

Efectos del aula invertida y la evaluación auténtica en el aprendizaje de la matemática universitaria en estudiantes de primer año de ingeniería

JORGE MALUENDA ALBORNOZ*

MARCELA VARAS CONTRERAS**

DIEGO CHACANO OSSES***

Universidad de Concepción – Chile

Recibido el 02-06-20; primera evaluación el 04-01-21;
segunda evaluación el 22-01-21; aceptado el 10-02-21

RESUMEN

Las altas tasas de reprobación en Introducción a la Matemática Universitaria (IMU) han llevado a que la presente investigación desarrolle una intervención basada en Flipped Classroom y Authentic Assessment para revertir esta situación. El objetivo del estudio fue evaluar el efecto de la implementación de estos dos enfoques en el desempeño de estudiantes de ingeniería de primer año en IMU. Se realizó un estudio cuasiexperimental con grupo de control e intervención, además de pre y post-test para medir cambios en el aprendizaje y la satisfacción de los estudiantes. Los resultados mostraron diferencias estadísticamente significativas favorables a la sección intervenida, reflejando un 18,5% de mejora en los resultados académicos. Además, se observó mayor satisfacción en esta sección, apreciándose la utilidad de la intervención para aportar en ambos aspectos.

Palabras clave: aula invertida, evaluación auténtica, aprendizaje, educación en ingeniería, matemática universitaria.

* Psicólogo educacional. Magíster en Política y Gobierno y candidato a doctor en Psicología. Coordinador de Desarrollo Curricular y docente de la Facultad de Ingeniería de la Universidad de Concepción. Correo electrónico: jorgemaluenda@udec.cl. <https://orcid.org/0000-0001-8148-4948>

** Ingeniera Civil Informática. Magíster en Ciencias de la Computación. Profesora asistente y directora del Departamento de Ingeniería Informática y Ciencias de la Computación, Universidad de Concepción. Correo electrónico: mvaras@udec.cl. <https://orcid.org/0000-0003-2721-9865>

*** Psicólogo educacional. Se desenvuelve en métodos y estrategias efectivas de aprendizaje para estudiantes de educación superior. Correo electrónico: dchacano@udec.cl. <https://orcid.org/0000-0002-2979-927X>



Effects on learning and satisfaction of using flipped classroom and authentic assessment in university maths for engineering freshmen

ABSTRACT

The high failure rates in Introduction to University Mathematics (IUM) caused the investigation to develop an intervention based on the incorporation of flipped classroom and authentic assessment to reverse this situation. The objective of the present study was to evaluate the effect of the implementation of these two approaches on the performance of engineering freshmen at IMU. A quasi-experimental study was carried out with a control and intervention group, as well as pre and post-test to measure learning and student satisfaction. The results showed statistically significant differences in the overall results and for each evaluation favorable to the changed section, reflecting an 18.5% improvement in academic results. In addition, greater general satisfaction was observed in this section, which shows a general utility of the intervention to contribute to both aspects.

Keywords: Flipped classroom, authentic assessment, learning, engineering education, University Mathematics.

Efeitos na aprendizagem e satisfação do uso da sala de aula invertida e avaliação autêntica em matemática da universidade para estudantes de engenharia do primeiro ano

RESUMO

As altas taxas de falhas em Introdução à Matemática Universitária (IMU) levaram a presente investigação a desenvolver uma intervenção baseada na classe invertida e na avaliação autêntica para reverter essa situação. O objetivo do estudo foi avaliar o efeito da implementação dessas duas abordagens no desempenho dos alunos do primeiro ano de engenharia da IMU. Um estudo quase experimental foi realizado com um grupo de controle e intervenção, além do pré e pós-teste para medir as mudanças na aprendizagem e na satisfação do aluno. Os resultados mostraram diferenças estatisticamente significantes favoráveis à seção interveniente, refletindo uma melhoria de 18,5% nos resultados acadêmicos. Além disso, foi observada maior satisfação nesta seção, valorizando a utilidade da intervenção em contribuir em ambos os aspectos.

Palavras-chave: classe invertida, avaliação autêntica, aprendizagem, educação em engenharia, universidade matemática.

1. INTRODUCCIÓN

En Chile, el acceso a la educación superior ha logrado un gran crecimiento (Acuña, 2012; Cox, 2012; Espinoza y González-Fiegehen, 2011) caracterizado por una ampliación en el acceso que ha permitido a jóvenes que antes se encontraban excluidos del sistema, emprender estudios de educación superior (Espinoza, 2013; Espinoza y González, 2011, 2015).

Esta ampliación en el acceso ha provocado una mayor diversificación del perfil estudiantil, produciendo el ingreso masivo de estudiantes denominados «primera generación». Estos estudiantes, corresponden a aquellos que ingresan a la Educación Superior sin que sus padres lo hayan hecho (Castillo y Cabezas, 2010; Jehangir, 2010; Thomas y Quinn, 2006).

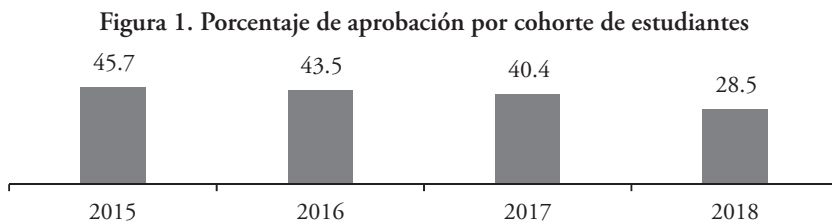
La OECD (2013) estima que 7 de cada 10 estudiantes se encuentran dentro de esta categoría. Al respecto, Flanagan (2017) plantea que ser un estudiante de primera generación y haber cursado la educación secundaria en un establecimiento público reduce las posibilidades de llegar bien preparado al nivel superior, permanecer en la carrera, y tener éxito en los estudios. Este último punto agrega una dificultad adicional vinculada con la desigualdad en la calidad educativa previa a la educación superior.

Estas dificultades se encuentran focalizadas en la capacidad de la educación escolar para desarrollar habilidades cognitivas y procedimentales en las áreas de Lenguaje y Matemática, especialmente, en la población socioeconómicamente más desfavorecida (Rodríguez y Jarpa, 2015). De este modo, el nivel socioeconómico, y más específicamente el capital social del estudiante, juega un rol gravitante en sus posibilidades de abordar con éxito los estudios superiores, más allá de su ingreso al sistema.

Sumado a esta dificultad, es importante consignar que las universidades han desarrollado sus políticas internas, normativas y estrategias pensando en un perfil estudiantil previo a la masificación de la educación superior. Por esta razón, la investigación sobre las variables que han adquirido importancia con el cambio del perfil estudiantil ha cobrado una importancia inusitada para comprender y enfrentar de mejor manera los problemas asociados a este fenómeno.

En el caso de la Facultad de Ingeniería, objeto de la presente investigación, existen algunos antecedentes relevantes a considerar. Se ha observado en cohortes anteriores de estudiantes la existencia de una brecha entre los conocimientos previos de los recién ingresados y los necesarios para la superación de los primeros años de la vida universitaria. Una de las asignaturas críticas es Introducción a la Matemática Universitaria, donde el porcen-

taje de aprobación obtenido por los estudiantes inscritos en esta asignatura obligatoria de segundo trimestre ha sido inferior, en general, al 50% (ver Figura 1).



Fuente: Elaboración propia a partir del sistema de control de gestión de la Universidad de Concepción.

Este tipo de asignaturas suele ser impartida con métodos centrados en el docente donde la clase expositiva y la resolución de ejercicios en la pizarra, son los principales abordajes del proceso de enseñanza-aprendizaje. Esto genera dos efectos principales: a) Una escasa participación activa del estudiante en su aprendizaje, con todas las consecuencias que ello conlleva sobre la motivación y el aprendizaje del estudiante, y b) Una orientación del tiempo docente en una actividad estándar que reduce las posibilidades de una vinculación efectiva entre docente estudiante, retroalimentación significativa y pormenorizada, así como orientación y corrección pertinente y oportuna.

Existen enfoques que proponen las formas para desarrollar el aprendizaje significativo y la vinculación efectiva. Uno es la clase invertida, que permitiría focalizar el tiempo para las actividades que son significativas para el aprendizaje como por ejemplo, el reforzamiento del material y el desarrollo de actividades prácticas (Mason, Shuman y Cook, 2013). De este modo, se pueden aprovechar de una forma más efectiva las horas pedagógicas para generar efectos más potentes en el aprendizaje de los estudiantes.

Otro enfoque es el de evaluación auténtica, que, entre otras cosas, permite orientar la evaluación no solo como medición de logros, sino también como una herramienta que favorezca el aprendizaje a través de la retroalimentación precisa, efectiva y contextualizada en el contexto profesional real. Lo anterior, permite la estimulación de los estudiantes hacia la resolución de problemas (Elliott y Higgins, 2005; Villarroel, Bloxham, Bruna, Bruna y Herrera-Seda, 2018) y el establecimiento de relaciones entre las nuevas ideas y el conocimiento previo de estos (Villarroel et al., 2018).

2. MARCO TEÓRICO

2.1. Clase invertida

Esta metodología fue concebida por Bergmann y Sams (2012) como una opción para organizar de forma óptima el tiempo de clases. En el aula invertida lo que es tradicionalmente hecho en clases y como tarea en el hogar es lo invertido (Gilboy, Heinerichs y Pazzaglia, 2015). Por ejemplo, en vez de asistir a una exposición de una clase sobre análisis de datos de tipo descriptivo para luego, en casa, resolver ejercicios relacionados con ella, los estudiantes estudian el material, ven videos, escuchan *podcasts* o realizan ejercicios antes de asistir a la clase y luego se involucran en estrategias de aprendizaje activo durante la clase; como debates, conferencias interactivas, resolución de problemas, experimentos de laboratorio, juegos de roles, diseño y creación colaborativa (Strayer, 2012).

Hay tres razones principales para usar una clase invertida (Mason et al., 2013):

- i. Las clases invertidas dejan tiempo libre para actividades interactivas y para reforzamiento del material del curso sin sacrificar su contenido,
- ii. Permiten a un educador presentar el material del curso en distintos formatos e involucrar los distintos estilos de aprendizaje de los estudiantes,
- iii. Alientan a los estudiantes a volverse autodidactas, ayudándolos a prepararse para aprender como ingenieros practicantes.

Por ello, distintos autores han postulado que un buen diseño de aula invertida aumenta la calidad y la eficiencia de los procesos de enseñanza y aprendizaje (Demski, 2013; Estes, Ingram y Liu, 2014).

Investigaciones recientes han mostrado que los estudiantes pueden mantenerse altamente comprometidos al usarse el modelo de clase invertida (Dove, 2013) permitiendo que los estudiantes adquieran mayor responsabilidad sobre su aprendizaje al controlar el ritmo de sus acciones (Alvarez, 2012).

A partir de una extensa revisión de literatura realizada por Karabulut-Ilgu, Jaramillo Cherez y Jahren, (2018) se ha propuesto que, uno de los beneficios más comúnmente citados del modelo de clase invertida es la flexibilidad. Esta ocurre gracias a que los estudiantes disponen de recursos didácticos del curso (en sus distintas formas) en todo momento, los cuales pueden utilizar para preparar, repasar y reforzar sus aprendizajes.

La dificultad más común del uso de aula invertida, es la necesidad de invertir una importante cantidad de tiempo para el diseño de los cursos, además, del tiempo de trabajo autónomo del estudiante (Argente, García-Fornes y Espinosa, 2016, en Torres y Cañadas, s.a).

Desde nuestra perspectiva, ambos puntos pueden tener una interpretación favorable:

- a) El tiempo de diseño es una inversión que se retribuye en la disponibilidad permanente de recursos didácticos de buena calidad que son útiles para el curso en cuestión y otros cursos.
- b) El tiempo que los estudiantes dedican es más eficiente dado que reemplaza horas de clase improductivas y tediosas.

Resultados empíricos de clase invertida en el contexto de ingeniería

En un estudio experimental realizado por Mason et al. (2013), quien tomó como muestra dos cursos de Sistemas de Control en Ingeniería Mecánica, se encontraron diferencias estadísticamente significativas en el desempeño en un grupo de problemas de la disciplina. El grupo en que se implementó la metodología de aula invertida mostró un desempeño significativamente mejor ($p < .05$). Además, este grupo mostró un mejor desempeño en problemas vinculados al diseño de soluciones ($p < .001$).

Otro estudio experimental realizado por Turra, Carrasco, González, Sandoval y Yáñez (2019) en tres cursos de matemáticas para estudiantes de ingeniería chilenos, analizó el impacto de la transformación de estos cursos basándose en modelos de aula invertida. Utilizando el Inventario de actitudes hacia las matemáticas –ATMI- se observó un mayor puntaje en el postest ($p < .05$) en las cuatro categorías del inventario (confianza personal, utilidad del contenido del tema, actitud del profesor y la percepción del tema como de dominio masculino). Este último factor mostró el cambio más amplio, lo que significa que los estudiantes cambiaron su percepción respecto al estereotipo de la disciplina matemática como un tema masculino, por una perspectiva más balanceada.

2.2. Evaluación auténtica

El término «evaluación auténtica» fue acuñado por Wiggins (1989) quien lo definió como una «prueba verdadera» del logro o habilidad intelectual de los estudiantes ya que requiere que demuestren un entendimiento profundo,

un pensamiento de orden superior, y una resolución de problemas complejos a través del desarrollo de tareas ejemplares (Koh, 2019).

Consiste en desarrollar evaluaciones que cumplan criterios de validez que permitan garantizar la adecuada coherencia entre los propósitos de la enseñanza y los métodos utilizados para evaluar. A su vez, se preocupa por resguardar que los propósitos de la enseñanza que son objeto de evaluación estén en estrecho vínculo con los aprendizajes requeridos para el ejercicio laboral real. De este modo, se diseñan evaluaciones pertinentes y coherentes que den cuenta efectiva de los logros de los aprendizajes y su ajuste respecto de los niveles esperados (Martínez, 2016).

La evaluación auténtica apunta a integrar lo que sucede en la sala de clases con el empleo, replicando las tareas y los estándares de desempeño típicamente enfrentados por profesionales en el mundo del trabajo (Wiggins, 1990).

En estas tareas, se requiere que los estudiantes produzcan o demuestren conocimiento y habilidades en actividades cercanas a la profesión. Por ejemplo, en una evaluación escrita se pueden incluir análisis de caso, resolución de problemas y preguntas de ensayo, que se usan como representación del mundo real (Villarroel et al., 2018). Sin embargo, la evaluación auténtica debe incluir además otras características, como ser cognitivamente desafiante (Ashford-Rowe, Herrington y Brown, 2014), pues debería estimular a los estudiantes a involucrarse en resolver problemas actuales aplicando conocimientos y tomando decisiones, llevándolos al desarrollo de habilidades cognitivas y metacognitivas (Elliott y Higgins, 2005; Villarroel et al., 2018).

La evaluación auténtica debería conducir a los estudiantes a establecer relaciones entre las nuevas ideas y el conocimiento previo, enlazando conceptos teóricos con la experiencia diaria, sacando conclusiones del análisis de datos, permitiéndoles examinar tanto la lógica de los argumentos presentes en la teoría como su alcance práctico (Villarroel et al., 2018).

En las tareas de evaluación auténtica, los estudiantes oscilan entre el rol de estudiante y futuro practicante mientras se dispersan por el ámbito académico y el mundo del trabajo (Lingard, Schryer, Garwood y Spafford, 2003; Field, Duffy y Huggins, 2013).

Las actividades de evaluación deberían alentar a los estudiantes a comprometerse con los estándares de un buen desempeño, juzgando su propia actuación y regulando su aprendizaje (Tai, Ajjawi, Boud, Dawson y Panadero, 2018; Villarroel et al., 2018). Este aspecto resulta importante debido a que permite que los estudiantes identifiquen áreas que necesitan mejoría, que puedan monitorear su progreso, y que desarrollen ideas con estándares aceptables de desempeño de calidad en su futura profesión (Dove, 2013; Sadler, 2005).

Resultados empíricos de la evaluación auténtica

En un estudio experimental realizado por James y Casidy (2018) con una muestra de 120 estudiantes de pregrado en un programa de negocios, se encontró que la evaluación auténtica estaba positivamente asociada a la satisfacción de los estudiantes con el programa de estudios ($\beta = 1.14$, $SE = .22$, $t = 5.08$, $p < .001$) y con la utilidad de la experiencia educativa para incrementar su potencial empleabilidad ($\beta = 1.08$, $SE = .20$, $t = 5.42$, $p < .001$).

En otro estudio, realizado por Gómez, Gracia y Robles (2019), de tipo cuasi experimental en donde se incluyó a 87 alumnos del curso de metodología de la investigación de la licenciatura de medicina, se determinó que el uso de una rúbrica realizada basada en evaluación auténtica produce mayor autocontrol del aprendizaje en el grupo intervenido ($p < .05$).

En un estudio experimental realizado por Bakr, Massey y Massa (2016) que tomó como muestra un curso de anatomía dental de primer año, se analizó en retrospectiva una intervención que aplicó evaluaciones basadas en el enfoque de evaluación auténtica. Los resultados mostraron mejoras significativas en el desempeño general de los estudiantes en los exámenes prácticos de mitad ($p < .05$) y final de semestre ($p < .05$) en relación con los estudiantes que trabajaron con métodos tradicionales de aprendizaje.

Los distintos antecedentes presentados conducen a que la presente investigación haya tenido por finalidad evaluar los resultados de la incorporación de ambos métodos, aula invertida y evaluación auténtica, para el favorecimiento del aprendizaje en matemáticas en estudiantes de ingeniería.

Los objetivos de la presente investigación fueron:

- a) Evaluar el efecto de la implementación de aula invertida y evaluación auténtica en el desempeño estudiantes de ingeniería de primer año que cursan la asignatura de Introducción a la Matemática Universitaria.
- b) Evaluar el efecto de la implementación de dichos métodos en la satisfacción de estos estudiantes.

3. METODOLOGÍA

3.1. Participantes

La muestra fue extraída a partir de un muestreo intencional que consideró la participación de 736 estudiantes de primer año y primer semestre de las 13 carreras de ingeniería de una universidad chilena. Participaron todos los estudiantes de Introducción a la Matemática Universitaria (IMU).

Se distribuyeron en 9 secciones, 8 secciones con enseñanza en su formato tradicional y una sección (sección 9 - 161 estudiantes) con las metodologías incorporadas. Las secciones tuvieron una mezcla aleatoria de los estudiantes: Los estudiantes de cada sección provienen de distintas carreras de ingeniería que el sistema informático de la universidad conforma por azar.

3.2. Diseño

Se realizó un estudio cuantitativo de tipo cuasiexperimental y de corte transversal (Ato, López y Benavente, 2013) para evaluar los efectos diferenciales del cambio metodológico realizado en la sección experimental (sección 9), en comparación con las demás secciones de la misma asignatura.

Los grupos de estudiantes no fueron conformados de manera aleatoria manteniendo su conformación natural. Como ya se indicó, esta conformación la genera automáticamente el sistema informático de la universidad de manera aleatoria, ubicando estudiantes de distintas especialidades en cada sección de clases.

El grupo de intervención corresponde al Plan Común de Ingeniería. Esta sección permite garantizar la heterogeneidad en las características de ingreso de los estudiantes (puntajes PSU, proporción hombres y mujeres). De este modo, se constituyó en el mejor representante del conjunto de estudiantes de ingeniería.

Los grupos de comparación corresponden a los estudiantes restantes distribuidos en las otras 8 secciones.

3.3. Instrumentos

Para la medición de los resultados académicos de los estudiantes se utilizó la evaluación académica tradicional (certámenes y test). Estos instrumentos han sido diseñados por la Facultad de Ciencias Físicas y Matemáticas de la universidad en cuestión, la cual, realiza regularmente en modalidad de «prestación de servicios» la asignatura IMU para los estudiantes de ingeniería de primer año, primer semestre.

Las evaluaciones tipo certamen contienen ejercicios de distinto tipo que recogen las unidades abordadas hasta ese punto del semestre en la asignatura. Incluyen ejercicios de análisis y razonamiento matemático, cálculo y toma de decisiones. Estos, son estándar semestre a semestre e idénticos para las distintas secciones de estudiantes que cursan el IMU.

Se decidió no intervenir ni modificar dichas evaluaciones por tres razones:

- a) Para no alterar la métrica que se ha utilizado regularmente en la evaluación de los estudiantes de modo de mantener la posibilidad de evaluar el registro histórico.
- b) Porque su diseño está en manos de expertos en las temáticas quienes tienen décadas de experiencia en la evaluación matemática para ingenieros.
- c) Para evitar evaluaciones disímiles entre secciones.

Todo lo anterior, fue realizado con la finalidad de que el instrumento fuese pertinente, ajustado y estándar, de modo de evitar que los resultados obtenidos se deban a la modificación del instrumento y no a los cambios metodológicos efectuados.

Adicionalmente, se agregaron evaluaciones formativas realizadas semanalmente (test breves), que fueron aplicados solamente a la sección piloto. Estas evaluaciones tuvieron la finalidad de obtener un complemento a la medición, pero también contribuir al aprendizaje de los estudiantes mediante su revisión, retroalimentación y análisis.

En cuanto a la evaluación de la satisfacción, se construyó un instrumento breve para consultar el nivel de conformidad de los estudiantes con su profesor, sus ayudantes y la asignatura. En cuanto al profesor, se consultó por la satisfacción del estudiante respecto de sus conocimientos sobre matemáticas, su forma de enseñanza y su disposición a ayudarles (resolver dudas, apoyar, etc.). Los ayudantes fueron evaluados en cuanto a sus conocimientos sobre matemáticas y su forma de enseñanza. Por último, la asignatura fue evaluada en cuanto a la satisfacción con las actividades prácticas, el método de evaluación y una evaluación de satisfacción global. En todos los casos, se utilizó una escala de respuesta de 1 a 5, donde 1 significa total insatisfacción y 5 máxima satisfacción.

3.4. Descripción de los cambios efectuados

Se desarrolló una modificación en la asignatura Introducción a la Matemática Universitaria (IMU). Esta asignatura, realizada regularmente en el primer trimestre para todas las carreras de ingeniería, imparte: lógica y conjuntos, números reales, geometría analítica, función exponencial y logarítmica, además de funciones circulares.

La modificación de la asignatura contempló el cambio de su planificación completa, sus métodos de enseñanza-aprendizaje y evaluación con el fin de ajustarlos a los principios de la Clase Invertida y Evaluación Auténtica. Estos cambios, implicaron que la intervención, se realizó durante todo el trimestre, comprendido por un total de 11 semanas de clases reales efectuadas.

En cuanto a la clase invertida, se reorganizó la planificación de la asignatura de modo tal que el contenido básico para cada módulo se administró de manera remota a través de una plataforma virtual en la forma de guías de trabajo (contenido y ejercicios) y videos explicativos sobre la comprensión y formulación de problemas matemáticos, además de la operatoria.

Se utilizó una guía de trabajo por módulo con una extensión de 22 páginas aproximadas por cada una. Su diseño fue en formato diapositiva el cual combinó texto explicativo, ejercicios fórmulas, etc. (ver Figura 2). Esta guía, se complementó con un listado de ejercicios propuestos y uno de ejercicios resueltos.

Figura 2. Ejemplo Guía de Trabajo

Operador Sumatoria

Definición

Dados n números reales indexados como $a_1, a_2, a_3, \dots, a_n$, la **sumatoria** de ellos se denota por:

$$\sum_{i=1}^n a_i = a_1 + a_2 + \dots + a_n$$

Propiedades

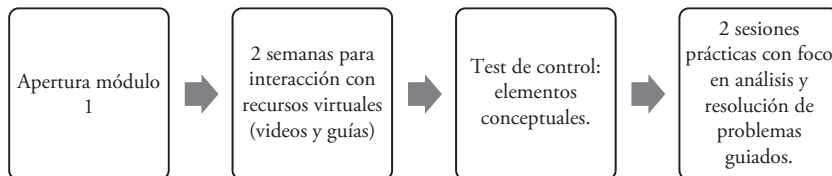
1. $\sum_{i=1}^n a_i = \sum_{j=1}^n a_j$
2. $\sum_{i=1}^n a_i = \sum_{i=0}^{n-1} a_{i+1} = \sum_{i=2}^{n+1} a_{i-1}$
3. $\sum_{i=1}^n a = na$

Los videos diseñados tuvieron una duración de entre 10 y 25 minutos y se focalizaron en conceptos clave de la temática de estudio, donde se explicó el contenido y se mostró un ejercicio vinculado.

El tiempo destinado al trabajo en el aula fue de 4 horas pedagógicas a la semana, divididas en dos sesiones prácticas. Estas sesiones se utilizaron para la explicación guiada por el docente, la revisión colectiva (en grupos) de los ejercicios de mayor complejidad y la resolución de dudas. Cada semana, antes de la clase los estudiantes debieron realizar un test (calificado) de los contenidos de clase con la finalidad de asegurar el dominio de los aspectos

teórico-conceptuales básicos necesarios para el aprovechamiento de las sesiones prácticas. La planificación de clases operó como se aprecia en la Figura 3.

Figura 3. Secuencia de Trabajo Virtual-presencial.



Fuente: Elaboración propia, hecha en base a la estructura del grupo intervenido.

Por otro lado, se implementaron los principios de la evaluación auténtica en los tests de la asignatura, tanto en aquellos formativos como en los que tuvieron calificación formal.

Se diseñaron tests orientados a la resolución de problemas matemáticos contextualizados en escenarios de ingeniería. Cada evaluación remota (aula virtual) contuvo una retroalimentación inmediata sobre los aspectos logrados y los errores, así como sugerencias de corrección.

En el caso de las evaluaciones presenciales, se aprovechó el trabajo de revisión colectiva en la clase presencial para la retroalimentación sobre los errores más comunes cometidos.

Tanto el diseño de las actividades y videos explicativos como de las evaluaciones de tests incorporadas fueron diseñadas por un equipo conformado por una docente especialista en matemáticas y un psicólogo educacional. La revisión de la correcta incorporación de estos insumos en el aula virtual fue también supervisada por este equipo.

3.5. Análisis de datos

Para el análisis de los datos, en primer lugar, se utilizó la estadística descriptiva para evaluar la distribución de los datos. Luego de la verificación de los supuestos asociados al uso de la técnica, se utilizó la prueba ANOVA (Analysis of Variance) de un factor para el análisis de los resultados en el desempeño de los estudiantes en IMU en función de su participación en los distintos grupos. Adicionalmente se calculó el tamaño del efecto de las diferencias encontradas. Para identificar específicamente entre qué grupos se producen estas diferencias, se utilizó el contraste de Bonferroni. Para todos los análisis, se utilizó el programa estadístico SPSS 21.

3.6. Procedimientos

La recolección de datos se realizó considerando los estudiantes de las 13 carreras que participaron del IMU 2019 tanto en su formato regular (no intervenido) como en el formato intervenido. Las evaluaciones fueron realizadas por los docentes de las asignaturas en su formato tradicional. Las calificaciones fueron ingresadas por ellos mismos al sistema interno de la universidad y la obtención de los datos por parte de los investigadores fue realizada mediante consulta a la secretaría académica de la facultad y el consentimiento de su uso para fines de investigación, de forma agregada y resguardando la confidencialidad de los estudiantes.

4. RESULTADOS

4.1. Resultados certamen 1

Al realizar la comparación entre los grupos en función de la calificación obtenida en el certamen 1 la prueba ANOVA de un factor reflejó la existencia de diferencias estadísticamente significativas entre las secciones ($F(8,99) = 3.56$, $P < .001$, $\eta^2 = .028$).

Las comparaciones múltiples de Bonferroni mostraron diferencias estadísticamente significativas entre la sección 9 y las secciones 1 ($t(270) = -3.16$, $p = .042$), 3 ($t(259) = -4.37$, $p < .001$), 4 ($t(266) = -3.36$, $p = .022$), 5 ($t(266) = -3.24$, $p = .033$) y 7 ($t(267) = -3.37$, $p = .021$). En todos los casos, incluso en aquellos en que no se observaron diferencias estadísticamente significativas, las calificaciones promedio fueron superiores para la sección 9.

Tabla 1. Resultados certamen 1 por sección

	1	2	3	4	5	6	7	8	9
Válidos	111	103	100	107	107	101	108	101	161
Perdidos	0	0	0	0	0	0	0	0	0
Media	4.28	4.40	4.02	4.23	4.25	4.57	4.23	4.59	4.88
Desviación estándar	1.61	1.58	1.56	1.42	1.49	1.51	1.56	1.73	1.51

Fuente: Elaboración propia a partir del sistema de gestión de control de Universidad de Concepción.

4.2. Resultados certamen 2

Al realizar la comparación entre los grupos en función de la calificación obtenida en el certamen 2 mediante la prueba ANOVA de un factor se observó

la existencia de diferencias estadísticamente significativas entre las secciones ($F(8,989) = 15.649$, $P < .001$, $\eta^2 = .112$).

Las comparaciones múltiples de Bonferroni mostraron diferencias estadísticamente significativas entre la sección 9 y las secciones 1 ($t(270) = -6.49$, $p < .001$), 2 ($t(262) = -7.16$, $p < .001$), 3 ($t(258) = -8.80$, $p < .001$), 4 ($t(266) = -7.16$, $p < .001$), 5 ($t(266) = -7.68$, $p < .001$), 6 ($t(260) = -6.75$, $p < .001$), 7 ($t(267) = -8.52$, $p < .001$) y 8 ($t(260) = -5.96$, $p < .001$). En todos los casos, incluso aquellos en que no hubo diferencias estadísticamente significativas, las calificaciones promedio fueron superiores para la sección 9.

Tabla 2. Resultados certamen 2 por sección

	1	2	3	4	5	6	7	8	9
Validos	111	103	99	107	107	101	108	101	161
Perdidos	0	0	1	0	0	0	0	0	0
Media	3.13	2.99	2.69	3.00	2.91	3.05	2.78	3.18	4.21
Desviación estándar	1.28	1.40	1.29	1.21	1.34	1.20	1.21	1.50	1.52

Fuente: Elaboración propia a partir del sistema de gestión de control de Universidad de Concepción

4.3. Resultados certamen 3

Al realizar la comparación entre los grupos en función de la calificación obtenida en el certamen 3 mediante la prueba ANOVA de un factor se observó la existencia de diferencias estadísticamente significativas entre las 9 secciones ($F(8,98) = 2.58$, $P = .009$, $\eta^2 = .021$).

Las comparaciones múltiples de Bonferroni mostraron diferencias estadísticamente significativas solamente entre la sección 1 y la sección 5 ($t(213) = 3.32$, $p < .033$). En todos los casos, incluso aquellos en que no hubo diferencias estadísticamente significativas, las calificaciones promedio fueron superiores para la sección 9, con excepción de la sección 1.

Tabla 3. Resultados certamen 3 por sección

	1	2	3	4	5	6	7	8	9
Validos	108	103	99	107	107	101	108	101	160
Perdidos	3	0	1	0	0	0	0	0	1
Media	3.65	3.22	3.01	3.34	2.97	3.13	3.10	3.18	3.50
Desviación estándar	1.48	1.63	1.55	1.49	1.41	1.34	1.54	1.63	1.46

Fuente: Elaboración propia a partir del sistema de gestión de control de Universidad de Concepción.

4.4. Resultados generales en los certámenes

Considerando los resultados obtenidos en cada uno de los certámenes, se tomó la media de las calificaciones obtenidas por las 8 secciones que actuaron como grupo control y se contrastaron con la media obtenida por los participantes del grupo intervenido en cada uno de ellos. Los resultados permiten observar un puntaje mayor en todas las evaluaciones en el grupo piloto en comparación con las demás secciones. Las diferencias fluctuaron en función del certamen entre 0,26 décimas y 1,18 décimas de nota.

Tabla 4. Diferencia de resultados entre el grupo de intervención y las secciones 1 a 8

Secciones	Media Certamen 1	Media Certamen 2	Media Certamen 3	Media Nota Final
Intervención (sección 9)	5.01	4.21	3.57	3.65
Control (secciones 1 a 8)	4.37	3.03	3.31	3.43
Diferencia de puntaje a favor de la intervención	+.64	+.18	+.26	+.22

4.5. Resultados de evaluación de satisfacción

Se realizó una evaluación de la satisfacción de los estudiantes con distintos aspectos vinculados a la mejora implementada, a partir de la consulta a estudiantes pertenecientes del grupo de intervención y estudiantes pertenecientes a las demás secciones. La evaluación de cada criterio fue en una escala de 1 a 5 puntos, donde 5 implicaba mayor nivel de satisfacción.

En las tablas 5, 6 y 7 se observan las medias por cada criterio evaluado, separando las ponderaciones realizadas por los estudiantes del grupo de intervención y los pertenecientes a las demás secciones. Los resultados muestran diferencias estadísticamente significativas en la satisfacción con los ayudantes ($p < .001$), con los prácticos ($p < .001$) y los métodos de evaluación ($p < .001$) y en la satisfacción con el propio rendimiento ($p < .05$), siendo el grupo de intervención superior en todas las áreas referidas.

Tabla 5. Medias de evaluación de satisfacción respecto a atributos del profesor

Secciones	Conocimiento sobre matemáticas	Forma de enseñar	Disposición	Global
Intervención (sección 9)	4.36	4.07	4.26	4.23
Control (secciones 1 a 8)	4.31	4.13	4.16	4.20

Tabla 6. Evaluación de satisfacción respecto a atributos del ayudante

Secciones	Conocimiento sobre matemáticas	Forma de enseñar	Global
Intervención (sección 9)	4.05	3.94	4.15
Control (secciones 1 a 8)	3.99	3.62	4.07

Tabla 7. Evaluación de satisfacción respecto a características de la asignatura

Secciones	Con los prácticos	Métodos de evaluación	Con la asignatura
Intervención (sección 9)	3.51	2.92	3.65
Control (secciones 1 a 8)	3.27	2.68	3.59

5. DISCUSIÓN

En general, los resultados obtenidos son congruentes con la literatura revisada respecto a los enfoques de enseñanza y evaluación con los cuales se rediseñó la asignatura. El desempeño en los estudiantes del grupo piloto resultó mayor en relación con los grupos de comparación, lo que significa un aumento derivado de los cambios realizados acorde a los encontrados en investigaciones previas (Bakr et al., 2016; Mason et al., 2013).

Estos cambios positivos estarían asociados a la mejor planificación de los recursos y esfuerzos de la clase con efectos sobre dos aspectos centrales: el tiempo y focalización de docentes y ayudantes, y la flexibilidad de organización de los tiempos de los estudiantes.

En cuanto al primer punto, se observa que la optimización del tiempo del docente y los ayudantes permitió focalizarlos en detectar y responder a las necesidades de los estudiantes de manera más precisa. Producto de la optimización del tiempo, el trabajo de clase pudo ser más personalizado y focalizado en los aspectos más complejos y críticos de la enseñanza. La flexibilidad entregada por la disponibilidad permanente del material favoreció que los estudiantes organizaran su estudio acorde a sus propias necesidades y estilos de trabajo. Además, permitió que revisaran más de una vez el material y repasaran las veces necesarias.

Por otra parte, la aplicación de la evaluación auténtica tuvo un efecto relevante sobre dos aspectos centrales: un efecto formativo y un efecto motivacional. El efecto formativo se produjo al entregar comentarios y correcciones inmediatas a cada respuesta alcanzada por los estudiantes en las evaluaciones tipo test. El efecto motivacional apareció al entregar una medida instantánea del desempeño y al presionar a los estudiantes a un estudio permanente debido a la presencia de evaluaciones constantes.

En el análisis de cada certamen, se observa que, para la primera evaluación, los resultados obtenidos por el grupo piloto son los mejores. Cabe destacar que los contenidos de este certamen corresponden, en general, a los mismos ya vistos durante la enseñanza media, por lo cual los estudiantes ya se encuentran familiarizados con estos.

En el segundo certamen, frente a los nuevos contenidos de carácter plenamente universitario, se observa un decaimiento en las calificaciones de todas las secciones excepto en el grupo piloto, que disminuye sus calificaciones en menor proporción que las demás secciones. Así, la intervención realizada parece tener un efecto protector en el grupo intervenido, disminuyendo la caída de las calificaciones de los estudiantes al enfrentarse al contenido más nuevo y complejo.

Una hipótesis para este efecto puede tener relación con la adquisición de herramientas para un mejor abordaje de la matemática, por ejemplo, mejor sistema de estudio y de comprensión. Otra hipótesis no excluyente con la anterior es la posible existencia de una mayor autoeficacia frente a la matemática en el grupo intervenido que favorezca una mejor disposición de trabajo frente a la adversidad y complejidad. En cualquier caso, este efecto debe ser estudiado específicamente para discriminar su naturaleza.

En el certamen 3, la sección intervenida mantiene mejores resultados que las otras 7 secciones con excepción de la sección 1. Se observa una reducción importante de la diferencia entre las calificaciones – antes amplia – entre la sección intervenida y las demás secciones, además de la caída del rendimiento en todos los grupos. Una hipótesis para explicar la combinación de estos efectos puede ser el componente motivacional.

En cuanto a la satisfacción en el grupo de intervención con respecto de la asignatura, se pudo notar una mejor satisfacción general de los estudiantes del grupo de intervención en contraste con sus pares.

Se encontraron diferencias significativas en cuanto a la satisfacción con los métodos de evaluación y con los prácticos a favor del grupo de intervención, lo que señala una buena recepción de los cambios realizados, probablemente debido a la retroalimentación inmediata y el repaso de los errores más comunes dentro de las clases. El hecho de contar con una mayor cantidad de evaluaciones parece no haber influido negativamente en la percepción de la satisfacción con la asignatura.

Por otro lado, el único aspecto en el que la metodología tradicional logró mayor satisfacción promedio fue en la forma de enseñar del profesor (sin diferencias estadísticamente significativas). Esto puede deberse a varios factores, dentro de los cuales el cambio inusual en la estructura de las clases,

más enfocado en la resolución de problemas que en la entrega de contenidos teóricos, parece el más explicativo. Los estudiantes no acostumbrados a la revisión del material en la plataforma previo a las clases teóricas pueden haberse sentido perdidos al asistir sin los contenidos necesarios para comprender los problemas planteados.

En términos generales, es posible indicar que existe evidencia a favor de la utilidad de diseños que incorporen enfoques de aula invertida y aprendizaje significativo sobre el aprendizaje de los estudiantes y su satisfacción con la asignatura. Sin embargo, será parte de investigación futura refinar y detallar las formas en cómo las trayectorias de rendimiento durante el curso se ven afectadas por potenciales variables intervinientes que no fueron foco de esta investigación.

6. CONCLUSIONES

A partir de los resultados obtenidos, es posible extraer las siguientes conclusiones:

- 1) La utilidad del cambio metodológico sobre el desempeño de los estudiantes. En las tres evaluaciones realizadas durante el curso, la media de calificaciones de los estudiantes del grupo piloto es superior a la media del conjunto de los grupos de comparación. En general, el grupo piloto tuvo un mejor desempeño en cada una de las evaluaciones respecto de cada uno de los demás grupos, pero no en todas las secciones esta diferencia fue estadísticamente significativa. Solo en el caso del grupo 1 para la evaluación 3 la media de calificaciones alcanza niveles superiores al grupo piloto.
- 2) La proporción de aumento en las calificaciones de los estudiantes no solo es indicador de un mayor aprendizaje de los mismos gracias a las modificaciones efectuadas, sino también, de una mayor tasa de aprobación. En el caso concreto de la sección piloto, se observa un 8,61% más de aprobación en comparación con el porcentaje de aprobación de las demás secciones. Lo anterior es un resultado de magnitud importante considerando que, si los resultados son extrapolables a la población de estudiantes de primer año de ingeniería en la facultad en cuestión, implicaría la aprobación de aproximadamente 95 estudiantes más cada año, tomando como referencia 1100 estudiantes cursando la asignatura cada semestre.

- 3) Por otro lado, los resultados obtenidos en la evaluación de satisfacción demuestran una satisfacción general superior en el grupo de intervención, principalmente en los métodos de evaluación aplicados, la satisfacción con los prácticos y la satisfacción con los ayudantes, en los cuales hay una diferencia estadísticamente significativa. Solo en la satisfacción con la forma de enseñar del profesor las secciones de comparación lograron un puntaje levemente mayor que el del grupo piloto.

REFERENCIAS BIBLIOGRÁFICAS

- Acuña, C. (2012). *Acceso y deserción en la educación superior: caso aplicado a Chile* (Tesis de Magíster). Facultad de Ciencias Físicas y Matemáticas, Universidad de Chile, Santiago.
- Alvarez, B. (2012). Flipping the classroom: homework in class, lessons at home. *Education Digest*, 77(8), 18-21.
- Argente, E., García-Fornes, A., y Espinosa, A. (2016). Aplicando la metodología Flipped-Teaching en el Grado de Ingeniería Informática: una experiencia práctica. En M. Torres y J. Cañadas (eds.), *Actas de las Jornadas sobre la Enseñanza Universitaria de la Informática* (pp. 221-229). Almería, España: Editorial Universidad de Almería.
- Ashford-Rowe, K., Herrington, J. y Brown, C. (2014). Establishing the critical elements that determine authentic assessment. *Assessment and Evaluation in Higher Education*, 39(2), 205-222. <https://doi.org/10.1080/02602938.2013.819566>
- Ato, M., López, J. J. y Benavente, A. (2013). Un sistema de clasificación de los diseños de investigación en psicología. *Anales de Psicología*, 29(3), 1038-1059. <https://doi.org/10.6018/analesps.29.3.178511>
- Bakr, M., Massey, W. y Massa, H. (2016). Flipping a Dental Anatomy Course: A Retrospective Study Over Four Years. *Education Research International*, 2016, 1-9. <https://doi.org/10.1155/2016/7097398>
- Bergmann, J. y Sams, A. (2012). *Flip your classroom: Reach every student in every class every day*. Washington DC: ISTE ASCD.
- Castillo, J. y Cabezas, G. (2010). Caracterización de jóvenes primera generación en educación superior. Nuevas trayectorias hacia la equidad educativa. *Calidad en la Educación*, 44(32). <https://doi.org/10.31619/caledu.n32.151>
- Cox, C. (2012). Política y políticas educacionales en Chile 1990-2010. *Revista Uruguaya de Ciencia Política*, 21(1), 13-42.
- Demski, J. (2013). *Managerial uses of accounting information*. Boston: Kluwer Academic Publishers.

- Dove, A. (2013). Students' Perceptions of Learning in a Flipped Statistics Class. (R. McBride y M. Searson, eds.), *Proceedings of SITE 2013-Society for Information Technology & Teacher Education International Conference* (pp. 393-398). New Orleans: Association for the Advancement of Computing in Education (AACE).
- Du, S.C., Fu, Z.T. y Wang, Y. (2014). The Flipped Classroom: Advantages and Challenges. En Wang (presidencia). *International Conference on Economic and Trade Cooperation* (pp. 17-20). International Conference on Economic Management and Trade Cooperation, Xi'an, China.
- Elliott, N. y Higgins, A. (2005). Self and peer assessment - Does it make a difference to student group work?. *Nurse Education in Practice*, 5(1), 40-48. <https://doi.org/10.1016/j.nepr.2004.03.004>
- Espinoza, O. (2013). *Equidad e inclusividad en el sistema de educación superior en Chile*. Santiago de Chile: UCINF.
- Espinoza, O. y González, E. (2011). *Acceso a instituciones de educación superior públicas y privadas: El caso de Chile*. Santiago de Chile: Universidad Diego Portales.
- Espinoza, O. y González, E. (2015). Equidad en el Sistema de Educación Superior de Chile: Acceso, Permanencia, Desempeño y Resultados. En A. Bernasconi (ed.), *Educación superior en Chile: Transformación, Desarrollo y Crisis* (pp. 517-580). Santiago, Chile: Ediciones de la Universidad Católica de Chile.
- Estes, M., Ingram, R. y Liu, J. (2014). A review of flipped classroom research, practice, and technologies. *International HETL Review*, 4(7), 1-8.
- Field, R., Duffy, J. y Huggins, A. (2013). Supporting transition to law school and student well-being: The role of professional legal identity. *The International Journal of the First Year in Higher Education*, 4(2), 15-25. <https://doi.org/10.5204/intjfyhe.v4i2.167>
- Flanagan, A. (2017). Experiences of first-generation students in Chilean universities: Realities and challenges. *Revista de la Educación Superior*, 46(183), 87-104. <https://doi.org/10.1016/j.resu.2017.06.003>
- Gilboy, M., Heinerichs, S. y Pazzaglia, G. (2015). Enhancing student engagement using the flipped classroom. *Journal of Nutrition Education and Behavior*, 47(1), 109-114. <https://doi.org/10.1016/j.jneb.2014.08.008>
- Gómez, V., Rosales, S. y Robles, A. (2019). The rubric as a learning strategy in research methodology in undergraduate medicine. *Investigación en Educación Médica*, 29(1), 30-35. <https://doi.org/10.22201/facmed.20075057e.2019.29.1769>
- James, L. y Casidy, R. (2018). Authentic assessment in business education: its effects on student satisfaction and promoting behaviour. *Studies in Higher Education*, 43(3), 401-415. <https://doi.org/10.1080/03075079.2016.1165659>

- Jehangir, R. (2010). *Higher Education and First-Generation Students: Cultivating Community, Voice, and Place for the New Majority*. Nueva York: Palgrave MacMillan. <https://doi.org/10.1057/9780230114678>
- Karabulut-Ilgu, A., Jaramillo, N. y Jähren, C. (2018). A systematic review of research on the flipped learning method in engineering education. *British Journal of Educational Technology*, 49(3), 398-411. <https://doi.org/10.1111/bjet.12548>
- Koh, K. (2019). Authentic Assessment. En G. Nobilt (ed.), *Oxford Research Encyclopedia of Education*. Oxford, US: Oxford.
- Lingard, L., Schryer, C., Garwood, K. y Spafford, M. (2003). «Talking the talk»: School and workplace genre tension in clerkship case presentations. *Medical Education*, 37(7), 612-620. <https://doi.org/10.1046/j.1365-2923.2003.01553.x>
- Martínez, L. (2016). *Más allá de la calificación: Instrumentos de evaluación para evaluar el aprendizaje*. Concepción, Chile: Editorial Universidad de Concepción.
- Mason, G., Shuman, T. y Cook, K. (2013). Comparing the effectiveness of an inverted classroom to a traditional classroom in an upper-division engineering course. *IEEE Transactions on Education*, 56(4), 430-435. <https://doi.org/10.1109/TE.2013.2249066>
- OECD. (2013). *Revisión de Políticas Nacionales de Educación: El Aseguramiento de la Calidad en la Educación Superior en Chile 2013*. Santiago de Chile: OECD Publishing.
- Orellana, A. (22 de agosto de 2018). 46,3 por ciento de los nuevos estudiantes de pregrado son primera generación universitaria de sus familias. *Diario Universidad de Chile*. Recuperado de <http://www.uchile.cl/noticias/146295/463-de-nuevos-estudiantes-de-pregrado-es-la-generacion-universitaria>
- Rodríguez, C. y Jarpa, C. (2015). Predictive ability of high school grades on performance at the national university selection test: The Chilean case. *Aula Abierta*, 43, 61-68. <https://doi.org/10.1016/j.aula.2015.03.002>
- Sadler, D. (2005). Interpretations of criteria-based assessment and grading in higher education. *Assessment and Evaluation in Higher Education*, 30(2), 175-194. <https://doi.org/10.1080/0260293042000264262>
- Strayer, J. (2012). How learning in an inverted classroom influences cooperation, innovation and task orientation. *Learning Environments Research*, 15(2), 171-193. <https://doi.org/10.1007/s10984-012-9108-4>
- Tai, J., Ajjawi, R., Boud, D., Dawson, P. y Panadero, E. (2018). Developing evaluative judgement: enabling students to make decisions about the quality of work. *Higher Education*, 76(3), 467-481. <https://doi.org/10.1007/s10734-017-0220-3>

- Thomas, L. y Quinn, J. (2006). *First generation entry into higher education: an international study*. Nueva York: McGraw Hill Education.
- Turra, H., Carrasco, V., González, C., Sandoval, V. y Yáñez, S. (2019). Flipped classroom experiences and their impact on engineering students' attitudes towards university-level mathematics. *Higher Education Pedagogies*, 4(1), 136-155. <https://doi.org/10.1080/23752696.2019.1644963>
- Villarroel, V., Bloxham, S., Bruna, D., Bruna, C. y Herrera-Seda, C. (2018). Authentic assessment: creating a blueprint for course design. *Assessment and Evaluation in Higher Education*, 43(5), 840-854. <https://doi.org/10.1080/02602938.2017.1412396>
- Wiggins, G. (1989). A True Test: Toward More Authentic and Equitable Assessment. *Phi Delta Kappan*, 70(9), 703-713.
- Wiggins, G. (1990). The case for authentic assessment. *Practical Assessment, Research & Evaluation*, 2(2). Recuperado de <http://PAREonline.net/getvn.asp?v=2&n=2>