación. También nos desplazamos desde programas centrados en el profesor a programas centrados en el que aprende, en el aprendizaje auto dirigido, en el que el modelo didáctico va alejándose de la clásica situación del aula o de la típica clase magistral o del currículo basado en ellas.

Competency-frameworks

 CanMeds <ul style="list-style-type: none"> • Medical expert • Communicator • Collaborator • Manager • Health advocate • Scholar • Professional 	 ACGME <ul style="list-style-type: none"> • Medical knowledge • Patient care • Practice-based learning & improvement • Interpersonal and communication skills • Professionalism • Systems-based practice 	 GMC <ul style="list-style-type: none"> • Good clinical care • Relationships with patients and families • Working with colleagues • Managing the workplace • Social responsibility and accountability • Professionalism
---	---	--

Seminario 1: Evaluar en el Grado
Gees van der Velden, Maastricht University

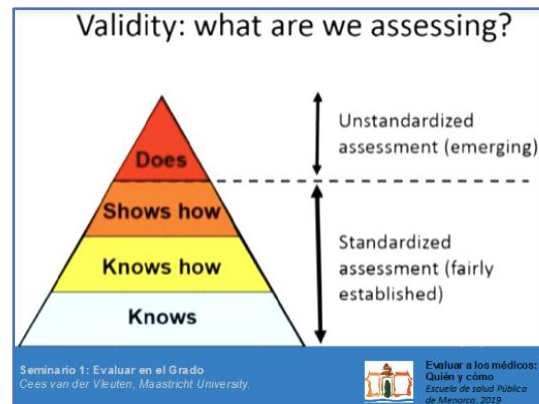
Pienso que esto es un movimiento mundial y si nos fijamos en los que está pasando en el, veremos que se están definiendo los resultados (outcomes). Aquí les muestro tres modelos de competencias, de los EU, de Canadá y del Reino Unido que se han definido hace ya varios años con la aportación de múltiples agentes, basándose en las necesidades futuras de salud que deberán afrontar los profesionales médicos de salud y en su definición.

Algunas cosas me llaman la atención en estos modelos. Si nos fijamos en ellos, vemos que el tipo de sus competencias son muy similares. El segundo aspecto que me llama la atención es que estas competencias no están relacionadas con los conocimientos. Si nos fijamos en estos documentos, vemos que se refieren a valores, aspectos relacionados con el profesionalismo. Estos son un tipo de habilidades que cada vez son más importantes. Son habilidades más blandas, comunicación, profesionalismo, trabajo en equipo, etc. Y son importantes porque sabemos que si las cosas no van bien en la práctica clínica, estas habilidades están implicadas, mucho más que las habilidades de conocimientos.

También sabemos que si algo va mal en la práctica clínica, en relación con estas habilidades, lo podríamos haber detectado ya en la formación médica de pregrado. Por ello tenemos la obligación de fijarnos en este tipo de habilidades y formar a los estudiantes en ellas. Pero ¿cómo hacerlo? ¿Como entrenar a los estudiantes en ellas? Tu puedes hacer un curso sobre comunicación por ejemplo durante dos semanas y luego haces una ECOE que certifica que eres un buen comunicador ¿para el resto de tu vida? ¡Esto no funciona así!

Este tipo de habilidades complejas se aprenden a lo largo del currículo, de forma longitudinal. ¿Tiene España un sistema de competencias? ¿Si, lo tienen? En los Países Bajos, hemos adoptado el CAN MEDS y hemos estructurado la educación de pregrado y postgraduada en los últimos 10 años, es decir, poniendo énfasis en la educación médica basada en competencias, tanto a nivel pregrado como postgrado.

Estas habilidades complejas no son fácilmente evaluadas de forma objetiva o estandarizada, por ejemplo, mediante un test de elección múltiple. Básicamente debemos evaluar esas habilidades a través de la observación y del juicio profesional. Básicamente debemos evaluar estas habilidades en la vida real, en situaciones no estandarizadas. Puede ser bueno tener una clase sobre comunicación, pero yo debo comunicar con un paciente o con otro miembro del equipo, y debo aprender a partir del *feedback* que me den sobre ello. Realmente, necesitaríamos evaluar estas habilidades con tecnología de la evaluación estandarizada, usando el juicio profesional holístico.



Messages from validity research

- There is no magic bullet; we need a mixture of methods to cover the competency pyramid
- We need BOTH standardized and non-standardized assessment methods
- For standardized assessment quality control around test development and administration is vital
- For unstandardized assessment the users (the people) are vital.

Seminario 1: Evaluar en el Grado
Cees van der Vleuten, Maastricht University

Evaluar a los médicos:
Quién y cómo
Escuela de salud Pública
de Menorca, 2019

A partir de la investigación sobre validez hemos aprendido que no hay una bala mágica. No existe un método que lo pueda hacer todo. Hemos introducido la ECOE y tendemos a pensar que esta lo puede hacer todo, pero no es así. Una sola medida no puede hacerlo todo.

Necesitamos cubrir todas competencias de la pirámide; necesitamos tanto tecnología de evaluación estandarizada como tecnología no estandarizada. Si tiene un

currículo, cuya evaluación consiste únicamente en evaluaciones estandarizadas del conocimiento, perdemos información.

Hemos aprendido que en la evaluación estandarizada, un poco de control de calidad tiene mucho impacto sobre la validez. Entonces si pones en marcha revisiones del desarrollo de estaciones o de preguntas test, si tienes un sistema de peer-review organizado, esto tendrá un mayor impacto sobre la calidad del examen. En los Países Bajos, actualmente tenemos una estrategia para compartir preguntas entre las facultades, reduciendo el coste de producir buenas preguntas. Por que tener un cajón con probablemente preguntas completamente elaboradas, ¿porque no compartirlas?

Sin embargo, en el vértice de la pirámide, para las evaluaciones no estandarizadas, no se trata tanto de control de calidad, sino de las personas que lo hacen. Si yo fuera un estudiante y él fuera mi supervisor clínico, observándome, viendo a un paciente, nos sentamos un par de minutos después de ese encuentro y él me dará algunos comentarios. La retroalimentación no se debe basarse en lo que está mal. La retroalimentación son

preguntas. ¿En qué estás pensando? ¿Cómo organizaste tu mente? ¿Notaste la reacción del paciente cuando dijiste esto? ¿Piensas qué lo harías diferente la próxima vez?

Entonces, la calidad de la interacción determina la validez de lo que obtienes y el control. El control de calidad no está tanto en el instrumento que se está usando para darme retroalimentación, sino en las personas, un dolor de cabeza habitual, porque si desea hacer un buen trabajo, debe preparar a las personas para que puedan hacerlo. El desarrollo del personal se vuelve realmente importante, ¿cómo dar retroalimentación adecuada en el entorno clínico? Estos son los mensajes de validez.

Method reliability as a function of testing time

Testing Time in Hours	MCQ ¹	Case-Based Short Essay ²	PMP ¹	Oral Exam ³	Long Case ⁴	OSCE ⁵	Mini CEX ⁶	Practice Video Assessment ⁷	In-cognito SPs ⁸
1	0.62	0.68	0.36	0.50	0.60	0.54	0.73	0.62	0.61
2	0.77	0.81	0.53	0.67	0.75	0.70	0.84	0.77	0.76
4	0.87	0.89	0.69	0.80	0.86	0.82	0.92	0.87	0.86
8	0.93	0.94	0.82	0.89	0.92	0.90	0.96	0.93	0.93

¹Norcini et al., 1985
²Stalenhoef-Halling et al., 1990
³Swanson, 1987
⁴Wass et al., 2001
⁵Van der Vleuten, 1988
⁶Norcini et al., 1999
⁷Ram et al., 1999
⁸Gortez, 2002

Pasemos a la fiabilidad, que es la precisión con la que podemos medir. Si quiero medir tus piernas y tengo una regla y las mido dos veces, esperarí una alta correlación entre la primera medida y la segunda medida; Eso es fiabilidad. Podemos expresar en una figura el coeficiente de fiabilidad, que puede variar de cero, no fiable en absoluto, a uno, fiabilidad perfecta. Se considera que necesitamos una realidad

mínima de 0.80 para un test o prueba sumativa; la prueba probará que tiene fiabilidad.

Ahora, voy a mostrarle una tabla de fiabilidad de los diferentes métodos de evaluación, que abarcan toda la Pirámide de Miller. Se expresa en función de las horas que dura una prueba, para que sean comparables a través de estos métodos (puedo evaluar muchos más elementos en la prueba de opción múltiple en una hora que estaciones puedo evaluar en 1 hora) y en orden para ser comparables, nos fijamos en el tiempo de prueba. Bueno, ¿qué podemos concluir de estas cifras? ¿Cuál es su conclusión?

Sí, de acuerdo. Con más tiempo obtienes más fiabilidad y la segunda conclusión es que la fiabilidad es muy similar. Esa es la segunda conclusión. Pero empecemos con la primera. Necesitas al menos tres o cuatro horas de tiempo de prueba para obtener un resultado confiable o, en otras palabras, nuestras pruebas educativas, la práctica educativa real no es tan confiable, lo que significa que tomaremos muchas decisiones falsas positivas y falsas negativas, simplemente debido al ruido en la medición y esto es una mala noticia. Esto se ha documentado una y otra vez en gran cantidad de estudios. Lo que subyace en esto es que nuestro rendimiento está vinculado al contexto. Si cambia el contexto de una actuación a menudo es muy diferente y debido a eso se necesita muchos contactos, se necesita una muestra correcta para llegar a la puntuación precisa. La mayoría de nuestras pruebas, utilizadas en la práctica médica, no son muy confiables.

Messages from reliability research

- Acceptable reliability is only achieved with large samples of test elements (contexts, cases) and assessors
- No method is inherently better than any other (that includes the new ones!)
- Objectivity is NOT equal to reliability
- Many subjective judgments are pretty reproducible/reliable.



Seminario 1: Evaluar en el Grado
Cees van der Vleuten, Maastricht University



Evaluar a los médicos:
Quién y cómo
Escuela de salud Pública
de Menorca, 2019

También hay buenas noticias en este aspecto, y es que no hay un método que por sí solo sea más fiable o, en otras palabras, que las evaluaciones más subjetivas pueden ser fiables, dependiendo de la muestras. Esto es una buena noticia porque si yo quiero evaluar estas habilidades complejas entonces debemos confiar en el juicio que es más subjetivo, si, pero si haces un buen trabajo con la muestra, obtendrás fiabilidad. Para

tener fiabilidad, no debes tener miedo del juicio profesional subjetivo, más moral.

Entonces, los mensajes que provienen de esa investigación es que la fiabilidad aceptable solo se logrará con un muestreo sustancial y, sobre todo, con la mayoría de los instrumentos en la práctica de educación médica. Sin embargo, método es mejor que cualquier otro método. Curiosamente, históricamente, originalmente, usamos estos métodos para la evaluación clínica. Pero no se estaba satisfecho con ello porque no eran fiables, por lo tanto, Harden en los años 70 propuso la ACOE. Fíjense en el título: examen clínico estructurado objetivo; para mejorar la fiabilidad de la ACOE se proponían listas de verificación objetivas (check-lists). Y al estandarizar, sin embargo, muchos años más tarde, se dispone de más investigación que nos indica que la fiabilidad no está en la objetivación ni en la estandarización, sino en el muestreo correcto. Y que la ACOE es tan buena o mala como nuestras evaluaciones subjetivas tradicionales. Por lo tanto, la objetividad no es lo mismo que la fiabilidad, por lo que se pueden obtener resultados de baja fiabilidad con una prueba objetiva. Y viceversa, se pueden tener resultados fiables con medidas subjetivas, todo dependiendo del muestreo. Y muchos juicios subjetivos son bastante robustos. Eso es validez.

Se dice que la evaluación impulsa el aprendizaje y eso es cierto; la relación es bastante compleja, pero a menudo también el impacto es bastante negativo. Los estudiantes empollan, memorizan para la ocasión; la mayor parte de nuestra evaluación es muy modular, por lo que tenemos una educación al final de cada módulo, tenemos una prueba sumativa, si la superas, puedes pasar al siguiente módulo y te vuelves inmune de por vida para aquella materia, porque no te vamos a evaluar más. Si falla, tiene que repetir la prueba, y si pasa, te vuelves inmune para toda la vida y si falla nuevamente, debe repetir el curso. No analizamos cuál es el problema, simplemente repetimos la prueba.

Educational impact: How does assessment drive learning?

- Relationship is complex (cf. Cilliers, 2011, 2012)
- But impact is often very negative
 - Poor learning styles
 - Grade culture (grade hunting, competitiveness)
 - Grade inflation (e.g. in the workplace)
- A lot of REDUCTIONISM!
 - Little feedback (grade is poorest form of feedback one can get)
 - Non-alignment with curricular goals
 - Non-meaningful aggregation of assessment information
 - Few longitudinal elements
 - Tick-box exercises (OSCEs, logbooks, work-based assessment).

En la práctica, en la práctica educativa, vemos estilos de aprendizaje en estudiantes consistentes en que se preparan, estudiando mucho de forma puntual para el examen, ya que piensan que lo más importante que han de hacer es prepararse para el examen. Y no estoy seguro si están familiarizados con la curva de olvido en psicología, pero algo que se ha aprendido solo para la ocasión se olvida el 50% en un par de semanas. Pero este es el modelo dominante en la evaluación.

Damos calificaciones numéricas, pero puedo decirle que las calificaciones son la forma más pobre de feedback que podemos dar. Si me calificas después de esta charla, en una escala de 10 puntos y me das un 7, gracias. ¿Qué debo hacer la próxima vez? Sí, obtener un 8. Pero la pregunta es si las calificaciones numéricas tienen sentido.

Particularmente, en la evaluación de las habilidades complejas, el peso de las calificaciones numéricas es bastante mínimo. Las calificaciones también proporcionan poca información. Yo veo un exceso de reduccionismo en los sistemas de evaluación. Se promueven estilos de aprendizaje que no reciben mucha retroalimentación y apenas miramos las cosas de forma longitudinal. La mayoría de los planes de estudio son muy modulares, no longitudinales y creo que estas competencias que definimos requieren un enfoque longitudinal, mucho más longitudinal. Necesitamos ver cómo las personas se desarrollan con el tiempo. Veo muchos ejercicios con check list como las ECOEs, y la probabilidad de superar una ECOE puede aumentar si se llega a memorizar la lista de verificación y veo mucho de eso. Veo muchas cosas incluso a nivel de posgrado y esto me causa perplejidad. Así como, voy a resumir cómo debería ser un enfoque dominante de evaluación dominante.

Messages learning impact research

- No assessment without (meaningful) feedback
- Narrative feedback has a lot more impact on complex skills than scores
- Provision of feedback is not enough (feedback is a dialogue)
- Longitudinal assessment is needed.



Seminario 1: Evaluar en el Grado
Cees van der Vleuten, Maastricht University



Evaluar a los médicos:
Quién y cómo
Escuela de salud Pública
de Mérida, 2019

Algunos mensajes. Yo diría que no debería haber evaluación sin feedback. Aprender se trata de obtener feedback sobre su el desempeño. Ahora sé que, en particular para las habilidades complejas, las palabras de la información narrativa contienen mucha más información que los puntajes de las calificaciones; así pues, usemos más narrativas en nuestra evaluación.

Hemos aprendido a través de la investigación, que la provisión de feedback no asegura el uso de este feedback. Particularmente en los sistemas sumativos tradicionales, los alumnos no están inclinados a recibir los comentarios cuando pasan la prueba. El uso del feedback debe estar apoyado, debe ser un andamio. Normalmente lo hacemos mediante la creación de un diálogo en torno al feedback, con los alumnos y reflexionar sobre el feedback y discutirlo con alguien.

Necesitamos una evaluación más longitudinal, lo cual es un desafío. Entonces, estos son algunos mensajes que nos llegan de la investigación.

Overview

- From practice to research
- **From research to theory**
- From theory to practice
- Conclusions



Seminario 1: Evaluar en el Grado
Cees van der Vleuten, Maastricht University.

Evaluar a los médicos:
Quién y cómo
Escuela de salud Pública
de Menorca, 2019

Pasemos a un enfoque más teórico de la evaluación utilizando estos mensajes.


La implicación de la investigación de la validez es que necesitamos una multitud de métodos diferentes. Las investigaciones de validez nos dicen que necesitamos mucho muestreo; necesitamos combinar información, debemos tener mucho cuidado al tomar

decisiones de alto riesgo basadas en pequeñas muestras y la investigación del impacto del aprendizaje dice que nuestra evaluación debería ser más funcional para el aprendizaje.

Y eso, para mí, es la base de la evaluación programática y discutiré algunas reglas básicas de la evaluación programática. Construyendo bloques. En la evaluación programática, cada evaluación se ve como un único punto de datos, nada más que eso. Es una medida y como una

Implications

- **Validity:** a multitude of methods needed
- **Reliability:** a lot of (combined) information is needed
- **Learning impact:** assessment should provide (longitudinal) meaningful information for learning



Programmatic assessment


Seminario 1: Evaluar en el Grado
Cees van der Vleuten, Maastricht University.

Evaluar a los médicos:
Quién y cómo
Escuela de salud Pública
de Menorca, 2019

medida realmente no es una medida, no tomamos las decisiones sobre una medida, no aprobamos ni suspendemos en base a una sola medida. Cada punto de datos no está

Building blocks for programmatic assessment 1

- Every assessment is but one data point (Δ)
- Every data point is optimized for learning
 - Information rich (quantitative, qualitative)
 - Meaningful
 - Variation in format
- Summative versus formative is replaced by continuum of stakes (stakes)
- N data points are proportionally related to the stakes of the decision to be taken.

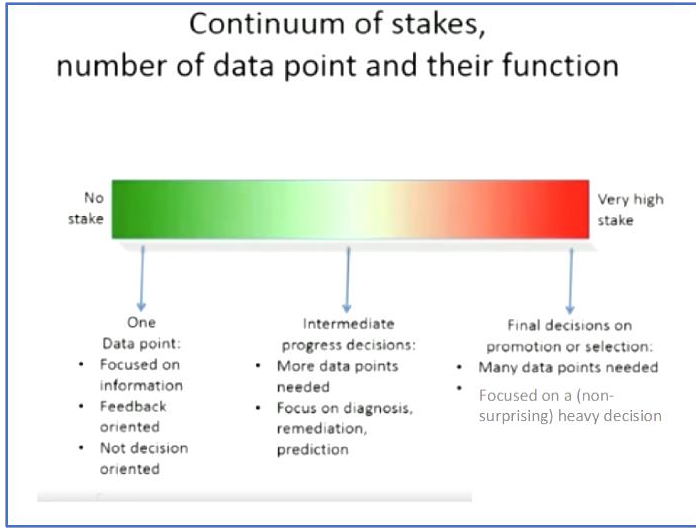


Seminario 1: Evaluar en el Grado
Cees van der Vleuten, Maastricht University.

Evaluar a los médicos:
Quién y cómo
Escuela de salud Pública
de Menorca, 2019

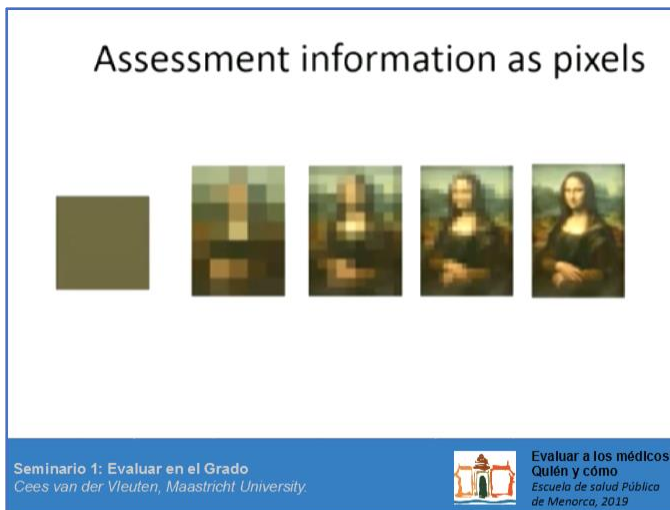
optimizado para la toma de decisiones, sino optimizado para dar feedback, optimizado para el aprendizaje, eso es para facilitar información enriquecedora tanto desde el punto de vista cualitativo como cuantitativo y debería ser comprensible para el aprendizaje de los estudiantes. Cada uno de los diferentes formatos tendrá una función diferente. No me importa, soy un poco agnóstico en términos de preferencia por un método de evaluación. No me importa, solo me importa su justificación para usar este

enfoque particular, y en este momento concreto del programa formativo y como se ajusta a sus objetivos educativos.

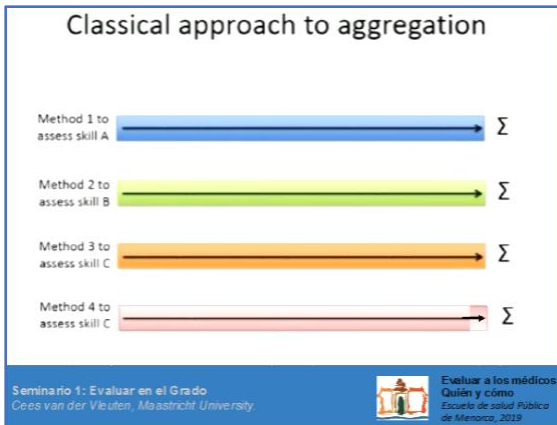


Reemplazamos la distinción tradicional entre evaluación sumativa y formativa, por un continuo de “marcas” (*stakes*). ¿Qué es una marca? Podemos tener decisiones a partir de evaluaciones de marcas de bajo nivel y a partir de evaluaciones de marcas de alto nivel. Claramente, cuanto mayor sea el “marca” o nivel, más puntos de datos necesitará, más información se necesitará. Este es un tema de punto de datos de

proporcionalidad relacionado con el nivel de la decisión que se este tomando. Si es una decisión de alto riesgo (*high-stake*), necesitaremos muchos puntos de datos, correcto. Entonces, en la evaluación individual continua, me centraré en el aprendizaje, me centraré en la retroalimentación, no en la toma de decisiones de bajo riesgo. Luego, en algún momento durante todo el año, tendremos que tomar una decisión de progreso intermedio, cómo estás, cómo te estás desarrollando, qué áreas necesitan más atención. Hasta que no sea una decisión de alta importancia, no tenemos consecuencias del tipo aprueba/suspende, tal vez haya consecuencias en términos de remediación, pero eso es todo y luego, por ejemplo, al final del año para pasar al próximo curso, necesitaré una decisión de alta importancia (*high-stake*). Esta decisión se basa en muchos puntos de datos, en mucha información.

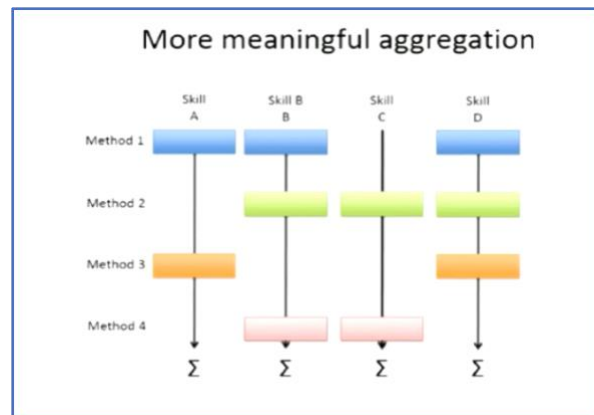


En cierta medida, se puede comparar esto con los píxeles en una fotografía. Si tengo un solo píxel, no veo la imagen. Al agregar píxeles a la fotografía, la imagen se aclara lentamente y luego entendemos lo que es. Cuando obtenemos una imagen clara, podemos tomar una decisión.



En el enfoque de la evaluación tradicional, agregamos información procedente de un método, por ejemplo, tenemos una ECOE y una estación es la estación de reanimación, en la siguiente estación se trata de dar malas noticias a una madre y la siguiente es el examen de las molestias abdominales. Luego agregamos, un promedio de la información en todas las estaciones y ese es su puntuación. Pero no tengo idea de lo que hay en común entre la reanimación y dar malas noticias a una madre.

En la evaluación programática, no evaluamos, no agregamos información dentro de un método, sino que agregamos a través de la información procedente de varios métodos, de cada una de las habilidades. Por lo tanto, podemos tener la información sobre comunicación procedente de la ECOE y triangulamos con la información que tenemos a partir de otros métodos de evaluación. Y al final del año, usted mira toda esa información y ve si ha habido un crecimiento suficiente en todas estas áreas diferentes. Esta es la teoría de la evaluación programática.



Overview

- From practice to research
- From research to theory
- **From theory to practice**
- From practice to research
- Conclusions

Seminario 1: Evaluar en el Grado
Cees van der Vleuten, Maastricht University

Evaluar a los médicos:
Quién y cómo
Escuela de salud Pública
de Menorca, 2019

Entiendo que todavía pueda verse muy radical, así que voy a pasar a la siguiente parte y daré una imagen de lo que es la evaluación programática en la Universidad de Maastricht.



La evaluación programática se está desarrollando en todo el mundo en este momento y se está implementando en muchos lugares diferentes, tanto a nivel de pregrado como de posgrado. En Canadá tienen programas de formación de posgrado que incluyeron una evaluación programática como parte fundamental de la educación médica basada en resultados.

Les voy a hablar del nuestro programa para graduados. Tenemos múltiples programas, pero les hablaré después. Nuestro programa para graduados es un programa de 4 años. Los alumnos provienen de un grado relevante (*Bachelor*) y estudian medicina en tres años. Tienen otro año dedicado a la investigación clínica. Es un programa basado en la estrategia PBL con casos o problemas escritos de pacientes en el primer año, con pacientes clínicos reales en el segundo año, de tal forma que los pacientes son el punto de partida para el aprendizaje en este segundo año. El tercer año se dedica a rotaciones clínicas en un área de preferencia, y el cuarto año dedicado a la investigación clínica. El área de investigación puede estar conectada con la rotación, es decir si quiere hacer investigación en cardiología y hacer una rotación en cardiología, es posible hacerlo.

Physician-clinical investigator program

- 4 year graduate entry program
- Competency-based (Canmeds) with emphasis on research
- PBL program
 - Year 1: classic PBL
 - Year 2: real patient PBL
 - Year 3: clerkship rotations
 - Year 4: participation in research and health care
- High expectations of students: in terms of motivation, promotion of excellence, self-directedness

Seminario 1: Evaluar en el Grado
Cees van der Vleuten, Maastricht University

Evaluar a los médicos:
Quién y cómo
Escuela de salud Pública
de Menorca, 2019

The assessment program

- Assessment in Modules: assignments, presentations, end-examinator etc.
- Longitudinal assessment: assignments, reviews, projects, progress tests, evaluation of professional behavior, etc.
- All assessment is informative and low stake formative
- The portfolio is central instrument

Module 1	Module 2	Module 3	Module 4
PT 1	PT 2	PT 3	PT 4
Longitudinal Module exceeding assessment of knowledge, skills and professional behavior			
Module exceeding assessment of knowledge in Progress Test			
Continuing meeting	Continuing meeting	portfolio	Continuing meeting

Seminario 1: Evaluar en el Grado
Cees van der Vleuten, Maastricht University

Evaluar a los médicos:
Quién y cómo
Escuela de salud Pública
de Menorca, 2019

El programa de evaluación se hace por evaluación de módulos, nada demasiado novedoso. Usamos una gran variedad de instrumentos, test de elección múltiple, exámenes tipo ensayo dependiendo de la justificación para usarlo. Algunas unidades se interrumpen para la evaluación, de tal forma que cada dos semanas se da feedback sobre las semanas previas y así tenemos una evaluación longitudinal



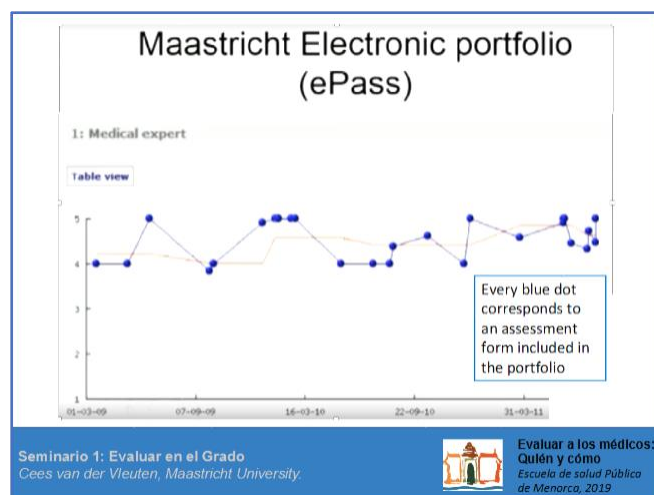
Por ejemplo, en el dominio de conocimientos, hacemos lo que llamamos un *progress-test*. El *progress-test* es un examen final, un examen final comprensivo de todas las materias del currículo, que se administra a los estudiantes, desde el primer año al cuarto y del primero al sexto en el programa de grado, cada 3 meses. Cuatro veces al año, todos los estudiantes pasan este examen escrito y se comprueba su progreso. Los estudiantes de primer año no son capaces de

contestar muchas preguntas, los de segundo año, más y así sucesivamente, pero podemos ver un interesante patrón de crecimiento. Es muy difícil estudiar para este examen ya que cada tres meses diseñamos un nuevo examen, por cierto, por la forma en que lo hacemos, en colaboración con otras facultades de medicina del país, para que podamos compartir los resultados e incluso comparar nuestro rendimiento como facultad de medicina.

Luego, también tenemos muchas evaluaciones de revisión por pares y evaluaciones de tutores sobre las competencias. Así yo soy un tutor y este fuera mi grupo de tutoría, les pediría que completasen un formulario que es muy abierto para la próxima tutoría, sobre su profesionalismo, y así en dicha próxima sesión de evaluación donde veremos lo que han escrito y si eso se alinea con su propia imagen de sí mismo.

Tenemos un nivel importante de comunicación; los estudiantes están en el mismo grupo todo el año y pueden dar *feedback*. La evaluación es formativa y “*low stake*”. No aprueban o suspenden.

El portfolio es un elemento central, que recoge cualquier evidencia del estudiante. Es parecido al historial de un paciente. Es el historial del estudiante. Es una visión general de los diferentes módulos; incluye los *progress-test* cuatro veces al año, tenemos las evaluaciones por pares, las evaluaciones de las habilidades en el laboratorio de habilidades, etc. Cada estudiante está conectado a un mentor, que denominamos consejero, aunque odio este término porque en inglés puede significar cosas diferentes.



Coaching by counselors

- Coaching is essential for successful use of reflective learning skills
- Counselor gives advice/comments (whether asked or not)
- He/she counsels if choices have to be made
- He/she guards and discusses study progress and development of competencies



Seminario 1: Evaluar en el Grado
Gees van der Vleuten, Maastricht University



Evaluar a los médicos:
Quién y cómo
Escuela de salud Pública
de Menorca, 2019

Así cada estudiante tiene un mentor y el mentor tiene reuniones periódicas con los estudiantes. El rol del estudiante es preparar estas reuniones, observando todas las evidencias, todos los datos que tienen del sistema de evaluación. Tienen que autoanalizarse, tienen que reflexionar y tienen que planificar lo que se está discutiendo con su mentor y así es como funciona y luego, al final del año, tenemos que tomar una decisión.

Tomamos esa decisión en un grupo de

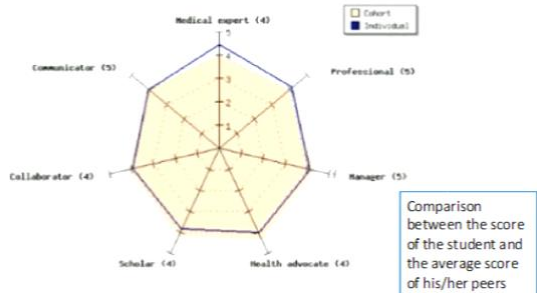
profesores y más adelante explicaré un poco más sobre eso.

Es imposible tener un promedio simple de toda la información facilitada por las evaluaciones, porque la información es cuantitativa y cualitativa. Es imposible agregarla de una manera simple, por lo que la decisión final nuevamente es un juicio profesional y lo

hacemos en grupo para que no haya sesgo. Ponemos mucho énfasis en el *feed-back*. Así en el *progress-test*, como cada 3 meses hay una nueva prueba, el estudiante puede ver el progreso. Así se puede ver el puntaje total de un estudiante, pero también puede ver sus puntajes de anatomía, fisiología o ciencias básicas y ver cómo se desarrolla. La computadora predice su futuro y si tiene una comprensión muy baja de desarrollo de sus conocimientos. Para otras evaluaciones, tenemos un sistema

de *feed-back* como el que se muestra para las diferentes competencias (entre paréntesis de ven los números) y se puede ver también como su desempeño en relación con la cohorte.

Maastricht Electronic portfolio (ePass)



Seminario 1: Evaluar en el Grado
Gees van der Vleuten, Maastricht University



Evaluar a los médicos:
Quién y cómo
Escuela de salud Pública
de Menorca, 2019

Podemos observar el rendimiento longitudinalmente, podemos hacer clic en cualquier punto y obtener la información cuantitativa y cualitativa de esta fecha. Una de las cosas más importantes, es que en el portafolio electrónico puedes ver la generación de feedback narrativa, y luego puedes conocer la historia del alumno y puedes filtrar por competencia si lo deseas, pero rápidamente obtienes una visión general del rendimiento de un individuo que es rico, que se puede utilizar con fines de aprendizaje y de evaluación, con fines de decisión. El asesoramiento es realizado por los consejeros. Este es un elemento muy esencial. La tutoría es realmente importante. Los estudiantes la aprecian. Es muy importante promover el aprendizaje auto dirigido. He estado implicado en PBL desde 1982 y he aprendido un hecho paradójico y es que el aprendizaje auto dirigido requiere mucha dirección.

La toma de decisiones se realiza en un comité. Se puede pensar que representa mucho trabajo, pero no es así ya que esos comités están preparados y en el 95-96% de los casos, no es necesaria la discusión. No necesitas discutir. No necesitas deliberar. En cada caso se llega a un consenso. Utilizamos estándares narrativos para cada año de capacitación para que el comité pueda usar nuestros estándares para llegar a esas decisiones. Básicamente, es un juicio profesional. Y has de poder justificar la decisión y debes tener argumentos. Si el resultado de la decisión del comité es una gran sorpresa

Decision-making by committee

- Committee of counselors and externals
- Decision is based on portfolio information & counselor recommendation, competency standards
- Deliberation is proportional to clarity of information
- Decisions are justified when needed; remediation recommendation may be provided



Seminario 1: Evaluar en el Grado
Cees van der Vleuten, Maastricht University



**Evaluar a los médicos:
Quién y cómo**
Escuela de salud Pública
de Menorca, 2019


para el alumno, entonces es que algo salió mal en el proceso. Por ello pienso que no deberían producirse sorpresas en este momento.

Para aplicar una determinada estrategia, deberíamos basarnos en los resultados que se deriven de la investigación cualitativa tanto como sea posible. No voy a explicarlo en detalle. Hemos introducido el enfoque programático en nuestro programa de Máster, que es de 60 estudiantes al año y los últimos tres años de Bachelor, con 342 estudiantes y vamos a tener una implementación completa en los primeros dos años en 2021.

State of implementation in Maastricht

- Graduate entry medicine program
- Master of Medicine (clinical phase year 4-6)
- Partial implementation in Bachelor of medicine (year 1-3; full implementation in preparation)
- Partial implementation in Bachelor of life sciences
- Full implementation in Bachelor 2021

Seminario 1: Evaluar en el Grado
Cees van der Vleuten, Maastricht University



**Evaluar a los médicos:
Quién y cómo**
Escuela de salud Pública
de Menorca, 2019

Overview

- From practice to research
- From research to theory
- From theory to practice
- From practice to research
- Conclusions



Seminario 1: Evaluar en el Grado
Cees van der Vleuten, Maastricht University



Evaluar a los médicos:
Quién y cómo
Escuela de salud Pública
de Menorca, 2019

Bien, debo terminar. Si quieren leer sobre estos temas, pueden ir a mi página personal donde encontrarán una serie de publicaciones sobre evaluación programática y los PDF. están allí. Pueden descargarlos si lo desean.

<https://www.ceesvandervleuten.com>

Los más divertido de la evaluación programática es su implementación. Y aquí lo importante son las personas. La gente necesita comprar esto. La gente necesita entenderlo. Aquí está el desafío porque tienes que convencer a mucha gente. Es lo mismo, que pasa con el PBL. Si se pasa de un sistema tradicional a PBL, necesitas un cambio mental, necesitas un cambio de cultura, y bueno, eso no es fácil de lograr.



<https://www.ceesvandervleuten.com>

Seminario 1: Evaluar en el Grado
Cees van der Vleuten, Maastricht University



Evaluar a los médicos:
Quién y cómo
Escuela de salud Pública
de Menorca, 2019

Sé que muchas facultades de medicina españolas son relativamente tradicionales y no es fácil lograr un cambio tan importante, pero espero que al menos les esté dando un poco de comida para la sal. No espero que implementen la evaluación programática mañana, pero espero que puedan tomar algunas de esas ideas y traerlas de vuelta a sus currículos. Todo se trata de la gestión del cambio, ¿verdad? De hecho, creo que la educación y la innovación en la educación se basan en la gestión del cambio. Tenemos muchos escenarios educativos, mucha metodología educativa, pero tenemos que cambiar a nuestros profesores de grado, de máster y de posgrado. Estamos frente a un desafío muy importante que hay que alcanzar.

Conclusions 1: The way forward

- We have to stop thinking in terms of individual assessment methods
- A systematic and programmatic approach is needed, longitudinally oriented
- Every method of assessment may be functional (old and new; standardized and unstandardized)
- Professional judgment is imperative (similar to clinical practice)
- Subjectivity is dealt with through sampling and procedural bias reduction methods (not with standardization or objectification).



Seminario 1: Evaluar en el Grado
Cees van der Vleuten, Maastricht University



Evaluar a los médicos:
Quién y cómo
Escuela de salud Pública
de Menorca, 2019

Como conclusión, creo que debemos alejarnos del enfoque de la instantánea, de todas las evaluaciones individuales. Necesitamos un enfoque más sistemático y longitudinal y, como ya dije, no me importa el método de evaluación. Nos hemos alejado de la ECOE, por ejemplo, con el aumento de la evaluación en lugar de trabajo, que hemos implementado, nos hemos alejado de la ECOE en la fase de Bachelor al aumentar las evaluaciones de los docentes dentro del laboratorio de

habilidades que estaba en el Centro de Simulación, que está conectado inmediatamente al feedback.

Creo que el juicio profesional es imperativo. Creo que les puedo explicar una discrepancia interesante. Ustedes, hacen juicios profesionales durante todo el día en su profesión, tienen información de pacientes, información cuantitativa, su información cualitativa y luego toman una decisión. Lo hacen todo el tiempo, ¿está sesgada esa decisión? Sí lo está. Es claro, tenemos sistemas de apoyo a la decisión, tenemos protocolos, pautas clínicas, y podemos tener segundas opiniones y creo que deberíamos hacer lo mismo en la evaluación. Deberíamos restablecer el juicio profesional nuevamente como base fundamental para evaluar a nuestros alumnos.

Conclusions 2: The way forward

- The programmatic approach to assessment optimizes:
 - The learning function (through information richness)
 - The pass/fail decision function (through the combination of rich information)



Seminario 1: Evaluar en el Grado
Cees van der Vleuten, Maastricht University



Evaluar a los médicos:
Quién y cómo
Escuela de salud Pública
de Menorca, 2019

Y creo que la subjetividad se trata mediante el muestreo y pensando en todo tipo de medidas de procedimiento para reducir el sesgo, y creo que este enfoque programático integra tanto las evaluaciones para el aprendizaje como la evaluación del aprendizaje. Optimiza tanto la función de aprendizaje como la función de toma de decisiones de la evaluación y ese es el final de mi viaje.

Ahora, déjenme conocer, sus asombros, tus preguntas, sus observaciones, sus comentarios críticos.

DISCUSION

Pilar Garrido

Mi pregunta es: ¿Cuántos estudiantes tienen en cada grupo, implicados en cada proceso? Ud. ha hablado de mentorías, ¿cuántos mentores tienen para cada caso?

CVV

En uno de los programas tenemos 60 alumnos y en el otro 340. La cuestión que está detrás es como podemos mejorar esto y cuáles son los costos.

Tomas Cobo

La pregunta que creo que quería hacer Pilar, es cuántos tutores se necesitan para 60 y 340 alumnos.

CVV

6 mentores y 34 mentores en cada programa. Están involucrados muchos médicos de los diferentes hospitales de nuestra zona. Les encanta, les encanta participar.

Tomas Cobo

¿Son remunerados?

CVV

Se paga por la educación en general, pero no se paga de forma individual. Lo que pagamos es por la educación clínica, pero ya sabes, eso no tiene ninguna consecuencia para el individuo.

La respuesta un poco más cínica sería que estamos financiados por el estado, no tenemos más fondos que cualquier otra Facultad en los Países Bajos, y hemos de mantenernos dentro del presupuesto y creo que si se quiere cambiar mudaría a algo como esto, teniendo en cuenta que ya gastamos bastante dinero en la evaluación, deberíamos redistribuirlo de una manera diferente. Y lo mismo vale para PBL, porque son necesarios muchos tutores. Pero yo he visto experiencias magníficas de PBL en Sudáfrica o en Mozambique que con muy pocos recursos y hacen un buen trabajo. Es mucho más una cuestión de mentalidad que de recursos.

Mila García Barbero

Gracias de nuevo. Creo que cuando te estamos escuchando, siento que estabas en un planeta diferente, no en una parte diferente de Europa sino en un planeta diferente.

Tenemos 4 personas que han preparado preguntas, pero permítanme comenzar con una mía. ¿Cuánta libertad tiene el decano (o quien sea) para organizar un plan de estudios, una evaluación o lo que sea que quiera implementar en Maastricht, considerando por un lado la universidad y por otro la regulación estatal o cualquier otra si la hay?

CVV

Cuando originalmente quisimos hacer una evaluación programática en el programa de grado, los abogados nos dijeron que no podíamos hacer esto, porque la ley requería otorgar sus créditos a cada una de las evaluaciones, por lo que originalmente se implementó créditos para cada evaluación individual, decisión fácil de tomar pero esto no funcionó. Así que fui al Rector de la universidad y le dije que debes cercar a tus abogados porque vamos a hacer esta innovación.

No fuimos contra la ley. Lo implementamos de tal manera para poder otorgar 60 créditos al final del año, y, sí, tenemos la autoridad si deseamos hacer esto. Sin embargo, se trata de personas, el problema es cómo hacer que la gente tan alejada de estas ideas haga ese cambio ahora. Nunca me preocupaban los estudiantes; los estudiantes son jóvenes, de todos modos, tenemos los mejores de lo mejores en medicina y, les irá bien a pesar de la educación. A menudo, el problema con muchos decanos es que gran parte de nuestra educación está en silos, y es muy difícil tocar esos silos; es muy difícil gestionar la educación en la mayoría de las facultades de medicina, cierto, pero sí, lo que realmente recomiendo es crear un staff con gran creatividad si haces un cambio importante, pero no estoy seguro de que puedas posponerlo.

Mila García Barbero

Las cuatro personas, discusores que ha llamado Arcadi que han hecho preguntas que creo que todo el mundo las tiene y por tanto no las voy a leer. Alguna la vamos a orientar de alguna otra forma, pero básicamente lo que quería decir en 5 puntos es lo que Cess ha considerado los puntos más importantes de la evaluación programática para centrar un poco la discusión,

1. Está basada en gran cantidad de información
2. Está basada en juicios de valor
3. Está basada en feedback
4. Es longitudinal
5. La utilización de métodos diferentes y el sexto punto es que están en el otro extremo del planeta con lo que hacemos nosotros.

Y dentro de lo que podemos hacer, aquí la legalidad y la administración, nos come por todos lados, es muy difícil hacerlo y realmente, como nos han dicho ellos, se enfrentaron a los abogados que decían que no se podía hacer. Un ejemplo es el papelito que nos ha dado Arcadi, que es un folio y medio para que digamos que sí que se nos puede gravar.

Entonces, de las preguntas que hay, yo diría, que la primera sería, ¿se puede empezar a hacer una evaluación programática a pequeña escala, de feedback, con diferentes métodos? Estamos usando diferentes métodos, pero en general, usamos la ECOEs por un lado, MCQ por otro. Pero no combinamos, los usamos de forma aislada, en bloques sin ninguna línea de continuidad.

¿Creemos, que se puede hacer en España, una evaluación programática o empezar a una pequeña escala, a un pequeño nivel? Tienen 360 alumnos, que tampoco son poquitos. Nosotros, somos pocos, si quitamos los repetidores, esta es otra. Aquí se pasa de curso con asignaturas, se llega a la graduación con asignaturas de segundo. Es que vivimos en otro mundo diferente, y hay que adaptar tantas cosas que se hace difícil. Pero bueno, aquí, vamos a dedicarnos al quien y al cómo, que son las dos preguntas que corresponde aquí. Por lo tanto, en principio;

¿Quién podría hacer estas evaluaciones, o quien podría empezar?, ¿Podríamos empezar con tutores, o con mentores, como ha dicho él?, independientemente de las funciones que les demos. ¿Debe ser la persona que ha impartido la docencia o la clínica, la que lo debe hacer o ha de ser una persona diferente? Por ejemplo, si queréis empezamos por el quién. Cualquier pregunta es bienvenida. Yo hago estas que nos han pasado los cuatros discusores que son:

Pablo Lara, presidente de la Conferencia Nacional de Decanos (CND) y decano de Málaga. Antonio Company, que es mi decano, de la Facultad de Medicina de la UMH. Además, los dos están en la conferencia de decanos. Luego está Jesús Millán, Catedrático de Medicina y vicedecano de la UCM y luego está Roger Ruiz que es profesor de la UFV en Madrid. Las preguntas que yo hago son en realidad las que han hecho ellos y que todos tenéis en vuestro librito.

Yo preguntaría, lo primero, ¿Debe ser el mismo profesor, suponemos el que haga la evaluación, debe ser el mismo que enseña y evalúa? ¿Debe ser el mismo profesor de Cirugía, de Medicina Interna o de Fisiología el que enseña y evalúa? Cess, crees que debería ser la misma persona la que enseña (profesor clínico, o de básicas) el que evalúa a los estudiantes?

CVV

Seguro, ¿por qué no?

Mila García Barbero

Entendía que utilizaban un comité

CVV

Si, es verdad, pero solo para los “hight stakes” para las decisiones de alto nivel. Puede ser una persona o dos, pero han der ser confiables.

Roger Ruiz

Me resulta difícil el papel del mentor en este proceso porque, en mi experiencia, si el mentor está involucrado de alguna manera en una decisión de alto riesgo, la relación con el mentorizado será diferente. ¿Como solucionas esto?

CVV

Debes separar estos dos roles, porque el mentor sigue al estudiante a lo largo de todos los años, construyendo una relación de confianza con el estudiante, pero no lo aprueba ni le suspende. Si estás en el comité, no eres el evaluador. En el gran programa de Máster, tenemos a los mentores también como el comité, pero si se discute sobre sus estudiantes, el no tiene otro papel que el de agregar información. Protegemos el papel del mentor. Toda la idea es hacer una relación de confianza. También creo que esto es muy beneficioso para evitar el estrés y el burnout que vemos entre los estudiantes de medicina y los residentes porque, si miras los datos, es realmente aterrador, alrededor del 25% de los estudiantes están con burnout, ¿sabes? Creo que les provocamos estrés al ir de un examen a otro. Nuestro sistema es muy punitivo.

Roger Ruiz

¿Cuáles son los profesores implicados en el comité? ¿Por qué la mayoría son clínicos, o no?

CVV

No necesariamente

Roger Ruiz

Todos ellos deberían tener una idea de la competencia muy similar sino esto podría ser un problema. ¿Qué pasa si tenemos una idea diferente de los demás? ¿Cómo os maneáis?

CVV

No lo hacemos porque es difícil estandarizar a todos y no creo que sea necesario. Creo que si eres un científico básico, incluso puedes ver a los estudiantes en los años clínicos y ver si ciertas áreas necesitan mejorar o no. El único problema es la triangulación. Creo que la vieja forma de tratar de estandarizar todo y objetivar es porque estamos tan inclinados a evaluar con objetividad. Pensamos que todo debe ser objetivo, pero en la práctica clínica no todo es una comedia objetiva como no debería serlo en la educación.

Pablo Lara

Muchas gracias por la charla. Mila dijo que estabas en otro planeta, yo diría que al menos en otra galaxia.

CVV

Bueno, hay muchas más estrellas en nuestra galaxia.

Pablo Lara

Hemos recibido de ti algunos consejos básicos porque nuestra situación es muy diferente. Tenemos una gran cantidad de estudiantes en cada grupo. Tenemos 120.

CVV

Para mí eso es pequeño. Desearía tener 120.

Pablo Lara

El número de docentes no es suficiente, nuestras asignaturas dependen de una solo área de conocimiento y cada una tiene una calificación numérica. La cultura del número está muy presente en la facultad de medicina. Obtener un lugar en la Facultad de Medicina, solo depende de los resultados de un examen teórico. El futuro del trabajo como médico depende del examen MIR, un examen teórico.

CVV

¿Qué estás haciendo, qué estás haciendo para cambiar tu sistema?

Pablo Lara

Lo intento, intento hacerlo lo mejor que puedo, pero lo que quiero decir es solo cómo comenzar.

CVV

En nuestro sistema anterior, teníamos departamentos y profesores individuales haciendo cada cosa. En los programas de educación más modernos, las cosas esta en línea con las nuevas ideas. Pero es difícil de conseguir en cualquier escuela de medicina.

Antonio Company

Estoy completamente de acuerdo con el viaje personal del Prof. C V V. Muchas gracias por su presentación y también por la presentación del Prof. A. Gual. Creo que esa es la mejor manera de empezar este seminario.

Los problemas es que Pablo y yo somos decanos y tenemos muchos problemas como usted, con los profesores, con los estudiantes, con el gobierno, con la ley, etc. No es la mejor manera de comenzar una evaluación programática. La evaluación programática es perfecta porque, es mejor tener 100 exámenes, que un examen. Es obvio. El problema es que como debemos comenzar en nuestras universidades.

Posiblemente la mejor manera es no cambiar todo al mismo tiempo. Lo mejor es paso a paso. Puede estar en un método, puede estar en una materia, pero para mí lo más importante es el necesario consenso político del personal y los estudiantes, porque, por ejemplo, en España es muy importante el examen para poder ingresar a la formación especializada. No estoy completamente seguro de que los estudiantes prefieran este tipo de evaluación. La evaluación programática no solo se realizó en pregrado, sino también en posgrado y luego en continuo. Es mi opinión.

CVV

Estoy completamente de acuerdo contigo. He estado mucho tiempo en tema del PBL, y si tomo eso como un ejemplo, te diré que igualmente es difícil cambiar de una facultad clásica a una facultad con PBL. Todo se trata de hacer política y hacer que la gente entienda. En mi opinión, creo que deberíamos capacitar a nuestros profesores mucho más. ¿De que disciplina eres?

Mila García Barbero

Es cirujano

CVV

Si tomo el control de su clínica de cirugía, mañana por la mañana, tendría muchos accidentes en la seguridad del paciente, sin embargo, al día siguiente, tu puedes ir a clase de cirugía y podrás enseñar. No tenemos conocimiento de la teoría educativa, de toda la evidencia que existe, totalmente ajena. ¿Qué tipo de práctica es esa? Y creo que para enseñar debemos estar cualificados.

Mila García Barbero

La mayoría de las personas que están aquí están de acuerdo porque son las convencidas.

CVV

Tenéis que comenzar por algún sitio, decanos.

Mila García Barbero

Hemos intentado desde la sociedad española de educación médica, implementar en nuestras facultades, programas de capacitación del profesorado.

CVV

He visto algunas implementaciones de PBL. No necesitan convencerme sobre el valor de PBL. Es un enfoque muy sólido, pero también he visto que las implementaciones de PBL fallan simplemente porque nadie la compraba. El problema sobre si hemos de ir paso a paso, de forma evolucionista, he de decirle que en mi experiencia, los enfoques revolucionarios tienen más sostenibilidad que los enfoques evolutivos, por lo que estoy a favor de un enfoque revolucionario que use la discusión en la línea longitudinal de la comunicación.

Arcadi Gual

¿Es el portfolio absolutamente necesario para la evaluación programática?

CVV

Sí, eso creo. Creo que es un principio.

Arcadi Gual

Entonces, si es necesario, ¿puede recomendar a los decanos cómo implementarlo?

CVV

No, porque es solo un signo. Se debe implementar, pero primero debes convencer a la gente y eso es lo primero. Hay muchos proveedores de portfolios, y se pueden comprar fácilmente. Nosotros originalmente lo hicimos en papel, no tan bueno como el electrónico. El portfolio electrónico es importante pero no es un elemento esencial.

Mila García Barbero

Cuando dijo que está haciendo la misma evaluación a lo largo de los 6 años, lo mismo para 1r, 2do, 3ero y 4rto, entonces, ¿la dificultad se va incrementando?

CVV

Sí, y los estándares aumentan a lo largo del tiempo. Veo sistemas, un sistema español con el MIR por ejemplo, muy orientado a los exámenes estandarizados, porque necesitan dar cuentas a la sociedad de la calidad. Pero nosotros en los Países Bajos hemos tomado un camino diferente en los Países Bajos, al tener este progres-test, junto con otras facultades de Medicina. La prueba no se trata solo de la prueba de medicina de Maastricht, sino de la prueba de medicina. Hemos invitado a participar a otras facultades y en ese momento unos 12,000 estudiantes toman la misma prueba al mismo tiempo en los Países Bajos. Falta unirse algunas facultades, pero lo harán dentro de dos años. Quiero llegar a un examen nacional de conocimiento, que será muy beneficioso para los estudiantes individuales porque recibes muchos comentarios, beneficioso para la cada facultad, porque puedes comparar el rendimiento con el resto y a muy bajo costo al hacerlo en colaboración.

Mila García Barbero

Tal vez podamos unirnos a ellos con un par de facultades.

CVV ¿Por que no?

José Ramón Repullo

Usted mencionó que en Canadá se ayuda a facultades de medicina para estos progres-test, pero el costo de estas actividades es muy alto para una facultad ¿Puede decirme algo más sobre este proceso?

CVV

En primer lugar, una aclaración, ¿te refieres al progres-test"? Esto es en los Países Bajos, no en el Canadá. El progres-test fue una respuesta para facilitar el aprendizaje auto dirigido que queríamos que hicieran los estudiantes y que originalmente teníamos al final de la prueba del módulo. De esta forma conseguíamos que los estudiantes no estudiaran específicamente para el examen.

Quizás ahora sería más fácil ayudarle a configurar algo. También formo parte de una organización europea, el *European Board of Medical Assessors* y de hecho estamos ayudando a las facultades de medicina a establecer un sistema de evaluación del tipo progres-test. Puede usar las preguntas que hacemos que las facilitamos o puede hacer preguntas. Estoy de acuerdo que ello requiere gobernanza nuevamente, que hay un sistema centralizado de elaboración de preguntas y que esto está en conflicto con los silos y con la independencia de los profesores individuales porque desarrollamos preguntas de diferentes disciplinas y luego les pedimos a otras disciplinas que revisen esas preguntas para mejorar la calidad. Esta es una organización central que es muy difícil de lograr en muchas facultades por el mismo problema de gobernanza.

Pablo Lara

Usted mencionó la importancia de la evaluación cualitativa, pero consume mucho más tiempo. Es muy difícil convertirla en un número. Lo sé, porque los estudiantes están interesados en el número y es un gran cambio de cultura. Tengo una asignatura optativa con 12 alumnos. No necesitan hacer ningún examen, o aprobar un examen, pero en el grupo más grande se requiere nuevamente la evaluación cuantitativa. Esto es así.

CVV

Si lo aceptamos, nunca lo cambiaremos.

Pablo Lara

Esa es nuestra situación.

CVV

Estoy en desacuerdo contigo.

Pablo Lara

Yo si estoy de acuerdo contigo. Quisiera saber su opinión sobre el presente y el futuro de la evaluación electrónica.

CVV.

¿Qué consideras como evaluación electrónica?

Pablo Lara

Bueno, hay métodos muy diferentes,

CVV

Veremos mucha más tecnología en el futuro. Tengo un estudiante de doctorado de los Estados Unidos que estaba buscando medidas del proceso para la residencia, porque en los EE.UU. lo tienen todo en la computadora, cada uno de los actores individuales en la computadora, en los registros electrónicos de salud y la pregunta es, si se puede tener en cuenta los registros de salud electrónicos para encontrar medidas de rendimiento de un residente individual. Los datos son bastante positivos, por lo que en el futuro podríamos tener un grupo de trabajo de rendimiento clínico.

Las simulaciones van bastante bien, ya sabes, nuestro nuevo programa de graduados comenzará en 2021. Todas las pruebas de los módulos se envían a la computadora y depende del estudiante inscribirse para la prueba. He estudiado una parte de fisiología, quiero examinarme y la computadora genera una prueba de fisiología. Depende del estudiante. No es para que lo califiquemos, es para su información y para la hoja informativa. Con el tiempo usaremos mucho mas la tecnología. El porfolio electrónico es bastante interesante y ahora estaría haciendo un tipo de investigación para dar sentido a los datos narrativos. Tomará bastantes años de investigación, pero sabemos que veremos más tecnología.

Beatriz Gal

Muchas gracias por su charla. Lo he pasado muy bien escuchándole y estoy de acuerdo con usted que el principal problema es convencer a nuestros profesores, no a los estudiantes, y realmente siento que debemos elegir cómo hacer este proceso. ¿Es un tema de entrenamiento o de hablar con ellos?

CVV

Te explico. Podría contar una larga historia sobre cómo llegamos a todo este tipo de cambios curriculares a lo largo de los años, pero, por lo general, tendemos a cambiar las cosas a lo largo de los años. Cambiar de acuerdo con el espíritu de la ley.

Vamos a renovar el programa de Bachelor para 2021 y rediseñamos el grupo, y el nuevo grupo de diseño está pensando en un sistema de capacitación completamente nuevo, pero al mismo tiempo, estamos involucrando al profesorado, esta es la idea general y tenemos unas sesiones que llamamos de co-creación. Se basa en pensar el diseño, y el profesorado me llama y me deja ver que problema particular tiene y como podemos resolverlo. Tenemos este caso particular, este es su problema, ¿cuántos ayudan a resolver este problema? Tenemos un montón de estas sesiones de co-creación. Los profesores maestros tendrán un nuevo papel en el nuevo programa y, por lo tanto, les invitaremos, a principios del próximo año a diseñar, ayudar a diseñar el programa, para que se sientan dueños del programa, por lo que es un enfoque paso a paso.

Jordi Palés

Como sabes, nuestras Facultades en España son muy tradicionales, con planes de estudio no integrados. Mi pregunta es: ¿es posible implementar una evaluación programática en cualquier Facultad de medicina con un plan de estudios no integrado o es realmente necesario tener un plan de estudios integrado para implementar de manera efectiva la evaluación programática?

CVV

No. Creo que el enfoque clásico de la educación es un enfoque de Maastricht basado en el conocimiento, que era inicialmente el de Maastricht. Posteriormente a partir de los datos procedentes de la investigación educativo se vio que esta no era una estrategia efectiva para el aprendizaje. Uno de los mayores problemas en la educación es el de la transferencia del aprendizaje. El estudiante ha de poder transferir ese conocimiento a situaciones profesionales, puede señalarlo. Los programas educativos modernos están más integrados y proporcionan tareas más auténticas, como punto de partida para el aprendizaje, es decir, para reducir la brecha de transferencia que tenemos en la mayoría de los currículos. Este enfoque moderno de la educación, la denominamos con el término constructivista. Los estudiantes tienen que construir el conocimiento, el aprendizaje no se basa tanto en la transmisión de información, sino en la comprensión de la información, y los estudiantes tienen que construirlos. Si no se acepta ese tipo de estrategia para la educación, la evaluación programática no tiene sentido. Entonces cualquier forma de enfoque constructivista integrado no funcionará.

Creo que funciona de maravilla en la educación de Máster, ya que mucho aprendizaje se desarrolla en el lugar de trabajo donde muchos de los elementos ya están allí y es más fácil de implementar en la educación de posgrado.

Me he referido a cuatro generaciones curriculares. Describí la primera generación, el antiguo plan de estudios basado en conferencias y disciplinas, luego la segunda generación curricular fue el plan de estudios integrado, en su mayoría en base a sistemas, que es difícil porque requiere gobernanza, requiere que los profesores se comuniquen entre sí. En tercer lugar, un currículo basado en problemas, ahora cambiamos el enfoque didáctico, y la cuarta generación son los programas basados en competencias. Entonces, si todavía están en el primer nivel, se debe empezar por la comprensión del problema.

Mila García Barbero

Bueno, tenemos dos minutos antes de que Arcadi me corte el cuello y solo quiero agradecerte todas tus explicaciones sobre la evaluación. Estás en una galaxia y nosotros estamos en la otra galaxia, la otra que existe. No sé a cuántos cientos de millones de años luz estamos. De todos modos, soy optimista, tenemos que empezar por algún lado. Creo que perdimos una gran oportunidad con Bolonia, tuvimos la gran excusa para presentar este tipo de cosas y al final acabamos con las mismas cosas que teníamos antes. Diez horas de conferencia fue 1 ECTS. Este fue el gran cambio

Creo que el problema es convencer, no a los decanos sino a las personas que estas por debajo de los decanos. Además, tenemos que llamar a los estudiantes porque los estudiantes son más flexibles que los profesores. Creo que son más convincentes. Ayer describiste el ejemplo de pasa/falla en lugar de calificaciones. Puede ser que puedas comenzar con un truco. En el primer año, pasar casi de ellos, al año siguiente, a casi todos. Quizás podrás llevar a los estudiantes a donde quieres.

Les agradezco a todos ustedes por estar aquí. Ha sido muy interesante con muchas ideas, difícil de implementar en el sistema español porque está altamente regulado por estructuras ajenas a la Facultad de Medicina, pero como dijo mi madre, si no lo intentas, nunca lo logras. Muchas gracias a todos. Son las 5:30 horas.