



Instructivo metodológico para la evaluación de resultados de aprendizaje en los programas de posgrado

SECRETARÍA TÉCNICA (ST)

DIRECCIÓN DE EXÁMENES (DGE)

DIRECCIÓN DE

INVESTIGACIÓN(DEI)

FEBRERO 2024



CONTENIDO

1.	<i>Antecedentes</i>	5
2.	<i>Base legal</i>	6
2.1	Constitución de la República del Ecuador	6
2.2	Ley Orgánica de Educación Superior	6
2.3	Normativa del CES sobre armonización de títulos	8
2.4	Reglamento interno del CACES	8
2.5	Reglamento de Evaluación externa con fines de acreditación para el aseguramiento de la calidad de los programas de posgrado de las universidades y escuelas politécnicas	9
2.6	Manual para el Diseño y Aplicación de Exámenes	10
3.	<i>Objetivo</i>	10
4.	<i>Glosario de términos</i>	10
5.	<i>Marco Teórico</i>	11
6.	<i>Metodología de evaluación de los resultados de aprendizaje (ERAP)</i>	15
6.1	Definir contenidos y estructura del examen	20
6.2	Elaboración de ítems o preguntas.....	22
6.3	Validación de preguntas.....	23
6.4	Pilotaje de preguntas o ítems	24
6.5	Diseño y ensamble del instrumento de evaluación	25
6.6	Aplicación del examen de evaluación de resultados del aprendizaje de programas de posgrado.....	31
6.6.1	Cohortes	32
6.6.2	Convocatoria para rendir el examen	34
6.6.3	Selección de Centro Autorizado para la Toma del Examen (CATE)	34
6.6.4	Confirmación de inscripción	34
6.6.5	Solicitud de actualización de información registrada en la inscripción	34
6.6.6	Aplicación del examen	35
6.6.7	Resultados definitivos del ERAP	35
6.6.8	Difusión y comunicación de resultados de aprendizaje	36
6.7	Análisis psicométrico.....	36
6.8	Mejora continua	37
6.9	Cronograma genérico para el desarrollo del ERAP	38
7.	<i>Bibliografía</i>	39

INDICE DE TABLAS

Tabla 1. Tamaños de muestra mínimos según la teoría de análisis.....	25
Tabla 2. Categorías de dictámenes según TCT.....	27
Tabla 3. Categorías de dictámenes según TRI.....	27
Tabla 4. Cronograma de construcción de la evaluación de resultados del aprendizaje.....	38

INDICE DE FIGURAS

Figura 1. Desarrollo del ERAP por campo de conocimiento	17
Figura 2. Diagrama de flujo del proceso ERAP	18
Figura 3. Diagrama de flujo de subprocesos del ERAP.....	19
Figura 4. Agrupación por campo amplio, específico y detallado	20
Figura 5. Ejemplo de curva característica de un ítem o pregunta	29
Figura 6. Dispersión esperada de resultados.....	31
Figura 7. Aplicación del ERAP	32

1. Antecedentes

El marco normativo ecuatoriano establece como una de las competencias específicas del Consejo de Aseguramiento de la Calidad de la Educación Superior (CACES) la evaluación y acreditación de las instituciones de educación superior, sus carreras y programas.

El “Reglamento de Evaluación Externa con fines de Acreditación para el Aseguramiento de la Calidad de los Programas de Posgrado de las Universidades y Escuelas Politécnicas” fue aprobado en abril de 2023, que en su artículo 5 establece que “(...) la evaluación externa de los programas de posgrado contempla dos procesos interdependientes y complementarios:

1. La evaluación del entorno de aprendizaje; y,
2. La evaluación de resultados de aprendizaje.”

Para la evaluación de los resultados de aprendizaje de los programas de posgrado (ERAP) se realizó una investigación interna titulada “Revisión y propuesta de mecanismos para la evaluación de los resultados de aprendizaje de los programas de posgrado de las Universidades y Escuelas Politécnicas”, ejecutada por la Dirección de Estudios e Investigación entre mayo y junio del 2023, la cual tuvo como finalidad sistematizar experiencias internacionales relativas a la evaluación de los resultados de aprendizaje de la educación superior.

Este insumo teórico y metodológico busca orientar la operativización de la evaluación mediante una propuesta para la evaluación de los resultados de aprendizaje de los Programas de Posgrado (ERAP) que, junto con la evaluación del entorno del aprendizaje, son los requisitos indispensables para la acreditación de los programas de posgrado.

2. Base legal

2.1 Constitución de la República del Ecuador

El artículo 353 de la Constitución de la República del Ecuador determina que: “El sistema de educación superior se regirá por: (...) 2. Un organismo público técnico de acreditación y aseguramiento de la calidad de instituciones, carreras y programas, que no podrá conformarse por representantes de las instituciones objeto de regulación.”

2.2 Ley Orgánica de Educación Superior

El artículo 93 de la LOES define al principio de calidad como: “(...) la búsqueda continua, auto-reflexiva del mejoramiento, aseguramiento y construcción colectiva de la cultura de la calidad educativa superior con la participación de todos los estamentos de las instituciones de educación superior y el Sistema de Educación Superior, basada en el equilibrio de la docencia, la investigación e innovación y la vinculación con la sociedad, orientadas por la pertinencia, la inclusión, la democratización del acceso y la equidad, la diversidad, la autonomía responsable, la integralidad, la democracia, la producción de conocimiento, el diálogo de saberes, y valores ciudadanos”.

El artículo 94 de la LOES que trata sobre el objeto del sistema interinstitucional de aseguramiento de la calidad determina que: “(...) tiene por objeto garantizar el efectivo cumplimiento del principio de calidad consagrado en la Constitución y en la presente ley, intervendrán como principales actores de este Sistema el Consejo de Educación Superior, el Consejo de Aseguramiento de la Calidad de la Educación Superior y las Instituciones de Educación Superior.

El artículo 95 de la LOES referente a los criterios y estándares para la acreditación determina que: “El Consejo de Aseguramiento de la Calidad de la Educación Superior establecerá modelos que incluirán criterios y estándares cuantitativos y cualitativos, que las instituciones de

educación superior, carreras y programas deberán alcanzar para ser acreditadas; entendiendo que el fin último es la calidad y no la acreditación (...);

El artículo 103 de la Ley Orgánica de Educación Superior establece que para la evaluación de resultados de aprendizaje de carreras y programas se establecerá: “(...) un examen u otros mecanismos de evaluación para estudiantes del último período académico. Los procesos de evaluación se realizarán sobre los conocimientos y de ser necesario según el perfil profesional se aplicará sobre otras competencias (...);

El artículo 171 de la LOES establece que el Consejo de Aseguramiento de la Calidad de la Educación Superior (CACES) es: “(...) el organismo público técnico, con personería jurídica y patrimonio propio, con independencia administrativa, financiera y operativa que tiene a su cargo la regulación, planificación y coordinación del sistema de aseguramiento de la calidad de la educación superior; tendrá facultad regulatoria y de gestión (...);

El artículo 173 de la LOES señala: “El Consejo de Aseguramiento de la Calidad de la Educación Superior normará la autoevaluación institucional, y ejecutará los procesos de evaluación externa, acreditación y apoyará el aseguramiento interno de la calidad de las instituciones de educación superior. Las instituciones de educación superior, tanto públicos como particulares, sus carreras y programas, deberán someterse en forma obligatoria a la evaluación externa y a la acreditación; además, deberán organizar los procesos que contribuyan al aseguramiento interno de la calidad.”;

El artículo 174 literales c) y m) de la LOES señala: “(...) c) Normar los procesos de evaluación integral, externa e interna, y de acreditación para la óptima implementación del Sistema de Aseguramiento de la Calidad conducente a obtener la Cualificación Académica de

Calidad Superior”; (...); y, “m) Ejecutar prioritariamente los procesos de evaluación, acreditación y clasificación académica de programas y carreras consideradas de interés público”.

2.3 Normativa del CES sobre armonización de títulos

Mediante Resolución RPC-SE-05-No.014-2023 el Consejo de Educación Superior resuelve emitir el “REGLAMENTO DE ARMONIZACIÓN DE LA NOMENCLATURA DE TÍTULOS PROFESIONALES Y GRADOS ACADÉMICOS QUE CONFIEREN LAS INSTITUCIONES DE EDUCACIÓN SUPERIOR DEL ECUADOR” mediante el cual se unifica y armoniza la nomenclatura de títulos profesionales, grados académicos y la codificación de los campos del conocimiento amplio, detallado y específico.

2.4 Reglamento interno del CACES

Mediante Resolución No. 011-SE-05-CACES-2019 de 11 de junio de 2019, el Pleno de este Organismo expidió el Reglamento Interno del CACES, que tiene por objeto regular el funcionamiento del Pleno del Consejo, los deberes y obligaciones de quienes lo integran, así como de las comisiones que sean conformadas para el desarrollo de las diferentes actividades propias de la institución.

El artículo 3 del citado Reglamento dispone: “El Pleno del Consejo es la máxima autoridad de deliberación y decisión del CACES”; Que el artículo 7 del Reglamento ibidem determina: “Son atribuciones y deberes del Pleno del Consejo, las siguientes: (...) e) Aprobar y modificar las normas, regulaciones, modelos y documentación técnica del CACES (...)”.

El artículo 8 del Reglamento ibidem determina: “El Pleno del Consejo aprobará sus actos administrativos y normativos por mayoría simple y en un solo debate, a través de resoluciones (...)”.

El artículo 27 de la norma ibidem prescribe: “Las comisiones del CACES serán permanentes u ocasionales y su funcionamiento se sujetará a lo dispuesto en el presente Reglamento. Son comisiones permanentes del CACES: (...) 4. Comisión de programas de posgrado (...)”.

El artículo 28 del Reglamento ibidem establece: “Son funciones generales de las comisiones permanentes y ocasionales del Consejo: (...) c) Proponer al Pleno del Consejo proyectos de actos administrativos, normativos y documentos técnicos, vinculados a su ámbito de acción (...)”.

2.5 Reglamento de Evaluación externa con fines de acreditación para el aseguramiento de la calidad de los programas de posgrado de las universidades y escuelas politécnicas

Mediante resolución Nro. 015-SE-07-CACES-2023 de 28 de abril de 2023, el Pleno del CACES resuelve expedir el “REGLAMENTO DE EVALUACIÓN EXTERNA CON FINES DE ACREDITACIÓN PARA EL ASEGURAMIENTO DE LA CALIDAD DE LOS PROGRAMAS DE POSGRADO DE LAS UNIVERSIDADES Y ESCUELAS POLITÉCNICAS En el numeral 6 del artículo 3 del reglamento mencionado en el párrafo anterior señala que “(...) La evaluación de los resultados de aprendizaje se establece a través de un examen u otro mecanismo de evaluación que se aplicará a los estudiantes del último período académico del programa evaluado. Este proceso se realizará sobre los conocimientos y competencias adquiridas durante el programa de posgrado.”

De la misma manera, el artículo 7 del reglamento mencionado establece que “La evaluación de los resultados de aprendizaje se establece a través de un examen u otro mecanismo de evaluación que se aplicará a los estudiantes del último período académico del programa evaluado. Este proceso se realizará sobre los conocimientos y competencias adquiridas durante el programa de posgrado.”

En la disposición transitoria segunda del mencionado reglamento se determina que “(...) la Secretaría Técnica deberá presentar la propuesta de Instructivo para la elaboración de los mecanismos que permitan la ejecución del ERAP a la Comisión Permanente de Programas de Posgrado (...).”

2.6 Manual para el Diseño y Aplicación de Exámenes

Mediante Resolución No. 030-SO-06-CACES-2022 de 22 de marzo de 2022, el Pleno del CACES resuelve aprobar el contenido del Manual para el Diseño y Aplicación de Exámenes, reformado mediante Resolución No. 031-SO-07-CACES-2023 de 21 de marzo de 2023.

3. Objetivo

Determinar la metodología de evaluación de resultados de aprendizaje de los programas de posgrado.

4. Glosario de términos

- **Rasgo:** En psicometría, se entiende como la habilidad o constructo que se pretende medir en una población mediante una evaluación (Muñiz J. , 2010).
- **Población objetivo:** Grupo poblacional que cumple con criterios de elegibilidad establecidos.
- **Punto de corte:** Nivel crítico absoluto o relativo a partir del cual se definen las categorías de los resultados de una evaluación.
- **Niveles de desempeño:** Categorías que describen los logros o niveles de cumplimiento de una población con base en el rendimiento en una evaluación.
- **TCT:** Teoría Clásica de los Test.
- **TRI:** Teoría de Respuesta al Ítem.

- **Parámetros psicométricos:** Especifican los niveles de dificultad, discriminación, pseudoadivinación u otros parámetros que describen el comportamiento de un ítem. Los parámetros pueden variar en función de la teoría psicométrica usada.

5. Marco Teórico

La evaluación educativa es un proceso fundamental en el ámbito de la educación que tiene como objetivo medir y valorar el aprendizaje y el desarrollo de los estudiantes, así como la efectividad de los métodos de enseñanza y los programas educativos en general. Es una herramienta clave para obtener información relevante sobre el progreso y los logros de los estudiantes, y para tomar decisiones informadas que promuevan la mejora continua de la calidad educativa. La evaluación educativa, se puede considerar como un instrumento para sensibilizar el quehacer académico y facilitar la innovación (Gonzáles & Ayarza, 1997).

En este sentido, la evaluación educativa abarca diferentes aspectos, desde la evaluación del desempeño individual de los estudiantes hasta la evaluación de los currículos, los programas escolares, los docentes y las instituciones educativas en su conjunto. Su propósito principal es proporcionar información objetiva y significativa que permita tomar medidas para optimizar el proceso de enseñanza-aprendizaje y garantizar que se cumplan los objetivos educativos establecidos.

Es importante destacar que la evaluación educativa no solo se enfoca en los resultados cuantitativos, como las calificaciones y los puntajes, sino también en aspectos cualitativos, como las habilidades, las competencias y las actitudes de los estudiantes. Se busca promover un enfoque integral que valore el proceso de aprendizaje en su totalidad y fomente el desarrollo integral de los estudiantes (López, Sánchez, Espinoza, & Carmona, 2013). Este fin, implica necesariamente hablar del agente evaluador que ejecute el proceso y garantice el enfoque integral de la evaluación.

Además, la evaluación puede clasificarse en función de quién la realiza y, entre otros tipos, tendríamos a la evaluación interna, la evaluación externa, la autoevaluación, la coevaluación y la heteroevaluación. En el presente documento se profundizará sobre la evaluación externa.

En síntesis, la evaluación interna se refiere a un proceso de evaluación realizado dentro de una institución, organización o sistema educativo para recopilar información y realizar análisis sobre el desempeño, el aprendizaje y otros aspectos relevantes. Este tipo de evaluación es llevado a cabo por los propios miembros de la institución, como docentes, directivos o equipos de evaluación interna designados. Tiene como fin mejorar la calidad de la educación, monitoreo o seguimiento de los aprendizajes y mejorar la toma de decisiones de las autoridades (García, 1997).

Es importante destacar que la evaluación en la educación superior no solo se realiza de manera interna, por parte de las instituciones mismas, sino que también existen instituciones externas e independientes encargadas de evaluar y monitorear la calidad de la educación superior a nivel nacional e internacional. Estas instituciones desempeñan un papel crucial en la promoción de estándares de calidad y en la mejora continua de las instituciones.

De esta manera, la evaluación externa es una conclusión del proceso de evaluación interna ya que su intención es valorar el grado en que la unidad académica o programa satisface los criterios de calidad establecidos (Quesada, 2006). El ERAP constituye una evaluación externa los programas de posgrado con fines de acreditación.

La evaluación en la educación superior es un proceso fundamental para asegurar la calidad y el cumplimiento de los estándares académicos en instituciones de educación superior, como universidades e institutos. Este tipo de evaluación se enfoca en diversos aspectos, que van

desde la evaluación de los programas y planes de estudio, hasta la evaluación de los docentes, los estudiantes y las infraestructuras.

En este sentido, un elemento clave en la evaluación en la educación superior es la acreditación. La acreditación es el resultado de un proceso de evaluación externa realizado por organismos especializados que verifican y reconocen la calidad de las instituciones y programas educativos. La acreditación garantiza que las instituciones, sus carreras y programas cumplan con los estándares establecidos en términos de currículo, recursos, infraestructura, docentes cualificados y resultados de aprendizaje, entre otros aspectos.

Otro componente de la evaluación en la educación superior es la evaluación del aprendizaje de los estudiantes. Esto implica medir y evaluar el progreso de los estudiantes en relación con los objetivos educativos establecidos. Se utilizan diferentes métodos de evaluación, como exámenes escritos, proyectos, presentaciones y trabajos prácticos, para evaluar el conocimiento, las habilidades y las competencias adquiridas por los estudiantes. Esta evaluación se realiza de manera periódica a lo largo del programa académico y tiene como objetivo brindar retroalimentación y orientación para el desarrollo continuo de los estudiantes (Haladyna & Downing, Handbook of Test Development (1st ed.), 2006).

En este sentido, la planificación clara y detallada de la construcción de una evaluación es el primer paso para lograr el objetivo de preparar, administrar, calificar y analizar una evaluación, de tal forma que presente evidencia razonable de validez para asegurar o refutar las inferencias tomadas a partir de sus resultados. La validez de una interpretación de resultados, a partir de un puntaje obtenido en una prueba, se basa en toda la evidencia relevante sobre la construcción de esa evaluación (Haladyna & Downing, 2006).

La aplicación de una evaluación facilita la obtención de información clave sobre los resultados del aprendizaje de los Programas de Posgrado en un determinado momento, sin embargo, es un proceso que implica una recolección sistemática de información para obtener un diagnóstico preciso del estado de dichos programas y tomar acciones enfocadas en su mejora.

De esta manera, un examen de alto impacto es aplicable para este tipo de evaluación, debido a que los resultados son empleados para tomar decisiones importantes que conllevan consecuencias significativas en quienes los sustentan. A diferencia de los exámenes a gran escala, los de alto impacto pueden o no ser administrados a grandes poblaciones. Por ejemplo, en el contexto educativo, los resultados de los exámenes de aptitud, certificación y dominio son frecuentemente empleados por las partes interesadas como criterios de egreso/ingreso tanto en la educación media superior como superior (Mendoza, 2015). Los instrumentos de evaluación deben ser diseñados considerando los parámetros y condiciones adecuadas de los ítems o preguntas para garantizar la confiabilidad y validez de sus resultados según el propósito de la evaluación (Haladyna & Downing, 2006).

Existen varias metodologías para evaluar los ítems o preguntas, entre ellas las dos corrientes teóricas y metodológicas principales son: la Teoría Clásica de los Test (TCT) y la Teoría de Respuesta al Ítem (TRI) (Chávez & Saade, 2010).

La Teoría Clásica de los Test (TCT), también conocida como teoría clásica del puntaje verdadero, es una corriente teórico-metodológica de análisis de ítems o preguntas que se centra en la prueba, sin poner énfasis en el comportamiento de los ítems o preguntas individuales (Chávez & Saade, 2010).

A diferencia de los enfoques tradicionales de evaluación, que se centran en el número total de respuestas correctas, la Teoría de Respuesta al Ítem (TRI) se enfoca en la relación entre las características de los ítems o preguntas y las habilidades de los individuos mediante el uso de modelos matemáticos. La Teoría de Respuesta al Ítem (TRI) asume que las respuestas a los ítems o preguntas son indicativas de las características subyacentes del individuo, y busca estimar estas características a partir de los patrones de respuestas observados (Chávez & Saade, 2010).

Las teorías de estos modelos permiten calcular estimaciones precisas de las habilidades de los individuos, así como evaluar la calidad de los ítems o preguntas utilizados en una evaluación en términos de su dificultad, su capacidad de discriminación y las posibilidades de pseudo adivinación. La TRI tiene como base 3 supuestos (Chávez & Saade, 2010):

- a) Se asume la existencia de una variable no observada, latente o rasgo, que explicaría las respuestas de las personas a una prueba o ítem.
- b) Unidimensionalidad del rasgo, es decir, la evaluación mide un solo rasgo.
- c) Independencia local, la respuesta a un ítem no puede estar influenciada por otro ítem en el mismo instrumento de evaluación.

6. Metodología de evaluación de los resultados de aprendizaje (ERAP)

La planificación clara y detallada de la construcción de una evaluación es el primer paso para lograr el objetivo de preparar, administrar, calificar y analizar una evaluación, de tal forma que presente evidencia razonable de validez para asegurar o refutar las inferencias tomadas a partir de sus resultados.

Para desarrollar una evaluación que genere resultados confiables es necesario considerar los siguientes componentes como parte del proceso de construcción de los instrumentos para cada programa a evaluar.

1. Definición de contenido y estructuras de evaluación

2. Desarrollo de preguntas y bancos robustos

3. Ensamble de instrumentos de evaluación

El proceso metodológico inicia con la definición de la estructura y contenido del examen que se va a evaluar, para lo cual se requiere la construcción participativa de los delegados de los programas de posgrado por campo de conocimiento¹.

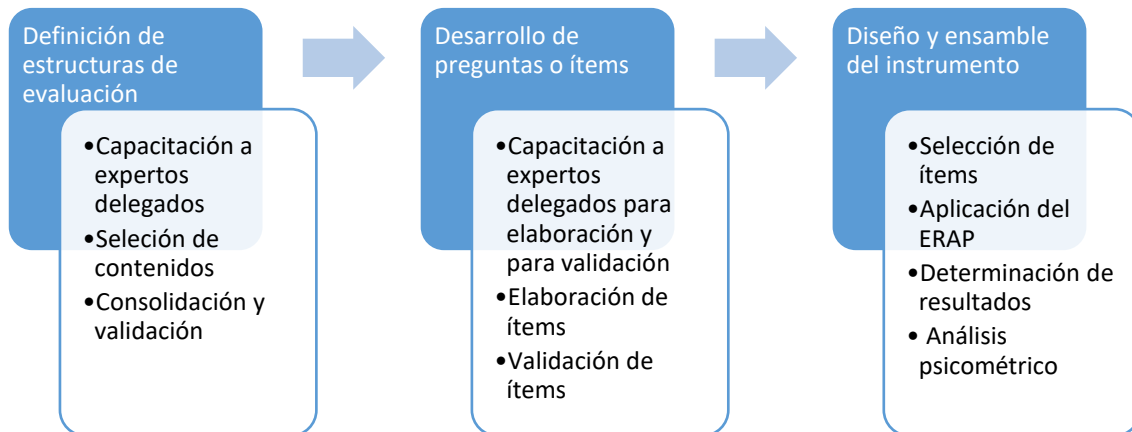
Una vez definida la estructura, se procede a elaborar y validar las preguntas que medirán los conocimientos o competencias requeridos. Este proceso lo ejecutará el CACES en conjunto con los delegados de los posgrados, se capacitará a los mismos para que elaboren preguntas de acuerdo con la metodología definida; estas preguntas serán revisadas en comités académicos de validación y serán piloteadas en escenarios similares al que se va aplicar el ERAP.

En todo este proceso, el CACES capacitará a los expertos delegados por los programas de posgrado en todos los aspectos necesarios para el desarrollo de la evaluación de resultados del aprendizaje y supervisará la entrega de documentos o productos necesarios para la construcción de la evaluación.

Una vez que se cuente con un banco de preguntas robusto, se realiza el ensamble y aplicación del examen conforme a características definidas para el comportamiento de este. Por último, se debe realizar un análisis psicométrico del comportamiento de las preguntas del examen, para determinar aspectos de mejora. El proceso del ERAP se indica en la figura 1.

¹ Definido para la evaluación del entorno de aprendizaje del programa.

Figura 1. Desarrollo del ERAP por campo de conocimiento



A continuación, se indica el flujo del proceso y se detallan las actividades para realizar la evaluación de resultados de aprendizaje en programas de posgrado:



Figura 2. Diagrama de flujo del proceso ERAP

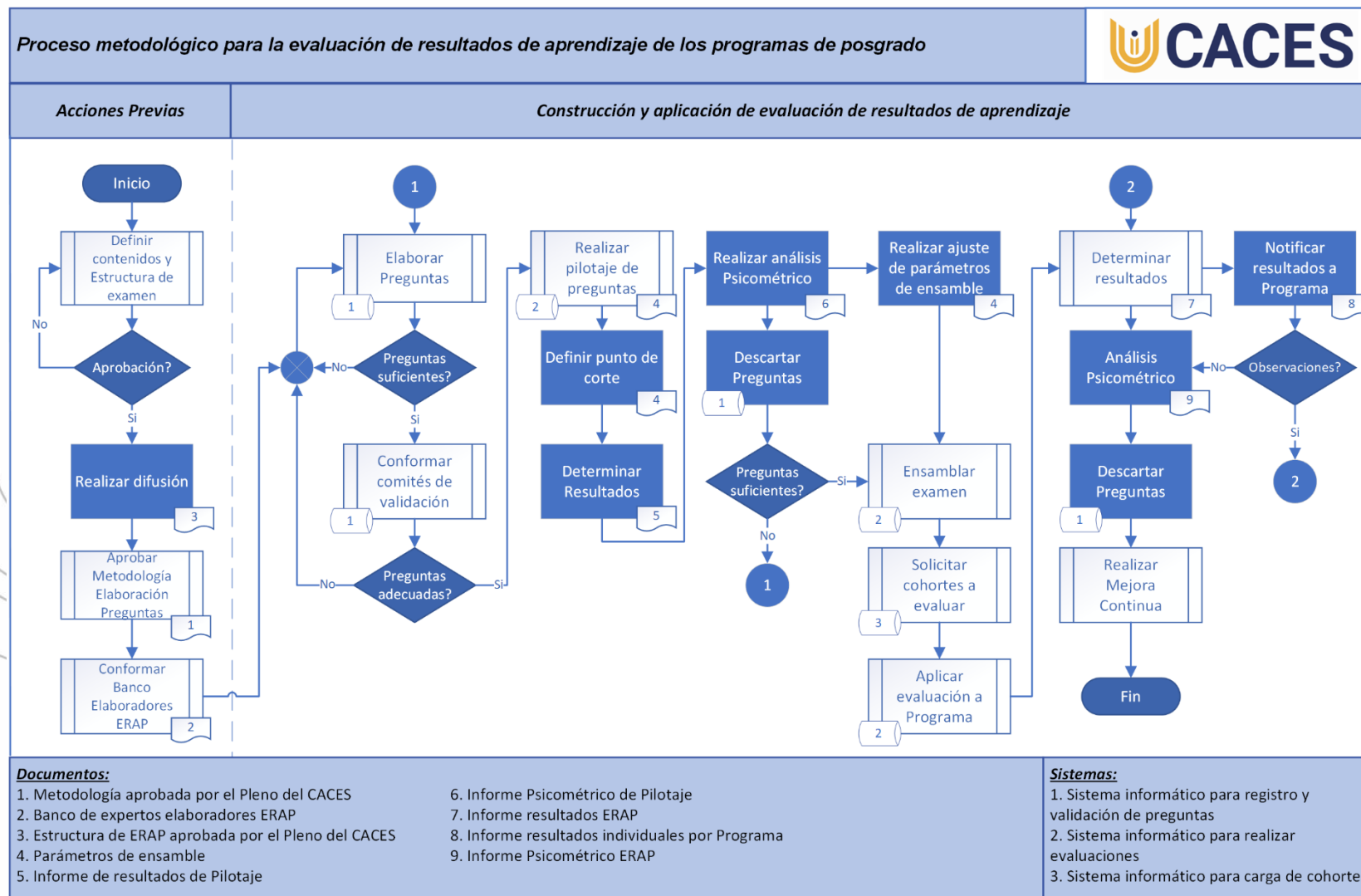
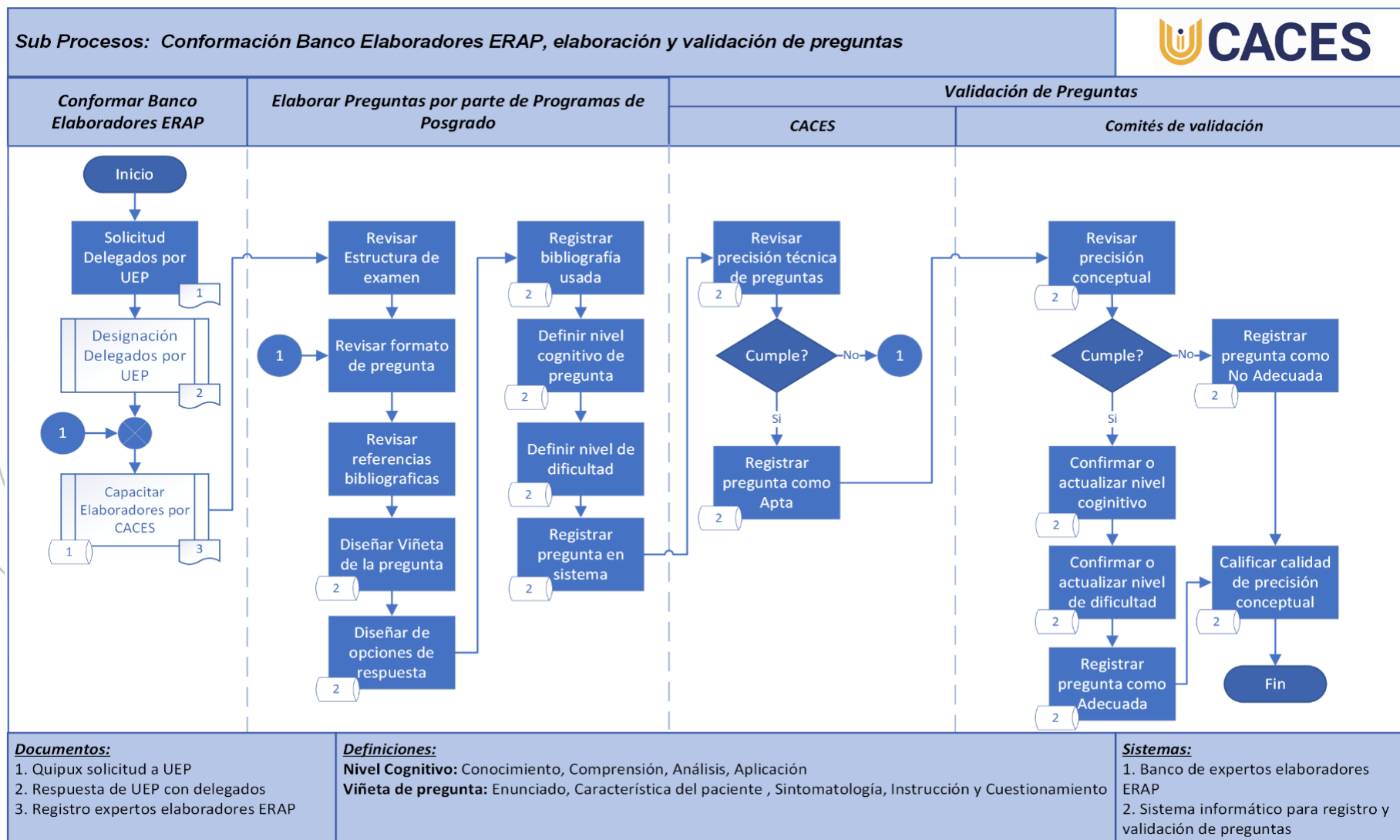


Figura 3. Diagrama de flujo de subprocesos del ERAP



6.1 Definir contenidos y estructura del examen

Conforme lo expuesto en el marco teórico, toda evidencia de validez está relacionada en gran medida con el contenido de una prueba (Haladyna & Downing, 2006). El contenido de la evaluación está condensado en la Estructura de Evaluación que detalla la desagregación del saber a evaluar en diferentes niveles.

Considerando la heterogeneidad de los contenidos abordados por los diferentes programas de posgrado, las estructuras de evaluación se agruparán en dos etapas, la primera por campo amplio para analizar contenidos comunes y la segunda por campo específico y detallado, de acuerdo con el modelo de evaluación del entorno de aprendizaje, para determinar las especificidades de la evaluación. En los casos que los contenidos comunes de campos amplios y específicos no coincidan, se considerará la secuencia que debe tener el campo amplio, el campo específico y el campo detallado tal como se muestra en la figura 4.

Figura 4. Agrupación por campo amplio, específico y detallado



Las especificaciones de la evaluación estarán contenidas en la ficha técnica del instrumento que deberá desarrollarse por campo de conocimiento. La ficha técnica contiene la información que permite planificar la ejecución de la evaluación y aplicarla a las cohortes definidas por los programas de posgrado, tales como:

- **Rasgo por medir:** Permite definir el objetivo de la evaluación y la interpretación adecuada de los resultados
- **Población objetivo:** Define las características de los individuos a evaluar.
- **Tipo de instrumento:** Define el tipo de examen que se utilizará, ya sea normativo o criterial, sumativo o formativo, etc.
- **Características y longitud del examen:** Especifican las condiciones del examen como tiempo, número de preguntas restricciones, material de apoyo, etc.
- **Uso de resultados:** Define si los resultados serán utilizados para aprobación, retroalimentación de los evaluados, acreditación del programa o demás usos que el programa considere necesario.
- **Punto de corte o niveles de desempeño:** Especifica el nivel crítico para delimitar la aprobación de los evaluados.
- **Periodicidad de la evaluación:** Se determina según el cronograma de evaluación con fines de acreditación.
- Demás información que se considere relevante para la evaluación.

Para el desarrollo de las estructuras de evaluación el CACES realizará la capacitación a los delegados de programas de posgrado y posteriormente ejecutará los comités académicos

para el desarrollo y validación de estructuras de evaluación. Los comités académicos estarán conformados por los delegados de los programas y un técnico del CACES participará en cada comité como facilitador. El desarrollo de la ficha técnica será responsabilidad del CACES y será difundida con los programas de posgrado.

Las estructuras de evaluación tendrán un desglose de contenidos en niveles definidos por los expertos delegados de los programas de tal manera que la estructura permita elaborar preguntas delimitadas a un determinado contenido.

6.2 Elaboración de ítems o preguntas

Los instrumentos de evaluación deberán ser diseñados considerando los parámetros y condiciones adecuadas de los ítems o preguntas para garantizar la confiabilidad y validez de sus resultados según el propósito de la evaluación (Haladyna & Downing, 2006).

Para elaborar un instrumento de evaluación es importante tener un banco de preguntas con parámetros psicométricos. En caso de no tener parámetros, es recomendable pilotear los ítems o preguntas antes de aplicarlos en su versión definitiva para descartar posibles imprecisiones técnicas. Este banco de preguntas estará conformado por aquellos que se ajusten adecuadamente a las necesidades del instrumento y permitirá equiparar la prueba a través de un análisis estadístico y psicométrico, además, permite mantener un repositorio de ítems lo suficientemente amplio para garantizar una evaluación sostenible.

Para el desarrollo del banco de preguntas, el CACES capacitará a los expertos delegados por los programas de posgrado sobre la metodología de elaboración de preguntas para definir un banco de elaboradores por cada campo de conocimiento.

El banco de elaboradores se encargará de construir las preguntas que permitan evaluar la totalidad de la estructura definida para cada programa, para ello, se desarrollarán talleres de

capacitación conforme la metodología de elaboración de preguntas establecida por el CACES en el Manual para el Diseño y Aplicación de Exámenes. El CACES asignará y solicitará la elaboración de preguntas a los programas, en caso de que no se complete el banco de preguntas, el CACES podrá contratar profesionales Expertos Elaboradores de Preguntas de acuerdo con la normativa legal vigente.

Por el uso que se les dará a los resultados del ERAP y las implicaciones que pueden tener sobre los programas de posgrado evaluados, se la considera como una evaluación de alto impacto y a gran escala. Dicho esto, las preguntas de opción múltiple son las más utilizadas en exámenes a gran escala por su alto grado de discriminación de las habilidades o conocimientos de los estudiantes o profesionales. Las preguntas de opción múltiple que se deberán diseñar están compuestas por un caso o situación que el profesional puede enfrentarse en una situación real, además deberá contar con cinco opciones de respuesta con sus respectivas argumentaciones (López, Sánchez, Espinoza, & Carmona, 2013).

La Metodología de Elaboración de Preguntas que apruebe el CACES aportará a profesores y comunidad académica para mejorar la calidad de preguntas mediante lineamientos técnicos con los que se deben construir las mismas para estos tipos de examen (Cómo elaborar preguntas para evaluaciones en el área de ciencias básicas y clínicas).

6.3 Validación de preguntas

Posterior a su elaboración, las preguntas deberán ser validadas por expertos afines al campo de conocimiento de los programas a evaluar. Los expertos validadores de preguntas serán vinculados por el CACES conforme la normativa legal vigente y respetando el perfil de contratación establecido. En caso de ser necesario, los expertos validadores podrán ser delegados, seleccionados del banco de elaboradores² según su desempeño o capacitados para

² Revisión y validación cruzada de preguntas.

realizar la validación de preguntas en comités académicos, los mismos conformarán el Banco de validadores. En caso de que no se valide la totalidad de preguntas elaboradas, el CACES podrá contratar Expertos Validadores de Preguntas de acuerdo con la normativa legal vigente.

Se definirá un banco de preguntas abierto para que los sustentantes puedan familiarizarse con la evaluación como un ejercicio de simulación y pilotaje.

6.4 Pilotaje de preguntas o ítems

El proceso de pilotaje de los ítems o preguntas permite asegurar un banco de ítems libre de errores técnicos o con características psicométricas no óptimas, es decir, con indicadores de dificultad o de discriminación no apropiados para la evaluación.

Cuando el impacto de una evaluación se considera bajo, como por ejemplo una evaluación de diagnóstico, un estudio o pilotaje de baja escala entre 10 a 30 evaluados puede ser suficiente para detectar problemas con los ítems. Sin embargo, cuando la evaluación es de alto impacto, como una acreditación, el pilotaje es un ejercicio mucho más exigente en términos de tamaño y representatividad de la muestra para que a partir de ella se estimen las métricas definitivas con las que los ítems serán incluidos en el banco (Chávez & Saade, 2010).

En la tabla 1 se muestran los tamaños de muestra mínimos que se necesitan para realizar un pilotaje adecuado según la teoría que se utilice.

Tabla 1. Tamaños de muestra mínimos según la teoría de análisis.

Modelo (Teoría)	Parámetros	Número de sustentantes requeridos
Teoría Clásica de los Test	N/A	Alrededor de 100
	Rasch (1 parámetro de dificultad)	100
Teoría de Respuesta al Ítem	2 parámetros (dificultad y discriminación)	250-500
	3 parámetros (dificultad, discriminación y pseudo adivinación)	Más de 1000

Fuente: (Ceneval, 2010)

El pilotaje de ítems será coordinado por el CACES en conjunto con los programas de posgrado y debe realizarse en condiciones que sean lo más parecidas posibles a las de una aplicación oficial o definitiva, esto garantiza estimaciones robustas.

6.5 Diseño y ensamble del instrumento de evaluación

La evaluación se operativiza mediante la selección de preguntas o ítems con mejores parámetros y su agrupación en un examen con un orden definido para medir los resultados de aprendizaje. La selección se analiza estadísticamente.

En esta fase se analiza el comportamiento psicométrico de cada ítem o preguntas y su inclusión en el instrumento de evaluación, contrastando sus parámetros individuales con el conjunto de ítems o preguntas seleccionados. La articulación entre los procesos del análisis cuantitativo con los del análisis cualitativo de los ítems o preguntas contribuye a garantizar la confiabilidad de los resultados de las evaluaciones para la adecuada toma de decisiones.

El CACES desarrollará el proceso de diseño y ensamble que permitirá generar los instrumentos de evaluación o examen para cada programa de posgrado. Para garantizar la

seguridad de la información, las preguntas seleccionadas no serán revisadas por expertos delegados de los programas de posgrado, sin embargo, se especificarán los contenidos a evaluar en la ficha técnica del instrumento.

a) Selección de ítems y agrupación de la prueba

Para realizar el diseño de un instrumento de evaluación, el CACES seleccionará los ítems o preguntas que han sido analizados y tienen parámetros psicométricos aceptables (óptimos y útiles) para que puedan ser parte del examen. Los parámetros que se analizan en los ítems o preguntas dependen de la teoría psicométrica que se pretenda utilizar para el desarrollo del diseño; en términos generales, se analiza la dificultad, la discriminación y la pseudoadivinación de los ítems o preguntas, sobre esta última, es necesario señalar que se analiza únicamente con TRI.

La dificultad del ítem determina el nivel de habilidad necesaria para responderlo; cuanto más difícil es el ítem, mayor nivel de habilidad será necesario para llegar adecuadamente a su respuesta. Este indicador se analiza de manera diferente en TCT que en TRI. Es así como, conocer la distribución de las dificultades de los ítems o preguntas permite diseñar pruebas que se parezcan entre sí. Se puede obtener versiones de una misma prueba con una dificultad “similar” eligiendo los ítems o preguntas con cuidado para que sus dificultades se distribuyan de una manera predeterminada en cada una de ellas.

Particularmente, si se utiliza TRI, la dificultad se estima como el nivel de habilidad requerido para obtener una probabilidad de acierto igual a 0,5. Los niveles de dificultad de los ítems o preguntas se estiman en el rango de entre -3 y 3 siendo -3 un ítem más fácil y 3 un ítem más difícil. Los tamaños poblacionales óptimos para utilizar las teorías psicométricas se muestran en las tablas 2 y 3.

Tabla 2. Categorías de dictámenes según TCT

Teoría	Parámetro	Dictamen	Rango
TCT	Dificultad	Descartado	$\text{dif_tct} < 10$ ó $\text{dif_tct} > 90$
		Útil	$10 \leq \text{dif_tct} < 20$ ó $80 < \text{dif_tct} \leq 90$
		Óptimo	$20 \leq \text{dif_tct} \leq 80$
	Discriminación	Descartado	$\text{dis_tct} < 0.15$
		Útil	$0.15 \leq \text{dis_tct} < 0.2$
		Óptimo	$0.2 \leq \text{dis_tct} \leq 1$

Fuente: CENEVAL, 2010

Tabla 3. Categorías de dictámenes según TRI

Teoría	Parámetro	Dictamen	Rango
TRI	Dificultad	Descartado	$\text{dif_tri} < -3$ ó $\text{dif_tri} > 3$
		Útil	$-3 \leq \text{dif_tri} < -2.5$ ó $2.5 < \text{dif_tri} \leq 3$
		Óptimo	$-2.5 \leq \text{dif_tri} \leq 2.5$
	Discriminación	Descartado	$\text{dis_tri} < 0.4$ ó $\text{dis_tri} > 2.8$
		Útil	$0.4 \leq \text{dis_tri} < 0.45$
		Óptimo	$0.45 \leq \text{dis_tri} \leq 2.8$

Fuente: CENEVAL, 2010

Aquellas preguntas o ítems óptimos y útiles son seleccionados para conformar el diseño del instrumento, verificando que se genere una dispersión homogénea de las dificultades o probabilidades de respuesta.

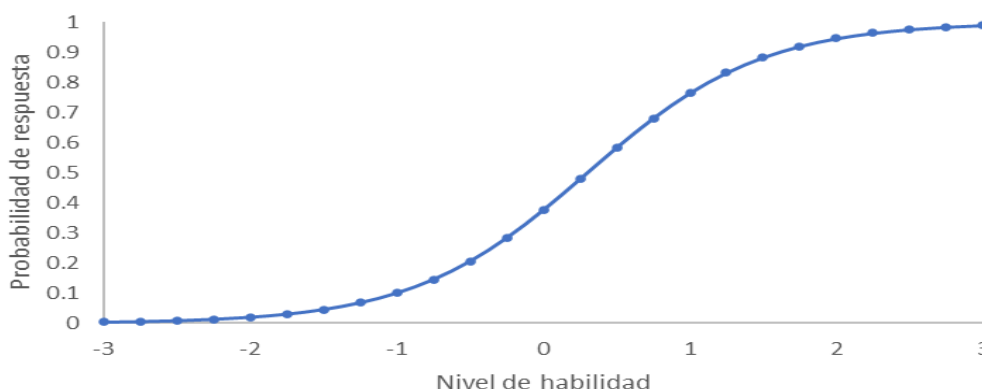
La dificultad y la discriminación de cada ítem son los principales parámetros de revisión, no obstante, se analizan los parámetros de pseudoadivinación y confiabilidad según la disponibilidad de información. El análisis de cada ítem necesariamente debe estar acompañado del análisis de rasgo latente o constructo estimado en la población, el cual es conocido también como habilidad, y determina la probabilidad de acertar o no a una pregunta en conjunto con sus parámetros (Muñiz & Fernández-Hermida, 2010).

La dificultad de cada ítem o pregunta es estimada en la misma dimensión que la habilidad de la población, por ello se puede decir que un sustentante domina un determinado ítem si su habilidad estimada es mayor que la dificultad.

La discriminación del ítem mide su capacidad para diferenciar a la población en función de su nivel de habilidad o rasgo latente (Hidalgo-Montesinos & French, 2016). El parámetro de pseudoadivinación indica la posibilidad de acertar a un ítem por azar.

La curva característica de un ítem muestra su comportamiento en diferentes niveles de habilidad y lo contrasta con sus probabilidades de respuesta, como se muestra en la figura 5.

Figura 5. Ejemplo de curva característica de un ítem o pregunta



La curva característica del ítem permite entender cómo funcionan los ítems o preguntas en diferentes niveles de habilidad y así definir un mismo nivel para el cual se calculan las probabilidades de respuesta. Centrar la habilidad en un punto común para todos los ítems o preguntas permitirá obtener un diseño equitativo.

b) Análisis diferencial, distractores y función de información

El análisis diferencial de los ítems o preguntas permitirá determinar los posibles sesgos que se presenten al momento de ser aplicados a una población, los cuales están relacionados con las características propias de la población, tales como:

- Territorio
- Género
- Tipo de sostenimiento
- Área (urbana o rural)
- Etnia o autoidentificación étnica

El análisis de distractores permite determinar si la pregunta o ítem funciona adecuadamente y facilita la separación de la población por el dominio o no del ítem. Los distractores son las opciones de repuesta incorrectas, pero mantienen un grado de plausibilidad, de tal manera que la selección de la respuesta correcta implique el dominio del constructo

evaluado y no esté afectado por factores externos. La función de información contribuye a entender el comportamiento de cada ítem en los diferentes niveles de habilidad y a comprender el funcionamiento del instrumento.

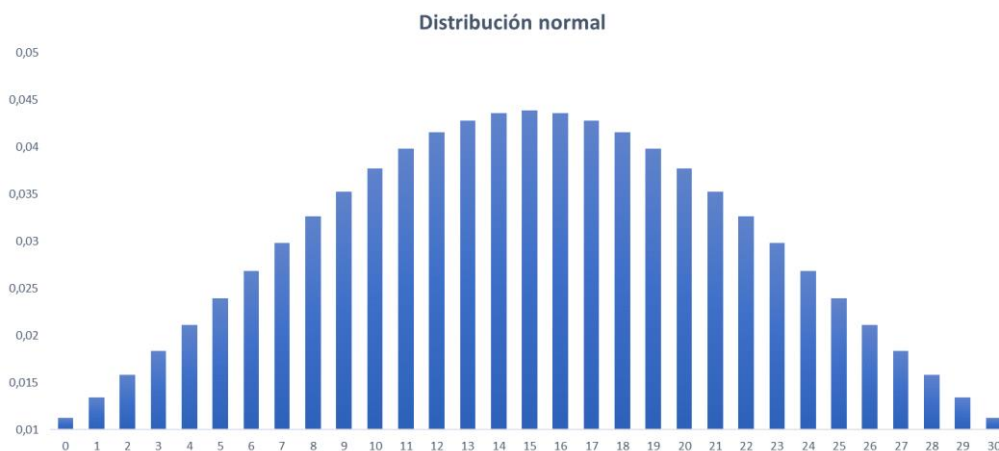
c) Distribución poblacional

La distribución poblacional refleja cómo se dispersan los puntajes obtenidos por cada sustentante. Al seleccionar los ítems o preguntas que conformarán el diseño del instrumento, se pretende generar una dispersión de parámetros homogénea que de origen a una distribución poblacional simétrica. Esta dispersión deberá aproximarse a una distribución normal, a fin de garantizar la igualdad de oportunidades (Fernández, 1997), sin embargo, en determinadas evaluaciones se buscará una distribución sesgada según el objetivo del proyecto. A continuación, se detalla los tipos de distribución:

- Distribución simétrica: es aquella que nos presenta uniformidad en la dispersión de datos y no presenta sesgos.
- Distribución asimétrica: puede sesgarse a la derecha o la izquierda. Es aquella cuyos datos pueden estar concentrados un lado de la distribución.

Al utilizar la distribución normal se espera obtener una selección de sustentantes que demuestren mejor desempeño, dependiendo del cumplimiento de las condiciones necesarias del Teorema Central del Límite. La figura 6 describe la dispersión esperada.

Figura 6. Dispersión esperada de resultados.

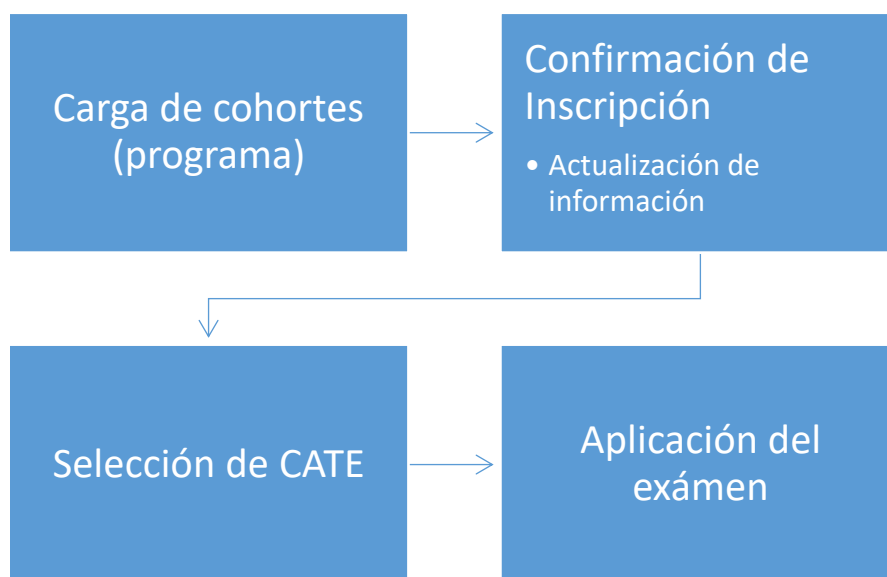


La dispersión simétrica debe generarse en torno a un punto central que puede ser la esperanza de aciertos o el punto de corte para aprobar la evaluación. De esta manera se garantiza una evaluación sin sesgos que afecten la confiabilidad de sus resultados. En caso de necesitar una distribución sesgada, lo ideal es modificar la esperanza de aciertos y no el punto de corte.

6.6 Aplicación del examen de evaluación de resultados del aprendizaje de programas de posgrado

Conforme lo establecido en el Manual para el Diseño y Aplicación de Exámenes vigente del CACES, a continuación, se adaptan las actividades operativas que deben desarrollarse para la implementación del ERAP. A continuación, se describe el flujo macro de la aplicación de la evaluación.

Figura 7. Aplicación del ERAP



6.6.1 Cohortes

La cohorte hace referencia a los convocados explícitamente para rendir un examen, en el caso del ERAP, los estudiantes del último nivel, conforme lo establecido en el numeral 6 del artículo 3 del “REGLAMENTO DE EVALUACIÓN EXTERNA CON FINES DE ACREDITACIÓN PARA EL ASEGURAMIENTO DE LA CALIDAD DE LOS PROGRAMAS DE POSGRADO DE LAS UNIVERSIDADES Y ESCUELAS POLITÉCNICAS”. Las cohortes de convocados se definirán una vez aprobado el cronograma de evaluación con fines de acreditación de programas de posgrado aprobado por el Pleno del CACES.

El CACES elaborará una proyección del tamaño de la cohorte o cantidad de estudiantes a evaluar para iniciar la planificación logística, para posteriormente realizar los ajustes necesarios sobre la misma. Dicha proyección se realizará en función de datos históricos, tendencia y con márgenes de error suficientemente amplios. Una vez se define el cronograma de evaluación, se solicita a los programas de posgrado que carguen las cohortes.

El Cronograma del examen contendrá al menos las fechas de ejecución de las siguientes actividades:

- a) Convocatoria para rendir el examen; z
- b) Inscripción
- b) Confirmación de inscripción para rendir el examen;
- c) Solicitud de actualización de información registrada en la inscripción para rendir el examen;
- d) Jornadas de acompañamiento;
- e) Aplicación del examen;
- f) Recepción de solicitudes de revisión académica
- g) Análisis de las solicitudes de revisión académica
- h) Determinación de resultados del análisis de las solicitudes de revisión académica
- i) Recepción de solicitudes de revisión del cometimiento de las prohibiciones establecidas en el Manual para el diseño y aplicación de exámenes(de ser el caso);
- j) Análisis de las solicitudes de revisión del cometimiento de las prohibiciones establecidas en el Manual para el diseño y aplicación de exámenes
- k) Determinación de resultados del análisis de las solicitudes de revisión del cometimiento de las prohibiciones establecidas en el Manual para el diseño y aplicación de exámenes
- l) Determinación de resultados definitivos; e,
- m) Informe de resultados definitivos y notificación a la UEP/Programa.

6.6.2 Convocatoria para rendir el examen

La convocatoria es el anuncio público y notificación individual a los estudiantes reportados en las cohortes por los programas de posgrado, que el CACES llevará a cabo el examen de resultados de aprendizaje y será publicada en la página web del CACES, y las redes sociales que la Unidad de Comunicación Social del CACES considere para una difusión oportuna, efectiva y eficiente, y será notificada a las instituciones pertinentes para su conocimiento y difusión.

6.6.3 Selección de Centro Autorizado para la Toma del Examen (CATE)

En la fase de inscripción, los postulantes serán informados sobre el CATE y jornada (día y horario) en la que rendirán el examen. En el caso de la Evaluación de Resultados del Aprendizaje en los Programas de Posgrado (ERAP), el CATE podrá ser el mismo centro evaluado en caso de cumplir con los requisitos tecnológicos y de infraestructura o podrá ser otro centro autorizado en la jornada y horario que determine el CACES.

6.6.4 Confirmación de inscripción

El proceso de confirmación de inscripción se debe realizar por los propios estudiantes en el sistema informático del CACES (GIIES). La Secretaría Técnica, en coordinación con la Unidad de Tecnologías de la Información y Comunicaciones del CACES, se encargará de mantener la confidencialidad, disponibilidad, integridad y autenticidad de la información receptada en el sistema GIIES por los postulantes al momento de la inscripción.

6.6.5 Solicitud de actualización de información registrada en la inscripción

La solicitud de modificación de información se podrá realizar para la confirmación de inscripción y se corresponderá exclusivamente para datos personales; los convocados tienen hasta tres (3) días término después de la finalización del periodo de confirmación de inscripción para presentar su solicitud a través del correo electrónico creado para el proceso de evaluación.

Para ello, deberán remitir el "formulario de solicitud de modificación de información", que estará disponible en la página web institucional del CACES, junto con copia de su documento de identidad.

No se aceptarán solicitudes de cambio de CATE, o de jornada, una vez que el postulante en su proceso de inscripción las haya seleccionado como sitio para rendir el examen. Con excepción de solicitudes por caso fortuito o fuerza mayor debidamente justificadas, por decisión de la Presidencia de la Comisión Permanente de Programas de Posgrado.

6.6.6 Aplicación del examen

Finalmente, se ejecuta la evaluación en el CATE validando a los estudiantes inscritos. Los estudiantes desarrollarán la evaluación en el marco de la normativa interna del CACES vigente para el desarrollo, aplicación y análisis de resultados de exámenes.

6.6.7 Resultados definitivos del ERAP

Previo a la entrega de resultados definitivos se realiza el trámite de solicitudes de revisión académica. Las solicitudes de revisión académica deberán ser canalizadas por el programa de posgrado conforme lo establecido en el Manual para el Diseño y Aplicación de Exámenes.

El trámite de cometimiento de prohibiciones se realizará conforme lo establecido en el Manual para el Diseño y Aplicación de Exámenes.

El informe de resultados definitivos del ERAP de cada programa debe ser aprobado por el Pleno del CACES e incluirá la información del rendimiento de cada sustentante y demás información estadística que se considere relevante.

6.6.8 Difusión y comunicación de resultados de aprendizaje

Como parte del proceso de enseñanza y aprendizaje, los planes de estudio y los resultados de las evaluaciones realizadas deben estar a disposición de los estudiantes, profesores y público en general.

Los programas de posgrado deben realizar procesos de difusión o comunicación de los resultados del aprendizaje con las partes interesadas de la comunidad académica para realizar retroalimentación y procesos de mejora continua. Cada programa debe ejecutar los procesos comunicativos de resultados de las evaluaciones del aprendizaje realizados con las partes interesadas y sobre los procesos de comunicación del plan de estudios con la comunidad académica.

Previo a la entrega de resultados definitivos se realiza el trámite de solicitudes de revisión académica y de cometimiento de prohibiciones conforme lo establecido en el Manual para el Diseño y Aplicación de Exámenes.

6.7 Análisis psicométrico

El análisis psicométrico corresponde al análisis matemático de las características de las preguntas de un examen, denominados reactivos o ítems, incluyendo su calidad y pertinencia. Existen varias metodologías para evaluar los reactivos, entre ellas las dos corrientes teóricas y metodológicas principales son: la teoría clásica de los test y la teoría de respuesta al ítem (Chávez & Saade, 2010).

La elección de la teoría a utilizar para el análisis de ítems depende de la coyuntura en la que se desarrolle la evaluación, es decir, las condiciones tanto teóricas como prácticas influyen en el tipo de análisis o la teoría que se utilizará. La Teoría Clásica (TCT) es recomendable por su simplicidad, sin embargo, la Teoría de Respuesta al Ítem (TRI) es recomendable cuando se

esperan diferencias importantes de habilidad entre distintas poblaciones a las que se aplique una evaluación, dicho de otro modo, cuando se desarrollan versiones distintas de un test se pueden obtener estimaciones comparables de las habilidades utilizando TRI (Chávez & Saade, 2010).

Mediante el análisis psicométrico se podrá evidenciar qué reactivos necesitan ser descartados o necesitan ajustarse para que su inclusión en la evaluación no afecte a la confiabilidad de los resultados.

6.8 Mejora continua

El proceso de mejora continua toma como base el análisis psicométrico y permite mejorar cada parte del proceso de desarrollo de los instrumentos de evaluación.

El proceso de mejora continua puede recomendar la revisión de las estructuras de evaluación ya sea para modificar o eliminar contenidos, también puede contribuir a la retroalimentación del proceso de elaboración y validación de preguntas sea para la inclusión o exclusión de preguntas, modificación de preguntas, redefinición o revisión del proceso de validación de preguntas.

6.9 Cronograma genérico para el desarrollo del ERAP

A continuación, se presenta el cronograma de construcción del ERAP.

Tabla 4. Cronograma de construcción de la evaluación de resultados del aprendizaje.

[illegible]

7. Bibliografía

- ACBSP. (30 de 08 de 2023). *Consejo de Acreditación de Escuelas y Programas de Negocios*. Obtenido de ACBSP: <https://acbsp.org/>
- Argibay, J. C. (2006). Técnicas psicométricas. Cuestiones de Validez y Confiabilidad. (U. d. Sociales, Ed.) *Subjetividad y Procesos Cognitivos*, 8, 15-33.
- Asamblea Nacional. (2018). *Ley Orgánica de Educación Superior*. Quito: Registro Oficial Suplemento 298 de 12-oct.-2010.
- Asamblea Nacional. (2023). *Ley Organica de Educación Superior*. Quito: Asamblea Nacional.
- CACES. (2023). *Reglamento de Evaluación Externa con fines de Acreditación para el Aseguramiento de la Calidad de los Programas de Posgrado de las Universidades y Escuelas Politécnicas*. Quito: Consejo de Aseguramiento de la Calidad de la Educación Superior.
- Cano García, E. (2005). *EL PORTAFOLIOS DEL PROFESORADO UNIVERSITARIO*. Barcelona, España: OCTAEDRO - ICE.
- Cano García, E. (2015). Las rúbricas como instrumento de evaluación de competencias en educación superior. *Profesorado. Revista de currículum y formación del profesorado*, 19, 266-280. Obtenido de <https://revistaseug.ugr.es/index.php/profesorado/article/view/18779/18277>
- CAPES. (2010). PLANO NACIONAL DE PÓS-GRADUAÇÃO (PNPG) 2011-2020. Brasilia.
- CAPES. (2017). regulamento da Avaliação. Obtenido de <https://www.capes.gov.br/historia-e-missao>
- Case, S., & Swanson, D. (2006). *Cómo elaborar preguntas para evaluaciones escritas en el área de ciencias básicas y clínicas*. Philadelphia: National Board of Medical Examiners.
- Ceneval. (2010). *Cuadernillo Técnico*. México D.F.: Centro Nacional de Evaluación para la Educación.
- Ceneval. (2018). *Estatutos Sociales del Centro Nacional de Evaluación para la Educación Superior A.C*. Ciudad de México: Centro Nacional de Evaluación para la Educación Superior .
- Ceneval. (2022). *La relevancia de la evaluación estandarizada externa: A 28 años de la creación de Ceneval*. Ciudad de México: Centro Nacional de Evaluación para la Educación Superior.
- CES. (2022). *Guía metodológica para la presentación de carreras y programas, ajustes curriculares sustantivos y no sustantivos* . Quito: Consejo de Educación Superior.
- Chávez, C., & Saade, A. (2010). *Procedimientos básicos para el análisis de reactivos*. México D.F.: Centro Nacional de Evaluación para la Educación Superior.
- CNA. (2010). "Autoevaluación con fines de acreditación de alta calidad de programas de maestría y doctorado: Guía de procedimiento" "Acreditación de alta calidad de Posgrados. Obtenido de <https://www.cna.gov.co/1741/article-186363.html>
- CNA. (2010). *Acreditación de programas de postgrado. "Lineamientos para la Acreditación de Alta Calidad de Maestrías y Doctorados"*. Obtenido de <https://www.cna.gov.co/1741/article-186363.html>

- CNA. (2019). *Consejo Nacional de Acreditación*. Obtenido de <https://www.cna.gov.co/1741/article-186363.html>
- CNA, Colombia. (5 de octubre de 2022). *ACTUALIZACIÓN DE LOS ASPECTOS POR*. Obtenido de [articles-412513_norma.pdf](#)
- CNA-Chile. (2013). *Resolución exenta DJ NRO 006-4 Aprueba criterios para la acreditación de programas de posgrado*. Obtenido de <https://www.cnachile.cl/Documentos%20de%20Paginas/Criterios%20vigentes%20para%20la%20Acreditación%20de%20programas%20de%20postgrado%20a%20partir%20del%2004%20de%20noviembre%20del%202013.pdf>
- CNA-Chile. (2014). *Procedimiento para el desarrollo de los procesos de acreditación de los programas de posgrado*. Obtenido de <https://www.cnachile.cl/Paginas/Criterios-procedimientos.aspx>
- CNA-Chile. (2016). *Operacionalización de Criterios de Evaluación para la acreditación de programas de posgrado: Doctorado, Magíster académico y Magister profesional*. Obtenido de <https://www.cnachile.cl/Paginas/Acreditacion-Postgrado.aspx>
- CNA-Chile. (2018). Aseguramiento de la calidad de programas de doctorado: convergencias y desafíos para Iberoamérica. *Serie estudios sobre acreditación - Comisión Nacional de Acreditación.*, 30-31.
- Conacyt. (2016). *Estatuto Orgánico del Consejo Nacional de Ciencia y Tecnología*. Ciudad de México: Consejo Nacional de Ciencia y Tecnología. Obtenido de <https://www.conacyt.gob.mx/index.php/el-conacyt>
- Conacyt. (2021a). *Terminos de Referencia para la evaluación y seguimiento de programas de nuevo ingreso*. Ciudad de México: Consejo Nacional de Ciencia y Tecnología.
- Conacyt. (2021b). *Terminos de Referencia para la Renovación y Seguimiento de Programas de Posgrado*. Ciudad de México: Consejo Nacional de Ciencia y Tecnología.
- CONEAU. (2013). *Ordenanza 059/13*. Obtenido de <https://www.coneau.gob.ar/coneau/acreditacion-de-carreras/carreras-de-posgrado/carreras-en-funcionamiento/normativa/>
- CONEAU. (2013). *RESOLUCIÓN N°: 116/13*. Obtenido de <https://www.coneau.gob.ar/archivos/resoluciones/Res116-13C30207.pdf>
- CONEAU. (2015). *Comisión Nacional de Evaluación y Acreditación Universitaria*. Obtenido de <https://www.coneau.gob.ar/coneau/que-es-la-coneau/>
- CONEAU. (2019). *Posgrados acreditados en la Republica de Argentina. CONEAU 2019*. Obtenido de https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&cad=rja&uact=8&ved=2ahUKEwim5IPP6fnoAhWDL-AKHSUxAM4QFjAAegQIAxAB&url=https%3A%2F%2Fwww.coneau.gob.ar%2Farquivos%2Fpublicaciones%2FCatalogoPosgrados_2019.pdf&usg=AOvVaw2qP-MygDRM5oPrTZhkunKa
- Fernández, A. (1997). Uso de la distribución normal en la evaluación del aprendizaje. *Estudios Pedagógicos*, 51-63. Obtenido de <https://dx.doi.org/10.4067/S0718-07051997000100005>

- Gonzáles, L., & Ayarza, H. (1997). Calidad, evaluación institucional y acreditación en la educación superior en la región Latinoamericana y del Caribe. *Documentos de la Conferencia Regional Políticas y Estrategias para la Transformación de la Educación Superior en América Latina y el Caribe*. La Habana: CRESALC-UNESCO.
- Grosso, M. (2013). *Curso Internacional sobre Cooperación para el desarrollo y fortalecimiento de la Calidad de los Estudios de Postgrado y Doctorado en Iberoamérica*. Obtenido de https://www.aui.org/images/stories/DATOS/PublicacionesOnLine/curso_internacional_trujillo/Ponencia_Marcela_Grosso_Trujillo_Espana.pdf
- Haladyna, T. M., & Downing, S. M. (2006). *Handbook of Test Development (1st ed.)*. New Jersey: Routledge.
- Haladyna, T. M., Downing, S. M., & Rodríguez, M. C. (2002). *A review of multiple-choice item-writing guidelines for classroom assessment (15 ed., Vol. 15)*. (I. Lawrence Erlbaum Associates, Ed.) London: APPLIED MEASUREMENT IN EDUCATION.
- Hernández, R., Fernández, C., & Baptista, M. (2014). *Metodología de la Investigación (Sexta ed.)*. México: McGrawHill.
- Hidalgo-Montesinos, & French. (2016). Una introducción didáctica a la Teoría de Respuesta al Ítem para comprender la construcción de escalas. *Revista de Psicología Clínica con niños y adolescentes*, 3.
- López, A., Sánchez, H., Espinoza, J., & Carmona, M. (2013). *Elaboración de Items de opción múltiple*. Quito: Instituto Nacional de Evaluación Educativa INEVAL.
- Lucio, R. (2009). *La construcción del saber y del saber hacer (Vol. 8 y 9)*. (U. d. Antioquia, Ed.) Medellín: Revista Educación y Pedagogía.
- MEN. (2023a). *Oferta de combinatorias para el examen Saber Pro*. Bogotá: Ministerio de Educación Nacional.
- MEN. (2023b). *Familiarízate con el examen saber pro ICFES*. Bogotá : Ministerio de Educación Nacional.
- MEN. (2023c). *Guía de orientación Saber Pro: Módulo de competencias genéricas*. Bogotá: Ministerio de Educación Nacional.
- Mendoza, A. (2015). La validez de los exámenes de alto impacto. *perfiles Educativos*, 37(149), 169-186.
- Meneses, J. (2013). *Psicometría*. Barcelona: Eureka Media S.L.
- Muñiz, J. (2010). LAS TEORÍAS DE LOS TESTS: TEORÍA CLÁSICA Y TEORÍA DE RESPUESTA A LOS ÍTEMS. (C. G. Psicólogos, Ed.) *Papeles del Psicólogo*, 31(1), 57-66. Recuperado el 22 de 1 de 2024, de <https://www.redalyc.org/articulo.oa?id=77812441006>
- Muñiz, J., & Fernández-Hermida, J. R. (2010). La opinión de los psicólogos españoles sobre el uso de los tests. *Papeles del Psicólogo*.
- National Board of Medical Examiners. (2006). *Cómo elaborar preguntas para evaluaciones en el área de ciencias básicas y clínicas*. Philadelphia: NBME.
- Prieto, G., & Delgado, A. (2010). Fiabilidad y Validez. *Papeles del Psicólogo*, 67-74.
- Quesada, A. (2006). Evaluación externa: características y vivencias del equipo de pares. *Revista Electrónica "Actualidades Investigativas en Educación"*, 6(2), 2-16. Obtenido de <https://www.redalyc.org/pdf/447/44760212.pdf>