Universitat Autònoma de Barcelona - Universitat de València
ENSEÑANZA DE LAS CIENCIAS 44-1 (2026), 67-84
Investigaciones didácticas
https://doi.org/10.5565/rev/ensciencias.6437
ISSN (impreso): 0212-4521 / ISSN (digital): 2174-6486

© del autor

Desarrollando la argumentación con modelo de Toulmin y formato WSDC en el nivel superior

Developing Argumentation with the Toulmin Model and the WSDC Format at an Advanced Level

Dení Stincer Gómez
Universidad Autónoma del Estado de Morelos, Facultad de Psicología, México
dstincerg@gmail.com
ORCID: https://orcid.org/0000-0003-3884-6805

Nanette Archer Svenson
Centro de Investigación Educativa, Ciudad de Panamá, Panamá
nanette.svenson@gmail.com
ORCID: https://orcid.org/0000-0003-3837-5900

Norma Miller
Universidad Tecnológica de Panamá, Ciudad de Panamá
norma.miller@utp.ac.pa
ORCID: https://orcid.org/0000-0002-1995-4299

Jane Saldaña
Independiente, Chiriquí, Panamá
janecaminando@gmail.com
ORCID: https://orcid.org/0000-0002-8079-2129

RESUMEN • La presente investigación, con diseño pretest-postest, tuvo el objetivo de desarrollar habilidades argumentativas en estudiantes recién ingresados a la universidad utilizando con un grupo el modelo argumentativo de Toulmin (MAT) y con otro el formato de debate WSDC. Cada intervención tuvo ocho sesiones. Para medir la calidad de la argumentación se usó el cuestionario y rúbrica de Lin y Mintzes. El puntaje para el grupo MAT aumentó 3.4 puntos en el postest con relación al pretest, y reveló una mejora en las cuatro habilidades argumentativas: razones a favor de la postura asumida, consideración de contraargumentos, razones suplementarias o refutaciones, y aporte de pruebas. En el grupo WSDC el aumento fue de 2.8 puntos, mostró mejoras solo en dos de estas habilidades, pero con significatividad estadística. Un análisis cualitativo permitió acceder a otros criterios de calidad.

PALABRAS CLAVE: Alfabetización científica; Debate escolar; Discurso argumentativo académico; Formato WSDC; Modelo argumentativo de Toulmin.

ABSTRACT • The purpose of this study was to foster argumentation skills in newly admitted university students using the Toulmin argumentation model with one group and the WSDC debate format with another. Both treatments lasted for eight weeks. Lin and Mintzes' questionnaire and rubric were used to measure argumentation quality. The score on the post-test for the Toulmin model group was 3.4 points higher than on the pretest, and showed improvement in all four argumentation skills: reasons in favor of the adopted position, consideration of counterarguments, supplementary reasons and rebuttals, and providing supporting evidence. The score increase for the WSDC group was 2.8 points, showing improvements only in two skills, but both statistically significant. A qualitative analysis allowed access to other quality criteria.

KEYWORDS: Scientific literacy; School debate; Academic argumentative discourse; WSDC format; Toulmin argumentation model.

Recepción: marzo 2025 • Aceptación: septiembre 2025 • Publicación: marzo 2026

Stincer Gómez, D., Miller, N., Archer Svenson, N. y Saldaña, J. (2026). Desarrollando la argumentación con modelo de Toulmin y formato WSDC en el nivel superior. *Enseñanza de las Ciencias, 44*(1), 67-84.
https://doi.org/10.5565/rev/ensciencias.6437

INTRODUCCIÓN

En la actualidad, la alfabetización científica y el pensamiento crítico son indispensables para el ejercicio de una conciencia independiente e informada. Para lograrlo, las nuevas tendencias educativas proponen una formación significativa en ciencia y tecnología.

Una estrategia pedagógica promovida por los investigadores en educación para formar estas habilidades es la práctica de la argumentación usando modelos argumentativos de las ciencias (Rodríguez Ortiz et al., 2023; Álvarez y García, 2023). Varias razones justifican esta recomendación: 1) es la forma de comunicación con la que los científicos producen y justifican conocimientos válidos y plausibles, 2) muestra el carácter social y constructivo del conocimiento, y 3) moviliza recursos cognitivos y metacognitivos complejos como la capacidad de inferir, razonar lógicamente y evaluar evidencias (Nussbaun et al., 2008). De acuerdo con Jiménez-Alexaindre y Erduran (2007), otorga a los estudiantes modelos de razonamiento científico, promueve el escepticismo, el pensamiento crítico y ofrece elementos para decidir qué creer y qué no creer a través de evaluaciones exhaustivas de evidencias (Osborne, 2000; Erduran, 2007; Duschl, 2007; Sandoval y Millwood, 2007). También incrementa la autonomía de los individuos, conductas de vida emancipatorias, un alto compromiso social y actitudes críticas de las imposiciones provenientes del poder (Lopes Piris, 2020). Visibiliza falacias y aprende a distinguir entre contenidos científicos y pseudocientíficos. Expone a los estudiantes a explicaciones plurales (Duschl y Osborne, 2002; Nussbaun, 2020) y fortalece la toma de decisiones fundamentadas (Felton et al., 2015; Foyler et al., 2009; Osborne et al., 2016). Diversos países han incorporado la formación de esta habilidad en sus currículos escolares (Jiménez-Alexaindre y Erduran, 2007) con resultados favorables.

El auge de la práctica argumentativa en aulas de educación básica y media superior comienza a finales del siglo xx. Una vertiente investigativa se ha centrado en probar la efectividad de intervenciones con estos modelos en la identificación de los componentes principales de un buen argumento y dar mayor número de razones a favor de las posiciones asumidas (Oyarzún et al., 2020), debatir sobre dilemas éticos y sociocientíficos, respetar posiciones contrarias y reconocer el valor de la crítica, o para el alcance de consensos (Kuhn y Udell 2003; Van Eemeren y Grootendorst, 2004).

También han identificado obstáculos comunes en la formación de esta habilidad, como el predominio del pensamiento intuitivo sobre el pensamiento lógico (Evans, 2010; Kuhn, 2000), que impide la evaluación exhaustiva de evidencias, lo que provoca saltos inferenciales no sustentados en procesos racionales. Otra dificultad es la dependencia entre calidad argumentativa y conocimientos previos y creencias epistemológicas del estudiante (Noroozi, 2022). Según Baytelman et al. (2020), cuanto más amplios y sofisticados sean estos mejor será la calidad del argumento. Kuhn (2000) apunta a la dependencia de la calidad del argumento de la etapa evolutiva epistemológica de los estudiantes, que transitan desde concepciones ingenuas a concepciones relativistas y, finalmente, más evaluativas en correspondencia con su edad. Estos resultados son consistentes con los de Osborne et al. (2016), que muestran que el aprendizaje de la habilidad argumentativa recorre un camino de lo simple a lo complejo; los estudiantes primero aprenden a identificar afirmaciones, justificaciones y pruebas, luego a relacionar componentes de la argumentación y, por último, evaluar, comparar y criticar. La tarea de exponer y recibir críticas, la más compleja de lograr, pertenece a la última etapa evaluativa delineada por Kuhn (2000).

Respecto a los logros de su incorporación en el aula, Davison y Carson (2020) concluyen que luego de una intervención los estudiantes muestran mayor conciencia de la necesidad de justificar sus decisiones con pruebas científicas. Ibacache y Merino (2021) hallan que un alto porcentaje de estudiantes logra construir una conclusión con base en pruebas de la disciplina, bien de naturaleza sociopolítica, ética o ambiental. Demircioglu et al. (2023) reportan enriquecimiento de los componentes del pensamiento crítico, por ejemplo, un mayor uso de evidencias para evaluar argumentos, la predisposición a

buscar razones y desafiar creencias que provienen de la autoridad, opiniones más independientes del grupo y una tendencia a realizar análisis críticos de otros discursos.

Por otro lado, Felton et al. (2015) aportan resultados relacionados con el ambiente de discusión propio de la argumentación. Los estudiantes expuestos a modelos argumentativos centrados en lograr consensos integran mejor las ideas de sus compañeros que quienes argumentan bajo modelos centrados en persuadir y competir. Esta práctica, denominada por Alexander (2020) enseñanza dialógica mediada por un experto para llegar a consensos racionales, está ocupando la atención de quienes consideran que el docente debe tener un buen dominio de esta habilidad si desea enseñarla (Gómez-Zacarelli et al., 2024).

Con respecto a las intervenciones que utilizan modelos de argumentación, se reportan resultados significativos en la calidad de los argumentos orales y escritos que elaboran los estudiantes de niveles medio-superior y superior. Por ejemplo, Magalaes (2024) reporta una disminución de las ambigüedades, las inconsistencias y los defectos de la estructura del texto. Pappas (2024) reporta la presencia del total de los componentes del modelo y un incremento de los argumentos persuasivos.

En suma, las investigaciones sobre intervenciones para la formación y desarrollo de la argumentación han obtenido resultados alentadores.

Modelo argumentativo de Toulmin y formato de debate WSDC

Dos modelos utilizados para la enseñanza y práctica de la argumentación son el modelo argumentativo de Toulmin (MAT) y el formato de debate World Schools Debating Championships (WSDC).

El MAT enfatiza la construcción de argumentos orales y escritos que contengan estos componentes básicos: conclusión (*claim*), evidencias (*data*), justificaciones (*warrants*), respaldos (*backing*) y refutadores (*rebuttals*). Este modelo prioriza el aspecto lógico y virtudes como la coherencia, la pertinencia y la congruencia entre premisas y conclusiones. Stincer y Blum (2017) consideran que el MAT incorpora, implícitamente, componentes dialógicos y sociales de la dimensión pragma-dialéctica de la argumentación propuesta por Van Eemeren y Grootendorst (2004). Sin embargo, estos son más aplicables en el formato WSDC.

De acuerdo con Stincer y Blum (2017) el trabajo mental que se lleva a cabo con el MAT es el siguiente: a partir de un conjunto de hechos, evidencias y datos (primer componente), nuestra mente concluye o genera una afirmación (segundo componente), una construcción mental-lingüística de lo que contienen los datos que, en general, no se percibe ni directa ni inmediatamente en ellos. Una vez elaborada la afirmación, es necesario justificarla (tercer componente), explicitando las razones, las reglas y los principios que permiten comprenderla como un enunciado cierto o plausible. Los respaldos aluden a explicaciones suplementarias de las razones principales o a la mención de las fuentes que «apoyan» las justificaciones proporcionadas. Los refutadores proveen interpretaciones alternas de los datos que contradicen o coexisten con la propia. Para Stincer y Blum (2017), el MAT actúa como un heurístico mental, esto es, un conjunto de instrucciones mentales que solucionan problemas fundamentalmente epistémicos. Toulmin sistematizó mediante su modelo las instrucciones y reglas que la mente debe seguir para estructurar un argumento.

Por otro lado, el formato WSDC es una propuesta inglesa de debate que fomenta «una batalla de mentes» (WSDC, 2024) por medio de unos actores con roles predefinidos a favor y en contra de una «moción», y sujetos a mecanismos de regulación de la discusión.

Su implementación parte de la conformación de dos equipos de tres miembros. Un equipo elabora razones a favor de la moción (equipo afirmativo) y el otro las razones en contra (equipo negativo). Cada equipo tiene tiempos precisos para su participación, que regula un moderador. El discurso presenta un orden en el que un miembro del equipo positivo propone y uno del negativo replica; el proceso se

itera con los demás participantes. Al final, el equipo negativo ofrece una última réplica seguido de la réplica de cierre del equipo positivo. Cada equipo es entrenado sobre los tipos de moción que existen, las clases de pruebas y su carga, así como los esquemas más efectivos para argumentar y refutar. Unos jueces determinan la calidad argumentativa.

La organización del WSDC ha logrado aceptación internacional (actualmente abarca más de 70 naciones) y los participantes, todos preuniversitarios, debaten sobre temas políticos, económicos, sociocientíficos, ambientales, entre otros. La iniciativa busca fortalecer el debate racional, la libertad de expresión, pero con respeto a las reglas pragmáticas que garantizan el buen curso de una discusión confrontativa. A los estudiantes que participan les ha permitido adquirir amplios conocimientos en diversas disciplinas, un aumento de la confianza en sí mismos y la excelencia en el debate.

Imán (2017) y Murrahmah et al. (2020) proporcionan algunas muestras de estos logros, como la mejora de la expresión oral, la fluidez, el uso correcto de la gramática, un vocabulario especializado, además de la cortesía en la exposición de razones a favor y en contra.

Desde nuestra perspectiva, la diferencia fundamental entre ambos modelos con respecto a la enseñanza es la dimensión teórica de la argumentación a la que privilegia. En el MAT prevalece la dimensión lógica (Vega, 2003), interesada en que el individuo utilice el modelo como heurístico mental para la elaboración de un producto oral y/o escrito con relaciones de implicación válidas, plausibles, coherentes y pertinentes. Por el contrario, el formato WSDC prioriza la dimensión social y pragmadialéctica de la argumentación (Mamberti, 2023). Esta última fue promovida por van Eemeren y Grootendorst (2004), quienes defendieron que una situación confrontativa con reglas explícitas de actuación contribuye con la producción de afirmaciones válidas o plausibles.

A diferencia del modelo de Toulmin, el formato WSDC es ampliamente conocido en Panamá, país donde se llevó a cabo la presente investigación, pues lo practican los clubes de debate escolares bajo la tutela de la Asociación Panameña de Debate (ASPADE).

La presente investigación tuvo como objetivo evaluar el desarrollo de habilidades argumentativas en estudiantes de primer ingreso a la educación superior expuestos al modelo de Toulmin o al formato WSDC.

MÉTODO

El presente estudio tuvo un diseño mixto, con una vertiente cuasiexperimental de pretest-postest, con dos grupos (G1 y G2) y una vertiente cualitativa que analizó la calidad del contenido de los argumentos emitidos en el pretest y el postest en ambos grupos. El G1 fue expuesto al MAT, mientras que el G2 se expuso al formato WSDC.

Las intervenciones consistieron en dos talleres, uno aplicando el modelo MAT y otro el formato WSDC, diseñados por una de las investigadoras. A ella también le correspondió capacitar a los facilitadores, dos estudiantes avanzados de ingeniería del Centro Regional de la Universidad Tecnológica de Panamá. Ambos talleres tuvieron 8 sesiones de 80 minutos cada una, impartidas simultáneamente en horario matutino y vespertino para facilitar la asistencia de los participantes al grupo que les correspondía. Los talleres fueron gratuitos, la participación voluntaria y el reclutamiento se realizó a través de redes sociales y correo electrónico del Centro Regional de Chiriquí, así como visitas a los salones donde se impartía un curso propedéutico a los estudiantes recién admitidos. La convocatoria tuvo un período de dos semanas. Se aceptaron como participantes a los interesados en participar. La asignación al G1 o al G2 se hizo mediante un generador de números aleatorios.

La variable en la que se pretendía incidir con las intervenciones era la calidad del discurso argumentativo. Por ende, se evaluó el impacto que cada modelo tuvo en las habilidades argumentativas esperadas.

Instrumento

Para la determinación de la calidad del discurso argumentativo antes y después de los talleres se utilizó el Cuestionario de Habilidades Argumentativas (CHA) de Lin y Mintzes (2010). Este instrumento supone la propuesta de una situación problemática familiar (preferiblemente de carácter social) sobre la que los estudiantes deben tomar una postura y responder a cuatro preguntas que evalúan las siguientes habilidades argumentativas: 1) hacer afirmaciones y justificaciones, 2) construir contraargumentos, 3) generar argumentos de apoyo, incluidas las refutaciones, y 4) generar pruebas o evidencias.

La situación que se eligió, familiar para todos los estudiantes de la región y con un impacto socioeconómico y ambiental, fue esta:

> La idea de construir una carretera entre Boquete y Cerro Punta, pasando a través del Parque Nacional Volcán Barú y del Parque Internacional La Amistad, data de los años 80. Hasta ahora el proyecto no se ha concretado, pero recientemente ha resurgido y está siendo considerado por el gobierno nacional y el gobierno provincial.

El mismo escenario se utilizó en el pretest y postest. Las preguntas que contestaron fueron:

1. ¿Estás de acuerdo o en desacuerdo con la idea de construir una carretera entre Boquete y Cerro Punta que pase a través del Parque Nacional Volcán Barú y el Parque Internacional La Amistad?
2. Si alguien no está de acuerdo con las opiniones que tú expresaste en la respuesta a la pregunta 1, es probable que tenga algunas razones. ¿Cuáles crees que podrían ser sus razones?
3. ¿Cómo convencerías a alguien que no estuviera de acuerdo contigo, si esa persona te hubiera dado las razones que escribiste en la pregunta 2?
4. Si te pidieran que aportaras pruebas para apoyar tus propias opiniones en las preguntas 1 y 3 ¿cuáles podrían ser esas pruebas?

Para la evaluación de las respuestas se utilizó la rúbrica de Lin y Mintzes (2010). La tabla 1 detalla los criterios y asignación de puntos que propone esta rúbrica:

Tabla 1.
Rúbrica del CHA para la asignación de puntos

Pregunta	*Categoría de respuesta*	*Puntuación*
P1 Afirmación y justificaciones	Sin respuesta o justificación no válida	0
	Afirmación aceptable, pero sin justificación	1
	Afirmación aceptable y justificación válida	1 + 1
	Afirmación aceptable y *n* justificaciones válidas	*n* + 1
P2 Contraargumentos	Sin respuesta o justificación no válida	0
	Justificaciones válidas	1+ Un punto por justificación

Pregunta	*Categoría de respuesta*	*Puntuación*
P3 Respaldos	Sin respuesta o justificación no válida	0
	Justificación elaborada válida (de P1)	1+ Un punto por cada razón elaborada
	Justificación suplementaria válida (de P1)	1+ Un punto por cada razón suplementaria de P1
	Refutación de contraargumento (de P2)	2+ Dos puntos por cada refutación
P4 Evidencias	Sin evidencias o explicación suplementaria	0
	Evidencia válida	1+ Un punto por cada evidencia

Fuente: traducción de la rúbrica de Lin y Mintzes (2010) por parte de las autoras.

En la aplicación de la rúbrica participaron dos jueces expertos, miembros del equipo investigador. Se realizó una revisión por separado, luego se cotejaron y consensuaron los puntos asignados en ambos momentos a cada participante. Para entender las diferencias en el puntaje se realizó un análisis de contenido de las respuestas con el procedimiento de Cáceres (2003). Se identificaron categorías que reflejaran las diferencias sustanciales en las respuestas de un mismo sujeto indicativas de un cambio hacia una mejor calidad argumentativa o no. Si el cambio resultaba favorable o no impactaba en la calificación.

Los datos cuantificados se analizaron con Stata (v 9.2) y Excel. Para su descripción se calcularon medias y desviaciones estándar. Sin embargo, dado el reducido tamaño de la muestra, las pruebas de significancia se realizaron con estadísticos no paramétricos como la prueba de Wilcoxon. El nivel de significación fue de 0.05.

RESULTADOS Y DISCUSIÓN

Inicialmente se reclutó a 27 estudiantes, 13 mujeres y 14 varones, entre 18 y 20 años, a los que se les aplicó el pretest. De estos, 9 se retiraron por motivos personales. La muestra quedó en 18 estudiantes, 7 mujeres y 11 varones, quienes terminaron la intervención y contestaron el postest. El grupo expuesto al modelo de Toulmin (G1) contó con 10 estudiantes, mientras que el expuesto al formato WSDC (G2) contó con 8.

Análisis cuantitativo

Una comparación del promedio del puntaje total (la suma de los puntos obtenidos en las cuatro preguntas del instrumento) de cada grupo en el pretest arrojó los siguientes resultados: 10.9 (±2.5) puntos para el grupo G1 y 10.5 (±4.6) puntos para el G2. La diferencia resultó no significativa, por lo que se pudo validar la comparabilidad inicial de los grupos.

La misma comparación se realizó luego con el postest. Esta vez el G1 obtuvo 14.3 ± (5.4) puntos, mientras que el G2 obtuvo 13.3 ± (3.7) puntos. Aunque el puntaje total promedio claramente aumentó para ambos grupos, la prueba de Wilcoxon resultó nuevamente no significativa, indicando que

la mejora para ambos tratamientos fue similar. En cuanto a la mejora a lo interno de cada grupo, la diferencia resultó no significativa al 5 % para el G1, pero significativa para el G2 (véase tabla 2).

Tabla 2.
Diferencias entre grupos según el tratamiento

Grupo	*n*	*Puntaje total (promedio)*		*Cambio*	*Probabilidad (Wilcoxon bilateral)*	*Conclusión*
		Pretest	*Postest*			
G1 Toulmin	10	10.9 (± 2.5)	14.3 (± 5.4)	3.4 (± 5.2)	p = 0.07	NS
G2 WSDC	8	10.5 (± 4.6)	13.3 (± 3.7)	2.8 (± 2.1)	p = 0.008	Significativo

Fuente: elaboración propia. Abreviatura: NS, no significativo.

Seguidamente, se comparó el puntaje total por pregunta para cada grupo (figura 1). Para el grupo G1 se observa una tendencia a mejorar en cada una de las competencias argumentativas, mientras que para el G2 la mejora se concentra en las competencias asociadas a las preguntas 3 y 4. Estos resultados sugieren que ambos modelos contribuyen al desarrollo de las habilidades argumentativas, pero que pudieran incidir de manera diferente en ellas.

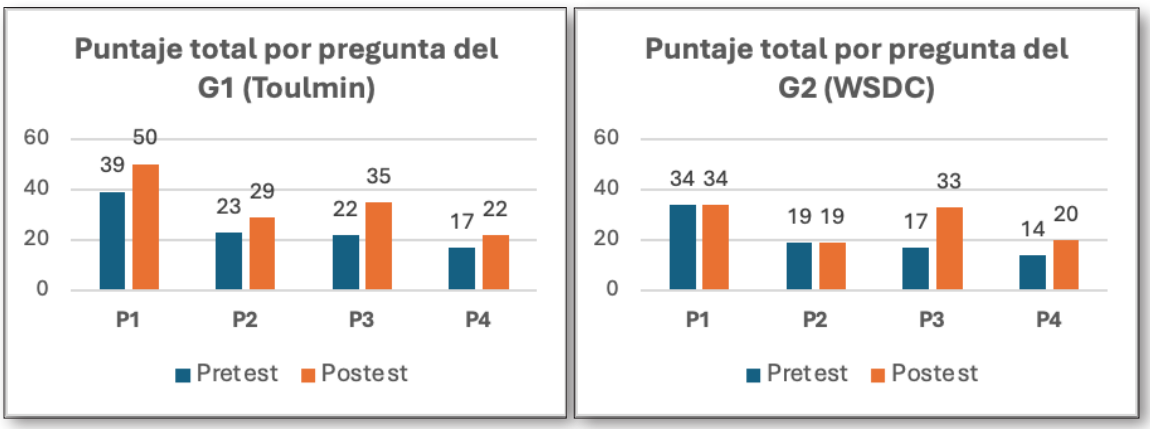


Fig. 1. Puntaje total por pregunta de cada grupo. *Fuente*: elaboración propia.

Las puntuaciones promedio por pregunta para cada grupo se presentan en la tabla 3, con una visible mejora en todas ellas. Al igual que con el puntaje total, no se detectaron variaciones significativas a nivel de pregunta entre los grupos en el pretest ni en el postest. Los grupos empezaron siendo equivalentes a nivel de pregunta y, aunque ambos mejoraron, se mantuvieron equivalentes después del tratamiento. Dentro de un mismo grupo, sin embargo, la mejora llegó a ser estadísticamente significativa para las preguntas 3 (argumentos de apoyo y refutación) y 4 (aportación de evidencias) en el G2.

Tabla 3.
Puntaje promedio por pregunta para cada grupo

Grupo	*Pregunta*	*Pretest*	*Postest*	*Probabilidad*	*Conclusión*
G1	P1	4.2 (± 1.1)	5.4 (± 2.8)	p = 0.296	NS
	P2	2.5 (± 0.8)	3.1 (± 2.0)	p = 0.351	NS
	P3	2.4 (± 1.2)	3.5 (± 1.8)	p = 0.264	NS
	P4	1.8 (± 1.0)	2.3 (± 1.3)	p = 0.272	NS
G2	P1	4.3 (± 1.7)	4.3 (± 1.4)	p = 0.825	NS
	P2	2.4 (± 0.9)	2.4 (± 0.5)	p = 0.682	NS
	P3	2.1 (± 1.4)	4.1 (± 1.2)	p = 0.011	Significativo
	P4	1.8 (± 1.6)	2.5 (± 1.4)	p = 0.027	Significativo

Fuente: elaboración propia. Abreviatura: NS, no significativo.

La mejora de la calidad de la argumentación en ambos grupos puede atribuirse a la naturaleza de las intervenciones. La intervención centrada en la aplicación del MAT hizo énfasis en la construcción lógica individual de un argumento. Cada estudiante construía su propio argumento apoyado en los componentes del MAT con la orientación del facilitador, que tenía que asegurarse de que las representaciones del estudiante se correspondiesen con la naturaleza del componente; más tarde se debatía en grupo. Esto puede explicar el incremento en cada una de las habilidades argumentativas. La intervención del G2 se centraba en debatir sobre un tema con los estudiantes organizados en dos «bandos», proponentes y oponentes. Sus intervenciones tenían el objetivo fundamental de proporcionar razones y pruebas a favor y en contra de la propuesta asignada y hacerlo dentro de un tiempo limitado. La propia construcción de la conclusión en correspondencia con los datos, su justificación y refutación a nivel individual no eran tareas en este grupo. Esto puede explicar el poco incremento en las dos primeras habilidades.

En cuanto a las diferencias por variables sociodemográficas, se contempló el tipo de escuela media (oficial o particular) de donde provenían los estudiantes, así como su sexo. El análisis por tipo de escuela no se pudo realizar porque la muestra final solo contaba con dos estudiantes de escuelas particulares. La variable sexo no reveló ninguna diferencia significativa entre mujeres y hombres.

Análisis cualitativo

Un análisis cualitativo del contenido de cada respuesta permite apreciar las manifestaciones y la evolución de la habilidad argumentativa lograda por los estudiantes. Se parte de considerar el cambio en el puntaje por pregunta y su dirección. Esto da lugar a tres categorías: 1) cambio positivo, 2) cambio negativo y 3) cambio nulo. La tabla 4 resume la distribución de estos cambios.

Tabla 4.
Distribución de cambios por pregunta

Cambio entre pretest y postest	*Pregunta 1*	*Pregunta 2*	*Pregunta 3*	*Pregunta 4*
Positivo	9 (50.0 %)	6 (33.3 %)	12 (66.7 %)	10 (55.6 %)
Nulo	4 (22.2 %)	7 (38.9 %)	4 (22.2 %)	6 (33.3 %)
Negativo	5 (27.8 %)	5 (27.8 %)	2 (11.1 %)	2 (11.1 %)

Fuente: elaboración propia.

Se esperaba un predominio de cambios positivos porque implicaba, dependiendo de la pregunta, un aumento del número de razones, contraargumentos, razones suplementarias y pruebas. Sin embargo, un análisis del contenido en el que el cambio fue negativo o nulo apunta a que no siempre un incremento del puntaje indicaba una mejor calidad argumentativa. En ocasiones, conservando el mismo valor o incluso disminuyéndolo se apreciaba una mejor calidad de contenido con respecto a la respuesta del pretest.

Para ilustrarlo se presentan ejemplos de respuestas con su respectivo análisis de contenido.

Análisis de contenido: pregunta 1

Según la rúbrica de Lin y Mintzes (2010), en la pregunta 1 se esperaba que el estudiante declarara su acuerdo o desacuerdo con la situación planteada y además proporcionara razones que justificaran su decisión. Si el estudiante declaraba una postura a favor o en contra se le asignaba un punto, y por cada razón que diera para justificar su posición también se le otorgaba un punto (véase tabla 1).

Como se aprecia en la tabla 4, entre el pretest y postest el 50 % de los estudiantes incrementaron las razones que respaldaban su posición. Es un dato positivo a favor del propósito de esta investigación, y por ello inferimos que la intervención desarrolló en los estudiantes la capacidad de búsqueda y asociación de razones convincentes.

Veamos un ejemplo (participante 10):

> **P10 (pretest)** No estaría de acuerdo. Una carretera como la descrita representa una ventaja turística importante ya que facilita a los visitantes conocer y disfrutar de la naturaleza de Boquete y Cerro Punta, disminuiría los costos en combustible (…), económicamente es una propuesta atractiva; sin embargo; ecológicamente las consecuencias no serán las mejores. (…) implica la pérdida de hábitats de la flora y fauna, aumentaría los niveles de contaminación acústica, del aire y también de los recursos acuáticos. (…) una inversión de capital tan grande como la que supondría la construcción de dicha carretera debería destinarse a la conservación de los bosques de estas zonas.
>
> **P10 (postest)** Yo no estaría de acuerdo con el desarrollo de esta construcción, esto debido a las siguientes razones: Daño a los hábitats: la construcción de dicha infraestructura requeriría de la tala de los bosques de las zonas en cuestión. Dicha actividad junto con la utilización de la maquinaria pesada o productos químicos (estabilización del terreno) que se puedan utilizar debilitarían la salud y el estado que los hábitats poseen. (…). Afectaciones a la flora y fauna: (…) debido a que la ausencia de nutrientes sólo saludables entre otros podría ocasionar que la vegetación enferme y desaparece, lo que a su vez ocasionaría que los animales que dependen de estas migran a otras áreas e incluso podrían correr el riesgo de extinción. aumento de las temperaturas: los automóviles producen grandes cantidades de dióxido de carbono, (…) y contribuye al efecto invernadero, y dependiendo del número de usuarios de esta carretera las emisiones anuales de CO_2 (…) (…) los niveles de infiltración de agua de lluvia disminuirían lo que aumente el riesgo de inundaciones. La pérdida de hábitat y la flora y fauna afectaría en la actividad turística de la región (…)

El ejemplo muestra un incremento significativo en las razones a favor de su posición, de 6 en el pretest a 11 en el postest. No solo las menciona, sino que las explica. Esto último se asocia a lo que Toulmin (1958) llama *respaldos* y Lin y Mintzes (2010) denominan *razones suplementarias*. El ejemplo representa lo idealmente esperado con las intervenciones. La mayor parte de los estudiantes, sin embargo, aumentó una o dos razones solamente.

En el caso de los estudiantes con cambios negativos, no se puede afirmar que la calidad de la respuesta haya sido precaria. Por ejemplo, aunque la participante 5 disminuye una razón, en el postest

muestra una implicación personal mayor, con elementos estéticos y morales que en el pretest aparecen con vaguedad. Veamos los extractos:

> **P5 (pretest)** Yo estoy en desacuerdo con la idea de construir una carretera de Boquete a Cerro Punta. El motivo principal de mi respuesta se basa en que tanto Boquete como Cerro Punta son lugares turísticos de nuestra provincia. (…) La idea de unir Boquete y Cerro Punta me gustaría, ya que como expliqué sería más fácil de transportarse y los dos lugares serían más visitados gracias a esta facilidad. (…) Además de que se conocería más a fondo la vegetación y naturaleza del Parque La Amistad y se apreciaría la bella fauna y flora del Parque Nacional Volcán Barú.
> **P5 (postest)** Lo primero que tengo en mente es la posibilidad de crear un acceso entre estos lugares tan importantes en nuestra provincia chiricana. Considerados puntos estratégicos de la provincia en sus diferentes aspectos, volcán (Cerro Punta), tierra de cultivos y parcelas, producción de tierra altas mientras que Boquete es lugar de turismo y bellos paisajes. Parque Nacional La Amistad es un lugar protegido por su flora, etc. todos los lugares antes mencionados tienen algo en común: la naturaleza que se encuentra inmersa en cada escenario (…). Una carretera que pasa a través de este tesoro verde tan preciado sería cambiar la armonía que ya poseen, hacer esa construcción hace cambios que todos y cada persona deberíamos evitar.

En la primera respuesta, reconoce que el transporte mejoraría, pero afectaría a la belleza del lugar. En la segunda respuesta continúa con este segundo argumento, pero agrega que «cada persona debería evitar», es decir, apela al compromiso moral de conservar la naturaleza. Decir «que pasa a través de este tesoro verde tan preciado sería cambiar la armonía que ya poseen» es aludir a argumentos afectivos-persuasivos, que en este tipo de dilema son válidos (Perelman y Olbrecht Tyteca, 2004). El argumento se enriquece con la introducción de estos elementos, ausentes anteriormente.

En esta pregunta también se manifiesta un «aparente declive» argumentativo, que se interpretó desde distintas perspectivas: 1) desmotivación al responder, 2) mayor cautela al argumentar, y 3) mayor tendencia a la síntesis. Hubo respuestas que presentaron en el pretest un mayor número de razones y en el postest un número menor, pero con poder de síntesis. Pereda (1994) considera la síntesis como una cualidad positiva de la argumentación, relacionada con la pertinencia de las razones, el poder «ir al grano», sin merodeos.

Lo que intentamos resaltar con este ejemplo es que los cambios negativos o nulos en la puntuación no necesariamente denotan una merma en la calidad argumentativa. Pueden aparecer nuevos elementos, como apreciaciones estéticas y morales o una mayor síntesis, no considerados explícitamente en la rúbrica de Lin y Mintzes (2010), pero válidos al argumentar sobre temas sociopolíticos (Dawson y Carson, 2020).

Otro aspecto que resaltar de la pregunta 1 es la diferencia en su tratamiento. La gráfica 1 muestra que el G2 mantuvo el mismo puntaje total en ambos cuestionarios, mientras que en el G1 hubo un incremento en la puntuación del postest con respecto al pretest. Esto puede significar que el MAT fortaleció la búsqueda y asociación de razones para apoyar la posición sostenida, en contraste con el formato WSDC, más centrado en el debate, el diálogo y el convencimiento.

Análisis de contenido: pregunta 2

En la pregunta 2, que solicitaba identificar los contraargumentos que pudieran surgir ante su afirmación, se observó un predominio de cambio nulo, y solo seis participantes aumentaron su puntuación. Sin embargo, consideramos que las respuestas no decaían en la calidad por la ausencia de uno o más contraargumentos adicionales. En ocasiones, como muestra el siguiente ejemplo, se aportaron contra-argumentos sustancialmente diferentes:

P4 (**pretest**) Algunas de las razones por la cual podrían estar en desacuerdo es que no son conscientes de la distancia que tienen que recorrer algunas personas para llegar a sus trabajos, sacrificando horas de sueño o comidas del día para así poder llegar a tiempo (…).
P4 (**postest**) El hecho del estado actual del país y la provincia no da para implementar la idea de construir una carretera de tal magnitud, cuando los mayores problemas actual exigen otro tipo de propuestas que ayuden al bienestar de comunidades. Teniendo en cuenta la deuda que el país lleva consigo es dudable que tal proyecto sea culminado.

En su primera respuesta aludió a que las personas no estarían de acuerdo con él/ella por desconocimiento, porque no son conscientes de las dificultades de otras personas para transportarse de un lugar a otro, mientras que en el postest apela a un contraargumento de naturaleza económica, el déficit presupuestario de su país para llevar a cabo un proyecto de tal envergadura. Se conserva el mismo número de razones, pero diferentes entre sí y ambas plausibles argumentativamente hablando. Este es un aspecto de calidad que no se toma en cuenta en los criterios de evaluación de la rúbrica y que parece implicar un puntaje desfavorable; desde nuestra perspectiva no lo es, pues ambos son plausibles.

En el próximo ejemplo también se mantiene la misma puntuación, pero no podría considerarse una disminución de la calidad argumentativa. La persona apela a los mismos contraargumentos, lo que podría indicar que estos son los que considera relevantes, y sostienen su posición inicial.

P6 (**pretest**) Razones por las que no estarían de acuerdo: La razón principal por la que creo que estén en desacuerdo es por toda la flora y fauna que saldría afectada por este proyecto. No soy conocedora de que tan grandes sean los parques por la que estaría supuesta a pasar la carretera, pero considero que si es una zona grande, afectaría de manera gigantesca al medio ambiente, ya que habría que talar kilómetros y kilómetros de vegetación lo que a su vez traería el desplazamiento y/o muerte de cientos de animales.
P6 (**postest**) La principal razón por la que una persona estaría en desacuerdo con la construcción de esta carretera sería por la gran perdida de ecosistema que habría. Para construcciones de este tamaño es necesario talar miles de kilómetros, lo cual implica una pérdida tanto de flora como de fauna.

Como ocurre con la pregunta 1, en la pregunta 2 tampoco hubo diferencia entre el puntaje total del pretest y postest para el G2, pero sí para el G1. Esto sugiere que el modelo de Toulmin estimula una mayor búsqueda tanto de las razones que respaldan una posición asumida como de los argumentos en su contra, un indicador importante del pensamiento crítico.

Quizás el menor avance en el desarrollo de esta habilidad argumentativa para ambos tratamientos pudiera deberse al escaso trabajo psicodidáctico enfocado en la disminución del compromiso afectivo acrítico que solemos asumir ante nuestras propias creencias. En el terreno argumentativo, esta condición se considera un obstáculo epistemológico por la imposibilidad que trae consigo de valorar la plausibilidad de otros puntos de vista. Es decir, en la formación de creencias y asunción de afirmaciones propias el componente afectivo está presente, mediado por nuestros gustos, preferencias e intereses personales que la sostienen (Vega, 2003; Walton, 1992). El movimiento de las creencias que sostienen nuestras percepciones del mundo genera un conflicto psíquico y malestar, por lo que tienden a evitarse (Polo et al., 2015). Una de las cualidades de la argumentación científica y la asunción de posiciones filosóficas críticas y escépticas es que promueve un compromiso mayor con la construcción de conocimientos plausibles más que con la demostración de certezas propias. Es posible que esta cualidad no se lograse lo suficiente tras ocho sesiones de intervención.

Análisis de contenido: pregunta 3

En la pregunta 3 se esperaba que el estudiante respaldara sus razones con otras suplementarias o que refutara los contraargumentos.

En los resultados predominan los cambios positivos. Ante los contraargumentos, los participantes aludieron a razones explicativas (o suplementarias) que fortalecían su posición o rebatían el contraargumento. De acuerdo con el contenido de las respuestas, los estudiantes tendieron a esforzarse en un debate crítico a favor de sus razones y a contraargumentar las posiciones contrarias, en ocasiones reconociendo explícitamente su valor. Veamos algunos extractos:

> **P12 (pretest)** La manera en la que convencería a esa persona sería darle otro punto de vista basado en mi opinión, pero sin afectar mucho la de la otra persona, trataría de hacerle ver que, aunque tal vez sea una ventaja para esa persona tener esa carretera, muchos no le darán el uso para el que fue realizada esa construcción.
> **P12 (postest)** Construir una carretera que pase por estos parques será algo tanto bueno como malo para algunas personas, trataría de hacerle ver a esta persona que una carretera que pase por un parque podría ocasionar una contaminación acústica, podría ocasionar que especies que habitan en la zona se tengan que ir debido a factores como el ruido, la contaminación muchos otros factores que se verán a largo plazo con la construcción de esta carretera.

Se puede observar cómo en la primera respuesta la persona no se compromete con razones explicativas y mantiene una postura más bien ambigua, trata de «no afectar» a la otra persona, como si argumentar generara malestar en el otro. Se trata de una actitud frecuente ante la crítica, que impide el devenir de la argumentación y la utilidad de la crítica en la toma de decisiones. Una vez que esta persona participó en la intervención, su respuesta mejoró cualitativamente porque apeló a razones concretas, diversas y ajenas al temor de molestar al otro.

En general, las respuestas del postest siguieron esta línea, la de la presencia de razones más realistas y técnicas. Veamos el próximo extracto:

> **P1 (pretest)** Viendo el desacuerdo de una persona que defendería un recurso tan valioso como la naturaleza, le expresaría no solo el daño que se ocasionaría a corto plazo, sino, los beneficios que tendríamos a futuro, contando en que podríamos recuperar la mayoría de nuestra naturaleza en años próximos, (…) al mismo tiempo nos encargaríamos de ayudar a nuestro ambiente, cuidando de que la mayoría de árboles y tierra procesada, sea nuevamente un área verde y hermosa como era en el pasado antes de dicha construcción.
> **P1 (postest)** El punto de vista que me has planteado es de suma importancia para dicha construcción, pero no únicamente veas las desventajas del proyecto, mira al futuro y piensa en la cantidad de beneficios que tendremos, por lo menos, el aumento de la economía en estos pueblos. Esta nueva carretera hará que las personas que están cerca de Boquete deseen ir hacia Cerro Punta, primeramente, porque tendrán que viajar menos para llegar allí, tienes que tomar en cuenta que gastaran menos gasolina y podrán visitar dos sitios turísticos en un solo día, si se trata de personas con mucho trabajo o que desean ir de viaje nocturno con su familia. Por otra parte, al ser una vía de transporte que genere menos tiempo de viaje, las empresas desearán crear nuevos mercados en un lugar así, en donde la accesibilidad a el sitio sea fácil, no solo generar interés en personas del mismo sitio por tener una nueva oportunidad de trabajo, también en personas de exterior para visitar esa nueva atracción (…)

Resulta interesante destacar que el G2 obtuvo una diferencia significativa con respecto al G1. Esto puede explicarse por el predominio en el formato WSDC para debatir: dialogar en equipos. Esta si-

tuación es estimulante y competitiva, de manera que promueve la búsqueda de una mayor cantidad de argumentos a favor y en contra para convencer al juez y a los escuchas. La intervención con el modelo de Toulmin pone énfasis en la construcción de la estructura lógica argumental más que en la situación de debate con figuras y propósitos tan claros.

Análisis de contenido: pregunta 4

Esta pregunta exigía una de las cualidades fundamentales de la argumentación: dar evidencias a favor de las razones a las que se apelan. Por el contenido de las respuestas, la mayor parte de los estudiantes entendió que debían acudir a las *fuentes* de donde se originan las pruebas en vez de a la prueba en sí. Por ello, en la asignación de puntos se consideró que, cuando el estudiante propusiera que se deben exigir pruebas que provengan de investigaciones técnicas y científicas, se asignara la puntuación requerida, lo mismo que si proporcionaba la prueba en sí.

Desde un punto de vista cualitativo se observó que el nivel de precisión y claridad de las pruebas ofrecidas fue mayor en el postest que en el pretest, donde resultaron ambiguas o escasas. Veamos los siguientes ejemplos:

> **P8 (pretest)** Estudios de áreas naturales en donde haya carreteras enfocándose en cómo afecta los alrededores a lo largo de su trayectoria.
> **P8 (postest)** Para comprobar el impacto del humano en la naturaleza sólo debemos ver las calles de nuestra ciudad o de la carretera Interamericana, donde apreciamos en varios puntos como está la basura tirada asimismo en esta carretera principal (la Interamericana) han ocurrido atropellos de animales principalmente de monos, que a pesar de que existen los puentes para que ellos crucen han sucedido estos lamentables incidentes.
>
> **P10 (pretest)** Presentaría las siguientes pruebas: 1. Medición de los niveles de contaminación ya existentes. 2) Evaluación de la salud de la flora y fauna de las zonas y 3) Estudio de impacto ambiental.
> **P10 (postest)** Para apoyar mi opinión lo primordial sería realizar un estudio de impacto ambiental para determinar si es viable la construcción para el ambiente. También se debería incluir un conteo de autos para evaluar a futuro las posibles cantidades de CO2 que se liberaría. Incluir también una evaluación técnica detallada que asegure si es o no seguro construir en estos terrenos que pueden ser muy frágiles y si la geología del terreno permite la construcción.

En ambos grupos se aprecia una mejoría al aportar pruebas a favor de sus posiciones (véanse gráfico 1 y tabla 3), atribuible al énfasis de la relevante presencia de pruebas en la argumentación en ambos modelos, sobre todo aquellas «que provengan de fuentes confiables». Tanto el MAT como el formato WSDC exigen acudir a pruebas y evidencias con un respaldo técnico y científico.

CONCLUSIONES

Las habilidades argumentativas apenas están presentes de manera sistemática en los currículos escolares de Latinoamérica[1], incluyendo Panamá. Los resultados obtenidos son alentadores porque la calidad de los argumentos emitidos después de las intervenciones mejoró. Esto significa que las intervenciones lograron que los estudiantes ofrecieran más razones a favor de su posición, identificaran contraargumentos, explicaran sus razones con elementos suplementarios y refutaran las razones contrarias, y que

1. Con excepción de Chile, que ha formalizado esta práctica como parte de sus materias curriculares.

proporcionaran más pruebas. La investigación evidencia que capacitar a los estudiantes en esta práctica intelectual compleja favorece su capacidad analítica y crítica, además de su expresión escrita.

De las diferencias obtenidas entre ambos grupos podemos concluir que los estudiantes expuestos al modelo de Toulmin tendieron a mejorar en todas las dimensiones del instrumento de Lin y Mintzes (2010), mientras que para los estudiantes expuestos al formato WSDC las mejoras se concentraron en las preguntas 3 y 4. Esto puede deberse al énfasis del MAT en la consistencia lógica entre datos, razones y evidencias que aúnan la identificación y construcción de los contraargumentos. El modelo WSDC enfatiza el aspecto dialógico, social, consensual y democrático del argumento.

Una mayor significatividad estadística pudo haberse logrado con una intervención de más tiempo; ocho sesiones no fueron suficientes, aun considerando las notorias deficiencias iniciales de los participantes. Reconocemos también las limitaciones relacionadas con el reducido tamaño de la muestra final (n = 18), producto de una deserción del 33 %, y la ausencia de un grupo control. Estas tres condiciones afectan a la validez externa del estudio, a la posibilidad de generalizar los resultados o a la determinación de que la intervención posea un carácter causal. Por lo que se reconoce que lo logrado forma parte de una tendencia a la mejora. No obstante, es por eso por lo que recurrimos a un análisis cualitativo de las respuestas, para entender mejor los cambios logrados por cada participante. Queda como importante recomendación ampliar la muestra y aplicar un diseño experimental más robusto.

Con respecto al instrumento de Lin y Mintzes (2010), concluimos que se enriquecería si la rúbrica tomara en cuenta no solo el aspecto cuantitativo, sino también el cualitativo. En ocasiones, la calidad del contenido de una razón puede tener un peso igual o mayor que la inclusión de más razones.

Los resultados tienen importantes implicaciones pedagógicas. Se constata empíricamente que una capacitación sistematizada apoyada en el MAT y en el formato WSDC desarrolla la habilidad de argumentar «lo más cercana posible» al proceder de los diálogos de contextos críticos. Esto contribuye al desarrollo de una ciudadanía crítica y capaz de tomar decisiones fundamentadas, tan deseado en las propuestas educativas actuales, sobre todo ante el amplio caudal de información al que están expuestos los jóvenes. Otra implicación relevante es la necesidad de capacitar a los docentes para el dominio conceptual y actitudinal de ambos modelos debido a su complejidad. Como bien señala Lopes Piris (2020), el poco desarrollo de esta habilidad en los docentes imposibilita su práctica real.

Es por ello por lo que los materiales y la experticia desarrollados para esta investigación se han vertido en una guía didáctica que pueda apoyar a docentes de todos los niveles e interesados en implementar actividades de argumentación en sus aulas.

Finalmente, algunas ideas concretas para las diversas actividades vinculadas con el fortalecimiento de las capacidades de argumentación y pensamiento crítico para estudiantes de varios niveles incluyen: la formulación de preguntas de investigación científica, con la correspondiente justificación para el estudio de estas preguntas; el diseño de campañas de mercadeo para un producto específico, con base en la evidencia; la redacción de artículos de opinión que citen cifras y datos sobre temas de actualidad para la prensa local; la preparación de ensayos que examinen opiniones opuestas sobre proyectos de ley que se encuentren actualmente en discusión; el desarrollo de planes de negocios utilizando datos y argumentación para justificar la inversión en ellos, entre muchas otras posibilidades. Una vez que los educadores y estudiantes tienen acceso a esta clase de herramientas y las incorporan con frecuencia al proceso del aprendizaje, las opciones son ilimitadas.

REFERENCIAS

Alexander, R. (2020). *A dialogic teaching companion*. Taylor & Francis.

Álvarez García, L. M. y García Martínez, A. (2023). Modelos de argumentación aplicados en la enseñanza de las ciencias: Una revisión sistemática. *Papeles*, *15*(29), *e1424.* https://doi.org/10.54104/papeles.v15n29.1424

Baytelman, A., Iordanou, K. y Constantinou, C. P. (2020). Epistemic beliefs and prior knowledge as predictors of the construction of different types of arguments on socioscientific issues. *J Res Sci Teach*, *57*(8), *1199 1227.* https://doi.org/10.1002/tea.21627

Cáceres, P. (2003). Análisis cualitativo de contenido: una alternativa metodológica alcanzable. *Psicoperspectivas*, *II*(1), 53-81.

Dawson, V. y Carson, K. (2020). Introducing argumentation about climate change socioscientific issues in a disadvantaged school. *Research in Science Education*, *50*(3), 863-883. https://doi.org/10.1007/s11165-018-9715-x

Demircioglu, T., Karakus, M. y Ucar, S. (2023). Developing students' critical thinking skills and argumentation abilities through augmented reality-based argumentation activities in science classes. *Science & Education*, 32, 1165-1195. https://doi.org/10.1007/s11191-022-00369-5

Driver, R., Newton, P. y Osborne, J. (2000). Establishing the norms of scientific argumentation in classrooms. *Science Education*, *84*, 287-312. https://doi.org/10.1002/(SICI)1098-237X(200005)84:3<287::AID-SCE1>3.0.CO;2-A

Duschl, R. y Osborne, J. (2002). Supporting and promoting argumentation discourse in science education. *Studies in Science Education*, *38*, 39-72. https://doi.org/10.1080/03057260208560187

Duschl, R. (2007). Quality argumentation and epistemic criteria. En S. Erduran y M. P. Jiménez-Alexaindre (Eds.), *Argumentation in science education: Perspective from classroom-based research* (pp. 159-177). Springer Science.

Erduran, S. (2007). Methodological foundations in the study of argumentation in science classrooms. En S. Erduran y M. P. Jiménez-Alexaindre (Eds.), *Argumentation in science education: Perspective from classroom-based research* (pp. 47-70). Springer Science.

Evans, J. S. B. T. (2010). Intuition and reasoning: A dual-process perspective. *Psychological Inquiry*, *21*(4), 313-326. https://doi.org/10.1080/1047840X.2010.521057

Felton, M. y Kuhn, D. (2001). The development of argumentative discourse skills. *Discourse Processes*, *32*, 135-153.

Felton, M., Garcia-Mila, M., Villarroel, C. y Gilabert, S. (2015). Arguing collaboratively: Argumentative discourse types and their potential for knowledge building. *Br J Educ Psychol*, *85*, 372-386. https://doi.org/10.1111/bjep.12078

García-Mila, M. y Andersen, C. (2007). Cognitive foundations of learning argumentation. En S. Erduran y M. P. Jiménez-Alexaindre (Eds.), *Argumentation in science education: Perspective from classroom-based research* (pp. 29-46). Springer Science.

Gómez Zaccarelli, F., Cándido Vendrasco, N. y Arriagada Jofré, V. (2024). Discusiones y argumentación en la enseñanza de las ciencias: Prácticas y desafíos docentes. *Enseñanza de las Ciencias*, *42*(2), 25-43. https://doi.org/10.5565/rev/ensciencias.5958

Hollihan, T. A. y Baaske, K. T. (2022). *Arguments and arguing: The products and process of human decision making* (4.ª ed.). Waveland Press.

Ibacache Plaza, M. y Merino Rubilar, C. (2021). Una propuesta de secuencia basada en el contexto, para la promoción de la argumentación científica en el aprendizaje de las reacciones químicas con estudiantes de educación media técnico profesional. *Revista Eureka sobre Enseñanza y Divulgación de las Ciencias*, *18*(1). https://www.redalyc.org/articulo.oa?id=92064232007

Iman, J. N. (2017). Debate instruction in EFL classroom: Impacts on the critical thinking and speaking skill. *International Journal of Instruction*, *10*(4), 87-108. https://doi.org/10.12973/iji.2017.1046a

Jiménez-Alexaindre, M. P. y Erduran, S. (2007). Argumentation in science education: An overview. En S. Erduran y M. P. Jiménez-Alexaindre (Eds.), *Argumentation in science education: Perspective from classroom-based research* (pp. 3-28). Springer Science.

Jiménez-Alexaindre, M. P. (2007). Designing argumentation learning environments. En S. Erduran y M. P. Jiménez-Alexaindre (Eds.), *Argumentation in science education: Perspective from classroom-based research* (pp. 91-116). Springer Science.

Kuhn, D. (1991). *The skills of argument*. Cambridge University Press.

Kuhn, D., Shaw, V. y Felton, M. (1997). Effects of dyadic interaction on argumentative reasoning. *Cognition and Instruction*, *15*, 287-315. https://doi.org/10.1207/s1532690xci1503_1

Kuhn, D. y Udell, W. (2003). The development of argument skills. *Child development*, *74*(5), 1245-1260.

Kuhn, D., Cheney, R. y Weinstock, M. (2000). The development of epistemological understanding. *Cognitive Development*, *15*(3), 309–328. https://doi.org/10.1016/S0885-2014(00)00030-7

Lin, S. y Mintzes, J. J. (2010). Learning argumentation skills through instruction in socioscientific issues: The effect of ability level. *International Journal of Science and Mathematics Education*, *8*, 993-1017.

Lopes Piris, E. (2020). (Im)posibilidades de enseñanza de la argumentación en la escuela. *Revista Iberoamericana de Argumentación*, 20, 30-56.

Noroozi, O. (2022). The role of students epistemic beliefs for their argumentation performance in higher education. *Innovations in Education and Teaching International*, *60*(4), 501-512. https://doi.org/10.1080/14703297.2022.2092188

Nussbaum, E. M. (2020). Critical integrative argumentation: Toward complexity in students' thinking. *Educational Psychologist*, *56*(1), 1-17. https://doi.org/10.1080/00461520.2020.1845173

Nussbaum, E. M., Sinatra, G. M. y Poliquin, A. (2008). Role of epistemic beliefs and scientific argumentation in science learning. *International Journal of Science Education*, *30*(15), 1977-1999. https://doi.org/10.1080/09500690701545919

Nurrahmah, N., Rukmini, D. y Yuliasri, I. (2020). The use of politeness strategies by Indonesian vs. Malaysian student debaters in the 2018 World Schools Debating Championship (WSDC). *English Education Journal*, 10(3), 282-291. https://doi.org/10.15294/eej.v10i1. 36310

Magalaes, A. L. (2024). Cómo desarrollar un argumento utilizando el modelo de Toulmin. *Revista H-Tec Humanidades e Tecnologia*, 10(2), 115-125.

Mamberti, J. (2023). Competitive debate formats analysed in light of the pragma-dialectical theory of argumentation's rules. RicercAzione. *Journal on Learning*, *Research and Innovation in Education*, 15(2 bis), 237-254.

Osborne, J. F., Henderson, J. B., MacPherson, A., Szu, E., Wild, A. y Yao, S.-Y. (2016). The development and validation of a learning progression for argumentation in science. *J Res Sci Teach*, *53*, 821-846. https://doi.org/10.1002/tea.21316

Pappas, S. (2024). Birds are not real: Exploring the Toulmin model of argumentation. *Communication Teacher*, 38(2), 105-110. https://doi.org/10.1080/17404622.2023.2300702

Pereda, C. (1994). *Vértigos argumentales*. Anthropos.

Perelman, C. y Olbrecht, L. (2006). *Tratado de la argumentación. La nueva Retórica*. Gredos.

Polo, C., Lund, K., Plantin, C. y Niccolai, G. P. (2016). Group emotions: The social and cognitive functions of emotions in argumentation. *International Journal of Computer-Supported Collaborative Learning*, 11, 123-156. https://doi.org/10.1007/s11412-016-9232-8

Rodríguez Ortiz, A. M., Ruiz Ortega, F. J. y Hernández Rodríguez, J. C. (2023). Modelos argumentativos: Análisis tipológico para su aplicabilidad en el aula. *Revista Guillermo de Ockham*, *21*(2), 639-656. https://doi.org/10.21500/22563202.5980

Sandoval, W. y Millwood, K. (2007). What can argumentation tell us about epistemology? En S. Erduran, y M. P. Jiménez-Alexaindre (Eds.), *Argumentation in science education: Perspective from classroom-based research* (pp. 71-89). Springer Science.

Stincer Gómez, D. y Blum Grynberg, B. (2017). El modelo argumentativo de Toulmin y la eficacia de titulación. *Revista electrónica de investigación educativa*, *19*(4), 9-19. https://doi.org/10.24320/redie.2017.19.4.1331

Toulmin, S. (1958). *The uses of argument*. Cambridge University Press.

Van Eemeren, F. H. y Grootendorst, R. (2004). *A systematic theory of argumentation. The pragma-dialectical approach*. Cambridge University Press.

Vega Reñón, L. (2003). *Si de argumentar se trata*. Montesinos.

Walton, D. (1992). *The place of emotion in argument*. Penn State University Press.

World Schools Debating Championships Ltd. (2025). *World Schools Debating Championships*. https://www.wsdcdebating.org

Developing Argumentation with the Toulmin Model and the WSDC Format at an Advanced Level

Dení Stincer Gómez
Universidad Autónoma del Estado de Morelos,
Facultad de Psicología, México
dstincerg@gmail.com
ORCID: https://orcid.org/0000-0003-3884-6805

Norma Miller
Universidad Tecnológica de Panamá,
Ciudad de Panamá
norma.miller@utp.ac.pa
ORCID: https://orcid.org/0000-0002-1995-4299

Nanette Archer Svenson
Centro de Investigación Educativa, Ciudad de Panamá, Panamá
nanette.svenson@gmail.com
ORCID: https://orcid.org/0000-0003-3837-5900

Jane Saldaña
Independiente, Chiriquí, Panamá
janecaminando@gmail.com
ORCID: https://orcid.org/0000-0002-8079-2129

The purpose of this study was to evaluate the development of argumentation skills in university freshmen that were exposed to one of two models, either the Toulmin argumentation model (MAT) or the World Schools Debating Championships (WSDC) format. Both have been used in the teaching and practice of argumentation, an epistemic skill associated with fostering a critical citizenry, capable of making informed decisions. MAT emphasizes the logical construction of arguments through its basic components: claim, evidence, warrants, backings and rebuttals, while the WSDC format highlights the pragmatic, confrontational and social aspects of argumentation (Van Eemeren y Grootendorst, 2004), promoting debate by means of actors with predefined roles in favor and against a «motion», and subject to discussion regulation mechanisms.

A mixed methods study was designed to test the effectiveness of these models. The quantitative part used a quasi-experimental, pretest-posttest design, with two groups, G1 and G2. The G1 group was exposed to Toulmin's model, whereas in G2 the WSDC format was employed. Interventions consisted of two workshops of 8 sessions, each one lasting 80 minutes. Participation in the study was voluntary, and participants were recruited from newly admitted first-year students at a public Panamanian university.

Lin & Mintzes (2010) Argumentation Skills Questionnaire was used for both the pretest and the posttest. To apply this instrument, a problematic social situation is proposed, with regards to which students must adopt a position and subsequently respond to four questions evaluating the following argumentation skills: 1) making claims and giving warrants, 2) constructing counterarguments, 3) generating supportive arguments, including rebuttals, and 4) providing evidence. Answers were scored using the scoring rubric developed by these authors. To understand score differences, a content analysis of the answers was conducted. Quantified data was analyzed using Stata (v 9.2) and Excel. Given the small sample size (10 students in G1 and 8 in G2), Wilcoxon's test was used to test statistical significance.

The mean total score (sum of scores on all four questions) for each group on the pretest was 10.9 (±2.5) points for G1 and 10.5 (±4.6) for G2, which indicates that initially the groups had similar argumentation skill levels. On the posttest, G1 obtained 14.3 ±(5.4) points, whereas G2 obtained 13.3 ±(3.7). Though the final mean total score increased for both groups, the difference was not significant, suggesting that both treatments resulted in similar gains. Moreover, improvements in G1 were observed in all four argumentation skills; in the G2 group, improvements were found only in two of the four.

The qualitative analysis of the participants' answers showed evidence of improved and evolving argumentation skills. Among those students whose scores increased, it was found that their answers not only included greater numbers of the expected elements (e.g., more reasons in favor of a given position, more evidence from reliable sources, better rebuttals of counterarguments), but also introduced persuasive, moral, synthetic and ethical elements.

This research has important psychopedagogical implications. It suggests that training students in argumentation skills helps develop their analytic and critical thinking abilities, as well as their written expression. Students trained in the Toulmin model tended to improve in all argumentation skills, though not at a statistically significant level; whereas students trained in the WSDC format improved significantly in skills most akin to debate: counterarguing and providing evidence for and against a position. Hence, this study empirically showed that systematic training both in Toulmin's model and the WSDC format develops the ability to argue in a manner «as close as possible» to critical context dialogues.