

REVISTA BOLETÍN REDIFE: 14 (10) OCTUBRE 2025 ISSN 2256-1536
RECIBIDO EL 18 DE MAYO DE 2025 - ACEPTADO EL 17 DE AGOSTO DE 2025

DEL RECONOCIMIENTO A LA EVALUACIÓN CRÍTICA: EFECTOS DE UNA ESTRATEGIA DE DESARROLLO PROFESIONAL EN ARGUMENTACIÓN CIENTÍFICA EN EDUCACIÓN SUPERIOR

FROM RECOGNITION TO CRITICAL EVALUATION: EFFECTS OF A PROFESSIONAL DEVELOPMENT STRATEGY ON SCIENTIFIC ARGUMENTATION IN HIGHER EDUCATION

Alejandra Martínez-Morales¹

Milany Gómez Betancur²

Universidad Antonio Nariño, Bogotá, Colombia.

Resumen

Este estudio presenta los resultados de una estrategia de desarrollo profesional para fortalecer la capacidad de docentes universitarios en la evaluación y retroalimentación de la argumentación científica. Participaron cinco profesores de ciencias en un proceso dividido en tres módulos: ensayos reflexivos, diseño

¹ Candidata a Doctora en Educación, Universidad Antonio Nariño. Docente Ciencias Naturales-Física, Gobernación de Cundinamarca. almartinez54@uan.edu.co. <https://orcid.org/0000-0002-1601-8374>

² Doctora en Marketing Político, actores e instituciones en las sociedades contemporáneas. Docente Facultad de Educación, Universidad Antonio Nariño. migomez23@uan.edu.co. <https://orcid.org/0000-0001-9362-3896>

de actividades argumentativas y aplicación del instrumento Del Conocimiento del Contenido Pedagógico de la Argumentación (PCK-A). El análisis combinó categorías cualitativas con estadísticas descriptivas. Los hallazgos muestran un avance gradual: se reconoce la importancia de argumentar en clase, se transforman situaciones cotidianas en problemas científicos y se incorpora la estructura de Toulmin para enriquecer el análisis y el debate, lo que finalmente conduce a decisiones que integran teoría y práctica. Se evidencian mejoras en el uso de evidencias y en el diálogo, aunque

persisten retos para explicar garantías y refutar sin criterios claros. Hacer explícita la estructura de Toulmin y emplear rúbricas por componente se plantea como clave para consolidar la práctica de la argumentación y convertir las debilidades en oportunidades de mejora.

Abstract

This study reports the results of a professional development strategy aimed at strengthening university instructors' ability to evaluate and provide feedback on scientific argumentation. Five science faculty members participated in a three-module process involving reflective essays, the design of argumentative activities, and the administration of the PCK-A instrument. The analysis combined qualitative category coding with descriptive statistics. Findings reveal a gradual progression: the importance of classroom argumentation is first acknowledged, everyday situations are then reframed as scientific problems, and finally Toulmin's argument structure is incorporated to deepen analysis and discussion, ultimately guiding decisions that integrate theory and practice. Improvements are evident in the use of evidence and in dialogic interactions, although challenges remain in articulating warrants and offering refutations when clear criteria are lacking. Making Toulmin's structure explicit and applying component-based rubrics emerge as key steps to consolidating argumentative practice and turning identified weaknesses into opportunities for improvement.

Palabras clave

Argumentación científica, Toulmin, evaluación formativa, educación superior, desarrollo profesional docente

Keywords

Scientific argumentation, Toulmin, formative assessment, higher education, teacher professional development

Introducción

En la enseñanza universitaria de las ciencias, la argumentación no es un adorno, sino una práctica esencial para razonar con evidencias (Briceño-Martínez & Benarroch, 2012; Briceño-Martínez, 2013; Cabrera Lanzo et al., 2022; Gibbs & Coffey, 2004; Knight, 2002). La literatura documenta que visibilizar los componentes del argumento, promover prácticas dialógicas con refutación y alinear la evaluación a criterios epistémicos eleva la calidad del razonamiento (Bağ & Çalik, 2017; Erduran et al., 2015; Jiménez-Aleixandre & Díaz, 2003; Molina, 2012; Osborne et al., 2004). Sin embargo, su implementación enfrenta obstáculos como la persistencia de enfoques expositivo-transmisivos, el tiempo lectivo limitado y las dificultades para explicitar cómo la evidencia respalda una afirmación o para confrontar alternativas con criterios claros (Briceño-Martínez & Benarroch, 2012; Fishman et al., 2017; Gray & Kang, 2014; González-Howard & McNeill, 2019; Larraín et al., 2017; McNeill & Knight, 2013; Nielsen, 2011; Osborne et al., 2013; Wess et al., 2023).

La Estrategia de Desarrollo Profesional Docente FAEUC se diseñó para fortalecer al profesorado universitario de ciencias en la promoción y evaluación de la argumentación científica, combinando un andamiaje conceptual, que integra aportes de Plantin, Toulmin y Walton, el Patrón Argumentativo de Toulmin (TAP) y el PCK-A, con herramientas didácticas y marcos dialógicos en modalidad 100 % virtual y asincrónica, alineados con principios de desarrollo profesional efectivo centrado en contenido, práctica y retroalimentación (Darling-Hammond et al., 2017).

La argumentación científica se concibe como una práctica epistémica contextualizada en la que justificar con evidencia y someter las afirmaciones al escrutinio de una comunidad es central (Bricker & Bell, 2008; Jiménez-Aleixandre & Erduran, 2007). El TAP explicita

componentes, afirmación, datos, garantía, respaldo, calificadores y refutación, visibilizando el nexo evidencia-conclusión y permitiendo enseñar el tratamiento de contraargumentos con criterios (Toulmin, 2007; Erduran et al., 2004; Kelly & Takao, 2002).

La interacción dialógica requiere normas claras sobre carga de la prueba, pertinencia y manejo de objeciones, así como esquemas conversacionales que ordenen turnos y cierres de desacuerdo (Walton, 1996). En el aula de ciencias, estas normas se concretan en tendencias discursivas que exigen evidencias, garantías, calificadores y consideración de alternativas (Osborne et al., 2004, 2013; Christodoulou & Osborne, 2014). El profesorado orienta el intercambio mediante preguntas auténticas, parafraseo productivo y exigencia de justificación, mientras que la interacción entre pares aumenta la exigencia argumentativa (Kim & Hand, 2015; Chen et al., 2019; Felton et al., 2022; González-Howard & McNeill, 2019).

El diseño de tareas debe ofrecer oportunidades para construir, comparar y revisar el vínculo evidencia-garantía-conclusión y explicitar calificadores y contraargumentos. Enfoques como *Argument-Based Inquiry*, debates con evidencia, estudios de caso y simulaciones cumplen este papel (Sampson et al., 2011; Osborne et al., 2004; Larraín et al., 2017; Fishman et al., 2017).

Para trasladar estos apoyos a la práctica, el PCK-A y el desarrollo profesional docente son claves: integrar saber disciplinar, comprensión del aprendizaje y decisiones pedagógicas permite materializar el conocimiento en intervenciones concretas (Osborne et al., 2013; McNeill et al., 2016; Ozdem-Yilmaz et al., 2017). La evidencia respalda que el desarrollo profesional efectivo combina foco en el contenido, práctica deliberada y retroalimentación específica, incluso en formatos virtuales asincrónicos

(Darling-Hammond et al., 2017; Wess et al., 2023).

Cuando se incorporan herramientas digitales, es fundamental articular el PCK con consideraciones tecnológicas para seleccionar recursos que hagan visible la relación evidencia-conclusión y faciliten la participación dialógica (Chai et al., 2013). Evaluar y retroalimentar con criterios claros eleva la calidad argumentativa y vincula la evidencia de aprendizaje con las decisiones de enseñanza (Sandoval, 2003; Hattie & Timperley, 2007; Sampson & Clark, 2006; Kelly & Takao, 2002; McNeill et al., 2016). Desde este marco, el análisis de productos docentes debe evidenciar la explicitación estructural de la argumentación científica y su distinción de la cotidiana, la solidez del vínculo evidencia-conclusión, la gestión del diálogo, el diseño de tareas con datos y oportunidades de contraste, y la evaluación y retroalimentación informadas por el PCK de la argumentación (Erduran & Jiménez-Aleixandre, 2007; Osborne et al., 2013; Sampson & Blanchard, 2012).

Metodología

Participantes y contexto

Participaron cinco docentes universitarios de ciencias naturales, tres de física, uno de química y uno de biología, que se incorporaron de manera voluntaria. Para resguardar la confidencialidad se emplearon seudónimos (D1-D5). La muestra estuvo compuesta por cuatro hombres y una mujer, con edades entre 29 y 36 años (media $\approx 32,4$). En términos de experiencia profesional general, dos docentes reportaron 11-20 años, dos 6-10 años y uno 0-6 años; en docencia universitaria, tres informaron 0-6 años y dos 6-10 años. Respecto a la formación, uno cuenta con doctorado y cuatro con maestría. En cuanto al nivel de enseñanza, tres se desempeñan en pregrado y dos en maestría. La estrategia de DPD se implementó como un curso breve (≈ 10 horas), 100% virtual y asincrónico, articulado en

tres módulos. El trabajo se realizó en español y se apoyó en guías descargables y tareas de análisis o producción.

Diseño e instrumentos

La Tabla 1 sintetiza, por módulo, los objetivos, contenidos, actividades, duración y productos de la estrategia implementada. Este diseño responde a la búsqueda de la progresión de

las concepciones y prácticas argumentativas de docentes en el contexto universitario y, por lo tanto, al fortalecimiento de la enseñanza de la argumentación científica. Los productos de cada módulo son analizados desde un enfoque mixto a partir de los elementos conceptuales identificados en el marco teórico de la investigación e incluido en las guías de la estrategia, o bajo los criterios planteados en el instrumento utilizado (McNeill et al., 2016).

Tabla 1. Estructura general y síntesis de la estrategia de DPD FAEUC

Modulo	Objetivo	Contenidos	Actividades	Duración	Recursos
1: Fundamentos de la argumentación científica	Proporcionar a los docentes una comprensión sólida de los conceptos clave en torno a la argumentación científica y su importancia en el aprendizaje de las ciencias.	Introducción a la argumentación científica.	Lectura y reflexión.	1h	Guía y Lectura
		Estructura de la argumentación científica.	Lectura y análisis.	1 h	Guía y Lectura
		La argumentación en el contexto universitario.	Lectura y ensayo reflexivo.	2 h	Guía y Lectura
Producto	Ensayo reflexivo (ER) integrador que analice el impacto de la argumentación en su práctica docente actual (1500 – 2000 palabras).				
2: Estrategias para promover la argumentación en el aula universitaria	Desarrollar estrategias efectivas que los docentes puedan aplicar en el aula para promover la construcción de argumentos sólidos y basados en evidencias.	Metodologías para promover la argumentación	Lectura y diseño de actividades.	1h	Guía y Lectura
		Estrategias dialógicas y diseño de actividades.	Lectura y diseño de actividades.	1 h	Guía y Lectura
		Promoción de la argumentación colaborativa y diseño de actividades.	Lectura y diseño de actividades.	2 h	Guía y Lectura
Producto	Actividad argumentativa (AA) específica para su disciplina, integrando estrategias de promoción, discusión y colaboración argumentativa (800 – 1200 palabras).				
3: Evaluación de la argumentación científica	Dotar a los docentes de herramientas para evaluar las habilidades argumentativas científicas de los estudiantes de manera eficaz y objetiva.	Criterios de evaluación de la argumentación Retroalimentación constructiva	Lectura	1 h	Lectura
		Aplicación de instrumento PCK argumentación.	Desarrollo instrumento	1 h	Instrumento
Producto	Desarrollo del instrumento para evaluar el Conocimiento Pedagógico del Contenido en Argumentación Científica (PCK-A) (McNeill et al., 2016).				

*Fuente: elaboración propia

Análisis de datos

Para todos los productos se siguió un enfoque mixto con codificación abierta y axial (Corbin & Strauss, 2008) asistida por ATLAS.ti, cuando fue pertinente, organizando categorías a partir de los propósitos de cada módulo, los contenidos trabajados y el marco teórico sobre argumentación científica. En los demás casos se trabajó con rubricas analíticas, sin requerir software de análisis cualitativo. Las categorías se construyeron de manera deductivo-inductiva, se parte de un conjunto a priori y se enriquecen con categorías emergentes identificadas en la lectura de los documentos, cuando así fue necesario.

En el caso de los ER del Módulo 1 (M1), el análisis se centró en las dimensiones centrales de la argumentación en la docencia universitaria, el reconocimiento y diferenciación de formas de argumentación, el tránsito hacia prácticas con criterios científicos, el uso de esquemas estructurales, la comprensión de la argumentación como práctica epistémica y la proyección de mejoras pedagógicas. La valoración consideró presencia, profundidad y coherencia por dimensión. La Tabla 2 presenta el detalle de códigos y definiciones empleados.

Tabla 2. Códigos utilizados en el análisis cualitativo de los ER Modulo 1

Código	Categoría
ARG_REC	Reconocimiento de la argumentación en la práctica docente
ARG_TRN	Transformación de la argumentación cotidiana en científica
MOD_TOU	Uso del modelo de Toulmin para analizar un tema científico
ARG_EPI	Comprensión crítica de la argumentación como práctica epistémica
MEJ_PRA	Proyección de mejoras en la práctica docente
INT_TPR	Integración teórico-práctica
HAB_CRIT	Reflexión crítica sobre el proceso de aprendizaje (Código emergente)
DES_LIM	Desafíos y limitaciones en la implementación de la argumentación científica (Código emergente)

*Fuente: elaboración propia

Dentro de las AA del Módulo 2 (M2), se realizó un análisis cualitativo guiado por las categorías descritas en la Tabla 3. Se valoraron presencia, profundidad y coherencia en cada una de ellas para identificar fortalezas, brechas y oportunidades de mejora.

Tabla 3. *Tabla de códigos para el análisis cualitativo de las AA del Módulo 2*

	Categoría	Subcategoría	Código
1	Pertinencia del tema o problema científico	1.1 Relevancia disciplinar	PR
		1.2 Potencial argumentativo	PP
		1.3 Contextualización	PC
2	Estrategia metodológica empleada	2.1 Estrategia explícita (TAP, ABI, SWH...)	EEE
		2.2 Estructura de la actividad	EEA
		2.3 Colaboración entre estudiantes	ECE
3	Integración del enfoque argumentativo	3.1 Uso del modelo de Toulmin u otro	IUM
		3.2 Promoción de evidencia y contraargumentos	IPE
		3.3 Visualización de argumentos	IVA
4	Evaluación de la calidad argumentativa	4.1 Criterios estructurales	EVCE
		4.2 Criterios epistémicos o dialógicos	EVCD
		4.3 Instrumentos de evaluación	EVIE
5	Promoción de habilidades colaborativas	5.1 Roles diferenciados en la actividad	PR
		5.2 Interacción y toma de decisiones	PI
		5.3 Clima de respeto y participación	PC
6	Reflexión del docente	6.1 Impacto esperado en el aprendizaje	RI
		6.2 Reconocimiento de retos o limitaciones	RR
		6.3 Aportes a la práctica docente	RA
7	Uso de recursos tecnológicos (si aplica)	7.1 Integración de TIC en la actividad	UI
		7.2 Finalidad del recurso tecnológico	UF

*Fuente: elaboración propia

Para el PCK-A aplicado en el Módulo 3 (M3), se empleó un análisis mixto, puntajes en ítems cerrados y rúbricas para respuestas abiertas. Se evaluaron las dimensiones de uso de evidencia (1A), razonamiento científico (1B), interacciones entre estudiantes (2A) y crítica de afirmaciones opuestas (2B). Todas las dimensiones se puntuaron y describieron conforme a las directrices, claves y ejemplos de codificación establecidos por las autoras del instrumento (McNeill et al., 2016). Este sistema permitió caracterizar aprendizajes, identificar patrones

y diferencias de apropiación, así como ofrecer retroalimentación formativa sobre fortalezas, brechas y oportunidades de mejora en el uso didáctico de la argumentación científica.

Resultados

De la intuición a la argumentación científica

Los ER del M1 muestran un proceso consistente, los docentes reconocen la diferencia entre explicaciones de sentido común y la argumentación científica, y ese reconocimiento

se traduce en acciones concretas en el aula. El TAP aparece como la estructura del trabajo para ordenar razones, evidencias y límites, la refutación y los calificadores se vuelven motores del cambio conceptual y la evaluación formativa, hace visible la calidad del argumento.

Esta lectura se apoya en el en el conjunto de códigos empleado (Tabla 2) y en los conteos por documento (Tabla 4), donde destacan ARG_REC (19 menciones), MEJ_PRA (18) y MOD_TOU (14).

Tabla 4. *Tabla de códigos por documento*

Código/ER	ER-D1	ER-D2	ER-D5	ER-D4	ER-D3	Totales
ARG_REC	9	5	3	1	1	19
MEJ_PRA	8	5	3	1	1	18
MOD_TOU	5	4	2	2	1	14
ARG_EPI	6	3	3	0	0	12
ARG_TRN	3	4	1	1	2	11
INT_TPR	1	4	2	1	0	8
HAB_CRI	1	3	1	0	1	6
DES_LIM	1	2	2	0	0	5
Totales	34	30	17	6	6	93

*Fuente: elaboración propia

En ARG_REC, emergen tres ideas con fortaleza. Primero, se reconoce la centralidad de la argumentación: “la argumentación científica es una herramienta pedagógica esencial” (ER-D1, p. 1) y “fundamental para fomentar la comprensión profunda” (ER-D3, p. 1). Segundo, se identifica la distinción explícita entre tipos de argumentación: “inicialmente no distinguía claramente entre la argumentación cotidiana y la científica” (ER-D2, p. 1). Y, por último, se manifiesta el razonamiento cotidiano (perceptivo e intuitivo): “basada en la percepción sensorial directa” (ER-D4, p. 1) y la mezcla entre ambos (ER-D2, p. 1). Este reconocimiento no solo se declara, sino que se identifican herramientas para conseguirlo, el uso del TAP (“se puede utilizar el modelo de Toulmin”, ER-D1, p. 1), la confrontación con evidencia (“confrontaron su intuición con la evidencia”, ER-D3, p. 2) y la modelización/medición de lo no observable (“el campo eléctrico... puede describirse, modelarse

y medirse”, ER-D4, p. 1). En conclusión, identificar la diferencia facilita construir la mediación entre formas de argumentar.

Ese tránsito se concreta en MEJ_PRA. Las líneas de acción que se identifican en los ER, van desde la estructuración con TAP: “usar el modelo de Toulmin” como dispositivo central (ER-D1, p. 1), hasta el acompañamiento cercano: “acompañamiento individualizado... tutorías” (ER-D2, p. 2). De la misma manera se reconocen las prácticas dialógicas: “un momento de debate... y retroalimentación” (ER-D2, p. 2); los puentes empírico-modalizadores: “confrontaron su intuición con la evidencia” (ER-D3, p. 2), “visualizar campos eléctricos” (ER-D4, p. 2); y la evaluación formativa: “evaluaciones formativas y retroalimentación continua” (ER-D1, p. 3); “autoevaluación y coevaluación” (ER-D4, p. 3).

El uso del TAP (código MOD_TOU) opera como articulación entre diagnóstico y mejora. Aparecen sistemáticamente la estructuración explícita (C/D/G/S/Q/R), la articulación teoría-evidencia, los cualificadores o condiciones, la refutación operativa y la contextualización disciplinar. Este patrón muestra distintos niveles de implementación. Desde el listado explícito de componentes TAP, hasta usos contextualizados que sostienen la lógica sin nombrar cada componente, lo que habilita intervención y retroalimentación en tiempo real.

En ARG_TRN, el proceso típico es claro. Se parte de la intuición disparadora (“un objeto pesado cae más rápido”, ER-D3, p. 2), a la reconfiguración con TAP (“afirmación, datos, respaldo, cualificador, refutación y base”, ER-D1, p. 1), luego a los puentes empírico-modelizadores (experimentos, simulaciones, datos de campo), para pasar a los cualificadores y o la refutación para romper generalizaciones, luego se pasa al lenguaje técnico y por último a la evaluación formativa para mantener el cambio. Las dificultades más recurrentes, la resistencia del sentido común, la alta abstracción y el tiempo didáctico, explican por qué el tránsito es gradual y necesita andamiaje.

El código ARG_EPI, agrega el para qué del proceso. La formación para la ciudadanía científica (“decisiones informadas y compromiso público”, ER-D5, p. 1), la cultura epistémica de aula (debate, criterios, retroalimentación), la trazabilidad evidencia-modelo-conclusión, la epistemología no absolutista (“siempre que...”, ER-D3, p. 2; “en algunas situaciones...”, ER-D2, p. 3), y la autonomía o metacognición (“enseñar ciencia implica enseñar a pensar científicamente”, ER-D4, p. 2). Este marco regula el tono del curso: justificar, refutar y cualificar como normas.

Las categorías emergentes precisan el panorama. HAB_CRIT concentra los procesos

metacognitivos (“antes... transmisión directa”, ER-D3, p. 3; “ha transformado mi manera de enseñar”, ER-D4, p. 2) que llevan a ajuste metodológico (debates, rúbricas) y a establecer evidencias de impacto (pre-post, retroalimentación). DES_LIM sitúa límites del diseño, persistencia del sentido común (“se ‘mezcla’ lo cotidiano con lo científico”, ER-D2, p. 1), los elementos abstractos (“el campo eléctrico no se ve”, ER-D4, p. 1), el tiempo y la capacidad de acompañamiento (ER-D2, p. 2; ER-D3, p. 3) y la continuidad del modelo centrado en la transmisión (ER-D5, p. 1).

El patrón de coocurrencias (Tabla 5) confirma dos núcleos funcionales. El operativo—ARG_REC, ARG_TRN, MOD_TOU, MEJ_PRA, INT_TPR—articula reconocimiento, tránsito, gramática TAP, mejora e integración; y el regulador—ARG_EPI, HAB_CRIT, DES_LIM—aporta criterios epistémicos, reflexión y límites de implementación. Se destaca el vínculo ARG_REC—MEJ_PRA (15; $J=0,68$): cuando el profesorado reconoce el problema, actúa (debates, informes con refutación, rúbricas, diagnósticos). También son fuertes ARG_TRN—MOD_TOU y MEJ_PRA—MOD_TOU ($J\approx 0,39$), la transformación ocurre cuando el TAP estructura el razonamiento y las tareas movilizan evidencia o la modelización. Se identifican dos brechas, por un lado, llevar HAB_CRIT al plano de la planificación (enlace débil con INT_TPR, $J=0,08$) y por otro traducir DES_LIM a requisitos de diseño (poca coocurrencia con MEJ_PRA, $J=0,05$).

Tabla 5. Coocurrencias de las categorías de análisis en los ER

	ARG_ EPI	ARG_ REC	ARG_ TRN	DES_ LIM	HAB_ CRI	INT_ TPR	MEJ_ PRA	MOD_ TOU
ARG_EPI	0* (1.00)	6 (0.24)	1 (0.05)	4 (0.31)	3 (0.20)	2 (0.11)	5 (0.20)	2 (0.08)
ARG_REC	6 (0.24)	0 (1.00)	8 (0.36)	2 (0.09)	4 (0.19)	3 (0.13)	15 (0.68)	8 (0.32)
ARG_TRN	1 (0.05)	8 (0.36)	0 (1.00)	2 (0.14)	4 (0.31)	4 (0.27)	8 (0.38)	7 (0.39)
DES_LIM	4 (0.31)	2 (0.09)	2 (0.14)	0 (1.00)	1 (0.10)	1 (0.08)	1 (0.05)	1 (0.06)
HAB_CRI	3 (0.20)	4 (0.19)	4 (0.31)	1 (0.10)	0 (1.00)	1 (0.08)	4 (0.20)	3 (0.18)
INT_TPR	2 (0.11)	3 (0.15)	4 (0.27)	1 (0.08)	1 (0.08)	0 (1.00)	4 (0.18)	6 (0.38)
MEJ_PRA	5 (0.20)	15 (0.68)	8 (0.38)	1 (0.05)	4 (0.20)	4 (0.18)	0 (1.00)	9 (0.39)
MOD_TOU	2 (0.08)	8 (0.32)	7 (0.39)	1 (0.06)	3 (0.18)	6 (0.38)	9 (0.39)	0 (1.00)

* conteo bruto de coocurrencias y, entre paréntesis, el índice normalizado de Jaccard. Fuente: elaboración propia

La trayectoria y el campo de formación y desempeño de los docentes participantes modulan la implementación. En química (D1), el TAP se ancla en marcos teóricos robustos y se institucionaliza la evaluación (“mediré el impacto... evaluaciones formativas”, p. 3). En física (D2), el foco está en refutación empírica e informes que exigen contraejemplos, sostenidos por tutorías (“acompañamiento individualizado”, p. 2). En física/matemáticas (D3, D4), el TAP se procedimentaliza (enumeración C/D/G/S/Q/R), con simulaciones para sostener contenidos abstractos y rúbricas para visibilizar calidad argumentativa. En Biología (D5), predomina la contextualización situada (datos de campo, perturbaciones reales) y la proyección a la ciudadanía y validación colectiva. Estas particularidades son complementarias, la experticia disciplinar, la lectura contextualizada del aula, el cambio evaluativo, la mediación tecnológica y el anclaje contextual convergen en robustecer la argumentación.

La evidencia de los ER constata y amplía elementos ya identificados en los referentes teóricos o metodológicos. Primero, el TAP funciona como un marco operativo, hace visibles los componentes datos–garantía–respaldo y obliga a calificar y refutar, lo que permite intervenir con precisión en la calidad del razonamiento (Toulmin, 2007; Erduran et al., 2004; Osborne et al., 2004, 2013). Segundo, las normas dialógicas y epistémicas estructuran la interacción, reglas claras sobre turnos, responsabilidad de aportar evidencia, su pertinencia y la respuesta a objeciones, elevan la calidad del diálogo y estabilizan la argumentación como práctica de aula, no como evento ocasional (Driver et al., 2000; Osborne et al., 2004, 2013; Walton, 1996). Tercero, una evaluación alineada con el TAP, soporta mejoras acumulativas y orienta el siguiente paso de enseñanza y aprendizaje (Kelly & Takao, 2002; Hattie & Timperley, 2007). Con estos tres pilares en juego, lo que muestran los ER deja de ser un conjunto de casos y se convierte en

una secuencia lógica de transformación de la enseñanza.

Igualmente, los límites reportados son esperables y orientan dos estrategias, (a) plantillas TAP que hagan obligatoria la refutación y los calificadores; (b) banco de actividades y pautas que precisen las premisas de aplicación del modelo y criterios mínimos de evaluación para asegurar que la relación evidencia–modelo–conclusión sea verificable. Así, HAB_CRIT deja de ser reflexión “a posteriori” y permite pasar a la planificación, y DES_LIM deja de ser una barrera y se convierte en una pauta de aplicación en el aula.

Se identifica un proceso claro—REC → TRN → (TAP) → MEJ → INT—que convierte la intención de argumentar mejor en prácticas verificables. Los estadísticos de frecuencia (Tabla 4) y coocurrencia (Tabla 5) respaldan ese proceso; las citas ejemplifican cómo se materializa en decisiones de aula; y los perfiles muestran vías complementarias para mantenerlo. El siguiente paso es cerrar las brechas detectadas, planificar con ítems TAP obligatorios, integrar la simulación y la medición con registros comparables y alinear la evaluación con evidencias de referencia. Con

ello, la argumentación científica deja de ser una consigna y se instituye como práctica cotidiana con criterios claros de calidad.

De la intención a la práctica: trayectorias de implementación

El conjunto de AA del M2 muestra un patrón consistente (Tabla 6), cuando la estrategia de aula está bien secuenciada (EEE/EEA altos), la argumentación emerge con evidencia y refutación (IPE alto) y la evaluación reconoce esa calidad discursiva (EVCD alto). La variable que explica las diferencias entre diseños es el grado de explicitación del patrón de TAP, tanto en el planteamiento como en el producto. En D3, D4 y D5 el TAP se exige de forma operativa (IUM alto), lo que permite valorar la estructura del argumento (EVCE alto). En D1 y D2, aunque hay uso sistemático de evidencia y contraargumento, el TAP queda implícito (IUM medio) y la evaluación privilegia lo dialógico sobre el esquema interno del razonamiento. En casi todos los casos la visualización del argumento (IVA) es una dificultad, abundan gráficos de contenido y simulaciones, pero rara vez se requiere un organizador Toulmin que haga visible C/D/G/S/Q/R.

Tabla 6. Matriz de síntesis comparativa M2

Categoría	Código	D1	D2	D3	D4	D5
Pertinencia	PR	Alto	Alto	Alto	Alto	Alto
	PP	Alto	Alto	Alto	Alto	Alto
	PC	Medio	Medio–Alto	Alto	Alto	Alto
Estrategia	EEE	Alto	Medio–Bajo	Alto	Alto	Alto
	EEA	Alto	Alto	Alto	Alto	Alto
	ECE	Alto	Alto	Alto	Alto	Alto
Integración argumentativa	IUM	Medio-bajo	Medio–Bajo	Alto	Alto	Alto
	IPE	Alto	Alto	Alto	Alto	Alto
	IVA	Bajo	Medio	Medio	Medio	Medio
Evaluación	EVCE	Medio	Medio	Alto	Alto	Alto
	EVCD	Alto	Alto	Alto	Alto	Medio
	EVIE	Medio-bajo	Alto	Medio	Medio	Alto

Categoría	Código	D1	D2	D3	D4	D5
Colaborativas	CPR	Alto	Alto	Alto	Alto	Medio-Bajo
	PI	Alto	Alto	Alto	Alto	Alto
	CPC	Alto	Alto	Medio	Medio-Alto	Alto
Reflexión docente	RI	Alto	Alto	Alto	Alto	Alto
	RR	Bajo	Bajo	Bajo	Bajo	Medio
	RA	Medio	Medio	Medio	Medio	Alto
TIC	UI	Bajo	Alto	Medio	Alto	Medio-Alto
	UF	Bajo	Alto	Alto	Alto	Alto

*Fuente: elaboración propia

El proceso estrategia–integración–evaluación es especialmente robusto cuando: (a) el planteamiento exige pasos TAP, (b) las tareas piden medir o simular y explicitar cualificadores y refutaciones, y (c) la evaluación dispone de criterios que distinguen entre calidad dialógica (EVCD) y calidad estructural (EVCE). En ese marco, la colaboración con roles definidos y debate intergrupal cataliza la refutación y el consenso, y el uso de TIC con función epistémica, permite identificar la relación entre evidencia, modelo y conclusión (UI/UF altos). Donde las TIC registran parámetros, unidades y condiciones, crecen IPE y EVCD, porque el discurso se apoya en pruebas comparables.

En evaluación, todos coinciden en criterios epistémicos (uso y validez de evidencia, capacidad de refutar, escucha y respeto). La diferencia está en la instrumentación, D2 y D5 publican rúbricas operativas (EVIE alto), D3 y D4 mantienen EVIE medio por falta de ponderaciones explícitas, D1 queda medio-bajo. Sin rúbricas por componente TAP y sin ejemplos de soporte, la trazabilidad entre lo que se enseña y lo que se valora se debilita. Respecto a la reflexión docente, predomina la intención formativa (RI alto), pero la anticipación de retos (RR) es baja. Se describen fines (pensamiento crítico, ciudadanía científica) más que condiciones de diseño (tiempo, heterogeneidad, manejo técnico) para lograrlos.

La manera en que cada docente llevó el enfoque argumentativo a su aula dibuja un gradiente claro. D1 ancla el patrón de Toulmin en marcos teóricos robustos y lo convierte en práctica evaluable, la estructura del argumento no solo se enseña, también se mide con criterios explícitos y retroalimentación, lo que acelera su estabilización en clase. D2, enfatiza el trabajo con datos y la refutación empírica; combina discusión guiada, roles y simulación para tensionar intuiciones, y acompaña con una rúbrica ponderada. Su fortaleza está en el corredor evidencia–modelo–conclusión, aunque el TAP aparece menos explicitado en el planteamiento y en el producto final. En D3 y D4 el TAP se vuelve altamente procedural, se enumeran componentes, se formulan cualificadores y escenarios de excepción y se apoya el tránsito en simulaciones y gráficas. Esa explicitación facilita la apropiación por parte del estudiantado y, cuando se traduce en criterios de evaluación, se incrementa la coherencia entre lo que se pide y lo que se valora. D5 liga el TAP con datos de campo y problemas socioambientales, subrayando ciudadanía científica y validación colectiva; la secuencia incluye refutación obligatoria y rúbrica con retroalimentación escrita, aunque se podría fortalecer la visualización del argumento y los roles formales.

A partir de lo anterior se puede afirmar que la experiencia modula la profundidad del TAP.

Quienes tienen más trayectoria (D1, D2) lo integran con mayor naturalidad al currículo y la evaluación; quienes están en etapas iniciales (D3, D4) lo hacen explícito paso a paso, lo que reduce la ambigüedad y dinamiza la transformación didáctica. El contexto disciplinar añade acentuaciones, en física, las leyes actúan como garantías y la simulación aporta evidencia trazable; en química, la base teórica sostiene los respaldos; en biología, los datos observacionales y los casos reales habilitan refutaciones situadas. La mejora se consolida cuando la estructura argumental es visible en el planteamiento y exigible en el producto, y cuando la evaluación diferencia componentes con ponderaciones.

Estos patrones, dialogan de forma consistente con el marco teórico de la estrategia. En primer lugar, donde el TAP es explícito en tareas y productos, es que la intervención docente gana precisión y la calidad del argumento se hace observable (Toulmin, 2007; Erduran et al., 2004; Osborne et al., 2003). Además, respecto a las normas dialógicas y epistémicas, cuando se dan reglas claras de participación, se enfatiza en la importancia de la fundamentación y el manejo riguroso de los contrargumentos, se eleva la calidad del intercambio y afianzan la argumentación dentro de la dinámica del aula; esto se aprecia en los diseños con roles, debate guiado y refutación contextualizada (Driver et al., 2000; Osborne et al. 2013; Walton, 1996). En tercer lugar, cuando se incluyen rúbricas por componente TAP, ejemplos de soporte y retroalimentación con pasos concretos, consolidan los avances y previene valoraciones difusas (Kelly & Takao, 2002; Hattie & Timperley; 2007). A estos elementos se suma la indagación basada en argumentos, pues cuando la secuencia obliga a formular hipótesis operacionales, producir o analizar datos y defender conclusiones con cualificadores, aumenta la calidad del proceso de argumentación y el pensamiento crítico

(Sampson & Blanchard, 2012; Sampson & Clark, 2006). Finalmente, los contextos auténticos y las controversias socio-científicas actúan como dinamizadores de la pertinencia y de la refutación fundamentada, en consonancia con la literatura sobre argumentación en ciencias (Jiménez-Aleixandre & Erduran, 2007; Sandoval & Millwood, 2007).

Los referentes teóricos, también ayudan a interpretar las brechas encontradas. La persistencia de un modelo centrado en la transmisión de contenidos y el tiempo limitado obstaculizan el trabajo puntual con C/D/G/S/Q/R. Por lo tanto, en los casos más sólidos se convierten límites en especificaciones didácticas, como las plantillas TAP obligatorias en la estrategia y el producto, rúbricas con ponderación por componente, y uso de TIC con una función epistémica para garantizar ruta de justificación que hace visibles datos, modelos y conclusiones (Driver et al., 2000; Osborne et al., 2003; Sampson & Blanchard, 2012).

A partir de ese contraste emergen cinco limitaciones recurrentes en las actividades analizadas, que derivan en implicaciones prácticas para el diseño de actividades en las que se promueva la argumentación científica. (1) TAP implícito, aun con debate y uso de evidencia, si la estructura no se exige en la actividad y en el producto, no es posible ponderar con precisión C/D/G/S/Q/R. (2) Visualización insuficiente, sin una proyección de Toulmin obligatoria, el razonamiento permanece opaco y la retroalimentación se vuelve genérica. (3) Instrumentos incompletos, faltan rúbricas con ponderaciones y criterios por componente, además de su evaluación con ejemplos de soporte para garantizar consistencia. (4) TIC sin función epistémica, cuando no se registran parámetros y condiciones, no dejan un registro de decisiones ni posibilitan refutaciones fundamentadas. Y (5) poca anticipación de desafíos, el diseño rara vez convierte los límites

(tiempo, abstracción, heterogeneidad) en condiciones explícitas de las actividades y su evaluación.

Se identifica entonces tras el análisis de los productos del M2, un núcleo operativo común, estrategia clara, evidencia con refutación y dinámica colaborativa. Y deja dos tareas compartidas para consolidar el proceso EEE/EEA ↔ IUM/IPE/IVA ↔ EVCE/EVCD/EVIE: (1) hacer visible la estructura del argumento y (2) estandarizar la evaluación. Donde estas condiciones se cumplen, la argumentación deja de ser una intención y se vuelve un camino coherente de cambio metodológico con impactos observables en la comprensión, el juicio crítico y la toma de decisiones fundamentada.

Fortalezas en el uso de evidencia y retos en la explicitación de garantías y refutaciones

Los resultados del PCK-A aplicado al cierre del M3 permiten evidenciar que el profesorado identifica y solicita evidencia con relativa solidez, pero aún debe hacer más explícito el puente conceptual que conecta esa evidencia con las afirmaciones y, además, estructurar mejor la interacción y la refutación entre pares. En términos globales, la categoría 1A (uso de evidencia) alcanza el desempeño más alto y estable, mientras que 1B (razonamiento científico que conecta evidencia y afirmación) es la principal dificultad; 2A (interacciones) y 2B (crítica a afirmaciones opuestas) se sitúan en niveles intermedios, con heterogeneidad entre docentes. Esta lectura se sustenta en la combinación de ítems cerrados y respuestas abiertas, y se aprecia de forma sintética en la Tabla 7.

Tabla 7. Resultados globales del PCK

D	1A			2A			1B			2B			x T	SD T	Total
	C	A	T	C	A	T	C	A	T	C	A	T			
D1	2	1	3	2	1	3	1	1	2	0	1	1	2,25	0,96	9
D2	1	0,75	1,75	0	0,5	0,5	0	0,75	0,75	4	0,75	4,75	1,94	1,95	7,75
D3	3	0,66	3,66	3	1	4	4	0,66	4,66	2	0,5	2,05	3,71	0,90	14,82
D4	2	0,67	2,67	1	0,5	1,5	0	0,33	0,33	0	1	1	1,38	0,99	5,5
D5	1	1	2	1	0,5	1,5	0	0,5	0,5	1	0,5	1,5	1,38	0,63	5,5
x	1,8	0,82	2,61	1,4	0,7	2,1	1	0,65	1,64	1,4	0,7	2,15	2,13	0,39	8,51
SD	0,84	0,17	0,77	1,14	0,27	1,38	1,73	0,25	1,8	1,67	0,25	1,57	1,38	0,45	3,83

*C: Items cerrados, A: Preguntas Abiertas, T: Total. Fuente: elaboración propia

A partir de esta síntesis, se confirman fortalezas en la identificación y solicitud de datos pertinentes (1A), tanto en decisiones cerradas como en descripciones abiertas, aunque con tendencia a no triangular todas las fuentes disponibles ni a justificar por qué la combinación de datos potencia la conclusión. En 2A, los docentes promueven la discusión, pero con frecuencia sin diseño de la interacción dialógica que establezca roles, turnos y mecanismos de apertura de

réplica; por eso la calidad del intercambio depende en exceso del papel del docente. En 1B se evidencian confusiones persistentes, se enuncian conceptos sin explicar cómo permiten el salto de los datos a la afirmación. En 2B, la crítica aparece, pero puede volverse retórica si no se sustenta en criterios y en la contrastación explícita de la calidad y suficiencia de la evidencia.

La lectura cualitativa de las respuestas abiertas ayuda a precisar estos matices. Se observan buenas prácticas concretas, como el uso de Afirmación-Evidencia-Razonamiento (A-E-R) con preguntas de soporte (D4 en 2A) o el reconocimiento de múltiples beneficios del trabajo con posturas rivales (D1 y D3 en 2B). También aparecen patrones recurrentes, estrategias genéricas sin mecanismo dialógico (D2 y D5 en 2A), identificación de evidencia sin justificar su pertinencia (varios casos en 1A) y carencias en explicitar debilidades o garantías (1B y 2B).

Con base en estos hallazgos, la mejora del proceso de enseñanza apunta a formalizar rutinas que transformen lo deseable en prácticas recurrentes. En 1A, resulta pertinente estandarizar un inventario de evidencias y concluir cada actividad con una matriz concisa de triangulación que haga explícita la relevancia, suficiencia y procedencia de las fuentes. En 2A, la discusión gana calidad si se hace con un guion de roles, un repertorio de desencadenadores de turno y rondas de réplica con tiempos y criterios visibles. En 1B, el foco es la garantía, se sugiere incluir plantillas que obliguen a escribir “porque, según el principio X, cuando... entonces...” y a incluir contraargumentos. En 2B, se propone normalizar la secuencia Contra afirmación–Evidencia–Principio–Conclusión y un protocolo comparativo para el cierre.

Al mirar los resultados por perfiles profesionales y trayectorias, emergen rutas diferenciadas de fortalecimiento. Quienes provienen de matemáticas organizan con facilidad la cadena A-E-R y sostienen interacciones eficaces; su siguiente paso es explicitar garantías y formalizar la comparación de posturas con criterios visibles. Perfiles de ingeniería e investigación aplicada muestran fortaleza en selección y triangulación de datos y en manejo de interacciones; su oportunidad es estandarizar la refutación epistémica. En física se observa claridad

conceptual con estructuras argumentativas irregulares; aquí ayudan plantillas A-E-R con casillas de justificación y rutinas dialógicas con roles. Perfiles pedagógicos reconocen beneficios y riesgos del trabajo con múltiples afirmaciones y dinamizan el aula; su foco es hacer visible el enlace evidencia-principio-conclusión y diseñar rúbricas que puntúen las garantías. En biología se selecciona evidencia aceptable y hay niveles medios en interacción y refutación, pero con respuestas abiertas muy generales; el reto es operacionalizar la interacción y nombrar el principio que conecta los datos.

El contraste teórico respalda estas prioridades. Los resultados convergen con Zohar (2007) y con el marco de Toulmin (2007) en que no basta con tener datos, pues hay que explicitar la garantía que hace que esos datos sostengan la conclusión y sus límites de aplicación. También se debe subrayar que el diálogo necesita diseño instruccional para que la comparación de posturas sea una práctica de indagación, no una conversación sin criterios (Sampson & Clark, 2006; Sandoval & Millwood, 2007). Finalmente, la retroalimentación con criterios claros y orientaciones específicas, en la línea de Hattie & Timperley (2007), potencia el desarrollo de habilidades críticas; de allí la utilidad de rúbricas centradas en evidencia, garantías y refutación. Aunque los resultados deben leerse con cautela por el tamaño muestral reducido y la heterogeneidad entre perfiles, la consistencia de los resultados da confianza a los patrones descritos. Esto implica que, en próximos procesos de formación, conviene priorizar tres frentes muy concretos, instalar plantillas y bancos de principios para 1B, organizar la interacción con roles e iniciadores para 2A, y normalizar una secuencia de refutación con criterios para 2B.

Conclusiones

La estrategia de desarrollo profesional mostró que la argumentación científica puede avanzar del reconocimiento inicial a una evaluación crítica cuando la estructura de Toulmin (TAP) se hace explícita y la evaluación se alinea con criterios epistémicos. Los docentes evidenciaron progresos al transformar situaciones cotidianas en problemas científicos, articular evidencias con garantías y calificadores, y cerrar con refutaciones fundamentadas, aunque persisten retos para explicitar el vínculo evidencia–afirmación y para diseñar interacciones y refutaciones con criterios claros. Entre las implicaciones destacan la necesidad de visibilizar el TAP mediante organizadores y rúbricas por componente, estructurar la interacción dialógica con roles definidos y marcadores de réplica, y dar a las TIC un papel epistémico que permita la trazabilidad de evidencias, modelos y conclusiones. Las principales limitaciones se relacionan con el tamaño reducido de la muestra, la heterogeneidad de perfiles y la instrumentación parcial de algunas tareas, así como con la evaluación dicotómica de respuestas abiertas.

Como proyección, se propone validar y ampliar el uso de organizadores Toulmin y rúbricas en contextos más diversos, desarrollar bancos de principios y contraejemplos transferibles entre disciplinas, evaluar el impacto longitudinal en el aprendizaje estudiantil y profundizar en el potencial de las TIC para sostener la calidad y permanencia de la práctica argumentativa en educación superior.

Referencias bibliográficas

- Bağ, H., & Çalik, M. (2017). A thematic review of argumentation studies at the K–8 level. *Eğitim ve Bilim*, 42(190), 281–303. <https://doi.org/10.15390/EB.2017.6845>
- Briceño-Martínez, J. J. (2013). *La argumentación y la reflexión en los procesos de mejora de los profesores universitarios colombianos de ciencia en activo: aplicación de estrategias formativas sobre ciencia, aprendizaje y enseñanza* [Tesis doctoral]. Universidad de Granada.
- Briceño-Martínez, J. J., & Benarroch, A. (2012). Concepciones y creencias sobre ciencia, aprendizaje y enseñanza de profesores universitarios de ciencias. *REIEC*, 8(1), 24–41. <https://doi.org/10.54343/reiec.v8i1.122>
- Bricker, L. A., & Bell, P. (2008). Conceptualizations of argumentation from science studies and the learning sciences and their implications for the practices of science education. *Science Education*, 92(3), 473–498. <https://doi.org/10.1002/sce.20278>
- Cabrera Lanzo, N., Guàrdia Ortiz, L., & Sangrà, A. (2022). Desarrollo profesional de docentes universitarios en línea: un análisis desde las ecologías de aprendizaje. *Educar*, 58(2), 321–336. <https://doi.org/10.5565/rev/educar.1528>
- Chai, C. S., Koh, J. H. L., & Tsai, C.-C. (2013). A review of technological pedagogical content knowledge. *Educational Technology & Society*, 16(2), 31–51. <https://www.jstor.org/stable/jeductechsoci.16.2.31>
- Chen, Y.-C., Benus, M. J., & Hernandez, J. (2019). Managing uncertainty in scientific argumentation. *Science Education*, 103(5), 1235–1276. <https://doi.org/10.1002/sce.21527>

- Corbin, J. & Strauss, A. (2008). Fundamentos de la investigación cualitativa: Técnicas y procedimientos para el desarrollo de la teoría fundamentada (3.ª ed.). Sage Publications, Inc. <https://doi.org/10.4135/9781452230153>
- Darling-Hammond, L., Hyster, M. E., & Gardner, M. (2017). *Effective teacher professional development*. Learning Policy Institute. <https://learningpolicyinstitute.org/product/effective-teacher-professional-development-report>
- Driver, R., Newton, P., & Osborne, J. (2000). Establishing the norms of scientific argumentation in classrooms. *Science Education*, 84(3), 287–312. [https://doi.org/10.1002/\(SICI\)1098-237X\(200005\)84:3<287::AID-SCE1>3.0.CO;2-A](https://doi.org/10.1002/(SICI)1098-237X(200005)84:3<287::AID-SCE1>3.0.CO;2-A)
- Erduran, S., & Jiménez-Aleixandre, M. P. (Eds.). (2007). *Argumentation in science education: Perspectives from classroom-based research*. Springer. <https://doi.org/10.1007/978-1-4020-6670-2>
- Erduran, S., Ozdem, Y., & Park, J. Y. (2015). Research trends on argumentation in science education: A journal content analysis from 1998–2014. *International Journal of STEM Education*, 2(1). <https://doi.org/10.1186/s40594-015-0020-1>
- Erduran, S., Simon, S., & Osborne, J. (2004). TAPping into argumentation: Developments in the use of Toulmin's Argument Pattern in studying science discourse. *Science Education*, 88(6), 915–933. <https://doi.org/10.1002/sce.20012>
- Felton, M., Levin, D. M., De La Paz, S., & Butler, C. (2022). Scientific argumentation and responsive teaching: Using dialog to teach science in three middle-school classrooms. *Science Education*, 106(6), 1354–1374. <https://doi.org/10.1002/sce.21740>
- Fishman, E. J., Borko, H., Osborne, J., Gomez, F., Rafanelli, S., Reigh, E., Tseng, A., Million, S., & Berson, E. (2017). A practice-based professional development program to support scientific argumentation from evidence in the elementary classroom. *Journal of Science Teacher Education*, 28(3), 222–249. <https://doi.org/10.1080/1046560X.2017.1302727>
- Gibbs, G., & Coffey, M. (2004). The impact of training of university teachers on their teaching skills, their approach to teaching and the approach to learning of their students. *Active Learning in Higher Education*, 5(1), 87–100. <https://doi.org/10.1177/1469787404040463>
- González-Howard, M., & McNeill, K. L. (2019). Teachers' framing of argumentation goals: Working together to develop individual versus communal understanding. *Journal of Research in Science Teaching*, 56(6), 821–844. <https://doi.org/10.1002/tea.21530>
- Gray, R., & Kang, N. H. (2014). The structure of scientific arguments by secondary science teachers: Comparison of experimental and historical science topics. *International Journal of Science Education*, 36(1), 46–65. <https://doi.org/10.1080/09500693.2012.715779>
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81–112. <https://doi.org/10.3102/003465430298487>
- Jiménez-Aleixandre, M. P., & Díaz, J. (2003). Discurso de aula y argumentación en la clase de ciencias: cuestiones teóricas y metodológicas. *Investigación Didáctica*, 21(3), 359–370. <http://hdl.handle.net/10347/20698>
- Jiménez-Aleixandre, M. P., & Erduran, S. (2007). Argumentation in science education: An overview. En S. Erduran & M. P. Jiménez-Aleixandre (Eds.), *Argumentation in Science Education: Perspectives from Classroom-Based Research* (pp. 3–27). Springer. https://doi.org/10.1007/978-1-4020-6670-2_1

- Kelly, G. J., & Takao, A. (2002). Epistemic levels in argument: An analysis of university oceanography students' use of evidence in writing. *Science Education*, 86(3), 314–342. <https://doi.org/10.1002/sce.10024>
- Kim, S., & Hand, B. (2015). An analysis of argumentation discourse patterns in elementary teachers' science classroom discussions. *Journal of Science Teacher Education*, 26(3), 221–236. <https://doi.org/10.1007/s10972-014-9416-x>
- Knight, P. (2002). *Being a Teacher in Higher Education*. McGraw-Hill Education.
- Larraín, A., Moreno, C., Grau, V., Freire, P., Salvat, I., López, P., & Silva, M. (2017). Curriculum materials support teachers in the promotion of argumentation in science teaching: A case study. *Teaching and Teacher Education*, 67, 522–537. <https://doi.org/10.1016/j.tate.2017.07.018>
- McNeill, K. L., González-Howard, M., Katsh-Singer, R., & Loper, S. (2016). Pedagogical content knowledge of argumentation: Using classroom contexts to assess high-quality PCK rather than pseudoargumentation. *Journal of Research in Science Teaching*, 53, 261–290. <https://doi.org/10.1002/tea.21252>
- McNeill, K. L., & Knight, A. M. (2013). Teachers' pedagogical content knowledge of scientific argumentation: The impact of professional development on K–12 teachers. *Science Education*, 97(6), 936–972. <https://doi.org/10.1002/sce.21081>
- Molina, M. E. (2012). Argumentar en clases de ciencias naturales: una revisión bibliográfica. *Actas III Jornadas de Enseñanza e Investigación Educativa en el Campo de las Ciencias Exactas y Naturales*, 553–554. <http://sedici.unlp.edu.ar/handle/10915/32091>
- Nielsen, J. A. (2011). Dialectical features of students' argumentation: A critical review of argumentation studies in science education. *Research in Science Education*, 43(1), 371–393. <https://doi.org/10.1007/s11165-011-9266-x>
- Osborne, J., Erduran, S., & Simon, S. (2004). Enhancing the quality of argumentation in school science. *Journal of Research in Science Teaching*, 41(10), 994–1020. <https://doi.org/10.1002/tea.20035>
- Osborne, J., Simon, S., Christodoulou, A., Howell-Richardson, C., & Richardson, K. (2013). Learning to argue: A study of four schools and their attempt to develop the use of argumentation as a common instructional practice and its impact on students. *Journal of Research in Science Teaching*, 50(3), 315–347. <https://doi.org/10.1002/tea.21073>
- Ozdem-Yilmaz, Y., Cakiroglu, J., Ertepinar, H., & Erduran, S. (2017). The pedagogy of argumentation in science education: Science teachers' instructional practices. *International Journal of Science Education*, 39(11), 1443–1464. <https://doi.org/10.1080/09500693.2017.1336807>
- Plantin, C. (1998). *La argumentación*. Ariel.
- Plantin, C. (2009). A place for figures of speech in argumentation theory. *Argumentation*, 23, 325–337. <https://doi.org/10.1007/s10503-009-9152-0>
- Sampson, V., & Blanchard, M. R. (2012). Science teachers and scientific argumentation: Trends in views and practice. *Journal of Research in Science Teaching*, 49(9), 1122–1148. <https://doi.org/10.1002/tea.21037>
- Sampson, V., & Clark, D. B. (2006). Assessment of argument in science education: A critical review of the literature. In S. A. Barab, K. E. Hay, & D. Hickey (Eds.), *Proceedings of the 7th International Conference of the Learning Sciences (ICLS)* (pp. 655–661). International Society of the Learning Sciences.
- Sampson, V., Grooms, J., & Walker, J. P. (2011). Argument-Driven Inquiry as a way to help students learn how to participate in scientific argumentation and craft written arguments: An exploratory study. *Science Education*,

95(2), 217–257. <https://doi.org/10.1002/sce.20421>

Sandoval, W. A. (2003). Conceptual and epistemic aspects of students' scientific explanations. *Journal of the Learning Sciences*, 12(1), 5–51. https://doi.org/10.1207/S15327809JLS1201_2

Sandoval, W., & Millwood, K. (2007). What Can Argumentation Tell Us About Epistemology?. En Erduran, S., Jiménez-Aleixandre, M.P. (Ed.) *Argumentation in Science Education*, (pp. 71-88). Springer, Dordrecht. https://doi.org/10.1007/978-1-4020-6670-2_4

Toulmin, S. (2007). *Los usos de la argumentación*. Península

Walton, D. (1996). *Argumentation schemes for presumptive reasoning*. Routledge. <https://doi.org/10.4324/9780203811160>

Wess, R., Priemer, B., & Parchmann, I. (2023). Professional development programs to improve science teachers' skills in the facilitation of argumentation in science classroom—A systematic review. *Disciplinary and Interdisciplinary Science Education Research*, 5(9), 1–22. <https://doi.org/10.1186/s43031-023-00076-3>

Zohar, A. (2007). Science Teacher Education and Professional Development in Argumentation. In M. del P. Jiménez-Aleixandre & S. Erduran (Ed.), *Argumentation in Science Education: Perspectives from Classroom-Based Research* (pp. 245–269). Springer. https://doi.org/10.1007/978-1-4020-6670-2_12