



Categoría: Education, Teaching, Learning and Assessment

ORIGINAL

Performance metrics of item developers for the assessment of learning outcomes

Métricas de desempeño de elaboradores de ítems para la evaluación de resultados de aprendizaje

Carlos Andrés Bravo Erazo¹  , Juan José Vizcaíno Figueroa¹  , Patty Janeth Guarnizo Cumbicus²  

¹Universidad Técnica de Cotopaxi. Latacunga, Ecuador.

²Instituto Tecnológico Superior Universitario ITI. Quito, Ecuador.

Citar como: Bravo Erazo CA, Vizcaíno Figueroa JJ, Guarnizo Cumbicus PJ. Performance metrics of item developers for the assessment of learning outcomes. Salud, Ciencia y Tecnología - Serie de Conferencias. 2024; 3:1006. <https://doi.org/10.56294/sctconf20241006>

Enviado: 16-02-2024

Revisado: 27-04-2024

Aceptado: 08-07-2024

Publicado: 09-07-2024

Editor: Dr. William Castillo-González 

ABSTRACT

The evaluation of learning results is a strategy that generates inputs for curricular feedback and continuous improvement of the academic management of vocational training programs. On the one hand, it contributes to determining the level of achievement of the graduation profile that students reach according to the progress of the study plan and, on the other, it allows identifying the level of performance of the teachers who participate during the design and application processes of the evaluation instruments. This article shows the application of three performance indicators: efficiency, quality and estimation, to identify the potential of teachers according to the role they must assume in an evaluation process. The study was carried out based on the learning results evaluation methodology developed by the Technical University of Cotopaxi in which 224 teachers participated, of which 143 met acceptable parameters of efficiency, quality and estimation, allowing 67 to be categorized as developers of items, 56 as technical reviewers and 20 as technical heads.

Keywords: Teaching Performance; Learning Results; Preparation of Exams; Continuous Improvement.

RESUMEN

La evaluación de resultados de aprendizaje es una estrategia que genera insumos para la retroalimentación curricular y mejora continua de la gestión académica de los programas de formación profesional. Por un lado, aporta a determinar el nivel de logro del perfil de egreso que los estudiantes alcanzan según el avance del plan de estudios y, por otro, permite identificar el nivel de desempeño de los docentes que participan durante los procesos de diseño y aplicación de los instrumentos de evaluación. Este trabajo muestra la aplicación de tres indicadores de desempeño: eficiencia, calidad y estimación, para identificar las potencialidades de los docentes según el rol que deben asumir en un proceso de evaluación. El estudio se realizó con base en la metodología de evaluación de resultados de aprendizaje desarrollada por la Universidad Técnica de Cotopaxi en la que participaron 224 docentes, de los cuales 143 cumplieron con parámetros aceptables de eficiencia, calidad y estimación, permitiendo categorizar a 67 como elaboradores de ítems, 56 como revisores técnicos y 20 como jefes técnicos.

Palabras clave: Desempeño Docente; Resultados de Aprendizaje; Elaboración de Exámenes; Mejora Continua.

INTRODUCCIÓN

La evaluación de resultados de aprendizaje en la educación superior se consolida como una estrategia que permite indagar sobre si los estudiantes conocen, comprenden y son capaces de hacer lo descrito en los perfiles de egreso de los planes de estudio bajo los cuales se forman profesionalmente.^(2,10,7) Al mismo tiempo, es un reto para las instituciones de educación superior, puesto que su aplicación demanda el diseño de metodologías e instrumentos que deben ser conocidos por los grupos de interés universitarios para que se obtengan resultados que favorezcan la toma de decisiones académicas.

Son varias las iniciativas que se han impulsado en los países para promover la evaluación de los resultados de aprendizaje, sin que tengan consensos en cómo y cuándo hacerlas. Por un lado, las agencias de acreditación han incorporado entre sus estándares o indicadores, referencias a que la calidad en la educación superior está directamente relacionada con el logro de los resultados de aprendizaje, infiriendo que si las condiciones institucionales, las estrategias pedagógicas, la pertinencia de la oferta académica y del modelo educativo, los planes estratégicos, el desarrollo de la investigación y otras estratégicas académicas funcionan bien, se garantiza la formación profesional de calidad.

Asimismo, algunas agencias profundizan el enfoque y no solo se centran en la importancia de alcanzar los resultados de aprendizaje sumativos, sino que establecen directrices para que se incorporen mediciones parciales en diferentes momentos de la formación, con el objetivo de que las instituciones de educación superior vayan incorporando mejoras académicas y pedagógicas de manera oportuna. En complemento a la evaluación formativa y sumativa de resultados de aprendizaje, las agencias también solicitan evaluaciones externas y comparativas.

En este contexto, la evaluación de resultados de aprendizaje se convierte en un escenario complejo y retador, puesto que constituye una oportunidad para promover la internacionalización del currículo y la homologación de competencias básicas (genéricas y específicas) que los profesionales deberían desarrollar en un mundo globalizado. No obstante, una excesiva estandarización, supone un condicionamiento al aprendizaje esperado basado en el enfoque conductista. Por ello, algunos autores prefieren hablar de un proceso evaluativo integral, que no solo tenga en cuenta competencias para el trabajo sino “el desarrollo de dimensiones humanas de forma holística teniendo en cuenta la perspectiva estética y ética de las mismas”.⁽⁵⁾

Coadyuvando a lo dicho, Camacho et al.⁽⁴⁾ resumen las tendencias curriculares en tres enfoques, que deben ser considerados cuando se habla de evaluación de resultados de aprendizaje. El enfoque técnico instrumental al tener énfasis en la orientación y el control, pretende demostrar que las competencias alcanzadas están en sintonía con el sector productivo. La evaluación curricular bajo este enfoque se realiza con normas y pruebas estandarizadas. El enfoque práctico, como su nombre lo indica, pone énfasis en la formación que se lleva a la práctica, aspirando a un mejoramiento continuo del propio proceso formativo y no solo del resultado alcanzado. La evaluación bajo este enfoque requiere la participación de todos los grupos de interés universitarios para contrastar el currículo escrito (documento aprobado), el currículo real (el que la institución defiende) y el currículo que se da en la acción.

El último enfoque que los autores citados plantean es el enfoque crítico, que apuesta por la autorreflexión como herramienta para la toma de decisiones que generen cambio en los procesos de enseñanza-aprendizaje orientados a la construcción de sociedades más justas y equitativas. La evaluación bajo este enfoque debe ser continua y permanente, evitando el sesgo único de la eficiencia sino realizada con una visión de interconexión entre todos los aspectos que influyen en el diseño e implementación del currículo. Paricio et al.⁽⁸⁾ señala que los resultados de aprendizaje deben evitar caer en el reduccionismo, sino más bien ser concretos y evaluables de algún modo, porque los efectos que justifican una asignatura o plan de estudios deben ser comprobados de alguna manera por lo que deben ser expresados de forma tal que sean observables, evaluables y alcanzables.

Precisamente, Bravo et al.⁽³⁾ proponen una metodología para evaluar los resultados de aprendizaje procurando mantener un equilibrio de los enfoques curriculares. Esta metodología tiene seis etapas operativas y se desarrolla con la construcción participativa y colectiva del instrumento de evaluación con énfasis en la autorreflexión profunda del perfil de egreso. La primera etapa es la definición de la estructura de evaluación, con base en cuatro niveles jerárquicos de desagregación: componentes, subcomponentes, temas y definiciones operacionales, que se obtienen del análisis del perfil de egreso y del plan de estudios.

En la segunda etapa se realizan capacitaciones grupales e individuales y se asignan las definiciones operacionales a los profesores seleccionados del programa o carrera para que inicien con la tercera etapa que es la de elaboración de preguntas, revisión técnica, validación de contenido y revisión de estilo. Posteriormente, en la cuarta etapa se diseña, valida y analiza cualitativamente el instrumento a aplicarse, para luego en la quinta etapa aplicarlo y finalmente en la sexta etapa analizar cuantitativamente los resultados y generar informes.

Esta metodología, que ha sido descrita de forma muy somera, está desarrollada pensando en los procesos, estimulando la reflexión antes que al análisis simplista de los resultados que se puedan obtener de la aplicación del instrumento para etiquetar a los evaluados. En tal sentido, cada etapa es procesual, “es decir, implicará el

seguimiento sistemático de los indicios que al ser interpretados a la luz de la teoría que subyace a la práctica evaluativa, arrojará una valoración integral de la persona evaluada, del proceso seguido y del evaluador”⁽⁵⁾ esperando como beneficio colateral el integrar el concepto de aprender a aprender como opción del cambio y evolución de las instituciones de educación superior para enfrentar el futuro.⁽⁹⁾

Justamente este trabajo tiene como objetivo analizar métricas de desempeño de los profesores que participaron durante los procesos de evaluación de resultados de aprendizaje, tomando en cuenta los reportes que genera la plataforma de evaluación que la Universidad Técnica de Cotopaxi (UTC) desarrolló y que la utiliza desde hace cuatro períodos académicos, generando un insumo para la metaevaluación de la Auditoría de Resultados de Aprendizaje, que se realiza en el marco del Sistema de Gestión del Aseguramiento de la Calidad que la UTC tiene implementado. Es importante señalar, además, que esta evaluación del desempeño de los profesores podría adscribirse, en el corto plazo, a lo recomendado por Kamandhari et al.⁽⁶⁾ en lo que respecta a utilizar múltiples evaluaciones indirectas del proceso de enseñanza para mejorar las experiencias de aprendizaje.

MÉTODO

La metodología y herramientas utilizadas por la UTC para la evaluación de resultados de aprendizaje permite generar un proceso con múltiples espacios de reflexión en los que participan los docentes con rol de elaboradores de ítems, generando información que se convierte en variables para calcular tres indicadores de desempeño:

- Eficiencia, en relación al aprovechamiento óptimo del tiempo durante la elaboración del ítem.
- Calidad, en relación a parámetros psicométricos óptimos de los ítems producidos por el elaborador.
- Estimación, de la dificultad de un ítem a partir del nivel de complejidad propuesto durante la elaboración de la pregunta.

El indicador de eficiencia se construye con tres variables: el número de preguntas que le son asignadas a cada elaborador, el número de preguntas asignadas que son aprobadas para ser incluidas en el instrumento de evaluación y un factor de penalización que se genera a partir de las preguntas asignadas que no fueron elaboradas. Para que una pregunta sea aprobada debe cumplir tres condiciones técnicas:

1. Ser entregada para revisión técnica en el tiempo establecido. Si una pregunta no es entregada a tiempo quedará en estado de `pregunta no elaborada`.
2. Cumplir los parámetros técnicos: enunciado, opciones de respuesta, nivel de complejidad, bibliografía y argumento, cuyo detalle se entrega a los elaboradores de preguntas mediante una ficha de orientación. Un revisor técnico analiza cada pregunta que es entregada por el elaborador y decide una de las siguientes alternativas: en caso de cumplir los parámetros técnicos la pregunta pasa a la etapa de validación conceptual y en caso de no cumplir es devuelta al elaborador para que realice ajustes a la pregunta, dándole nuevamente un tiempo máximo para hacerlo. Este proceso se repite las veces que sean necesarias matizando la retroalimentación y capacitación constante que le permitan al elaborador mejorar sus preguntas. Si el ajuste de la pregunta no se realiza en el tiempo establecido pasa a estado de `pregunta no elaborada`.
3. Ser validada conceptualmente por los validadores o grupo colegiado de especialistas, quienes darán el dictamen final para que la pregunta sea aprobada por ser conceptualmente correcta, elaborada a partir de bibliografía actualizada y cumplir con los criterios de pertinencia, objetividad, confiabilidad y validez en relación al grupo objetivo de estudiantes a ser evaluados. Si una pregunta no aprueba la validación conceptual se descarta y pasa a estado de `pregunta no elaborada`.

Este indicador es importante porque cada pregunta en estado de `pregunta no elaborada` representa un gasto de recursos significativos, considerando que a cada docente por asumir el rol de elaborador de ítems se le asignan horas en su carga horaria semanal y se destina un técnico para su acompañamiento.

La fórmula con la que se calcula la eficiencia es la siguiente:

$$\text{Eficiencia} = \frac{\text{Número de preguntas aprobadas}}{\text{Número de preguntas asignadas} + \text{Factor de penalización}}$$

Factor de penalización = (0,75)* (Preguntas no elaboradas).

Una vez calculados los valores de eficiencia se califica a cada docente según la escala que se describe a continuación:

Tabla 1. Dictamen de eficiencia de los elaboradores de ítems

Dictamen Eficiencia	Indicador de Eficiencia
Alta	0,71 - 1,00
Media	0,41 - 0,70
Baja	0,00 - 0,40

Fuente: plataforma E.R.A-UTC, 2022

El indicador de calidad se fundamenta en los parámetros psicométricos obtenidos de la calibración de las preguntas elaboradas, utilizando los criterios de dificultad tabla 2 y discriminación tabla 3 de la Teoría Clásica del Test (TCT), entendiendo como dificultad de una pregunta la proporción de personas que responden correctamente un reactivo de una prueba, siendo de menor dificultad mientras mayor sea esta proporción, mientras que la discriminación es la expresión numérica de la medida en que una pregunta separa a los examinados de más alto rendimiento de los de más bajo rendimiento.⁽³⁾

Tabla 2. Catalogación de la dificultad de un ítem

Nivel Dificultad TCT	Rango Dificultad TCT	Rango Dictamen Ítem	Dictamen Ítem
Muy fácil	0,90 - 1,00	0,90 - 1,00	LS(Descartar)
Fácil	0,60 - 0,89	0,80 - 0,89	Útil
Moderado	0,40 - 0,59	0,20 - 0,79	Óptimo
Difícil	0,10 - 0,39	0,10 - 0,19	Útil
Muy difícil	0,00 - 0,09	0,00 - 0,09	LI(Descartar)

Fuente: plataforma E.R.A-UTC, 2022

Tabla 3. Catalogación de la discriminación de un ítem

Nivel Discriminación TCT	Rango Discriminación TCT	Rango Dictamen Ítem	Dictamen Ítem
Discrimina muy bien	0,40 - 1,00	0,30 - 1,00	Óptimo
Discrimina bien	0,30 - 0,39		
Discriminación baja	0,20 - 0,29	0,00 - 0,29	Útil
No discrimina	0,00 - 0,19		
Discriminación opuesta	Menor a 0,00	Menor a 0,00	LI(Descartar)

Fuente: plataforma E.R.A-UTC, 2022

Las preguntas catalogadas con criterio de descarte en los parámetros de dificultad o discriminación son excluidas por no alcanzar la calidad psicométrica que debe tener una pregunta. Con base en los parámetros de calibración obtenidos de cada pregunta aprobada se aplica la siguiente fórmula para calcular el indicador de calidad:

$$Calidad = \left(\frac{DO + DU}{TP} \right) * \left[\frac{(0,25 * Nd) + (0,50 * db) + (0,75 * dB) + dMB}{TP} \right]$$

TP = Total preguntas calibradas.

DO = Número de preguntas con dictamen de dificultad óptima.

DU = Número de preguntas con dictamen de dificultad útil.

Nd = Número de preguntas que no discriminan.

db = Número de preguntas con discriminación baja.

dB = Número de preguntas que discriminan bien.

dMB = Número de preguntas que discriminan muy bien.

Una vez calculados los valores de calidad se califica a cada docente según la escala que se describe a continuación:

Tabla 4. Dictamen de calidad de los elaboradores de ítems

Dictamen Calidad	Indicador de Calidad
Alta	0,51 - 1,00
Media	0,31 - 0,50
Baja	0,00 - 0,30

Fuente: plataforma E.R.A-UTC, 2022

El indicador de estimación está estructurado con dos variables: la sumatoria del factor de concordancia de las preguntas elaboradas por el elaborador de ítems y el total de preguntas del elaborador que fueron aplicadas y cuentan con parámetros psicométricos. El factor de concordancia se genera a partir de una relación simple entre la complejidad estimada por el elaborador de ítems para la pregunta y el nivel de dificultad tabla 5 que obtuvo la pregunta en la calibración.

Tabla 5. Relación simple entre dificultad y complejidad de un ítem

Dificultad TCT	Rango TCT	Complejidad
Muy fácil	0,90 - 1,00	Baja
Fácil	0,60 - 0,89	
Moderado	0,40 - 0,59	Media
Difícil	0,10 - 0,39	Alta
Muy difícil	0,00 - 0,09	

Se debe tener en cuenta no confundir la complejidad con el nivel de dificultad, puesto que “la complejidad es un atributo del rasgo latente que se establece a priori, por juicio de expertos, mientras que la dificultad (o el grado de dificultad) es un parámetro que se calcula a posteriori en función de las respuestas de los sujetos ante un estímulo o ítem dado”

Durante la elaboración de la pregunta el elaborador de ítems estima en tres niveles la complejidad que tendrá la pregunta: alto, medio o bajo, conforme los criterios establecidos en la ficha de orientación. En la etapa de análisis cualitativo del instrumento, los expertos confirman o ajustan la complejidad propuesta de cada pregunta para que puedan ser aplicadas y obtener el parámetro de dificultad.

Una buena estimación del elaborador de preguntas se produce cuando la complejidad se relaciona directamente con su dificultad, es decir, si el elaborador establece que la pregunta tiene una complejidad baja en la aplicación la pregunta debe resultar fácil o muy fácil para la población objetivo de estudiantes, si el elaborador dice que la pregunta tiene una complejidad media, en la aplicación la pregunta debe resultar en el nivel moderado y si el elaborador dice que la pregunta tiene una complejidad alta en la aplicación la pregunta debe resultar difícil o muy difícil. Con base en este criterio se genera el factor de concordancia tabla 6 de cada pregunta, siendo 1 la concordancia relacionalmente directa, 0,5 si la relación es media y cero si la relación es nula.

Tabla 6. Factor de concordancia por pregunta

Dificultad / Complejidad	Baja	Media	Alta
Muy fácil	1,0	0,5	0,0
Fácil	1,0	0,5	0,0
Moderado	0,5	1,0	0,5
Difícil	0,0	0,5	1,0
Muy difícil	0,0	0,5	1,0

Fuente: plataforma E.R.A-UTC, 2022

Con base en los criterios descritos se calcula el indicador de estimación mediante la siguiente fórmula que se aplica a cada elaborador de ítems.

$$\text{Estimación} = \frac{\sum \text{Factor de concordancia de las preguntas}}{\text{Total preguntas}}$$

Una vez calculados los valores de estimación se califica a cada docente según la escala que se describe a continuación:

Tabla 7. Dictamen de estimación de los elaboradores de ítems

Dictamen Estimación	Indicador de Estimación
Alta	0,71 - 1,00
Media	0,21 - 0,70
Baja	0,00 - 0,20

Fuente: plataforma E.R.A-UTC, 2022

DISCUSIÓN Y CONCLUSIÓN

Una de las características de la metodología de evaluación de resultados de aprendizaje aplicada por la UTC es que cada contacto con los docentes se convierte en un momento de capacitación. Ya sea de forma individual o grupal o encuentros de inducción o socialización, siempre se intercambian experiencias y se aprende por la propia interacción que se desarrolla a lo largo del trabajo. En total se reconocieron 1 411 momentos de capacitación tabla 8 durante las seis etapas, en las que participaron 224 docentes de 16 programas de grado.

Tabla 8. Momentos de capacitación

Etapas	Inducción	Socialización	Grupal	Individual
1	1	1	41	28
2	6	-	49	-
3	-	-	-	1 227
4	-	-	20	-
5	-	2	-	-
6	-	-	36	-
Total	7	3	146	1 255

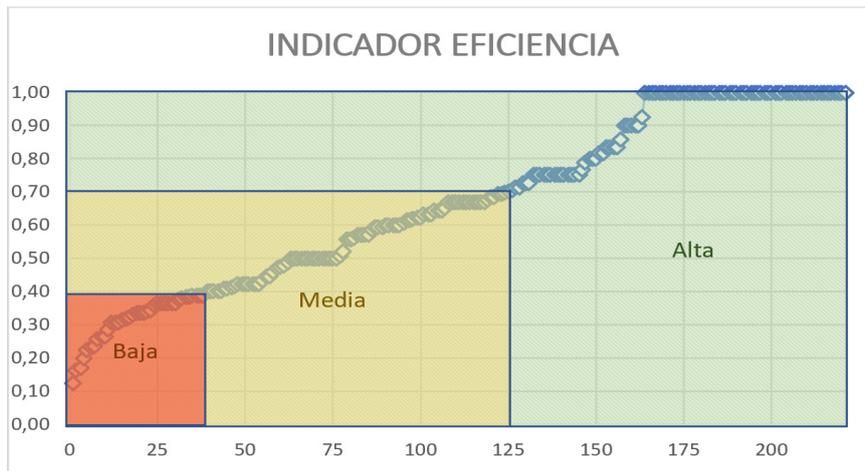
Si bien todos los momentos de capacitación son importantes para el proceso, durante la tercera etapa es cuando se generan los datos para el cálculo de los indicadores de desempeño, los mismos que se gestionan a través de la plataforma de evaluación. El indicador de eficiencia es el único que se puede calcular antes de la aplicación de la evaluación puesto que no requiere parámetros psicométricos. En la tabla 9 se presentan los resultados de la eficiencia a nivel de programa evaluado, lo que evidencia la heterogeneidad del comportamiento de los docentes en el cumplimiento de los tiempos establecidos para entregar las preguntas que les asignaron.

Tabla 9. Datos para el cálculo del indicador de eficiencia

Programa	PAS	PAP	PNE	Eficiencia	Dictamen
Administración de Empresas	212	104	108	0,35	Baja
Agroindustria	66	48	18	0,60	Media
Agronomía	117	78	39	0,53	Media
Comunicación	71	37	34	0,38	Baja
Contabilidad y Auditoría	158	108	50	0,55	Media
Diseño Gráfico	125	67	58	0,40	Media
Educación Básica	141	96	45	0,55	Media
Educación Inicial	193	124	69	0,51	Media
Electricidad	62	46	16	0,62	Media
Electromecánica	126	70	56	0,42	Media
Ingeniería Ambiental	133	94	39	0,58	Media
Ingeniería Industrial	206	94	112	0,32	Baja
Medicina Veterinaria	65	36	29	0,41	Media
Pedagogía de los Idiomas Nacionales y Extranjeros	100	90	10	0,84	Alta
Sistemas de Información	146	93	53	0,50	Media
Turismo	56	42	14	0,63	Media
Total	1977	1227	750	0,48	Media

PAS: preguntas asignadas; PAP: preguntas aprobadas; PNE: preguntas no elaboradas

Una vez calculado el indicador de eficiencia para cada docente se realizó una distribución de dichos valores (Ver Figura 1) para determinar el nivel de desempeño.



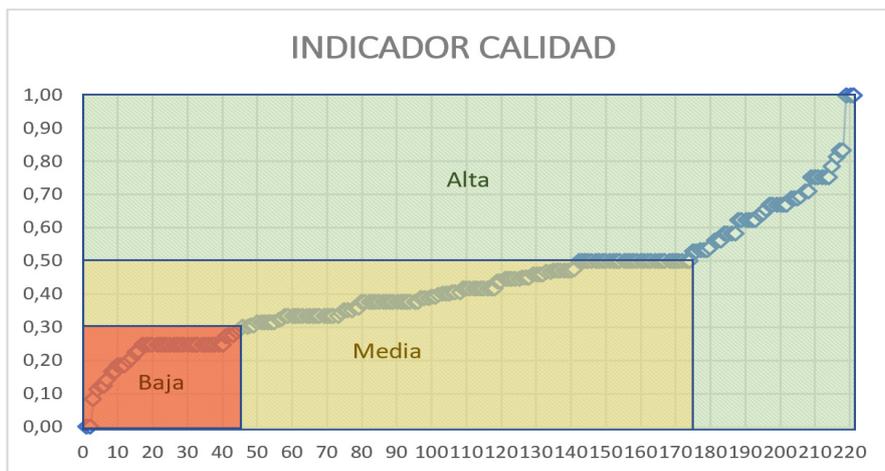
Fuente: plataforma E.R.A-UTC, 2022
 Figura 1. Distribución del indicador de eficiencia

Con base en el indicador de eficiencia de cada docente se pudo evidenciar que el impacto que genera una concentración de docentes con bajo desempeño en el indicador general del programa es mayor al de los programas que tienen docentes con alto desempeño, considerando que de acuerdo a los resultados presentados en la tabla 10 prevalece un desempeño alto entre los docentes de la UTC. Esta situación puede deberse a lo que Wambua et al.⁽¹²⁾ denominan ansiedad por exámenes, que se presenta principalmente cuando se utilizan medios tecnológicos como computadoras para las pruebas, tal como sucede con la plataforma que los docentes debieron utilizar para elaborar preguntas.

Tabla 10. Nivel de desempeño de los docentes por indicador de eficiencia

Eficiencia	Elaboradores
Baja	38
Media	87
Alta	96
Total	221*

*3 docentes fueron excluidos de la valoración del desempeño porque en la mitad del proceso se retiraron para atender otras funciones asignadas por las autoridades universitarias



Fuente: plataforma E.R.A-UTC, 2022
 Figura 2. Distribución del indicador de calidad

Para calcular los indicadores de calidad y estimación se descartaron 52 preguntas por tener errores conceptuales y epistemológicos detectados en las validaciones realizadas con expertos, quedando 1 175 preguntas a las que se les aplicaron los criterios de las tablas 2 y 3, determinando que 57 de ellas tenían parámetros psicométricos en los límites de descarte por TCT, procediendo nuevamente a validarlas conceptualmente por expertos, concluyendo que todas ellas tenían errores en su redacción y preceptos epistemológicos que causaron confusión en los estudiantes y por lo tanto fueron descartadas definitivamente. En la tabla 11 se presentan los datos obtenidos para calcular el indicador de calidad, que luego fueron distribuidos para determinar el nivel de desempeño figura 2, cuyos resultados se presentan en la tabla 12.

Programa	TP	DO	DU	Nd	db	dB	MB
Administración de Empresas	94	19	71	56	16	6	6
Agroindustria	47	11	30	21	10	7	3
Agronomía	70	20	48	22	11	8	16
Comunicación	30	9	16	9	6	4	4
Contabilidad y Auditoría	108	36	61	36	21	14	20
Diseño Gráfico	65	19	43	20	14	14	11
Educación Básica	96	27	64	41	12	15	14
Educación Inicial	120	36	70	57	22	19	5
Electricidad	45	6	33	25	9	5	3
Electromecánica	69	27	38	16	5	4	27
Ingeniería Ambiental	91	26	61	50	16	14	1
Ingeniería Industrial	91	34	54	40	28	8	4
Medicina Veterinaria	35	13	20	20	5	9	1
Pedagogía de los Idiomas Nacionales y Extranjeros	89	21	63	39		20	16
Sistemas de Información	86	32	51	42	20	8	10
Turismo	39	14	21	14	4	12	6
Total	1175	350	744	508	199	167	147

Calidad	Elaboradores
Baja	45
Media	129
Alta	47
Total	221

Para el cálculo del indicador de estimación se procedió a determinar la relación simple de cada una de las 1 175 preguntas validadas entre el nivel de dificultad planteado por el docente y la complejidad obtenida luego de la aplicación y con base en esa relación se estableció el factor de concordancia, los mismos que se sumaron y dividieron para el número de preguntas de cada docente elaborador. Posteriormente, se realizó la distribución de los factores de concordancia (Ver Figura 3) para establecer el nivel de desempeño en el indicador de estimación tabla 13.

Estimación	Elaboradores
Baja	7
Media	145
Alta	69
Total	221



Fuente: plataforma E.R.A-UTC, 2022

Figura 3. Distribución del indicador de estimación

Las métricas de desempeño permitieron iniciar con el proceso de institucionalización de la evaluación de resultados de aprendizaje en la UTC, puesto que se pudo determinar cuáles docentes tenían mejores aptitudes para asumir los diferentes roles: jefe técnico, revisor técnico y elaboradores. Asimismo, se evidenció lo propuesto por Álvarez y López (2018) al referirse a que para que la educación superior se centre en resultados de aprendizaje es necesario cambiar los modelos y estrategias que emplean los docentes incorporando tareas diferentes a la impartición de clases para el aprendizaje del alumnado, como orientación y tutorización, que también son necesarias para elaborar instrumentos de evaluación contextualizados y fiables.

De esta manera se pudo identificar a 143 docentes de los 221 que participaron, que cumplieron con niveles aceptables de desempeño en eficiencia, calidad y estimación, que fueron categorizados 67 como elaboradores de ítems, 56 como revisores técnicos y 20 como jefes técnicos. Con este equipo de trabajo se espera fortalecer el enfoque de cambio del aprendizaje centrado en contenidos a uno orientado al logro del resultado de ese aprendizaje, de los tecnicismos teóricos a los prácticos con ejemplos contextualizados y de los resultados de la evaluación punitivos y numéricos a los orientados a la mejora de los aprendizajes.

REFERENCIAS BIBLIOGRÁFICAS

1. Álvarez-Pérez PR, López-Aguilar D. Competencias genéricas y resultados de aprendizaje en los estudios de grado de Pedagogía. *Roja U* [Internet]. 2018 [consultado el 20 de junio del 2024];16(1):137. Disponible en: <https://polipapers.upv.es/index.php/REDU/article/view/8895>
2. Pérez Gamboa AJ, Díaz-Guerra DD. Artificial Intelligence for the development of qualitative studies. *LatIA*. 2023;1:4.
3. Educa.co. [consultado el 20 de junio de 2024]. Disponible en: https://repositorio.itm.edu.co/bitstream/handle/20.500.12622/6235/Buenas_Pra%CC%81cticas_2023.pdf?sequence=5&isAllowed=y#page=64
4. Edu.co. [cited 2024 Jun 20]. Available from: https://ww2.ufps.edu.co/public/archivos/oferta_academica/a00ca7dc282007f8c767802321fc3761.pdf
5. Cano CAG, Troya ALC. Artificial Intelligence applied to teaching and learning processes. *LatIA* 2023;1:2-2. <https://doi.org/10.62486/latia20232>.
6. Educa.co. [consultado el 21 de junio de 2024]. Disponible en: <https://repositorio.itm.edu.co/handle/20.500.12622/5781?locale-attribute=en>
7. Kamandhari HH, Lavandera Ponce S. Guidelines for the alignment of indirect measures of teaching performance: Triangular perspectives of students, peer faculty, and external reviewers. *Int J Assess Eval* [Internet]. 2021;28(2):91-117. Available from: <https://www.proquest.com/openview/5b5952766b23906f962e4d0a692832a9/1?pq-origsite=gscholar&cbl=5528231>
8. Gómez Cano CA, Colala Troya AL. Artificial Intelligence applied to teaching and learning processes. *LatIA*. 2023;1:2.

9. Escuelasdearte.es. [consultado el 21 de junio de 2024]. Disponible en: <http://www.escuelasdearte.es/news/2012almeria/stuff/guiasdocentes.pdf>

10. Mario Luis P, Alejandra Elena M, Néstor Horacio B. El desafío del aprendizaje en las organizaciones: La necesidad de aprender a aprender para enfrentar el futuro [Internet]. Cgscholar.com. [consultado el 21 de junio de 2024]. Disponible en: https://cgscholar.com/bookstore/works/el-desafio-del-aprendizaje-en-las-organizaciones?category_id=cgrn-es

11. Gamboa AJP, Díaz-Guerra DD. Artificial Intelligence for the development of qualitative studies. LatIA 2023;1:4-4. <https://doi.org/10.62486/latia20234>.

12. UNESCO. Declaración de Incheon. Educación 2030: Hacia una educación inclusiva y equitativa de calidad y un aprendizaje a lo largo de la vida para todos. 2015 [cited 2024 Jun 21]; Available from: <https://bibliotecadigital.mineduc.cl/handle/20.500.12365/18066>

13. Vizcaíno JJ, Rojas N, Cisneros D. LA EVALUACIÓN DEL DESEMPEÑO DOCENTE. ENFOQUES Y MODELOS. Prospectivas UTC “Revista de Ciencias Administrativas y Económicas” [Internet]. 2022 [cited 2024 Jun 21];13-21. Available from: <http://investigacion.utc.edu.ec/index.php/prospectivasutc/article/view/475>

14. Wambua R, Mwaura P, Dinga J. Psychometric properties of a test anxiety scale for use in computer-based testing in Kenya. Int J Assess Eval [Internet]. 2023;31(1):1-18. Available from: <https://www.proquest.com/openview/46948fd44ccb3871e02dd4b86bc8a23b/1?pq-origsite=gscholar&cbl=5528231>

FINANCIACIÓN

Ninguna.

CONFLICTO DE INTERÉS

Ninguno.

CONTRIBUCIÓN DE AUTORÍA

Conceptualización: Carlos Andrés Bravo Erazo, Juan José Vizcaíno Figueroa, Patty Janeth Guarnizo Cumbicus.

Curación de datos: Carlos Andrés Bravo Erazo, Juan José Vizcaíno Figueroa, Patty Janeth Guarnizo Cumbicus.

Investigación: Carlos Andrés Bravo Erazo, Juan José Vizcaíno Figueroa, Patty Janeth Guarnizo Cumbicus.

Administración del proyecto: Carlos Andrés Bravo Erazo, Juan José Vizcaíno Figueroa, Patty Janeth Guarnizo Cumbicus.

Recursos: Carlos Andrés Bravo Erazo, Juan José Vizcaíno Figueroa, Patty Janeth Guarnizo Cumbicus.

Supervisión: Carlos Andrés Bravo Erazo, Juan José Vizcaíno Figueroa, Patty Janeth Guarnizo Cumbicus.

Redacción - borrador original: Carlos Andrés Bravo Erazo, Juan José Vizcaíno Figueroa, Patty Janeth Guarnizo Cumbicus.