



METODOLOGÍAS PARA LA EVALUACIÓN DE IMPACTO EN LA EDUCACIÓN SUPERIOR

Handbook n°2

Mayra Alejandra Lozano Rodríguez
Luisa Fernanda Rojas Barreto
Jesús Gabalán Coello
Claudia Aponte González
Gonzalo Raúl Fonseca Grandón
Miriam Camacho Zenteno



OEI



METODOLOGÍAS PARA LA EVALUACIÓN DE IMPACTO EN LA EDUCACIÓN SUPERIOR

Handbook n°2

Mayra Alejandra Lozano Rodríguez
Luisa Fernanda Rojas Barreto
Jesús Gabalán Coello
Claudia Aponte González
Gonzalo Raúl Fonseca Grandón
Miriam Camacho Zenteno



© 2026 Universidad Privada del Valle – Univalle
Todos los derechos reservados.

Queda prohibida la reproducción total o parcial de esta obra,
por cualquier medio o procedimiento, sin autorización previa
y por escrito del titular de los derechos.

Título: METODOLOGÍAS PARA LA EVALUACIÓN DE IMPACTO EN LA EDUCACIÓN SUPERIOR
Handbook No. 2
Primera edición
Obra elaborada por la Universidad Privada del Valle – UNIVALLE

ISBN: 978-99974-349-2-0
Depósito Legal: 2-1-3987-2026

Publicado por: Universidad Privada del Valle – Univalle
Cochabamba, Bolivia

AUTORES

Mayra Alejandra Lozano Rodríguez¹
Luisa Fernanda Rojas Barreto²
Jesús Gabalán Coello³
Claudia Aponte González⁴
Gonzalo Raúl Fonseca Grandón⁵
Miriam Camacho Zenteno⁶

Ing. Pamela Gil
**VICERRECTORA SEDE ACADÉMICA
SANTA CRUZ**

M.B.A. Carlos Torricos Mérida
**VICERRECTOR SEDE ACADÉMICA
SUCRE**

ORGANIZACIÓN DE ESTADOS IBEROAMERICANOS

Donatella Montaldo
DIRECTORA OEI COLOMBIA

Claudia Johana Parra Pulido
ESPECIALISTA DE CTEI

Rodolfo Barrere
**COORDINADOR DEL
OBSERVATORIO DE CIENCIA,
TECNOLOGÍA Y SOCIEDAD, OEI
ARGENTINA**

M.Sc. Franklin Nestor Rada
**VICERRECTOR SEDE ACADÉMICA LA
PAZ**

M.Sc. Martha Mejía Fayer
**VICERRECTORA SEDE ACADÉMICA
TRINIDAD**

Sandro Guerrero Morales Ph.D.
VICERRECTOR DE POSTGRADO

M.Sc. Jorge Carlos Ruiz De la Quintana
**DIRECTOR NACIONAL DE
INVESTIGACIÓN**

Jaime Rodríguez Coariti Ph.D.
**DIRECTOR DE INVESTIGACIÓN
SEDE LA PAZ**

AUTORIDADES DE LA UNIVERSIDAD PRIVADA DEL VALLE

M.Sc. Gonzalo Vicente Ruiz Ostría
RECTOR NACIONAL

Diego Alonso Villegas Zamora, Ph.D.
**VICERRECTOR ACADÉMICO
NACIONAL**

M.Sc. Sandra Marcela Ruiz Ostría
**VICERRECTORA INTERACCIÓN
SOCIAL Y DIFUSIÓN CULTURAL**

Lic. Daniela Zambrana Grandy
SECRETARIA GENERAL

EDICIÓN DIAGRAMACIÓN:

Ing. Giubell Melanie Mercado Franco
**COORDINADORA DE DIFUSIÓN
CIENTÍFICA Y FORMACIÓN EN
INVESTIGACIÓN**

M.Sc. Luis Marco Fernández Sandoval
**COORDINADOR DE PRODUCCION
AUDIOVISUAL Y DIGITAL**

¹ Director Corporación PENSER (Washington, USA – Colombia), direccion@penser.org, <https://orcid.org/0000-0001-7674-8849>

² Directora General de Investigación Educativa Corporación PENSER (Washington, USA – Colombia), dirinvest@penser.org, <https://orcid.org/0000-0001-8780-3882>

³ Profesional Contratista en la Dirección de Información y Estadísticas de la Secretaría Distrital de Planeación (Colombia), lurojas.barreto@gmail.com, <https://orcid.org/0000-0001-8820-876X>

⁴ Académico Facultad de Educación, Universidad de Concepción (Chile), gfonseca@udec.cl, <https://orcid.org/0000-0003-1063-1319>

⁵ Consultora en Educación Superior (Colombia), apontemac@gmail.com, <https://orcid.org/0000-0002-8163-4499>

⁶ Directora Nacional de Pedagogía, Universidad del Valle (Bolivia), mcamachoz@univalle.edu, <https://orcid.org/009-001-6088-8879>

PRÓLOGO

Donatella Montaldo, Directora de OEI Colombia

Para la Organización de Estados Iberoamericanos, es un honor colaborar en iniciativas como el Proyecto de Investigación Transnacional “Construcción de Referentes Metodológicos para la Evaluación de Impacto del Perfil de Egreso en Programas de América Latina”, cuyos aportes enriquecieron la comprensión de las realidades regionales y favorecieron la construcción de una mirada integral sobre el impacto, basándose en la colaboración de 18 universidades de 6 países de Iberoamérica.

Este proyecto constituye un ejemplo del valor que la cooperación científica cuando se concibe como un espacio de intercambio, aprendizaje mutuo y construcción conjunta, a partir de la generación de herramientas que permitan fortalecer la calidad de la educación superior y el acceso al conocimiento, la innovación institucional, y la integración regional fomentando una educación superior más pertinente y articulada con las necesidades sociales y de los países iberoamericanos.

En esta línea, además de ofrecer un marco conceptual sólido, estos handbooks brindan orientaciones operativas que permiten a las instituciones contextualizar, adaptar y aplicar los referentes metodológicos en sus propios procesos de aseguramiento de la calidad. La sistematización de prácticas, la integración de evidencias y la claridad de las rutas metodológicas aquí presentadas constituyen un aporte significativo para quienes lideran procesos de autoevaluación, acreditación, rediseño curricular, internacionalización y proyección social o vinculación con el medio.

Así, estos handbooks no solo compilan conocimiento, sino que lo transforman en una guía accesible, adaptable y científicamente fundamentada para fortalecer la toma de decisiones estratégicas en el ámbito de la educación superior.

En conjunto, esta obra invita a avanzar hacia una cultura de evaluación más rigurosa, reflexiva e innovadora, que reconozca el valor del impacto como criterio central para garantizar la pertinencia social, la calidad académica y el compromiso de nuestras universidades con el desarrollo de la región. Asimismo, reafirma el potencial de la cooperación iberoamericana como motor de transformación educativa y como una vía para construir un espacio común de educación superior más sólido, colaborativo y orientado al bienestar de nuestras sociedades

CONTENIDO

1. Introducción.....	1
2. Glosario.....	2
3. Metodologías para la evaluación de impacto en la educación superior.....	4
3.1. Tema 1. Enfoques teóricos y conceptuales en la evaluación de impacto.....	4
3.1.1. Hacia una definición del impacto de la educación superior.....	6
3.1.2. ¿Cómo evaluar el impacto de la educación superior?.....	6
3.2. Tema 2. Métodos cuantitativos experimentales y cuasi-experimentales para la evaluación de impacto.....	7
3.2.1. Métodos cuantitativos experimentales.....	8
3.2.1.1. Randomized Controlled Trial - RCT.....	8
3.2.1.1.1. ¿Cómo se determina el número requerido de participantes en un estudio aleatorizado?.....	9
3.2.2. Métodos cuasiexperimentales.....	10
3.2.2.1. Diferencias en Diferencias (DiD - Differences-in-Differences).....	10
3.2.2.2. Emparejamiento por Puntaje de Propensión (Propensity Score Matching - PSM).....	11
3.2.2.3. Regresión Discontinua (Regression Discontinuity Design - RDD).....	12
3.2.2.4. Variables Instrumentales (Instrumental Variables - IV).....	12
3.2.2.5. Serie Temporal Interrumpida (Interrupted Time Series - ITS).....	14
3.2.2.6. Matching Exacto o Coarsened Exact Matching (CEM).....	15
3.2.3. Instrumentos cuantitativos.....	18
3.3. Tema 3. Métodos cualitativos en la evaluación de impacto.....	18
3.3.1. Instrumentos de evaluación cualitativa.....	19
3.3.2. ¿Cómo se determina el número requerido de participantes en un estudio no aleatorizado?.....	20
3.3.3. Proceso del Método Delphi.....	22
3.4. Tema 4. Métodos mixtos.....	22
3.4.1. Tipos de Diseños Mixtos.....	23
3.5. Teoría de Cambio.....	23
3.6. Tema 5. Instrumentos y herramientas digitales para la evaluación de impacto.....	25
4. Actividades Prácticas.....	27
4.1. Caso 1.....	27
4.2. Caso 2.....	28
4.3. Caso 3.....	29
4.4. Caso 4.....	30
5. Recursos Complementarios.....	32
6. Evaluación.....	33
7. Referencias Bibliográficas.....	36

Lista de tablas:

Tabla 1. Dimensiones del Impacto Social.....	5
Tabla 2. Métodos cuasiexperimentales.....	16
Tabla 3. Instrumentos cuantitativos para la evaluación de impacto.....	18
Tabla 4. Instrumentos de evaluación cualitativa.....	19
Tabla 5. Herramientas digitales para la evaluación de impacto.....	
Tabla 6. Caso 1 de evaluación de impacto.....	
Tabla 7. Caso 2 de evaluación de impacto.....	
Tabla 8. Caso 3 de evaluación de impacto.....	
Tabla 9. Caso 4 de evaluación de impacto.....	
Tabla 10. Recursos complementarios para la evaluación de impacto.....	

Lista de gráficas:

Gráfica 1. Definición del contrafactual.....	8
Gráfica 2. Pasos para establecer el número de participantes en un estudio aleatorizado.....	9
Gráfica 3. Pasos para determinar el número de participantes en un estudio no aleatorizado.....	21
Gráfica 4. Pasos de la teoría de cambio y unificar el tipo de letra con la que se está usando el texto.....	25

1. Introducción

La educación superior desempeña un papel fundamental en la formación de individuos y en la transformación de las sociedades, cumpliendo funciones que van desde la creación de capital humano hasta la promoción de la equidad. En primer lugar, forma profesionales con conocimientos especializados, competencias prácticas y habilidades socioemocionales necesarias para su inserción laboral (Luciano & Velázquez, 2022). Además, impulsa la investigación y la innovación como motores del desarrollo económico y social, mediante entornos académicos que favorecen la generación de soluciones a problemas emergentes (Ñáñez-Rodríguez et al., 2019). Igualmente, la educación superior contribuye al desarrollo social y cultural, al formar ciudadanos críticos que promueven la inclusión y los valores democráticos (Gómez et al., 2023). Este compromiso se amplía al escenario tecnológico, en el que las universidades deben adaptarse a los cambios mediante el uso de tecnologías de la información y la formación docente en nuevas metodologías pedagógicas (García-Ruiz et al., 2010). Asimismo, tiene la responsabilidad de contextualizar el conocimiento en relación con su entorno sociocultural, lo que fortalece las capacidades locales y permite una vinculación activa con los desafíos globales. Finalmente, la educación superior debe garantizar la inclusión y la equidad mediante políticas que aseguren el acceso y la permanencia de estudiantes de diversos contextos socioeconómicos (Morales-Maza, 2020), consolidando así su función como agente transformador y democratizador. En este contexto, la evaluación de impacto en la educación superior se configura como una herramienta estratégica que permite valorar, de forma sistemática y rigurosa, los efectos que producen las instituciones universitarias en los individuos, las comunidades y la sociedad en su conjunto.

Este handbook surge como una respuesta a la necesidad creciente de las instituciones, investigadores y formuladores de políticas públicas de contar con marcos metodológicos claros, pertinentes y adaptados a los desafíos contemporáneos de la educación superior. Frente a la expansión de la cobertura, la diversificación de la oferta académica, los procesos de acreditación, entre otros aspectos, la evaluación de impacto cobra un nuevo sentido: ya no se trata solamente de verificar resultados, sino de comprender los procesos de transformación que produce la educación en diversos contextos diversos.

En las últimas décadas, los enfoques evaluativos han avanzado hacia modelos más completos que combinan herramientas cuantitativas y cualitativas, capaces de captar tanto los efectos medibles como las dimensiones más subjetivas y contextuales del cambio. Así, han proliferado estudios que utilizan métodos experimentales, cuasiexperimentales, técnicas econométricas, análisis narrativos, reconstrucción de trayectorias y construcción de indicadores participativos. Este abanico de posibilidades ha enriquecido las perspectivas de análisis, pero también ha generado nuevos retos en términos de rigurosidad metodológica, pertinencia contextual y uso ético de la información.

La evaluación de impacto en educación superior requiere superar enfoques lineales y reduccionistas para dar paso a una mirada multidimensional. Esto implica reconocer que el impacto no es unívoco ni inmediato: se manifiesta en distintos niveles (micro, meso y macro), abarca múltiples ámbitos (educativo, social, económico, cultural) y

responde a lógicas causales complejas. Evaluar impacto significa entonces indagar por la calidad del aprendizaje, la empleabilidad, la equidad en el acceso, la contribución al desarrollo territorial, la transformación de vidas, el fortalecimiento de competencias ciudadanas y la incidencia en políticas públicas.

Este handbook recoge marcos teóricos y metodologías que han demostrado su utilidad en la evaluación de programas, proyectos e intervenciones en educación superior. A lo largo del documento se presentan métodos cuantitativos rigurosos como los diseños aleatorizados controlados (RCT), las diferencias en diferencias (DiD), el Propensity Score Matching (PSM), la regresión discontinua (RDD), las variables instrumentales (IV) y las series temporales interrumpidas (ITS). Asimismo, se profundiza en métodos cualitativos como entrevistas, historias de vida, cartografía social y análisis documental, y se propone la integración de ambas perspectivas mediante diseños de métodos mixtos articulados con la teoría de cambio.

Dirigido a académicos, gestores institucionales, tomadores de decisiones y estudiantes de posgrado, este handbook invita a repensar la evaluación como un proceso reflexivo, colaborativo y situado, que permite valorar no solo lo que las universidades hacen, sino lo que logran transformar en la sociedad.

Los objetivos de aprendizaje que acompañan cada tema permitirán orientar la apropiación práctica y crítica para el diseño e implementación de evaluaciones de impacto que sean útiles para la toma de decisiones, la mejora continua y la rendición de cuentas con sentido ético y transformador. Estos son:

- Comprender los fundamentos conceptuales, las dimensiones y niveles del impacto de la educación superior, y su relación con las funciones misionales de las universidades.
- Identificar los principales métodos cuantitativos (RCT, DiD, PSM, RDD, IV, ITS y CEM) y seleccionar los pertinentes para estimar efectos causales de programas e intervenciones en educación superior, valorando sus ventajas, limitaciones y condiciones de uso.
- Diseñar y emplear estrategias cualitativas (entrevistas, grupos focales, historias de vida, cartografía social, análisis documental, entre otras) para analizar de manera contextualizada los procesos, trayectorias y significados asociados al impacto educativo.
- Integrar enfoques cuantitativos y cualitativos en diseños de métodos mixtos, para evaluar el impacto de las intervenciones en educación superior.
- Seleccionar e implementar herramientas digitales pertinentes para optimizar la recolección, análisis y comunicación de resultados en evaluaciones de impacto.

2. Glosario

Análisis cualitativo: Método de interpretación de datos no numéricos (entrevistas, narrativas, documentos) que busca comprender significados, percepciones y contextos, a través de técnicas como la codificación temática y el análisis narrativo.

Análisis cuantitativo: Proceso de procesamiento estadístico de datos numéricos que permite identificar patrones, asociaciones o efectos causales mediante métodos descriptivos e inferenciales.

Contrafactual: Situación hipotética que representa lo que habría ocurrido con una población objetivo si no hubiera participado en una intervención. Su estimación es fundamental en la evaluación de impacto.

Correlación: Medida que indica el grado de relación entre dos variables. Una correlación alta no implica causalidad, pero sí puede ser un indicio de asociación que merece ser analizada.

Diseño experimental: Método de evaluación en el cual los participantes son asignados aleatoriamente a un grupo de tratamiento o de control, lo que permite estimar efectos causales con alta validez interna.

Diseño cuasiexperimental: Método de evaluación que no utiliza aleatorización, pero sí emplea grupos de comparación o mediciones antes y después de la intervención para estimar efectos causales aproximados.

Evaluación de impacto: Proceso metodológico que busca identificar, medir y atribuir los efectos o cambios producidos por una intervención, programa o política en su población objetivo.

Efecto causal / Causalidad: Relación en la que un cambio en una variable genera un cambio en otra. A diferencia de la correlación, aquí hay implicación de causa y efecto. La causalidad debe ser inferida con métodos específicos.

Error estándar: Medida de la precisión de una estimación. Cuanto menor es el error estándar, más confiable es el valor estimado de un parámetro.

Indicador: Variable cuantitativa o cualitativa que permite medir o describir el estado, avance o efecto de un fenómeno, facilitando la toma de decisiones y la evaluación de resultados.

Intervalo de confianza (IC): Rango dentro del cual se espera que se ubique el valor real de una estimación, con un nivel de confianza (usualmente 95%). Brinda una idea del grado de incertidumbre de los resultados.

Métodos mixtos: Enfoque metodológico que combina técnicas cualitativas y cuantitativas en un mismo estudio para obtener una visión más completa del fenómeno analizado.

Muestreo: Proceso de selección de una parte de la población para recolectar datos representativos o relevantes en una investigación. Puede ser probabilístico (aleatorio) o no probabilístico (intencionado).

Multicolinealidad: Ocurre cuando dos o más variables independientes están altamente correlacionadas entre sí en un modelo de regresión, lo que puede distorsionar los resultados.

Teoría de cambio: Marco lógico y conceptual que describe cómo y por qué se espera que una intervención produzca determinados resultados, a través de una secuencia causal de actividades, productos, resultados e impactos.

Transferencia de conocimiento: Proceso mediante el cual las universidades comunican, adaptan o aplican los resultados de su investigación o innovación en contextos sociales, económicos o institucionales.

Triangulación: Estrategia metodológica que integra diversas fuentes, métodos o enfoques para contrastar y validar hallazgos, fortaleciendo la credibilidad y profundidad del análisis.

Variable dependiente: es el resultado o efecto que se quiere explicar (por ejemplo, tasa de graduación, ingreso laboral, satisfacción estudiantil).

Variable independiente: es el factor que puede estar influyendo en ese resultado (como tipo de programa, edad, género, modalidad de estudio).

3. Metodologías para la evaluación de impacto en la educación superior

3.1. Tema 1. Enfoques teóricos y conceptuales en la evaluación de impacto.

Si bien no existe una definición única y consensuada sobre lo que constituye el impacto de la educación superior, múltiples estudios y aproximaciones teóricas han intentado delimitar sus alcances a partir de su función social y su razón de ser en las sociedades contemporáneas. Estas aproximaciones coinciden en señalar que las universidades deben generar transformaciones que trasciendan lo académico y respondan a las demandas del entorno en diversos ámbitos, niveles y dimensiones.

En este sentido, se han identificado al menos tres grandes categorías desde las cuales puede analizarse el impacto de la educación superior:

- **Impacto educativo**

El impacto educativo se refiere a los efectos que tienen las intervenciones educativas, ya sean programas, políticas o prácticas, en el aprendizaje y desarrollo de habilidades y competencias de los estudiantes. Este concepto abarca una variedad de dimensiones, como la calidad del aprendizaje (Villota et al., 2024; Otero-Saborido & Ramos, 2019; Pastor, 2011), el desarrollo social y emocional de los estudiantes (Arango et al., 2024; Martínez-Padilla & Pérez-González, 2008), la inclusión educativa y el uso de tecnologías en la enseñanza (Estévez et al., 2021; Almeida & Montenegro, 2023).

- **Impacto social**

El impacto social se refiere a los efectos que una intervención, proyecto o programa tiene sobre la sociedad. La caracterización del impacto social puede abordarse desde distintos niveles (Cruz et al., 2020):

- a. Macrosocial: se refiere a los efectos a gran escala que afectan estructuras amplias de la sociedad, como políticas públicas, normativas, marcos institucionales o dinámicas económicas y culturales nacionales.
- b. Mesosocial: corresponde a los efectos a nivel institucional o comunitario, donde se evidencian cambios en organizaciones, colectivos o grupos intermedios.
- c. Microsocial: hace referencia a los efectos en individuos o familias, en sus prácticas, actitudes, conocimientos, relaciones o calidad de vida.

Así mismo, el impacto puede producirse en diferentes dimensiones (Vanclay, F. 2003):

Tabla 1. Dimensiones del Impacto Social

Dimensiones	Descripción
Forma de vida de las personas	Cómo viven, trabajan, juegan e interactúan unas con otras en el quehacer cotidiano
Cultura	Sus creencias, costumbres, valores e idioma o dialecto
Su comunidad	Su cohesión, estabilidad, carácter, servicios e instalaciones
Sus sistemas políticos	El grado al que las personas pueden participar en las decisiones que afectan sus vidas, el nivel de democratización que está teniendo lugar y los recursos suministrados para ese fin.
Su entorno	La calidad del aire y el agua que utiliza la población, la disponibilidad y calidad de los alimentos que consume, el nivel de peligro o riesgo, polvo y ruido al que está expuesta, la idoneidad del saneamiento, su seguridad física y su acceso a y control sobre los recursos.
Su salud y bienestar	Desde el punto de vista físico, mental, social y espiritual, y no solamente la ausencia de enfermedad
Sus derechos	Tanto personales como a la propiedad; especialmente si las personas se ven económicamente afectadas o si sufren desventajas personales que pueden incluir la violación de sus libertades civiles.
Sus temores y aspiraciones	Sus percepciones acerca de su propia seguridad, sus temores acerca del futuro de su comunidad y sus aspiraciones tanto en lo que respecta a su propio futuro como al de sus hijos.

Fuente: (Vanclay, F. 2003)

Por último, el Banco Mundial (2003) define la evaluación de impacto como la medición de los cambios en el bienestar de los individuos que pueden ser atribuidos a un programa o a una política específica. Su propósito general es determinar la efectividad de las políticas, programas o proyectos ejecutados.

Desde una perspectiva investigativa, es fundamental que las evaluaciones de impacto social definan de manera precisa tanto el nivel (macro, meso o micro) como las dimensiones específicas (bienestar, calidad de vida, vida cotidiana, cultura, salud, derechos, entre otras) en las que se espera generar efectos. Esta doble caracterización permite construir marcos de análisis más robustos, establecer indicadores adecuados

y comprender con mayor profundidad la magnitud y naturaleza de los cambios provocados por una intervención. Asimismo, esta claridad favorece la toma de decisiones basada en evidencia, el diseño de políticas públicas y la mejora continua de programas sociales y educativos.

- **El impacto a través de la transferencia de conocimiento**

El impacto a través de la transferencia de conocimiento se concreta en los procesos de articulación entre la universidad, la empresa y el Estado, mediante la investigación aplicada, la innovación, la consultoría y otras formas de extensión que promuevan el desarrollo sostenible y la solución de problemáticas sociales relevantes.

Según Castillo et al., una gestión efectiva del conocimiento implica no solo su adquisición, sino también su transferibilidad hacia el sector empresarial y el Estado, lo que es esencial para incrementar la competitividad en un entorno global en constante cambio (Castillo et al., 2024). Esto implica que las universidades deben formular políticas que fomenten la investigación aplicada y la colaboración con la industria, asegurando que los resultados de sus investigaciones sean utilizados en la práctica (Castillo et al., 2024).

Además, la empresa tiene un papel activo en este proceso. Las organizaciones deben estar dispuestas a colaborar con las universidades para aprovechar el conocimiento académico. Según Bernal et al., existe una necesidad crítica de establecer una comunicación fluida entre las universidades y las empresas, facilitando así que los proyectos de investigación se alineen con las necesidades del sector (Bernal et al., 2023). En este contexto, no solo se producen beneficios mutuos, sino que también contribuye a la capacitación de profesionales calificados que pueden servir mejor a la economía (Flores-Cortez et al., 2023). Así, la transferencia de conocimiento se convierte en un puente que conecta la teoría con la práctica, optimizando la formación y el aprendizaje continuado (Castillo et al., 2024).

3.1.1 Hacia una definición del impacto de la educación superior

El impacto de la educación superior puede entenderse como el conjunto de transformaciones significativas que las instituciones universitarias generan en los individuos, las comunidades y la sociedad en su conjunto, como resultado de sus funciones. Se reconoce que dicho impacto trasciende lo estrictamente académico y se manifiesta en distintos ámbitos (educativo, social y de transferencia de conocimiento), niveles (micro, meso y macro) y dimensiones (como la calidad de vida, la participación, la cultura, la salud, entre otras). Estas transformaciones evidencian la capacidad de las universidades para responder a las necesidades del entorno, promover el desarrollo humano sostenible y contribuir al mejoramiento del bienestar, desde su razón de ser como actor social.

3.1.2. ¿Cómo evaluar el impacto de la educación superior?

La evaluación de impacto implica la adopción de metodologías rigurosas que permiten una comprensión más amplia y profunda de los efectos de la educación superior en

distintos niveles. Entre estas metodologías, se destacan los métodos cuantitativos experimentales y cuasi-experimentales, que permiten establecer relaciones causales entre la participación en programas educativos y ciertos resultados observables, mediante el uso de grupos de control, diseños longitudinales y técnicas estadísticas avanzadas. Complementariamente, los métodos cualitativos ofrecen una mirada comprensiva de las trayectorias educativas y vitales, al explorar las percepciones, motivaciones, barreras estructurales y transformaciones subjetivas que experimentan los estudiantes y egresados.

En el ámbito de la educación superior, se han desarrollado metodologías que integran estos enfoques cualitativos y cuantitativos para capturar las diversas dimensiones del impacto, abarcando desde indicadores sociales hasta evaluaciones internas de desempeño académico (Rocha et al., 2020; Rivera et al., 2024). Las universidades han reconocido la necesidad de incorporar estos métodos de evaluación para fortalecer su misión educativa y responder a las demandas sociales contemporáneas (Espinosa et al., 2023). Por ello, se destaca la importancia de utilizar instrumentos analíticos robustos que logren reflejar la complejidad inherente al ejercicio educativo superior.

Los modelos de evaluación de impacto se sustentan en una base teórica que combina elementos de gestión de la calidad y análisis social, lo que permite una aproximación multidimensional a los procesos educativos (Rocha et al., 2020; Cely-Salazar & Urquijo, 2022). De esta manera, la evaluación del impacto también incluye un análisis de los procesos de autorregulación y mejoras internas (Olmos-Gómez et al., 2022; Delgado & Reyes, 2022).

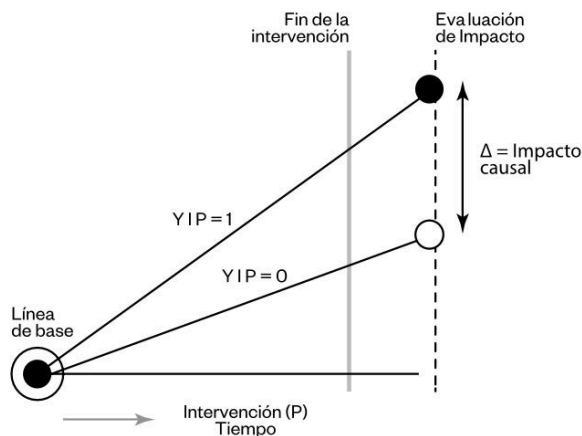
3.2. Tema 2. Métodos cuantitativos experimentales y cuasi-experimentales para la evaluación de impacto.

Una evaluación de impacto permite determinar si un programa o política educativa genera un efecto causal en una o varias variables de interés. Es decir, busca establecer si los resultados observados pueden atribuirse directamente a la intervención educativa implementada.

El reto fundamental radica en que solo se puede observar lo que realmente ocurrió, sin acceso a lo que habría sucedido en ausencia del programa, lo que se conoce como el contrafactual. Comprender este contrafactual es esencial, ya que el impacto de la intervención se define como la diferencia entre el resultado observado con el programa y el resultado hipotético sin él. (Cardona-Arias, 2020).

Dado que no es posible observar directamente el contrafactual, la evaluación de impacto debe construir una estimación de este escenario no observado. Para ello, se utiliza comúnmente un grupo de control o de comparación, compuesto por personas o entidades que no participaron en el programa. Este grupo sirve como referencia para contrastar los resultados con los del grupo de tratamiento, conformado por quienes sí participaron en la intervención.

Grafica 1. Definición del contractual



Fuente: Bayona-Umbarila, et. al, 2022

3.2.1. Métodos cuantitativos experimentales

El diseño experimental se caracteriza por la asignación aleatoria de participantes a grupos de intervención y control, lo que permite controlar las variables extrañas y garantizar la validez interna del estudio (Ato et al., 2013; Longva & Foss, 2018). Esta aleatorización es útil para comprobar hipótesis causales con rigurosidad, eliminando sesgos sistemáticos derivados del proceso de asignación (Ato et al., 2013; Longva & Foss, 2018).

3.2.1.1 Randomized Controlled Trial - RCT

El Experimento Aleatorizado Controlado (RCT) es considerado una de las metodologías más rigurosas para evaluar la efectividad de intervenciones en diversos campos. Este método se basa en la asignación aleatoria de los participantes a grupos de tratamiento y control, lo que minimiza sesgos y permite establecer relaciones causales con mayor claridad (Rey-Fuentes et al., 2021).

Una de las características destacadas de los RCT es su diseño, el cual incluye la aleatorización que asegura que los grupos no difieran significativamente en ninguna característica relevante al inicio del experimento.

Un desafío fundamental de los RCT en el contexto educativo es la selección adecuada de los participantes y la identificación de grupos de control. Pampaka et al. indican que, para evaluar qué intervenciones funcionan, su diseño requiere un cuidado meticuloso en la injerencia de variables como la auto-selección y el sesgo, que pueden influir en los resultados (Pampaka et al., 2016). La aleatorización debe ser implementada eficientemente para evitar que la naturaleza del contexto educativo afecte de manera desproporcionada a ciertos grupos, lo que puede llevar a conclusiones erróneas sobre el impacto de las intervenciones (Pampaka et al., 2016).

La validez interna —la capacidad de un estudio para establecer una relación causal entre la intervención y los resultados— es un elemento que puede verse comprometido en los RCT. Se requiere un diseño cuidadoso que minimice las variables de confusión.

Adicionalmente, la sostenibilidad de las intervenciones evaluadas a través de RCT es una preocupación que debe abordarse. Haerizadeh y M. comentan que las evaluaciones deben considerar no solo la eficacia inmediata de las intervenciones, sino también su capacidad para influir en el sistema educativo a largo plazo (Haerizadeh & M, 2019). Esto es especialmente relevante en el ámbito de la educación superior, donde las reformas basadas en evidencias pueden requerir tiempo para manifestar mejoras en los resultados.

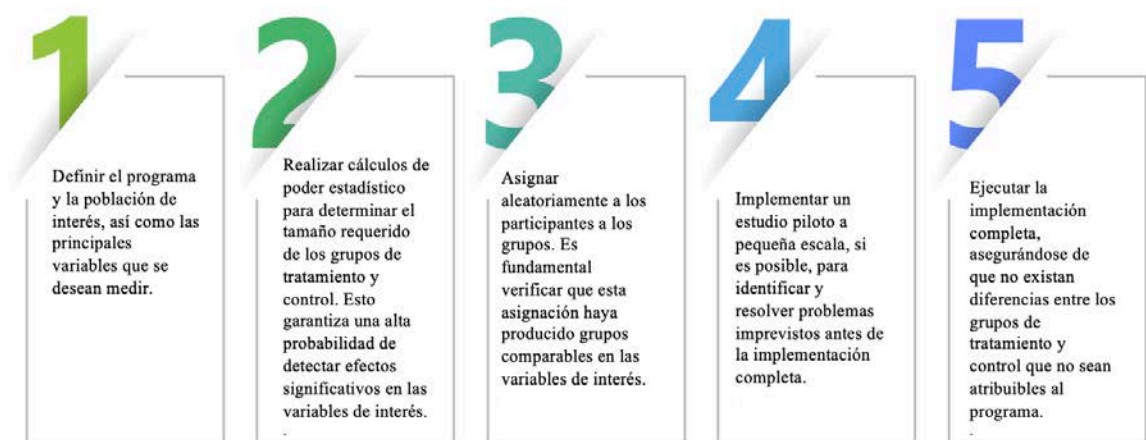
Por último, la replicabilidad de los RCT en contextos diferentes es una cuestión vital. Los hallazgos obtenidos en un contexto educativo pueden no ser aplicables a otro debido a las diferencias culturales y estructurales. Blenker et al. discuten cómo las variaciones en los métodos de investigación pueden influir en la generalización de los resultados, lo que resalta la necesidad de un enfoque adaptativo en el uso de RCT en diversas instituciones de educación superior (Blenker et al., 2014).

3.2.1.1.1 ¿Cómo se determina el número requerido de participantes en un estudio aleatorizado?

Tal y como lo menciona Pomeranz (2011), mientras mayor sea el número de participantes en un estudio, más probable es que los grupos de tratamiento y control sean comparables. Esto hace que el tamaño de la muestra sea un aspecto fundamental en el diseño del estudio.

Así mismo, la autora plantea los siguientes pasos para diseñar un estudio aleatorizado:

Grafica 2. Gráfica 2. Pasos para establecer el número de participantes en un estudio aleatorizado



Fuente: Pomeranz (2011)

3.2.2 Métodos cuasiexperimentales

Por otro lado, los métodos cuasiexperimentales aparecen como alternativas metodológicas en contextos en los que la asignación aleatoria resulta inviable por razones éticas, prácticas o logísticas. En estos diseños, se utilizan grupos de comparación no equivalentes o se implementan evaluaciones pretest-postest, de modo que se puedan comparar los efectos de la intervención a través de medidas repetidas en el mismo grupo o entre grupos que presentan diferencias iniciales (Harris et al., 2006). La ausencia de aleatorización exige el uso de técnicas adicionales, tales como el matching o el propensity score matching, con el fin de mitigar las diferencias basalmente existentes entre los grupos y reducir amenazas a la validez interna (Mardones, 2017). Además, la aplicación de estos métodos permite hallar un equilibrio entre la rigurosidad metodológica y la viabilidad práctica en entornos educativos reales (García et al., 2016).

3.2.2.1 Diferencias en Diferencias (DiD - Differences-in-Differences)

El método de Diferencias en Diferencias (DiD) es una técnica econométrica utilizada para evaluar el impacto de una intervención o tratamiento en un grupo, comparando cambios en resultados entre un grupo de tratamiento y un grupo de control en distintos momentos del tiempo. Este enfoque es especialmente útil en contextos donde no se puede realizar un Experimento Aleatorizado Controlado (RCT). A través del método DiD, los investigadores pueden controlar por factores no observables que permanecen constantes en el tiempo, permitiendo así aislar el efecto de la intervención (Feijoo-Criollo & Guaraca, 2023).

El principio básico del DiD se basa en comparar la evolución de resultados en el grupo que recibe el tratamiento con la evolución en un grupo similar que no lo recibe. Para aplicar el DiD, se requiere tener datos de antes y después de la intervención tanto para el grupo de tratamiento como para el grupo de control. La idea es observar las tendencias en los resultados antes de la intervención para asegurarse de que ambos grupos estaban en un camino similar antes de que el tratamiento se implementara. Si las tendencias antes de la intervención son paralelas, cualquier diferencia observada después de la intervención puede interpretarse como resultado del tratamiento (Corrales-Liévano et al., 2019).

Un aspecto crítico para la robustez de este método es la validación de la suposición de tendencias paralelas, que es la base del modelo DiD. Si las tendencias no son paralelas, las estimaciones del efecto del tratamiento pueden estar sesgadas (Feijoo-Criollo & Guaraca, 2023).

Además, el método puede ser ampliado para abordar variaciones en el tratamiento a lo largo del tiempo a través de la implementación de técnicas de pesaje, como el Propensity Score Matching, lo que permite mejorar la estimación del efecto causal al equilibrar las características observables entre los grupos (Feijoo-Criollo & Guaraca, 2023). Esta combinación mejora la capacidad del análisis para ofrecer conclusiones más precisas sobre el impacto de políticas o tratamientos.

3.2.2.2 Emparejamiento por Puntaje de Propensión (Propensity Score Matching - PSM)

El método de Emparejamiento por Puntaje de Propensión (Propensity Score Matching - PSM) es una técnica utilizada en estudios observacionales para estimar el efecto causal de un tratamiento o intervención. La fundamentación del PSM radica en la creación de un puntaje que representa la probabilidad de que un individuo reciba un tratamiento específico dado un conjunto de covariables observadas. Este puntaje se utiliza para emparejar individuos en grupos de tratamiento y control, buscando establecer grupos comparables que minimicen sesgos en la estimación de efectos (Austin, 2010).

El proceso de emparejamiento implica varias etapas:

- 1. Estimación del puntaje de propensión:** Se modela la probabilidad de tratamiento usando regresiones, considerando todas las covariables relevantes, asegurando que las variables que puedan influir en la decisión de recibir tratamiento estén incluidas (Austin et al., 2006).
- 2. Emparejamiento:** Individuos en el grupo de tratamiento son emparejados con aquellos en el grupo de control de acuerdo a sus puntajes de propensión. Se pueden utilizar diferentes métodos de emparejamiento, como el emparejamiento más cercano (nearest neighbor), donde se empareja cada sujeto del grupo tratado con el sujeto del grupo no tratado cuya puntuación de propensión es más cercana (Desai et al., 2017).
- 3. Comparación de resultados:** Una vez que se han creado los grupos emparejados, se procede a comparar los resultados de interés, como cambios en ingreso, salud o educación, entre los grupos para estimar el efecto del tratamiento (Austin, 2010).

Entre las principales ventajas del PSM se encuentran:

- **Reducción del sesgo de selección:** Al emparejamiento, se eliminan muchas de las diferencias entre los grupos que podrían influir en los resultados, proporcionando un estimador más limpio del efecto causal (Austin, 2012).
- **Flexibilidad y aplicabilidad:** PSM es aplicable en diversas áreas y es útil en investigaciones donde la aleatorización no es ética o práctica (Caliendo & Kopeinig, 2008).

No obstante, el PSM también tiene sus limitaciones. Una de las críticas más comunes es que solo ajusta para covariables observadas; cualquier sesgo que esté relacionado con variables no observadas no puede ser corregido (Desai et al., 2017). Además, la validez del método depende en gran medida de la capacidad para modelar correctamente el puntaje de propensión y de la disponibilidad de datos de calidad sobre las covariables relevantes (Austin & Stuart, 2015; Austin, 2013).

Otra preocupación es la elección del método de emparejamiento, donde diferentes algoritmos pueden conducir a estimaciones de efectos significativamente diferentes. Por ejemplo, algunas técnicas de emparejamiento pueden mejorar el balance entre

grupos comparados con otras (Austin, 2009), lo que implica que el investigador debe ser cuidadoso al seleccionar el enfoque adecuado (Austin, 2006).

3.2.2.3 Regresión Discontinua (Regression Discontinuity Design - RDD)

El diseño de Regresión Discontinuidad (RDD) es un método estadístico utilizado en la evaluación de políticas y programas, que permite hacer inferencias causales a partir de datos observacionales. Este método se basa en un umbral o punto de corte que determina la elegibilidad para un tratamiento o intervención específica. Aquellos individuos que se encuentran justo por encima y por debajo de este punto de corte se consideran comparables, lo que permite a los investigadores evaluar el impacto causal de la intervención (Carriel & Mendoza, 2021)(Gallego et al., 2021).

El RDD se fundamenta en la premisa de que, cuando una variable de tratamiento es asignada en función de un criterio específico (como un puntaje de elegibilidad), los individuos que están cerca del umbral son similares en sus características observables y no observables. Este diseño aprovecha el hecho de que pequeñas diferencias en la variable que determina el tratamiento (por ejemplo, ingresos, calificaciones, etc.) no deberían llevar a grandes diferencias en los resultados, exceptuando el efecto del tratamiento mismo (Carriel & Mendoza, 2021)(Gallego et al., 2021; Abad & León, 2023).

Una de las principales ventajas del RDD es que permite un control efectivo de los sesgos de selección al comparar sujetos que son casi idénticos en características observables y no observables. Esto significa que, bajo la suposición de que no hay manipulación del puntaje de elegibilidad alrededor del umbral, se puede obtener una estimación causal más robusta del efecto del tratamiento (Abad & León, 2023; Carrera, 2014).

Otra ventaja es la posibilidad de generar estimaciones locales del efecto del tratamiento (LATE). Esto significa que se puede obtener información específica sobre cómo la intervención afecta a aquellos individuos que están marginalmente por encima y por debajo del umbral de elegibilidad (Acuña et al., 2019).

En todo caso y a pesar de sus ventajas, el RDD tiene limitaciones importantes. Primero, depende de la existencia de un umbral bien definido (Carrera, 2014, Battocchio, 2018). Además, el enfoque RDD puede no capturar el efecto global de la intervención en toda la población, ya que se centra solo en individuos cerca del umbral, lo que puede limitar su generalización (Arela-Bobadilla, 2023; Mera & Quintero-Rendón, 2019).

3.2.2.4 Variables Instrumentales (Instrumental Variables - IV)

El método de Variables Instrumentales (IV) es una técnica estadística utilizada en la econometría para abordar problemas de endogeneidad en modelos de regresión. La endogeneidad se presenta cuando una variable independiente está correlacionada con el término de error, lo que puede sesgar las estimaciones y hacer que los resultados

sean no confiables. El uso de instrumentos permite separar la variación exógena de la variación endógena en los modelos, permitiendo así estimar efectos causales de manera más adecuada (Anatolyev & Skolkova, 2019; Rose & Stone, 2011; Goenner, 2013).

En otras palabras, imaginemos que queremos saber si estudiar más horas mejora las notas de los estudiantes. Parece lógico que sí, pero también puede pasar algo curioso: los estudiantes que van mal podrían estudiar más (porque les va mal), y eso complica ver si estudiar más realmente causa mejores notas. A esto le llamamos un *problema de endogeneidad* —cuando no sabemos si A causa B, o si B también afecta a A, o si hay algo más escondido que afecta a ambos.

Ante este tipo de problemas, el enfoque de variables instrumentales introduce una tercera variable, llamada “instrumento”, que está relacionada con la variable independiente (en este caso, las horas de estudio), pero no está relacionada directamente con el resultado final (las notas). De esta manera, se puede “aislar” el efecto real de la variable de interés, como si se estuviera realizando un experimento controlado.

Un buen instrumento debe cumplir dos condiciones fundamentales: debe ser relevante (fuertemente correlacionado con la variable independiente endógena) y debe ser exógeno (no debe estar correlacionado con el término de error de la ecuación de resultados) (Chen & Hsiang, 2019; Rose & Stone, 2011). Esta doble condición asegura que cualquier variación en el resultado se deba a cambios en la variable independiente mediada por el instrumento, en lugar de por factores ocultos o sesgos.

Un ejemplo clásico en el que se aplica el método IV es el análisis de los efectos de la educación en los ingresos. Si se sospecha que la educación está correlacionada con variables no observadas (como la motivación o el acceso a redes laborales), se puede utilizar un instrumento, como la ubicación geográfica o las políticas de admisión en universidades, que afecta la educación pero no directamente el ingreso (Rose & Stone, 2011; Sekula et al., 2016).

El proceso de estimación IV generalmente se lleva a cabo en dos etapas:

1. **Primera etapa:** Se regresa la variable endógena sobre el instrumento y otras covariables. Se estima la relación entre el instrumento y la variable endógena. Este paso genera lo que se denomina el “valor predicho” o “fitted value” de la variable endógena.
2. **Segunda etapa:** Se utiliza el valor predicho de la variable endógena en lugar de la variable original en la ecuación de resultados. Esto permite obtener una estimación del efecto causal de la variable endógena en el resultado observado (Rose & Stone, 2011; Goenner, 2013).

El método de variables instrumentales también presenta desafíos y limitaciones. La elección de un buen instrumento es crucial; si el instrumento no es relevante o es endógeno, las estimaciones pueden resultar más sesgadas que en el modelo original (Mabeba, 2024; Yang et al., 2019). Esto plantea un desafío significativo en la práctica, ya que identificar instrumentos adecuados puede ser complicado.

Además, el efecto estimado mediante IV suele ser “local”, lo que significa que se refiere a un grupo específico de individuos que están influenciados por la variable instrumentada. Esto limita la generalización de los resultados a otras poblaciones (Imbens, 2014; Tsiboe & Turner, 2023).

3.2.2.5 Serie Temporal Interrumpida (Interrupted Time Series - ITS)

El método de Interrupción de Serie Temporal (ITS) es una técnica valiosa utilizada para evaluar los efectos de una intervención o tratamiento en datos observacionales obtenidos a lo largo del tiempo. Este método se basa en el análisis de series temporales antes y después de una intervención, permitiendo observar cambios en las tendencias o niveles de la variable de interés. El ITS es particularmente útil en contextos donde no es posible realizar ensayos controlados aleatorios debido a consideraciones éticas o logísticas (Cassiano & Menezes, 2018).

La serie temporal se analiza antes de la intervención (la fase pre-intervención) y después de la intervención (la fase post-intervención). Los cambios observados en la media o la tendencia de la variable después de la intervención se comparan con la tendencia esperada si no hubiera ocurrido la intervención.

La estructura básica del análisis ITS incluye tres componentes:

1. **Tendencias previas:** Evaluación de la tendencia de la serie temporal antes de la intervención.
2. **Punto de interrupción:** Identificación del periodo en el que ocurrió la intervención, donde se espera que la serie muestre un cambio.
3. **Tendencias posteriores:** Medición de la evolución de la variable después de la intervención y comparación con la tendencia esperada.

Las principales ventajas del diseño de Interrupción de Serie Temporal son:

- **Capacidad de detectar cambios en tendencias y niveles:** A diferencia de otros métodos que pueden centrarse únicamente en cambios en promedios, el ITS permite observar no solo cambios en el nivel de la serie, sino también en sus tendencias a largo plazo (Tao et al., 2018).
- **Contexto natural:** Al utilizar datos de series temporales observacionales, el ITS puede proporcionar estimaciones en condiciones naturales sin requerir un entorno experimental estricto.
- **Análisis longitudinal:** Este enfoque permite observar patrones a lo largo del tiempo, facilitando la identificación de efectos latentes que podrían no ser evidentes en estudios transversales (Hao & Fang-ai, 2024).

No obstante, el método ITS también enfrenta limitaciones. Uno de los desafíos es el riesgo de confusión por tendencias externas. Por ejemplo, si otros eventos o políticas se implementan simultáneamente a la intervención analizada, puede volverse complicado atribuir cambios en la serie a la intervención misma (Sax & Steiner, 2013).

Además, debe haber un adecuado período de observación tanto antes como después de la intervención para captar adecuadamente los efectos (Liu et al., 2020).

3.2.2.6 Matching Exacto o Coarsened Exact Matching (CEM)

El método de Coarsened Exact Matching (CEM) es una técnica estadística utilizada para emparejar grupos en estudios observacionales, especialmente en el contexto de inferencias causales. Este método tiene como objetivo controlar el sesgo y mejorar la validez de los resultados al asegurar que los grupos comparados sean similares en aspectos observables, a pesar de que no se hayan asignado aleatoriamente a las intervenciones.

El CEM consiste en tres pasos fundamentales:

1. **Coarsificación (Coarsening):** Las covariables que se desea equilibrar se transforman en categorías más amplias o “coarsened”. Este proceso implica agrupar la información original en intervalos o categorías significativas, lo cual facilita el emparejamiento posterior (Yang et al., 2024). Por ejemplo, si se dispone de datos de edad, en lugar de usar valores precisos, se pueden crear categorías como “18-30”, “31-45” y “46 y más”.
2. **Emparejamiento (Matching):** Después de la coarsificación, se utiliza el emparejamiento exacto para crear pares de tratamientos y controles que son similares en base a las covariables coarsenadas. Este paso busca formar grupos comparables utilizando las categorías predefinidas, donde cada unidad en el grupo de tratamiento tiene un “match” en el grupo de control (Burden et al., 2017).
3. **Estimation (Estimación):** Una vez que se ha realizado el emparejamiento, se procede a analizar los resultados de interés, utilizando las unidades emparejadas para estimar el efecto del tratamiento. El CEM permite obtener estimaciones que son menos dependientes del modelo y más eficaces en la reducción del sesgo en comparación con otros métodos como el Propensity Score Matching (PSM) (Wagner et al., 2022).

Al igual que en los demás métodos explicados, el CEM no está exento de limitaciones:

1. **Pérdida de Información:** La coarsificación de variables puede llevar a una pérdida de precisión, especialmente si las categorías son demasiado amplias y ocultan variaciones importantes dentro de las clases (Sunarwibowo et al., 2024).
2. **Dependencia de la Selección de Covariables:** La efectividad del CEM depende de la elección de covariables relevantes al inicio del proceso. Si se omiten covariables importantes que podrían influir en el resultado, el sesgo permanece (Iacus et al., 2012).
3. **Tamaño de la Muestra:** En algunos casos, la coarsificación puede resultar en grupos demasiado pequeños después del emparejamiento, lo que podría afectar la capacidad de generalizar los resultados (Dym & Lipman, 2017).

En síntesis:

Tabla 2. Metodos cuasiexperimentales

Método Cuantitativo	¿Qué es?	Ventajas (lo bueno)	Desventajas (lo difícil)	Ejemplo práctico en educación superior
1. Experimento Aleatorizado Controlado (RCT)	Es como lanzar una moneda para decidir quién recibe un programa (grupo tratamiento) y quién no (grupo control). Todo se asigna al azar.	Es el método más confiable para saber si algo realmente funciona. Evita que otros factores influyan en los resultados.	Puede ser costoso, difícil de aplicar en entornos reales (como universidades) y puede generar conflictos éticos (por negar el beneficio a algunos).	Una universidad quiere saber si un nuevo curso de nivelación académica mejora la permanencia. Se asigna aleatoriamente a unos estudiantes para tomarlo y se compara su avance con quienes no lo tomaron.
2. Diferencias en Diferencias (DiD)	Compara qué tanto cambia un grupo que recibe el programa frente a otro que no, antes y después de que se aplique.	Permite ver el efecto de la intervención incluso si los grupos no eran iguales al inicio. Usa datos que muchas instituciones ya tienen.	No sirve si los dos grupos estaban cambiando de forma muy diferente antes del programa.	El gobierno da incentivos para ampliar cupos a jóvenes rurales en algunas universidades. Se compara cómo cambió la matrícula rural en estas frente a otras universidades sin el incentivo.
3. Propensity Score Matching (PSM)	Empareja personas que son muy parecidas entre sí (edad, nivel socioeconómico, colegio, etc.), y que solo se diferencian en que unas recibieron el programa y otras no.	Mejora la comparación entre grupos. Se basa en características reales y conocidas.	No puede tener en cuenta diferencias que no se ven en los datos (por ejemplo, motivación personal).	Para saber si una beca mejora la graduación, se comparan estudiantes beneficiados con otros muy similares (pero sin beca) para ver quiénes terminan más.
4. Regresión Discontinua (RDD)	Se aplica cuando el programa se entrega solo a quienes superan cierto punto (por ejemplo, un puntaje mínimo en un examen). Se compara a los que están justo por encima y por debajo del punto.	Muy útil cuando la asignación al programa depende de un número. Da resultados muy confiables cerca del umbral.	Solo funciona para los estudiantes que están cerca del punto de corte; no dice nada de los que están muy lejos.	Un fondo universitario entrega ayuda económica a quienes obtienen más de 300 puntos en una prueba estandarizada de Estado. Se compara el rendimiento entre quienes sacaron 299 y 301.

Método Cuantitativo	¿Qué es?	Ventajas (lo bueno)	Desventajas (lo difícil)	Ejemplo práctico en educación superior
5. Variables Instrumentales (IV)	Se usa cuando no se puede hacer aleatorización, y hay una “tercera variable” que influye en participar en el programa, pero no en los resultados directamente.	Permite corregir errores cuando los estudiantes se “autoseleccionan” para participar, como cuando los más motivados se inscriben solos.	Difícil encontrar esa “tercera variable” que sea válida. Y los resultados solo aplican a quienes cambiaron de decisión por esa variable.	Una universidad quiere medir el impacto de estudiar en la sede principal. Usa como “instrumento” la distancia a la sede para ver si vivir más cerca afecta el ingreso y los resultados.
6. Serie de Tiempo Interrumpida (ITS)	Analiza los datos de varios años antes y después de aplicar un programa, buscando un “cambio brusco” tras la intervención.	Es útil cuando no hay grupo de comparación, pero sí datos históricos. Bueno para políticas a nivel nacional.	Puede confundirse si pasaron otras cosas al mismo tiempo que también afectaron los resultados.	Se analiza si, después de implementar una ley que obliga a orientar mejor al estudiante sobre carreras universitarias, mejora la tasa de retención en universidades públicas en los años siguientes.
7. Coarsened Exact Matching (CEM)	Es una forma de emparejar personas de forma más precisa, agrupando las variables (por ejemplo, poner a todos los de 18–20 años juntos). Luego se compara solo entre los grupos que quedan bien emparejados.	Se logra mayor precisión en los emparejamientos. Mejora la calidad del análisis.	Puede dejar fuera a muchas personas si no se encuentran buenos emparejamientos.	Se compara el impacto de estudiar en universidades acreditadas frente a no acreditadas, emparejando estudiantes con similares características: edad, puntaje en pruebas estandarizada de Estado, zona de residencia, etc.

3.2.3 Instrumentos cuantitativos

Tabla 3. Instrumentos cuantitativos para la evaluación de impacto

Instrumento	Definición	Recomendaciones Psicométricas
Encuestas y Cuestionarios	Herramientas estructuradas utilizadas para recolectar datos de grandes grupos. Usan preguntas cerradas como escalas de Likert, opción múltiple o dicotómicas. González et al. (2018).	Validar la confiabilidad (α de Cronbach > 0.70), realizar prueba piloto, revisar consistencia interna y claridad semántica.
Pruebas Estándar	Instrumentos diseñados para medir competencias, conocimientos o habilidades específicas. Se aplican con protocolos establecidos y permiten comparaciones entre grupos.	Utilizar pruebas validadas o adaptadas con análisis de validez de contenido y constructo. Reportar medias, desviaciones y percentiles.
Escalas de Medición	Instrumentos que cuantifican actitudes, creencias o percepciones. Comúnmente emplean escalas de tipo Likert para valorar el nivel de acuerdo con afirmaciones específicas. (Manzanares & Llamazares, 2021).	Evaluar la fiabilidad (α de Cronbach o McDonald's ω), realizar análisis factorial exploratorio/confirmatorio y análisis de ítems.
Registro de Datos	Sistemas automatizados o manuales que capturan información de manera sistemática, útil para estudios longitudinales (ej., desempeño académico, deserción).	Garantizar la integridad de los datos, establecer reglas de validación y verificar consistencia temporal y coherencia de las series.
Análisis Estadístico	Procesamiento matemático de datos recolectados para identificar patrones, asociaciones o diferencias. Incluye análisis descriptivo e inferencial.	Usar software validado (SPSS, Stata, R). Reportar supuestos estadísticos, tamaño del efecto y significancia. Validar resultados extremos.

3.3 Tema 3. Métodos cualitativos en la evaluación de impacto.

Los métodos cualitativos en la evaluación de impacto se han posicionado como herramientas fundamentales para comprender en profundidad los procesos, contextos y experiencias subyacentes que generan efectos en las intervenciones educativas. La aplicación de estos instrumentos se centra en identificar y analizar las percepciones, actitudes y vivencias de los actores involucrados, lo que resulta indispensable para evaluar el impacto de un programa o política. En este sentido, se han utilizado instrumentos como entrevistas semiestructuradas, grupos focales, análisis documental

y observación participante, que permiten explorar el “cómo” y el “por qué” de los cambios producidos, aportando una visión contextualizada que contrasta con el enfoque exclusivamente numérico de los métodos cuantitativos (López-Meseguer & Valdés, 2020).

La aplicación de técnicas cualitativas en la evaluación de impacto se destaca por su capacidad para interpretar la complejidad de los escenarios educativos, donde el contexto, la cultura y la subjetividad son elementos centrales en el proceso de cambio. Asimismo, el enfoque cualitativo facilita la identificación de indicadores emergentes y permite una mayor flexibilidad interpretativa, facilitando la articulación de recomendaciones pertinentes para la mejora de las políticas e intervenciones educativas (Román, 2024).

3.3.1 Instrumentos de evaluación cualitativa

Los instrumentos de evaluación cualitativa son herramientas fundamentales que permiten recolectar y analizar datos no numéricos, facilitando la comprensión de fenómenos sociales a través de la exploración de experiencias, percepciones y procesos.

Tabla 4. Instrumentos de evaluación cualitativa

Instrumento	Definición	Recomendaciones para su uso en evaluación de impacto en educación superior
Entrevistas semiestructuradas	Conversaciones con un propósito investigativo que combinan preguntas preparadas con flexibilidad para explorar nuevas ideas emergentes. Permiten obtener descripciones del mundo vivido por los participantes. (Galindo-Lozano & Correa, 2019; Agudelo & Loaiza, 2022).	Diseñar una guía temática alineada con los objetivos de impacto; capacitar a los entrevistadores; grabar y transcribir con consentimiento informado; analizar mediante codificación temática.
Grupos focales	Interacción de 6 a 12 personas bajo la guía de un moderador para discutir un tema específico. Favorecen la exploración de percepciones colectivas y dinámicas grupales. (Trad, 2009; Lizcano & Amaya, 2014).	Utilizar cuando se busca contrastar puntos de vista sobre el impacto de políticas o programas; seleccionar perfiles diversos; garantizar un espacio seguro para el diálogo.
Observación participativa	Método que implica la presencia activa del investigador en el contexto a estudiar, observando comportamientos, interacciones y dinámicas sociales. (Venegas & Crispín, 2018).	Ideal en estudios de impacto institucional; registrar notas de campo sistemáticamente; complementar con entrevistas; reflexionar críticamente sobre el rol del investigador.
Historias de vida	Relatos personales que reconstruyen trayectorias de vida en relación con la experiencia educativa, brindando profundidad sobre procesos de transformación individual. (Conde et al., 2022; Brítez & Chung, 2022).	Útiles para evaluar impacto en trayectorias de egresados; seleccionar casos representativos; establecer confianza; analizar desde enfoques narrativos.

Instrumento	Definición	Recomendaciones para su uso en evaluación de impacto en educación superior
Cartografía social	Técnica participativa que utiliza mapas como herramientas de investigación y comunicación para representar y analizar la realidad social de un territorio desde la perspectiva de las comunidades que lo habitan. Esta práctica va más allá de la mera representación geográfica y se enfoca en la inclusión de voces, historias y significados que las comunidades atribuyen a su entorno (Mesa et al. (2021; Vélez & Torres et al., 2012).	Adecuada para comunidades universitarias diversas; facilita análisis de percepciones territoriales del impacto educativo; requiere facilitación clara y materiales visuales.
Diarios reflexivos	Son herramientas metodológicas utilizadas en contextos educativos y profesionales para fomentar la reflexión sobre experiencias, desempeños y procesos de aprendizaje. Son documentos escritos donde los individuos registran sus pensamientos, emociones, reacciones y aprendizajes relacionados a situaciones vividas, con el objetivo de evaluar y mejorar continuamente su práctica (Campana (2023; Gallardo et al., 2020; Boszko & Rosa, 2020).	Puede integrarse en procesos longitudinales de evaluación; pedir reflexiones orientadas; analizar cambios en discurso y autopercepción; resguardar privacidad.
Análisis documental	Revisión sistemática de documentos institucionales, políticas, proyectos o testimonios escritos para comprender contextos, decisiones y resultados. (Jiménez & Tarazona, 2024).	Útil para triangulación; aplicar matrices de análisis; identificar coherencia entre intenciones y resultados declarados; considerar evolución histórica del documento.

3.3.2 ¿Cómo se determina el número requerido de participantes en un estudio no aleatorizado?

El muestreo no probabilístico es un método de selección de muestras que no se basa en la aleatoriedad. Esto significa que los participantes no tienen la misma probabilidad de ser elegidos, lo que permite al investigador decidir quiénes se incluirán en la muestra basándose en criterios específicos. Este tipo de muestreo es útil en situaciones donde se busca obtener información detallada de un grupo particular, pero puede limitar la capacidad de generalizar los resultados a toda la población debido a su naturaleza sesgada (Faúndez-Casanova, 2018).

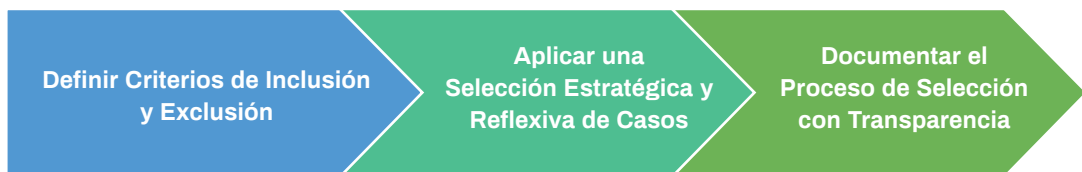
Existen diferentes tipos de muestreo no probabilístico:

1. **Muestreo por Conveniencia:** Consiste en seleccionar a los participantes que son más accesibles o fáciles de reclutar. Este método es práctico, pero puede no representar adecuadamente a la población (Faúndez-Casanova, 2018).

2. **Muestreo Intencionado (o objetivo):** Se eligen a los participantes con base en características específicas que son relevantes para el estudio. Este muestreo es útil cuando se busca profundizar en un tema particular, pero puede introducir sesgos debido a la selección subjetiva de los participantes (Campoverde-Luque et al., 2023).
3. **Muestreo por Cuotas:** Se establece un criterio específico para seleccionar participantes hasta completar un número predeterminado en diferentes subgrupos. Esto permite asegurar la inclusión de diferentes perspectivas, pero puede ser menos representativo de la población general (Paz-Bailey et al., 2013).
4. **Muestreo en Cadena (Bola de Nieve):** Este método se utiliza frecuentemente en poblaciones ocultas o difíciles de alcanzar. Los participantes iniciales son utilizados para identificar y referir a otros participantes, creando así una “bola de nieve” en el proceso de reclutamiento (Barrientos et al., 2014).

En este contexto se recomienda seguir los siguientes pasos:

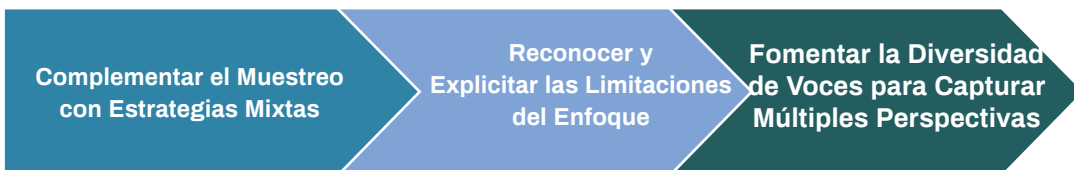
Grafica 3. Pasos para determinar el numero de participantes en un estudio aleatoreo



Antes de iniciar el proceso de selección, es crucial establecer criterios claros, alineados con los objetivos evaluativos y/o la teoría del cambio del programa o intervención educativa. Esto permitirá identificar perfiles estratégicos (egresados, docentes, beneficiarios directos o indirectos) cuya experiencia aporte a la comprensión del impacto. Estos criterios deben ser explícitos y justificados (Muñoz et al., 2016).

Utilizar un enfoque intencionado y fundamentado en la lógica del caso típico, caso extremo, o caso contrastante, según lo requiera la evaluación. Esta selección debe maximizar la variabilidad conceptual y enriquecer la comprensión del fenómeno de impacto. Se recomienda justificar cada inclusión a partir de su valor informativo (Faúndez-Casanova, 2018).

Mantener una bitácora o matriz de selección en la que se consignen los criterios aplicados, la procedencia de los participantes, y los argumentos para su elección. Esta trazabilidad fortalece la credibilidad del estudio, facilita auditorías metodológicas y permite futuras replicaciones o comparaciones (Bae et al., 2017).



La integración con técnicas cuantitativas o el uso de métodos mixtos fortalece la triangulación de datos. Por ejemplo, puede combinarse una encuesta amplia con entrevistas a profundidad a casos seleccionados, o emplear observaciones y análisis documental para enriquecer la interpretación de los efectos identificados (Baldizzoni, 2014).

Dado que el muestreo no probabilístico no busca representatividad estadística, es necesario discutir con claridad sus límites en términos de generalización. Sin embargo, su fortaleza radica en la profundidad analítica y en la generación de comprensión contextualizada, elementos fundamentales para explicar impactos en poblaciones diversas (Arceo-Gómez & Campos-Vázquez, 2014).

Especialmente en educación superior, donde las trayectorias educativas se cruzan con factores de género, etnia, territorio, discapacidad características socioeconómicas, es fundamental garantizar la inclusión de voces diversas. Esto permite mapear los efectos diferenciales y atender a los principios de equidad y justicia educativa en la evaluación (Ramírez-Moreno & Ramírez-Villegas, 2011).

3.3.3. Proceso del Método Delphi

El método Delphi es una técnica de investigación cualitativa que se utiliza para obtener consenso sobre un determinado tema o problema mediante la colaboración de un panel de expertos (Wang et al., 2022). El proceso típico incluye las siguientes etapas:

- **Selección de Expertos:** Se elige un grupo de expertos que tengan conocimiento relevante sobre el tema en cuestión. La representatividad y diversidad de opiniones son necesarias para asegurar un consenso informado.
- **Rondas de Consulta:** Se llevan a cabo múltiples rondas en las que se solicita a los participantes que respondan a una serie de preguntas abiertas y cerradas. Después de cada ronda, se compila la información y se proporciona un resumen de las respuestas a los participantes para fomentar el debate y la reflexión.
- **Análisis de Resultados:** Tras varias rondas de recolección de datos, se analizan las respuestas para identificar aspectos de consenso o desacuerdo, así como para refinarlas y llegar a conclusiones sobre el tema en cuestión.
- **Retroalimentación Continua:** Este enfoque permite que los participantes revisen sus opiniones con base en el feedback proporcionado, estimulando así la discusión y el refinamiento continuo de ideas.

3.4 Tema 4. Métodos mixtos

Dado que ni los métodos cuantitativos ni los cualitativos por sí solos pueden captar la complejidad del fenómeno educativo, cobra especial relevancia la implementación de metodologías mixtas en los procesos de evaluación de impacto. Estas metodologías

integran los beneficios de ambos enfoques, permitiendo no solo medir los resultados obtenidos, sino también comprender los mecanismos y contextos que los generan. Así, se potencia la capacidad diagnóstica, explicativa y propositiva de la evaluación, aportando evidencia más robusta para la toma de decisiones y la mejora continua. Los diseños metodológicos mixtos han ganado popularidad en la investigación contemporánea por su capacidad para combinar enfoques cualitativos y cuantitativos, permitiendo a los investigadores explorar temas complejos de manera más holística. Estos diseños permiten integrar diferentes tipos de datos, lo que contribuye a una comprensión más rica y completa del fenómeno en cuestión (Johnson & Onwuegbuzie, 2004; Igiebor & Okonmah, 2022).

3.4.1 Tipos de Diseños Mixtos

- **Diseño Secuencial Explanatorio:** Este diseño implica realizar primero la recolección y análisis de datos cuantitativos, seguido de la recolección y análisis de datos cualitativos. Permite utilizar los resultados cuantitativos para informar y profundizar en la comprensión de los hallazgos cualitativos (Thannimalai, 2021).
- **Diseño Secuencial Exploratorio:** En este diseño, se inicia con la recolección y análisis de datos cualitativos, que se utilizan para informar posteriormente la recolección de datos cuantitativos. Es útil cuando se busca explorar un fenómeno antes de medirlo (Saputra et al., 2023).
- **Diseño Convergente Paralelo:** También conocido como diseño concurrente, en este enfoque se recopilan datos cualitativos y cuantitativos de manera simultánea y se analizan por separado, para luego integrar los resultados en la interpretación final. Este diseño permite comparar y contrastar los resultados de ambos tipos de datos de manera equitativa (Saputra, 2021).
- **Diseño Integrado:** Este diseño implica la inclusión de un tipo de datos dentro de otro. Por ejemplo, se puede realizar un estudio principalmente cuantitativo en el que se integren algunos componentes cualitativos para enriquecer la interpretación de los resultados (Saputra et al., 2023).
- **Diseño Triangulado:** Este diseño combina múltiples métodos de recolección de datos y enfoques de análisis en un solo estudio. La triangulación ofrece una manera de validar los resultados mediante la comparación de diferentes fuentes de datos (Saputra, 2021).

3.5 Teoría de Cambio

Si bien la Teoría de Cambio no es en sí un método de recolección de datos, puede estructurar y guiar una investigación con enfoque mixto. De hecho, es ideal para investigaciones que combinan lo cuantitativo y lo cualitativo.

La teoría de cambio en el contexto de la educación superior se refiere a un mecanismo conceptual que permite a las instituciones educativas planificar, implementar y evaluar sus acciones y programas. Este enfoque facilita la comprensión de cómo y por qué se esperan cambios en un contexto específico mediante la identificación de las relaciones causales entre las actividades realizadas y los resultados deseados (Limaico et al.,

2022). A continuación, se describen los pasos fundamentales de la teoría de cambio aplicados a la educación superior:

- **Definición de la Visión y Objetivos a Largo Plazo**

El primer paso en la teoría de cambio es definir claramente la visión o el impacto esperado. En el contexto educativo, esto puede incluir objetivos como mejorar la calidad de la enseñanza, aumentar la tasa de graduación o preparar a los estudiantes para el mercado laboral. Este paso es fundamental ya que establece un marco claro sobre lo que se quiere lograr (Valle et al., 2023; Méndez-Mantuano et al., 2021).

- **Identificación de Resultados Intermedios**

Una vez que se han establecido los objetivos a largo plazo, el siguiente paso es identificar los resultados intermedios que serán necesarios para alcanzarlos. Estos resultados suelen ser más específicos y medibles, como aumentar la participación estudiantil en actividades académicas, mejorar las calificaciones en algunos cursos o elevar la satisfacción estudiantil. Es importante que estos resultados sean factibles y se basen en la evidencia existente (Luciano & Velázquez, 2022).

- **Desarrollo de Actividades e Intervenciones**

En este paso, se diseñan las actividades e intervenciones que se llevarán a cabo para lograr los resultados intermedios. Las actividades deben ser coherentes y alineadas con los objetivos. Por ejemplo, para mejorar la calidad educativa, se pueden implementar programas de formación docente o metodologías de enseñanza innovadoras, como el aula invertida (Bastida-Molina et al., 2020; López, 2023). Es esencial que estas actividades sean realistas y adaptadas al contexto específico de la institución educativa.

- **Establecimiento de Supuestos y Condiciones Necesarias**

En este paso, se identifican los supuestos que sustentan la teoría de cambio, así como las condiciones necesarias para que las actividades tengan éxito. Esto implica reconocer los factores externos que pueden influir en la implementación, como el apoyo institucional, la financiación adecuada, la participación de los estudiantes y la disponibilidad de recursos (Valle et al., 2023; Méndez-Mantuano et al., 2021). También se deben considerar las posibles barreras que podrían dificultar el logro de los objetivos deseados.

- **Recolección de Datos y Evaluación**

La recopilación de datos es un componente fundamental de la teoría de cambio. Se deben definir indicadores claros para medir el progreso hacia los resultados intermedios y los objetivos finales. Esto puede incluir encuestas de satisfacción estudiantil, análisis del rendimiento académico, entre otros (Luciano & Velázquez, 2022). La evaluación continua de estos datos permitirá realizar ajustes en las estrategias si es necesario.

- **Análisis y Reflexión**

Finalmente, la teoría de cambio implica un proceso de análisis y reflexión sobre los resultados obtenidos. Esto no solo ayuda a verificar si se han alcanzado los objetivos,

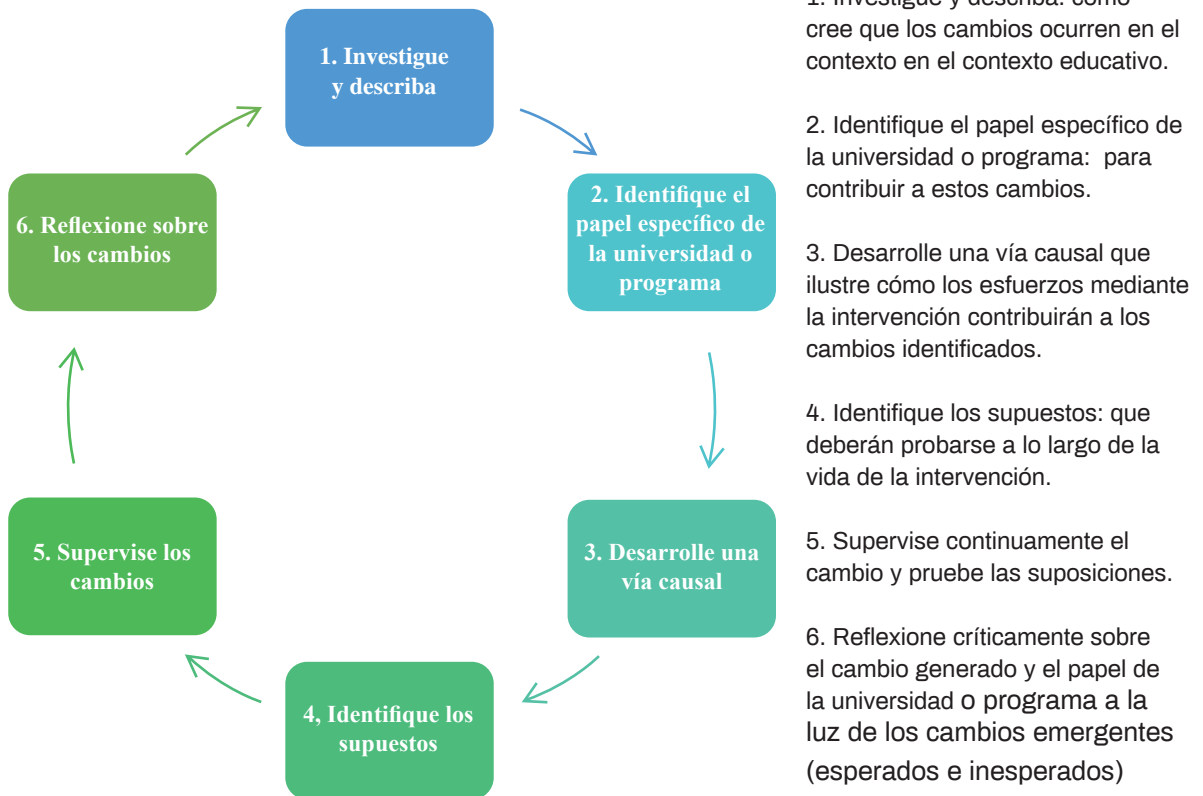
sino que también permite aprender de la experiencia. Las lecciones extraídas del proceso pueden informar futuras intervenciones y estrategias en la institución educativa (Urunaga & Cabrera, 2024; Gaul et al., 2021).

Como marco orientador en investigaciones mixtas:

- Permite vincular datos cuantitativos (indicadores, encuestas, estadísticas) con datos cualitativos (entrevistas, grupos focales, estudios de caso).
- Ayuda a diseñar una estrategia metodológica donde las técnicas cualitativas expliquen el “cómo” y “por qué” de los resultados identificados con métodos cuantitativos.
- Posibilita analizar si los resultados observados corresponden a los resultados esperados, o si surgieron efectos no anticipados.

Global Focus (2015) proponen seis pasos a tener en cuenta en el proceso de planificación de la teoría de cambio, así:

Grafica 3. Pasos para de la teoria de cambio



Fuente: Adaptado de Global Focus (2015:10)

3.6 Tema 5. Instrumentos y herramientas digitales para la evaluación de impacto.

En el contexto actual de transformación digital y expansión del acceso a tecnologías de información, la evaluación de impacto en educación superior puede beneficiarse significativamente del uso de herramientas digitales. Estas tecnologías permiten recopilar, analizar y visualizar información de manera más eficiente, accesible y escalable, favoreciendo la toma de decisiones informada, la transparencia en la gestión educativa y la mejora continua de los programas formativos.

El siguiente cuadro presenta una selección de instrumentos y herramientas digitales que pueden ser utilizadas en procesos de evaluación de impacto, especialmente en contextos educativos. Cada herramienta se describe brevemente, se señala su uso recomendado en evaluaciones aplicadas a la educación superior, y se ofrecen recomendaciones metodológicas para su implementación:

Tabla 5. Herramientas digitales para la evaluación de impacto

Herramienta Digital	Definición y características	Uso recomendado	Recomendaciones metodológicas
Plataformas de encuestas (Qualtrics, Google Forms, SurveyMonkey, Typeform)	Permiten diseñar, distribuir y analizar cuestionarios en línea, con escalas cerradas o abiertas, lógica de ramificación y exportación de datos.	Recolección masiva de datos cuantitativos o cualitativos con cobertura nacional o internacional.	Validar escalas antes de aplicar; cuidar la tasa de respuesta; garantizar accesibilidad y protección de datos.
Software de análisis estadístico (SPSS, Stata, R, Jamovi)	Herramientas para el análisis estadístico descriptivo e inferencial, análisis de regresión, diseño experimental, entre otros.	Procesamiento de datos cuantitativos para medir el impacto en variables académicas, sociales o laborales.	Aplicar pruebas de confiabilidad; usar modelos adecuados al diseño de la evaluación (pre-post, regresión, etc.).
Software de análisis cualitativo (ATLAS.ti, NVivo, MAXQDA)	Permiten codificar, categorizar y analizar grandes volúmenes de datos textuales, visuales o sonoros.	Análisis de entrevistas, grupos focales, relatos de vida, y triangulación de datos.	Definir un marco analítico claro; capacitar en el uso del software; documentar el proceso de codificación.
Plataformas de visualización (Power BI, Tableau, Google Data Studio)	Herramientas interactivas que transforman los datos en visualizaciones dinámicas y dashboards.	Comunicación de hallazgos a tomadores de decisión y públicos no expertos.	Priorizar indicadores; usar filtros interactivos para análisis por grupo; actualizar los datos en tiempo real.
Formularios con grabación de audio/video (Voiceform, VideoAsk)	Herramientas que permiten que el encuestado grabe respuestas orales o visuales, útil para entrevistas asincrónicas y testimonios.	Recogida de historias de vida, percepción de impacto, testimonios audiovisuales.	Proveer instrucciones claras; verificar calidad del audio/video; garantizar consentimiento informado audiovisual.

Herramienta Digital	Definición y características	Uso recomendado	Recomendaciones metodológicas
Redes sociales y minería de datos sociales (Twitter API, Netlytic, NodeXL)	Permiten extraer y analizar datos de redes sociales para identificar tendencias, percepciones o impactos simbólicos de programas o políticas.	Análisis de percepción pública de universidades, programas académicos o reformas educativas.	Evaluar la pertinencia ética del análisis; usar criterios claros de recolección; considerar el sesgo algorítmico.
Plataformas de gestión del aprendizaje (Moodle, Canvas, Blackboard)	Sistemas donde se pueden extraer datos de comportamiento, desempeño y participación de los estudiantes en cursos virtuales o híbridos.	Evaluación del impacto de programas de formación en línea o estrategias pedagógicas digitales.	Usar registros de actividad, finalización y evaluación; complementar con encuestas o entrevistas.
Blockchain educativo	Tecnología emergente para registrar y verificar logros educativos de forma descentralizada y segura.	Validación de competencias adquiridas y trazabilidad del egreso en programas con certificación digital.	Considerar su implementación en alianzas institucionales; analizar legalidad y aceptación en el país.

4. ACTIVIDADES PRÁCTICAS

A continuación, se presentan cuatro casos prototípicos de evaluación de impacto, con posibles soluciones metodológicas para su abordaje:

4.1. Caso 1.

La Universidad Nacional de un país latinoamericano, con presencia regional en zonas rurales, ha implementado durante la última década un conjunto de estrategias orientadas a la inclusión educativa: becas, programas de acompañamiento psicosocial, prácticas comunitarias y formación con enfoque territorial. Hoy, egresan jóvenes que son primera generación en acceder a la educación superior.

La institución quiere saber si estos esfuerzos han logrado incidir realmente en la movilidad social de los egresados, en su capacidad para generar empleo, retornar a sus territorios con soluciones aplicadas, y mejorar su calidad de vida y la de sus familias. Se reconoce que el acceso no garantiza per se la transformación.

¿Hasta qué punto la formación universitaria está cambiando las trayectorias de vida de los egresados más vulnerables?

Tabla 6. Caso 1 de evaluación de impacto

Ámbito de la educación superior	Posibles dilemas de la evaluación	Información disponible en la Institución	Propuesta metodológica de evaluación de impacto
Impacto de los egresados	<p>¿Cómo aislar el impacto de la universidad frente a otros factores (familia, género, territorio)?</p> <p>¿Cómo captar efectos de largo plazo en trayectorias vitales?</p> <p>¿Cómo evitar el sesgo de autoselección (los mejores ingresan a mejores universidades)?</p> <p>¿Qué hacer cuando no existe grupo de comparación directo?</p>	<p>Encuestas institucionales de egresados (trazabilidad 1 a 5 años).</p> <p>Bases de datos de rendimiento académico, becas, origen geográfico y socioeconómico.</p> <p>Registros de prácticas, empleo, participación en proyectos institucionales.</p> <p>Algunos testimonios o estudios de caso sobre transformación personal.</p>	<p>Diseño mixto: Secuencial Explanatorio.</p> <p>Método cuantitativo: Propensity Score Matching (PSM). Comparar egresados de contextos similares que accedieron o no a un programa institucional (ej. becas + mentoría).</p> <p>Instrumento: Encuesta de egresados ampliada (variables de ingreso, calidad de vida, autonomía económica, liderazgo).</p> <p>Método cualitativo: Historias de vida en profundidad de egresados representativos por región, género y trayectoria (ascenso social, retorno a territorio, transformación familiar).</p> <p>Teoría de cambio aplicada: El acompañamiento integral y la educación de calidad impulsan la movilidad social, el liderazgo juvenil y el bienestar familiar.</p>

4.2. Caso 2.

La Universidad Federal ha desarrollado en los últimos cinco años un programa estratégico de ciencia aplicada llamado “Investigación con Sentido”, que financia proyectos de innovación social y tecnológica liderados por grupos de investigación interdisciplinarios, en alianza con gobiernos locales, MIPYMES y organizaciones comunitarias.

Uno de los objetivos centrales del programa ha sido conectar la producción de conocimiento con la solución de problemas reales: desde tecnologías para el acceso a agua potable hasta políticas públicas en salud mental rural. Se han financiado más de 100 proyectos.

La universidad quiere demostrar que su inversión ha producido un impacto tangible más allá de las publicaciones.

¿La investigación ha mejorado prácticas, políticas o procesos productivos en los territorios? ¿Qué efectos han percibido las comunidades aliadas?

Tabla 7. Caso 2 de evaluación de impacto

Ámbito de la educación superior	Posibles dilemas de la evaluación	Información disponible en la Institución	Propuesta metodológica de evaluación de impacto
Impacto de la investigación en comunidades y sectores productivos	<p>¿Cómo medir el impacto cuando los efectos de la investigación son difusos o de largo plazo?</p> <p>¿Cómo separar el aporte universitario de otras intervenciones en el territorio?</p> <p>¿Qué significa “impacto” en sectores distintos (salud, agro, políticas públicas)?</p> <p>¿La producción académica se transforma en valor social?</p>	<p>Proyectos con financiación y aliados externos (informes técnicos, productos).</p> <p>Patentes, consultorías, materiales aplicados.</p> <p>Convenios universidad-gobierno-empresa.</p> <p>Premios o reconocimientos por impacto social de proyectos.</p>	<p>Diseño mixto: Convergente paralelo.</p> <p>Método cuantitativo: Interrupción de Series Temporales (ITS) en variables comunitarias o sectoriales (por ejemplo, reducción de enfermedades, mejoras en productividad, acceso a servicios).</p> <p>Método cualitativo: Entrevistas semiestructuradas a líderes sociales, empresarios, funcionarios y usuarios sobre cambios derivados de la investigación.</p> <p>Instrumento complementario: Análisis documental de políticas, publicaciones, informes de proyecto.</p> <p>Teoría de cambio aplicada: La investigación orientada a la solución de problemas reales, con enfoque participativo, transforma capacidades institucionales, prácticas productivas y bienestar territorial.</p>

4.3. Caso 3.

Una, con presencia en zonas de frontera con poblaciones indígenas y campesinas, implementa desde hace una década un modelo de proyección social crítica con enfoque intercultural y territorial. Estudiantes y docentes co-diseñan intervenciones con comunidades para fortalecer la autonomía alimentaria y la soberanía territorial. Recientemente, los equipos rectorales desean sistematizar si estas experiencias han generado cambios significativos en las comunidades más allá de la actividad puntual (taller, brigada, campaña).

¿La relación universidad-comunidad ha transformado prácticas, empoderado actores o incidido en políticas locales? ¿Cómo perciben las comunidades estos impactos?

Tabla 8. Caso 3 de evaluación de impacto

Ámbito de la educación superior	Posibles dilemas de la evaluación	Información disponible en la Institución	Propuesta metodológica de evaluación de impacto
Impacto de la proyección social en comunidades	<p>¿Cómo atribuir cambios comunitarios a la acción universitaria y no a otros factores?</p> <p>¿Qué se entiende como “impacto” desde la mirada de las comunidades?</p> <p>¿Cómo no caer en una lógica asistencialista o de “intervención extractiva”?</p> <p>¿Cómo garantizar que los indicadores incluyan la voz comunitaria?</p>	<p>Proyectos documentados en sistemas de extensión y proyección social.</p> <p>Materiales producidos con las comunidades (cartillas, audiovisuales, memorias).</p> <p>Encuestas o diagnósticos participativos aplicados antes o durante los proyectos.</p> <p>Contacto activo con líderes comunitarios, organizaciones sociales y gobiernos locales.</p>	<p>Diseño mixto:</p> <p>Triangulado con enfoque participativo.</p> <p>Método cuantitativo: Diferencias en Diferencias (DiD) para evaluar cambios en comunidades con y sin intervención universitaria (ej., mejora en prácticas agrícolas, organización barrial, participación política).</p> <p>Método cualitativo: Cartografía social y grupos focales con miembros de la comunidad para reconstruir procesos y resignificar cambios.</p> <p>Instrumento complementario: Diarios reflexivos de estudiantes o actores universitarios en campo.</p> <p>Teoría de cambio aplicada: Las relaciones horizontales y continuas entre universidad y comunidad permiten generar apropiación del conocimiento, transformación cultural y fortalecimiento del tejido social.</p>

4.4. Caso 4

En 2019, la Facultad de Educación de una universidad realizó una reforma curricular profunda para todos sus programas de licenciatura. Se adoptó un modelo por competencias, se integraron TIC, se fortaleció el componente de prácticas pedagógicas y se incluyó un módulo obligatorio de pedagogías decoloniales.

El rediseño fue impulsado por exigencias de acreditación y por la necesidad de actualizar los perfiles de egreso. Hoy, en 2025, egresa la tercera cohorte bajo el nuevo plan, y surgen preguntas desde la decanatura:

¿Las reformas han mejorado efectivamente la calidad del aprendizaje?
¿Los nuevos egresados están mejor preparados para los retos escolares contemporáneos?

La facultad quiere evaluar si el rediseño ha tenido efectos diferenciales en el aprendizaje, la inserción laboral y la percepción de los actores educativos.

Tabla 9. Caso 4 de evaluación de impacto

Ámbito de la educación superior	Posibles dilemas de la evaluación	Información disponible en la Institución	Propuesta metodológica de evaluación de impacto
Impacto de las reformas curriculares	<p>¿Cómo aislar los efectos de la reforma curricular de otros cambios institucionales o contextuales?</p> <p>¿La percepción de mejora se traduce en aprendizaje real?</p> <p>¿Se afectó la equidad o inclusión con la nueva malla?</p> <p>¿Cómo incluir la voz del estudiantado en la evaluación?</p>	<p>Planes de estudio antes y después de la reforma.</p> <p>Evaluaciones internas y externas de aprendizaje.</p> <p>Registros de prácticas, evaluaciones docentes, egreso, empleabilidad.</p> <p>Encuestas a egresados y empleadores.</p>	<p>Diseño mixto: Secuencial exploratorio.</p> <p>Método cualitativo: Entrevistas y análisis documental para identificar los cambios intencionados y sus supuestos. Se analizan resistencias, adecuaciones y percepciones.</p> <p>Método cuantitativo: Regresión Discontinua (RDD) si la reforma se implementó a partir de una cohorte específica. Se comparan resultados académicos, empleabilidad o desempeño en pruebas entre cohortes.</p> <p>Instrumento complementario: Encuestas de autoevaluación de competencias por parte de estudiantes y egresados.</p> <p>Teoría de cambio aplicada: Las reformas pedagógicas alineadas con competencias y pertinencia social elevan el aprendizaje significativo, la formación integral y la empleabilidad de los graduados.</p>

5. Recursos Complementarios

Tabla 10. Recursos complementarios para la evaluación de impacto

Herramienta Digital	Definición y características	Uso recomendado	Recomendaciones metodológicas
Plataformas de gestión de datos (REDCap, KoBoToolbox, ODK)	Plataformas diseñadas para la recolección estructurada, el almacenamiento seguro y el manejo longitudinal de datos en campo.	Estudios con múltiples instrumentos, recolección en zonas sin conectividad o investigaciones multicéntricas.	Establecer protocolos de respaldo; configurar lógica de formularios; capacitar al equipo en sincronización segura.
Aplicaciones de georreferenciación (ArcGIS, QGIS, Mapbox)	Herramientas de mapeo digital que permiten vincular datos con ubicaciones geográficas.	Evaluación de impacto territorial, mapeo de cobertura, análisis de brechas geoespaciales.	Definir variables espaciales con anticipación; capacitar en SIG; asegurar confidencialidad de localizaciones sensibles.
Sistemas de panel en línea (Alchemer, Panelbit, Prolific)	Plataformas que permiten acceder a paneles de usuarios preseleccionados para responder encuestas representativas.	Estudios con necesidad de representatividad nacional o regional rápida.	Asegurar criterios de segmentación del panel; validar consistencia de respuestas; verificar trazabilidad y consentimiento.
Herramientas de análisis automatizado de texto (Leximancer, Voyant Tools, ChatGPT en modo análisis)	Programas que extraen temas, relaciones y patrones de texto de manera automatizada.	Análisis exploratorio de grandes volúmenes de texto en narrativas o encuestas abiertas.	Utilizar como complemento, no sustituto, del análisis cualitativo profundo; validar resultados con codificación manual.
Plataformas de colaboración (Miro, Mural, Notion, Google Jamboard)	Herramientas visuales para co-construcción remota de mapas, procesos y consensos.	Aplicación del método Delphi, desarrollo de teoría de cambio participativa, análisis colaborativo de resultados.	Establecer roles claros en las sesiones colaborativas; registrar iteraciones; combinar con actas o análisis de contenido.
Grabadores móviles de experiencia (EthnoApp, dscout, EpiCollect5)	Aplicaciones móviles para registrar datos etnográficos o diarios de campo en tiempo real por los propios participantes.	Estudios sobre trayectorias estudiantiles, experiencia institucional o seguimiento de prácticas pedagógicas.	Brindar formación previa a participantes; garantizar privacidad; definir con claridad los momentos y formatos de registro.

6. Evaluación

1. Una universidad implementa un sistema de alertas tempranas para identificar estudiantes en riesgo académico. Desea saber si este sistema ha mejorado la tasa de retención estudiantil al año siguiente. No fue posible asignar estudiantes aleatoriamente. ¿Qué método es el más adecuado?
 - A. Delphi
 - B. Propensity Score Matching
 - C. Observación participante
 - D. Grupo focal
2. Una facultad busca evaluar si la introducción de aulas híbridas ha incrementado el rendimiento académico en asignaturas teóricas. Dispone de notas académicas de cinco años antes y después de la implementación. ¿Qué diseño es el más pertinente?
 - A. Serie de Tiempo Interrumpida
 - B. Entrevistas a docentes
 - C. Regresión discontinua
 - D. Prueba de hipótesis t de Student
3. Una universidad desea conocer cómo la participación en proyectos de investigación afecta la empleabilidad de sus estudiantes, considerando que quienes participan pueden tener mejores promedios. ¿Qué técnica permite abordar este sesgo?
 - A. Diseño exploratorio cualitativo
 - B. Variables Instrumentales
 - C. RCT
 - D. Encuesta cerrada
4. Un programa de posgrado en innovación social busca evaluar el cambio en la percepción de liderazgo transformador en sus egresados a tres años de graduados. ¿Qué instrumento cualitativo sería más útil?
 - A. Cuestionario con escala de Likert
 - B. Prueba estandarizada de liderazgo
 - C. Historias de vida
 - D. Cartografía social
5. Una universidad pública implementó una política de gratuidad parcial en matrículas. Se busca evaluar si esto aumentó el ingreso de mujeres cabeza de hogar, comparando con otras universidades que no aplicaron la medida. ¿Qué método es ideal?
 - A. RCT
 - B. Diferencias en Diferencias
 - C. Encuesta de percepción
 - D. Delphi

6. Una universidad lanza una cátedra de ciudadanía digital. Al finalizar el curso, quiere conocer el cambio en la comprensión de derechos digitales. ¿Cuál instrumento es el más apropiado para medir este cambio con enfoque cuantitativo?
- A. Portafolio reflexivo
 - B. Encuesta con escalas validadas
 - C. Cartografía social
 - D. Grupo focal
7. Un grupo de investigación desea medir cómo la participación en asociaciones estudiantiles impacta el desarrollo de competencias ciudadanas. Se quiere evitar el sesgo de autoselección. ¿Qué método conviene usar?
- A. Variables Instrumentales
 - B. ITS
 - C. Entrevistas semiestructuradas
 - D. Observación estructurada
8. Una universidad inicia una estrategia de mentorías entre pares en carreras STEM. Se quiere identificar cómo esta práctica afecta el sentido de pertenencia de estudiantes mujeres. ¿Qué instrumento cualitativo es ideal?
- A. Escala de Likert
 - B. Grupo focal
 - C. Diarios reflexivos
 - D. Análisis documental
9. Una universidad que forma docentes rurales implementa un programa de aula invertida. Se busca evaluar si la nueva estrategia mejora la participación de los estudiantes en clase. ¿Qué diseño experimental podría aplicar si se puede controlar el entorno?
- A. Serie de Tiempo Interrumpida
 - B. RCT
 - C. DiD
 - D. Matching exacto
10. Un programa de extensión universitaria implementa talleres con comunidades étnicas sobre autonomía alimentaria. Para representar los aprendizajes colectivos y el impacto en las prácticas comunitarias, ¿qué técnica sería más pertinente?
- A. Prueba objetiva
 - B. Grupo focal
 - C. Cartografía social
 - D. Análisis de regresión

Retroalimentación:

Pregunta 1. Respuesta correcta: B. Propensity Score Matching. El PSM permite construir grupos comparables a partir de variables observadas para estimar el impacto del sistema de alertas sin necesidad de aleatorización.

Pregunta 2. Respuesta correcta: A. Serie de Tiempo Interrumpida. La ITS es útil para detectar cambios en la tendencia del rendimiento a lo largo del tiempo, antes y después de la implementación del aula híbrida.

Pregunta 3. Respuesta correcta: B. Variables Instrumentales. El uso de un instrumento (por ejemplo, asignación automática a semilleros según disponibilidad) permite corregir el sesgo de selección no aleatoria.

Pregunta 4. Respuesta correcta: C. Historias de vida. Las historias de vida permiten reconstruir trayectorias y transformaciones subjetivas a lo largo del tiempo, aportando profundidad al análisis del impacto del programa.

Pregunta 5. Respuesta correcta: B. Diferencias en Diferencias. Este método permite comparar cambios en las tasas de ingreso entre universidades tratadas y no tratadas antes y después de aplicar la política, estimando el efecto de la gratuidad.

Pregunta 6. Respuesta correcta: B. Encuesta con escalas validadas. Las encuestas estructuradas con escalas psicométricas permiten medir cambios en conocimientos o actitudes de forma cuantitativa y confiable.

Pregunta 7. A. Variables Instrumentales. Dado que la participación en asociaciones puede depender de motivaciones previas, un instrumento externo ayuda a estimar el impacto eliminando la endogeneidad.

Pregunta 8. Respuesta correcta: C. Diarios reflexivos. Los diarios permiten captar la evolución emocional y cognitiva de las participantes, siendo útiles para entender el impacto en el sentido de pertenencia.

Pregunta 9. Respuesta correcta: B. RCT. Si hay posibilidad de asignación aleatoria de docentes o grupos, el RCT es el método más riguroso para establecer causalidad en el efecto de la metodología innovadora.

Pregunta 10. Respuesta correcta: C. Cartografía social. La cartografía social permite visualizar, desde la perspectiva comunitaria, los cambios en el entorno, prácticas y saberes relacionados con la intervención.

7. Referencias Bibliográficas

Abad, F. and León, J. (2023). Inferencia causal en investigación educativa: análisis de la causalidad en estudios observacionales de carácter transversal. *Relieve - Revista Electrónica De Investigación Y Evaluación Educativa*, 29(2). <https://doi.org/10.30827/relieve.v29i2.26843>

Acuña, H., Carrasco, D., Carrasco, M., & Soto, A. (2019). El efecto del partido político del alcalde sobre variables de gasto municipal. *El Trimestre Económico*, 86(342), 343-404. <https://doi.org/10.20430/ete.v86i342.697>

Agudelo, H. and Loaiza, M. (2022). Aproximación a la gestión de los residuos de construcción y demolición en el área metropolitana del valle de aburrá. *Revista Cea*, 8(18), e2129. <https://doi.org/10.22430/24223182.2129>

Almeida, S. and Montenegro, D. (2023). La pizarra digital como herramienta de aprendizaje en el área de matemáticas.. *Mqinvestigar*, 7(1), 1971-1988. <https://doi.org/10.56048/mqr20225.7.1.2023.1971-1988>

Anatolyev, S. and Skolkova, A. (2019). Many instruments: implementation in stata. *The Stata Journal Promoting Communications on Statistics and Stata*, 19(4), 849-866. <https://doi.org/10.1177/1536867x19893627>

Arango, C., Torres, P., & Gonzáles, R. (2024). The pandemic as a factor for the exclusion of students in a situation of vulnerability at the upc. *Espoch Congresses the Ecuadorian Journal of S T E a M*, 3(2), 11-20. <https://doi.org/10.18502/espoch.v4i1.15796>

Arela-Bobadilla, R. (2023). Impacto de los linchamientos de delinquentes en la frecuencia de delitos denunciados en Perú durante 2017. *Debates en Sociología*, (56), 33-55. <https://doi.org/10.18800/debatesensociologia.202301.002>

Ato, M., García, J., & Benavente, A. (2013). Un sistema de clasificación de los diseños de investigación en psicología. *Anales De Psicología*, 29(3). <https://doi.org/10.6018/analesps.29.3.178511>

Austin, P. (2006). The performance of different propensity score methods for estimating marginal odds ratios. *Statistics in Medicine*, 26(16), 3078-3094. <https://doi.org/10.1002/sim.2781>

Austin, P. (2009). Some methods of propensity score matching had superior performance to others: results of an empirical investigation and monte carlo simulations. *Biometrical Journal*, 51(1), 171-184. <https://doi.org/10.1002/bimj.200810488>

Austin, P. (2010). The performance of different propensity-score methods for estimating differences in proportions (risk differences or absolute risk reductions) in observational studies. *Statistics in Medicine*, 29(20), 2137-2148. <https://doi.org/10.1002/sim.3854>

Austin, P. (2012). The performance of different propensity score methods for estimating marginal hazard ratios. *Statistics in Medicine*, 32(16), 2837-2849. <https://doi.org/10.1002/sim.5705>

Austin, P. (2013). A comparison of 12 algorithms for matching on the propensity score. *Statistics in Medicine*, 33(6), 1057-1069. <https://doi.org/10.1002/sim.6004>

Austin, P., Grootendorst, P., & Anderson, G. (2006). A comparison of the ability of different propensity score models to balance measured variables between treated and untreated subjects: a monte carlo study. *Statistics in Medicine*, 26(4), 734-753. <https://doi.org/10.1002/sim.2580>

Austin, P. and Stuart, E. (2015). Estimating the effect of treatment on binary outcomes using full matching on the propensity score. *Statistical Methods in Medical Research*, 26(6), 2505-2525. <https://doi.org/10.1177/0962280215601134>

Battocchio, M. (2018). La ventaja del oficialismo en las elecciones municipales de la provincia de buenos aires. *Económica*, 133-171. <https://doi.org/10.24215/18521649e006>

Bae, S., Lim, Y., Lee, Y., Shin, C., Oh, S., & Hong, Y. (2017). Maternal urinary bisphenol a concentration during midterm pregnancy and children's blood pressure at age 4. *Hypertension*, 69(2), 367-374. <https://doi.org/10.1161/hypertensionaha.116.08281>

Baldizzoni, E. (2014). Propuesta de proceso de transformación de datos para proyectos de explotación de información. *Revista Latinoamericana De Ingenieria De Software*, 1(2), 56. <https://doi.org/10.18294/relais.2013.56-70>

Barrientos, J., Cárdenas-Castro, M., & Gómez, F. (2014). Características sociodemográficas, bienestar subjetivo y homofobia en una muestra de hombres gay en tres ciudades chilenas. *Cadernos De Saúde Pública*, 30(6), 1259-1269. <https://doi.org/10.1590/0102-311x00108413>

Bastida-Molina, P., Vargas-Salgado, C., Montuori, L., & Alcázar-Ortega, M. (2020). Aplicación de la metodología clase inversa en la enseñanza de máquinas eléctricas avanzadas.. <https://doi.org/10.4995/inred2020.2020.11987>

Bayona-Umbarila, J. M., Rodríguez-Hernández, C. F., & Peña Ortega, L. O. (2022). Evaluación de Impacto en Instituciones de Educación Superior [Documento de Trabajo]. *Tecnológico de Monterrey*. <https://hdl.handle.net/11285/649985>.

Bernal, S., Abad, C., Bacuilima, W., & López, D. (2023). La vinculación con la sociedad como factor clave para acrecentar los niveles de calidad en la universidad ecuatoriana. *Debate Universitario*, 13(22), 39-53. <https://doi.org/10.59471/debate202346>

Blenker, P., Elmholdt, S., Frederiksen, S., Korsgaard, S., & Wagner, K. (2014). Methods in entrepreneurship education research: a review and integrative framework. *Education + Training*, 56(8/9), 697-715. <https://doi.org/10.1108/et-06-2014-0066>

Boszko, C. and Rosa, C. (2020). Diários reflexivos: definições e referenciais norteadores. *Revista Insignare Scientia - Ris*, 3(2), 18-35. <https://doi.org/10.36661/2595-4520.2020v3i2.11135>

Brítez, M. and Chung, C. (2022). Historia de vida como técnica de recolección de datos en ciencias administrativas. *Revista Científica Estudios E Investigaciones*, 11(2), 92-105. <https://doi.org/10.26885/rcei.11.2.92>

Burden, A., Roche, N., Miglio, C., Hillyer, E., Postma, D., Herings, R., ... & Price, D. (2017). An evaluation of exact matching and propensity score methods as applied in a comparative effectiveness study of inhaled corticosteroids in asthma. *Pragmatic and Observational Research*, Volume 8, 15-30. <https://doi.org/10.2147/por.s122563>

Caliendo, M. and Kopeinig, S. (2008). Some practical guidance for the implementation of propensity score matching. *Journal of Economic Surveys*, 22(1), 31-72. <https://doi.org/10.1111/j.1467-6419.2007.00527.x>

Campana, C. (2023). Diários reflexivos como aliados da aprendizagem e da avaliação. *Revista Brasileira De Casos De Ensino Em Administração*, a1. <https://doi.org/10.12660/gvcasosv13nespeciala1>

Campoverde-Luque, R., Almeida, O., Macias, G., Amador, M., & Morán, E. (2023). Aprendo contabilidad desde la reutilización de prendas de vestir: una propuesta para el proyecto de grado contable. *Ciencia Latina Revista Científica Multidisciplinar*, 7(3), 6128-6161. https://doi.org/10.37811/cl_rcm.v7i3.6616

Cardona Arias, J. (2020). Evaluación del impacto económico de programas sociales contra la pobreza: una revisión de estudios aleatorizados en la obra de Esther Duflo. *Revista Facultad Nacional De Salud Pública*, 38(2), 1-14. <https://doi.org/10.17533/udea.rfnsp.e338856>

Carrera, I. (2014). Efecto incumbente en elecciones municipales: un análisis de regresión discontinua para Guatemala. *Revista De Análisis Económico*, 29(2), 113-150. <https://doi.org/10.4067/s0718-88702014000200005>

Carriel, V. and Mendoza, C. (2021). La inserción escolar en el Ecuador y el papel del bono de desarrollo humano. *Revista Economía Y Política*, 86-103. <https://doi.org/10.25097/rep.n33.2021.05>

Cassiano, K. and Menezes, M. (2018). Analisando o desempenho da remoção de ruídos de uma série temporal de vazão de afluentes usando dbSCAN na abordagem análise espectral singular. *Pesquisa Operacional Para O Desenvolvimento*, 10(2), 93-107. <https://doi.org/10.4322/podes.2018.007>

Castillo, M., Álvarez, A., Prats, G., & Vidal, V. (2024). Knowledge management: strategies for the acquisition, management and transfer of knowledge in organizations. *Publicaciones E Investigación*, 18(1). <https://doi.org/10.22490/25394088.7588>.

Cely-Salazar, M. and Urquijo, A. (2022). Revisión sistemática de las características de evaluación curricular en programas académicos de pregrado a través del método prisma-nma. *Revista Electrónica Calidad en La Educación Superior*, 13(2), 150-174. <https://doi.org/10.22458/caes.v13i2.4415>

Chen, J. and Hsiang, C. (2019). Causal random forests model using instrumental variable quantile regression. *Econometrics*, 7(4), 49. <https://doi.org/10.3390/econometrics7040049>

Conde, J., Villamar, J., & Toledo, L. (2022). Relatos de vida de profesores universitarios. alteridad y otredad en las prácticas universitarias. Franz Tamayo - *Revista De Educación*, 4(10), 45-65. <https://doi.org/10.33996/franztamayo.v4i10.877>

Corrales-Liévano, J., Ruiz-Medina, C., & Angulo-Camargo, M. (2019). Evaluación de impacto del comercio electrónico en las ventas de las pymes en colombia: un análisis econométrico. *Clío América*, 13(26), 328-339. <https://doi.org/10.21676/23897848.3558>

Cruz Casallas, N. E., Álvarez Cardona, N., & Cárdenas Torrado, L. (2020). Impacto de la investigación dentro del proceso de formación profesional. *Cultura Educación Sociedad*, 11(2), 145–160. <https://doi.org/10.17981/culteducsoc.11.2.2020.09>

Delgado, C. and Reyes, M. (2022). Experiencias de evaluación docente durante la pandemia. voces de académicas en educación superior. *Revista Mais Educação*, (6 - Agosto 2022), 846-858. <https://doi.org/10.51778/2595-9611.v5i6p846-858>

Desai, R., Rothman, K., Bateman, B., Hernández-Díaz, S., & Huybrechts, K. (2017). A propensity-score-based fine stratification approach for confounding adjustment when exposure is infrequent. *Epidemiology*, 28(2), 249-257. <https://doi.org/10.1097/ede.0000000000000595>

Dym, N. and Lipman, Y. (2017). Exact recovery with symmetries for procrustes matching. *Siam Journal on Optimization*, 27(3), 1513-1530. <https://doi.org/10.1137/16m1078628>

Espinosa, G., Ángeles, J., & Flores, A. (2023). Capítulo 1: la política de evaluación de los planes de estudio en la educación superior., 23-36. <https://doi.org/10.59899/ges-cono-60-c1>

Estévez, E., Suárez, C., & Núñez, R. (2021). Proceso educativo en educación superior apoyado con recursos tic en tiempos de no presencialidad: perspectiva de los estudiantes de una facultad de educación, artes y humanidades. *Revista Boletín*.

Faúndez-Casanova, C. (2018). Estudio longitudinal del estado nutricional y de la condición física de estudiantes de enseñanza media de un liceo municipal de Talca, Chile. *Revista Ciencias De La Actividad Física*, 19(2), 1-10. <https://doi.org/10.29035/rcaf.19.2.4>

Feijoo-Criollo, E. and Guaraca, L. (2023). Política pública y bienestar social: efectos del programa barrio para mejor vivir en Cuenca-Ecuador, 2013-2017. *Estado & Comunes*, 1(16), 103-124. https://doi.org/10.37228/estado_comunes.v1.n16.2023.290

Flores-Cortez, O., Cruz, V., Callejas, M., & Vaquerano, C. (2023). Diseño e implementación de modelos a escala para el Museo Nacional de Aviación. un proyecto de extensión social universitaria. *Entorno*, 1(76), 11-20. <https://doi.org/10.5377/entorno.v1i76.17216>

Galindo-Lozano, D. and Correa, R. (2019). Lectura, escritura y oralidad en la escuela desde la perspectiva sociocultural. *Revista De Investigación Desarrollo E Innovación*, 10(1), 163-176. <https://doi.org/10.19053/20278306.v10.n1.2019.10020>

Gallardo, M., Palomar, M., García, E., & Palomar, M. (2020). Educar para la vida y la muerte a los estudiantes de enfermería: “del miedo al amor”. *Enfermería Cuidados Humanizados*, 9(1), 63-79. <https://doi.org/10.22235/ech.v9i1.2167>

Gallego, J., Hoffmann, B., Ibararán, P., Medina, M., Pecha, C., Romero, O., ... & Vera-Cossio, D. (2021). Impactos del programa ingreso solidario frente a la crisis del COVID-19 en Colombia. <https://doi.org/10.18235/0003261>

García, Y., Orozco, L., & Martín, G. (2016). Comparación de dos procedimientos de enseñanza universitaria: un ejemplo de interteaching. *Psicología Escolar E Educativa*, 20(3), 493-501. <https://doi.org/10.1590/2175-3539201502031029>

García-Ruiz, R., Liaño, S., Fernández, N., & Arregui, E. (2010). Estudio exploratorio de las percepciones del profesorado universitario respecto a la gestión de la docencia. *Educación Xx1*, 13(2). <https://doi.org/10.5944/educxx1.13.2.242>

Gaul, M., Vargas, C., & Olmo, P. (2021). Autogestión de la evaluación virtual en el campo de la programación. *Revista Iberoamericana De Tecnología en Educación Y Educación en Tecnología*, (28), e13. <https://doi.org/10.24215/18509959.28.e13>

Global (2015). *Theory of Change, How to navigate towards positive change in complex social settings.*

Goenner, C. (2013). Mission accomplished: a reply to Reuveny and Keshk. *Conflict Management and Peace Science*, 30(1), 19-23. <https://doi.org/10.1177/0738894212470792>

Gómez, S., Gallardo, B., & Vidal, P. (2023). Influencia del entorno físico y social en el rol de estudiantes universitarios, en modo virtual de aprendizaje, en el contexto de COVID-19. *Revista Chilena De Terapia Ocupacional*, 24, 17. <https://doi.org/10.5354/0719-5346.2023.65581>

González, A., Hincapié, M., Builes, G., & Molano-Builes, P. (2018). Tendencias de evaluación en promoción de la salud. actualización del debate en la década 2005-2015. *Hacia La Promoción De La Salud*, 24(1), 123-137. <https://doi.org/10.17151/hpsal.2019.24.1.11>

Haerizadeh, M. and M, V. (2019). Impacts of lean six sigma on improving a higher education system: a case study. *International Journal of Quality & Reliability Management*, 36(6), 983-998. <https://doi.org/10.1108/ijqrm-07-2018-0198>

Hao, J. and Fang-ai, L. (2024). Improving long-term multivariate time series forecasting with a seasonal-trend decomposition-based 2-dimensional temporal convolution dense network. *Scientific Reports*, 14(1). <https://doi.org/10.1038/s41598-024-52240-y>

Harris, A., McGregor, J., Perencevich, E., Furuno, J., Zhu, J., Peterson, D. & Finkelstein, J. (2006). The use and interpretation of quasi-experimental studies in medical informatics. *Journal of the American Medical Informatics Association*, 13(1), 16-23. <https://doi.org/10.1197/jamia.m1749>

Iacus, S., King, G., & Porro, G. (2012). Causal inference without balance checking: coarsened exact matching. *Political Analysis*, 20(1), 1-24. <https://doi.org/10.1093/pan/mpr013>

Igiebor, G. And Okonmah, E. (2022). The application and implication of combining quantitative and qualitative data in the social sciences: a third methodological movement in context. *Wilberforce Journal of the Social Sciences*, 7(2), 72-91. <https://doi.org/10.36108/wjss/2202.70.0250>

Imbens, G. (2014). Instrumental variables: an econometrician's perspective. *Statistical Science*, 29(3). <https://doi.org/10.1214/14-sts480>

Jiménez, L. and Tarazona, M. (2024). Practicas docentes en aulas multigrados del sector rural. *CIE*, 1(16). <https://doi.org/10.24054/cie.v1i16.2428>

Johnson, R. and Onwuegbuzie, A. (2004). Mixed methods research: a research paradigm whose time has come. *Educational Researcher*, 33(7), 14-26. <https://doi.org/10.3102/0013189x033007014>

Limaico, S., Jaramillo, R., Saquicela, M., & Salguero, J. (2022). Procesos de convergencia en américa latina: el papel de la internacionalización de la universidad. *Revista Scientific*, 7(24), 344-360. <https://doi.org/10.29394/scientific.issn.2542-2987.2022.7.24.18.344-360>

Liu, Y., Gong, C., Yang, L., & Chen, Y. (2020). Dstp-rnn: a dual-stage two-phase attention-based recurrent neural network for long-term and multivariate time series prediction. *Expert Systems With Applications*, 143, 113082. <https://doi.org/10.1016/j.eswa.2019.113082>

Lizcano, P. and Amaya, A. (2014). Percepciones ambientales de los habitantes del parque natural páramo de miraflores en colombia. estudio de caso en la vereda las mercedes, municipio de garzón. Entornos, (28), 13. <https://doi.org/10.25054/01247905.519>

Longva, K. and Foss, L. (2018). Measuring impact through experimental design in entrepreneurship education: a literature review and research agenda. Industry and Higher Education, 32(6), 358-374. <https://doi.org/10.1177/0950422218804912>

López, I. (2023). Nuevas metodologías y recursos en el aula universitaria de enseñanzas de segundas lenguas: alemán. EstFA, 9, 13-32. <https://doi.org/10.21071/estfa.v9i.15831>

(2017). La estructura curricular en la formación universitaria en psicología en colombia., 289-313. <https://doi.org/10.35985/9789588920795>

López-Meseguer, R. and Valdés, M. (2020). La evaluación comprensiva de programas educativos: ¿un nuevo paradigma teórico?. Revista Iberoamericana De Evaluación Educativa, 13(2), 85-105. <https://doi.org/10.15366/riee2020.13.2.005>

Luciano, C. and Velázquez, E. (2022). Oferta de formación en competencias socioemocionales para la empleabilidad en estudiantes de turismo. Revista De Innovación Y Buenas Prácticas Docentes, 11(1), 22-34. <https://doi.org/10.21071/ripadoc.v11i1.13578>

Mabeba, M. (2024). The effect of military expenditure on economic prosperity in croatia., 4-20. <https://doi.org/10.57125/fs.2024.06.20.01>

Manzanares, M. and Llamazares, M. (2021). Efectividad de los métodos mixtos en investigación contextual en salud y educación. New Trends in Qualitative Research, 28-40. <https://doi.org/10.36367/ntqr.5.2021.28-40>

Mardones, C. (2017). ¿alternativa privada a programas sociales de capacitación para población vulnerable?. Horizontes Empresariales, 15(2), 4-29. <https://doi.org/10.22320/hem.v15i2.2809>

Martínez-Padilla, J. and Pérez-González, J. (2008). Efecto de la trayectoria académica en el desempeño de estudiantes de ingeniería en evaluaciones nacionales. Formación Universitaria, 1(1). <https://doi.org/10.4067/s0718-50062008000100002>

Méndez-Mantuano, M., Morán, M., Caviedes, E., Jiménez, K., & Robles, D. (2021). Metodología para evaluación y actualización de los contenidos del syllabus. European Scientific Journal Esj, 17(32), 90. <https://doi.org/10.19044/esj.2021.v17n32p90>

Mera, S. and Quintero-Rendón, D. (2019). Impactos iniciales de la financiación de la educación superior con recursos del presupuesto participativo: evidencia para medellín. Sociedad Y Economía, (37). <https://doi.org/10.25100/sye.v0i37.7826>

Mesa, C., Arango, V., & López, L. (2021). La escuela es territorio. cartografía social de experiencias pedagógicas en instituciones educativas de medellín y bello, colombia. Territorios, (44-Esp.). <https://doi.org/10.12804/revistas.urosario.edu.co/territorios/a.9024>

Morales–Maza, J. (2020). Un acercamiento multidisciplinario al rol orientador del docente en el contexto comunitario y educativo. Conocimiento Educativo, 7, 39-59. <https://doi.org/10.5377/ce.v7i0.10029>

Muñoz, D., Muñoz, D., & Ramírez-López, A. (2016). El problema del voceador con incertidumbre paramétrica. Información Tecnológica, 27(4), 183-192. <https://doi.org/10.4067/s0718-07642016000400020>

Ñáñez-Rodríguez, J., Guerrero, J., & Bernal-Castillo, E. (2019). Ambientes digitales de aprendizaje en educación a distancia para la formación inicial de docentes: percepciones acerca de su pertinencia. Revista De Investigación Desarrollo E Innovación, 10(1), 107-119. <https://doi.org/10.19053/20278306.v10.n1.2019.10015>

Olmos-Gómez, M., Luque-Suárez, M., & Mohand, L. (2022). La calidad en educación superior. comparativa entre profesorado y alumnado. International Journal of Developmental and Educational Psychology Revista Infad De Psicología, 1(1), 361-366. <https://doi.org/10.17060/ijodaep.2022.n1.v1.2392>

Otero-Saborido, F. and Ramos, F. (2019). Validación de un instrumento para la valoración de la globalidad de los estándares curriculares: estudio de caso del currículo de educación física en andalucía. Revista Infancia Educación Y Aprendizaje, 5(2), 379. <https://doi.org/10.22370/ieya.2019.5.2.1543>

Pampaka, M., Williams, J., & Homer, M. (2016). Is the educational ‘what works’ agenda working? critical methodological developments. International Journal of Research & Method in Education, 39(4), 345-348. <https://doi.org/10.1080/1743727x.2016.1205817>

Pastor, V. (2011). El papel de la evaluación formativa en la evaluación por competencias: aportaciones de la red de evaluación formativa y compartida en docencia universitaria. Redu Revista De Docencia Universitaria, 9(1), 159. <https://doi.org/10.4995/redu.2011.6185>

Paz-Bailey, G., Miller, W., Shiraishi, R., Jacobson, J., Abimbola, T., & Chen, S. (2013). Reaching men who have sex with men: a comparison of respondent-driven sampling and time-location sampling in guatemala city. Aids and Behavior, 17(9), 3081-3090. <https://doi.org/10.1007/s10461-013-0589-7>

Pomeranz, D. (2011). Métodos de Evaluación. Harvard Business School. Consultado en: https://www.hbs.edu/ris/Supplemental%20Files/Metodos-de-Evaluacion-de-Impacto_50067.pdf

Rey-Fuentes, R., Jiménez-Maldonado, M., Arroyo-Medrano, M., & Villaseñor-Cabrera, T. (2021). El uso de programas computarizados y su efectividad en la rehabilitación de funciones ejecutivas en daño cerebral adquirido. *Revista Ecuatoriana De Neurología*, 30(1), 135-144. <https://doi.org/10.46997/revecuatneurol30100135>

Rivera, F., Carrasco, P., & Vargas, C. (2024). Aplicación de una metodología cualitativa de evaluación de impacto en la investigación aplicada e innovación, en una institución de educación superior técnico-profesional. *Calidad en La Educación*, (61). <https://doi.org/10.31619/caledu.n61.1480>

Rocha, J., Arrondo, V., & Morilla, M. (2020). Apropiación social del conocimiento y bien común: una propuesta para evaluar el impacto de la educación superior. *Indagare*, (8). <https://doi.org/10.35707/indagare/805>

Román, J. (2024). Métodos para evaluar políticas e intervenciones públicas. *Revista De Gobierno Administración Y Políticas De Salud*, 3, 2. <https://doi.org/10.5944/gaps.2024.00002>

Rose, R. and Stone, S. (2011). Instrumental variable estimation in social work research: a technique for estimating causal effects in nonrandomized settings. *Journal of the Society for Social Work and Research*, 2(2), 76-88. <https://doi.org/10.5243/jsswr.2011.4>

Saputra, R. (2021). Policy evaluation of regional expansion and determination regional boundaries in banten and west java province. *Sosiohumaniora*, 23(3), 400. <https://doi.org/10.24198/sosiohumaniora.v23i3.34597>

Saputra, K., Artini, L., & Utami, I. (2023). Parental support towards young learner's english achievement. *Jurnal Pendidikan Bahasa Inggris Undiksha*, 11(1), 83-90. <https://doi.org/10.23887/jpbi.v11i1.37412>

Sax, C. and Steiner, P. (2013). Temporal disaggregation of time series. *The R Journal*, 5(2), 80. <https://doi.org/10.32614/rj-2013-028>

Sekula, P., M, F., Pattaro, C., & Köttgen, A. (2016). Mendelian randomization as an approach to assess causality using observational data. *Journal of the American Society of Nephrology*, 27(11), 3253-3265. <https://doi.org/10.1681/asn.2016010098>

Sunarwibowo, R., Ikhsan, M., Mahi, B., & Wisana, I. (2024). Efficiency of agricultural cooperative members in indonesia. *Asian Economic and Financial Review*, 14(5), 367-388. <https://doi.org/10.55493/5002.v14i5.5070>

Tao, Q., Gao, T., Hu, L., & Wang, Z. (2018). Optimal selection and application analysis of multi-temporal differential interferogram series in stamps-based sbas insar. *European Journal of Remote Sensing*, 51(1), 1070-1086. <https://doi.org/10.1080/22797254.2018.1542977>

Thannimalai, T. (2021). Teachers' intention to continue using virtual learning environment (vle). *International Journal of Academic Research in Progressive Education and Development*, 10(2). <https://doi.org/10.6007/ijarped/v10-i2/10135>

Trad, L. (2009). Grupos focais: conceitos, procedimentos e reflexões baseadas em experiências com o uso da técnica em pesquisas de saúde. *Physis Revista De Saúde Coletiva*, 19(3), 777-796. <https://doi.org/10.1590/s0103-73312009000300013>

Tsiboe, F. and Turner, D. (2023). Econometric identification of crop insurance participation. *Agricultural and Resource Economics Review*, 52(3), 476-497. <https://doi.org/10.1017/age.2023.13>

Urunaga, S. and Cabrera, T. (2024). Estrés, ansiedad y depresión en los estudiantes de la facultad de ciencias de la salud - upap luque. *Revista Científica Upap*, 4(1), 14-25. <https://doi.org/10.54360/rcupap.v4i1.165>

Valle, J., Berdún, C., Capitán, J., Pavón, F., Romero-Cuevas, M., & Jiménez-Navarro, M. (2023). Evaluación de la adquisición de competencias en sistema cardiovascular en medicina: autopercepción, asistencia a clase y rendimiento académico. *Revista Española De Educación Médica*, 4(1). <https://doi.org/10.6018/edumed.548881>

Vanclay, F. (2003). International Principles For Social Impact Assessment. *Impact Assessment and Project Appraisal*, 21(1), 5–12. <https://doi.org/10.3152/147154603781766491>.

Vélez-Torres, I., Gaona, S., & Corredor, D. (2012). Cartografía social como metodología participativa y colaborativa de investigación en el territorio afrodescendiente de la cuenca alta del río cauca. *Cuadernos De Geografía Revista Colombiana De Geografía*, 21(2), 59-73. <https://doi.org/10.15446/rcdg.v21n2.25774>

Venegas, G. and Crispín, Á. (2018). Fricción del espacio y externalidades territoriales de la movilidad turística terrestre en américa central. *Pasos Revista De Turismo Y Patrimonio Cultural*, 16(4), 1023-1036. <https://doi.org/10.25145/j.pasos.2018.16.071>

Villota, J., Rosero, Á., & Urbano, L. (2024). Evaluación de resultados de aprendizaje rap en la licenciatura en educación básica primaria. *Fedumar Pedagogía Y Educación*, 11(1), 167-175. <https://doi.org/10.31948/fpe.v11i1.4254>

Wagner, B., Lorscheider, J., Wiencierz, A., Blackham, K., Psychogios, M., Bolliger, D., & Wegener, S. (2022). Endovascular treatment for acute ischemic stroke with or without general anesthesia: a matched comparison. *Stroke*, 53(5), 1520-1529. <https://doi.org/10.1161/strokeaha.121.034934>

Wang, Q., Liu, X., Wang, T., Zhu, Z., Yang, L., Guo, S., ... & Sun, Q. (2022). Primary care provider's job satisfaction and organizational commitment after covid-19 restrictions ended: a mixed-method study using a mediation model. *Frontiers in Psychology*, 13. <https://doi.org/10.3389/fpsyg.2022.873770>

Yang, M., McFowland, E., Burtch, G., & Adomavičius, G. (2019). Achieving reliable causal inference with data-mined variables: a random forest approach to the measurement error problem. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.3339983>

Yang, P., Ang, Z., & Berney, C. (2024). Restoration of intestinal continuity following emergency sigmoid colectomy for sigmoid volvulus: an american college of surgeons national surgical quality improvement program analysis using coarsened exact matching. *World Journal of Surgery*, 48(5), 1252-1260. <https://doi.org/10.1002/wjs.12138>



SEDE CENTRAL COCHABAMBA

Bolivia - Sur América

Campus Universitario Tiquipaya
c. Guillermina Martínez s/n
Telf: (591 – 4) 4318800

Torre Académica América

Av. América N°165 entre Túpac Amaru y Av. Libertador Bolívar
Telf: (591 – 4) 4150300

Edif. Polifuncional Ayacucho

Av. Ayacucho N°256
Telf: (591 – 4) 4150200

SEDE ACADÉMICA LA PAZ

Campus Miraflores
Av. Argentina N° 2083 esq. Nicaragua
Telf: (591 – 2) 2246725/6/7

SEDE ACADÉMICA SUCRE

Campus Las Delicias
Pasaje Guillermina de Ruiz N° 1 (Zona Bajo Delicias)
Telf: (591 – 4) 6441664

SEDE ACADÉMICA TRINIDAD

Campus El Gran Paititi
Av. Reyes s/n
Telf: (591 – 3) 4621238

SEDE ACADÉMICA SANTA CRUZ

Campus Eco Smart
Av. Banzer – Séptimo anillo y Av. Juan Pablo II