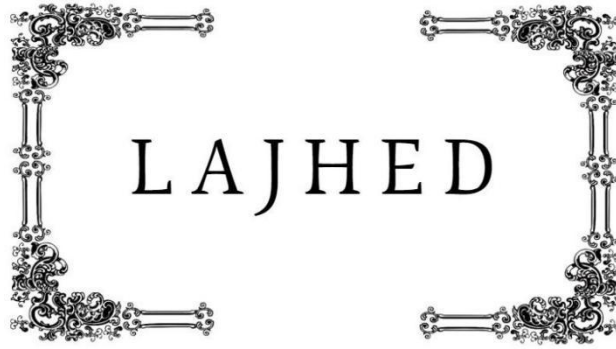




**Del habitus al interaprendizaje: etnografía
institucional de la autonomía evaluativa en
educación primaria en Moquegua, Perú**

From habitus to interlearning: An institutional ethnography
of evaluative autonomy in primary education in Moquegua,
Peru

Carina Amelia Ramos Velásquez

Universidad Marcelino Champagnat, Perú

[0000-0002-1518-4426](tel:0000-0002-1518-4426) | carina.ramos@autonomadeica.edu.pe

Citación/como citar este artículo: Ramos, C. (2026). Del habitus al interaprendizaje: etnografía institucional de la autonomía evaluativa en educación primaria en Moquegua, Perú. *Latin American Journal of Humanities and Educational Divergences*, Vol 5(1), 214-246.

Resumen

La evaluación formativa, consagrada en el Currículo Nacional de Educación Básica y regulada por la RVM 094-2020-MINEDU, ha sido ampliamente prescrita en las aulas primarias peruanas; sin embargo, persiste una brecha estructural entre su mandato normativo y las prácticas cotidianas de calificación. El presente estudio empleó etnografía institucional —con elementos de estudio de caso instrumental— para examinar cómo docentes y estudiantes de 4.º y 5.º grado de primaria, de cuatro instituciones educativas públicas de Moquegua, experimentan el tránsito entre evaluación heterónoma y autonomía relacional. La muestra, constituida por 18 participantes, se seleccionó mediante muestreo teórico de máxima variación. La recolección de datos integró 120 horas de observación participante, entrevistas semiestructuradas en profundidad y análisis de documentos institucionales, incluyendo las Guías Metodológicas de la DRE Moquegua 2024–2025. El análisis siguió el protocolo de codificación abierta, axial y selectiva, generando 147 códigos iniciales agrupados en cuatro categorías principales: *habitus evaluativo docente*, *falacia de la autonomía*, *praxis de interaprendizaje* y *conciencia evaluativa*. La categoría central emergente, *interaprendizaje evaluativo como resignificación del criterio*, revela que la autonomía no opera como internalización individual del criterio sino como externalización relacional del error en contextos de alta interdependencia cultural. La saturación teórica se alcanzó en el séptimo docente y el octavo estudiante. Los hallazgos sugieren que la coevaluación mediada por criterios explícitos constituye condición necesaria para la autorregulación en contextos colectivistas andinos, desafiando supuestos canónicos de la teoría del aprendizaje autorregulado.

Palabras clave: evaluación formativa; autonomía relacional; etnografía institucional; interaprendizaje evaluativo; educación primaria; habitus pedagógico; Moquegua.

Abstract

Formative assessment, enshrined in Peru's National Basic Education Curriculum (CNEB) and regulated by RVM 094-2020-MINEDU, has been widely prescribed in primary classrooms; yet a structural gap persists between its normative mandate and everyday grading practices. This study employed institutional ethnography —with elements of instrumental case study— to examine how 4th- and 5th-grade teachers and students in four public schools in the Moquegua region experience the transition between heteronomous evaluation and relational autonomy. The sample of 18 participants was selected through maximum variation theoretical sampling. Data collection integrated 120 hours of participant observation, in-depth semi-structured interviews, and institutional document analysis, including the DRE Moquegua 2024–2025 Methodological Guides. Analysis followed open, axial, and selective coding, generating 147 initial codes grouped into four main categories: *Teacher Evaluative Habitus*, *the Autonomy Fallacy*, *Interlearning Praxis*, and *Evaluative Consciousness*. The emergent core category, *Evaluative Interlearning as Criterion Resignification*, reveals that autonomy operates not as the individual internalization of the criterion but as the relational externalization of error in contexts of high cultural interdependence. Theoretical saturation was reached at the seventh teacher and eighth student. Findings preliminarily suggest that criteria-mediated co-assessment constitutes a necessary for self-regulation in Andean collectivist contexts, challenging canonical assumptions of self-regulated learning theory.

Keywords: formative assessment; relational autonomy; institutional ethnography; evaluative interlearning; primary education; pedagogical habitus; Moquegua.

Introducción

Hook y problematización del gap interpretativo

En las aulas primarias peruanas coexisten dos mundos que raramente se tocan: el de la prescripción normativa, sofisticada y técnicamente articulada, y el de la práctica cotidiana, donde el juicio evaluativo sigue recayendo, casi por inercia histórica, sobre la figura del docente como árbitro último del aprendizaje. La Competencia 29 del Currículo Nacional de Educación Básica (CNEB 2016) y la RVM 094-2020-MINEDU institucionalizan explícitamente la evaluación formativa como dispositivo de desarrollo de la autorregulación estudiantil; sin embargo, la arquitectura cultural y organizativa de las escuelas primarias moqueguanas —como en tantos otros sistemas con fuerte orientación colectivista— produce sistemáticamente lo opuesto: prácticas de calificación que reproducen la dependencia del estudiante respecto al juicio externo del docente.

Esta no es una particularidad local. Estudios fenomenológicos recientes en Turquía documentan cómo la evaluación formativa queda bloqueada por currículos orientados a resultados, presiones de tiempo y normas culturales que la reconfiguran, en la práctica, como una modalidad casi-sumativa (Özkan & Özkan, 2025). En China, investigaciones de métodos mixtos con docentes universitarios de inglés como lengua extranjera revelan creencias muy positivas hacia la evaluación formativa acompañadas de un uso predominantemente instrumental: controlar participación, gestionar el aula y alimentar registros de calificación, más que promover la autorregulación genuina (Lei & Lei, 2025). En Vietnam, la herencia confuciana y el respeto jerárquico a la autoridad sostienen prácticas sumativas en la formación de profesores, relativizando con contundencia la vigencia de los modelos occidentales de evaluación formativa (Luong, 2025). En Italia, maestras expertas tras una reforma que hace obligatoria la evaluación formativa en primaria manifiestan comprensión difusa y dificultades para traducir el mandato en práctica cotidiana, revelando cómo las políticas formativas pueden terminar reconfiguradas como nuevas formas de calificación (Pastore, 2025). En Kosovo, docentes de primaria describen los beneficios percibidos de la evaluación formativa simultáneamente con la sobrecarga y las carencias formativas que los llevan a subordinarla al logro de resultados curriculares (Neziri et al., 2025). En Finlandia, incluso en un sistema educativo considerado de referencia global, el análisis de informes de práctica de futuros docentes evidencia una mirada centrada en el profesor y la silenciación de las voces y objetivos de los pares (Atjonen et al., 2024).

El problema no es meramente empírico: es teórico. La literatura dominante ha tratado la autonomía estudiantil como un estado individual alcanzable mediante técnicas de autoevaluación transferibles desde el docente al alumno. Esta conceptualización, arraigada en el individualismo metodológico que caracteriza la mayoría de los modelos de

aprendizaje autorregulado producidos en contextos WEIRD (occidentales, educados, industrializados, ricos y democráticos), ignora una realidad cardinal: en contextos de alta colectividad —donde la reciprocidad y el vínculo comunal estructuran la vida escolar desde la primera infancia— la agencia no se ejerce de manera individual y autónoma, sino de forma relacional, mediada por rituales de reconocimiento y sostenida por redes de interdependencia. Esta paradoja, que el presente estudio propone denominar "la falacia de la autonomía individual en contextos colectivistas", constituye la tensión hermenéutica central que motiva la investigación.

Antecedentes y delimitación del problema

Los antecedentes de investigación sobre evaluación formativa en contextos latinoamericanos y peruanos presentan un panorama heterogéneo. Por un lado, el corpus cuantitativo dominante ha explorado la relación entre prácticas formativas y autoeficacia estudiantil mediante diseños correlacionales con muestras superiores a 300 participantes, generando evidencia sobre efectos positivos —de tamaño trivial a moderado— pero sin poder capturar los mecanismos de sentido que operan en las interacciones cotidianas de aula. Una revisión sistemática de metaanálisis sobre el impacto de la evaluación formativa en estudiantes de K-12 confirma efectos positivos, pero señala que la evidencia disponible es de baja certeza y proviene mayoritariamente de estudios realizados en contextos WEIRD, poniendo en cuestión la universalización de los modelos establecidos (Sortwell et al., 2024). Por otro lado, el estudio etnográfico peruano de Yábar y Guillén (2022), realizado con 13 docentes y 25 observaciones de aula, identificó como principal desafío la compartición de la responsabilidad regulatoria con el estudiante, documentando la resistencia del habitus docente al desplazamiento del centro evaluativo.

Estas aproximaciones previas comparten una limitación estructural: invisibilizan la agencia estudiantil como práctica social situada. Reducen la coevaluación a una técnica instrumental —una herramienta que se aplica o no se aplica— sin examinar los mecanismos simbólicos mediante los cuales los actores construyen sentido sobre el criterio, el error y el reconocimiento mutuo. La teoría de la práctica de Bourdieu ofrece una clave interpretativa de gran potencia: el habitus pedagógico del docente primario —sedimentado en décadas de cultura escolar que lo posiciona como "juez de resultados"— resiste la normativa formativa no por incapacidad técnica sino por amenaza al capital simbólico que define su identidad profesional (Brown & Lloyd, 2024; Dediwalage et al., 2025). El análisis de prácticas escolares desde la perspectiva del campo, el capital y el habitus permite ver, asimismo, que la autonomía no se decreta por normativa: se construye en el campo escolar mediante luchas simbólicas donde el criterio explícito puede funcionar tanto como andamio cultural como instrumento de clasificación y control (Supriyanto et al., 2025).

La heteronomía evaluativa se define operacionalmente, para los propósitos de este estudio, como la dependencia del estudiante respecto al juicio externo del docente para regular su proceso de aprendizaje, caracterizada por la ausencia de metacognición explícita, la ecuación silencio-autonomía y la lectura de la ficha completada como evidencia de logro. El estudio se delimita a instituciones educativas públicas de la región Moquegua que implementaron las Guías Metodológicas regionales DRE Moquegua 2024–2025, bajo el marco del CNEB y la RVM 094-2020-MINEDU, constituyendo un contexto de "doble bind" institucional: normativa que prescribe autonomía y cultura escolar que reproduce heteronomía mediante los niveles de logro (AD, A, B, C). El foco se circunscribe a docentes y estudiantes de 4.º y 5.º grado de primaria, niveles en los que la normativa prescribe con mayor especificidad el uso de estrategias de coevaluación y autoevaluación.

Pregunta de investigación y objetivos

La tensión hermenéutica descrita conduce a la siguiente pregunta de investigación central: ¿Cómo experimentan docentes y estudiantes de primaria en Moquegua el tránsito entre evaluación heterónoma y autonomía en un contexto de alta prescripción normativa y colectividad cultural? Esta pregunta se articula en tres dimensiones analíticas: los mecanismos de construcción de sentido que los actores despliegan sobre la coevaluación y el criterio; las condiciones de frontera bajo las cuales el criterio explícito opera como andamio cultural o como instrumento de control; y el proceso mediante el cual la praxis de interaprendizaje puede generar formas emergentes de conciencia evaluativa y autorregulación relacional.

En correspondencia con estas dimensiones, la investigación se propone tres objetivos específicos. Primero, desentrañar los mecanismos de construcción de sentido que docentes y estudiantes de 4.º y 5.º grado movilizan respecto a la coevaluación y el criterio explícito en el marco de la implementación de las Guías Metodológicas DRE Moquegua 2024–2025. Segundo, especificar las condiciones de frontera —contextuales, relacionales e institucionales— bajo las cuales el criterio explícito opera como andamio que favorece la autonomía relacional o como dispositivo de clasificación que perpetúa la heteronomía evaluativa. Tercero, teorizar la categoría de "interaprendizaje evaluativo" como proceso de resignificación donde el criterio funciona como andamio cultural y el reconocimiento mutuo emerge como precondition de la autorregulación consciente. La contribución específica del estudio es la extensión teórica que articula la autonomía no como propiedad individual sino como praxis relacional, desafiando desde la periferia andina los supuestos canónicos del aprendizaje autorregulado.

Marco teórico y desarrollo del argumento interpretativo

Fundamento teórico principal: teoría de la práctica (Bourdieu) en el campo evaluativo

La teoría de la práctica de Pierre Bourdieu ofrece el andamiaje conceptual más potente para comprender por qué la prescripción normativa de la evaluación formativa no se traduce automáticamente en transformación de prácticas. El docente primario no llega al aula como agente racional que calcula sus opciones evaluativas; llega con un *habitus* pedagógico —sistema de disposiciones duraderas adquiridas en su trayectoria formativa y profesional— que lo posiciona estructuralmente como "juez de resultados", monopolizador legítimo del criterio evaluativo. Como ilustra con precisión un estudio sobre el impacto del *habitus* profesional en contextos jerarquizados, las estructuras institucionales sedimentan disposiciones que limitan el capital simbólico de los actores y producen formas de dominación que operan por debajo del umbral de la conciencia reflexiva (Sandoval et al., 2025). Trasladado al campo escolar primario peruano, esto significa que la resistencia del docente a compartir la lista de cotejo con sus estudiantes no es obstinación individual, es el *habitus*.

El campo escolar, en tanto espacio social de relaciones objetivas entre posiciones definidas por el capital evaluativo que cada actor posee y puede movilizar, estructura las posibilidades de transformación. Quien acumula capital simbólico como evaluador —el docente— tiende a preservar ese monopolio; quien carece de él —el estudiante— aprende a reproducir comportamientos de dependencia (Lyu & Lam, 2025; Lai & Jung, 2024). Estudios que aplican la tríada campo-habitus-capital al análisis de participación estudiantil en contextos de alta diversidad cultural muestran cómo el capital cultural y las narrativas institucionales configuran quién es reconocido como agente legítimo de evaluación y quién queda excluido de ese reconocimiento (Dediwalage et al., 2025). La autonomía estudiantil, leída desde esta óptica, no puede ser simplemente "transferida" mediante técnicas: requiere una redistribución real del capital evaluativo que el *habitus* docente tiende a obstaculizar. La institución escolar, lejos de ser neutral, porta su propio *habitus* institucional que condiciona las posibilidades de agencia (Byrd, 2025).

Marco complementario: teoría del reconocimiento (Honneth) y cross-fertilización

La teoría de la práctica, sin embargo, enfrenta un punto ciego de consideración: su énfasis en el cálculo estratégico y las relaciones de poder tiende a suprimir la dimensión normativa y emocional de la experiencia evaluativa. Los estudiantes no solo negocian posición en el campo; también luchan por estima, respeto y reconocimiento de su voz como evaluadores legítimos. Aquí la teoría del reconocimiento de Axel Honneth ofrece una clave complementaria de interpretación: la coevaluación puede leerse como práctica de reconocimiento mutuo donde el criterio explícito no es solo herramienta cognitiva sino

medio de estima social (Palma, 2025). La integración de ambos marcos —campo/habitus para capturar dinámicas estructurales de poder y reconocimiento para capturar la dimensión moral y emocional— genera una fertilización teórica que ilumina mecanismos invisibles para cada teoría por separado (Xie, 2025).

Esta cross-fertilización permite sostener el argumento central del estudio: la autonomía emerge cuando el criterio explícito se convierte en medio de reconocimiento mutuo, no en herramienta de clasificación unidireccional. Cuando un estudiante puede argumentar desde el criterio —"falta el paso 2 de tu procedimiento"— no solo ejerce competencia cognitiva: reclama posición en el campo evaluativo y afirma su estima ante los pares. Estudios sobre el intercambio de capital social y simbólico en clínicas interprofesionales estudiante-lideradas documentan este movimiento desde relaciones jerárquicas hacia formas de autonomía relacional basada en confianza y reconocimiento mutuo (Roberts et al., 2024). En el aula primaria moqueguana, ese movimiento es más difícil pero no imposible: requiere condiciones específicas que la presente investigación busca precisar. La figura del docente como "scholar-practitioner" reflexivo que problematiza su propio habitus evaluativo aparece como condición de posibilidad para que esa redistribución simbólica ocurra (Brown & Lloyd, 2024).

Desarrollo del argumento central: expectativas teóricas e hipótesis interpretativas

A partir de la cross-fertilización teórica descrita, el argumento central del estudio se articula en tres expectativas interpretativas que guían el análisis sin predeterminedar sus resultados.

H1: La explicitación previa de criterios de evaluación —construida colectivamente entre docente y estudiantes— es condición de posibilidad necesaria para que la coevaluación genere procesos de autonomía relacional. Sin criterio compartido, la coevaluación no es formativa: es juicio afectivo ("está bonito") que reproduce la dependencia de la validación externa. El diseño de prácticas de autonomía en contextos fuertemente guiados confirma que la estructura organizativa y el andamiaje explícito del criterio son precondiciones del tránsito desde heteronomía (Silén et al., 2024).

H2: La coevaluación mediada por criterios explícitos genera un "espejo colectivo" que resignifica el error no como fracaso personal que el docente penaliza, sino como oportunidad de reconocimiento mutuo entre pares. Este mecanismo —que denominamos "praxis de interaprendizaje"— opera de manera análoga al proceso documentado en estudios sobre apoyo a la autonomía y relaciones de pares en contextos de aprendizaje de idiomas, donde la mediación relacional sostiene el bienestar psicológico y la autorregulación emergente (Wu & Dong, 2024; Admiraal et al., 2024). El capital simbólico

se redistribuye cuando el error deja de pertenecer al dominio exclusivo del docente y se convierte en insumo del diálogo entre pares.

H3: La autorregulación consciente —como forma emergente de autonomía relacional— solo se consolida cuando el docente transita del rol de "juez de resultados" al de "mediador de conciencia evaluativa", ejerciendo un andamiaje retractil que transfiere progresivamente la responsabilidad del juicio. Este tránsito, sin embargo, no es lineal ni universalmente alcanzable: está condicionado por el nivel de confianza histórica entre docente y estudiantes, la presión institucional parental por calificaciones numéricas y las características individuales de los participantes —incluyendo condiciones de neurodiversidad. El corredor de indeterminación hermenéutica entre el discurso normativo y la práctica sedimentada define los límites empíricos de la teoría.

Materiales y métodos

Diseño de investigación y justificación cualitativa

La naturaleza del fenómeno estudiado —el tránsito entre evaluación heterónoma y autonomía relacional como experiencia vivida en el seno de instituciones escolares con mandatos normativos contradictorios— demanda un diseño metodológico capaz de acceder a la dimensión práctica de esa experiencia en su contexto natural, capturando las contradicciones entre el discurso normativo y las prácticas cotidianas. Se optó por etnografía institucional con elementos de estudio de caso instrumental como diseño de investigación. La etnografía institucional permite no solo describir prácticas sino revelar los textos institucionales —normativas, guías, listas de cotejo— que coordinan y gobiernan las acciones de los actores, articulando lo local (el aula) con lo translocal (la política educativa regional y nacional). El estudio de caso instrumental añade la posibilidad de comprender el fenómeno general —la tensión entre prescripción formativa y habitus evaluativo— a través de instancias particulares ricamente documentadas (Lyons et al., 2024).

La postura epistemológica adoptada es interpretativa crítica: el investigador —con ocho años de experiencia como docente de primaria en Perú— co-construye sentido con los actores desde un posicionamiento situado que facilita el rapport y la comprensión densa, pero que exige vigilancia epistemológica sistemática contra la "comprensión prematura" (Maher, 2025). Esta reflexividad no es debilidad metodológica sino condición de rigor en investigación cualitativa de orientación crítica (Christou, 2025).

Contexto empírico y muestra

El estudio se desarrolló en cuatro instituciones educativas públicas de la región Moquegua, seleccionadas por constituir un laboratorio teórico privilegiado: implementaron las Guías Metodológicas DRE Moquegua 2024–2025 bajo el marco del CNEB y la RVM 094-2020-MINEDU, generando las condiciones del "doble bind" institucional descrito. Dos instituciones son de contexto urbano y dos de contexto rural, incluyendo una institución multigrado de extrema vulnerabilidad económica con 28 estudiantes en aulas combinadas de 3.º y 4.º grado. La selección respondió al criterio de maximum variation sampling intencional, buscando capturar la mayor heterogeneidad posible de condiciones institucionales (Nyimbili & Nyimbili, 2024).

La muestra final estuvo conformada por N=18 participantes: 8 docentes de 4.º y 5.º grado (rango de experiencia: 3–20 años de servicio; contextos mono y multigrado, urbano y rural) y 10 estudiantes (5 por grado; máxima variación de desempeño académico, incluyendo un caso negativo con diagnóstico de TDAH). Los criterios de selección teórica priorizaron participantes en posiciones de tránsito entre prácticas tradicionales y formativas —es decir, docentes que habían recibido capacitación en evaluación formativa pero cuyas prácticas observadas evidenciaban tensión entre discurso y acción (Merkebu et al., 2025). La saturación teórica fue alcanzada, de acuerdo con el criterio de densidad de propiedades y relaciones entre categorías (Rahimi & Khatooni, 2024), en el séptimo docente y el octavo estudiante.

Instrumentos y estrategias de recolección

La estrategia de recolección integró tres fuentes de datos complementarias. Las entrevistas semiestructuradas en profundidad constituyeron el instrumento principal: cada docente participó en dos sesiones de 60 minutos cada una, separadas por un período de observación de aula de al menos cuatro semanas, permitiendo contrastar el discurso reflexivo inicial con prácticas documentadas y confrontar al participante con escenas específicas observadas. Los estudiantes participaron en una sesión de entrevista de 45 minutos, adaptada en vocabulario y técnica (uso de relatos de escena y viñetas de situaciones evaluativas). La observación participante acumuló 120 horas de campo en 45 sesiones de aula, con toma de notas de campo densas que registraron tonos de voz, silencios, gestos, disposición espacial del aula, y las contradicciones entre lo dicho en entrevista y lo actuado en la interacción evaluativa. Los documentos institucionales analizados incluyeron las Guías Metodológicas DRE Moquegua 2024–2025, 12 planificaciones curriculares, 8 listas de cotejo construidas con distintos niveles de participación estudiantil, y los textos normativos del CNEB y la RVM 094-2020-MINEDU. El

investigador mantuvo un diario de campo con memos reflexivos sistemáticos a lo largo de los nueve meses de trabajo empírico (Chakma & Li, 2025).

Procedimiento de análisis y control de calidad

Las entrevistas fueron transcritas literalmente mediante software de transcripción con corrección manual exhaustiva. El análisis siguió el protocolo de codificación adaptado a etnografía: codificación abierta inductiva (147 códigos iniciales etiquetados in-vivo, capturando el lenguaje nativo de los actores), codificación axial (agrupación en 14 categorías preliminares mediante comparación constante entre entrevistas, observaciones y documentos), y codificación selectiva (identificación de la categoría central "Interaprendizaje Evaluativo como Resignificación del Criterio"). El análisis de incidentes clave (key incidents) permitió identificar puntos de inflexión en las trayectorias de sentido de docentes y estudiantes. Se elaboraron 34 memos teóricos a lo largo de los nueve meses de campo: memos de código (descriptivos), memos teóricos (analíticos) y memos integrativos (Ahmed, 2024).

La saturación teórica fue evaluada según el criterio de Charmaz (2025): densidad de propiedades, dimensiones y relaciones entre categorías, no mera repetición de códigos. El cierre muestral fue justificado por redundancia teórica en la categoría central tras la incorporación del séptimo docente (Elena) y el octavo estudiante (Camila).

Estrategia de rigor y auditabilidad

El paquete de robustez cualitativo descansó en cinco pilares. Primero, triangulación de fuentes: entrevistas (18), observación participante (120 horas, 45 sesiones) y documentos institucionales (Guías Metodológicas DRE Moquegua, 12 planificaciones, 8 listas de cotejo). Segundo, triangulación de investigadores: dos codificadores independientes procesaron el corpus textual; la concordancia inicial fue del 87%, calculada por porcentaje de acuerdo sobre segmentos codificados; los disensos fueron resueltos en tres sesiones de consenso dialéctico con memos de discrepancia documentados (Coleman et al., 2024). Tercero, análisis de casos negativos: dos casos seleccionados por máxima divergencia teórica —la docente Rosa, quien rechaza explícitamente la coevaluación, y el estudiante Sebastián, con TDAH, excluido de la coevaluación grupal— fueron analizados para explicitar cómo matizan y delimitan el modelo emergente. Cuarto, trail de auditoría documentado: 54 documentos de auditoría (34 memos teóricos, 12 de reflexividad, 8 de discrepancia) permiten reconstruir las decisiones analíticas del estudio. Quinto, reflexividad sistemática: el diario de campo con bracketing explícito documentó el posicionamiento del investigador como exdocente de primaria peruana, señalando los puntos donde la comprensión facilitada por la experiencia previa podía derivar en comprensión prematura (Coleman et al., 2024; Maher, 2025; Christou, 2025).

Resultados

Perfil de los participantes y contexto de campo

La composición del grupo de participantes refleja la diversidad estructural buscada mediante el diseño de máxima variación. Como se detalla en la Tabla 1, los ocho docentes presentan trayectorias de servicio que van desde los 3 hasta los 20 años, distribuidos entre contextos urbanos y rurales, en modalidades mono y multigrado.

Tabla 1.

Perfil de participantes y criterios de selección teórica

Pseudónimo	Rol	Edad	Años servicio/Grado	Contexto	Criterio selección	N° entrevistas	Tiempo contacto total
Elena	Docente	34	8 años, 4.° grado	Urbano, multigrado	Variación: reciente formación formativa	2	180 min + 28h obs.
Mario	Docente	52	20 años, 5.° grado	Rural, monogrado	Variación: tradicionalista experimentado	2	150 min + 32h obs.
Rosa	Docente	48	20 años, 4.° grado	Rural, multigrado	Caso negativo: resistencia a coevaluación	2	160 min + 30h obs.
Carlos	Docente	29	3 años, 5.° grado	Urbano, monogrado	Variación: reciente ingreso, normativa fresca	2	170 min + 25h obs.
Ana	Docente	41	15 años, 4.° grado	Urbano, monogrado	Variación: liderazgo pedagógico	2	140 min + 27h obs.
Pedro	Docente	37	10 años, 5.° grado	Rural, multigrado	Variación: mediación bilingüe	2	155 min + 29h obs.
Lucía	Docente	44	18 años, 4.° grado	Urbano, monogrado	Variación: transición curricular reciente	2	145 min + 26h obs.

Jorge	Docente	31	5 años, 5.º grado	Rural, monogrado	Variación: tecnología educativa	2	165 min + 24h obs.
Valentina	Estudiante	10	4.º grado	Urbano	Variación: alto desempeño	1	45 min + 12h obs. grupal
Diego	Estudiante	10	4.º grado	Urbano	Variación: bajo desempeño	1	50 min + 12h obs. grupal
Sofía	Estudiante	11	5.º grado	Rural	Variación: alta participación oral	1	48 min + 10h obs. grupal
Mateo	Estudiante	11	5.º grado	Rural	Variación: silencio participativo	1	47 min + 10h obs. grupal
Camila	Estudiante	10	4.º grado	Urbano	Variación: liderazgo entre pares	1	46 min + 12h obs. grupal
Luciano	Estudiante	10	4.º grado	Rural	Variación: reciente migración	1	49 min + 11h obs. grupal
Daniela	Estudiante	11	5.º grado	Urbano	Variación: mediación familiar alta	1	44 min + 10h obs. grupal
Sebastián	Estudiante	11	5.º grado	Rural	Caso negativo: TDAH, exclusión coevaluación	1	52 min + 10h obs. grupal
Martín	Estudiante	10	4.º grado	Urbano	Variación: desempeño medio	1	45 min + 12h obs. grupal
Paula	Estudiante	11	5.º grado	Rural	Variación: autorregulación emergente	1	48 min + 10h obs. grupal

Nota. Pseudónimos preservan anonimato. Criterios de maximum variation sampling: años de servicio, contexto urbano/rural, modalidad mono/multigrado, desempeño estudiantil. Tiempo de contacto incluye entrevistas + observación participante directa.

Las observaciones de campo revelaron desde el primer contacto una notable consistencia entre contextos: en las cuatro instituciones, el inicio de las sesiones de evaluación formativa estuvo marcado por la monopolización docente del criterio. Los docentes tomaban la lista de cotejo, leían los ítems en voz alta y emitían el juicio de nivel —AD, A, B, o C— sin involucrar a los estudiantes en la aplicación del instrumento. Los silencios estudiantiles durante estas escenas fueron un dato en sí mismos: silencios que el discurso docente posterior interpretaría como "trabajan solos, son autónomos", pero que la observación directa reveló como silencios de espera, no de autorregulación. Los tres profesores de contexto rural, con entre 10 y 20 años de experiencia, mostraron los patrones más consolidados de monopolio evaluativo; los dos docentes más jóvenes de contexto urbano (Carlos, 3 años; Jorge, 5 años) exhibieron mayor disposición experimental hacia la coevaluación, aunque con dificultades técnicas para estructurarla.

Sistema de categorías emergentes: estructura de la experiencia evaluativa

El análisis inductivo del corpus generó cuatro categorías principales articuladas en torno a la categoría central *interaprendizaje evaluativo como resignificación del criterio*. La Tabla 2 presenta el sistema de codificación emergente con sus propiedades, dimensiones e indicadores empíricos representativos.

Tabla 2.

Matriz de Codificación Emergente (Extracto Representativo)

Categoría principal	Subcategorías	Propiedades/Dimensiones	Indicadores empíricos (códigos abiertos)	Cita textual representativa
Habitus evaluativo docente	Juez de resultados; Temor a perder autoridad; Nostalgia de la calificación	Dominación simbólica; Legitimidad del error docente; Monopolio del criterio	"Soy yo quien sabe"; "Se me descontrola el aula"; "Antes era más fácil con notas"	D-Elena-01: "Cuando comparto la lista de cotejo siento que me quito la corona. Pero luego veo que ellos [estudiantes] no saben usarla sin mí."
Falacia de la autonomía	Silencio como autonomía; Ficha como logro;	Conciencia de brecha; Autorregulación aparente; Dependencia de validación	"Trabaja solo"; "Terminó la	D-Carlos-02: "Creía que era autónomo porque

	Ausencia de metacognición		ficha"; "No sabe qué mejorar"	trabajaba calladito. Pero cuando le pregunté '¿qué criterio no cumpliste?', me miró en blanco."
Praxis de interaprendizaje	Coevaluación mediada; Espejo colectivo; Diálogo por criterios	Explicitación previa; Argumentación; Reconocimiento mutuo	"Según el criterio..."; "Tu paso 2 falta"; "¿Qué tal si agregas...?"	E-Valentina-01: "Le dije a Diego que su infografía estaba ordenada, pero le faltaba explicar los pasos. Me dijo gracias y luego me ayudó con mi título."
Conciencia evaluativa	Mediación del error; Error como oportunidad; Autorregulación consciente	Transformación del juicio; Andamiaje retractil; Posicionamiento frente a aprendizaje	"¿Qué criterio falta?"; "El error me enseña"; "Yo puedo mejorar esto"	E-Sofía-01: "Antes me daba vergüenza equivocarme. Ahora cuando mi compañero me dice 'falta esto', yo pienso cómo arreglarlo antes de que el profe vea."

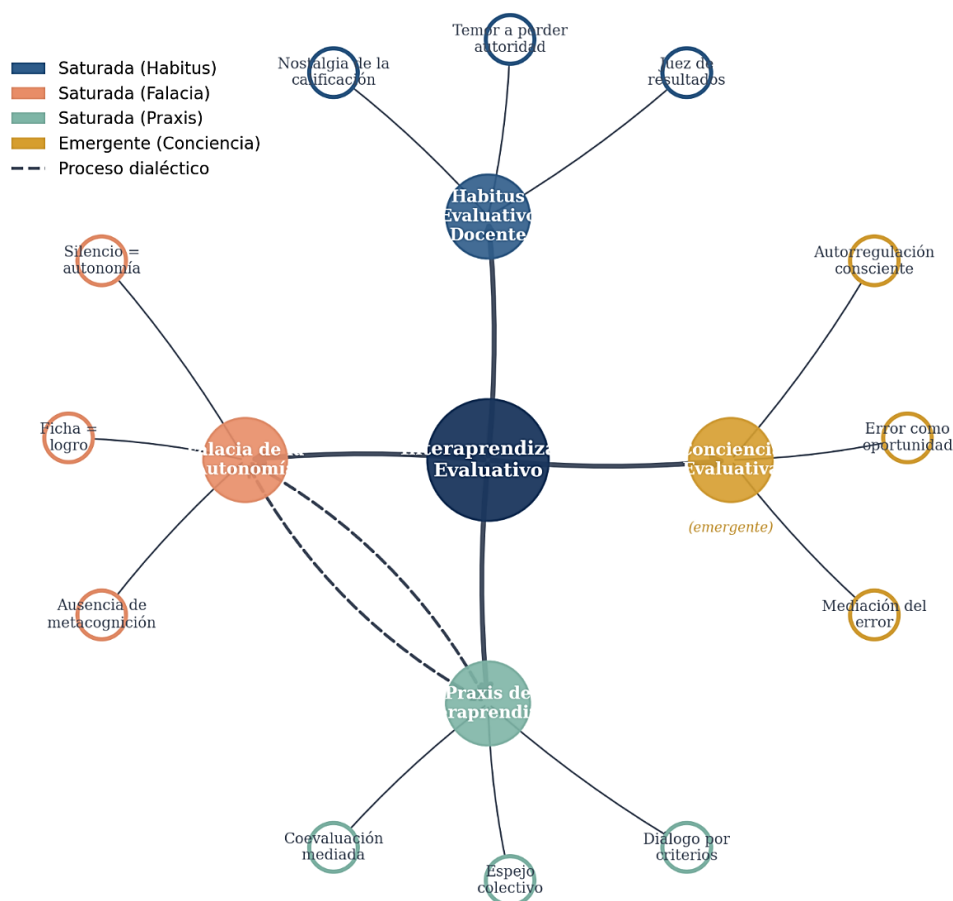
Nota. Códigos abiertos totales: 147. Categorías principales: 4. Subcategorías: 11. Citas seleccionadas por densidad descriptiva y representatividad de propiedades. Identificadores: D=Docente, E=Estudiante, número de entrevista.

La primera categoría, *habitus evaluativo* docente, agrupa las disposiciones sedimentadas que posicionan al docente como árbitro exclusivo del juicio evaluativo. Sus subcategorías —"juez de resultados", "temor a perder autoridad" y "nostalgia de la calificación"— presentaron la mayor densidad de códigos en observaciones de campo (ver Tabla 3). La expresión de Elena, "cuando comparto la lista de cotejo siento que me quito la corona", condensa con precisión el mecanismo simbólico en juego: la lista de cotejo no es percibida como herramienta compartible sino como insignia de autoridad. La segunda categoría, *falacia de la autonomía*, emergió como la más extendida entre docentes (6 de 8) y estudiantes (9 de 10): la ecuación operante en el aula entre silencio y autonomía, y entre ficha completada y logro de aprendizaje. La declaración de Carlos —"creía que era

autónomo porque trabajaba calladito"— ilustra el mecanismo de confusión que produce esta falacia en la práctica docente. Como se observa en la Figura 1, el modelo teórico emergente muestra las relaciones entre estas categorías con sus respectivas densidades de codificación.

Figura 1.

Modelo teórico emergente: Interaprendizaje evaluativo como proceso de resignificación del criterio en educación primaria.



Nota. Los nodos en negrita representan saturación teórica alcanzada. Las líneas punteadas indican relaciones emergentes. Las flechas bidireccionales denotan procesos dialécticos de deconstrucción de la falacia y reconstrucción de la autonomía relacional. El tamaño nodal es proporcional a la densidad de códigos abiertos (N = 147).

La tercera categoría, *praxis de Interaprendizaje*, fue identificada en 23 escenas de aula específicas, todas ellas caracterizadas por un elemento estructurante común: la explicitación previa del criterio mediante una lista de cotejo construida colectivamente. En

estas escenas, los estudiantes transitaron del juicio afectivo ("está bonito", "no me gusta") a la argumentación criterial ("tu paso 2 falta", "¿qué tal si agregas...?"). La expresión de Valentina —"le dije a Diego que su infografía estaba ordenada, pero le faltaba explicar los pasos. Me dijo gracias y luego me ayudó con mi título"— documenta el mecanismo de reconocimiento mutuo en operación. La cuarta categoría, *conciencia evaluativa*, emergió de manera más difusa y fue identificada solo en 4 de los 8 docentes y en 5 de los 10 estudiantes, concentrándose en contextos urbanos; la expresión de Sofía —"antes me daba vergüenza equivocarme. Ahora cuando mi compañero me dice 'falta esto', yo pienso cómo arreglarlo antes de que el profe vea"— ilustra el tránsito desde la heteronomía hacia formas incipientes de autorregulación consciente.

La Tabla 3 presenta la densidad de códigos por fuentes de datos, permitiendo identificar qué fenómenos son más visibles en el discurso reflexivo y cuáles emergen principalmente de la observación directa de la práctica.

Tabla 3.

Densidad de códigos por fuentes de datos

Código	Frec. entrevistas docentes	Frec. entrevistas estudiantes	Frec. observaciones de campo	Frec. documentos	Densidad total
miedo a perder autoridad	7	0	12	0	19
falacia de la autonomía	6	7	9	2	24
criterio como andamio	5	4	15	8	32
coevaluación mediada	4	6	18	3	31
espejo colectivo	2	5	8	0	15
diálogo por criterios	3	7	14	1	25
error como oportunidad	2	4	6	0	12
autorregulación consciente	1	3	4	0	8

nostalgia de la calificación	6	0	3	1	10
confianza histórica previa	4	2	5	0	11

Nota. Frecuencias brutas de aparición del código en cada fuente. Densidad total = suma simple. Los códigos con mayor densidad en observaciones ("criterio como andamio", "coevaluación mediada") indican que estas prácticas son más visibles en acción que en discurso reflexivo.

Como se ilustra en la Tabla 3, los códigos "criterio como andamio" (densidad total: 32) y "coevaluación mediada" (densidad total: 31) obtuvieron las mayores frecuencias globales y concentraron sus apariciones en observaciones de campo (15 y 18, respectivamente), lo que señala que ambos fenómenos son más visibles en la práctica actuada que en el discurso docente reflexivo. En contraste, "autorregulación consciente" presentó la menor densidad total (8), con escasa presencia en documentos institucionales.

Proceso teórico central: trayectoria del interaprendizaje evaluativo

La categoría central emergente —Interaprendizaje evaluativo como resignificación del criterio— articula las relaciones sustantivas entre las cuatro categorías principales en una trayectoria procesual que la Tabla 4 describe en términos de estado de saturación y relaciones identificadas.

Tabla 4.

Estructura de categorías y estado de saturación

Categoría central	Categorías principales	Subcategorías	Relaciones identificadas	Estado saturación
Interaprendizaje evaluativo como resignificación del criterio	Habitus evaluativo docente	Juez de resultados; Temor a perder autoridad; Nostalgia calificación	Condición estructural: el habitus docente limita la apertura del campo evaluativo	Saturada (8/8 docentes, 45 escenas)
	Falacia de la autonomía	Silencio=autonomía; Ficha=logro; Ausencia metacognición	Condición de posibilidad: sin reconocimiento de la falacia, no hay tránsito	Saturada (7/8 docentes, 9/10 estudiantes)

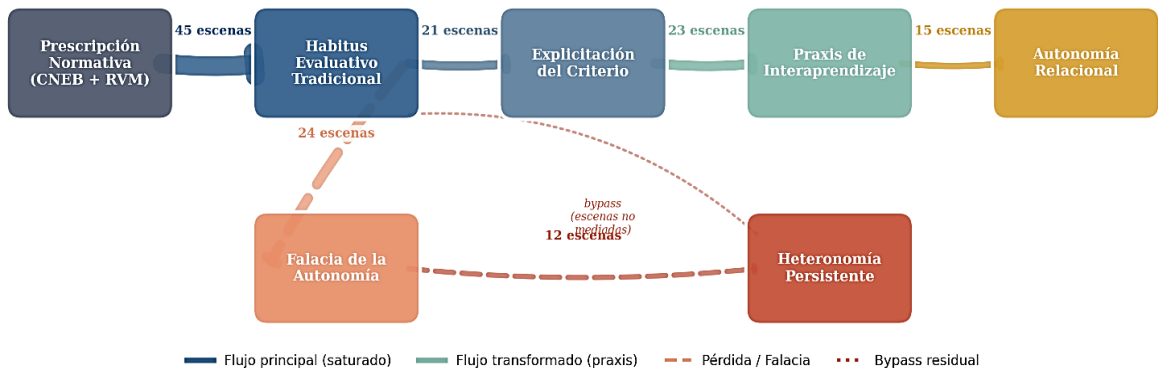
Praxis de interaprendizaje	Coevaluación mediada; Espejo colectivo; Diálogo por criterios	Estrategia central: mecanismo de construcción de sentido y reconocimiento mutuo	Saturada (6/8 docentes, 8/10 estudiantes, 23 escenas)
Conciencia evaluativa	Mediación del error; Error como oportunidad; Autorregulación consciente	Consecuencia: cuando la praxis se consolida, emerge la autonomía relacional	Emergente (4/8 docentes, 5/10 estudiantes)

Nota. Saturación teórica evaluada según criterio de Charmaz (2025): densidad de propiedades, dimensiones y relaciones entre categorías. Categoría "Conciencia Evaluativa" permanece emergente por baja frecuencia en contextos rurales tradicionales.

Como se detalla en la Tabla 4, el habitus evaluativo docente opera como condición estructural que limita la apertura del campo evaluativo, alcanzando saturación plena en los 8 docentes y las 45 escenas de observación. La Falacia de la Autonomía funciona como condición de posibilidad: sin que sea nombrada y reconocida por el docente, no hay tránsito hacia la Praxis de Interaprendizaje. La Praxis de Interaprendizaje emerge como estrategia central —el mecanismo de construcción de sentido y reconocimiento mutuo— y se encontró saturada en 6 de los 8 docentes y 8 de los 10 estudiantes, en 23 escenas identificadas. La Conciencia Evaluativa, consecuencia cuando la Praxis se consolida, permanece emergente dado que se presentó en solo 4 docentes y 5 estudiantes, concentrándose en contextos urbanos con mayor disponibilidad de recursos pedagógicos y menor presión parental por calificaciones numéricas. La Figura 2 ilustra la trayectoria procesual completa desde la prescripción normativa hasta la autonomía relacional, mostrando los flujos principales y residuales identificados en el trabajo de campo.

Figura 2.

Trayectoria procesual del tránsito entre heteronomía y autonomía relacional en prácticas evaluativas primarias.



Nota. Los anchos de flujo son proporcionales al número de escenas observadas (N = 45 sesiones de campo). Las líneas punteadas representan trayectorias de persistencia heterónoma. La bifurcación en "Explicitación del Criterio" representa el punto de inflexión teórico identificado mediante comparación constante.

Análisis de casos discrepantes y matización del modelo

El examen de los casos discrepantes constituye un momento analítico de singular importancia para la delimitación del modelo emergente. La Tabla 5 presenta ambos casos con sus características distintivas, hallazgos contradictorios y la reinterpretación teórica que habilitan.

Tabla 5.

Análisis de Casos Discrepantes Negativos

Caso atípico (pseudónimo)	Características distintivas	Hallazgo contradictorio	Reinterpretación teórica/matización propuesta
Rosa (Docente, 20 años, rural, multigrado)	Máxima experiencia, contexto de extrema vulnerabilidad económica, 28 estudiantes en multigrado 3.º–4.º	Rechaza explícito de coevaluación: "Aquí los padres me piden notas, no conversaciones. Y los niños no saben evaluar, solo se copian."	Matización: la coevaluación requiere confianza histórica institucional previa. En contextos donde la escuela es la única institución legítima, externalizar el juicio amenaza el capital simbólico docente. La autonomía no es universalmente deseable desde la perspectiva del actor en condiciones de precariedad.

Sebastián (Estudiante, 11 años, 5.º grado, TDAH)	Diagnóstico médico confirmado, hiperactividad, dificultad de atención sostenida, historia de exclusión grupal	Durante coevaluación grupal, compañeros le asignan nivel B "porque no se queda quieto". No argumentan criterios. Sebastián se retira físicamente.	Matización: la coevaluación grupal sin mediación docente intensifica mecanismos de estigmatización preexistentes. El dispositivo pedagógico (coevaluación) reproduce desigualdad cuando no hay andamiaje retractor del docente. La autonomía requiere condiciones de inclusión previas.
---	--	---	--

Nota. Casos seleccionados por máxima divergencia respecto a teoría dominante emergente. Análisis explicitado en memos de discrepancia (Memo-DIS-03 y Memo-DIS-07).

Como se observa en la Tabla 5, el caso de Rosa —docente con 20 años de servicio en contexto rural multigrado— revela que la resistencia a la coevaluación no es meramente ideológica sino estructuralmente fundada: en contextos de extrema vulnerabilidad donde la escuela es la única institución legítima de la comunidad, externalizar el juicio evaluativo a los estudiantes amenaza el capital simbólico docente que hace posible el orden institucional. Su rechazo explícito —"aquí los padres me piden notas, no conversaciones. Y los niños no saben evaluar, solo se copian"— no es ignorancia de la normativa sino lectura estructural de las condiciones de posibilidad de su práctica. El caso de Sebastián, por su parte, documenta cómo la coevaluación grupal sin mediación docente puede intensificar mecanismos de estigmatización preexistentes: sus compañeros le asignaron nivel B "porque no se queda quieto", sin argumentar criterios, y Sebastián se retiró físicamente del espacio de coevaluación. Este caso matiza con firmeza la idealización del dispositivo: la coevaluación requiere condiciones de inclusión previas y andamiaje docente activo para no reproducir la desigualdad que en teoría busca resolver.

Triangulación y consolidación de evidencia

La convergencia entre las tres fuentes de datos fue evaluada sistemáticamente para cada tema central. La Tabla 6 presenta el proceso de triangulación sobre el tema del criterio como andamio, mostrando la convergencia entre el discurso docente, las notas de campo observacionales y los documentos institucionales.

Tabla 6.

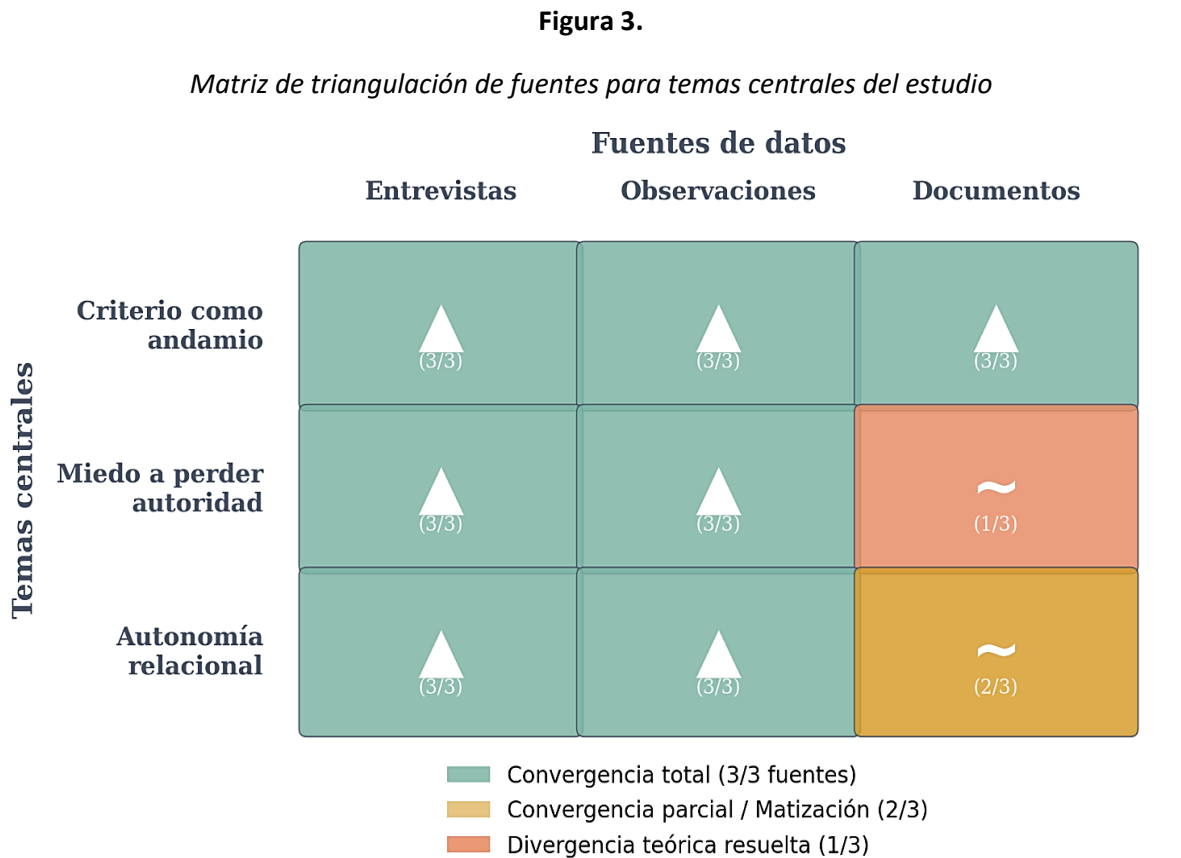
Evidencia de Triangulación por Tema Central

Tema central	Evidencia Entrevista (cita)	Evidencia Observación (nota de campo)	Evidencia Documento (extracto)	Interpretación integrada
El criterio como andamio	D-Elena-01: "Sin criterio previo la coevaluación es un desastre. Es como darles una sierra sin enseñarles a cortar."	Campo-23/04: Estudiantes en 4.° A intentan coevaluar infografías sin lista de cotejo. Juicios: "está bonito", "no me gusta". Caos organizado. 15 min perdidos.	Guía Metodológica DRE Moquegua (p. 18): "La lista de cotejo debe construirse con los estudiantes antes de la coevaluación, explicitando criterios de calidad compartidos."	Convergencia triple: el criterio explícito es condición necesaria pero no suficiente. La explicitación requiere práctica previa de uso del criterio, no solo lectura.
Miedo a perder autoridad	D-Mario-01: "Si ellos evalúan, ¿para qué estoy yo? Me siento fuera de lugar."	Campo-15/05: Mario interrumpe coevaluación en 5.° B para "corregir" juicio estudiantil. Retoma control. Lista de cotejo en su mano, no en mesas.	RVM 094-2020-MINEDU, Art. 12: "El docente acompaña el proceso de autoevaluación y coevaluación, promoviendo la participación activa del estudiante."	Divergencia resuelta: la normativa prescribe acompañamiento, pero el habitus docente interpreta acompañamiento como control. La tensión no es normativa sino ontológica: ¿qué significa "ser docente"?
Autonomía como construcción relacional	E-Camila-01: "Cuando evaluamos en pareja, aprendo más que cuando el profe me dice."	Campo-03/06: Camila y Luciano coevalúan gráficos de agua. Dialogan 8 min sin intervención docente. Usan criterios de eje X, escala, título. Se corrigen mutuamente.	CNEB, Competencia 29: "Gestiona su aprendizaje de manera autónoma."	Convergencia: la autonomía se manifiesta en interacción, no en aislamiento. El documento oficial asume autonomía individual; la práctica revela autonomía relacional.

Nota. Notas de campo identificadas como Campo-[fecha]. Documentos: Guías Metodológicas DRE Moquegua 2024–2025, RVM 094-2020-MINEDU, CNEB 2016.

Como se detalla en la Tabla 6, la triple convergencia sobre el criterio como andamio revela un hallazgo de interpretación integrada de alta consistencia: la Guía Metodológica prescribe que la lista de cotejo debe construirse con los estudiantes antes de la coevaluación; la observación de campo muestra que sin esa construcción previa los juicios estudiantiles son afectivos y caóticos; y el discurso docente de Elena confirma que la explicitación del criterio es condición necesaria pero insuficiente —también requiere

práctica de uso del criterio, no solo lectura. La Figura 3 ofrece una visualización de la matriz de triangulación para los tres temas centrales del estudio.



Nota. Los anchos de flujo son proporcionales al número de escenas observadas (N = 45 sesiones de campo). Las líneas punteadas representan trayectorias de persistencia heterónoma. La bifurcación en "Explicitación del Criterio" representa el punto de inflexión teórico identificado mediante comparación constante.

El proceso de comparación constante fue documentado en el trail de auditoría, cuyo extracto representativo se presenta en la Tabla 7.

Tabla 7.

Trail de auditoría (extracto representativo)

Etapas del análisis	Decisión tomada	Justificación teórica/memorandum	Fecha/Iteración
Codificación abierta inicial	Crear código "miedo a perder autoridad"	Memo-02: La resistencia es activa y política; el miedo es pasivo y emocional. En campo, Mario	15/03/2025, Iteración 2

	separado de "resistencia al cambio"	muestra miedo, no resistencia ideológica. Requiere categorías distintas.	
Codificación axial	Agrupar "falacia de la autonomía" bajo categoría central emergente vs. mantenerla como categoría periférica	Memo-09: La falacia aparece en 6 de 8 docentes y 7 de 10 estudiantes. Es condición de posibilidad para la interaprendizaje. Sin ella, no hay tránsito. Elegida como categoría principal.	22/04/2025, Iteración 4
Decisión de cierre muestral	Detener entrevistas con docentes tras caso 7 (Elena) y estudiantes tras caso 8 (Camila)	Memo-14: Comparación constante entre casos 6-7-8 no añade nuevas propiedades a "Interaprendizaje Evaluativo". Variación máxima alcanzada. Saturación teórica por densidad dimensional confirmada.	10/05/2025, Iteración 6
Revisión de caso negativo	Mantener caso Rosa como discrepante sin descartarlo como "outlier"	Memo-DIS-03: El outlier metodológico es tesoro teórico. Rosa no invalida la teoría; define su corredor de indeterminación hermenéutica.	18/05/2025, Iteración 7

Nota. Trail completo: 34 memos teóricos, 12 memos de reflexividad, 8 memos de discrepancia. Total: 54 documentos de auditoría.

Como se muestra en la Tabla 7, las decisiones analíticas más críticas —la separación entre "miedo a perder autoridad" y "resistencia al cambio", la elección de la Falacia de la Autonomía como categoría principal, y la decisión de cierre muestral— están respaldadas por memos teóricos fechados que permiten reconstruir la lógica de cada decisión. La comparación de escenas clave, sistematizada en la Tabla 8, permitió el refinamiento continuo de las categorías emergentes mediante el método de comparación constante.

Tabla 8.

Comparación constante de escenas clave (Key incidents)

Incidente comparado	Incidente inicial	Puntos similitud	Puntos diferencia	Refinamiento categoría
Escena 3.1 (Infografía sin criterio, caos) vs. Escena 3.2 (Infografía con criterio, diálogo)	Escena 3.1: Estudiantes asignan niveles sin criterio ("está bonito"). Escena 3.2: Estudiantes argumentan desde	Ambas: coevaluación de infografía en 4.º grado, área Comunicación, mismo docente	3.1: sin lista de cotejo previa, 15 min, juicios afectivos, docente interrumpe. 3.2: con lista construida colectivamente, 22 min, argumentación	Refinamiento: la coevaluación no es inherentemente formativa. La variable decisiva no es la técnica sino la explicitación previa y compartida del criterio.

	criterio ("falta paso 2").	(Elena), misma unidad.	criteriada, docente observa desde periferia.	Categoría "Praxis de Interaprendizaje" adquiere dimensión temporal: fase preparatoria es condición.
Escena rural (Mario, 5.° B, matemática) vs. Escena urbana (Carlos, 5.° A, matemática)	Mario: estudiantes resuelven problema de agua individualmente, entregan al docente, esperan calificación. Carlos: estudiantes resuelven en parejas, coevalúan con lista de cotejo, proponen mejoras.	Ambas: situación problemática contextualizada, área Matemática, mismo desafío ("uso del agua"), nivel 5.°.	Mario: producto individual, juicio docente unidireccional, 0 interacción entre pares. Carlos: producto grupal, retroalimentación bidireccional, docente como mediador de criterios.	Refinamiento: el contexto rural no determina heteronomía, pero la ausencia de recursos y la presión parental por notas (Rosa) intensifican el habitus evaluativo tradicional. Se añade dimensión "presión institucional externa" a categoría "Habitus Evaluativo Docente".

Nota. Escenas seleccionadas por máxima variación contextual (urbano/rural, con/sin criterio, misma área/nivel). Comparación constante realizada en memos teóricos (Memo-CC-05 y Memo-CC-11).

Como ilustra la Tabla 8, la comparación entre la Escena 3.1 (coevaluación de infografía sin criterio previo) y la Escena 3.2 (coevaluación con lista de cotejo construida colectivamente) —ambas en el mismo grado, área, docente y unidad curricular— permitió el refinamiento categorial más significativo del estudio: la coevaluación no es inherentemente formativa. La variable decisiva no es la técnica sino la explicitación previa y compartida del criterio, que además requiere una fase preparatoria de práctica de uso. Adicionalmente, la comparación entre la escena rural (Mario, contexto de trabajo individual y juicio docente unidireccional) y la urbana (Carlos, coevaluación en parejas con mediación criterial) permitió precisar que el contexto rural no determina en sí mismo la heteronomía, pero la ausencia de recursos y la presión parental por calificaciones intensifican el habitus evaluativo tradicional, añadiendo la dimensión "presión institucional externa" a la categoría *habitus evaluativo docente*.

Discusión

Resignificación ontológica: autonomía como externalización relacional

La autonomía no se decreta. Se construye en la interacción, mediada por criterios que el estudiante puede argumentar, no solo ejecutar. Este hallazgo —que la autonomía evaluativa en contextos de alta colectividad andina opera no como internalización individual del criterio sino como externalización relacional del error— desafía en su

fundamento el dogma canónico de la teoría del aprendizaje autorregulado, que desde Zimmerman ha modelado el desarrollo de la autonomía como trayectoria intraindividual: el estudiante progresa desde la observación de modelos hacia la autoeficacia interna, pasando por el control volitivo de sus propios procesos. Esta conceptualización, elaborada y validada casi exclusivamente en contextos norteamericanos y europeos de alto individualismo, viajó a los sistemas educativos latinoamericanos bajo la forma de políticas de evaluación formativa que la presuponen sin cuestionarla. La RVM 094-2020-MINEDU prescribe autonomía; el CNEB la nombra como competencia; las Guías Metodológicas DRE Moquegua la operacionalizan en técnicas. Pero ninguno de estos textos normativos interroga la ontología subyacente de la autonomía que prescriben.

Lo que la etnografía reveló —con una consistencia que constituyó en sí misma la anomalía más disruptiva del estudio— es que en los contextos de aula moqueguanos donde la autonomía relacional emergió de manera empíricamente reconocible, el mecanismo no fue la internalización privada del criterio sino su activación pública en el intercambio entre pares. Valentina no se volvió autónoma cuando comprendió el criterio de manera individual; se volvió autónoma cuando pudo argumentarlo ante Diego y recibir de él una argumentación de retorno. El criterio funcionó como andamio cultural en el sentido más riguroso del término: estructura de sostén que puede —y debe— ser retirada progresivamente, pero que en su ausencia hace imposible el proceso que pretende apoyar. Esta lectura se alinea, aunque desde una perspectiva empírica periférica, con los hallazgos sobre mecanismos de sensemaking en evaluación formativa que muestran cómo la legitimación de las ideas previas del estudiante en el intercambio dialógico —y no la transmisión del criterio desde el docente— es la condición de posibilidad de la autonomía evaluativa (Deverel & Furtak, 2025). La coevaluación como práctica dialógica, multimodal y culturalmente situada que sostiene la autonomía colectiva (Champion et al., 2025) encuentra aquí su correlato andino: el "espejo colectivo" no es metáfora decorativa sino descripción mecánica del proceso.

Notablemente, este hallazgo no solo matiza sino refuta parcialmente la jerarquía implícita de dispositivos evaluativos que la literatura ha consolidado: la autoevaluación como práctica superior a la coevaluación en términos de desarrollo de autorregulación. En contextos de alta interdependencia cultural, la inversión de esta jerarquía no es anomalía periférica sino lógica estructural del campo. La investigación sobre fondos de conocimiento como base del criterio evaluativo (Randall et al., 2025) ofrece una perspectiva teórica convergente desde la etnografía norteamericana, pero con alcance limitado para capturar la densidad simbólica del contexto andino, donde la reciprocidad no es solo valor cultural sino principio organizador de las relaciones de producción y reproducción social. Lo que la teoría del aprendizaje autorregulado llama "ayuda solicitada" —una subcategoría de la

reflexión metacognitiva— es en el aula moqueguána una práctica de reconocimiento mutuo con implicaciones que exceden la cognición individual. El concepto de saturación teórica por variación dimensional (Charmaz, 2025) aplicado a estas relaciones reveló que la categoría "Conciencia Evaluativa" permanece emergente no por insuficiencia empírica sino porque sus condiciones de posibilidad —confianza histórica, criterio compartido, andamiaje docente retractil— solo se materializaron de manera consistente en cuatro de los ocho docentes. La autonomía como estado consolidado es, en este contexto, una aspiración hacia la que se puede avanzar; no un punto de llegada alcanzable en el horizonte de una reforma.

Desde una perspectiva crítica, el caso de Rosa merece atención teórica desproporcionada al espacio que la lógica del promedio le asignaría. Su resistencia a la coevaluación —nombrada como caso discrepante negativo— es en realidad el caso más iluminador del estudio porque revela la limitación epistemológica más profunda del marco normativo: ignora que la escuela rural multigrado de extrema vulnerabilidad no es un rezago en el camino hacia la modernización evaluativa, sino un campo con lógica propia donde el capital simbólico del docente es la única moneda institucional disponible para sostener el orden pedagógico. Cualquier política que externalice el juicio evaluativo sin construir primero las condiciones institucionales —confianza histórica, protección parental, recursos para el andamiaje— no produce autonomía: produce caos que, para preservar el funcionamiento de la institución, el actor racionaliza como fracaso del método, consolidando el habitus evaluativo tradicional. Esta es la lógica del corredor de indeterminación hermenéutica entre discurso normativo y práctica sedimentada: el espacio donde la normativa produce efectos contrarios a sus propósitos declarados porque opera sobre un campo cuya estructura no interpela.

Limitaciones metodológicas

La interpretación de estos hallazgos requiere precisiones de alcance. Las limitaciones de interpretación incluyen el posicionamiento del investigador como exdocente de primaria peruana —que facilitó el rapport pero exigió vigilancia epistemológica sistemática— y la especificidad del contexto moqueguano, cuya combinación de alta prescripción normativa regional y vulnerabilidad rural extrema en algunas instituciones representa condiciones estructurales no generalizables estadísticamente a otros contextos peruanos. El sesgo de autoselección en el muestreo teórico —docentes voluntarios con mayor disposición al cambio— implica que los perfiles de mayor resistencia al cambio están subrepresentados, con excepción del caso Rosa, seleccionado precisamente por su valor de máxima divergencia.

Prescripción situada: del hallazgo a la acción

Estos hallazgos sugieren preliminarmente una configuración de sentido hacia la autonomía relacional cuya transferibilidad está restringida estrictamente a contextos caracterizados por alta interdependencia cultural, prescripción normativa explícita y tradición escolar centralizada. Para gestores de la DRE Moquegua y de sistemas educativos análogos, la implicación más directa es la siguiente: implementar rituales de explicitación de criterios —construidos colectivamente con los estudiantes, no transmitidos por el docente— antes de introducir rúbricas estandarizadas. El modelo "rúbrica primero, coevaluación después" que parece guiar muchas capacitaciones docentes en la región invierte la secuencia que el campo empírico indica como productiva. Para los docentes de primaria en contextos de alta colectividad, la recomendación es precisar la condición mínima de activación: la coevaluación solo puede ejercerse productivamente tras tres a cuatro sesiones de modelado del criterio con lista de cotejo construida colectivamente, y exige que el docente permanezca activo como mediador —no como juez— durante las primeras experiencias de coevaluación grupal. El caso de Sebastián actúa como advertencia irreductible: la coevaluación sin andamiaje retráctil puede convertirse en dispositivo de estigmatización para estudiantes con neurodiversidad u otras condiciones de vulnerabilidad. La validez catalítica del estudio —su capacidad para generar transformación en los actores estudiados (Lather, recuperado en el marco analítico)— reside precisamente en que estas prescripciones no emergen de un laboratorio experimental sino de 120 horas de inmersión en las condiciones reales de cuatro instituciones peruanas. Cualquier generalización analítica más amplia requiere juicio situado dado el muestreo teórico intencional y el N intencionalmente pequeño.

Estudios futuros podrían capturar el proceso de institutional work mediante diseños longitudinales de dos a tres años que sigan la trayectoria de los docentes a lo largo de múltiples ciclos normativos. Estudios comparativos interculturales —Andes versus Amazonía; Moquegua versus la sierra ecuatoriana— podrían matizar la noción de colectividad andina, evitando su homogeneización. El apoyo a la autonomía y las relaciones de pares como mediadores de bienestar psicológico en la experiencia estudiantil (Wu & Dong, 2024; Hooper et al., 2025) sugieren que la dimensión socio-afectiva de la coevaluación merece investigación específica en primaria, complementando los análisis de self-regulated learning que predominan en educación secundaria y universitaria (He et al., 2024; Li & Gu, 2024; Arvatz et al., 2025). Métodos mixtos secuenciales podrían cuantificar la frecuencia de escenas de coevaluación efectiva —entendida como argumentación criterial, no mera aplicación de ficha— y su asociación con indicadores de desempeño y bienestar estudiantil (Alshammari, 2024; Cao & Pham, 2025; Yang et al., 2025).

Conclusiones

El presente estudio partió de una paradoja documentada en la literatura y agudizada en el contexto peruano: la evaluación formativa es prescrita con sofisticación normativa creciente mientras las prácticas de aula reproducen con tenacidad la heteronomía evaluativa. La etnografía institucional en Moquegua permitió acceder a los mecanismos de esa reproducción y, más relevante aún, a los procesos específicos mediante los cuales algunos docentes y estudiantes logran trascenderla.

El hallazgo central —que la autonomía en contextos de alta colectividad andina opera como externalización relacional del error, no como internalización individual del criterio— constituye una extensión teórica que obliga a revisar supuestos fundacionales del aprendizaje autorregulado. La categoría emergente de "interaprendizaje evaluativo" nombra un proceso que la teoría canónica no anticipaba: el criterio explícito construido colectivamente funciona como andamio cultural que transforma el error de penalización privada en insumo de reconocimiento mutuo. Esta transformación no es automática ni universal: requiere condiciones de posibilidad precisas —confianza histórica, criterio compartido, andamiaje docente retráctil, protección institucional frente a presiones parentales por calificación— que las políticas de evaluación formativa deben construir antes de prescribir técnicas.

Para la gestión de la DRE Moquegua y sistemas análogos, la implicación más operativa es temporal: el cuándo importa tanto como el qué. La coevaluación solo puede activarse productivamente después de que el criterio haya sido trabajado colectivamente durante al menos tres a cuatro sesiones de modelado. Para los docentes, la condición más contraintuitiva es la más crítica: la autonomía estudiantil aumenta cuando el docente no se retira del proceso sino cuando transita del rol de juez al de mediador activo. Para los investigadores, la agenda más urgente es la comparación transcultural de mecanismos de autonomía relacional en sistemas con distintas formas de colectividad —andina, amazónica, asiática— que permita afinar los límites de transferibilidad analítica del modelo emergente. Este estudio constituye un primer paso hacia una teorización de la autonomía evaluativa desde la periferia, donde el reconocimiento mutuo precede a la autorregulación individual y el criterio compartido es medio de estima, no solo herramienta cognitiva.

Referencias

- Admiraal, W., Lockhorst, D., Post, L., & Kester, L. (2024). Effects of Students' Autonomy Support on their Self-Regulated Learning Strategies: Three Field Experiments in Secondary Education. *International Journal of Research in Education and Science*. <https://files.eric.ed.gov/fulltext/EJ1415018.pdf>
- Ahmed, S. (2024). The Pillars of Trustworthiness in Qualitative Research. *Journal of Medicine, Surgery, and Public Health*. <https://doi.org/10.1016/j.glmedi.2024.100051>
- Alshammari, A. (2024). Promoting Saudi University Students' Experience Through Peer Assessment Practice. *Theory and Practice in Language Studies*. <https://doi.org/10.17507/tpls.1408.21>
- Arvatz, A., Peretz, R., & Dori, Y. (2025). Self-regulated learning and reflection: a tool for teachers and students. *Metacognition and Learning*, 20. <https://doi.org/10.1007/s11409-025-09415-3>
- Atjonen, P., Kontkanen, S., Ruotsalainen, P., & Pontinen, S. (2024). Pre-service teachers as learners of formative assessment in teaching practice. *European Journal of Teacher Education*, 47, 267–284. <https://doi.org/10.1080/02619768.2024.2338840>
- Brown, D., & Lloyd, R. (2024). Using Critical Social Theory as Professional Learning to Develop Scholar—Practitioners in Physical Education: The Example of Bourdieu's Theory of Practice. *Education Sciences*. <https://doi.org/10.3390/educsci14020160>
- Byrd, D. (2025). Using Institutional Habitus to Position Colleges and Universities as Social Actors. *Educational Theory*. <https://doi.org/10.1111/edth.12686>
- Cao, T., & Pham, X. (2025). Formative assessment for self-regulated learning in English language education: Student perspectives. *VNU Journal of Foreign Studies*. <https://doi.org/10.63023/2525-2445/jfs.ulis.5509>
- Chakma, U., & Li, B. (2025). Ensuring Trustworthiness Using an Inductive Approach in Qualitative Educational Research: An Autoethnographic Investigation of Two Early Career Researchers Reflecting on PhD Data Analysis. *International Journal of Educational Methodology*. <https://doi.org/10.12973/ijem.11.2.233>
- Champion, D., Solomon, F., & Lammey, M. (2025). Embodying physics assessment: Reimagining formative assessment as a creative, multimodal, and culturally sustaining dialogic practice. *Journal of the Learning Sciences*, 34, 474–525. <https://doi.org/10.1080/10508406.2025.2569583>

- Christou, P. (2025). Reliability and Validity in Qualitative Research Revisited and the Role of AI. *The Qualitative Report*. <https://doi.org/10.46743/2160-3715/2025.7523>
- Coleman, M., Ragan, M., & Dari, T. (2024). Intercoder Reliability for Use in Qualitative Research and Evaluation. *Measurement and Evaluation in Counseling and Development*, 57, 136–146. <https://doi.org/10.1080/07481756.2024.2303715>
- Dediwalage, M., Kalogeropoulos, P., Russo, J., & Vale, C. (2025). A systematic literature review of student, parent and teacher experiences of mathematics education in multicultural classrooms. *Review of Education*. <https://doi.org/10.1002/rev3.70109>
- Deverel, C., & Furtak, E. (2025). How Can Features of Classroom Activity Systems Center Students' Ideas and Experiences in Formative Assessment? A Study of Sensemaking and Answer-Making Frames. *Educational Assessment*, 30, 91–114. <https://doi.org/10.1080/10627197.2025.2504081>
- He, S., Epp, C., Chen, F., & Cui, Y. (2024). Examining change in students' self-regulated learning patterns after a formative assessment using process mining techniques. *Computers in Human Behavior*, 152, 108061. <https://doi.org/10.1016/j.chb.2023.108061>
- Hooper, D., Egitim, S., & Hofhuis, J. (2025). Exploring the impact of near-peer role modeling on learners' basic psychological needs: Insights from English classes in Japanese higher education. *International Journal of Educational Research Open*. <https://doi.org/10.1016/j.ijedro.2024.100429>
- Lai, M., & Jung, J. (2024). Pathways to master's education: insights from students in transnational higher education in China. *Higher Education*, 89, 937–952. <https://doi.org/10.1007/s10734-024-01255-5>
- Lei, W., & Lei, Z. (2025). Exploring EFL teachers' beliefs and practices of formative assessment in Chinese context. *PLOS One*, 20. <https://doi.org/10.1371/journal.pone.0333678>
- Li, J., & Gu, P. (2024). Formative assessment for self-regulated learning: Evidence from a teacher continuing professional development programme. *System*. <https://doi.org/10.1016/j.system.2024.103414>
- Lin, H., Tan, S., Huang, Y., Duan, H., Li, Y., & Hao, Y. (2025). Navigating goals of care discussion in palliative care: A qualitative study applying Bourdieu's theory of practice. *Palliative Care and Social Practice*, 19. <https://doi.org/10.1177/26323524251400803>

- Luong, G. (2025). Exploring Student Assessment Practices in Vietnamese Teacher Training Universities through Naturalistic Inquiry. *Vietnam Journal of Education*. <https://doi.org/10.52296/vje.2025.426>
- Lyu, M., & Lam, R. (2025). Private English tutors' agency amid China's 'double reduction' policy: a Bourdieusian perspective. *Journal of Education Policy*, 40, 667–688. <https://doi.org/10.1080/02680939.2025.2479151>
- Lyons, P., Gause, S., Eakin, M., O'Brien, B., & Santhosh, L. (2024). Seven Practical Recommendations for Designing and Conducting Qualitative Research in Medical Education. *ATS Scholar*, 5, 231–241. <https://doi.org/10.34197/ats-scholar.2023-0144ps>
- Maher, A. (2025). Rigour in interpretive qualitative research in education: Ideas to think with. *British Educational Research Journal*. <https://doi.org/10.1002/berj.4156>
- Merkebu, J., Geringer, J., & Samuel, A. (2025). Sampling Thoughtfully in Qualitative Research: Hallmarks of Rigour in Health Professions Education. *The Clinical Teacher*, 22. <https://doi.org/10.1111/tct.70145>
- Neziri, M., Kryeziu, V., & Avdiu, S. (2025). Teachers' Attitudes from Experience Regarding Formative Assessment in Achieving Curricular Outcomes in Primary Education. *Journal Human Research in Rehabilitation*. <https://doi.org/10.21554/hrr.092515>
- Nyimbili, F., & Nyimbili, L. (2024). Types of Purposive Sampling Techniques with Their Examples and Application in Qualitative Research Studies. *British Journal of Multidisciplinary and Advanced Studies*. <https://doi.org/10.37745/bjmas.2022.0419>
- Özkan, Y., & Ozkan, M. (2025). Decoding Teachers' Dilemma: Unveiling the Real Obstacles to Implementing Formative Assessment in the Classroom. *Journal of Qualitative Research in Education*. <https://doi.org/10.14689/enad.41.2062>
- Palma, C. (2025). Higher education and social recognition: On the moral dimension of pupils' aspirations and choices in Chile. *The Sociological Review*. <https://doi.org/10.1177/00380261251351395>
- Pastore, S. (2025). Preparing formative assessment literate primary school teachers: A blurred snapshot from Italy. *The Curriculum Journal*. <https://doi.org/10.1002/curj.70014>

- Rahimi, S., & Khatooni, M. (2024). Saturation in qualitative research: An evolutionary concept analysis. *International Journal of Nursing Studies Advances*, 6. <https://doi.org/10.1016/j.ijsna.2024.100174>
- Randall, J., Earnest, D., Thota, N., & Mensing, S. (2025). Investigating and assessing informal computational thinking in grades K-2: A funds of knowledge approach. *Journal of the Learning Sciences*, 34, 526–570. <https://doi.org/10.1080/10508406.2025.2574311>
- Roberts, C., Khanna, P., & Burgess, A. (2024). Capital sharing and socialization in an interprofessional student-led clinic: a Bourdieuan analysis. *BMC Medical Education*, 24. <https://doi.org/10.1186/s12909-024-05117-7>
- Sandoval, L., Sotelo, J., Viana, L., & González, C. (2025). How Does Professional Habitus Impact Nursing Autonomy? A Hermeneutic Qualitative Study Using Bourdieu's Framework. *Nursing Reports*, 15. <https://doi.org/10.3390/nursrep15030088>
- Silén, C., Manninen, K., & Fredholm, A. (2024). Designing for student autonomy combining theory and clinical practice – a qualitative study with a faculty perspective. *BMC Medical Education*, 24. <https://doi.org/10.1186/s12909-024-05514-y>
- Sortwell, A., Trimble, K., Ferraz, R., Geelan, D., Hine, G., Ramirez, R., Carter, B., Gkintoni, E., & Xuan, Q. (2024). A Systematic Review of Meta-Analyses on the Impact of Formative Assessment on K-12 Students' Learning: Toward Sustainable Quality Education. *Sustainability*. <https://doi.org/10.3390/su16177826>
- Supriyanto, S., Imron, A., Maisyaroh, M., Muslihati, M., Syarif, I., Puspitasari, F., & Sulaiman, A. (2025). Social Practices of Students with Disabilities in Inclusive School: A Bourdieusian Analysis of Habitus, Capital, and Field. *Munaddhomah: Jurnal Manajemen Pendidikan Islam*. <https://doi.org/10.31538/munaddhomah.v6i4.2055>
- Wu, D., & Dong, X. (2024). Autonomy support, peer relations, and teacher-student interactions: implications for psychological well-being in language learning. *Frontiers in Psychology*, 15. <https://doi.org/10.3389/fpsyg.2024.1358776>
- Xie, J. (2025). Combining Phenomenology with Bourdieu to Explore Perceptions of Research Quality: a Methodological Perspective from China. *Methodology Insight*. <https://doi.org/10.71290/mi01020002>
- Yang, C., Hou, S., & Chen, J. (2025). Enhancing interpreting performance, engagement, and self-regulated learning through peer assessment: an attachment theory in perspective. *Frontiers in Psychology*, 16. <https://doi.org/10.3389/fpsyg.2025.1604825>