



**Universidad del
Rosario**



Pontificia Universidad
JAVERIANA
Bogotá

El fenómeno de “failure to fail” en los posgrados de medicina, una mirada desde la motivación de los docentes.

Autores

Paola Karina García, Isabel Cristina Cuéllar, Angela María Soler

Tutor

Francisco Manuel Olmos

Trabajo para optar por el título de:

Magíster en educación para profesionales de la salud

Escuela de Medicina y Ciencias de la Salud – Facultad de Medicina

Maestría en educación para profesionales de la salud

Universidad del Rosario – Pontificia Universidad Javeriana

Bogotá - Colombia

2024

Tabla de contenido

RESUMEN	3
INTRODUCCIÓN	4
MARCO TEÓRICO	5
METODOLOGÍA	8
Diseño Metodológico	8
Contexto.....	8
Recolección de datos y análisis de la información.	9
Reflexividad.....	11
RESULTADOS	12
Interés y compromiso del docente al momento de evaluar.....	13
La importancia de un sistema de evaluación y la capacitación docente	14
El costo para el docente al momento de evaluar	16
DISCUSIÓN	19
Expectativa reducida de éxito en la evaluación	20
El valor de la tarea evaluativa en la formación de profesionales competentes.	21
El alto costo de la evaluación como factor desmotivador.	21
LIMITACIONES DEL ESTUDIO	23
IMPLICACIONES PARA LA PRÁCTICA	23
CONCLUSIÓN	24
APÉNDICE S1	25
APÉNDICE S2	25
BIBLIOGRAFÍA	27

RESUMEN

Introducción

El "Failure to Fail" es un fenómeno en la educación que ocurre cuando los docentes evitan reprobado a estudiantes de bajo rendimiento, permitiendo que se gradúen sin cumplir con las competencias necesarias. Este estudio explora mediante la Teoría Expectativa-Valor-Costo (TEVC) las motivaciones de los docentes al momento de evaluar, para comprender los factores que perpetúan este fenómeno.

Metodología

Se realizó un estudio cualitativo bajo el enfoque constructivista de la teoría fundamentada. Se realizaron 12 entrevistas semiestructuradas a docentes de especialidades médico-quirúrgicas. Los hallazgos se construyeron a través de un método iterativo de recolección y análisis de datos, comparación constante y muestreo teórico.

Resultados

Desde el punto de vista de la expectativa, los docentes expresaron dificultades para realizar evaluaciones justas debido a la falta de formación y la poca claridad en los procesos evaluativos incluyendo el del profesionalismo. Aunque el valor de querer formar profesionales competentes es alto, no compensa los costos asociados a la evaluación, que incluyen la presión de mantener adecuadas relaciones interpersonales, buena reputación, favorecer logros institucionales, además de los costos relacionados con el esfuerzo y tiempo invertido en los estudiantes, así como el impacto psicológico al considerar las implicaciones económicas y emocionales para los estudiantes.

Conclusión

Los factores motivacionales que influyen en la evaluación incluyen la falta de capacitación docente y los altos costos asociados. Se recomienda implementar procesos de evaluación estructurados y realizar formación docente continua. Abordar estos factores puede reducir el "Failure to Fail" y optimizar la calidad de la educación.

INTRODUCCIÓN

Los docentes en el área de la salud enfrentan un desafío importante al momento de tener que reprobar un estudiante con mal desempeño en la evaluación en el sitio de trabajo, lo que se conoce como el fenómeno de “failure to fail”, también definido como el fracaso a fracasar (Gingerich et al., 2022). Este se presenta “cuando el docente teme dar una calificación no aprobatoria a un estudiante del área de la salud con bajo rendimiento (Barrett et al., 2016; Scarff et al., 2019); prefiriendo callar que enfrentar la conversación difícil sobre su inadecuado desempeño (efecto MUM) (Gingerich et al., 2020; Mak-van Der Vossen, 2019; Scarff et al., 2019; Yepes-Rios et al., 2016). Las estrategias para minimizarlo están bien descritas, sin embargo, este fenómeno persiste (Carr et al., 2010; Gingerich et al., 2020). Existe una amplia literatura que estudia el fenómeno de “failure to fail” y describe muy bien los factores de tipo personal del docente, del estudiante y de las instituciones (Audétat et al., 2017; Daly et al., 2022; Gingerich et al., 2022) que explican las razones para que este fenómeno se dé, sin embargo, no se ha explorado de manera suficiente el rol que tiene la motivación de los docentes al momento de evaluar el desempeño de los residentes y por tanto los factores que influyen en la perpetuación de esta conducta. Indagar estas motivaciones puede mejorar la comprensión de la importancia de discutir el rendimiento insatisfactorio de un estudiante contribuyendo a la enseñanza y aprendizaje en el sitio de trabajo y, por tanto, la atención segura de los pacientes.

Los docentes reconocen deficiencias en el desempeño de los estudiantes, pero no lo reflejan en las evaluaciones formales y se muestran reacios a calificar negativamente, lo que resulta en una promoción no merecida en su programa académico (Gingerich et al., 2022; Rekman et al., 2016). Esto es un problema complejo en la educación médica que impide que los supervisores actúen según sus juicios profesionales sobre el rendimiento de los estudiantes (Gingerich et al., 2020); además, trae como consecuencias una formación deficiente que se refleja en un riesgo para la seguridad del paciente y deterioro de los servicios de salud (Yepes-Rios et al., 2016).

Se han identificado factores promotores del fenómeno de “failure to fail”, que se pueden clasificar en tres grandes grupos: los relacionados con el docente, con el estudiante y los factores asociados al sistema (Daly et al., 2022). Los docentes pueden sentirse influenciados por diversas circunstancias como el temor a perder el prestigio profesional, perder su trabajo, enfrentarse a litigios, además de la carga emocional que implica lidiar con las reacciones del estudiante; adicionalmente el tiempo limitado en los entornos clínicos dificulta el proceso de evaluación (Mak-van Der Vossen, 2019; Rutkowski, 2007).

Por otro lado, los estudiantes también contribuyen al problema, al temer las repercusiones que esta decisión puede tener en su carrera y autoestima, así como su reacción emocional ante el fracaso, lo que genera preocupación en los docentes (Carr et al., 2010; Gingerich et al., 2020). A nivel institucional, la falta de herramientas de evaluación adecuadas, la escasa capacitación del docente, la presión por evitar altas tasas de fracaso que irían en contra del prestigio institucional, la falta de planes estructurados para los estudiantes que reprueban y el temor a perder la “mano de obra” en el sitio de trabajo, contribuyen a perpetuar el fenómeno (Daly et al., 2022). En resumen, el “failure to fail” resulta de las tensiones provenientes de los docentes, el estudiante influenciado por situaciones del sistema y de la cultura institucional; sin

embargo, no se conoce cómo estos factores interactúan entre sí para perpetuar este fenómeno y su intervención podría evitar el “failure to fail. (Yepes-Rios et al., 2016).

En la literatura Daly et al., (2022) plantean 4 estrategias para mitigar este fenómeno: a nivel del estudiante, recomiendan un enfoque basado en etapas centrado en el alumno, logrando que tengan autogestión y los docentes sean más objetivos, al tener más observaciones claramente definidas (Ellaway et al., 2018). En cuanto al desarrollo del cuerpo profesoral, autores como Larocque y Luhanga (2014); Yepes-Rios et al., (2016) proponen que capacitar a los docentes en funciones de supervisión y liderazgo, los ayudaría en adquirir confianza y destreza en habilidades como dar retroalimentación en situaciones difíciles y planificar el aprendizaje basado en el sitio de trabajo (Luhanga et al., 2014). Desde el punto de vista del sistema, Van der Vleuten et al., (2018) sugieren implementar una evaluación programática, lo que favorece el crecimiento y el desarrollo del estudiante y al docente le da una comprensión clara del rendimiento de éste a lo largo del tiempo y no limitarse a una evaluación final. Por último, la estrategia que aborda las tensiones entre el personal y el aprendizaje es la más retadora porque debería alinear la mano de obra y el aprendizaje en un escenario de práctica en el que el supervisor docente, el estudiante y el entorno clínico establecen relaciones complejas ante los modelos de atención de los hospitales (Sanfey, 2011).

El fenómeno “failure to fail” está ampliamente estudiado en la literatura, y se han descrito diferentes estrategias para mitigarlo, sin embargo, no se contemplan específicamente los aspectos motivacionales desde el punto de vista de los docentes al momento de evaluar, así como la relación entre ellos, para comprender los factores que perpetúan este fenómeno.

MARCO TEÓRICO

La teoría escogida para un trabajo de investigación es vital en todas las etapas de un estudio cualitativo, pues esta debe servir de método y como herramienta para abordar desde la pregunta de investigación, el análisis de los datos hasta la elaboración del informe, lo cual guía todo el proceso. De esta forma se pueden obtener perspectivas importantes y novedosas para el estudio. Por lo tanto, es fundamental seleccionarla adecuadamente, pues su aplicación y su conocimiento servirán como método para comprender mejor el fenómeno estudiado (McOwen et al., 2023).

A lo largo del tiempo, han surgido diversas teorías que intentan comprender la motivación. En este contexto, los enfoques más recientes consideran la motivación como la interacción de diferentes dimensiones que influyen en cómo actuamos de manera particular (Loh, 2019). Una de las teorías utilizadas que explica muy bien las razones motivacionales por las cuáles un ser humano toma una decisión es la Teoría Expectativa Valor- Coste (TEVC), donde Eccles y Wigfield (2002) y posteriormente Barrón & Hulleman (2015), adoptan una versión contemporánea para comprender mejor las elecciones, la persistencia y el rendimiento de los estudiantes y docentes.

El componente de la expectativa responde a la pregunta: ¿Puedo hacer la tarea?, refiriéndose a la habilidad que tiene una persona para completar una tarea o si se cree capaz de realizarla en el futuro (Eccles y Wigfield, 2002). El componente de valor responde a la pregunta: ¿Quiero

hacer la tarea? e incluye tres tipos de valores que influyen en la motivación de una persona; el valor intrínseco, que se refiere al placer que la persona experimenta al realizar una tarea por sí misma; el valor de utilidad (extrínseco) que se refiere a la importancia de alcanzar la tarea para lograr otros objetivos y el valor de logro, que refleja cómo la tarea refuerza aspectos de la identidad del individuo y satisface necesidades importantes para él (Eccles & Wigfield, 2002). En cuanto al componente de coste, este lo añadió Eccles en 1983 como un cuarto tipo de valor, definiéndolo como la afectación negativa de una tarea por los costes asociados a su realización, pero Barron & Hulleman (2015) lo elevan a la categoría de la expectativa y valor como un tercer componente conformando la TEVC (Ver Figura 1).

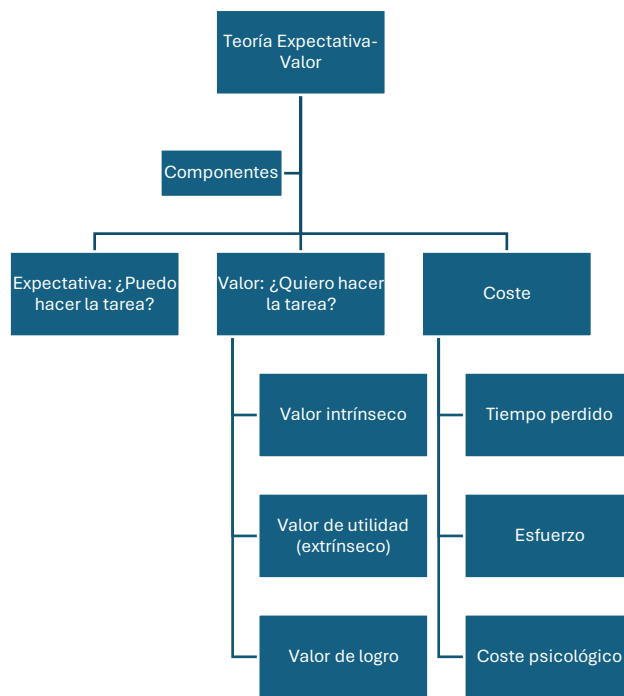


Figura 1. Componentes de la teoría Expectativa-valor.

Los costes se relacionan al tiempo perdido para realizar otras actividades valiosas, la cantidad de esfuerzo necesaria para realizar la tarea con éxito y la lucha o fracaso en esta actividad que corresponden a los costes psicológicos. Sin embargo, esta categoría aún requiere mayor exploración y validación, como si las tienen las de expectativa y valor. Con lo anterior podemos observar que, en la motivación, la expectativa es el factor más potente de los resultados de **rendimiento** y el valor refleja resultados de **interés**, mientras que el coste es un predictor negativo tanto del interés como del rendimiento (Wigfield & Eccles, 2000) (Ver Figura 2).

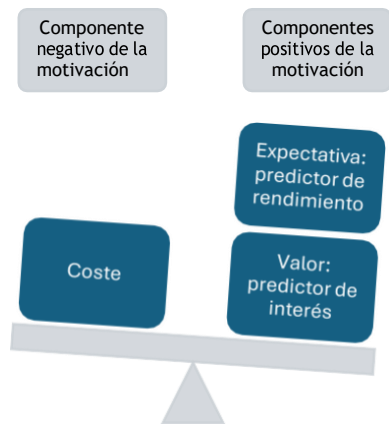


Figura 2: Relación de los componentes de la Teoría Expectativa-valor-coste

La TEVC se ha utilizado para comprender el comportamiento motivacional en la educación, demostrado en varios estudios que permiten entender mejor algunas decisiones de los profesores y los estudiantes. Van Maele et al.,(2021) utilizaron esta misma teoría, para estudiar la intención de investigar de los estudiantes de último año de medicina; otros investigadores (Chan & Hu, 2023) evaluaron la motivación en la decisión de los estudiantes de educación superior de adoptar la inteligencia artificial (IA) en su metodología de estudio. Así mismo, otro estudio muestra su aplicación en estudiantes de posgrados medico quirúrgicos en el que demuestran como la TEVC puede de forma innovadora aplicarse a los residentes de cirugía en su ambiente de trabajo, para promover la motivación a través de un buen desempeño en las tareas programadas, lo cual se reflejaba en un aumento en la exposición a casos quirúrgicos y que de alguna manera este contexto sería semejante al que se enfrentarán en su vida laboral (Shweiki et al., 2015).

Mak-van der Vossen et al., (2018) emplearon la TEVC para explorar la motivación de los estudiantes de medicina frente a las faltas de profesionalismo, considerando expectativas de éxito, valor y costes. Su enfoque abre la puerta para comprender por qué los educadores podrían no sentirse motivados para ofrecer retroalimentación negativa a los estudiantes, argumentando que el efecto del mensaje no es visible para el educador (expectativa), la falta de interés personal (valor) y el esfuerzo involucrado (coste) pueden influir en lo que se enfrentarán en su vida laboral.

La TEVC en nuestro estudio nos servirá para comprender las motivaciones de los docentes al momento de evaluar en el fenómeno del “failure to fail” y más concretamente como el coste percibido sobre la expectativa y el valor de la tarea, determina la intención de los docentes de no reprobar a los estudiantes con bajo rendimiento e indagar sobre los factores interpersonales y del sistema que puedan modificarse para cambiar esta conducta y diseñar intervenciones educativas para aumentar la motivación en los docentes (Eppich et al., 2019).

METODOLOGÍA

Diseño Metodológico

Aplicamos el diseño de la Teoría Fundamentada Constructivista (TFC), siguiendo sus principios que incluyen la recolección y análisis iterativo de los datos, la comparación constante y el muestreo teórico. La TFC es utilizada en educación en salud para explorar el cómo, el qué y el porqué de fenómenos sociales, de esta manera, este diseño nos permitió explorar y entender el fenómeno “Failure to Fail”, el cual no ha sido lo suficientemente estudiado teniendo en cuenta no solo las vivencias de los participantes sino de los investigadores, mediante la reflexión de sus creencias y percepciones que pueden influir en el proceso de investigación permitiendo la construcción de nuevo conocimiento. El proceso analítico implicó varias fases que evolucionaron hasta alcanzar códigos de más alto nivel para llegar a una codificación axial que nos permitió entender mejor la situación planteada (Eppich et al., 2019).

Contexto

La investigación se desarrolló en una Universidad con acreditación institucional multicampus, en Colombia, que cuenta con 43 posgrados de especialidades médico-quirúrgicas que tienen como escenario de práctica principal un Hospital Universitario de alto nivel de complejidad. En Colombia los programas de residencias médicas tienen una duración que va de 3 a 5 años dependiendo de la especialidad, las cuales se pueden cursar después una formación en pregrado en medicina de 6 años y un año de servicio social. Para el estudio se escogieron las especialidades de pediatría, anestesiología, dermatología, cirugía general y medicina de urgencias, ya que contaban con estrategias definidas de evaluación, como por ejemplo rúbricas, Mini-Cex, retroalimentación, entre otras.

- El programa de pediatría tiene más de 60 años de creación y cuenta con acreditación de alta calidad, tiene una duración de 3 años divididos en 6 semestres, con 48 residentes y alrededor de 50 docentes en su principal sitio de práctica, la mayoría son egresados del programa. Cuenta con 3 docentes con maestría y 2 con especialización en educación en salud.
- El programa de anestesiología cuenta con 50 años de fundación, tiene una duración de 3 años divididos en 6 semestres, aproximadamente 33 docentes, la mayoría de ellos egresados del programa. Varios de sus profesores cuentan con formación en educación médica lo que hace que sus procesos de evaluación sean más rigurosos y estructurados.
- El programa de dermatología tiene 7 años, su duración es de 3 años, tiene 12 cupos aprobados y 4 profesores de los cuales uno es especialista en educación para profesionales de la salud.
- El programa de Cirugía General tiene 50 años de creación, la especialización tiene una duración de 4 años, con acreditación de alta calidad y 15 profesores.
- El programa de Medicina de Urgencias tiene 10 años, con una duración de 3 años, con alrededor de 10 docentes.

Recolección de datos y análisis de la información.

Realizamos entrevistas a 12 docentes de 5 especialidades médico - quirúrgicas donde exploramos los aspectos motivacionales al momento de evaluar sus residentes en el sitio de trabajo, para comprender mejor los factores que facilitan el fenómeno del "Failure to Fail". Inicialmente mediante un muestreo propositivo de máxima variación, invitamos a participar a docentes de 3 especialidades médico-quirúrgicas que tuvieran diferentes edades, años de experiencia en su labor docente, variabilidad en los tiempos de vinculación con la institución, experiencia formal en educación y escalafón docente, con el fin de que estos pudieran aportar diversos matices de sus experiencias al momento de evaluar al residente (tabla 1).

TABLA 1 Características de los participantes

CARACTERÍSTICA	No DE PARTICIPANTES
Género	
Femenino	9
Masculino	3
Rango de edades	
31-35 años	2
36-40	2
41-45	2
46-50	5
56-60	1
Especialidad	
Anestesiología	3
Cirugía General	1
Dermatología	3
Medicina de Urgencias	1
Pediatría	4
Rango de años de labor docente	
1-5 años	1
6-10	3
11-15	3
16-20	2
21-25	3
Tipología docente	
Clínico adjunto (contratación hospital)	2
Profesor titular (contratación hospital y universidad)	1
Profesor asociado contratación hospital y universidad)	1
Profesor asistente contratación hospital y universidad)	4
Profesor instructor contratación hospital y universidad)	4
Formación docente (maestría o especialización en educación)	
Si	4
No	8
TOTAL DE PARTICIPANTES	12

Previo firma de consentimiento informado, se realizaron inicialmente 10 entrevistas semiestructuradas (ver apéndice S1) a docentes de pediatría, dermatología y anestesiología, que permitieron indagar el fenómeno estudiado de forma dirigida y precisa. Dado que se iban a tocar temas sensibles relacionados con el fenómeno, se definió realizar entrevistas individuales, pues se consideró que un grupo focal podría cohibir la participación espontánea y directa de los profesores. Se eligieron inicialmente estos tres programas académicos ya que contaban con rúbricas de evaluación estructuradas, para facilitar la construcción interpretativa del fenómeno estudiado y tener una comprensión más amplia, permitiendo la transferibilidad

de los resultados. Los docentes ofrecieron una visión clara y transparente de sus experiencias reconociendo situaciones difíciles al momento de hacer una evaluación.

Se concretaron reuniones del equipo de investigación de forma semanal, para realizar el análisis iterativo de los datos con la ayuda de los memos obtenidos en el transcurso de las entrevistas. A continuación, realizamos 2 entrevistas para explorar información adicional relacionada con la evaluación y la retroalimentación, por lo cual se decidió incluir mediante un muestreo teórico, dos docentes, uno de cirugía general y otro de medicina de urgencias con un ajuste al guion de preguntas (ver apéndice S2).

Las entrevistas se realizaron de forma virtual a través de la plataforma Zoom® lo que permitió un espacio neutral para la realización de estas, logrando también la grabación de video y audio con la herramienta FATHOM® con la cual se garantizó la legitimidad de la transcripción de cada una de ellas. Durante todas las entrevistas las investigadoras hicieron memos con el fin de mantener un proceso de reflexividad continua y tener claro el papel de cada una en la construcción de los resultados (Olmos-Vega et al., 2023).

Inicialmente PG, AS, IC y FO realizaron una codificación abierta de los datos de las primeras cinco entrevistas, las cuales fueron codificadas por mínimo dos investigadores de manera cruzada, utilizando el software Quirkos®, estableciendo alrededor de 28 códigos. Se analizaron los datos de forma iterativa y mediante una comparación constante, decidimos realizar ajustes a las entrevistas como lo determina la TFC, ya que se identificaron áreas adicionales de exploración. Posteriormente, al recopilar 5 entrevistas adicionales, se obtuvieron en total 52 códigos, que mediante una codificación focalizada se consolidaron en 5 categorías más generales (Eppich et al., 2019). Finalmente, se realiza una codificación axial con las categorías obtenidas, logrando agruparlas en tres temas principales, permitiendo proponer una teoría emergente fundamentada en los hallazgos analizados, teniendo en cuenta los elementos de la TEVC (Eccles & Wigfield, 2002) .

Después de 12 entrevistas, alcanzamos la suficiencia teórica cuando nuestras categorías emergentes fueron capaces de explicar el fenómeno estudiado y definimos que realizar nuevas entrevistas aportarían los mismos resultados (Varpio et al., 2017).

Todos los hallazgos encontrados permitieron entender nuevas interrelaciones de los factores que favorecen el “Failure to Fail” y explicaron la motivación de los docentes basados en la TEVC.

Reflexividad

La reflexividad hace referencia al examen constante y crítico del rol del investigador y su influencia en el proceso, teniendo en cuenta que en los estudios cualitativos tanto los participantes como los investigadores construyen los resultados. La reflexividad implica un esfuerzo consciente por entender y minimizar la influencia del investigador sobre los datos y en la interpretación de los fenómenos estudiados, siendo ésta un determinante fundamental de la TFC (Olmos-Vega et al., 2023).

Los investigadores nos desempeñamos como docentes de estudiantes de posgrado, con una trayectoria de más de 15 años, en las áreas de anestesiología, dermatología, pediatría y nefrología. En el proceso de la reflexividad individual, cada una manifestó tener experiencias de situaciones difíciles, al evaluar residentes con bajo rendimiento en el sitio de trabajo. En una reunión de investigadores respondimos la entrevista para conocer la posición de cada uno frente al fenómeno estudiado pudiendo comparar nuestras respuestas y llevando a cabo un proceso de reflexión con respecto a las posiciones de IC, PG y AS como directora de posgrado y coordinadora de dermatología, directora de Educación y coordinadora académica del programa de pediatría respectivamente, cuidando de no ejercer algún grado de poder o intimidación por lo que cada una entrevistó a participantes con los que tenía la menor relación de poder posible.

Durante el estudio las reflexiones de cada investigador se discutieron en el grupo considerando los preconceptos y experiencias previas, siempre reflexionando sobre si los roles de poder podían haber influido en las respuestas de los participantes. Siempre se mantuvo una reflexión interpersonal o colaboración reflexiva para mantener una activa y continua interacción entre las motivaciones, expectativas y suposiciones de los miembros del equipo (Olmos-Vega et al., 2023).

Teniendo en cuenta que el proyecto puede tener un impacto intencionado, al tener respuestas que exploran en profundidad motivaciones, expectativas, valores y costos sobre la evaluación de un estudiante con bajo rendimiento académico, luego de completar las entrevistas, dimos a conocer los resultados a los participantes, quienes estuvieron de acuerdo y no hubo objeción a los mismos, lo que dejó ver que su opinión estaba representada en ellos (Olmos-Vega et al., 2023).

Después de cada entrevista cada una de las investigadoras realizó memos como parte de la reflexividad que buscaron crear consciencia de las preconcepciones que se tenían y adicionalmente hicieron parte de las evidencias que demuestran un trabajo transparente lo que mejora confirmabilidad y transferibilidad del estudio (Frambach & Durning, s. f.)

RESULTADOS

Los resultados de las entrevistas realizadas a los docentes muestran que, en determinadas circunstancias, no se opta por realizar una evaluación justa y estricta que incluya reprobar cuando es necesario; esto se debe principalmente a la combinación de una baja expectativa, debido a la falta de herramientas adecuadas para realizar la evaluación, y un coste elevado, tanto emocional, como burocrático. Lo anterior hace que el valor positivo refiriéndose al interés de evaluar adecuadamente, no sea suficiente para superar estas dificultades. En estas situaciones el fenómeno del 'Failure to Fail' se convierte en una respuesta comprensible al desbalance entre estos factores, ya que evitar el fracaso del estudiante puede parecer una opción menos costosa y arriesgada para los docentes.

Este análisis desde la teoría de la expectativa, valor y coste ofrece una comprensión profunda de cómo y por qué los profesores en estos programas optan por evitar realizar una evaluación

que demuestre el real desempeño del estudiante, lo cual perjudica a largo plazo al programa académico y a la atención en salud.

Interés y compromiso del docente al momento de evaluar

En TEVC el valor corresponde a la voluntad y el gusto de hacer la tarea, de tener un propósito que lo estimula a cumplirla, relacionado con el sentido de identidad profesional, la autorrealización y el crecimiento personal (valor intrínseco, valor extrínseco, valor del logro).

El valor intrínseco que en este caso es querer hacer una evaluación justa y objetiva ya que existe un compromiso de formar buenos especialistas para la comunidad, *“no puedo olvidar esa responsabilidad tan grande que tengo de graduar buenos especialistas. Entonces pienso que ha sido una motivación grande cuando lo hemos tenido que hacer, si bien no es algo de todos los días...esos casos son difíciles, duros, pero lo hemos hecho” (P4); “yo sí pienso que uno debe ser exigente con los residentes porque es estamos hablando de personas que tratan seres humanos y que tienen mucha responsabilidad en sus manos” (P5).*

Algunos participantes consideran que la evaluación debe ser un proceso consciente *“pienso que debemos ser muy responsables con el proceso de evaluación, cada vez que ponemos una nota es darle un valor al esfuerzo de alguien, a los sueños de alguien, a los anhelos de una familia, de una sociedad, pero también tenemos que ser muy conscientes que tenemos una exigencia y tenemos que garantizar que esa persona salga con unas habilidades (P2); “yo creo que hay que reflexionar en la responsabilidad que uno tiene con la evaluación, de hacerla justa” (P3).* Lo cual refleja que, a mayor valor intrínseco de los docentes, mejor será la motivación para dar una calificación objetiva.

En el valor extrínseco se destacan las motivaciones externas que animan al docente a hacer una buena evaluación, como, por ejemplo, las aspiraciones de la institución, de los colegas y de los estudiantes que influyen en la conducta y las decisiones en el momento de evaluar. Los entrevistados manifiestan responsabilidad con la universidad, con el buen nombre del programa académico *“nosotros tenemos que cuidar nuestra institución, pues nuestros egresados son el sello” (P1);* sin embargo, también esto puede ejercer una influencia negativa *“..yo siempre he pensado que aquí nosotros sobre calificamos a las personas ...yo creo que en la medida que se han ido dando premios, como seleccionar el internado dependiendo de notas, todos nos hemos visto envueltos en una presión de poner buenas notas” (P1)* para satisfacer las necesidades institucionales de demostrar calidad y mantener la "imagen" de éxito del programa.

El valor del logro se ve reflejado cuando el docente observa que su residente se gradúa alcanzando las competencias necesarias para ejercer en su vida profesional y reconocen la importancia y el beneficio que se obtiene al reprobar un residente, *“es un error gravísimo, pasarlo, porque en ese momento no estaba cumpliendo el objetivo y los residentes también creo que reflexionan mucho y llegan con una actitud diferente cuando tienen que repetir” (P2); “está la responsabilidad de que el profesional se gradúe con las habilidades mínimas para el*

ejercicio y para su paciente, para el nombre de la universidad, el programa, velar por el mínimo viable” (P7). Un participante relata que también observa actitudes de resiliencia en un residente que repite una rotación y sale adelante, “desafortunadamente el residente no cumplió los objetivos, se quedó, arrancó seis meses después con nosotros, y fue el más amado de todo su grupo de rotación” (P2). El participante 12 comenta una experiencia donde se observa el valor del logro al obtener los resultados esperados en el residente “al principio pues hubo un poquito de resistencia por parte del residente, pero digamos, esta persona en particular era una persona que tenía introspección, yo creo que él en algún momento lo entendió y vio que lo que estábamos haciendo básicamente era.... fortalecer alguna de sus competencias, perdió esta rotación, cuando la repitió e hizo los refuerzos, el comentario final de todo el mundo es «Ven lo, logramos» .

Otro de los aspectos relacionados con el valor de logro es la satisfacción que experimentan los docentes al enseñar y guiar a los estudiantes, en el que la mayoría refieren que disfrutan su rol de educador “yo me siento muy contenta, tranquila preparo mis cosas” (P1), “ha sido bien gratificante, bien interesante todo lo que he trabajado” (P2), “descubrí cómo me ayudaba a estructurar mis conocimientos y mi experiencia” (P11). Los docentes desde que son residentes consideran que empiezan a ejercer como docentes disfrutando su labor “yo me acuerdo perfectamente de que en residencia disfrutaba darle los seminarios que me asignaban los profesores para los estudiantes y sentía que lograba de alguna manera conectar con los estudiantes” (P4). Por otro lado, ser docente contribuye a su educación continua “siento que enseñar es la mejor forma de retroalimentarse uno, para mí es un requisito estar en la formación docente como para yo sentir que estoy actualizada” (P5).

Si bien el proceso de evolución profesional lleva a una satisfacción personal y una mejor comprensión de las dinámicas de evaluación, puede generar dilemas cuando se enfrentan a la decisión de reprobar a un estudiante con mal desempeño, pues este puede verse relacionado con su incapacidad para enseñar de forma apropiada, de esta manera podemos ver como aun teniendo “el valor “para hacer las tareas, las expectativas no son suficientes para alcanzar el objetivo de hacer una buena evaluación.

La importancia de un sistema de evaluación y la capacitación docente

El análisis de la expectativa dentro de la TEVC se enfoca en la percepción que tienen los docentes sobre su capacidad para realizar evaluaciones justas y objetivas. De acuerdo con los hallazgos de las encuestas encontramos lo siguiente:

Existe una preocupación entre los docentes por no contar con las herramientas necesarias para evaluar de manera adecuada. Como lo mencionan varios entrevistados, algunos reconocen que podrían haber cometido errores al no reprobar a estudiantes en su momento debido a esta falta de recursos: “el sabor que nos dejó en ese momento es que si todos hubiéramos estado de acuerdo y pues hubiera reprobado la asignatura quizás no íbamos a tener esa dificultad en su último año de residencia y todos sabemos que es muy grave que lleguen a un tercer año de residencia con unos conceptos deficientes o sin haber cumplido los objetivos” (P2), lo cual puede traer repercusiones negativas en su desempeño laboral.

Los profesores destacan la importancia de contar con esquemas de retroalimentación eficientes, ya que esto permitiría no solo una evaluación más justa, sino también un proceso de autoevaluación y aprendizaje continuo para los residentes. Sin embargo, los mismos docentes reconocen que en ocasiones no han realizado una retroalimentación adecuada, lo que dificulta su capacidad de evaluar objetivamente: *“una retroalimentación con los residentes mejora mucho lo que nosotros pretendemos y saben qué es lo que queremos de ellos”* (P2), *“al no tener, digamos, la referencia clara de cómo retroalimentar al residente, pues se cometen muchos errores...”* (P12). Reconociendo entonces que evaluar requiere saber retroalimentar.

Igualmente, el entrenamiento docente es un aspecto fundamental para tener las herramientas para realizar una evaluación objetiva *“a mí me generaba muchísima ansiedad y me estresaba bastante, por mi personalidad ... pues obviamente recién graduada insegura en el día a día de estar con los residentes me sentía intimidada, pararme en un salón antes de entrar a la clase me daba taquicardia me sentía angustiada, insegura”* (P1); otro participante refiere *“creo que nos hace falta saber un poco más de evaluación; por supuesto, en los últimos años hemos venido puliendo ese proceso y hoy en día yo diría que estamos haciéndolo muchísimo mejor”* (P3) y coinciden en la importancia de la *capacitación de los docentes: “no todos tienen que hacer una maestría, pero si usted le da principios básicos de pedagogía, ¿qué es una competencia?, ¿qué es un ambiente de aprendizaje?, ¿qué es el aprendizaje en lugar de trabajo?, ¿qué es realmente la curva de aprendizaje? ... yo tengo que profundizar en esto y eso me motivó a estudiar, y empecé a entender como es el proceso de aprendizaje y de hablar y entendí que había una forma más adecuada de hacerlo”* (P11). Cada vez se hace más necesaria la educación en docencia para profesionales de la salud, *“hemos aprendido de generación en generación cómo enseñar y eso pues ya no puede ser más así”* (P4). En este sentido, la formación docente no solo mejora la calidad de la enseñanza, sino que también permite a los docentes tener las herramientas para realizar una buena evaluación que contribuya al éxito de sus estudiantes.

Algunos docentes manifiestan dificultades al evaluar aspectos en las rúbricas que facilitan poner una nota mayor a la que realmente merece, en especial en áreas difíciles de abordar *“como por ejemplo hay un ítem en la rúbrica que es relaciones interpersonales, y ahí me cuesta, no hay razón para ponerle una nota bajita, pienso yo. O sea, como que ahí todo el mundo saca cinco a menos que haya sucedido algo, siempre termino poniendo ahí buena nota”* (P7). Mostrando entonces que objetivar la evaluación no resuelve por completo el fenómeno del “Failure to Fail”, puesto que se pueden enmascarar problemas en áreas como la ética y el profesionalismo, permitiendo al residente aprobar la asignatura.

Varios de los entrevistados consideraron que una de las causas del “Failure to Fail” es la falta de documentación adecuada en los procesos de evaluación, ya que en ocasiones los docentes se ven limitados al justificar formalmente una reprobación debido a la ausencia de evidencias concretas. Los profesores consideran que se deben tener *“las pruebas para sustentar que alguien no lo logro, si yo he hecho el seguimiento, pues porque voy a temer, me daría miedo si yo lo rajo sin saber por qué”* (P1); para otro docente esto genera consecuencias como graduar a alguien sin las competencias necesarias *“...este estudiante es muy complicado, ya en este momento no hay nada que hacer, nos tocó graduarlo, nos pasó. Creo que a partir de esa*

experiencia es que hemos empezado a ser más acuciosos con nuestros residentes, a hacer evaluaciones semanales para que precisamente no lleguen al final sin evidencias” (P12). La probabilidad del fenómeno del “Failure to Fail” se disminuye si existen procesos de evaluación claramente documentados.

En los departamentos donde los procesos de evaluación y retroalimentación se llevan a cabo de forma rigurosa y periódica, el proceso de reprobar a un residente con mal desempeño se hace menos difícil, *“emitimos un concepto, pero habitualmente en esos casos ya ese residente ha sido alertado y ya se han hecho varias reuniones, igual llenamos nuestro formato de retroalimentación que ellos firman y se hace un plan de mejoramiento y unos compromisos de parte y parte” (P2).*

En las entrevistas se evidencia que existe un esfuerzo muy grande por parte de los docentes en ir evolucionando de forma positiva en su carrera pedagógica mejorando las estrategias de enseñanza y evaluación: *“uno va cambiando como su rol a lo largo del tiempo no solo por la experiencia sino porque también accede a información y a recursos que te permiten hacerlo” (P1); “poco a poco hemos venido mejorando ese proceso a lo que me refiero es que estoy evaluando puntualidad, estoy evaluando compromiso con el paciente, estoy evaluando cómo se relaciona con su familia, estoy evaluando ya por otro lado conocimientos” (P3). “He madurado, eso es un proceso de identificación, todo nace desde cuando uno ya se vuelve como un experto, como que tiene identidad en la profesión que hace” (P8).*

En resumen, la expectativa de los docentes se ve afectada por la falta de herramientas adecuadas, insuficiente retroalimentación, escasa documentación de las evaluaciones y falta de capacitación docente. Estos factores disminuyen la confianza de los profesores en su capacidad para realizar evaluaciones objetivas y, por lo tanto, aumentan la probabilidad de evitar reprobar a estudiantes con bajo rendimiento, perpetuando el fenómeno del *“Failure to Fail”*.

El costo para el docente al momento de evaluar

En la TEVC, este último se refiere a las barreras para cumplir con la tarea, y se describen costos relacionados con la cantidad de esfuerzo, el tiempo perdido y el aspecto psicológico y emocional que conlleva reprobar a un residente.

Los docentes entrevistados temen ser juzgados negativamente por sus jefes, colegas y estudiantes pues cuando un docente reprueba un residente con bajo desempeño le puede generar un costo no solo en términos de la relación con los estudiantes, sino también en cuanto a su reputación y percepción en el entorno académico. Según el participante 2, este proceso puede afectar *“el prestigio del profesor”* perdiendo *“el favoritismo entre los residentes” (P2), “¿pienso que dirán a la espalda de uno?” (P9), “... hay un escarnio público del profesor...se rotula como problemático” (P11),* lo que repercute en la percepción general que se tiene del docente. Otro entrevistado considera que la asignatura donde se produce esta situación también se va a ver afectada *“no va a ser la rotación más querida por el semestre” (P2).* Todo lo anterior puede generar tensiones en el ambiente de aprendizaje, especialmente en sistemas donde la reprobación de los estudiantes se ve como un fracaso del propio proceso docente o

de la institución lo que lleva al “Failure to Fail” con el fin de evitar ser juzgados de forma negativa.

El proceso de reprobar a un residente con mal desempeño requiere una inversión adicional de tiempo y esfuerzo para los docentes no solo para justificar la decisión, sino también para darle seguimiento al proceso de recuperación y evaluación, lo que es mencionado por varios de los entrevistados:

- *“creo que es trabajo extra, tengo esa sensación como que tendría uno que asumir las explicaciones, la reunión, y la justificación” (P7).*
- *“pocas veces hay el espacio protegido de sentarse a pensar en estas cosas, digamos saquemos 10 minutos para hacer esta herramienta, inventemos un formato, no tenemos incluso dentro de la facultad espacios para sentarnos de verdad a pensar en eso” (P1).*
- *“la persona que no rinde muchas veces sale beneficiada, cuando me dedico a acompañar a alguien que no rinde, requiere más tiempo y esfuerzo” (P1).*

Todo lo anterior, produce un considerable costo en términos de tiempo para los docentes.

El costo psicológico que enfrentan los docentes en el momento de reprobar a un estudiante con mal desempeño es un aspecto crítico que perpetúa el fenómeno del “Failure to Fail”, pues este problema genera una carga emocional ya que se siente atrapado entre la empatía y la obligación de ser objetivo, lo que provoca un desgaste emocional considerable, *“tenemos de alguna manera una conexión personal con ese residente, que además conocemos su historia familiar desde que lo entrevistamos en el proceso de admisión... que de pronto nos generó inclusive sentimientos de admiración desde el punto de vista de una persona que se ha superado... y por eso mismo quiero ayudarlo” (P4).* Los evaluadores enfrentan mayores dificultades al dar notas no satisfactorias a residentes con quienes establecen vínculos cercanos y duraderos *“tengo la sensación de que de aquí a 12 horas somos colegas y nos seguimos viendo... hacen casi que parte de la familia” (P7).* Factores como la empatía, lazos personales o favores recibidos pueden influir en los evaluadores para otorgar calificaciones más altas de las merecidas.

Otro de los aspectos que conduce a un costo psicológico es el temor de los docentes de ser los responsables de truncar la carrera de un estudiante *“uno sabe el impacto que tiene en la residencia perder algo y se siente uno como culpable” (P7).* *“Al reprobar a un estudiante también te da pesar, además vamos a tener unos días difíciles y le puede a uno afectar” (P9).* Inclusive uno de los participantes considera que otro factor es el temor por su integridad al reprobar a los estudiantes *“uno no sabe de quién sea hijo, ni qué influencias tiene, termina uno con problemas, no sé, con su familia, porque reprobó un residente. Eso es un arma, la verdad, en este país que pereza.... o va y se queje con alguien” (P5).*

Por otra parte, el acto de reprobar a un residente no solo tiene implicaciones académicas, sino también económicas, al considerar el impacto financiero en el estudiante, que puede generar mayor presión para el docente en el momento de evaluar a un residente *“la parte económica, a veces lo pone a pensar más, lo que realmente no se debería hacer, pensar en los papás que*

están pagando ese mes... porque yo le ponga dos en su rotación por su relación con el paciente, pierda la materia y que por perder la materia tenga que pagar otro semestre” (P5); “todo lo que ellos tienen que sufrir para hacer eso, me pongo en el lugar de ellos, el pago la matrícula, se extiende en el tiempo, se retrasan del grupo, entonces se da una nota más baja que el promedio de los otros y ya” (P10); “volver a pagar, eso es una cosa real.....si eso no lo impactara económicamente....” (P8). Estos factores de alargar el tiempo de estudio y aumentar la carga económica no solo afecta al residente, sino que genera en los docentes un costo psicológico que influye en el momento de reprobar al residente.

Sumado a esto, el fenómeno de “Failure to Fail” se agrava cuando los docentes no reciben el respaldo adecuado por parte de las instituciones. En muchas ocasiones, los docentes se sienten solos en la toma de decisiones difíciles, sin contar con una estructura clara que los apoye en el manejo de casos de bajo rendimiento *“el residente reprobó la asignatura, e hice todo el proceso, ¿qué pasó? que después otra universidad lo pasó y se graduó” (P5), “ya uno no sabe cómo abordar las situaciones y necesita como un apoyo desde la universidad para fortalecer el abordaje de los estudiantes” (P10), otro participante cuenta su experiencia al reprobar a un residente: “el reprobar a la residente me generó de pronto un poquito de decepción pensar que en ese momento mi jefe no estuviera teniendo en cuenta esa evaluación que yo estaba tratando de hacer bien.....yo describiría también este proceso como muy exigente, eso ha implicado que estemos todo el tiempo actualizándonos”* lo que evidencia un costo relacionado con la falta de apoyo de las instituciones educativas y al esfuerzo para llevar a cabo las evaluaciones de forma adecuada.

Un docente refiere que el motivo principal para no reprobar un residente, *“es evitar la consecuencia de ir a meterme en un proceso me van a llamar de posgrados a ver qué pasó y el decano me va a decir, ¿pero este estudiante, que pasó? en fin este tipo de cosas creo que influyen radicalmente” (P4)* o en otros casos, prefieren evitar ver al residente nuevamente repitiendo la rotación *“yo no me lo aguanto otro mes, si uno llega y le ve la cara de yo no quiero esta consulta, pues yo no voy a querer seguir viendo esa cara por un mes más” (P5); “yo les digo que repita la rotación, y respondieron, no aquí no la volvemos a recibir, no la volvemos a ver en su vida” (P8).* Otro entrevistado plantea también esa presión que sienten los docentes y es que las instituciones deben asegurar las finanzas *“por asegurarnos “x” millones de pesos de matrículas...se nos daña el indicador de deserción, entonces no podemos permitir que les vaya mal “(P11).* De esa manera, optan por el camino más fácil de dar una nota aprobatoria y caer en el “Failure to Fail”.

Finalmente, situaciones que involucran maltrato o acoso, con acusaciones bidireccionales pueden generar en los docentes mayores costos al enfrentarse con un residente que tenga un mal desempeño *“al no tener, la referencia clara de cómo retroalimentar al residente, pues se comentan muchos errores.....y digamos que esos errores pues claramente se pueden malinterpretar o se pueden interpretar como actitudes de acoso o de maltrato.... la gente con miedo entonces va a decir no, yo prefiero, no decirle nada para que no haya problemas” (P12); “¿qué le digo o no le digo? porque que tal que lo va a tomar como maltrato” (P11).* Para otro entrevistado el ser claro y directo en las evaluaciones puede generar situaciones difíciles que ponen en riesgo enfrentarse a acusaciones de maltrato *“la confianza de decirle al residente que*

se está equivocando es mal interpretada y se convierte en quejas con las autoridades de la universidad, se pasan cartas donde se evidencia que las quejas son las retroalimentaciones que les hacemos a ellos” (P11).

Esta dimensión del coste engloba aspectos como los descritos anteriormente que dan peso en la balanza de la teoría EVC hacia favorecer el fenómeno del “Failure to Fail” y permiten entender ese comportamiento motivacional de los docentes en su proceso de evaluar (Ver Figura 3).

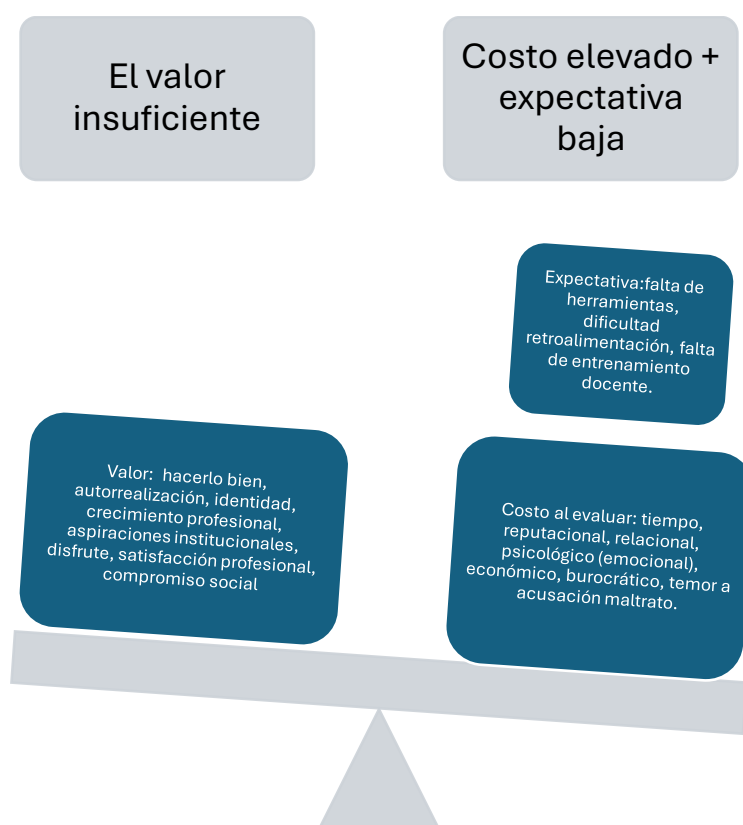


Figura 3. Resultados basados en la Teoría Expectativa, valor y coste en el fenómeno del “Failure to Fail”

DISCUSIÓN

En este estudio pudimos observar cuáles son los principales aspectos motivacionales de los docentes al momento de evaluar y cómo interactúan entre sí para favorecer el fenómeno del “Failure to Fail”. Basados en la TEVC, encontramos 3 puntos importantes: una baja expectativa de los docentes al contar con pocas herramientas y capacidad para evaluar a los estudiantes; un valor refiriéndose al interés de los docentes en conservar el compromiso de cumplir con la labor docente de manera transparente al realizar una evaluación pero que es insuficiente al tener en cuenta los costos elevados que implican un mayor esfuerzo, el tiempo invertido, las implicaciones emocionales y los problemas a los que pueden verse enfrentados.

La TEVC establece que la expectativa de éxito, el valor de la tarea y los costos asociados determinan la motivación de los individuos para comprometerse con una tarea específica (Eccles y Wigfield, 2002). En el contexto del "Failure to Fail", los docentes enfrentan un conflicto en el que la baja expectativa en su capacidad para realizar evaluaciones justas y objetivas, unida a los altos costos percibidos, contrarresta el valor significativo de realizar adecuadamente la tarea de evaluar con el fin de formar profesionales competentes. Esto resulta en una resistencia a reprobar estudiantes con bajo rendimiento, a pesar de su deseo de contribuir a la educación de calidad en salud. Estudios como el de Yepes-Ríos et al., (2016) y Scarff et al., (2019) destacan que el valor de la tarea pierde fuerza cuando los costos percibidos son altos y las expectativas de éxito son bajas. Los hallazgos del estudio revelan que el "Failure to Fail" responde a una combinación de expectativas, valores y costos percibidos por lo docentes al momento de evaluar a los residentes y específicamente en los casos donde el desempeño del residente no ha sido el esperado.

Expectativa reducida de éxito en la evaluación

La expectativa, en este contexto, se relaciona con la confianza que los docentes tienen en su capacidad para llevar a cabo una evaluación justa y efectiva. Estudios como el de Audétat et al., (2017) sugieren que las barreras individuales, como la falta de herramientas evaluativas contribuyen a la tendencia a evitar evaluaciones negativas. En consonancia, nuestro estudio confirma que los docentes enfrentan dificultades para acceder a herramientas de evaluación objetivas y adecuadas, un problema que parece estar presente en múltiples contextos académicos. Los docentes expresaron que debido a la falta de herramientas y de formación en métodos evaluativos, sienten que no cuentan con el respaldo suficiente para aplicar criterios rigurosos sin ambigüedad o dudas y como consecuencia, perciben que no pueden hacer una evaluación justa y eficaz, lo cual disminuye la expectativa de éxito en la tarea. La literatura respalda esta conclusión, indicando que la baja expectativa de éxito se relaciona con la falta de capacitación y claridad en los procesos de evaluación (Barrett et al., 2016; Mak-van Der Vossen et al., 2018). Estudios como el de Carr et al., (2010) demuestran que la baja expectativa se intensifica en contextos donde los docentes no tienen una estructura formal de evaluación que facilite la recopilación de datos suficientes para justificar una calificación negativa. De esta forma, las limitaciones en las herramientas y el proceso evaluativo generan una sensación de incompetencia al evaluar, que disuade a los docentes de emprender acciones que, aunque necesarias, perciben como riesgosas o poco efectivas.

En nuestro estudio particularmente encontramos que no solo existe una barrera en la expectativa por la falta de herramientas de formación en métodos evaluativos de los docentes, también se evidencia que en algunas especialidades a pesar de contar con métodos de evaluación estructurados, existe dificultad en evaluar algunas dimensiones específicas del desempeño de los estudiantes como el profesionalismo, la colaboración y las habilidades de comunicación, asignando una nota sobrevalorada permitiendo que el estudiante apruebe, exacerbando el fenómeno del "Failure to Fail". Ong Y.T et al., (2020), en una revisión de alcance plantean que una evaluación global del profesionalismo tiene retos en 4 dimensiones como definir qué es, cuáles son sus dominios, qué enfoque se adopta sobre éste en la formación y las barreras o facilitadores en la enseñanza de este; enfatizando que la falta de una

definición clara de profesionalismo impacta de forma importante en los diferentes currículos de enseñanza y por lo tanto en evaluarlo en un escenario de práctica. Sin embargo, ante estos retos diversas organizaciones como la National Board of Medical Examiners, el Royal College Of Physicians and Surgeons of Canadá, el Accreditation Council for Graduate Medical Education entre otros han realizado esfuerzos para que estas habilidades no se pierdan en la medicina y han elaborado herramientas para enseñar y evaluar estos conceptos de manera formal (Iserson, 2019)

El valor de la tarea evaluativa en la formación de profesionales competentes.

Los docentes en este estudio expresaron un valor intrínseco y extrínseco alto al momento de evaluar. Desean contribuir significativamente a la formación de futuros profesionales competentes y reconocen que, aprobar a estudiantes con bajo rendimiento, podría comprometer la calidad del servicio en salud, lo que va en contra de su responsabilidad social. Este valor es consistente con estudios previos que destacan el compromiso de los docentes con la formación de profesionales de la salud altamente cualificados (Gingerich et al., 2022). Los docentes experimentan un sentido de responsabilidad hacia los estudiantes y los pacientes, y esto refleja un valor de identidad en el que se consideran guardianes de la calidad educativa y profesional para el beneficio de la sociedad. Según el meta análisis de Da Rosa M et al., (2024) los médicos que mostraron comportamientos poco profesionales durante su formación médica tenían 2.54 veces más probabilidades de enfrentar acciones disciplinarias en su vida profesional.

En el contexto del "Failure to Fail", este valor intrínseco y extrínseco no se traduce en una acción evaluativa rigurosa, debido a que no se compensa por los bajos niveles de expectativa y los altos costos percibidos. En palabras de Mak-van Der Vossen et al., (2018), cuando el valor no logra superar los costos percibidos, incluso un alto compromiso en realizar bien la tarea de evaluar, asociado al deseo del crecimiento profesional, la satisfacción personal, las aspiraciones institucionales y el compromiso social, resultan insuficientes para generar una evaluación crítica y objetiva. Esto pone de relieve la necesidad de no solo aumentar el valor de la evaluación en términos de formación ética y profesional, sino también de desarrollar estrategias que eleven las expectativas de éxito y reduzcan los costos asociados.

El alto costo de la evaluación como factor desmotivador.

El costo asociado a realizar una evaluación justa que refleje el desempeño del estudiante es percibido como una barrera emocional significativa que promueve el "Failure to Fail". Los docentes expresaron que, al considerar reprobar a un estudiante, enfrentan altos costos emocionales y relacionales: desde el temor de dañar la relación con el estudiante, hasta el esfuerzo emocional que implica la confrontación directa. Además, el tiempo invertido en estudiantes con bajo rendimiento se considera un costo negativo que afecta la disposición para evaluar críticamente. Como concluye Gingerich et al., (2020), el costo emocional percibido es un factor desmotivador central en la evaluación negativa, y nuestro estudio respalda esta conclusión, sugiriendo que los costos psicológicos y relacionales pueden llevar a los docentes a evitar reprobar al estudiante.

Además de los costos emocionales, en el estudio se identificaron costos profesionales, tales como el riesgo de afectar la reputación y el temor a represalias dentro del entorno institucional. Los docentes manifestaron enfrentar incertidumbre sobre las repercusiones sociales, laborales e inclusive legales que podrían surgir de una calificación negativa, especialmente cuando no tienen certeza sobre el apoyo institucional en estos casos. La literatura ha mencionado estos costos, especialmente en contextos donde los docentes sienten que una evaluación negativa puede tener repercusiones laborales (Audétat et al., 2017). Asimismo, Barrett et al., (2016) y Yepes-Ríos et al., (2016) describen que el temor a las repercusiones emocionales y relacionales afecta la objetividad de los docentes.

Los entrevistados manifiestan enfrentar presiones institucionales y culturales para mantener altas tasas de éxito académico, limitando aún más la disposición de los docentes para evitar reprobar a estudiantes que no cumplen con los estándares requeridos. Los estudios de Makvan Der Vossen et al., (2018) y de Liu et al., (2024), señalan el papel de las políticas institucionales en la perpetuación del "Failure to Fail", revelando que la cultura institucional también refuerza las prácticas de evaluación permisivas, promoviendo que los estudiantes culminen sus estudios aún sin cumplir con los estándares de competencias clínicas.

Uno de los factores que identificamos en el estudio que propicia el "Failure to Fail", fue el temor a pensar que una retroalimentación pueda confundirse o malinterpretarse con situaciones de maltrato. Lind et al., (2020) manifiestan que los docentes a menudo tienen miedo de presionar demasiado y ser percibidos como ofensivos o autoritarios. Este temor contribuye a tergiversar la diferencia entre el maltrato real y un entorno de aprendizaje subóptimo. El estudio de Lind et al., (2020) consistió en escuchar las voces de estudiantes y docentes en una estrategia denominada "Ending Mistreatment Task Force" (EMTF), lo cual identificó que la colaboración entre ambos grupos reforzó el compromiso de mantener un entorno de aprendizaje y atención clínica respetuoso y libre de maltrato. Además, recomiendan la discusión permanente para aclarar la definición de maltrato y distinguirlo de otras experiencias negativas que no califican como tal. En este sentido, Dart J et al., (2023) plantean que asegurar momentos de retroalimentación donde se conozca el contexto de cada estudiante, se identifiquen tempranamente los problemas y se creen espacios seguros para conversaciones constructivas que aumente la confianza en los docentes para reprobar un estudiante por fallas significativas.

También observamos que los docentes optan por disminuir el "costo" que tiene reprobar un residente y verlo nuevamente, prefiriendo darle una nota aprobatoria. Tiana et al., (2008) defienden la necesidad de equilibrar las prácticas de promoción automática con un enfoque que no tenga miedo de reconocer el fracaso cuando sea necesario, para así proteger la calidad educativa y mantener altos estándares académicos. Por eso recomiendan que las escuelas y los docentes reevalúen sus políticas de promoción y adopten un enfoque más matizado que considere tanto el bienestar del estudiante como la necesidad de asegurar un aprendizaje efectivo.

Los hallazgos de nuestro estudio ofrecen una contribución significativa al campo de la educación médica, abordando de manera novedosa las motivaciones y los obstáculos que enfrentan los docentes al evaluar en el contexto del "Failure to Fail". La incorporación de la TEVC, al lado de una metodología cualitativa bajo la teoría fundamentada constructivista,

permite una comprensión más amplia de las dinámicas motivacionales de los docentes al momento de evaluar, proporcionando una estructura teórica robusta para futuras investigaciones.

LIMITACIONES DEL ESTUDIO

Nuestro estudio se realizó en una Institución Educativa Universitaria privada en un contexto clínico de un Hospital Universitario, de Bogotá Colombia en donde se realizaron entrevistas a docentes de 5 programas de especializaciones médico-quirúrgicas; aunque el fenómeno del "Failure to Fail" es reconocido globalmente, los lectores deberán identificar la aplicabilidad de los resultados en sus propios entornos y contextos.

El muestreo intencional buscó diversidad en experiencia y formación de los participantes, sin embargo, la transferibilidad de estos resultados a otras instituciones o contextos puede ser limitada, dado que el fenómeno del "Failure to Fail" podría manifestarse de forma distinta según la cultura institucional y el soporte que cada entidad ofrece a sus docentes. Por otra parte, la mayoría de los docentes entrevistados eran egresados del mismo programa, lo que podría haber influido en sus respuestas, dado que comparten una formación similar y pueden tener puntos de vista alineados respecto a las expectativas y barreras en la evaluación. Este estudio exploró la evaluación en un contexto de programas con rúbricas estructuradas que abarcan diversas dimensiones al momento de evaluar; lo que limita los hallazgos en contextos de programas sin estas herramientas.

La recolección de datos se llevó a cabo durante un período de tiempo limitado, lo cual restringió el análisis longitudinal del fenómeno. Esto impidió observar cómo las decisiones de evaluación de los docentes pueden cambiar a lo largo de los ciclos académicos o en respuesta a intervenciones pedagógicas. Un estudio longitudinal permitiría explorar cómo la experiencia acumulada y la exposición a programas de capacitación afectan las decisiones evaluativas de los docentes a largo plazo.

Las limitaciones presentadas no invalidan los hallazgos, sino que enfatizan la importancia de interpretarlos dentro de su contexto. Reconocer estos desafíos permite que futuras investigaciones expandan y profundicen en otras áreas y contextos para medir el impacto de factores culturales e institucionales en la motivación de los docentes al momento de evaluar.

IMPLICACIONES PARA LA PRÁCTICA

Los resultados sugieren que los docentes se encuentran motivados y comprometidos en formar talento humano de alta calidad con valores que resaltan el profesionalismo de su especialidad, sin embargo se evidencia la necesidad de un análisis y revisión de las prácticas y políticas institucionales para reducir los factores que perpetúan el "Failure to Fail", tales como una capacitación docente en evaluación, inmersa en un sistema sólido de seguimiento, acompañamiento y retroalimentación en el proceso de formación del estudiante.

Este sistema de evaluación fortalecido debe enfocarse en la documentación y argumentación veraz y en línea con las diversas fuentes de información que tiene cada programa para gestionar temprana y eficazmente un estudiante con bajo rendimiento, de tal manera que permita compartir, analizar y reflexionar, en un ambiente de confianza la situación para que se dé un acompañamiento efectivo para lograr alcanzar las competencias y habilidades requeridas.

En la tarea de retroalimentar, como parte de la evaluación se requiere un esfuerzo colectivo y cultural que busque espacios realmente seguros y constructivos en conversaciones difíciles y para esto se sugieren talleres de formación en técnicas de retroalimentación enfocadas en estudiantes de bajo rendimiento, favoreciendo oportunidades de mejoramiento incorporadas en los currículos que disminuyan la carga emocional tanto de estudiantes como docentes.

Finalmente, nuestros hallazgos destacan la importancia de evaluar correctamente habilidades como el profesionalismo y la comunicación basados en definiciones claras en el contexto institucional que se evidencie en el currículo oculto.

CONCLUSIÓN

En conclusión, este estudio no solo corrobora aspectos previamente explorados sobre el "Failure to Fail", sino que también amplía la comprensión de las interrelaciones de este fenómeno al destacar los costos emocionales y las influencias culturales que afectan las decisiones evaluativas. Estos hallazgos sugieren la necesidad de desarrollar políticas educativas con sistemas de evaluación robustos, capacitaciones a los docentes, fomentando una cultura de transparencia y objetividad que, en última instancia, beneficiará tanto a los estudiantes como a la atención en salud.

APÉNDICE S1

Formato inicial de entrevistas

1. Cuéntenos, ¿usted hace cuanto es docente?,
2. ¿Cómo se siente desempeñándose como docente?
3. ¿Cuéntenos que estrategias utiliza para evaluar o retroalimentar un residente, y cómo ha sido su experiencia?
4. ¿Cuál ha sido su impresión en su ejercicio de evaluación utilizando rúbricas?
5. ¿Cómo ha sido su experiencia en evaluar un residente con mal desempeño en las actividades del lugar de trabajo, podría dar ejemplos?
6. ¿Ha tenido la experiencia de poner una calificación superior a la que refleja el desempeño del estudiante? ¿En qué casos o en que aspecto de la evaluación o ítem (o de la rúbrica que tenga) le ha pasado?
7. ¿Qué criterios tiene en cuenta para reprobar un residente?
8. ¿Cuáles serían las razones para preferir dar una nota aprobatoria a un residente que no lo merece? ¿Puede contarnos, si le ha sucedido?
9. ¿Cuáles cree usted que pueden ser los beneficios que puede tener en su labor docente y para el programa, el reprobar un residente con mal desempeño?
10. ¿Qué consecuencias cree que puede traer para usted y al residente si lo reprueba?
11. ¿Qué elementos pesan en su decisión de reprobar un residente?
12. ¿Quisiera agregar algo más?
13. Si tiene que escoger lo más importante de la conversación, ¿qué destacaría?

APÉNDICE S2

Formato de preguntas modificado

1. Cuéntenos, ¿usted hace cuanto es docente?,
2. ¿En el momento actual, cómo se siente desempeñándose como docente?
3. ¿Cuéntenos que estrategias utiliza para evaluar o retroalimentar un residente, y cómo ha sido su experiencia?
4. ¿Cómo ha sido su experiencia en evaluar un residente con mal desempeño en las actividades del lugar de trabajo, podría dar ejemplos?
5. Los ambientes clínicos en los que nos desenvolvemos son espacios enseñanza y aprendizaje que ocurren en los escenarios clínicos y que se relacionan de forma directa con la atención de pacientes y familia. Es en estos ambientes en los que los estudiantes

y los residentes aprenden a ser médicos o especialista; ¿cómo cree que el área donde se desempeña como docente, influye en el momento de evaluar?

6. ¿Qué criterios tienen en cuenta para reprobar un residente?
7. ¿Cuáles serían las razones para preferir dar una nota aprobatoria a un residente que no lo merece? ¿Puede contarnos, si le ha sucedido?
8. ¿De qué manera las quejas de maltrato pueden influir en la decisión de un profesor de no reprobar a un residente que lo merece?
9. ¿Cómo cree que las quejas de maltrato de los residentes hacia los profesores dificultan la retroalimentación objetiva y honesta sobre el desempeño de los residentes?
10. ¿Qué cree que podría ayudar a mitigar el impacto de las acusaciones de maltrato hacia los residentes en la evaluación justa y objetiva de los residentes?
11. ¿Qué consecuencias cree que puede traer para usted y al residente si lo reprueba?
12. ¿Quisiera agregar algo más?
13. Si tiene que escoger lo más importante de la conversación, ¿qué destacaría?

BIBLIOGRAFÍA

- Audétat, M.-C., Laurin, S., Dory, V., Charlin, B., & Nendaz, M. R. (2017). Diagnosis and management of clinical reasoning difficulties: Part I. Clinical reasoning supervision and educational diagnosis. *Medical Teacher*, 39(8), 792-796.
<https://doi.org/10.1080/0142159X.2017.1331033>
- Barrett, A., Galvin, R., Steinert, Y., Scherpbier, A., O'Shaughnessy, A., Horgan, M., & Horsley, T. (2016). A BEME (Best Evidence in Medical Education) review of the use of workplace-based assessment in identifying and remediating underperformance among postgraduate medical trainees: BEME Guide No. 43. *Medical Teacher*, 38(12), 1188-1198.
<https://doi.org/10.1080/0142159X.2016.1215413>
- Barron, K. E., & Hulleman, C. S. (2015). Expectancy-Value-Cost Model of Motivation. En *International Encyclopedia of the Social & Behavioral Sciences* (pp. 503-509). Elsevier. <https://doi.org/10.1016/B978-0-08-097086-8.26099-6>
- Carr, J., Heggarty, H., Carr, M., Fulwood, D., Goodwin, C., Walker, W., & Whittingham, K. (2010). Reflect for success: Recommendations for mentors managing failing students. *British Journal of Community Nursing*, 15(12), 594-596. <https://doi.org/10.12968/bjcn.2010.15.12.594>
- Chan, C. K. Y., & Hu, W. (2023). Students' voices on generative AI: Perceptions, benefits, and challenges in higher education. *International Journal of Educational Technology in Higher Education*, 20(1), 43.
<https://doi.org/10.1186/s41239-023-00411-8>

- Da Rosa, M. I., Ceretta, L. B., Martins, M. A., Colonetti, L., Colonetti, T., Grande, A. J., & Tempski, P. (2024). The effects of academic unprofessional behaviour on disciplinary action by medical boards: Systematic review and meta-analysis. *The Clinical Teacher*, 21(4), e13740.
<https://doi.org/10.1111/tct.13740>
- Daly, M., Lucas, G., & Wilkinson, T. J. (2022). Learning, judgement and workforce tensions impact failure to fail in a medical postgraduate setting: A supervisor perspective. *Medical Teacher*, 44(10), 1092-1099.
<https://doi.org/10.1080/0142159X.2022.2058385>
- Dart, J., Rees, C., Ash, S., McCall, L., & Palermo, C. (2023). Shifting the narrative and practice of assessing professionalism in dietetics education: An Australasian qualitative study. *Nutrition & Dietetics*, 80(3), 240-252.
<https://doi.org/10.1111/1747-0080.12804>
- Eccles, J. S., & Wigfield, A. (2002). Motivational Beliefs, Values, and Goals. *Annual Review of Psychology*, 53(1), 109-132.
<https://doi.org/10.1146/annurev.psych.53.100901.135153>
- Ellaway, R. H., Chou, C. L., & Kalet, A. L. (2018). Situating Remediation: Accommodating Success and Failure in Medical Education Systems. *Academic Medicine*, 93(3), 391-398.
<https://doi.org/10.1097/ACM.0000000000001855>
- Eppich, W. J., Olmos-Vega, F. M., & Watling, C. J. (2019). Grounded Theory Methodology: Key Principles. En D. Nestel, J. Hui, K. Kunkler, M. W. Scerbo, & A. W. Calhoun (Eds.), *Healthcare Simulation Research* (pp. 127-133).

Springer International Publishing. https://doi.org/10.1007/978-3-030-26837-4_18

Frambach, J. M., & Durning, S. J. (s. f.). *AM Last Page: Quality Criteria in Qualitative and Quantitative Research*.

Gingerich, A., Sebok-Syer, S. S., Larstone, R., Watling, C. J., & Lingard, L. (2020).

Seeing but not believing: Insights into the intractability of failure to fail.

Medical Education, 54(12), 1148-1158.

<https://doi.org/10.1111/medu.14271>

Gingerich, A., Sebok-Syer, S. S., Lingard, L., & Watling, C. J. (2022). The shift from

disbelieving underperformance to recognising failure: A tipping point

model. *Medical Education*, 56(4), 395-406.

<https://doi.org/10.1111/medu.14681>

Iserson, K. V. (2019). Talking About Professionalism Through the Lens of

Professional Identity. *AEM Education and Training*, 3(1), 105-112.

<https://doi.org/10.1002/aet2.10307>

Lind, K. T., Osborne, C. M., Badesch, B., Blood, A., & Lowenstein, S. R. (2020).

Ending student mistreatment: Early successes and continuing challenges.

Medical Education Online, 25(1), 1690846.

<https://doi.org/10.1080/10872981.2019.1690846>

Liu, C., Rajagopalan, J., Wainman, B., Wojkowski, S., Pierazzo, J., & Akhtar-Danesh,

N. (2024). Why do instructors pass underperforming students? A Q-

methodology study. *BMC Medical Education*, 24(1), 1135.

<https://doi.org/10.1186/s12909-024-06126-2>

- Loh, E. K. Y. (2019). What we know about expectancy-value theory, and how it helps to design a sustained motivating learning environment. *System*, 86, 102119. <https://doi.org/10.1016/j.system.2019.102119>
- Luhanga, F., Larocque, S., MacEwan, L., Gwekwerere, Y., & Danyluk, P. (2014). Exploring the Issue of Failure to Fail in Professional Education Programs: A Multidisciplinary Study. *Journal of University Teaching and Learning Practice*, 11(2). <https://doi.org/10.53761/1.11.2.3>
- Mak-van Der Vossen, M. (2019). 'Failure to fail': The teacher's dilemma revisited. *Medical Education*, 53(2), 108-110. <https://doi.org/10.1111/medu.13772>
- Mak-van Der Vossen, M., Teherani, A., Van Mook, W. N. K. A., Croiset, G., & Kusurkar, R. A. (2018). Investigating US medical students' motivation to respond to lapses in professionalism. *Medical Education*, 52(8), 838-850. <https://doi.org/10.1111/medu.13617>
- McOwen, K. S., Varpio, L., & Konopasky, A. W. (2023). How to ... use theory as method in HPE research. *The Clinical Teacher*, 20(6), e13615. <https://doi.org/10.1111/tct.13615>
- Olmos-Vega, F. M., Stalmeijer, R. E., Varpio, L., & Kahlke, R. (2023). A practical guide to reflexivity in qualitative research: AMEE Guide No. 149. *Medical Teacher*, 45(3), 241-251. <https://doi.org/10.1080/0142159X.2022.2057287>
- Ong, Y. T., Kow, C. S., Teo, Y. H., Tan, L. H. E., Abdurrahman, A. B. H. M., Quek, N. W. S., Prakash, K., Cheong, C. W. S., Tan, X. H., Lim, W. Q., Wu, J., Tan, L. H. S., Tay, K. T., Chin, A., Toh, Y. P., Mason, S., & Radha Krishna, L. K. (2020). Nurturing professionalism in medical schools. A systematic scoping review

- of training curricula between 1990–2019. *Medical Teacher*, 42(6), 636-649.
<https://doi.org/10.1080/0142159X.2020.1724921>
- Rekman, J., Gofton, W., Dudek, N., Gofton, T., & Hamstra, S. J. (2016).
Entrustability Scales: Outlining Their Usefulness for Competency-Based
Clinical Assessment. *Academic Medicine*, 91(2), 186-190.
<https://doi.org/10.1097/ACM.0000000000001045>
- Rutkowski, K. (2007). Failure to fail: Assessing nursing students' competence
during practice placements. *Nursing Standard*, 22(13), 35-40.
<https://doi.org/10.7748/ns2007.12.22.13.35.c6299>
- Sanfey, H. (2011). Service or Education: In the Eye of the Beholder. *Archives of
Surgery*, 146(12), 1389. <https://doi.org/10.1001/archsurg.2011.292>
- Scarff, C. E., Bearman, M., Chiavaroli, N., & Trumble, S. (2019). Keeping mum in
clinical supervision: Private thoughts and public judgements. *Medical
Education*, 53(2), 133-142. <https://doi.org/10.1111/medu.13728>
- Shweiki, E., Beekley, A., Jenoff, J., Koenig, G., Kaulback, K., Lindenbaum, G., Patel,
P., Cohen, M., Weinstein, M., Zubair, M., Rosen, M., & Martin, N. (2015).
Applying Expectancy Theory to residency training: Proposing opportunities
to understand resident motivation and enhance residency training.
Advances in Medical Education and Practice, 339.
<https://doi.org/10.2147/AMEP.S76587>
- Tiana, A. (2008). Should 'failing' students repeat a grade/the year?: Answer by
Alejandro Tiana (Professor of History of Education and Comparative
Education at the Spanish Distance Teaching University and former
Secretary General for Education in the Ministry of Education and Science of

Spain). *Journal of Educational Change*, 9(4), 391-394.

<https://doi.org/10.1007/s10833-008-9088-4>

Van Maele, L., Devos, C., Guisset, S., Leconte, S., & Macq, J. (2021). *Expectancy-Value-Cost motivational theory to explore final year medical students' research intentions and past research experience: A multicentre cross-sectional questionnaire study.*

<https://doi.org/10.1101/2021.01.04.20244038>

Varpio, L., Ajjawi, R., Monrouxe, L. V., O'Brien, B. C., & Rees, C. E. (2017). Shedding the cobra effect: Problematising thematic emergence, triangulation, saturation and member checking. *Medical Education*, 51(1), 40-50.

<https://doi.org/10.1111/medu.13124>

Vleuten, C., Lindemann, I., & Schmidt, L. (2018). Programmatic assessment: The process, rationale and evidence for modern evaluation approaches in medical education. *Medical Journal of Australia*, 209(9), 386-388.

<https://doi.org/10.5694/mja17.00926>

Wigfield, A., & Eccles, J. S. (2000). Expectancy-Value Theory of Achievement Motivation. *Contemporary Educational Psychology*, 25(1), 68-81.

<https://doi.org/10.1006/ceps.1999.1015>

Yepes-Rios, M., Dudek, N., Duboyce, R., Curtis, J., Allard, R. J., & Varpio, L. (2016). The failure to fail underperforming trainees in health professions education: A BEME systematic review: BEME Guide No. 42. *Medical Teacher*, 38(11), 1092-1099. <https://doi.org/10.1080/0142159X.2016.1215414>