

EVALUACIÓN DEL EFECTO DE LA CAPACITACION DOCENTE EN PROGRAMAS STEM. ANALISIS DE LA PUNTUACION EN MATEMATICAS OBTENIDA POR LOS ALUMNOS MODALIDAD TECNICO/PROFESIONAL EN LAS PRUEBAS NACIONALES 2025.



Implicaciones para el diseno de una politica de
evaluacion de impactos de los programas de
capacitacion y formacion docente

**Informe Técnico elaborado por el LABCISFAM para ser
sometido a la consideración del** Instituto Nacional de
Formación y Capacitación del Magisterio (INAFOCAM) y
otras instituciones relacionadas.

□

Version preliminar no revisada por pares

Índice de Contenido

RESUMEN EJECUTIVO	4
1. INTRODUCCIÓN.....	5
2. ANTECEDENTES DEL PROGRAMA.....	5
3. MARCO CONCEPTUAL	6
3.1 La evaluación de programas de formación docente	6
3.2 Evaluación de resultados y evaluación de impacto	7
3.3 La teoría del cambio aplicada a la evaluación de impacto de la formación docente....	8
3.4 La transferencia pedagógica como mecanismo de cambio	9
3.5 El aprovechamiento institucional	9
3.6 Hacia un modelo integrado de evaluación.....	10
3.7 Implicaciones para el presente estudio	10
4. MARCO METODOLÓGICO.....	10
4.1 Diseño de la investigación.....	10
4.2 Modelo general de evaluación.....	11
4.3 Unidad de análisis.....	12
4.4 Variables del estudio	12
4.5 Variables derivadas.....	13
4.6 Hipótesis de investigación.....	13
4.7 Estrategia de análisis estadístico.....	14
4.8 Construcción del Índice de Aprovechamiento STEM (IAS).....	15
4.9 Criterios de clasificación	16
4.10 Limitaciones metodológicas	16
5. ANÁLISIS DESCRIPTIVO DE LOS RESULTADOS	17
5.1 Introducción.....	17
5.2 Descripción de la muestra	17
5.3 Distribución general de los resultados.....	18
5.4 Estadísticos descriptivos.....	18
5.5 Comparación con el promedio distrital	23
5.6 Comparación con el promedio regional.....	24
5.7 Comparación con el promedio nacional	26
5.8 Comparación integrada de los tres referentes.....	27
5.9 Heterogeneidad del desempeño institucional	28
5.10 Síntesis del análisis descriptivo	29
6. ANÁLISIS INFERENCIAL DE LOS RESULTADOS	29

6.1	Introducción.....	29
6.2	Hipótesis de investigación.....	29
6.3	Comparación con el promedio distrital	30
6.4	Comparación con el promedio regional.....	30
6.5	Comparación con el promedio nacional	31
6.6	Magnitud del efecto	32
6.7	Discusión de los resultados inferenciales	32
6.8	Limitaciones del análisis inferencial	33
6.9	Transición hacia el Índice de Aprovechamiento STEM (IAS).....	33
7.	ÍNDICE DE APROVECHAMIENTO STEM (IAS)	33
7.1	Introducción.....	33
7.2	Fundamentación conceptual	34
7.3	Principios metodológicos	34
7.4	Construcción del Índice Ponderado de Desempeño Relativo (IPDR).....	35
7.5	Normalización del índice.....	36
7.6	Clasificación del IAS.....	36
7.7	Resultados del IAS	36
7.8	Interpretación de los resultados	37
7.9	Aplicaciones del IAS	38
7.10	Limitaciones del índice.....	38
7.11	Validación y líneas futuras	38
8.	SISTEMA INTEGRADO PARA LA EVALUACIÓN DE PROGRAMAS DE FORMACIÓN DOCENTE (SIEPFD).....	39
8.1	Introducción.....	39
8.2	Principios del sistema	39
8.3	Arquitectura del sistema	40
8.4	Arquitectura funcional del SIEPFD.....	43
8.5	Flujo de información	43
8.6	Aplicaciones institucionales	44
8.7	Hoja de ruta para la implementación.....	44
9.	DISCUSIÓN DE LOS RESULTADOS	45
9.1	Introducción.....	45
9.2	El desempeño observado no constituye evidencia de impacto	45
9.3	La heterogeneidad institucional como principal hallazgo	46
9.4	El aporte del Índice de Aprovechamiento STEM (IAS)	46

9.5 Implicaciones para la gestión del INAFOCAM	46
9.6 Hacia una evaluación basada en evidencia.....	47
9.7 Aportes metodológicos del estudio	47
9.8 Limitaciones y agenda de investigación.....	47
9.9 Síntesis del capítulo	48
10. CONCLUSIONES Y RECOMENDACIONES	48
10.1 Conclusiones.....	48
10.2 Recomendaciones	50
10.3 Propuesta para la implementación del SIEPFD.....	51
10.4 Agenda de investigación.....	51
10.5 Consideraciones finales	52
Bibliografía citada	53

índice de Anexos

ANEXO 1: Resultados parametricos y no parametricos	54
ANEXO 2: Desarrollo metodológico y validación del Índice de Aprovechamiento STEM (IAS)	55
ANEXO 3	58
Desarrollo metodológico del Índice de Transferencia Pedagógica (ITP).....	58
ANEXO 4	60
Sistema Integrado para la Evaluación de Programas de Formación Docente (SIEPFD)	60

RESUMEN EJECUTIVO

El fortalecimiento de las competencias docentes constituye uno de los principales instrumentos de política pública para mejorar la calidad de los aprendizajes en los sistemas educativos. En este contexto, el Instituto Nacional de Formación y Capacitación del Magisterio (INAFOCAM) implementó un programa de capacitación docente en metodología STEM con énfasis en Matemática, dirigido a docentes de catorce politécnicos de la República Dominicana, y ejecutados por la UAFAM

El presente informe tiene como propósito evaluar el desempeño alcanzado por los politécnicos participantes, utilizando como indicador de referencia los resultados obtenidos por sus estudiantes en las Pruebas Nacionales de Matemática correspondientes al año 2025. Para ello, se comparan los promedios alcanzados por los centros intervenidos con los promedios de sus respectivos distritos educativos, regionales y del conjunto del sistema educativo nacional.

El estudio adopta un enfoque cuantitativo de carácter descriptivo e inferencial. Inicialmente se analizan las diferencias observadas entre los politécnicos y sus tres niveles de referencia. Posteriormente se aplican pruebas *t* para muestras pareadas, pruebas no paramétricas de Wilcoxon y estimaciones del tamaño del efecto mediante el estadístico de Cohen, con el propósito de determinar si las diferencias observadas presentan significación estadística.

Los resultados muestran que, en promedio, los politécnicos participantes presentan un desempeño ligeramente inferior al promedio de sus respectivos distritos y regionales, mientras que superan el promedio nacional. Sin embargo, las diferencias observadas no alcanzan significación estadística al nivel de confianza del cinco por ciento, por lo que no es posible atribuir, con la información disponible, un efecto estadísticamente verificable al programa de capacitación.

No obstante, el análisis revela una importante heterogeneidad entre los centros participantes. Algunos politécnicos superan simultáneamente los tres niveles de referencia, mientras otros permanecen por debajo de ellos. Esta variabilidad sugiere que la capacidad institucional para transformar la capacitación docente en mejoras del aprendizaje difiere significativamente entre los centros educativos.

Como principal innovación metodológica, el informe propone el **Índice de Aprovechamiento STEM (IAS)**, un indicador sintético diseñado para medir el desempeño relativo de cada politécnico respecto de sus referentes territoriales y nacionales. A diferencia de una comparación simple de medias, el IAS incorpora ponderaciones diferenciadas para el distrito, la regional y el promedio nacional, permitiendo identificar distintos niveles de aprovechamiento institucional de la capacitación recibida.

Finalmente, el informe presenta una propuesta de **Modelo Integrado de Evaluación de Programas de Formación Docente (MIEPFD)**, concebido como un marco metodológico para futuras evaluaciones del INAFOCAM. Este modelo integra cuatro niveles de análisis: implementación, transferencia pedagógica, aprovechamiento institucional e impacto educativo, proporcionando una arquitectura conceptual para el seguimiento y evaluación de programas de desarrollo profesional docente.

1. INTRODUCCIÓN

La mejora de la calidad de la educación constituye uno de los principales desafíos de los sistemas educativos contemporáneos. Entre los múltiples factores que influyen sobre el aprendizaje de los estudiantes, la calidad de la enseñanza ocupa un lugar central. En consecuencia, la formación inicial y continua del profesorado se ha consolidado como una de las estrategias más utilizadas para fortalecer las competencias docentes y promover mejores resultados educativos.

En la República Dominicana, el Instituto Nacional de Formación y Capacitación del Magisterio (INAFOCAM) desempeña un papel fundamental en el diseño e implementación de programas orientados al desarrollo profesional de los docentes. Durante los últimos años, el fortalecimiento de las competencias en Ciencia, Tecnología, Ingeniería y Matemática (STEM) ha adquirido especial relevancia como respuesta a la necesidad de desarrollar capacidades científicas y tecnológicas acordes con las demandas del siglo XXI.

En este contexto se implementó un programa de capacitación docente en metodología STEM con énfasis en Matemática, dirigido a docentes de catorce politécnicos del país. La intervención tuvo como propósito fortalecer las competencias pedagógicas para la enseñanza de la Matemática mediante metodologías activas, resolución de problemas, aprendizaje basado en proyectos e integración interdisciplinaria.

La presente investigación constituye una evaluación del desempeño alcanzado por los politécnicos participantes utilizando como indicador los resultados obtenidos por sus estudiantes en las Pruebas Nacionales de Matemática correspondientes al año 2025. El estudio compara dichos resultados con los promedios de los respectivos distritos educativos, regionales y del conjunto del sistema educativo nacional.

Es importante precisar que el presente trabajo **no corresponde a una evaluación de impacto en sentido estricto**. La ausencia de una línea de base previa a la intervención y de un grupo de comparación equivalente impide estimar el efecto causal atribuible al programa. En consecuencia, el análisis se concentra en el desempeño observado y en la identificación de diferencias relativas entre los centros participantes.

A pesar de esta limitación, el estudio incorpora una innovación metodológica mediante el desarrollo del **Índice de Aprovechamiento STEM (IAS)**, concebido como un indicador del desempeño relativo de los politécnicos participantes y de su capacidad institucional para traducir la capacitación docente en mejores resultados de aprendizaje.

2. ANTECEDENTES DEL PROGRAMA

El fortalecimiento de las competencias en Matemática constituye una prioridad estratégica para el sistema educativo dominicano, tanto por su incidencia en el desempeño académico como por su importancia para el desarrollo de competencias científicas, tecnológicas y de innovación.

En respuesta a esta necesidad, el INAFOCAM diseñó un programa de capacitación docente basado en el enfoque STEM, orientado a fortalecer las competencias pedagógicas de los docentes de Matemática.

En el caso de la UAFAM, programa se aplicó a docentes pertenecientes a catorce centros educativos distribuidos en diferentes regionales y distritos del país. Su diseño incorporó metodologías activas de enseñanza, estrategias de resolución de problemas, integración curricular y enfoques orientados al aprendizaje por competencias.

Desde una perspectiva conceptual, la intervención parte de la premisa de que la mejora de las competencias docentes constituye un insumo esencial para transformar las prácticas de enseñanza y, en consecuencia, mejorar los aprendizajes de los estudiantes.

Sin embargo, la literatura especializada reconoce que la relación entre capacitación docente y resultados educativos no es automática. Entre ambos extremos intervienen múltiples factores asociados al contexto institucional, al liderazgo pedagógico, a la cultura organizacional, a la disponibilidad de recursos y a la capacidad de los docentes para incorporar efectivamente las metodologías aprendidas en su práctica cotidiana.

Por esta razón, la evaluación de programas de formación docente requiere considerar no solo los resultados finales, sino también los procesos de transferencia pedagógica y las capacidades institucionales que facilitan o limitan el aprovechamiento de la intervención.

3. MARCO CONCEPTUAL

3.1 La evaluación de programas de formación docente

La formación continua del profesorado constituye uno de los pilares fundamentales para el mejoramiento de la calidad educativa. Durante las últimas décadas, organismos internacionales como la Organización para la Cooperación y el Desarrollo Económicos (OCDE), la UNESCO y el Banco Mundial han señalado que el desarrollo profesional docente representa una de las intervenciones con mayor potencial para fortalecer los aprendizajes, siempre que la capacitación responda a necesidades reales de los docentes, promueva cambios en la práctica pedagógica y cuente con mecanismos efectivos de seguimiento y evaluación.

La evidencia internacional muestra que la calidad de los sistemas educativos depende, en gran medida, de la capacidad de los docentes para transformar el currículo en experiencias significativas de aprendizaje. En consecuencia, la formación continua no debe entenderse como un conjunto de actividades aisladas, sino como un proceso sistemático de actualización profesional orientado a mejorar la práctica pedagógica y los resultados de los estudiantes.

Los estudios e investigaciones sobre los efectos e impactos de formación docente no son muy abundantes. En Desimone (2009) se presenta uno de los trabajos metodológicos más influyentes sobre evaluación de programas de desarrollo profesional docente. Propone una teoría de cambio que relaciona la formación docente, la modificación de las prácticas pedagógicas y los resultados de aprendizaje, proporcionando el fundamento conceptual del Índice de Transferencia Pedagógica (ITP).

En otro estudio pionero en la materia, (Yoon et al., 2007), se demuestra que los programas de desarrollo profesional docente suficientemente intensivos producen mejoras significativas en el rendimiento académico de los estudiantes.

No obstante, también existen diversos estudios han demostrado que la participación en programas de capacitación no garantiza, por sí sola, mejoras en el rendimiento académico. Los efectos esperados dependen de múltiples factores relacionados con la apropiación de los contenidos, el contexto institucional, el liderazgo pedagógico, el acompañamiento posterior a la formación y la capacidad de los docentes para incorporar nuevas estrategias de enseñanza en el aula. Al respecto, Timperley (et al., 2007) sintetizan la evidencia internacional sobre aprendizaje profesional docente y demostrando que la formación tiene mayor impacto cuando está centrada en el aprendizaje de los estudiantes y vinculada directamente a la práctica pedagógica.

En el ámbito nacional, el estudio o referencia más importante de los efectos de los programas de capacitación docente, y específicamente, los ejecutados por el INAFOCAM, es el de Ayala de la Cruz (2016). El mismo analiza la calidad del sistema de formación continua del INAFOCAM, su articulación con las políticas educativas y sus efectos sobre el desarrollo profesional y el desempeño docente. Además, propone avanzar hacia un sistema institucional de evaluación de la formación permanente, lo que constituye un antecedente directo del SIEPFD desarrollado en este informe.

A partir de lo anteriormente expuesto se puede concluir que existe un desafío metodológico importante para las instituciones responsables de la formación docente: **¿cómo evaluar si una capacitación produjo realmente cambios en los procesos de enseñanza y en los aprendizajes de los estudiantes?** Responder a esta pregunta constituye el eje central del presente estudio.

3.2 Evaluación de resultados y evaluación de impacto

En la literatura especializada se distinguen dos enfoques complementarios para valorar la efectividad de un programa educativo: la evaluación de resultados y la evaluación de impacto. La **evaluación de resultados** describe los cambios observados una vez finalizada la intervención. Su propósito consiste en determinar si los beneficiarios presentan un desempeño superior al de determinados grupos o referentes de comparación. Este enfoque resulta particularmente útil cuando no se dispone de información previa a la intervención o cuando las condiciones operativas impiden establecer grupos de control.

Por su parte, la **evaluación de impacto** busca estimar el efecto atribuible exclusivamente al programa. Para ello compara la situación observada con el escenario hipotético que habría ocurrido si los beneficiarios no hubiesen participado en la intervención. Este tipo de evaluación requiere diseños metodológicos más robustos, tales como experimentos aleatorios, diseños cuasi experimentales, modelos de diferencias en diferencias o métodos de emparejamiento estadístico.

La evaluación de impacto es quizás una de las herramientas fundamentales y más importantes para la toma de decisiones de asignación de recursos a los programas presupuestarios que reflejan la naturaleza y objetivos de las políticas públicas. Una de las referencias de mayor influencia se puede encontrar en Gertler (et al., 2016), en la cual se presenta la lógica metodológica utilizada para valorar los efectos de los programas de formación docente.

Otra publicación de mucha influencia se puede encontrar en Rossi, Lipsey y Henry, G. T. (2019), una obra clásica sobre evaluación de programas públicos, la cual proporciona

fundamento metodológico general del Sistema Integrado para la Evaluación de Programas de Formación Docente (SIEPFD).

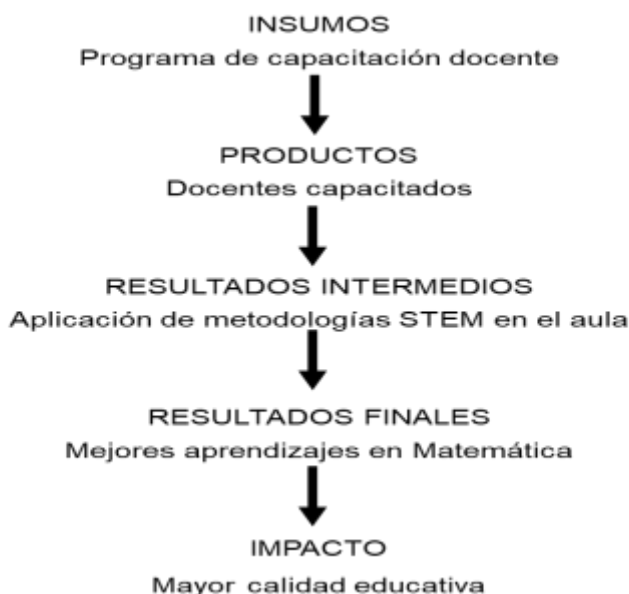
No obstante, es necesario aclarar que el presente estudio se inscribe dentro del primer enfoque de los dos mencionados al principio, es decir, evaluación de resultados. Los resultados obtenidos por los catorce politécnicos participantes se comparan con los promedios de sus respectivos distritos educativos, regionales y del conjunto del sistema nacional. Si bien esta estrategia permite identificar diferencias de desempeño, no posibilita establecer relaciones causales entre la capacitación docente y los resultados alcanzados.

Por esta razón, las conclusiones del informe se presentan como evidencia sobre el desempeño relativo de los centros participantes y no como una demostración del impacto del programa.

3.3 La teoría del cambio aplicada a la evaluación de impacto de la formación docente

Toda intervención educativa se fundamenta en una teoría del cambio, es decir, en un conjunto de supuestos que explican cómo determinadas acciones generan los resultados esperados, lo cual a su vez está íntimamente vinculado a la teoría de la evaluación de impacto. Al respecto, puede consultarse a White (2009), quien desarrolla el enfoque de evaluación basado en teoría del cambio, la que inspira la secuencia conceptual del SIEPFD: implementación, transferencia pedagógica, aprovechamiento institucional, e impacto. En el caso del programa de capacitación STEM, la teoría del cambio puede representarse como una cadena de resultados en la que cada etapa constituye la condición necesaria para la siguiente.

Figura 3.1. Teoría del cambio del programa STEM



Este esquema pone de manifiesto que la capacitación constituye únicamente el punto de partida de un proceso mucho más complejo. Entre la formación recibida por los docentes y

los resultados obtenidos por los estudiantes intervienen numerosos factores que pueden fortalecer o debilitar los efectos esperados de la intervención.

3.4 La transferencia pedagógica como mecanismo de cambio

Uno de los conceptos centrales del presente estudio es el de **transferencia pedagógica**.

Se entiende por transferencia pedagógica el proceso mediante el cual los conocimientos, habilidades y estrategias adquiridas durante la capacitación son incorporados efectivamente a la práctica docente cotidiana.

La transferencia constituye el puente entre la formación recibida y la mejora de los aprendizajes. Sin ella, incluso los programas de capacitación mejor diseñados difícilmente producirán cambios observables en el desempeño de los estudiantes.

La literatura internacional identifica diversos factores que favorecen la transferencia pedagógica (Visscher, et al., 2025), entre ellos:

- liderazgo directivo comprometido con la innovación;
- acompañamiento y retroalimentación posterior a la capacitación;
- trabajo colaborativo entre docentes;
- disponibilidad de recursos didácticos;
- tiempo para la planificación conjunta;
- cultura institucional orientada al aprendizaje y la mejora continua.

Estos factores ayudan a explicar por qué centros educativos que reciben una misma capacitación pueden alcanzar resultados muy diferentes.

3.5 El aprovechamiento institucional

El concepto de **aprovechamiento institucional** constituye uno de los aportes centrales del presente informe.

Se define como la capacidad de un centro educativo para convertir una intervención de desarrollo profesional docente en mejoras observables de los resultados de aprendizaje de sus estudiantes.

Desde esta perspectiva, el desempeño alcanzado por un politécnico depende no solo de la calidad de la capacitación recibida, sino también de las condiciones institucionales que facilitan la apropiación y aplicación efectiva de los conocimientos adquiridos.

El aprovechamiento institucional constituye, por tanto, una propiedad emergente de la organización escolar, resultado de la interacción entre liderazgo, gestión pedagógica, cultura institucional, trabajo colaborativo y compromiso docente.

Este concepto servirá de fundamento para el desarrollo del **Índice de Aprovechamiento STEM (IAS)** a ser presentado en capítulos posteriores.

3.6 Hacia un modelo integrado de evaluación

El análisis precedente pone de manifiesto que la evaluación de programas de formación docente no puede limitarse a la comparación de resultados académicos.

Un modelo integral debe considerar, al menos, cuatro dimensiones complementarias:

1. **Implementación del programa**, para verificar el cumplimiento de las actividades previstas.
2. **Transferencia pedagógica**, para evaluar la incorporación de los aprendizajes de la capacitación a la práctica docente.
3. **Aprovechamiento institucional**, para medir la capacidad de los centros educativos de transformar la formación en mejores resultados.
4. **Impacto educativo**, para estimar el efecto atribuible al programa mediante diseños de evaluación rigurosos.

Estas cuatro dimensiones constituyen la base conceptual del **Modelo Integrado de Evaluación de Programas de Formación Docente (MIEPFD)** que se desarrolla en los capítulos finales del informe.

3.7 Implicaciones para el presente estudio

A la luz del marco conceptual desarrollado, el análisis de los resultados de las Pruebas Nacionales debe interpretarse como una medida del desempeño observado de los politécnicos participantes y no como una prueba definitiva del impacto de la capacitación.

Sin embargo, la identificación de diferencias sistemáticas entre los centros participantes constituye una oportunidad para explorar nuevas herramientas de evaluación institucional. En este contexto surge el **Índice de Aprovechamiento STEM (IAS)**, concebido como un indicador que sintetiza el desempeño relativo de los centros y permite identificar distintos niveles de aprovechamiento de la intervención.

De este modo, el presente estudio no solo aporta evidencia empírica sobre los resultados del programa STEM, sino que propone un marco conceptual y metodológico susceptible de ser utilizado en la evaluación de futuras intervenciones de formación docente impulsadas por el INAFOCAM.

4. MARCO METODOLÓGICO

4.1 Diseño de la investigación

El presente estudio corresponde a una investigación **cuantitativa, no experimental, transversal** y de alcance **descriptivo-inferencial**, cuyo propósito consiste en evaluar el desempeño relativo de los politécnicos participantes en el Programa de Capacitación Docente STEM en Matemática, utilizando como variable de resultado los puntajes obtenidos por sus estudiantes en las Pruebas Nacionales de Matemática correspondientes al año 2025.

El diseño es **no experimental**, dado que las variables analizadas no fueron manipuladas por los investigadores y los datos corresponden a registros oficiales generados con posterioridad a la implementación del programa.

Es igualmente un estudio **transversal**, ya que analiza información correspondiente a un único período de observación.

Finalmente, posee un alcance **descriptivo e inferencial**, porque combina el análisis de los niveles de desempeño observados con procedimientos estadísticos destinados a determinar la significación de las diferencias entre los politécnicos participantes y sus respectivos referentes educativos.

Debe reiterarse que el presente estudio **no constituye una evaluación de impacto** en sentido estricto. La ausencia de una línea de base previa a la intervención y de un grupo de comparación equivalente impide estimar el efecto causal atribuible al programa de capacitación. En consecuencia, los resultados deben interpretarse como evidencia sobre el desempeño relativo de los centros participantes.

4.2 Modelo general de evaluación

La estrategia metodológica adoptada se estructura en cuatro niveles complementarios de análisis, cada uno orientado a responder una pregunta específica de evaluación.

Nivel I. Evaluación descriptiva

Pregunta:

¿Cómo se comportaron los politécnicos participantes respecto a sus referentes?

Indicadores:

- Promedio del politécnico.
- Promedio distrital.
- Promedio regional.
- Promedio nacional.
- Brechas de desempeño.

Nivel II. Evaluación inferencial

Pregunta:

¿Las diferencias observadas son estadísticamente significativas?

Herramientas utilizadas:

- prueba t para muestras pareadas;
- prueba de Wilcoxon;
- tamaño del efecto (Cohen's d);
- intervalos de confianza del 95 %.

Nivel III. Evaluación del aprovechamiento institucional

Pregunta:

¿Existen diferencias entre los politécnicos en la capacidad para transformar la capacitación en mejores resultados de aprendizaje?

Herramienta:

- Índice de Aprovechamiento STEM (IAS).

Nivel IV. Interpretación institucional

Pregunta:

¿Qué implicaciones tienen los resultados para futuras políticas de formación docente del INAFOCAM?

Producto:

- recomendaciones institucionales;
- Modelo Integrado de Evaluación de Programas de Formación Docente (MIEPFD).

4.3 Unidad de análisis

La unidad de análisis está constituida por los **14 politécnicos** participantes en el Programa de Capacitación Docente STEM con énfasis en Matemática.

Cada politécnico representa una unidad independiente de observación.

Para cada centro educativo se dispone de la siguiente información:

- promedio obtenido en las Pruebas Nacionales de Matemática;
- promedio del distrito educativo correspondiente;
- promedio de la regional educativa;
- promedio nacional.

4.4 Variables del estudio

Variable dependiente

Resultado del politécnico

Corresponde al promedio obtenido por los estudiantes del centro educativo en las Pruebas Nacionales de Matemática 2025.

Escala: Cuantitativa continua.

Variables de comparación

Se consideran tres variables de referencia.

Distrito: Promedio obtenido por el conjunto de centros pertenecientes al distrito educativo.

Regional: Promedio correspondiente a la regional educativa.

Nacional: Promedio nacional de Matemática.

4.5 Variables derivadas

Con el propósito de medir el desempeño relativo de cada politécnico se calcularon tres diferencias:

Brecha Distrital

$$D = PP - PD$$

donde:

PP = promedio del politécnico.

PD = promedio del distrito.

Brecha Regional

$$R = PP - PR$$

donde:

PR = promedio regional.

Brecha Nacional

$$N = PP - PN$$

donde:

PN = promedio nacional.

Las tres brechas pueden adoptar valores positivos o negativos. Un valor positivo indica que el politécnico supera el referente correspondiente. Un valor negativo indica un desempeño inferior.

4.6 Hipótesis de investigación

El estudio evalúa tres hipótesis.

Hipótesis descriptiva

Los politécnicos participantes presentan resultados iguales o superiores a sus respectivos referentes educativos.

Hipótesis inferencial

Las diferencias observadas entre los politécnicos y los referentes educativos son estadísticamente significativas.

Hipótesis institucional

Los politécnicos que presentan mejores resultados relativos evidencian una mayor capacidad institucional para transformar la capacitación docente en mejores aprendizajes de los estudiantes.

Esta tercera hipótesis constituye el fundamento conceptual del Índice de Aprovechamiento STEM.

4.7 Estrategia de análisis estadístico

La estrategia analítica comprende cuatro etapas sucesivas.

Primera etapa

Estadística descriptiva.

Se calcularon:

- medias;
- máximos;
- mínimos;
- desviaciones estándar;
- brechas de desempeño.

Segunda etapa

Comparación de medias.

Se utilizó la prueba t para muestras pareadas.

La hipótesis evaluada fue:

H_0 : No existen diferencias entre los politécnicos y el referente correspondiente.

H_1 : Los politécnicos presentan resultados superiores.

Tercera etapa

Validación no paramétrica.

Debido al reducido tamaño de la muestra ($n = 14$), se aplicó igualmente la prueba de rangos con signo de Wilcoxon, menos sensible al incumplimiento del supuesto de normalidad.

Cuarta etapa

Magnitud del efecto.

Se estimó el tamaño del efecto mediante el estadístico de Cohen (d), con la finalidad de evaluar la importancia práctica de las diferencias observadas.

La interpretación utilizada fue:

Valor	Interpretación
<0.20	Muy pequeño
0.20–0.49	Pequeño
0.50–0.79	Moderado
≥ 0.80	Grande

4.8 Construcción del Índice de Aprovechamiento STEM (IAS)

El IAS constituye el principal aporte metodológico del presente estudio.

Su finalidad consiste en sintetizar el desempeño relativo de cada politécnico respecto de los tres niveles de referencia.

Paso 1

Cálculo de las brechas.

- $D = PP - PD$
- $R = PP - PR$
- $N = PP - PN$

Paso 2

Construcción del Índice Ponderado de Desempeño Relativo (IPDR).

Se propone la siguiente expresión:

$$IPDR = (0.50 \times D) + (0.30 \times R) + (0.20 \times N)$$

Las ponderaciones responden al criterio de proximidad institucional.

Referente Peso

Distrito	50 %
Regional	30 %

Referente Peso

Nacional 20 %

El mayor peso asignado al distrito se justifica porque constituye el entorno educativo más comparable con el politécnico y refleja condiciones administrativas, territoriales y, en muchos casos, socioeducativas similares.

Paso 3

Normalización.

Con el propósito de facilitar la interpretación del indicador, el IPDR se transforma a una escala de 0 a 100 mediante la siguiente expresión (Organisation for Economic Co-operation and Development (OECD) y Joint Research Centre (JRC). (2008):

$$\text{IAS} = 100 \times [(\text{IPDR} - \text{IPDR mínimo}) / (\text{IPDR máximo} - \text{IPDR mínimo})]$$

En consecuencia:

- **100** representa el mayor desempeño relativo observado.
- **0** representa el menor desempeño relativo de la muestra.
- Los demás centros se distribuyen proporcionalmente entre ambos extremos.

4.9 Criterios de clasificación

Para facilitar la interpretación del IAS se propone la siguiente escala:

IAS Nivel de aprovechamiento

80–100 Muy alto

60–79.9 Alto

40–59.9 Medio

20–39.9 Bajo

0–19.9 Muy bajo

Esta clasificación tiene un propósito analítico y de gestión. No constituye un juicio sobre la calidad global de los centros educativos, sino una herramienta para identificar diferencias en el desempeño relativo y orientar acciones de acompañamiento.

4.10 Limitaciones metodológicas

Los resultados del presente estudio deben interpretarse considerando las siguientes limitaciones:

1. La evaluación se basa en una única medición posterior a la intervención, por lo que no es posible estimar cambios respecto a una situación inicial.

2. No se dispone de un grupo de comparación que permita atribuir causalmente las diferencias observadas al programa de capacitación.
3. El reducido tamaño de la muestra limita la potencia estadística para detectar diferencias pequeñas o moderadas.
4. No se incorporan variables contextuales —como nivel socioeconómico, experiencia docente, liderazgo directivo o recursos institucionales— que podrían influir en el desempeño de los centros.

En consecuencia, las conclusiones deben entenderse como evidencia sobre el desempeño relativo de los politécnicos participantes y como una base para el diseño de evaluaciones de impacto más robustas en futuras cohortes.

5. ANÁLISIS DESCRIPTIVO DE LOS RESULTADOS

5.1 Introducción

Este capítulo presenta el análisis descriptivo de los resultados obtenidos por los catorce politécnicos participantes en el Programa de Capacitación Docente STEM en Matemática, utilizando como indicador los puntajes alcanzados por sus estudiantes en las Pruebas Nacionales de Matemática 2025.

El análisis descriptivo constituye la primera etapa de la evaluación y tiene como finalidad caracterizar el comportamiento de los centros participantes antes de aplicar procedimientos de inferencia estadística. Para ello se examinan las diferencias entre los puntajes obtenidos por los politécnicos y los promedios de sus respectivos distritos educativos, regionales y del conjunto del sistema educativo nacional.

Asimismo, se analizan las brechas de desempeño y la distribución de los resultados con el propósito de identificar patrones generales y posibles diferencias entre los centros participantes.

5.2 Descripción de la muestra

La muestra está conformada por **catorce politécnicos** que participaron en el Programa de Capacitación Docente STEM con énfasis en Matemática.

Cada politécnico constituye una unidad independiente de observación.

Para cada centro se dispone de cuatro indicadores:

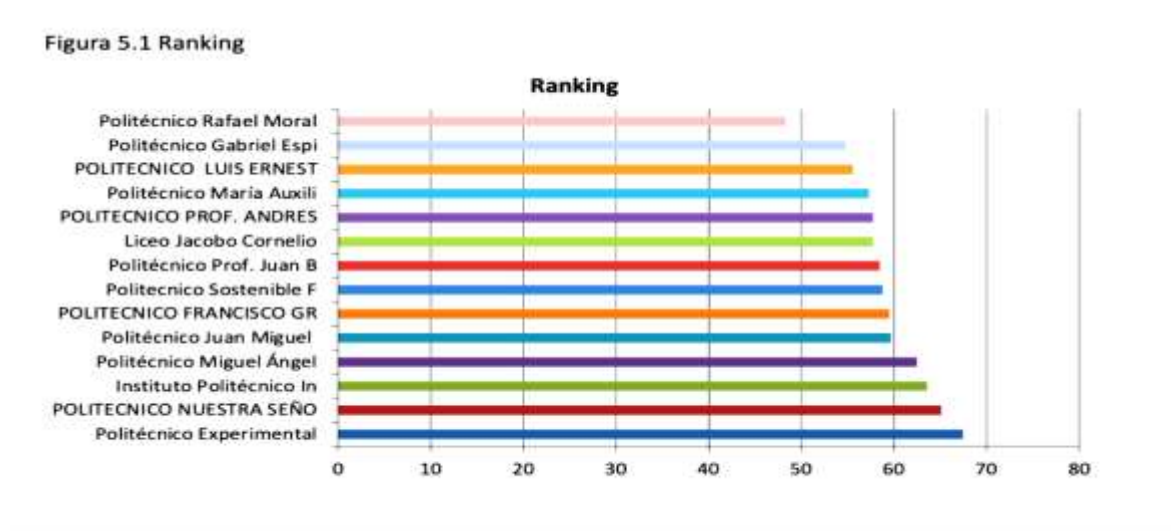
- Promedio del politécnico en las Pruebas Nacionales de Matemática.
- Promedio del distrito educativo correspondiente.
- Promedio de la regional educativa.
- Promedio nacional.

En conjunto, estos indicadores permiten realizar comparaciones a tres niveles territoriales y construir medidas del desempeño relativo de cada institución.

5.3 Distribución general de los resultados

La Figura 5.1 presenta la distribución de los puntajes obtenidos por los politécnicos participantes.

Figura 5.1. Distribución de los resultados de Matemática en los politécnicos participantes



La distribución evidencia una apreciable variabilidad entre los centros educativos. Mientras algunos politécnicos alcanzan resultados claramente superiores a los promedios de referencia, otros presentan desempeños inferiores.

Esta dispersión constituye un primer indicio de que la respuesta institucional a la intervención no fue homogénea.

5.4 Estadísticos descriptivos

La Tabla 5.1 resume los principales estadísticos descriptivos correspondientes a los puntajes obtenidos por los politécnicos participantes.

Tabla 5.1. Estadísticos descriptivos de los resultados en Matemática

Estadístico	Valor	indicador	Valor
Número de observaciones	14.00%	Politécnicos por encima del distrito	6.00%
Promedio	59.04%	Politécnicos por debajo del distrito	8.00%
Mediana	58.62%	% por encima	42.92%
Desviación estándar	4.74%	% por debajo	57.10%
Mínimo	48.33%	Brecha promedio	0.24%
Máximo	67.45%	Brecha máxima	6.84%
Rango	19.12%	Brecha mínima	-7.95%
Coeficiente de variación	8.03%		

La amplitud entre el valor máximo y el mínimo confirma la existencia de diferencias importantes en el desempeño de los centros participantes, lo que justifica realizar un análisis individual por politécnico.

5.5 Comparación con el promedio distrital

El distrito educativo constituye el referente territorial más próximo y, por tanto, el criterio de comparación más exigente.

La Tabla 5.2 resume las diferencias observadas entre los resultados de cada politécnico y el promedio de su distrito.

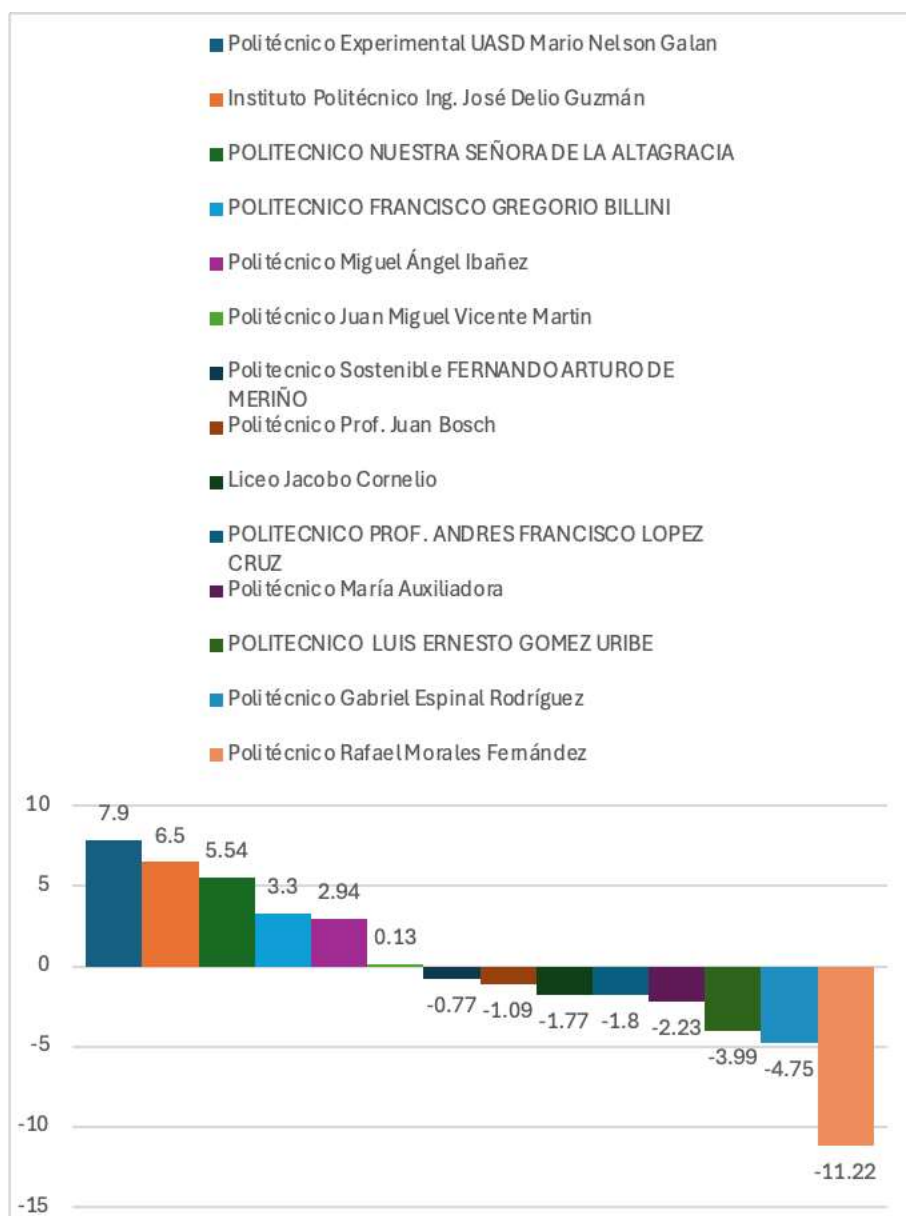
Tabla 5.2. Comparación con el promedio distrital

Indicador	Valor
Politécnicos por encima del distrito	6
Politécnicos por debajo del distrito	8
Diferencia promedio	-0.24 puntos

En términos porcentuales, el **42.9 %** de los politécnicos supera el promedio de su distrito, mientras que el **57.1 %** obtiene resultados inferiores.

Este comportamiento indica que el desempeño de los centros participantes es heterogéneo y que la capacitación no se tradujo en ventajas uniformes respecto al contexto distrital, lo cual queda reflejado en la figura 5.2 que se presenta más abajo.

Figura 5.2. Brechas Politécnico – Distrito



5.6 Comparación con el promedio regional

Según la gráfica 5.3 que sigue a continuación, la regional educativa representa un nivel intermedio de comparación. Los resultados muestran un patrón muy similar al observado respecto al distrito.

Figura 5.3. Brechas Politécnico – Regional

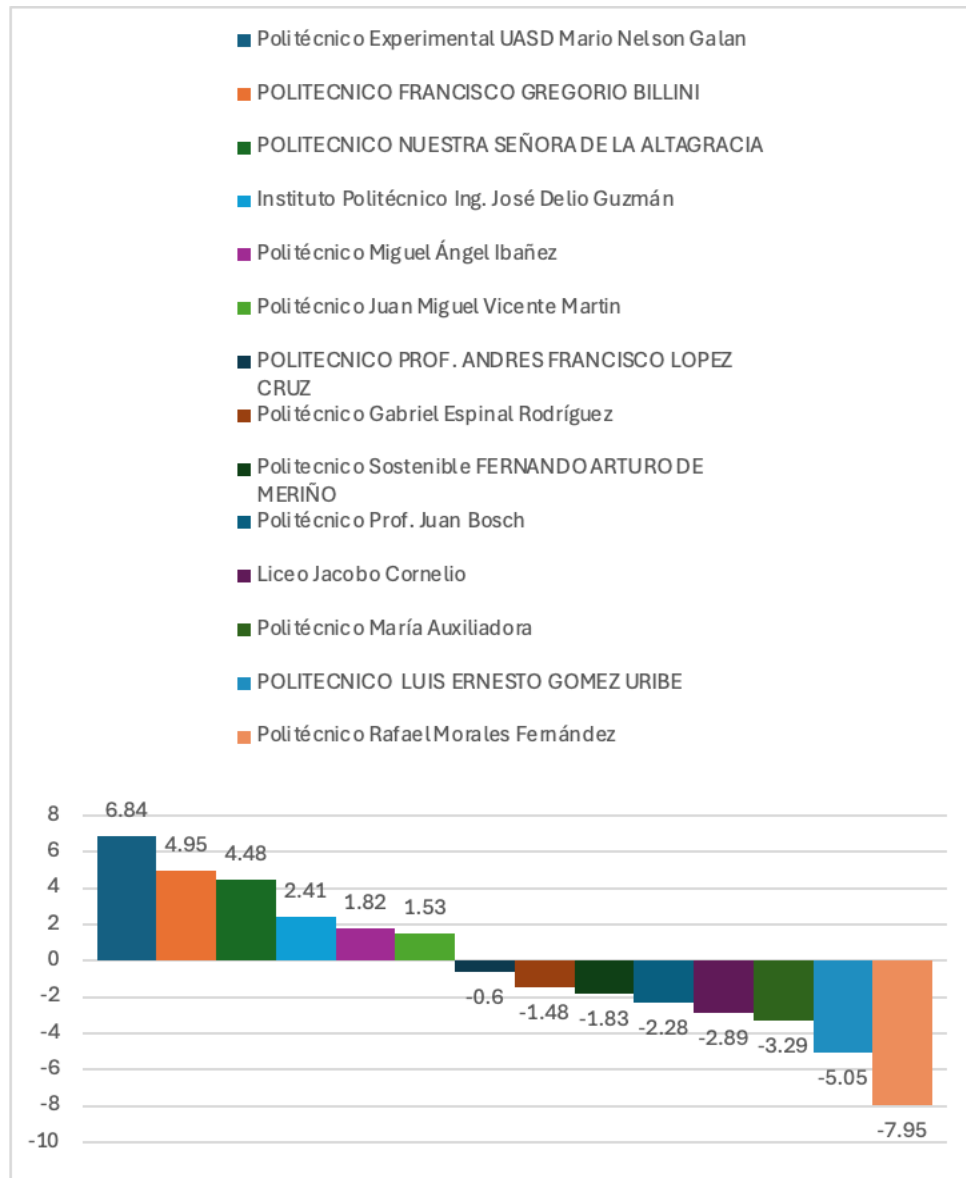


Tabla 5.3. Comparación con la regional educativa

Indicador	Valor
Politécnicos por encima de la regional	6
Politécnicos por debajo de la regional	8
Diferencia promedio	-0.09 puntos

Según la tabla 5.3, la diferencia promedio es prácticamente nula, lo que indica que, en conjunto, los politécnicos participantes presentan un desempeño equivalente al promedio regional.

Sin embargo, la distribución individual revela nuevamente una considerable variabilidad entre centros.

5.7 Comparación con el promedio nacional

La comparación con el promedio nacional presenta un comportamiento distinto.

Tabla 5.4. Comparación con el promedio nacional

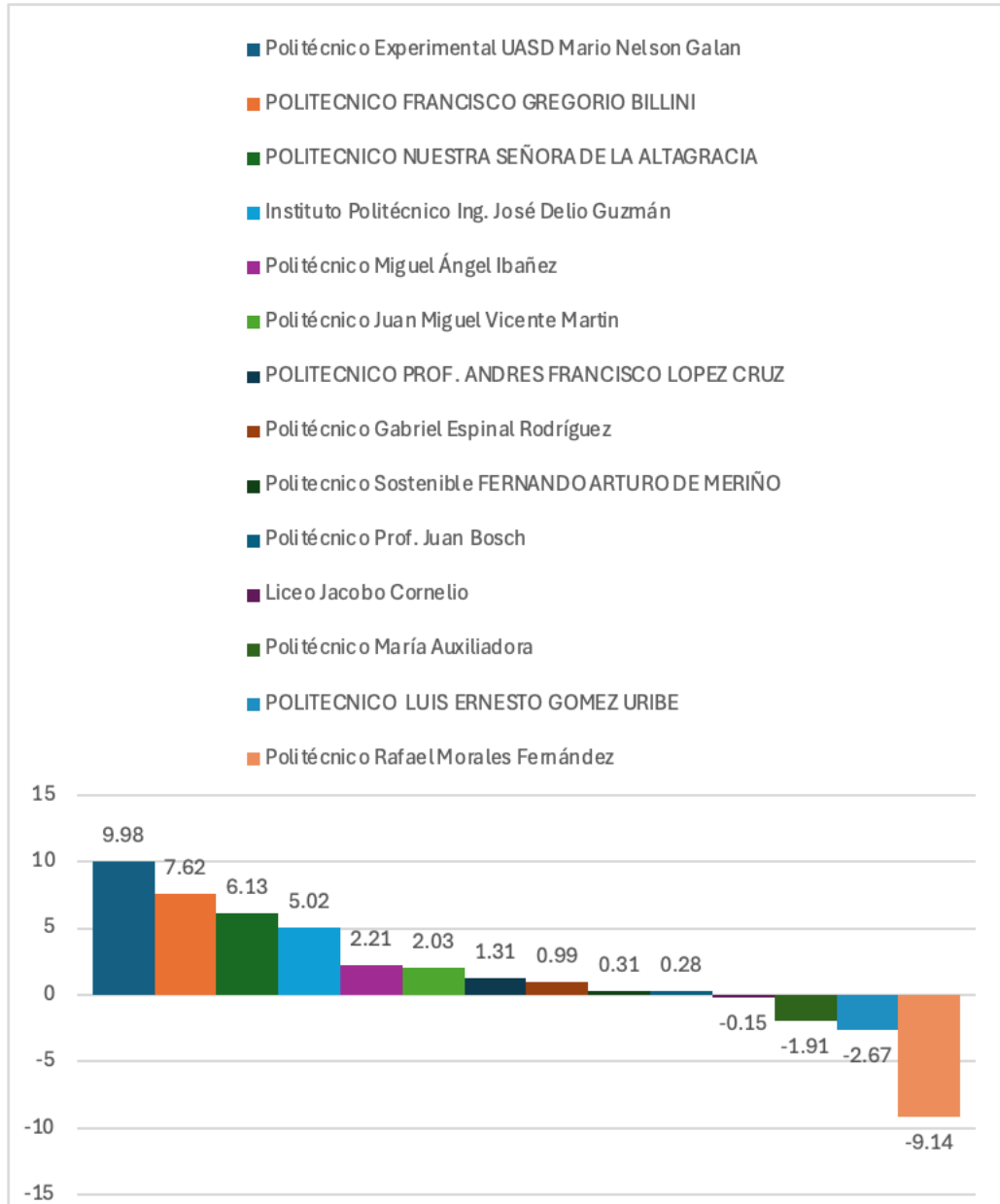
Indicador	Valor
Politécnicos por encima del promedio nacional	10
Politécnicos por debajo del promedio nacional	4
Diferencia promedio	+1.57 puntos

En este caso, el **71.4 %** de los politécnicos supera el promedio nacional, según se puede observar en la tabla 5.4 y la figura 5.4.

Esta constituye la evidencia descriptiva más favorable del estudio y sugiere que los centros participantes presentan, en conjunto, un desempeño superior al promedio del sistema educativo.

No obstante, la significación estadística de esta diferencia será analizada en el capítulo siguiente.

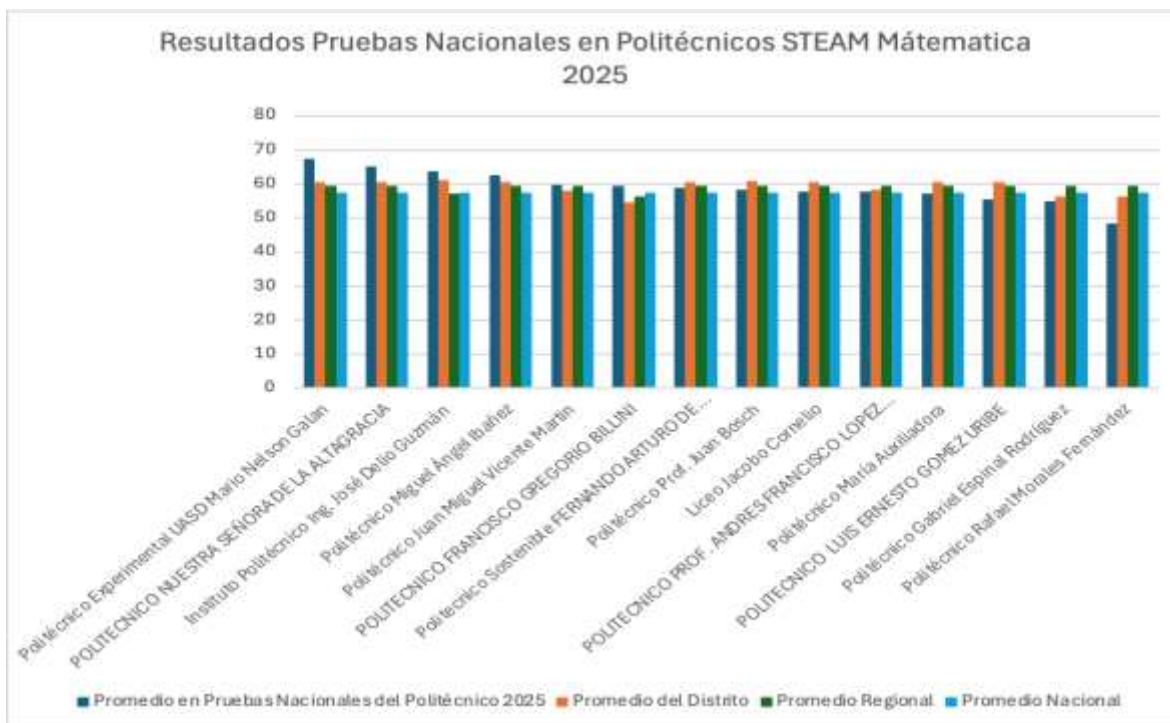
Figura 5.4. Brechas Politécnico – Promedio Nacional



5.8 Comparación integrada de los tres referentes

Con el propósito de sintetizar el comportamiento de los politécnicos, la Figura 5.5 presenta simultáneamente las diferencias respecto a los tres niveles de referencia.

Figura 5.5. Desempeño relativo de los politécnicos respecto a los tres niveles de comparación



La comparación integrada muestra tres patrones claramente diferenciados:

- Politécnicos que superan simultáneamente los tres referentes.
- Politécnicos que solo superan el promedio nacional.
- Politécnicos que permanecen por debajo de todos los niveles de referencia.

Esta clasificación constituye el punto de partida para el desarrollo del Índice de Aprovechamiento STEM presentado en el Capítulo 7.

5.9 Heterogeneidad del desempeño institucional

Uno de los hallazgos más importantes del análisis descriptivo es la elevada heterogeneidad observada entre los centros participantes.

Mientras algunos politécnicos presentan ventajas consistentes frente a los tres niveles de referencia, otros registran resultados inferiores incluso al promedio nacional.

Esta dispersión sugiere que la capacidad institucional para convertir la capacitación docente en mejoras del aprendizaje no fue uniforme entre los centros educativos.

Desde una perspectiva de gestión, este resultado resulta especialmente relevante, ya que indica la conveniencia de complementar el análisis de medias con herramientas que permitan identificar distintos niveles de aprovechamiento institucional.

5.10 Síntesis del análisis descriptivo

El análisis realizado permite destacar cinco hallazgos preliminares:

1. Los resultados de los politécnicos muestran una variabilidad considerable.
2. La mayoría de los centros no supera los promedios de sus respectivos distritos y regionales.
3. Diez de los catorce politécnicos obtienen resultados superiores al promedio nacional.
4. Las diferencias observadas varían significativamente entre centros educativos.
5. La heterogeneidad identificada justifica la incorporación de indicadores complementarios de desempeño relativo.

En consecuencia, el análisis descriptivo proporciona una primera aproximación al comportamiento de los politécnicos participantes y establece las bases para el análisis inferencial que se desarrolla en el capítulo siguiente.

6. ANÁLISIS INFERENCIAL DE LOS RESULTADOS

6.1 Introducción

El análisis descriptivo desarrollado en el capítulo anterior permitió identificar diferencias entre los resultados obtenidos por los politécnicos participantes y los promedios de sus respectivos distritos educativos, regionales y el promedio nacional. Sin embargo, las diferencias observadas en las medias no permiten determinar, por sí mismas, si dichas diferencias obedecen a un patrón sistemático o si pueden atribuirse a la variabilidad propia de los datos.

Con el propósito de fortalecer la evidencia empírica del estudio, se aplicaron procedimientos de estadística inferencial orientados a contrastar las hipótesis planteadas en el marco metodológico. En particular, se utilizaron tres herramientas complementarias:

- Prueba **t** para muestras pareadas.
- Prueba no paramétrica de **Wilcoxon** para rangos con signo.
- Estimación del **tamaño del efecto** mediante el estadístico de **Cohen (d)**.

La combinación de estos procedimientos permite evaluar no solo la existencia de diferencias estadísticas, sino también la magnitud práctica de las mismas.

6.2 Hipótesis de investigación

Para cada uno de los tres niveles de comparación se plantearon las siguientes hipótesis.

Hipótesis nula (H_0): No existen diferencias estadísticamente significativas entre el promedio obtenido por los politécnicos participantes y el promedio del nivel de referencia correspondiente.

Hipótesis alternativa (H_1): Los politécnicos participantes presentan resultados superiores al promedio del nivel de referencia correspondiente.

Las hipótesis fueron evaluadas para tres comparaciones:

- Politécnico vs. Distrito.
- Politécnico vs. Regional.
- Politécnico vs. Promedio Nacional.

6.3 Comparación con el promedio distrital

La primera comparación evalúa si los politécnicos participantes obtuvieron un desempeño superior al promedio de sus respectivos distritos educativos.

Tabla 6.1

Resultados de la prueba t para muestras pareadas

Estadístico	Valor
Diferencia promedio	-0.24 puntos
Estadístico t	-0.22
Valor p	0.832

El valor de p es considerablemente superior al nivel de significación convencional ($\alpha = 0.05$). En consecuencia, **no se rechaza la hipótesis nula**, lo que indica que no existe evidencia estadística suficiente para afirmar que los politécnicos obtuvieron un desempeño superior al promedio de sus respectivos distritos.

Con el propósito de verificar la robustez de este resultado, se aplicó la prueba no paramétrica de Wilcoxon.

Tabla 6.2

Resultados de la prueba de Wilcoxon

Estadístico	Valor
Valor p	0.808

La prueba de Wilcoxon conduce a la misma conclusión, reforzando la evidencia de que las diferencias observadas respecto al promedio distrital no son estadísticamente significativas.

6.4 Comparación con el promedio regional

La segunda comparación analiza el desempeño de los politécnicos respecto al promedio de sus correspondientes regionales educativas.

Tabla 6.3
Resultados de la prueba t

Estadístico	Valor
Diferencia promedio	-0.09 puntos
Estadístico t	-0.07
Valor p	0.946

El análisis muestra una diferencia promedio prácticamente nula. El elevado valor de p confirma que el desempeño de los politécnicos participantes es estadísticamente equivalente al promedio regional.

Tabla 6.4
Resultados de la prueba de Wilcoxon

Estadístico	Valor
Valor p	1

Nuevamente, la prueba no paramétrica confirma la ausencia de diferencias estadísticamente significativas.

6.5 Comparación con el promedio nacional

La tercera comparación presenta un comportamiento diferente.

Tabla 6.5
Resultados de la prueba t

Estadístico	Valor
Diferencia promedio	+1.57 puntos
Estadístico t	1.24
Valor p	0.237

En términos descriptivos, los politécnicos participantes superan el promedio nacional en **1.57 puntos**. Sin embargo, el valor p continúa siendo superior al nivel de significación del cinco por ciento, por lo que **no es posible rechazar la hipótesis nula**. La prueba de Wilcoxon produce resultados consistentes:

Tabla 6.6
Resultados de la prueba de Wilcoxon

Estadístico	Valor
Valor p	0.153

Aunque el desempeño de los politécnicos es superior al promedio nacional, la evidencia estadística disponible no permite afirmar que esta diferencia sea significativa con el tamaño muestral utilizado.

6.6 Magnitud del efecto

La significación estadística depende, entre otros factores, del tamaño de la muestra. Por esta razón, resulta conveniente complementar el análisis mediante una estimación del tamaño del efecto.

El estadístico de Cohen (d) permite valorar la importancia práctica de las diferencias observadas.

Tabla 6.7
Tamaño del efecto

Comparación	Cohen (d)	Interpretación
Politécnico – Distrito	-0.06	Efecto prácticamente nulo
Politécnico – Regional	-0.02	Efecto prácticamente nulo
Politécnico – Nacional	0.33	Efecto pequeño–moderado

Los resultados muestran que únicamente la comparación con el promedio nacional presenta un tamaño del efecto de magnitud apreciable, aunque insuficiente para alcanzar significación estadística con la muestra disponible.

6.7 Discusión de los resultados inferenciales

El análisis inferencial confirma que las diferencias observadas entre los politécnicos participantes y los tres niveles de referencia no alcanzan significación estadística al nivel convencional del 95 %.

No obstante, esta conclusión debe interpretarse con cautela por varias razones.

En primer lugar, la muestra está conformada únicamente por catorce politécnicos, lo que reduce la potencia estadística de las pruebas aplicadas. En muestras pequeñas, diferencias de magnitud moderada pueden no alcanzar niveles convencionales de significación.

En segundo lugar, el análisis descriptivo mostró una elevada heterogeneidad entre los centros participantes. Algunos politécnicos presentan desempeños consistentemente superiores a todos los referentes, mientras otros obtienen resultados inferiores. Esta variabilidad reduce la capacidad de las pruebas basadas en promedios para detectar patrones diferenciados.

Finalmente, la diferencia positiva observada respecto al promedio nacional y el tamaño del efecto estimado mediante el estadístico de Cohen sugieren la existencia de una tendencia favorable que merece ser analizada mediante herramientas adicionales.

6.8 Limitaciones del análisis inferencial

Las pruebas estadísticas aplicadas evalúan el comportamiento promedio del conjunto de politécnicos participantes. Sin embargo, este enfoque presenta dos limitaciones relevantes.

La primera es que resume la información en una única media, ocultando diferencias importantes entre los centros educativos.

La segunda es que no incorpora información sobre la magnitud relativa de las brechas respecto a los distintos niveles de referencia.

Estas limitaciones justifican el desarrollo de un indicador complementario capaz de sintetizar el desempeño relativo de cada politécnico de manera individual.

6.9 Transición hacia el Índice de Aprovechamiento STEM (IAS)

Los resultados obtenidos muestran que la comparación de medias constituye una aproximación útil, pero insuficiente para comprender plenamente el comportamiento de los politécnicos participantes.

La heterogeneidad observada entre los centros educativos sugiere que algunos lograron transformar con mayor efectividad la capacitación docente en mejores resultados de aprendizaje, mientras otros no alcanzaron el mismo nivel de desempeño.

Con el propósito de capturar estas diferencias individuales, el capítulo siguiente presenta el **Índice de Aprovechamiento STEM (IAS)**, un indicador sintético diseñado para medir el desempeño relativo de cada politécnico respecto de sus referentes distritales, regionales y nacionales, y para apoyar la toma de decisiones en materia de seguimiento y fortalecimiento institucional.

7. ÍNDICE DE APROVECHAMIENTO STEM (IAS)

7.1 Introducción

Los capítulos anteriores demostraron que la comparación de medias constituye una herramienta útil para describir el comportamiento general de los politécnicos participantes.

Sin embargo, dicho enfoque presenta una limitación importante: resume el desempeño de todos los centros mediante un único promedio, ocultando diferencias significativas entre instituciones.

El análisis descriptivo mostró que algunos politécnicos superan simultáneamente los promedios de su distrito, su regional y el promedio nacional, mientras otros permanecen por debajo de dichos referentes. Esta heterogeneidad sugiere que la capacidad institucional para transformar la capacitación docente en mejores resultados de aprendizaje no fue uniforme.

Con el propósito de capturar estas diferencias individuales, el presente estudio propone el **Índice de Aprovechamiento STEM (IAS)**, un indicador sintético diseñado para medir el desempeño relativo de cada politécnico respecto de tres niveles de referencia y apoyar los procesos de seguimiento, evaluación y mejora continua de los programas de formación docente.

El IAS constituye una herramienta de análisis institucional y **no una medida de impacto causal**. Su finalidad consiste en identificar patrones de desempeño relativo y orientar la toma de decisiones del INAFOCAM.

7.2 Fundamentación conceptual

El IAS se sustenta en el concepto de **aprovechamiento institucional**, entendido como la capacidad de un centro educativo para transformar una intervención de desarrollo profesional docente en mejoras observables del aprendizaje de sus estudiantes.

Desde esta perspectiva, el resultado obtenido por un politécnico depende de dos componentes:

1. La calidad y pertinencia de la capacitación recibida.
2. La capacidad institucional para incorporar esa capacitación en la práctica pedagógica cotidiana.

En consecuencia, el IAS no mide directamente la calidad del programa de formación, sino el grado en que cada institución logró traducir dicha formación en un desempeño relativo superior.

7.3 Principios metodológicos

La construcción del índice responde a cuatro principios.

Comparabilidad

Todos los politécnicos deben evaluarse utilizando los mismos criterios.

Proximidad

Los referentes territoriales más cercanos poseen mayor capacidad para reflejar el contexto institucional del centro.

Objetividad

El índice se construye exclusivamente a partir de indicadores cuantitativos derivados de los resultados oficiales de las Pruebas Nacionales.

Interpretabilidad

Los resultados deben expresarse mediante una escala sencilla que facilite su comprensión por parte de los responsables de la toma de decisiones.

7.4 Construcción del Índice Ponderado de Desempeño Relativo (IPDR)

La primera etapa consiste en calcular tres brechas de desempeño.

Brecha distrital

$$D = PP - PD$$

Brecha regional

$$R = PP - PR$$

Brecha nacional

$$N = PP - PN$$

donde:

- **PP** = Promedio del Politécnico.
- **PD** = Promedio del Distrito.
- **PR** = Promedio Regional.
- **PN** = Promedio Nacional.

Cada brecha expresa la diferencia entre el desempeño del politécnico y el referente correspondiente.

Índice Ponderado de Desempeño Relativo (IPDR)

Las tres brechas se integran mediante la siguiente expresión:

$$IPDR = (0.50 \times D) + (0.30 \times R) + (0.20 \times N)$$

Las ponderaciones responden al principio de proximidad institucional.

Referente Peso Justificación

Distrito 50 % Comparador más próximo y exigente.

Regional 30 % Contexto territorial ampliado.

Referente Peso Justificación

Nacional 20 % Referente general del sistema educativo.

7.5 Normalización del índice

Para facilitar la comparación entre centros, el IPDR se transforma a una escala de 0 a 100 mediante una normalización lineal.

$$IAS = 100 \times [(IPDR - IPDR \text{ mínimo}) / (IPDR \text{ máximo} - IPDR \text{ mínimo})]$$

Esta transformación garantiza que:

- el politécnico con mayor desempeño relativo obtenga un IAS de 100;
- el politécnico con menor desempeño relativo obtenga un IAS de 0;
- los demás centros se distribuyan proporcionalmente entre ambos extremos.

7.6 Clasificación del IAS

Con fines analíticos se propone la siguiente clasificación.

IAS	Nivel de aprovechamiento
80 – 100	Muy alto
60 – 79.9	Alto
40 – 59.9	Medio
20 – 39.9	Bajo
0 – 19.9	Muy bajo

Esta clasificación permite identificar distintos niveles de desempeño institucional y orientar estrategias diferenciadas de acompañamiento.

7.7 Resultados del IAS

Tabla 7.1. Ranking del Índice de Aprovechamiento STEM

Politécnico	IAS	Nivel
Politécnico Experimental UASD Mario Nelson Galán	100	Muy alto
Politécnico Nuestra Señora de la Altagracia	87	Muy alto
Instituto Politécnico Ing. José Delio Guzmán	82	Muy alto
Politécnico Francisco Gregorio Billini	73	Alto
Politécnico Miguel Ángel Ibañez	72	Alto
Politécnico Juan Miguel Vicente Martín	61	Alto
Politécnico Sostenible Fernando Arturo de Meriño	51	Medio
Politécnico Prof. Andrés Francisco López Cruz	49	Medio
Politécnico Prof. Juan Bosch	49	Medio
Liceo Jacobo Cornelio	45	Medio
Politécnico María Auxiliadora	43	Medio
Politécnico Gabriel Espinal Rodríguez	37	Bajo
Politécnico Luis Ernesto Gómez Uribe	33	Bajo
Politécnico Rafael Morales Fernández	0	Muy bajo

7.8 Interpretación de los resultados

La distribución del IAS pone de manifiesto que la respuesta institucional al programa de capacitación fue heterogénea. Seis politécnicos se ubican en las categorías de **aprovechamiento alto o muy alto**, lo que sugiere una mayor capacidad para traducir la capacitación docente en mejores resultados de aprendizaje. Cinco centros presentan un **aprovechamiento medio**, mientras que tres se sitúan en las categorías de **aprovechamiento bajo o muy bajo**, evidenciando oportunidades de fortalecimiento institucional.

Estos resultados no deben interpretarse como una clasificación de la calidad de los centros educativos, sino como una herramienta para identificar patrones de desempeño relativo y orientar acciones de acompañamiento diferenciadas.

7.9 Aplicaciones del IAS

El Índice de Aprovechamiento STEM puede utilizarse para:

- monitorear futuras cohortes del programa;
- comparar el desempeño entre regiones y distritos;
- identificar centros con buenas prácticas pedagógicas;
- focalizar el acompañamiento técnico del INAFOCAM;
- seleccionar casos para estudios cualitativos;
- alimentar tableros de control institucional;
- complementar evaluaciones de impacto.

7.10 Limitaciones del índice

El IAS presenta las siguientes limitaciones:

1. No estima el impacto causal del programa.
2. Depende de la calidad de la información disponible.
3. No incorpora variables contextuales como nivel socioeconómico, infraestructura o experiencia docente.
4. Las ponderaciones propuestas responden a un criterio técnico que deberá ser validado empíricamente en futuras aplicaciones.

7.11 Validación y líneas futuras

Para consolidar el IAS como un instrumento institucional del INAFOCAM se recomienda desarrollar un proceso de validación en tres etapas:

1. **Validación de contenido**, mediante consulta a expertos en evaluación educativa, formación docente y estadística aplicada.
2. **Validación empírica**, aplicando el índice a nuevas cohortes y programas de formación en otras áreas curriculares.
3. **Análisis de sensibilidad**, comparando diferentes esquemas de ponderación (por ejemplo, 40–40–20, 50–25–25 o pesos derivados de técnicas multivariadas) para evaluar la estabilidad del indicador.

Asimismo, futuras investigaciones podrían incorporar variables contextuales y explorar modelos más avanzados —como análisis factorial o modelos multinivel— para refinar el índice sin perder su sencillez operativa.

8. SISTEMA INTEGRADO PARA LA EVALUACIÓN DE PROGRAMAS DE FORMACIÓN DOCENTE (SIEPFD)

8.1 Introducción

Los capítulos precedentes demostraron que la evaluación de programas de formación docente no puede limitarse al análisis de los resultados obtenidos por los estudiantes en pruebas estandarizadas. Si bien dichos resultados constituyen un indicador fundamental de desempeño, representan únicamente el último eslabón de una cadena mucho más compleja que comienza con el diseño de la capacitación y culmina con la mejora de los aprendizajes.

La experiencia internacional muestra que muchos programas de desarrollo profesional docente son evaluados únicamente en términos de cobertura, número de participantes o niveles de satisfacción, dejando sin respuesta preguntas esenciales relacionadas con la transferencia de la capacitación al aula, la capacidad institucional para incorporar nuevas prácticas pedagógicas y el impacto efectivo sobre el aprendizaje de los estudiantes.

Con el propósito de superar estas limitaciones, el presente informe propone un **Sistema Integrado para la Evaluación de Programas de Formación Docente (SIEPFD)**, concebido como una arquitectura metodológica que articula diferentes niveles de evaluación y permite dar seguimiento al ciclo completo de una intervención de formación docente.

El SIEPFD integra el **Índice de Aprovechamiento STEM (IAS)** desarrollado en el capítulo anterior dentro de un sistema más amplio de monitoreo, evaluación y aprendizaje institucional, proporcionando al INAFOCAM una herramienta para la gestión basada en evidencia.

8.2 Principios del sistema

El SIEPFD se fundamenta en seis principios.

Integralidad

La evaluación debe abarcar todas las etapas del programa, desde la planificación hasta los resultados de aprendizaje.

Continuidad

La evaluación debe realizarse antes, durante y después de la implementación de la capacitación.

Multinivel

Debe considerar indicadores del programa, de los docentes, de los centros educativos y de los estudiantes.

Comparabilidad

Los resultados deben expresarse mediante indicadores que permitan comparaciones entre cohortes, regiones, instituciones y programas.

Aprendizaje institucional

La información generada debe utilizarse para fortalecer el diseño de futuras intervenciones y mejorar la calidad de la oferta formativa.

Toma de decisiones basada en evidencia

Los resultados del sistema deben apoyar la planificación, asignación de recursos y priorización de acciones de acompañamiento por parte del INAFOCAM.

Eficiencia

La evaluación debe valorar no solo si los programas producen resultados, sino también la relación entre los recursos utilizados y los logros alcanzados. Un programa más eficiente es aquel que obtiene mayores mejoras en el aprendizaje utilizando una menor cantidad de recursos por docente, por centro educativo o por estudiante beneficiado.

8.3 Arquitectura del sistema

El SIEPFD se estructura en **cuatro niveles secuenciales** de evaluación.

Nivel I. Evaluación de la implementación

Este nivel responde a la pregunta:

¿El programa se ejecutó conforme a lo planificado?

Indicadores propuestos:

- porcentaje de ejecución del programa;
- cobertura de docentes participantes;
- tasa de finalización;
- asistencia;
- cumplimiento del cronograma;
- satisfacción de los participantes;
- calidad percibida de la capacitación.

Producto esperado: Informe de implementación.

Nivel II. Evaluación de la transferencia pedagógica

Pregunta central:

¿Los docentes incorporaron al aula las metodologías desarrolladas durante la capacitación?

Este nivel constituye el puente entre la formación recibida y los resultados educativos.

Indicadores sugeridos:

- observación estructurada de clases;
- aplicación de metodologías STEM;
- utilización de recursos didácticos;
- planificación por competencias;
- evaluación auténtica;
- trabajo colaborativo;
- innovación pedagógica.

Como parte del sistema se propone desarrollar un nuevo indicador:

Índice de Transferencia Pedagógica (ITP)

Este índice permitirá medir el grado de incorporación de las competencias adquiridas durante la capacitación.

Nivel III. Evaluación del aprovechamiento institucional

Este nivel responde a la pregunta:

¿Los centros educativos lograron transformar la capacitación docente en mejores resultados académicos?

La herramienta central de este nivel es el **Índice de Aprovechamiento STEM (IAS)**.

El IAS sintetiza el desempeño relativo del centro educativo respecto a:

- distrito;
- regional;
- promedio nacional.

Este nivel permite comparar instituciones, identificar buenas prácticas y focalizar acciones de fortalecimiento institucional.

Nivel IV. Evaluación del impacto educativo y de la eficiencia

Preguntas centrales:

¿Qué cambios pueden atribuirse al programa de formación docente?

Este nivel requiere metodologías de evaluación de impacto más robustas.

Entre ellas:

- Diferencias en Diferencias.
- Modelos multinivel.
- Modelos longitudinales.
- Ensayos controlados aleatorios, cuando sean viables.

Este componente permitirá estimar el efecto neto de la capacitación sobre los aprendizajes de los estudiantes.

¿Qué programas producen mayores resultados por cada peso invertido?

Entre los indicadores posibles se encuentran:

- costo por docente capacitado
- costo por centro beneficiado
- costo por estudiante beneficiado
- costo por punto adicional obtenido en las Pruebas Nacionales
- costo por unidad de aumento del IAS
- costo por unidad del futuro ITP

Entre los métodos a considerar se encuentran los siguientes:

- Análisis costo–eficiencia
- Análisis costo–efectividad
- Razón incremental de costo–efectividad
- Benchmarking entre instituciones formadoras

El análisis costo/eficiencia y costo/efectividad tienen fuertes vinculaciones con la evaluación de impactos de los programas de formación y capacitación docente. Van den Brande y Zuccollo, J. (2021) presentan el que probablemente sea la referencia más importante sobre análisis beneficio–costo de la capacitación docente. Estima que ofrecer 35 horas anuales de desarrollo profesional de alta calidad a todos los docentes de Inglaterra tendría un costo cercano a £4 mil millones en diez años, pero generaría beneficios sociales netos estimados en aproximadamente £61 mil millones, principalmente por el aumento del aprendizaje estudiantil y sus efectos sobre los ingresos futuros. El estudio demuestra que la formación docente de calidad puede representar una de las inversiones públicas con mayor rentabilidad social.

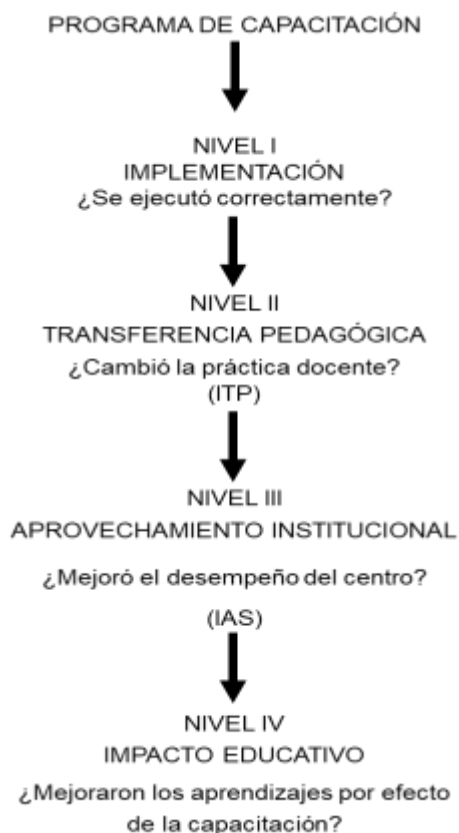
Popova (et al.,2018). Presentan un importante estudio en el que se sintetiza la evidencia internacional sobre programas de desarrollo profesional docente y concluye que la efectividad depende menos del volumen de recursos invertidos que de la calidad del diseño pedagógico. Además de analizar los impactos, ofrece criterios útiles para valorar la eficiencia de la inversión, al mostrar qué características generan mejores resultados por unidad de recursos invertidos.

La incorporación del análisis de costo–eficiencia y costo–beneficio responde a la creciente necesidad de que las evaluaciones de programas públicos no solo determinen si las intervenciones producen resultados, sino también si dichos resultados justifican la inversión realizada. En el ámbito de la formación docente, este enfoque permite identificar aquellas

modalidades de capacitación que generan mayores mejoras en los aprendizajes por cada unidad monetaria invertida, fortaleciendo la asignación eficiente de los recursos públicos y la sostenibilidad de las políticas educativas.

8.4 Arquitectura funcional del SIEPFD

Figura 8.1. Arquitectura del Sistema Integrado para la Evaluación de Programas de Formación Docente



8.5 Flujo de información

El sistema genera información para distintos niveles de decisión.

Nivel Producto		Usuario principal
I	Informe de implementación	Coordinación del programa
II	Índice de Transferencia Pedagógica (ITP)	Equipos pedagógicos
III	Índice de Aprovechamiento STEM (IAS)	INAFOCAM y centros educativos
IV	Evaluación de impacto	Alta dirección y formuladores de política

Esta estructura permite que cada actor disponga de información pertinente para la toma de decisiones.

8.6 Aplicaciones institucionales

El SIEPFD puede utilizarse para:

- monitorear programas de formación docente en Matemática, Ciencias, Lengua Española y Tecnología;
- comparar resultados entre instituciones formadoras;
- evaluar cohortes sucesivas de un mismo programa;
- identificar centros educativos con alto potencial para convertirse en referentes de buenas prácticas;
- orientar procesos de acompañamiento pedagógico;
- apoyar la asignación de recursos y la planificación institucional.

Asimismo, el sistema facilita la construcción de tableros de control y reportes ejecutivos para la alta dirección del INAFOCAM.

8.7 Hoja de ruta para la implementación

Se propone una implementación gradual en tres fases.

Fase I. Diseño y validación

- Validación técnica del IAS.
- Diseño del Índice de Transferencia Pedagógica (ITP).
- Definición de protocolos de recolección de datos.

Fase II. Piloto institucional

- Aplicación del sistema en nuevos programas de formación docente.
- Ajuste de indicadores y procedimientos.
- Capacitación de los equipos técnicos del INAFOCAM.

Fase III. Institucionalización

- Incorporación del SIEPFD al sistema de monitoreo y evaluación del INAFOCAM.
- Elaboración de informes periódicos.
- Integración con los procesos de planificación y mejora institucional.

8.8 Contribución metodológica

La principal contribución del presente estudio consiste en proponer un sistema que integra la evaluación de la implementación, la transferencia pedagógica, el aprovechamiento institucional y el impacto educativo en una sola arquitectura metodológica. El **IAS** deja de ser un indicador aislado para convertirse en uno de los componentes del sistema, complementado por el futuro **Índice de Transferencia Pedagógica (ITP)** y por

metodologías de evaluación de impacto aplicables cuando se disponga de diseños adecuados.

Este enfoque permite al INAFOCAM pasar de una evaluación centrada en actividades y productos a una evaluación orientada a resultados, aprendizaje institucional y mejora continua, fortaleciendo la toma de decisiones basada en evidencia.

9. DISCUSIÓN DE LOS RESULTADOS

9.1 Introducción

El propósito de este capítulo es interpretar los resultados obtenidos a la luz del marco conceptual desarrollado y discutir sus implicaciones para la evaluación de programas de formación docente. Mientras los capítulos anteriores presentaron la evidencia empírica y los desarrollos metodológicos del estudio, esta sección busca responder tres preguntas fundamentales:

1. ¿Qué significan los resultados observados desde la perspectiva de la evaluación educativa?
2. ¿Qué aportes ofrece el Índice de Aprovechamiento STEM (IAS) para comprender las diferencias entre los centros educativos?
3. ¿Qué implicaciones tienen estos hallazgos para el diseño y evaluación de futuras políticas de desarrollo profesional docente del INAFOCAM?

La discusión se organiza en torno a los principales hallazgos del estudio y su relación con la literatura especializada y con los desafíos de la gestión educativa.

9.2 El desempeño observado no constituye evidencia de impacto

El primer resultado relevante del estudio es que las diferencias entre los politécnicos participantes y sus referentes educativos no alcanzan significación estadística al nivel convencional del 95 %. Este hallazgo obliga a interpretar los resultados con prudencia.

Desde una perspectiva metodológica, la ausencia de significación estadística **no implica necesariamente que el programa haya sido ineficaz**. La capacidad de una prueba estadística para detectar diferencias depende, entre otros factores, del tamaño de la muestra, de la variabilidad de los datos y de la magnitud del efecto. En este estudio, el reducido número de politécnicos participantes limita la potencia estadística y aumenta la probabilidad de que diferencias relevantes desde el punto de vista educativo no alcancen significación estadística.

Asimismo, el diseño de la investigación corresponde a una evaluación de resultados y no a una evaluación de impacto. La ausencia de una línea de base y de un grupo de comparación impide estimar el efecto atribuible exclusivamente al programa de capacitación. En consecuencia, los hallazgos deben entenderse como evidencia sobre el desempeño relativo de los centros participantes y no como una demostración del impacto causal de la intervención.

9.3 La heterogeneidad institucional como principal hallazgo

Más allá de la comparación de medias, el estudio pone de manifiesto un resultado de mayor interés para la gestión educativa: la existencia de una marcada heterogeneidad entre los politécnicos participantes.

Algunos centros superan simultáneamente los promedios distritales, regionales y nacionales, mientras otros presentan desempeños inferiores incluso al promedio nacional. Esta variabilidad sugiere que la respuesta institucional a una misma intervención no fue uniforme.

Desde la perspectiva de la teoría del cambio presentada en el Capítulo 3, este hallazgo resulta consistente con la idea de que la capacitación constituye una condición necesaria, pero no suficiente, para mejorar los resultados de aprendizaje. Entre la formación recibida y el desempeño de los estudiantes intervienen factores como el liderazgo pedagógico, la organización escolar, la colaboración docente, la disponibilidad de recursos y la capacidad de gestión del centro educativo.

En consecuencia, la heterogeneidad observada constituye una evidencia indirecta de que el contexto institucional desempeña un papel determinante en el aprovechamiento de las oportunidades de desarrollo profesional.

9.4 El aporte del Índice de Aprovechamiento STEM (IAS)

El desarrollo del IAS responde precisamente a la necesidad de capturar esta heterogeneidad.

Mientras las pruebas estadísticas sintetizan el comportamiento de todos los politécnicos mediante una media, el IAS permite evaluar el desempeño relativo de cada centro considerando simultáneamente tres referentes territoriales. Este enfoque proporciona una visión más rica de las diferencias entre instituciones y facilita la identificación de patrones de desempeño.

No obstante, es importante precisar el alcance del indicador. El IAS **no mide la calidad absoluta de un centro educativo ni el impacto causal del programa**, sino su desempeño relativo dentro del contexto analizado. Su utilidad radica en ofrecer un criterio objetivo para priorizar acciones de acompañamiento, identificar experiencias exitosas y orientar la mejora institucional.

9.5 Implicaciones para la gestión del INAFOCAM

Los resultados obtenidos tienen implicaciones relevantes para la planificación y evaluación de los programas de formación docente.

En primer lugar, evidencian la conveniencia de complementar los indicadores tradicionales de cobertura y participación con indicadores orientados a resultados y desempeño institucional.

En segundo lugar, sugieren que las futuras intervenciones deberían incorporar mecanismos sistemáticos de seguimiento a la transferencia pedagógica, de manera que sea posible

identificar oportunamente las dificultades que enfrentan los docentes para aplicar en el aula las metodologías desarrolladas durante la capacitación.

En tercer lugar, ponen de manifiesto la importancia de diferenciar las estrategias de acompañamiento según el nivel de aprovechamiento institucional de cada centro educativo. Los resultados indican que una intervención homogénea puede producir efectos distintos dependiendo de las capacidades organizacionales de cada institución.

9.6 Hacia una evaluación basada en evidencia

Uno de los aportes más importantes del presente estudio consiste en proponer una transición desde una evaluación centrada en actividades hacia una evaluación basada en evidencia.

Este cambio implica que el éxito de un programa de formación docente no debe medirse únicamente por el número de participantes, las horas impartidas o el grado de satisfacción de los docentes, sino también por su capacidad para producir cambios observables en las prácticas pedagógicas y, posteriormente, en los resultados de aprendizaje.

El Sistema Integrado para la Evaluación de Programas de Formación Docente (SIEPFD) responde a esta necesidad al articular indicadores de implementación, transferencia pedagógica, aprovechamiento institucional e impacto educativo dentro de un mismo marco metodológico.

9.7 Aportes metodológicos del estudio

El presente informe realiza tres contribuciones principales al campo de la evaluación de programas de formación docente.

La primera consiste en incorporar el concepto de **aprovechamiento institucional** como una dimensión específica del análisis, reconociendo que la efectividad de una intervención depende también de las capacidades organizacionales de los centros educativos.

La segunda es el desarrollo del **Índice de Aprovechamiento STEM (IAS)**, un indicador sintético diseñado para medir el desempeño relativo de los centros participantes y facilitar la identificación de buenas prácticas y necesidades de fortalecimiento.

La tercera corresponde al diseño del **Sistema Integrado para la Evaluación de Programas de Formación Docente (SIEPFD)**, que propone una arquitectura institucional para organizar los procesos de monitoreo y evaluación del INAFOCAM.

En conjunto, estos aportes amplían el alcance del estudio y ofrecen herramientas que pueden ser adaptadas a otros programas de formación docente.

9.8 Limitaciones y agenda de investigación

Como toda investigación, el presente estudio presenta limitaciones que deben ser consideradas al interpretar sus resultados.

La principal limitación radica en la ausencia de información previa a la intervención y de un grupo de comparación equivalente. Esto restringe la posibilidad de establecer relaciones causales entre la capacitación y los resultados observados.

Asimismo, el tamaño reducido de la muestra limita la potencia de las pruebas estadísticas y aconseja interpretar los resultados con prudencia.

Finalmente, el estudio no incorpora variables contextuales relacionadas con las características de los estudiantes, los docentes o los centros educativos, que podrían contribuir a explicar parte de la heterogeneidad observada.

Estas limitaciones abren una agenda de investigación orientada a fortalecer futuras evaluaciones mediante diseños longitudinales, metodologías cuasi experimentales y la incorporación de indicadores de contexto que permitan comprender mejor los mecanismos a través de los cuales la formación docente influye sobre los aprendizajes.

9.9 Síntesis del capítulo

La discusión desarrollada permite extraer cinco conclusiones principales:

1. Los resultados obtenidos constituyen evidencia sobre el desempeño relativo de los politécnicos participantes, pero no permiten establecer el impacto causal del programa.
2. La heterogeneidad observada entre los centros educativos representa el hallazgo empírico más relevante del estudio.
3. El IAS complementa el análisis estadístico tradicional al proporcionar una medida del desempeño relativo de cada institución.
4. El SIEPFD ofrece un marco metodológico para fortalecer la evaluación de los programas de formación docente del INAFOCAM.
5. La incorporación de evaluaciones basadas en evidencia permitirá mejorar progresivamente el diseño, seguimiento y efectividad de las políticas de desarrollo profesional docente.

10. CONCLUSIONES Y RECOMENDACIONES

10.1 Conclusiones

El presente estudio tuvo como propósito evaluar el desempeño de los catorce politécnicos participantes en el Programa de Capacitación Docente STEM con énfasis en Matemática, utilizando como referencia los resultados obtenidos por sus estudiantes en las Pruebas Nacionales de Matemática 2025.

A partir del análisis realizado pueden formularse las siguientes conclusiones.

Primera conclusión

Los politécnicos participantes presentan, en conjunto, un desempeño superior al promedio nacional, aunque no muestran diferencias estadísticamente significativas respecto a los promedios distritales y regionales.

El análisis inferencial indica que las diferencias observadas no alcanzan significación estadística al nivel convencional del cinco por ciento. Sin embargo, este resultado debe interpretarse considerando el reducido tamaño de la muestra y las limitaciones propias del diseño metodológico utilizado.

Segunda conclusión

La principal evidencia empírica del estudio es la marcada heterogeneidad existente entre los politécnicos participantes.

Mientras algunos centros superan simultáneamente los tres niveles de referencia, otros permanecen por debajo de ellos, lo que indica que la respuesta institucional a la capacitación docente fue diferenciada.

Esta heterogeneidad constituye un hallazgo relevante para la gestión educativa, ya que pone de manifiesto que intervenciones similares pueden producir resultados distintos según las capacidades organizacionales de cada institución.

Tercera conclusión

La comparación de promedios resulta insuficiente para comprender plenamente el desempeño de los centros educativos.

El análisis estadístico tradicional permite establecer el comportamiento promedio del conjunto de instituciones, pero no identifica las diferencias individuales en la capacidad de aprovechamiento de la intervención.

En consecuencia, resulta conveniente incorporar indicadores complementarios que permitan analizar el desempeño relativo de cada centro educativo.

Cuarta conclusión

El Índice de Aprovechamiento STEM (IAS) constituye una herramienta útil para complementar la evaluación tradicional de programas de formación docente.

El IAS sintetiza el desempeño relativo de cada politécnico respecto a sus referentes distritales, regionales y nacionales, permitiendo identificar distintos niveles de aprovechamiento institucional y facilitando la focalización de acciones de acompañamiento.

No obstante, el indicador no debe interpretarse como una medida de impacto causal, sino como un instrumento de análisis comparativo y apoyo a la gestión institucional.

Quinta conclusión

El Sistema Integrado para la Evaluación de Programas de Formación Docente (SIEPFD) proporciona una arquitectura metodológica que fortalece la capacidad institucional del INAFOCAM para evaluar sus programas de desarrollo profesional docente.

La integración de indicadores de implementación, transferencia pedagógica, aprovechamiento institucional e impacto educativo permite avanzar hacia un enfoque de evaluación basado en evidencia y orientado a la mejora continua.

10.2 Recomendaciones

Recomendaciones para el INAFOCAM

1. Institucionalizar un sistema integral de evaluación. Incorporar progresivamente el SIEPFD como marco de referencia para el monitoreo y evaluación de todos los programas de formación docente, adaptándolo a las distintas modalidades y áreas curriculares.

2. Incorporar indicadores de transferencia pedagógica. Diseñar e implementar el **Índice de Transferencia Pedagógica (ITP)** como complemento del IAS, mediante observaciones de aula, análisis de planificaciones y otras evidencias de práctica docente.

3. Utilizar el IAS como herramienta de gestión. Emplear el índice para identificar centros con alto potencial de liderazgo pedagógico y para priorizar acciones de acompañamiento en aquellos con menores niveles de aprovechamiento.

4. Fortalecer la evaluación de impacto. En futuras cohortes, incorporar diseños metodológicos que incluyan línea de base, grupos de comparación y seguimiento longitudinal para estimar el efecto atribuible a los programas de formación.

5. Desarrollar tableros de control institucional. Integrar los indicadores del SIEPFD en plataformas de monitoreo que faciliten el seguimiento oportuno de los programas y apoyen la toma de decisiones.

Recomendaciones para las instituciones formadoras

Las instituciones responsables de la ejecución de programas de formación docente deberían:

- fortalecer los mecanismos de acompañamiento posterior a la capacitación;
- promover comunidades profesionales de aprendizaje;
- incorporar estrategias sistemáticas de seguimiento a la implementación de las metodologías desarrolladas;
- documentar y compartir experiencias exitosas entre centros educativos.

Recomendaciones para los centros educativos

Los resultados sugieren la conveniencia de:

- consolidar equipos de liderazgo pedagógico;
- promover espacios permanentes de reflexión sobre la práctica docente;
- fortalecer el trabajo colaborativo entre docentes;
- utilizar los resultados de las evaluaciones externas como insumo para la mejora continua.

10.3 Propuesta para la implementación del SIEPFD

Se recomienda que el INAFOCAM adopte una estrategia gradual de implementación.

Primera etapa

Diseño de instrumentos.

- Validación técnica del IAS.
- Desarrollo del Índice de Transferencia Pedagógica.
- Elaboración de protocolos de evaluación.

Segunda etapa

Aplicación piloto.

- Implementación del sistema en nuevos programas de Matemática.
- Ajuste metodológico.
- Capacitación de equipos técnicos.

Tercera etapa

Escalamiento institucional.

- Incorporación del sistema al proceso regular de monitoreo y evaluación.
- Integración con la planificación institucional.
- Publicación periódica de informes de resultados.

10.4 Agenda de investigación

El presente estudio abre diversas líneas para investigaciones futuras.

Se recomienda desarrollar estudios orientados a:

- validar empíricamente el IAS mediante nuevas cohortes;
- analizar la sensibilidad del índice frente a distintos esquemas de ponderación;
- incorporar variables contextuales relacionadas con estudiantes, docentes y centros educativos;
- evaluar la relación entre el IAS y medidas de transferencia pedagógica;
- estimar el impacto causal de los programas utilizando metodologías cuasi experimentales o experimentales;
- adaptar el SIEPFD a programas de formación en otras áreas curriculares.

Estas investigaciones permitirán consolidar progresivamente un sistema de evaluación basado en evidencia y fortalecer la capacidad analítica del INAFOCAM.

10.5 Consideraciones finales

La formación continua del profesorado representa una de las inversiones más importantes para mejorar la calidad de los sistemas educativos. Sin embargo, su efectividad depende no solo de la calidad de los programas ofrecidos, sino también de la capacidad institucional para evaluar sus resultados, aprender de la evidencia y ajustar continuamente las estrategias de intervención.

El presente informe demuestra que la evaluación de programas de formación docente puede trascender la medición de actividades y cobertura para orientarse hacia el análisis del desempeño institucional y de los resultados de aprendizaje. La incorporación del **Índice de Aprovechamiento STEM (IAS)** y del **Sistema Integrado para la Evaluación de Programas de Formación Docente (SIEPFD)** constituye una propuesta metodológica que busca fortalecer esa transición.

Más allá del análisis del programa STEM, el estudio plantea un cambio de enfoque: pasar de una lógica centrada en la ejecución de la capacitación a una cultura institucional de evaluación basada en evidencia, orientada al aprendizaje organizacional y a la mejora continua. En este sentido, el informe no solo ofrece una valoración de un programa específico, sino también una ruta para consolidar un sistema moderno de evaluación que contribuya al fortalecimiento de las políticas de desarrollo profesional docente en la República Dominicana.

Bibliografía citada

Ayala de la Cruz, M. E. (2016). *Evaluación de la formación permanente del profesorado en la República Dominicana: El caso del INAFOCAM* (Tesis doctoral). Universidad de Murcia.

Desimone, L. M. (2009). Improving impact studies of teachers' professional development: Toward better conceptualizations and measures. *Educational Researcher*, 38(3), 181–199.

Gertler, P. J., Martínez, S., Premand, P., Rawlings, L. B., & Vermeersch, C. M. J. (2016). *Impact evaluation in practice* (2nd ed.). World Bank

Organization for Economic Co-operation and Development (OECD), & Joint Research Centre (JRC). (2008). *Handbook on constructing composite indicators: Methodology and user guide*. OECD Publishing.

Popova, A., Evans, D. K., Breeding, M. E., & Arancibia, V. (2018). Teacher professional development around the world: The gap between evidence and practice. *World Bank Research Observer*, 37(1), 107–13

Rossi, P. H., Lipsey, M. W., & Henry, G. T. (2019). *Evaluation: A systematic approach* (8th ed.). SAGE.

Timperley, H., Wilson, A., Barrar, H., & Fung, I. (2007). *Teacher professional learning and development: Best evidence synthesis iteration (BES)*. Ministry of Education.

Van den Brande, J., & Zuccollo, J. (2021). *The effects of high-quality professional development on teachers and students: A cost-benefit analysis*. Education Policy Institute.

Visscher, A. J., Dmoshinskaia, N., Pellegrini, M., & Rey-Naizaque, A. (2025). (When) do teacher professional development interventions improve student achievement? A meta-analysis of 128 high-quality studies. *Educational Research Review*, 49, 100742.

White, H. (2009). Theory-based impact evaluation: Principles and practice. *Journal of Development Effectiveness*, 1(3), 271–284.

Yoon, K. S., Duncan, T., Lee, S. W.-Y., Scarloss, B., & Shapley, K. (2007). *Reviewing the evidence on how teacher professional development affects student achievement* (Issues & Answers Report, REL 2007-No. 033). Institute of Education Sciences, U.S. Department of Education.

ANEXOS

ANEXO 1: Resultados parametricos y no parametricos

1.1 Resultados de la prueba t para muestras relacionadas

Resultados de la prueba t pareada

Comparación	Diferencia media	Estadístico t	Valor p	IC 95 % de la diferencia	Decisión
Politécnico Distrito	-0.239	-0.216	13 0.832	(-2.619; 2.142)	No se rechaza H_0
Politécnico Regional	-0.094	-0.069	13 0.946	(-3.013; 2.825)	No se rechaza H_0
Politécnico Nacional	1.572	1.241	13 0.237	(-1.165; 4.309)	No se rechaza H_0

Interpretación

Las comparaciones de medias entre los politécnicos participantes y sus referentes territoriales no evidencian diferencias estadísticamente significativas al nivel de significación del 5 %. Los tres valores p son superiores a 0.05 y los intervalos de confianza del 95 % incluyen el valor cero, lo que indica que las diferencias observadas pueden atribuirse a la variabilidad muestral y no proporcionan evidencia suficiente para rechazar la hipótesis nula.

1.2 Resultados de la prueba de Wilcoxon

Resultados de la prueba de rangos con signo de Wilcoxon

Comparación	Estadístico W	Valor p	Decisión
Politécnico – Distrito	48.0	0.808	No se rechaza H_0
Politécnico – Regional	52.0	1.000	No se rechaza H_0
Politécnico – Nacional	29.0	0.153	No se rechaza H_0

Interpretación

La prueba no paramétrica confirma íntegramente los resultados obtenidos mediante la prueba t . En ninguna de las comparaciones se detectan diferencias estadísticamente significativas entre los politécnicos y los referentes utilizados, reforzando la consistencia de la evidencia inferencial.

1.3 Resultados del tamaño del efecto

Tamaño del efecto (Cohen's d)

Comparación	Cohen's <i>d</i>	Magnitud
Politécnico – Distrito	-0.058	Efecto prácticamente nulo
Politécnico – Regional	-0.019	Efecto prácticamente nulo
Politécnico – Nacional	0.332	Efecto pequeño

Interpretación

Aunque ninguna de las comparaciones alcanzó significación estadística, la comparación con el promedio nacional presentó un tamaño del efecto pequeño ($d = 0.332$), indicando una ventaja descriptiva de los politécnicos participantes respecto al promedio nacional. En contraste, las diferencias respecto al distrito y a la regional fueron prácticamente nulas.

Interpretación integrada

La evidencia obtenida mediante las pruebas paramétricas y no paramétricas converge en una misma conclusión: no existen diferencias estadísticamente significativas entre el desempeño promedio de los politécnicos participantes y los promedios de los distritos, las regionales y el nivel nacional. No obstante, la comparación con el promedio nacional muestra una diferencia media positiva de 1.57 puntos y un tamaño del efecto pequeño, lo que sugiere una tendencia favorable que no alcanza significación estadística, probablemente debido al reducido tamaño de la muestra (14 politécnicos). Esta situación pone de manifiesto las limitaciones de las comparaciones basadas exclusivamente en pruebas de hipótesis y justifica la incorporación del **Índice de Aprovechamiento STEM (IAS)** como una medida complementaria para valorar el desempeño institucional de manera más integral.

ANEXO 2: Desarrollo metodológico y validación del Índice de Aprovechamiento STEM (IAS)

2.1 Introducción

El análisis inferencial desarrollado en el Anexo D mostró que las comparaciones de medias mediante pruebas paramétricas y no paramétricas constituyen herramientas indispensables para evaluar la existencia de diferencias estadísticamente significativas entre grupos. Sin embargo, dichas pruebas presentan limitaciones cuando el tamaño de la muestra es reducido, existe una elevada heterogeneidad entre las unidades de análisis o el objetivo consiste en valorar el desempeño relativo de instituciones individuales más que contrastar medias poblacionales.

En el presente estudio, aunque las pruebas *t* y de Wilcoxon no evidenciaron diferencias estadísticamente significativas entre los politécnicos participantes y sus referentes territoriales, la comparación con el promedio nacional mostró una ventaja descriptiva y un tamaño del efecto pequeño. Esta situación motivó el desarrollo del **Índice de Aprovechamiento STEM (IAS)** como un indicador sintético destinado a medir el grado en que cada politécnico aprovechó la intervención, considerando simultáneamente su desempeño respecto al distrito educativo, la regional y el promedio nacional.

El IAS no sustituye las pruebas estadísticas tradicionales. Por el contrario, constituye un instrumento complementario que permite evaluar el desempeño institucional desde una

perspectiva comparativa, facilitar la clasificación de los centros educativos y orientar la toma de decisiones basada en evidencia.

2.2 Fundamento conceptual

La evaluación tradicional responde una pregunta:

¿Existen diferencias estadísticamente significativas?

El IAS responde otra:

¿Qué tan favorable fue el desempeño relativo de cada politécnico frente a sus referentes educativos?

2.3 Construcción del indicador y calculo brechas

Definición de las variables.

PP = Promedio del Politécnico

PD = Promedio Distrital

PR = Promedio Regional

PN = Promedio Nacional

Cálculo de las brechas

$$D = PP - PD$$

$$R = PP - PR$$

$$N = PP - PN$$

2.4 Índice Ponderado de Desempeño Relativo (IPDR)

La ecuación fundamental es la siguiente:

$$IPDR = 0.50D + 0.30R + 0.20N$$

Donde D es el Distrito, R la Regional, N el nivel Nacional

2.5 Normalización

Para explicar adecuadamente los valores se requiere convertir el IPDR a una escala común:

$$IAS = 100 \left(\frac{IPDR - IPDR_{min}}{IPDR_{max} - IPDR_{min}} \right)$$

La transformación preserva el orden relativo de los centros educativos y facilita la interpretación del índice.

2.6 Clasificación

IAS Nivel

80–100 Muy alto

60–79 Alto

40–59 Medio

20–39 Bajo

0–19 Muy bajo

2.7 Propiedades del IAS

El índice posee varias propiedades deseables:

- Integra múltiples niveles de comparación.
- Es independiente de la escala original de las pruebas.
- Facilita comparaciones entre instituciones.
- Es fácilmente interpretable.
- Puede aplicarse a distintas áreas curriculares.
- Permite elaborar rankings y categorías.
- Puede incorporarse a sistemas de monitoreo institucional.

2.8 Fortalezas y limitaciones

Fortalezas

- Sintetiza información compleja.
- Facilita la toma de decisiones.
- Complementa el análisis inferencial.
- Permite comparaciones longitudinales.

Limitaciones

- Depende de las ponderaciones adoptadas.
- No incorpora variables contextuales.
- No mide directamente la transferencia pedagógica.
- Requiere validación con nuevas cohortes.

2.9 Posibles desarrollos futuros

El IAS puede evolucionar hacia una familia de indicadores:

Por ejemplo:

- IAS Matemática.

- IAS Ciencias.
- IAS Lectura.
- IAS Competencias Digitales.
- IAS Multidimensional.

ANEXO 3

Desarrollo metodológico del Índice de Transferencia Pedagógica (ITP)

3.1 Introducción

La mejora de los resultados de aprendizaje constituye el objetivo último de los programas de formación docente; sin embargo, entre la capacitación recibida por el profesor y el desempeño de los estudiantes existe una etapa intermedia que con frecuencia no es observada: la **transferencia pedagógica**. Un docente puede completar exitosamente un programa de formación sin que ello implique necesariamente la incorporación efectiva de los nuevos conocimientos, metodologías y recursos al proceso de enseñanza.

El Índice de Transferencia Pedagógica (ITP) se propone medir precisamente ese proceso. Su finalidad es estimar el grado en que las competencias desarrolladas durante la capacitación son aplicadas de manera sistemática en la práctica docente y se traducen en mejoras observables en la gestión del aula. El ITP complementa al IAS, ya que mientras este último evalúa el aprovechamiento institucional expresado en los resultados de aprendizaje, el primero permite analizar el mecanismo mediante el cual la formación docente genera dichos resultados.

3.2 Fundamento conceptual

El ITP se sustenta en el supuesto de que la cadena causal de un programa de formación docente comprende, al menos, cuatro etapas sucesivas:

Capacitación → Aprendizaje docente → Transferencia pedagógica → Resultados de aprendizaje.

La transferencia constituye el vínculo entre la formación recibida y el desempeño institucional, razón por la cual su medición resulta indispensable para comprender el impacto real de la intervención.

3.3 Dimensiones del ITP

El índice se estructura en cinco dimensiones:

1. **Aplicación metodológica:** uso de estrategias activas, resolución de problemas, enfoque STEM, aprendizaje basado en proyectos.
2. **Planificación didáctica:** incorporación de los contenidos y recursos desarrollados durante la capacitación.
3. **Práctica de aula:** observación de clases, interacción docente–estudiante, evaluación formativa.

4. **Innovación y uso de recursos:** utilización de tecnologías, laboratorios y materiales didácticos.
5. **Reflexión y mejora continua:** autoevaluación, comunidades de aprendizaje y retroalimentación profesional.

3.4 Construcción del índice

Cada dimensión se mediría mediante un conjunto de indicadores valorados en una escala de 0 a 4. El índice general se obtendría como una combinación ponderada de las dimensiones:

$$ITP = \sum_{i=1}^5 w_i D_i$$

donde D_i representa la puntuación de cada dimensión y w_i su ponderación, cumpliéndose:

$$\sum_{i=1}^5 w_i = 1$$

En una primera aplicación puede asignarse igual ponderación a todas las dimensiones ($w_i = 0.20$), dejando abierta la posibilidad de ajustar los pesos mediante validaciones posteriores.

3.5 Interpretación

Se propone una escala de cinco niveles:

ITP	Nivel
80–100	Muy alta transferencia
60–79	Alta
40–59	Moderada
20–39	Baja
0–19	Muy baja

3.6 Instrumentos de medición

El ITP puede construirse a partir de diversas fuentes:

- observación estructurada de clases;
- rúbricas de desempeño docente;
- revisión de planificaciones;
- portafolios profesionales;
- entrevistas y grupos focales;
- encuestas a docentes y directores;

- evidencias de aprendizaje estudiantil.

3.7 Relación entre ITP e IAS

El ITP mide la **transferencia de la capacitación a la práctica docente**; el IAS mide el **aprovechamiento institucional reflejado en los resultados de aprendizaje**. Ambos indicadores son complementarios y permiten analizar la cadena completa de generación de impacto.

ANEXO 4

Sistema Integrado para la Evaluación de Programas de Formación Docente (SIEPFD)

4.1 Introducción

El SIEPFD constituye una propuesta metodológica para evaluar integralmente programas de formación docente, articulando información sobre implementación, transferencia pedagógica, aprovechamiento institucional, impacto educativo y eficiencia del gasto público. Su propósito es ofrecer un marco sistemático que permita a instituciones como el INAFOCAM monitorear, evaluar y mejorar continuamente sus programas de desarrollo profesional docente.

4.2 Principios del sistema

1. Integralidad.
2. Continuidad.
3. Multinivel.
4. Comparabilidad.
5. Aprendizaje institucional.
6. Decisiones basadas en evidencia.
7. **Eficiencia**, entendida como la búsqueda de los mayores resultados posibles con el uso más racional de los recursos públicos.

4.3 Arquitectura del sistema

El SIEPFD se organiza en cuatro niveles secuenciales:

Nivel I. Implementación del programa

- Cobertura.
- Participación.
- Finalización.
- Satisfacción.

Nivel II. Transferencia pedagógica

- Aplicación en el aula.
- Innovación didáctica.
- Prácticas docentes.

- **Indicador principal: ITP.**

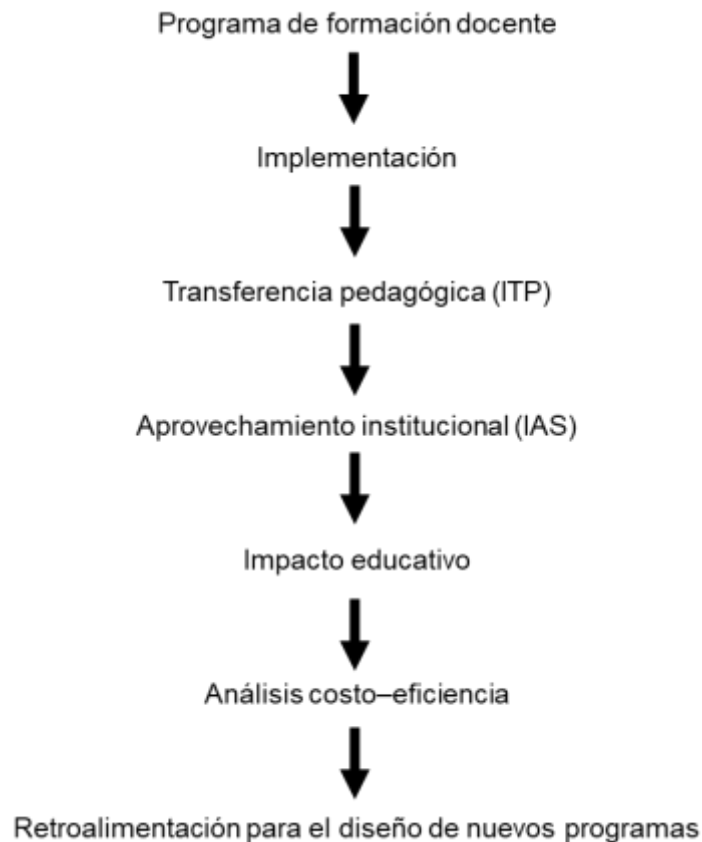
Nivel III. Aprovechamiento institucional

- Resultados en aprendizaje.
- Comparación con referentes.
- Clasificación institucional.
- **Indicador principal: IAS.**

Nivel IV. Impacto y eficiencia

- Evaluación de impacto.
- Costo–eficiencia.
- Costo–efectividad.
- Costo por punto adicional del IAS.
- Costo por docente con alta transferencia pedagógica.

4.4 Flujo de información



4.5 Productos del sistema

El SIEPFD permite generar:

- indicadores comparables entre programas;
- rankings institucionales;
- informes de desempeño;
- tableros de control;
- evaluaciones de impacto;
- análisis de eficiencia del gasto;
- recomendaciones para la mejora continua y la planificación de nuevas intervenciones.

4.6 Agenda de desarrollo

- validación del ITP en diferentes contextos y áreas curriculares;
- aplicación del IAS e ITP a nuevas cohortes;
- incorporación de modelos multinivel y análisis longitudinales;
- integración con sistemas de información del INAFOCAM y del MINERD;
- desarrollo de paneles de seguimiento en tiempo real;
- incorporación gradual de analítica de aprendizaje e inteligencia artificial para apoyar la toma de decisiones.

Una observación estratégica

Los anexos aquí presentados conforman una **arquitectura metodológica completa** para la evaluación de programas de formación docente. El IAS responde a la pregunta **qué tanto mejoró la institución**; el ITP explica **cómo la capacitación se transformó en práctica pedagógica**; y el SIEPFD integra ambas dimensiones con la implementación y la eficiencia para ofrecer un sistema de evaluación replicable y escalable. Esa articulación constituye, el aporte más original y de mayor proyección del estudio.