

Evaluación del egreso de nivel licenciatura en la Universidad Autónoma de Baja California

JOSÉ ALFONSO JIMÉNEZ MORENO* | JOAQUÍN CASO NIEBLA**
ANDREA ROMERO MOJICA***

El artículo aborda la experiencia de desarrollo de un mecanismo para la evaluación del egreso de nivel licenciatura de nueve programas educativos de baja matrícula de la Universidad Autónoma de Baja California: Artes plásticas, Bioingeniería, Docencia de la matemática, Docencia de la lengua y la literatura, Ingeniería en energías renovables, Enseñanza de lenguas, Oceanología, Sociología y Traducción. Para su desarrollo se consideró una metodología de evaluación comprensiva, capaz de resaltar la participación de las comunidades académicas y de reflejar la perspectiva institucional de cada disciplina para obtener un ejercicio de evaluación del egreso, inédito para las licenciaturas en cuestión. Se concluye la necesidad de velar por el uso formativo de los resultados y considerar este ejercicio como un insumo de mejora curricular.

The article discusses the experience of developing a mechanism for the assessment of the undergraduate degree of nine low-enrollment educational programs of the Autonomous University of Baja California: Plastic Arts, Bioengineering, Teaching Mathematics, Teaching Language and Literature, Renewable Energy Engineering, Language Teaching, Oceanology, Sociology, and Translation. For its development a comprehensive evaluation methodology was considered, able to highlight the participation of academic communities and to reflect the institutional perspective of each discipline to obtain an unprecedented graduate evaluation exercise for the degrees in question. The need to ensure the formative use of the results and to consider this exercise as an input for curriculum improvement is concluded.

Palabras clave

Evaluación educativa
Evaluación del aprendizaje
Pruebas de aprovechamiento
Egresados
Formación profesional

Keywords

Educational evaluation
Learning evaluation
Achievement tests
Graduates
Vocational training

Recepción: 2 de junio de 2023 | Aceptación: 29 de noviembre de 2023

DOI: <https://doi.org/10.22201/iisue.24486167e.2024.184.61383>

* Profesor de la Facultad de Ciencias Políticas y Sociales (FCPyS) de la Universidad Nacional Autónoma de México (UNAM) (México). Doctor en Pedagogía. Línea de investigación: evaluación educativa. CE: alpsic@gmail.com
ORCID: <https://orcid.org/0000-0003-0704-7883>

** Investigador del Instituto de Investigación y Desarrollo Educativo (IIDE) de la Universidad Autónoma de Baja California (UABC). Doctor en Psicología. Línea de investigación: evaluación educativa. CE: jcaso@uabc.edu.mx
ORCID: <https://orcid.org/0000-0002-3557-1722>

*** Profesora de la Facultad de Ciencias Administrativas y Sociales (FCAyS) de la Universidad Autónoma de Baja California (UABC). Maestra en Pedagogía. Línea de investigación: evaluación y ciudadanía. CE: andrea.romero.mojica@uabc.edu.mx ORCID: <https://orcid.org/0000-0002-7906-4546>

INTRODUCCIÓN

La evaluación es uno de los procesos educativos de mayor auge en las últimas décadas; existen diversas vertientes, cada una con determinados objetivos, metodologías y alcances. Una de las evaluaciones que mayor impacto tiene para los estudiantes, así como para la vida institucional, es la que se dirige al aprendizaje, ya que se centra en los conocimientos y habilidades adquiridos en la formación profesional. La evaluación del egreso a nivel superior en nivel licenciatura es un ejercicio valioso, que cuenta con diversas experiencias a nivel internacional (Huamán *et al.*, 2020); esto debido a que se considera como un mecanismo capaz de proveer información para la toma de decisiones formativas y sumativas para los estudiantes, los programas educativos y las universidades en general.

En el caso de México, la evaluación del egreso tiene al menos 30 años de existencia (Gago, 2000; Vidal, 2013), sin embargo, ha sido poco aplicada desde las perspectivas institucionales. El tema del presente artículo es la evaluación de los aprendizajes al egreso de la formación profesional, específicamente bajo el contexto e implementación de una institución pública, la Universidad Autónoma de Baja California (UABC).

La UABC está ubicada al noroeste de México. Es una universidad preocupada por asegurar una formación profesional acorde con los contextos actuales y, como tal, cuenta con diversos mecanismos para robustecer sus procesos formativos. Parte de las herramientas que utiliza con esa finalidad es la búsqueda constante de actualización curricular, así como el desarrollo de procesos de evaluación de diversos componentes (aprendizaje, docentes, entre otros).

En años recientes, la UABC ha fomentado fuertemente la evaluación del egreso de nivel licenciatura. Este ejercicio se encuentra normado en el Estatuto escolar (UABC, 2018), que en su artículo 84 establece lo que se pretende con este tipo de evaluaciones:

- I. Identificar la medida en que los egresados de los programas de licenciatura cuentan con los conocimientos y habilidades que son esenciales para el inicio del ejercicio profesional;
- II. Conocer el nivel de efectividad de los programas de licenciatura;
- III. Contar con información académica útil para la evaluación de la operación y conducción del programa educativo, y
- IV. Obtener información oportuna que contribuya a la actualización o modificación de los planes del estudio (UABC, 2018: 9).

Con base en estos objetivos, la UABC recurre al uso de mecanismos de evaluación externos, particularmente el Examen General para el Egreso de la Licenciatura (EGEL), instrumento de tradición psicométrica en México, desarrollado desde 1994 por el Centro Nacional de Evaluación para la Educación Superior (Ceneval) (Gago, 2000). El EGEL ofrece una alternativa de evaluación del egreso entre los 42 instrumentos con los que cuenta, cada uno enfocado en una disciplina en particular.¹ De esta manera, a través de este instrumento, la UABC, así como cientos de universidades públicas y privadas en el país, valoran los conocimientos y habilidades clave alcanzados por los estudiantes una vez concluida su formación profesional. Sin embargo, existen diversos programas de la UABC que no cuentan con mecanismos para realizar las evaluaciones del egreso, dado que instancias como el Ceneval no atienden a este tipo de disciplinas. Estas

¹ Biología, Ciencias agrícolas, Enfermería, Medicina general, Medicina veterinaria y zootecnia, Nutrición, Odontología, Psicología, Química, Química clínica, Químico farmacéutico biólogo, Profesional técnico en enfermería, Arquitectura, Ciencias computacionales, Diseño gráfico, Informática, Ingeniería biomédica, Ingeniería civil, Ingeniería computacional, Ingeniería de *software*, Ingeniería eléctrica, Ingeniería electrónica, Ingeniería en alimentos, Ingeniería en biotecnología, Ingeniería industrial, Ingeniería mecánica, Ingeniería mecánica eléctrica, Ingeniería mecatrónica, Ingeniería química, Administración, Ciencia política y administración pública, Ciencias de la comunicación, Comercio/negocios internacionales, Contaduría, Derecho, Economía, Finanzas, Gastronomía, Mercadotecnia, Pedagogía-ciencias de la educación, Relaciones internacionales, Trabajo social y Turismo (Ceneval, 2023).

formaciones, por lo tanto, no cuentan con posibilidades institucionales para cumplir con las intenciones establecidas en el Estatuto escolar; y carecen de información sobre el logro educativo alcanzado por sus estudiantes al finalizar la formación profesional, lo cual se considera un insumo importante para la retroalimentación de los programas educativos y los procesos que los constituyen.

Como antecedente para abordar esta necesidad, en 2018, la UABC promovió el desarrollo del examen para el egreso de la Licenciatura en Actividad Física y Deportes (LAFyD), primer instrumento en México en su tipo elaborado para una profesión de esta naturaleza (Jiménez *et al.*, 2019). Para ello se adoptó una perspectiva comprensiva que guiara la interpretación de los resultados, así como sus usos formativos (Stake, 2006); ésta establece que la evaluación debe buscar reflejar una realidad en particular y no centrarse exclusivamente en los elementos técnicos asociados a los instrumentos. Esta perspectiva permitió diversos logros, entre ellos:

1. Dar un peso importante al currículo de la formación profesional. No se trataba únicamente de reflejar los contenidos u objetivos explícitos de un plan de estudio dentro de un examen, sino que los contenidos del instrumento estarían matizados por la experiencia de los docentes del programa, la visión institucional de la formación en cuestión, así como por las particularidades de los diferentes campus en los que se ofrece el programa.
2. Considerar lo efímero de la propuesta evaluativa. Dado que la evaluación pretende reflejar una realidad en particular, se asume que ésta es dinámica; por tanto, los instrumentos requieren ser considerados como limitados en cuanto a su alcance y vigencia.

3. Derivado de los elementos previamente descritos, se debería resaltar el valor de los juicios profesionales de quienes desarrollan y operan el plan de estudio, ya que es esa población la que tiene un conocimiento directo de la realidad que la evaluación pretende reflejar.

Así, con el antecedente de implementación de la evaluación del egreso de la LAFyD, la UABC se dio a la tarea de establecer un esquema para el desarrollo de exámenes de egreso para más programas que carecían de referentes como los que ofrece el Ceneval. Esto permitiría que aquéllos que no cuentan con este tipo de mecanismos: 1) cumplan con las demandas estatutarias; 2) cuenten con un instrumento desarrollado por su comunidad para el reflejo de logros de aprendizaje del currículo que opera en sus aulas; y 3) cuenten con resultados del logro de sus estudiantes al momento de concluir su formación profesional; de esta manera se busca potenciar el análisis de las condiciones asociadas al trayecto formativo, los servicios de apoyo estudiantil y servicio social, entre otros componentes curriculares.

A partir de ello, el presente artículo se enfoca en la difusión de la experiencia institucional de la UABC en el desarrollo e implementación de nueve exámenes para el egreso de nivel licenciatura, en la búsqueda de contar con elementos de retroalimentación de sus programas educativos. La intención última de este ejercicio institucional se fundamenta en la búsqueda de mecanismos de retroalimentación y mejora de los programas educativos desde la perspectiva comprensiva; por lo mismo, el artículo aborda el proceso de construcción de los nueve instrumentos. Para ello se describen los antecedentes de la evaluación del egreso, posteriormente se retoman los referentes teórico-conceptuales utilizados, para continuar con la descripción de la metodología implementada, seguida de los resultados obtenidos y las conclusiones derivadas de la experiencia.

ANTECEDENTES DE LA EVALUACIÓN DEL EGRESO

La evaluación de conocimientos y habilidades obtenidas en la formación profesional es un tema de gran potencial formativo para las instituciones educativas, ya que permite la valoración de los logros que se han alcanzado como producto de varios años de formación. Al respecto, existen tendencias para realizar trabajos de esta naturaleza; desde un acercamiento a la opinión de los egresados respecto a la formación recibida (a través de cuestionarios y entrevistas), o bien, a través de la implementación de mecanismos asociados a procesos de evaluación del aprendizaje (Huamán *et al.*, 2020). Respecto del análisis en los ejercicios particulares que requieren el desarrollo de instrumentos de medición del aprendizaje o logro educativo, hay un incipiente desarrollo a nivel mundial. La recopilación realizada por Jiménez (2020) refleja no más de 15 ejercicios de esta naturaleza en países como Australia, Brasil, Canadá, Reino Unido, México, Estados Unidos y Colombia.

Interesa resaltar el caso del Ceneval, pionero en la evaluación del egreso, tanto en México como a nivel internacional. Derivado de políticas modernas, enfocadas en la valoración sumativa y de rendición de cuentas de los logros alcanzados por la formación profesional, a mediados de los noventa, la Asociación Nacional de Universidades e Instituciones de Educación Superior (ANUIES) fomentó la creación del Ceneval (Gago, 2000); institución que, hasta la fecha, se encarga de desarrollar instrumentos de evaluación dirigidos a otorgar información que facilite la toma de decisiones. Entre sus aportes se encuentra el EGEL, un instrumento cuyo uso anual alcanza a más de 660 mil estudiantes.

A través del Ceneval, en México, la evaluación del egreso se ha implementado de manera sistemática desde 1994 (Gago, 2000). Con los EGEL, las instituciones de educación superior mexicanas han contado con un mecanismo

externo para la valoración de la formación al egreso, que contempla el análisis de los conocimientos y habilidades que subyacen al ejercicio de 42 disciplinas; todas ellas caracterizadas por ser formaciones profesionales de alta demanda. Esto ha derivado en una serie de estudios sobre los factores asociados a los resultados que obtienen los estudiantes (Aguilar y Díaz, 2021; Cornejo *et al.*, 2020; Méndez *et al.*, 2022; Velázquez-Núñez *et al.*, 2019); pero, sobre todo, ha facilitado la implementación de una cultura evaluativa, reflejada en la alta aceptación e implementación de estos exámenes en diversas universidades de carácter nacional. En el ámbito metodológico, los EGEL han sido punta de lanza para una evaluación a gran escala; utilizan reactivos de opción múltiple para cumplir con la intención de una evaluación objetiva.

No obstante lo anterior, y a pesar de los aportes que realiza el Ceneval a través del EGEL, se carece de una política pública encaminada a aprovechar los resultados de estos instrumentos a nivel institucional y nacional, salvo algunos estudios en los que se identifican elementos asociados a éstos; o bien, en los que se analizan de manera crítica (Vera y González-Ledesma, 2018). Adicionalmente, en México son escasos los ejercicios de evaluación del egreso a través del acercamiento a los conocimientos y habilidades de los estudiantes que se realizan como iniciativas propias dentro de las universidades. Ejemplo de ello es el caso de la Universidad Nacional Autónoma de México (UNAM), la cual aplica un instrumento de opción múltiple como uno de los medios para la titulación de sus estudiantes de Medicina (Delgado-Maldonado y Sánchez-Mendiola, 2012). Otro ejemplo más reciente es el trabajo de la LAFyD de la UABC (Jiménez *et al.*, 2019), el cual no sólo responde a la normatividad que rige a dicha institución respecto a las diversas maneras de retroalimentar la formación en este programa, sino que resulta un ejercicio inédito a nivel internacional de una evaluación del egreso para un programa en actividad física, caracterizado, además, por ser de

baja matrícula. En resumen, la evaluación del egreso, generada a través de los conocimientos y habilidades obtenidos en la formación profesional, es un ejercicio escasamente abordado. Dados los antecedentes disponibles, así como la necesidad de que la UABC establezca mecanismos de evaluación del aprendizaje, en el año 2020 se dio a la tarea de dar continuidad a los trabajos institucionales iniciados en 2018 con la evaluación del egreso de la LAFyD. Esta experiencia permitió sentar las bases para plantear la posibilidad de abarcar más programas de la UABC que no contaran con referentes externos de evaluación, y con base en ello se decidió desarrollar los exámenes de egreso de los siguientes programas educativos: Artes plásticas, Bioingeniería, Docencia de la matemática, Docencia de la lengua y la literatura, Ingeniería en energías renovables, Enseñanza de lenguas, Oceanología, Sociología y Traducción. Todos ellos se caracterizan por ser de baja matrícula (menos de 100 estudiantes por generación en todos los casos), así como por no tener mecanismos de evaluación propuestos por instancias externas; de esta manera, la UABC afrontó un importante reto académico y operativo. Bajo estos antecedentes, a continuación se describe el proceso metodológico llevado a cabo.

METODOLOGÍA

Se replicó el modelo utilizado en la experiencia de la LAFyD, el cual se fundamenta en la búsqueda de una evaluación que recurra a los principios de la evaluación comprensiva (Stake, 2006). El modelo comprensivo surgió como un elemento propio para la evaluación de programas, en respuesta a la necesidad de contar con esquemas de evaluación más abarcativos y profundos que los tradicionalmente conocidos, como el centrado en objetivos (propuesto por Tyler), o bien, el modelo basado en el consumidor (establecido por Scriven) (Alcaraz, 2015; García, 2010). Stake propuso una visión que implica dar relevancia a las

particularidades de la realidad que la evaluación pretende reflejar, más que enfocarse en demasía en los estándares que otros modelos de evaluación promueven. Así, se entiende que la evaluación que se realice bajo esta perspectiva buscará resaltar la manera en que se recopilan las particularidades del programa u objeto de evaluación de interés; de esta manera, los resultados que se emitan procurarán reflejar una realidad en particular —en este caso, la experiencia de haber egresado de un programa educativo de nivel profesional, observado a través de las habilidades obtenidas— (Jiménez, 2020). De acuerdo con Stake (2006), un trabajo de esta naturaleza representa más una actitud que un esquema, por lo tanto, la evaluación será un ejercicio guiado por la receptividad e intereses de los agentes involucrados, y no sólo basado en elementos técnicos asociados con la medición.

El modelo comprensivo suele asociarse fuertemente a metodologías y técnicas de carácter interpretativo, sin embargo, esto no necesariamente es así, pues el modelo promueve la diversidad técnica para la recolección de información. El aspecto clave reside en no considerar las técnicas utilizadas como el eje del ejercicio evaluativo, sino en poner especial atención a las experiencias y valores de quienes conforman las comunidades asociadas con el objeto de evaluación y la experiencia formativa de interés (Stake, 2006). En este sentido, el enfoque adoptado por la UABC procura reflejar la experiencia formativa de sus estudiantes en los nueve programas educativos en cuestión bajo la perspectiva particular de sus comunidades académicas. Frente a ello, resulta relevante la visión de las y los docentes de cada programa, ya que su dominio disciplinar ofrecerá un reflejo importante de la formación que se pretende analizar; otro insumo importante es el plan de estudio, que articula los discursos, las intenciones y las habilidades esperadas.

Si bien el enfoque adoptado no pone en el centro de los esfuerzos a los instrumentos elaborados, sí reconoce su relevancia como

una herramienta que deberá reflejar una realidad en particular mediante una estrategia de cuantificación basada en la visión de quienes dan vida a la formación profesional que interesa analizar. A su vez, se da especial atención a la emisión de resultados y su comunicación a diversos agentes educativos, incluidos los estudiantes participantes. De esta manera, la perspectiva comprensiva adoptada no se orienta hacia intenciones técnicas, como la búsqueda de refinamiento de procesos métricos (como la validez y la confiabilidad); la intención de la UABC en esta materia es dar un lugar preponderante a la impronta que las comunidades académicas imprimen a los productos desarrollados, a las características

particulares de los planes de estudio de los programas de interés, así como al establecimiento de mecanismos de comunicación de resultados y análisis para la mejora de la formación profesional.

Como se comentó, los programas educativos participantes fueron seleccionados a partir del criterio primordial de que fueran de baja matrícula y que no contaran con instrumentos externos para la evaluación del egreso; un segundo criterio fue que las comunidades académicas estuvieran abiertas a colaborar en este proyecto. Diez programas educativos se sumaron a la iniciativa. En la Tabla 1 de enlistan los programas, las unidades académicas y los campus involucrados.²

Tabla 1. Programas educativos participantes en la evaluación

Programa educativo evaluado	Unidades académicas participantes	Campus participantes	Número de personal académico participante
Artes plásticas	Facultad de Artes	Mexicali Tijuana Ensenada	21
Bioingeniería	Facultad de Ingeniería Facultad de Ingeniería, Arquitectura y Diseño Facultad de Ciencias de la Ingeniería y Tecnología	Mexicali Tijuana Ensenada	21
Docencia de la matemática	Facultad de Pedagogía e Innovación Educativa	Mexicali Tijuana	24
Docencia de la lengua y la literatura	Facultad de Humanidades y Ciencias Sociales		13
Enseñanza de lenguas	Facultad de Idiomas	Mexicali Tijuana Ensenada	13
Traducción			5
Oceanología	Facultad de Ciencias Marinas	Ensenada	9
Sociología	Facultad de Ciencias Humanas Facultad de Humanidades y Ciencias Sociales Facultad de Ciencias Administrativas y Sociales	Mexicali Tijuana Ensenada	22
Ingeniería en energías renovables	Facultad de Ciencias de la Ingeniería y Tecnología	Mexicali Tijuana	11
Ingeniería aeroespacial*	Facultad de Ingeniería		12

*La sobrecarga académica de las comunidades asociadas no permitió la conclusión del proceso en este programa.

Fuente: elaboración propia.

² La UABC organiza a sus unidades académicas en tres campus: Tijuana, Mexicali y Ensenada.

El desarrollo de cada instrumento implicó la participación de una coordinación general del proyecto, junto con la colaboración de un coordinador específico para cada dos programas, así como un técnico asistente en cada caso. Por su parte, el personal académico participante fue personal docente de tiempo

completo y de asignatura de las unidades académicas participantes, todos ellos/ellas con experiencia formativa en los programas educativos de interés. De manera general, la evaluación, para todos los programas educativos, unidades y campus involucrados, implicó una serie de fases que se describen en la Tabla 2.

Tabla 2. Resumen de las fases del proyecto

Fase	Actividades asociadas	Productos generados	Periodo**
1. Delimitación conceptual	<ul style="list-style-type: none"> • Análisis colegiado del plan de estudio • Identificación colegiada de los conocimientos y habilidades por evaluar, a través de los perfiles de egreso y las unidades de aprendizaje de cada plan de estudio • Identificación colegiada del nivel taxonómico de cada conocimiento y habilidad identificado • Elaboración de tabla de especificaciones • Validación colegiada de tabla de especificaciones* 	<ul style="list-style-type: none"> • Delimitación conceptual del objeto de medida • Establecimiento de tabla de especificaciones 	Octubre 2020 a febrero 2021
2. Desarrollo de reactivos/ ítems***	<ul style="list-style-type: none"> • Capacitación para la elaboración de reactivos de opción múltiple (dicotómicos, con cuatro opciones de respuesta) • Desarrollo de reactivos de opción múltiple por parte de las comunidades académicas • Validación de reactivos* • Revisión lingüística de reactivos aprobados en validación 	<ul style="list-style-type: none"> • Reactivos validados y revisados lingüísticamente 	Febrero a octubre 2021
3. Pilotaje	<ul style="list-style-type: none"> • Elaboración de guías de estudio para estudiantes participantes • Preparación de <i>software</i> para la aplicación de los instrumentos • Desarrollo de la aplicación piloto de nueve instrumentos, bajo la organización de diversas instancias y unidades académicas universitarias en los tres campus de la institución 	<ul style="list-style-type: none"> • Guías de estudio para estudiantes participantes • Base de datos con respuestas de estudiantes participantes 	Abril 2022
4. Análisis de la aplicación piloto	<ul style="list-style-type: none"> • Implementación de análisis psicométricos a través de la teoría clásica de los test (TCT) (Muñiz, 2010) y calibración de los nueve instrumentos piloteados 	<ul style="list-style-type: none"> • Estructura operativa del instrumento • Establecimiento de banco de reactivos para la conformación de versiones paralelas 	Junio 2022

Tabla 2. Resumen de las fases del proyecto

(continuación)

Fase	Actividades asociadas	Productos generados	Periodo**
5. Definición de puntos de corte y emisión de resultados	<ul style="list-style-type: none"> • Establecimiento de puntos de corte para cada área de los nueve instrumentos, a través de la metodología Angoff (Cizek, 2011) • Definición de criterios de desempeño para cada instrumento • Elaboración de informes de resultados de cada estudiante participante, entregados a la Coordinación General de Formación Profesional**** 	<ul style="list-style-type: none"> • Puntos de corte y criterios de emisión de desempeño 	Agosto a diciembre 2022

*Cada proceso de validación se realizó bajo lineamientos y formatos específicos retomando las sugerencias de AERA/APA/NCME (2014).

**El periodo mostrado muestra el calendario general, sin embargo, hubo particularidades para cada programa, en función de las agendas de las diversas comunidades involucradas.

***Para el caso de este documento, los términos “ítems” y “reactivos” se utilizan de manera indistinta.

**** Oficina universitaria responsable de los procesos institucionales en materia de formación y evaluación.

Fuente: elaboración propia.

La Tabla 2 describe, de manera general, el proceso de construcción de los nueve instrumentos. En cada caso se procuró un trabajo colegiado con el fin de cumplir con las demandas técnicas sugeridas por la literatura internacional en materia métrica, por ejemplo, la sistematización de las especificaciones y reglas unificadas en los procesos de construcción y redacción de reactivos (AERA/APA/NCME, 2014). Con la intención de velar por la implementación del enfoque comprensivo dentro del proceso, se buscó que la labor realizada no se considerara únicamente como un proceso técnico, sino como una base que permitiera reflexionar a profundidad a las comunidades académicas participantes sobre los logros esperados en los planes de estudio, el trabajo didáctico que se realiza en cada facultad y las diferentes posibilidades de retroalimentación de los programas educativos (Stake, 2006).

En términos operativos, la elaboración de los instrumentos implicó una fuerte carga para las y los académicos participantes, además de una intensa organización de los equipos de trabajo para asegurar: a) el desarrollo de instrumentos acordes con las sugerencias técnicas de la literatura especializada (AERA/APA/NCME, 2014); b) el aprovechamiento de la experiencia, visión y perspectiva profesional

de la comunidad participante; c) la representatividad de los contenidos propuestos para cada instrumento respecto al plan de estudio; así como para la diversidad de unidades académicas y campus en los programas que se ofertan en más de una facultad (por ejemplo, Bioingeniería o Docencia de la matemática, entre otros). La atención en estos procesos se cuidó de manera transversal, lo que permitió trabajar bajo la guía técnica de la psicometría sin perder de vista que la esencia del proyecto se desarrollaba bajo el modelo de evaluación comprensiva (Stake, 2006; Jiménez, 2020); es decir que los instrumentos debían reflejar el plan de estudio, la experiencia del personal académico y los matices de enseñanza disciplinar entre facultades y campus.

Es importante resaltar que, como parte de las fases 1 y 2, todos los instrumentos se trabajaron con comités académicos, cada ejercicio con un número distinto de personas, en función del número de personal disponible en cada facultad (Tabla 1). De esta manera, si bien se contaba con una metodología eje en común para todos los instrumentos, al igual que de formatos institucionales, los comités se encargaron de establecer una perspectiva particular en cada contenido y reactivo. Para esta tarea, los participantes eran convocados por

las facultades, capacitados por las coordinaciones de cada instrumento y dotados del correspondiente apoyo técnico; asimismo, cada producto desarrollado fue sometido a procesos de validación a través de jueces (García *et al.*, 2017; INEE, 2017). Cabe resaltar que la validación de los reactivos procuró un proceso ciego, en el que tres académicos del programa educativo analizaban cada una de las propuestas sin conocer el nombre del autor; esto permitió asegurar la transparencia del proceso, así como la solidez disciplinar y técnica de cada uno de los reactivos elaborados.

En el ámbito psicométrico, la mayor limitante fue el número de sustentantes estimado que participaría en la aplicación piloto. El trabajo psicométrico suele ser más robusto conforme se cuenta con una mayor población participante (Bazán, 2004); sin embargo, el proyecto nació para desarrollar instrumentos para programas de baja matrícula. Si bien para la evaluación comprensiva, los datos psicométricos no son el eje que fundamenta una evaluación, era importante velar por las evidencias de validez y confiabilidad de los instrumentos, con el fin de cumplir con requerimientos técnicos internacionales en esta materia. En ese sentido, se estipuló trabajar el análisis psicométrico de la aplicación piloto desde la teoría clásica de los tests (TCT), además de acumular evidencia de cada instrumento durante un año, con el fin de desarrollar análisis más robustos, basados en la teoría de respuesta al ítem (TRI); así también, se analizó la estructura de cada prueba a través, inicialmente, del análisis factorial exploratorio, no sólo por el bajo número de participantes en cada caso, sino también debido a que los objetos de medida que se construyeron en este proceso carecen de antecedentes que permitan orientar posibilidades teóricas de conformación de estructura factorial más robusta.

Ahora bien, para la realización de la aplicación piloto se contó con el número de estudiantes que se presenta en la Tabla 3, todos ellos en el último semestre de su formación profesional en su respectivo programa educativo.

Tabla 3. Número de estudiantes participantes en la aplicación piloto

Instrumento / programa educativo	Número de sustentantes
Artes plásticas	24
Bioingeniería	41
Docencia de la matemática	16
Docencia de la lengua y la literatura	19
Enseñanza de lenguas	24
Energías renovables	52
Oceanología	5
Sociología	14
Traducción	32
Total	227

Fuente: elaboración propia.

Por otra parte, para establecer las reglas de calificación y emisión de resultados se decidió hacer uso de la metodología Angoff (Cizek, 2011), lo que permitió la consideración de dos importantes insumos: los resultados obtenidos a través de la aplicación piloto; y la visión profesional de las comunidades académicas, quienes, con su experiencia, dan sentido y orientación a los indicadores empíricos recolectados.

En términos operativos, la organización de la aplicación piloto requirió una compleja planificación. El proyecto se desarrolló en su totalidad a través de medios digitales debido a que se realizó bajo las restricciones impuestas durante la pandemia de COVID-19. Esto implicó el uso de plataformas digitales; en este sentido, la complicación radicó principalmente en la organización de la amplia cantidad de académicos en las diferentes facultades de los tres campus. En el transcurso del proyecto, en la UABC, al igual que en el resto del mundo, el retorno a las aulas se realizó de manera progresiva entre 2021 y 2022, lo que permitió la implementación de un esquema de aplicación digitalizado, pero presencial. De esta manera, la UABC desarrolló un aplicativo que permitió que la prueba piloto se desarrollara de manera digital, con lo cual se facilitó la operación aprovechando la infraestructura universitaria.

Bajo estas consideraciones metodológicas, así como gracias a un esfuerzo importante de la comunidad universitaria, se logró el desarrollo de las nueve evaluaciones para el egreso de licenciatura. A continuación se describen los principales resultados.

RESULTADOS

Los resultados de este ejercicio de evaluación se presentan en función de las cinco fases que lo conformaron.

Fase 1. Delimitación conceptual

Como parte de las primeras actividades se realizó un análisis de los planes de estudio

de cada programa para el cual se enfatizaron los siguientes componentes: a) la visión institucional de la disciplina; b) la organización curricular y de créditos, incluyendo las competencias de los programas de unidad de aprendizaje (PUA); y, c) el perfil de egreso. Dados estos tres elementos se elaboró un documento con la descripción conceptual de la disciplina, mismo que incluyó el análisis de los tres componentes descritos; y se estableció una propuesta de áreas del examen en función de las competencias declaradas en el perfil de egreso y su relación con los PUA.

Derivado de estos dos productos se estableció un trabajo colegiado con académicos de cada programa, con el fin de construir la

Tabla 4. Áreas de cada examen

Artes plásticas	Bioingeniería	Docencia de la matemática	Docencia de la lengua y la literatura
<ol style="list-style-type: none"> 1. Producción de obra artística profesional que refleje congruencia entre el discurso y la factura. 2. Fortalecer la educación artística a través del diseño e implementación de proyectos educativos que integren las artes plásticas. 3. Gestionar proyectos culturales para impulsar las artes en la sociedad. 4. Preparar textos divulgativos que reflejen el conocimiento artístico y críticas a las artes en la comunidad. 	<ol style="list-style-type: none"> 1. Elaborar dispositivos biomédicos, biomateriales y biocatalizadores para mejorar la producción industrial y la calidad de vida. 2. Incorporación de sistemas tecnológicos para uso biomédico en espacios físicos para contribuir a mejorar el servicio de salud e industrial. 3. Implementar estrategias para el desarrollo sustentable. 4. Participar en procesos administrativos en el área de bioingeniería para ayudar en el desarrollo económico y empresarial. 	<ol style="list-style-type: none"> 1. Analizar elementos pedagógicos y matemáticos de los planes y programas de estudio que orientan la enseñanza de las matemáticas en secundaria y media superior. 2. Desarrollar en los estudiantes las competencias matemáticas esperadas en la educación secundaria y media superior a través de propuestas didácticas específicas. 3. Aplicar la normatividad escolar y los valores relacionados a la práctica docente en secundaria y media superior. 4. Desarrollar proyectos de investigación e intervención educativa orientados a mejorar el proceso de enseñanza-aprendizaje de la matemática propios de la educación secundaria y media superior 5. Llevar a cabo procesos de análisis teórico-metodológico de la práctica docente para la mejora continua. 	<ol style="list-style-type: none"> 1. Desarrollar habilidades discursivas orales y escritas. 2. Dominar contenidos canónicos y representativos de la literatura universal. 3. Reconocer elementos contextuales para la implementación de estrategias didácticas. 4. Resolver problemáticas sobre el proceso de enseñanza-aprendizaje en el aula con base en las teorías psicopedagógicas. 5. Emplear modelos de la didáctica de la disciplina para el fortalecimiento de la enseñanza de la lengua y la literatura.

tabla de especificaciones. Para su construcción se realizó una decantación nomológica-deductiva de las áreas del examen propuestas (Castillo, 2012). La derivación se realizó en tres niveles, a los cuales se les denominó *competencia*, *acciones* y *habilidades*, respectivamente. Estos tres niveles permitieron un claro posicionamiento pragmático que orientó a los comités académicos para que evitaran consignar contenidos conceptuales; con ello se buscaba asegurar que el instrumento orientara la medición de conocimientos y habilidades cognitivos propios del ejercicio profesional. Además de esta decantación, cada habilidad resultante se clasificó en un nivel taxonómico en particular. La clasificación utilizada fue la taxonomía

de Bloom revisada (Krathwohl, 2002); la racionalidad de su uso versó, en primer lugar, en su utilización previa en el ejercicio de evaluación del egreso de la LAFyD; y, en segundo lugar, en su cercanía para el trabajo docente en la UABC, ya que era reconocida por las diversas disciplinas participantes en el proyecto.

La propuesta generada se revisó a través de un proceso de validación en cada una de las áreas disciplinares; los comités se conformaron por académicos de la propia UABC de las diferentes facultades participantes en el proyecto, lo que dio como resultado una tabla de especificaciones en la que se operacionaliza la intención de medida. La Tabla 4 muestra el resumen de las áreas de los exámenes.

Tabla 4. Áreas de cada examen

(continuación)

Energías renovables	Enseñanza de lenguas	Oceanología	Sociología
<ol style="list-style-type: none"> 1. Generar estrategias que permitan resolver los problemas de abastecimiento de energía. 2. Evaluar el recurso energético para determinar la viabilidad de la implementación de sistemas de generación de energía. 3. Evaluar el impacto ambiental en la generación y uso de energéticos. 4. Desarrollar proyectos sostenibles. 5. Implementar tecnologías y procesos acordes a la disponibilidad y la demanda energética. 6. Aplicar sistemas de gestión para la sustentabilidad energética. 7. Implementar programas de ahorro y uso eficiente de la energía. 	<ol style="list-style-type: none"> 1. Comunicación oral y escrita en español, en inglés y en una tercera lengua. 2. Elaboración de propuestas didácticas para la enseñanza de lenguas. 3. Evaluación del proceso de enseñanza-aprendizaje de lenguas. 4. Diseño e implementación de proyectos de intervención educativa en el campo de la enseñanza de lenguas. 	<ol style="list-style-type: none"> 1. Describir y modelar la variabilidad y dinamismo de los procesos y fenómenos del sistema marino. 2. Elaborar diagnósticos para el uso sustentable y solución de problemas en el sistema marino. 3. Proponer medidas para el uso y conservación de recursos renovables y no renovables en el sistema marino. 4. Proponer esquemas para mitigar daños al ambiente y a las actividades humanas generados por la introducción de agentes externos a los sistemas marinos. 5. Planear soluciones a problemas ambientales y humanos generados por fenómenos naturales. 	<ol style="list-style-type: none"> 1. Examinar fenómenos sociales locales, regionales, nacionales e internacionales. 2. Intervenir en el ámbito social aplicando teoría y métodos de intervención sociológica. 3. Gestionar para la implementación de planes, programas, proyectos y políticas aplicando herramientas teórico-metodológicas y habilidades de negociación. 4. Atender necesidades educativas en diversas situaciones y sectores sociales por medio de la aplicación de metodologías de las ciencias sociales.
Traducción			
<ol style="list-style-type: none"> 1. Expresión y comunicación oral de mensajes en español y en la segunda lengua. 2. Traducción de textos, tanto al español como a la segunda lengua. 3. Interpretación de textos, tanto al español como a la segunda lengua. 4. Uso y aplicación de herramientas tecnológicas como apoyo en los procesos de traducción escrita y oral. 5. Creación e implementación de empresas para la oferta de servicios de traducción. 			

Fuente: elaboración propia.

Las áreas de cada examen pretenden orientarse hacia el ejercicio profesional, más que agruparse en función de contenidos; de esta manera, los niveles de decantación de cada objeto de medida procuraron reflejar las acciones y habilidades particulares que permiten el desarrollo de estas áreas de trabajo profesional, en concordancia con los planes de estudio vigentes, así como con la mirada y experiencia de los comités académicos de cada facultad. Así, el establecimiento de la tabla de especificaciones representa ya, para cada caso, un importante logro para la UABC y la disciplina, en tanto que delimita una aproximación sin antecedentes del ejercicio profesional esperado por el plan de estudio y las comunidades participantes. Además, representa los límites y alcances interpretativos de los juicios que derivan de la evaluación en cada programa educativo, lo cual subraya la intención del uso comprensivo de los puntajes obtenidos a partir de la implementación de los instrumentos.

Fase 2. Desarrollo de reactivos

Una vez que se contó con las tablas de especificaciones se procedió a capacitar a las comunidades académicas de las facultades participantes para la elaboración de reactivos de opción múltiple. La capacitación, realizada a través de medios digitales, consideró lo siguiente:

- El uso de las habilidades y nivel taxonómico como eje para su elaboración.
- Los PUA relacionados con cada área del examen como referente de las competencias que se esperan formar en el trayecto formativo.
- El uso de reactivos dicotómicos de opción múltiple con cuatro opciones de respuesta.
- La propuesta de cinco formatos distintos para estructurar el planteamiento de los reactivos.

Cada reactivo fue revisado técnicamente por parte de la coordinación de cada instrumento, a través de un proceso de retroalimentación constante. Una vez que el reactivo dejó de tener observaciones técnicas o de formato se consideró listo para validación. Para el proceso de validación se cuidó que cada reactivo fuera analizado por tres docentes de manera ciega, junto con el uso de una lista de verificación enfocada a verificar requerimientos técnicos de los reactivos, así como posibles sesgos culturales y errores disciplinares (INEE, 2017). Al finalizar, cada reactivo fue sujeto a un proceso de revisión de estilo por parte de un experto externo a la UABC en materia de redacción de reactivos de opción múltiple (con experiencia en instancias nacionales de evaluación). Como resultado, se obtuvo el número de reactivos listos para pilotear que se describe en la Tabla 5. Cabe mencionar que, en función de la disciplina, algunos reactivos incluían imágenes, tablas y gráficas.

Tabla 5. Número de reactivos elaborados y validados

Instrumento	Reactivos aceptados en validación
Artes plásticas	254
Bioingeniería	162
Docencia de la matemática	276
Docencia de la lengua y la literatura	104
Enseñanza de lenguas	300
Energías renovables	119
Oceanología	284
Sociología	106
Traducción	296
Total	1,901

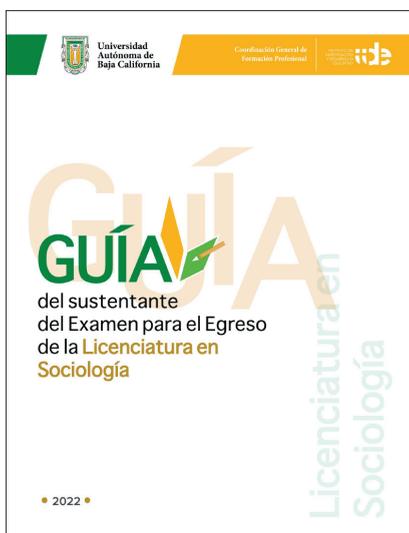
Fuente: elaboración propia.

Fase 3. Pilotaje

El pilotaje permitió poner a juicio empírico la propuesta desarrollada, así como valorar las cuestiones operativas asociadas al proyecto. Este proceso se planificó una vez que los estudiantes volvieron a las aulas de manera paulatina, al término del confinamiento provocado por la pandemia.

Con la finalidad de orientar a los estudiantes participantes y ofrecer bibliografía de consulta se elaboró una guía de estudio para cada programa, la cual fue difundida de manera digital entre las diversas comunidades. La Fig. 1 muestra la portada de una de estas guías.

Figura 1. Ejemplo de portada de la “Guía del sustentante”



Fuente: UABC (2022).

Derivado del pilotaje, se obtuvieron los datos necesarios para el análisis respectivo; a la par, se identificaron mejoras a los instrumentos, por ejemplo, respecto de la calidad de algunas imágenes poco legibles. En cuanto a lo operativo, se consideró que el *software* desarrollado por la institución permitía la aplicación del instrumento, sin embargo, requería mejoras para facilitar la carga de reactivos y la administración de los mismos.

Fase 4. Análisis de la aplicación piloto

Considerando los alcances y limitaciones metodológicas descritas, se realizó una calibración psicométrica bajo los preceptos de la TCT. Ello implicó observar las siguientes características en cada reactivo: correlación punto-biserial, dificultad, varianza, desviación estándar, análisis de distractores, error típico de medida y consistencia interna. En coherencia con la lógica de elaboración de cada instrumento, este análisis se realizó de manera independiente para cada área del examen (Tabla 2), todo lo cual derivó en los siguientes resultados (Tablas 6-14).

Tabla 6. Resultados psicométricos. Artes plásticas

Característica	Área 1	Área 2	Área 3	Área 4
Número medio de aciertos	30.50	50.25	39.67	16.17
Error típico de la media	0.75	1.76	2.91	1.11
Varianza	13.08	71.19	194.81	28.14
Desviación típica	3.62	8.44	13.96	5.30
Facilidad media	0.71	0.53	0.47	0.49
Correlación punto-biserial media	0.20	0.20	0.37	0.36
Fiabilidad (alfa de Cronbach)	0.523	0.747	0.926	0.791
Error típico de medida	2.50	4.24	3.81	2.42

Fuente: elaboración propia.

Tabla 7. Resultados psicométricos. Bioingeniería

Característica	Área 1	Área 2	Área 3	Área 4
Número medio de aciertos	22.20	34.59	15.63	1.05
Error típico de la media	0.52	0.78	0.53	0.12
Varianza	10.84	24.34	11.26	0.53
Desviación típica	3.29	4.93	3.36	0.73
Facilidad media	0.40	0.52	0.40	0.52
Correlación punto-biserial media	0.13	0.16	0.19	0.76
Fiabilidad (alfa de Cronbach)	-0.07	0.49	0.33	0.27
Error típico de medida	3.40	3.53	2.76	0.63

Fuente: elaboración propia.

Tabla 8. Resultados psicométricos. Docencia de la matemática

Característica	Área 1	Área 2	Área 3	Área 4	Área 5
Número medio de aciertos	33.19	31.69	16.31	26.94	29.69
Error típico de la media	1.84	1.62	0.65	1.89	1.60
Varianza	50.90	39.46	6.34	53.68	38.46
Desviación típica	7.13	6.28	2.52	7.33	6.20
Facilidad media	0.50	0.50	0.65	0.44	0.49
Correlación punto-biserial media	0.24	0.21	0.23	0.26	0.23
Fiabilidad (alfa de Cronbach)	0.77	0.67	0.32	0.79	0.70
Error típico de medida	3.43	3.59	2.08	3.38	3.40

Fuente: elaboración propia.

Tabla 9. Resultados psicométricos. Docencia de la lengua y la literatura

Característica	Área 1	Área 2	Área 3	Área 4	Área 5
Número medio de aciertos	14.89	5.16	6.53	6.21	14.89
Error típico de la media	0.63	0.38	0.36	0.40	0.61
Varianza	7.25	2.66	2.35	2.90	6.62
Desviación típica	2.69	1.63	1.53	1.70	2.57
Facilidad media	0.44	0.43	0.54	0.48	0.45
Correlación punto-biserial media	0.18	0.31	0.30	0.28	0.19
Fiabilidad (alfa de Cronbach)	0.21	0.30	0.12	0.06	0.03
Error típico de medida	2.39	1.36	1.44	1.66	2.53

Fuente: elaboración propia.

Tabla 10. Resultados psicométricos. Energías renovables

Característica	Área 1	Área 2	Área 3	Área 4	Área 5	Área 6	Área 7
Número medio de aciertos	11.67	4.67	7.31	8.33	4.15	5.77	1.46
Error típico de la media	0.55	0.27	0.29	0.31	0.23	0.25	0.14
Varianza	15.22	3.64	4.33	5.03	2.71	3.14	0.94
Desviación típica	3.90	1.91	2.08	2.24	1.65	1.77	0.97
Facilidad media	0.39	0.42	0.33	0.38	0.28	0.44	0.24
Correlación punto-biserial media	0.28	0.37	0.21	0.22	0.25	0.32	0.39
Fiabilidad (alfa de Cronbach)	0.61	0.39	0.15	0.13	-0.01	0.28	0.03
Error típico de medida	2.45	1.50	1.92	2.09	1.66	1.50	0.96

Fuente: elaboración propia.

Tabla 11. Resultados psicométricos. Enseñanza de lenguas

Característica	Área 1	Área 2	Área 3	Área 4
Número medio de aciertos	66.13	59.25	17.50	12.13
Error típico de la media	1.98	2.89	1.02	1.47
Varianza	90.28	192.69	23.92	49.86
Desviación típica	9.50	13.88	4.89	7.06
Facilidad media	0.61	0.53	0.46	0.29
Correlación punto-biserial media	0.21	0.27	0.27	0.37
Fiabilidad (alfa de Cronbach)	0.81	0.89	0.67	0.86
Error típico de medida	4.13	4.65	2.82	2.65

Fuente: elaboración propia.

Tabla 12. Resultados psicométricos. Oceanología

Característica	Área 1	Área 2	Área 3	Área 4	Área 5
Número medio de aciertos	29.20	25.20	30.80	23.60	26.80
Error típico de la media	1.36	1.02	1.83	1.50	2.03
Varianza	7.36	4.16	13.36	9.04	16.56
Desviación típica	2.71	2.04	3.66	3.01	4.07
Facilidad media	0.54	0.47	0.49	0.45	0.45
Correlación punto-biserial media	0.10	0.08	0.13	0.11	0.15
Fiabilidad (alfa de Cronbach)	-0.04	-1.14	0.28	0.17	0.43
Error típico de medida	2.77	2.98	3.10	2.74	3.07

Fuente: elaboración propia.

Tabla 13. Resultados psicométricos. Sociología

Característica	Área 1	Área 2	Área 3	Área 4
Número medio de aciertos	18.64	14.93	12.43	7.79
Error típico de la media	1.04	1.04	0.70	0.49
Varianza	14.09	14.07	6.39	3.17
Desviación típica	3.75	3.75	2.53	1.78
Facilidad media	0.50	0.48	0.54	0.52
Correlación punto-biserial media	0.22	0.27	0.25	0.25
Fiabilidad (alfa de Cronbach)	0.53	0.61	0.34	0.08
Error típico de medida	2.57	2.35	2.05	1.71

Fuente: elaboración propia.

Tabla 14. Resultados psicométricos. Traducción

Característica	Área 1	Área 2	Área 3	Área 4
Número medio de aciertos	54.75	20.56	8.88	6.06
Error típico de la media	2.44	2.17	1.19	1.08
Varianza	184.44	146.50	44.17	36.43
Desviación típica	13.58	12.10	6.65	6.04
Facilidad media	0.38	0.41	0.30	0.22
Correlación punto-biserial media	0.21	0.52	0.50	0.53
Fiabilidad (alfa de Cronbach)	0.86	0.95	0.91	0.91
Error típico de medida	5.05	2.81	2.00	1.76

Fuente: elaboración propia.

Como muestran los resultados, algunos instrumentos presentan datos empíricos mucho más favorables que otros. En cualquier caso, es importante recalcar que estos resultados reflejan sólo un primer acercamiento, y que deberán ser complementados con más evidencia empírica que permita, a su vez, valorar la estructura factorial de cada instrumento. A partir de estos resultados, se realizó un proceso de depuración de reactivos con características poco favorables, como: baja correlación biserial-puntual, dificultad menor a .10 y mayor a .90.

Adicionalmente, con la intención de discutir sobre la base científica de la evaluación realizada (Stake, 2006), los resultados obtenidos se revisaron en lo particular con cada grupo de académicos. En el caso de Bioingeniería y Energías renovables se identificaron problemas conceptuales en los reactivos planteados, lo cual derivó en el ajuste de la estructura; en el caso de Oceanología se depuraron problemas en las imágenes y gráficas presentadas en el piloto. Como resultado, Bioingeniería quedó con tres áreas (en lugar de cuatro, como inicialmente se planteó), al igual que Energías renovables, que quedó con tres áreas (en lugar de las siete iniciales). Estos ajustes se realizaron de manera colegiada con los académicos participantes, atendiendo, así, a la búsqueda de instrumentos acordes con las necesidades y perspectivas de las propias comunidades beneficiarias del ejercicio evaluativo, en equilibrio con los resultados psicométricos de la aplicación piloto.

En el proceso de identificación de problemas en cada instrumento fue vital la consideración de las diversas comunidades que participaron, ya que su experiencia en el programa, su conocimiento de la población estudiantil, así como la sensibilidad en materia de evaluación que desarrollaron durante los dos años del proyecto permitió la identificación de problemas de planteamiento que se ajustaron con la experiencia del pilotaje. Si bien ello no sustituye la necesidad de mayor evidencia

empírica, resulta una aproximación valiosa para la consolidación de la propuesta de evaluación, sobre todo bajo un modelo comprensivo. Con estos ajustes, no sólo se procuraba la precisión del instrumento, sino que éste mismo fuera una herramienta capaz de dar sentido a las comunidades para su utilización con fines formativos, con el propósito de acercar las métricas obtenidas a una interpretación colectiva de los resultados obtenidos.

Fase 5. Definición de puntos de corte y emisión de resultados

Dada la experiencia institucional suscitada con la evaluación del egreso de la LAFyD, se retomó la importancia de utilizar un método que permitiera: a) considerar a cada área del examen como relativamente independiente; b) resaltar la visión de las comunidades académicas participantes (como medio de reflejo de las realidades implicadas en cada programa educativo); y c) los resultados empíricos de la aplicación piloto. Derivado ello, se optó por el uso del método Angoff (Cizek, 2011), perspectiva que permite tomar como insumo principal los resultados empíricos, pero que también da un valor preponderante a la visión de los expertos de las comunidades académicas.

En coincidencia con este método, se invitó a las comunidades académicas a identificar, para cada reactivo y de manera individual, la proporción de estudiantes que idealmente darían respuesta correcta. Se realizó de esta forma con la finalidad de establecer de manera ideal al “sustentante mínimamente competente”³ (Cizek, 2011), con el fin de establecer un estándar de competencia mínimo a lo largo del instrumento. Se seleccionó esta metodología debido a que permite el aprovechamiento del conocimiento disciplinar y la experiencia pedagógica de los comités para la determinación de un estándar acorde con las características propias de estudiantes de la UABC; es decir, se dio énfasis al juicio analítico

³ En inglés: *minimally competent examinee*.

dado por las comunidades académicas. El resultado conjunto de los juicios de los participantes se consideró el punto de corte de cada una de las áreas del instrumento en términos de porcentaje de respuestas correctas. Posterior a este ejercicio, se discutieron los resultados y se establecieron condicionantes para otorgar un resultado global; esto se hizo con la intención de contar con un resultado que permitiera a la institución la comprensión relativamente inmediata del desempeño mostrado por cada estudiante. La clasificación genérica de dicho resultado es la siguiente:

- Insuficiente. Los sustentantes cuentan con habilidades y conocimientos elementales relativos a la formación profesional. Implica considerar robustecer la formación recibida en algunas de las áreas evaluadas.

- Competente. Los sustentantes muestran conocimientos y habilidades esperados para la formación en algunas de las áreas que comprende la prueba.
- Avanzado. Los sustentantes muestran conocimientos y habilidades esperados de la formación profesional en la mayoría de las áreas que comprende la prueba.

De esta forma, cada grupo de participantes estableció, de manera colegiada, el mínimo de áreas en las que cada estudiante necesitaba superar el punto de corte para obtener un resultado general en particular. La Tabla 15 muestra el resumen de los puntos de corte de cada área/instrumento (expresados en porcentaje de aciertos), así como las condicionantes para el otorgamiento del resultado general.

Tabla 15. Resumen de puntos de corte y condicionantes de resultados

Instrumento	Puntos de corte					Condicionantes*		
	Área 1	Área 2	Área 3	Área 4	Área 5	Insuficiente	Competente	Avanzado
Artes plásticas	70	70	70	70	—	0-1	2-3	4
Bioingeniería	78.75	70	67.5	—	—	0-1	2	3
Docencia de la lengua y literatura	55	70	75	70	—	0-1	2-3	4
Docencia de la matemática	85	85	72.5	80	57.5	0-2	3	4-5
Energías renovables	60	60	72.5	—	—	0-1	2	3
Enseñanza de lenguas	90.5	75	85.75	75	—	0-1	2	3-4
Oceanología	77.5	82.5	82.5	90	85	0-1	2-3	4-5
Sociología	73.75	75	80	80	—	0-2	3	4
Traducción	70	80	80	50	80	0-1	2-3	4-5

*Áreas que cada estudiante necesita acreditar para cada nivel de competencia.

Fuente: elaboración propia.

A partir de estas decisiones se consideraron dos acciones. En la primera de ellas, se envió a los académicos que participaron en el proyecto un resumen del porcentaje de aciertos de todos los estudiantes en cada una de las

áreas del examen en la que participó, así como el resultado global en cada caso; esta información se proporcionó a la UABC para su difusión entre los responsables de cada facultad y programa educativo. Se espera que esto sea

un insumo para la reflexión en materia curricular y didáctica, en función de los logros esperados declarados en el documento con descripción conceptual de la disciplina, así como en el propio plan de estudio.

Como una segunda acción se elaboró un formato de resultados dirigido a cada estudiante, el cual le permite analizar el logro alcanzado en función de los criterios que delimita cada instrumento. A través de este formato (Fig. 2), se da retroalimentación a cada estudiante y se dejan en claro los alcances y

limitaciones del instrumento para evitar conclusiones radicales que suelen caracterizar a las evaluaciones preordenadas (Stake, 2006; López-Meseguer y Valdés, 2020). En este sentido, resulta vital el acompañamiento docente en la interpretación y análisis de los resultados obtenidos, no sólo para que los estudiantes tengan un formato con resultados, sino para que también cuenten con una guía de mejora de la mano de sus docentes y los responsables de los programas educativos.

Figura 2. Ejemplo de formato de resultados para estudiantes participantes



Universidad Autónoma de Baja California
Coordinación General de Formación Profesional

XXXXXXXXXX
Facultad de Pedagogía e Innovación Educativa

Esta evaluación tuvo por objeto identificar el logro de los egresados sobre las competencias establecidas en el plan de estudios de la Licenciatura en Docencia de la Lengua y Literatura. Para esta valoración se realizó un juicio sobre tus respuestas respecto a problemas típicos de la formación en docencia de la lengua y literatura, organizados en cuatro áreas de competencia profesional. Los resultados no muestran todo tu posible desempeño como profesional, sólo una muestra de tus habilidades, de acuerdo con lo que establece el perfil de egreso del programa educativo de Licenciatura en Docencia de la Lengua y Literatura.

Los resultados muestran el porcentaje de aciertos obtenido en cada área de la prueba, así como un resultado global que pretende reflejar el nivel alcanzado respecto al nivel mínimo esperado por el programa de Licenciatura en Docencia de la Lengua y Literatura.

Resultados por área

Área de ejercicio profesional	Porcentaje de aciertos obtenido	Porcentaje de aciertos mínimo esperado ¹
1. Desarrollar habilidades discursivas orales y escritas.	X	55
2. Dominar contenidos canónicos y representativos de la literatura universal.	X	70
3. Reconocer elementos contextuales para la implementación de estrategias didácticas.	X	75
4. Emplear Modelos de la Didáctica de la Disciplina para el fortalecimiento de la enseñanza de la lengua y la literatura.	X	70
Resultado global	X	

Las condiciones de obtención que determinaron tu resultado global son las siguientes:

Número de áreas en las que se requiere pasar el mínimo de aciertos ¹	
Insuficiente	0 o 1
Competente	2 o 3
Avanzado	4

Los resultados mostrados no pretenden ser determinantes de tu formación profesional, sino un insumo de reflexión sobre tu propio avance. Te invitamos a analizar tus resultados con apoyo de tus docentes, valorando tus áreas de interés profesional y otras habilidades desarrolladas durante tu formación.

Muchas gracias por tu participación y felicidades por el esfuerzo mostrado en este ejercicio. Tu esfuerzo y resultados son de gran relevancia para retroalimentar el plan de estudios de la Licenciatura en Docencia de la Lengua y Literatura.

¹ Definido por el comité académico del programa de Licenciatura en Docencia de la Lengua y Literatura comisionado para la emisión de puntos de corte, el día 16 de noviembre de 2022.



Licenciatura en Docencia de la Lengua y Literatura
Evaluación del egreso
Reporte individual de resultados

Fuente: UABC (2022).

Se pretende que este ejercicio evaluativo aporte a varias líneas de acción, entre las que se encuentran:

- Ser un insumo para la revisión colegiada de los reactivos con propiedades psicométricas poco favorables, con el fin de valorar su pilotaje con ajustes.
- Considerarla como información relevante para hacer una revisión colegiada de los resultados obtenidos, con la intención de identificar acciones de mejora curricular, áreas de oportunidad en materia de actualización docente y posibles espacios de mejora en los PUA.

A la par, corresponde a la universidad planificar la revisión y consolidación de los instrumentos, a través de la elaboración de nuevos reactivos para pilotear y crear versiones paralelas, recolección de mayor evidencia empírica para robustecer las evidencias de validez y confiabilidad, así como considerar posibles acciones de metaevaluación de este ejercicio, no sólo en materia psicométrica, sino en función de sus alcances pedagógicos y utilidad formativa.

CONCLUSIONES

La experiencia permite valorar el esfuerzo de una institución pública para la obtención de información que permita reflexionar sobre sus procesos formativos en diversas disciplinas. La UABC ha establecido como una prioridad la evaluación del nivel de aprendizaje de sus egresados de nivel licenciatura; en coincidencia, y con la valiosa participación de más de 150 docentes, se logró la implementación de un proyecto ambicioso que aporta información institucional que refleja la formación profesional que se vive en sus aulas. Los resultados del estudio son una muestra del aprendizaje alcanzado a lo largo de varios años de formación respecto de las consideraciones pedagógicas establecidas en los planes de estudio, con

las particularidades formativas que se presentan en las diferentes facultades de la UABC.

El proyecto se enfocó en rescatar las recomendaciones internacionales en materia de construcción de instrumentos, por lo que siempre se dio un lugar preponderante al acopio de aquella información que pudiera utilizarse de forma comprensiva, a favor de la reflexión sobre los programas educativos, así como de los resultados de los estudiantes al concluir su formación profesional. En este recorrido se enfrentaron problemáticas de orden metodológico, principalmente debido a que los programas se caracterizan por ser de baja matrícula, lo cual pone en un predicamento cualquier análisis de corte cuantitativo. Esto derivó en la sensibilización constante de las comunidades respecto a las limitaciones de las inferencias que se asocian a los resultados emitidos. A pesar de ello, resulta un ejercicio inicial valioso, que permite retomar experiencias académicas e institucionales previas en esta materia, sin olvidar la enorme tarea de organización entre todas las facultades involucradas —junto con la participación de su personal docente—, que permitió establecer un ejercicio sin precedentes en México, al menos en universidades públicas estatales, así como para las diferentes disciplinas participantes. De esta manera, la UABC no sólo avanzó en la satisfacción interna de sus necesidades evaluativas, sino que también ofreció —a la comunidad académica interesada en el ámbito de la evaluación educativa— una práctica bajo una mirada comprensiva, así como un ejercicio inédito para disciplinas sin experiencia de este tipo.

Con la intención de mantenerla como una evaluación comprensiva, las labores que se requieren a partir de este ejercicio deben orientarse hacia la retroalimentación de los diferentes programas y comunidades, de tal suerte que los resultados obtenidos sean un insumo para robustecer la formación profesional que se ofrece en la UABC. Esta intención no deberá considerarse completa con el solo

envío de la información a las autoridades de la universidad y las respectivas facultades, sino que se deberán buscar mecanismos para ajustar los PUA y planes de estudio, así como generar espacios de discusión con la comunidad docente, sin olvidar el establecimiento de mecanismos de retroalimentación a los estudiantes participantes. Lograr esta articulación permitirá consolidar la intención comprensiva del estudio, pero, sobre todo, generar insumos valiosos de reflexión que aporten a la transformación constante en materia formativa dentro de la UABC.

Por otra parte, además del potencial en materia investigativa que representa este proyecto es importante realizar un ejercicio de metaevaluación que permita identificar las mejoras conceptuales, metodológicas y operativas necesarias para ejercicios de esta naturaleza. Así, la experiencia pretende, a su vez, abonar a una discusión teórico-metodológica en materia de

evaluación del egreso. En este sentido, algunas interrogantes que quedan pendientes de analizar son, por ejemplo: ¿cuáles son las mejores experiencias y perspectivas para el uso de la información generada?, ¿de qué manera los resultados pueden abonar a robustecer las modificaciones de los diversos planes de estudio?, ¿existen posibilidades de abonar, en lo inmediato, a la formación de los estudiantes que se encuentran cursando los diversos programas educativos?, ¿existen otras estrategias metodológicas que faciliten la evaluación del egreso para poblaciones de baja matrícula?

Estas interrogantes representan una oportunidad para buscar estrategias metodológicas para la mejora de la enseñanza en educación superior pública, a través del refinamiento de procesos de evaluación más comprensivos, así como de definición de acciones de retroalimentación a profundidad para los programas educativos.

REFERENCIAS

- AGUILAR Vera, R. y J. Díaz Mendoza (2021), "Análisis del desempeño en la última década de los egresados en Ciencias de la Computación de la FMAT-UADY en el EGEL-COMPU", *Tecnología Educativa Revista CONAIC*, vol. 8, núm. 3, pp. 131-142.
- ALCARAZ, Noelia (2015), "Aproximación histórica a la evaluación educativa: de la generación de la medición a la generación ecléctica", *Revista Iberoamericana de Evaluación Educativa*, vol. 8, núm. 1, pp. 11-25.
- American Educational Research Association (AREA)/ American Psychological Association (APA)/ National Council on Measurement in Education (NCME) (2014), *Standards for the Educational and Psychological Testing*, Washington, AERA.
- BAZÁN, Jorge (2004), "Introducción al modelo psicométrico de la teoría clásica de los tests (parte 1)", *Pro Mathematica*, vol. 18, núm. 35-36, pp. 79-107.
- CASTILLO, Katherine (2012), "La posible aplicación del modelo explicativo nomológico-deductivo de Carl Hempel a los hechos institucionales de John Searle", *Revista Disertaciones*, vol. 3, núm. 3, pp. 104-115.
- Centro Nacional para la Evaluación de la Educación Superior (Ceneval) (2023), *Examen General para el Egreso de la Licenciatura*, México, Ceneval, en: <https://ceneval.edu.mx/examenes-egreso-egel/> (consulta: 14 de agosto de 2023).
- CIZEK, Gregory (2011), "Standard Setting", en Steven Downing y Thomas Haladyna (eds.), *Handbook of Test Development*, Londres, Lawrence Erlbaum Associates, pp. 225-258.
- CORNEJO, Fernando, Sergio Franco y Claudia Islas (2020), "Análisis descriptivo del examen EGEL-COMPU y su testimonio de desempeño", *Transdigital*, vol. 4, núm. 7, pp. 1-16. DOI: <https://doi.org/10.56162/transdigital154>
- DELGADO-Maldonado, Laura y Melchor Sánchez-Mendiola (2012), "Análisis del examen profesional de la Facultad de Medicina de la UNAM: una experiencia de evaluación objetiva del aprendizaje con la teoría de respuesta al ítem", *Investigación en Educación Médica*, vol. 1, núm. 3, pp. 130-139.
- GAGO, Antonio (2000), "El Ceneval y la evaluación externa de la educación en México", *Revista Electrónica de Investigación Educativa*, vol. 2, núm. 2, pp. 106-114.

- GARCÍA, Benilde (2010), “Modelos teóricos e indicadores de evaluación educativa”, *Sinéctica*, núm. 35, pp. 1-17, en: <https://sinectica.iteso.mx/index.php/SINECTICA/issue/view/38> (consulta: 14 de agosto de 2023).
- GARCÍA, Moisés, Felipe Martínez, Graciela Cordero y Joaquín Caso (2017), “Evolución del concepto de validez en la medición educativa”, en Edna Luna y Graciela Cordero (coords.), *Contribuciones a la evaluación educativa desde la formación doctoral*, Guadalajara, UdeG/UABC, pp. 15-46.
- HUAMÁN, Ludencino, Teresa Pucuhuaranga y Nora Hilario (2020), “Evaluación del logro del perfil de egreso en grados universitarios: tendencias y desafíos”, *Revista Iberoamericana para la Investigación y el Desarrollo Educativo*, vol. 11, núm. 21. DOI: <https://doi.org/10.23913/ride.v11i21.691>
- Instituto Nacional para la Evaluación de la Educación (INEE) (2017), *Criterios técnicos para el desarrollo, uso y mantenimiento de instrumentos de evaluación*, México, INEE.
- JIMÉNEZ, Alfonso (2020), *Evaluación a gran escala del egreso de la licenciatura en México: construcción de una propuesta bajo un modelo comprensivo*, Mexicali, UABC.
- JIMÉNEZ, Alfonso, Juan Machado, Joaquín Caso y Emilio Arrayales (2019), “Evaluación del egreso de la licenciatura en actividad física y deportes de la UABC: un ejercicio comprensivo”, *Revista Iberoamericana sobre Calidad, Eficacia y Cambio en la Educación*, vol. 17, núm. 2, pp. 123-146, en: <https://revistas.uam.es/index.php/reice/article/view/reice2019.17.2.007/10899> (consulta: 14 de agosto de 2023).
- KRATHWOHL, David (2002), “A Revision of Bloom’s Taxonomy: An overview”, *Theory into Practice*, vol. 41, núm. 4, pp. 212-218. DOI: https://doi.org/10.1207/s15430421tip4104_2
- LÓPEZ-MESEGUER, Rafael y Manuel Valdés (2020), “La evaluación comprensiva de programas educativos: ¿un nuevo paradigma teórico?”, *Revista Iberoamericana de Evaluación Educativa*, vol. 13, núm. 2, pp. 85-105. DOI: <https://doi.org/10.15366/riee2020.13.2.005>
- MÉNDEZ, Saúl, Eva Olivia Martínez, Maricela Sevilla, Ricardo Ching y Adelaida Figueroa (2022), “Análisis de los resultados del EGEL-Ceneval 2017-2019 de la Licenciatura en Informática de la UABC”, *Revista Iberoamericana de Producción Académica y Gestión Educativa*, vol. 9, núm. 17, pp. 84-107.
- MUÑOZ, José (2010), “Las teorías de los tests: teoría clásica y teoría de respuesta a los ítems”, *Papeles del Psicólogo: Revista del Colegio Oficial de Psicólogos*, vol. 31, núm. 1, pp. 57-66.
- STAKE, Robert (2006), *Evaluación comprensiva y evaluación basada en estándares*, Madrid, Graó.
- Universidad Autónoma de Baja California (UABC) (2018), *Estatuto escolar de la Universidad Autónoma de Baja California*, Mexicali, UABC.
- Universidad Autónoma de Baja California (UABC) (2022), *Guía del sustentante del examen de egreso de la Licenciatura en Sociología*, Mexicali, UABC.
- Universidad Autónoma de Baja California (UABC) (2023), *Resultados de exámenes de egreso. Documento interno*, Mexicali, UABC.
- VELÁZQUEZ-NÚÑEZ, Jorge-Javier, Sandra-Eloína Campos-López y María-Magdalena Pelayo-Cortés (2019), “Habilidades profesionales de los alumnos de la Licenciatura en Contaduría Pública del CUCSUR-UDG en el área de finanzas y examen EGEL-Ceneval”, *Contaduría. Universidad de Antioquia*, núm. 75, pp. 115-137. DOI: <https://doi.org/10.17533/udea.rc.n75a05>
- VERA, Héctor y Miguel González-Ledesma (2018), “Calidad y evaluación: matrimonio del cielo y el infierno”, *Perfiles Educativos*, vol. 40, núm. especial, pp. 53-97. DOI: <https://doi.org/10.22201/iisue.24486167e.2018.especial.59180>
- VIDAL, Rafael (2013), “Measurement of Learning Outcomes in Higher Education: The case of Ceneval in Mexico”, en Sigrid Blömeke, Olga Zlatkin-Troitschanskaia, Christiane Khun y Judith Fege (eds.), *Modeling and Measuring Competencies in Higher Education: Tasks and challenges*, Rotterdam, Sense, pp. 137-146.