

RMPC-Vol.1. N1. 007

**Analítica multimodal del aprendizaje para la evaluación inclusiva:
inteligencia artificial, comportamiento y nuevas métricas educativas**

*Multimodal learning analytics for inclusive assessment: artificial intelligence,
behavior, and new educational metrics*

Autores:

Graciela Mamani Torres
Centro de Investigación, Desarrollo e Innovación JOGB
Oruro - Bolivia
docentetutor2020@gmail.com
<https://orcid.org/0009-0004-6155-2164>

Wilson Ángel Gutiérrez Rodríguez
Centro de Investigación, Desarrollo e Innovación JOGB
Oruro - Bolivia
vicedecanoft@uto.edu.bo
<https://orcid.org/0000-0001-8188-1441>

Giovanni Juan Maldonado Aguirre
Centro de Investigación, Desarrollo e Innovación JOGB
Oruro - Bolivia
<https://orcid.org/0009-0007-8627-8087>
gmaldonadoda@gmail.com

Autor de correspondencia: Graciela Mamani Torres, docentetutor2020@gmail.com

Recepción: 26-09-2024

Aceptación: 26-11-2024

Publicación: 20-12-2024

Cómo citar este artículo:

Mamani Torres, G., Maldonado Aguirre, GJ, & Gutiérrez Rodríguez, W. Ángel. (2024). Analítica multimodal del aprendizaje para la evaluación inclusiva: inteligencia artificial, comportamiento y nuevas métricas educativas. Revista Multidisciplinaria Perspectiva Científica, 1 (1), 1-22.
<https://rmpc.editorialgutierrez.com/index.php/rmpc/article/view/10>

RESUMEN

La analítica multimodal del aprendizaje (MMLA) integra datos procedentes de múltiples fuentes registros digitales, audio, video, movimiento, posición, mirada y señales fisiológicas para comprender procesos educativos que no pueden representarse mediante una única métrica. Esta revisión narrativa tuvo como objetivo analizar, entre 2019 y 2023, las oportunidades y limitaciones de la MMLA apoyada por inteligencia artificial para una evaluación inclusiva, con especial atención a la interpretación del comportamiento, la validez de nuevas métricas, la retroalimentación y los riesgos éticos. Se integró evidencia sobre diseño de sistemas, fusión multimodal, evaluación formativa, aprendizaje colaborativo, simulación, privacidad, equidad e inteligencia artificial. Los resultados muestran que la MMLA puede ampliar la evaluación más allá de calificaciones y clics, haciendo visibles procesos espaciales, comunicativos y colaborativos y facilitando retroalimentación basada en evidencia. Sin embargo, la literatura continúa concentrada en educación superior, muestras reducidas y entornos experimentales; además, persisten problemas de generalización, transparencia, privacidad, consentimiento y posible confusión entre conducta observable y aprendizaje. Se concluye que una evaluación multimodal inclusiva requiere combinar métricas pedagógicamente válidas, inteligencia artificial interpretable, análisis desagregado, participación de estudiantes y docentes, minimización de datos y supervisión humana. Las nuevas métricas deben utilizarse para ampliar las formas de demostrar aprendizaje y no para imponer patrones conductuales únicos como norma de rendimiento.

Palabras clave: analítica multimodal del aprendizaje; evaluación inclusiva; inteligencia artificial; comportamiento; métricas educativas.

ABSTRACT

Multimodal learning analytics (MMLA) integrates data from multiple sources including digital traces, audio, video, movement, positioning, gaze, and physiological signals to understand educational processes that cannot be represented through a single metric. This narrative review examined, between 2019 and 2023, the opportunities and limitations of artificial intelligence-supported MMLA for inclusive assessment, with particular attention to behavioral interpretation, the validity of emerging metrics, feedback, and ethical risks. Evidence on system design, multimodal fusion, formative assessment, collaborative learning, simulation, privacy, fairness, and artificial intelligence was integrated. Findings indicate that MMLA can extend assessment beyond grades and clicks by making spatial, communicative, and collaborative processes visible and enabling evidence-based feedback. However, research remains concentrated in higher education, small samples, and experimental settings, while important challenges persist regarding generalizability, transparency, privacy, consent, and the potential conflation of observable behavior with learning. The review concludes that inclusive multimodal assessment requires pedagogically valid metrics, interpretable artificial intelligence, disaggregated analysis, meaningful participation of students and teachers, data minimization, and human oversight. Emerging metrics should expand the ways learners can demonstrate learning rather than impose a single behavioral pattern as the norm for academic performance.

Keywords: multimodal learning analytics; inclusive assessment; artificial intelligence; behavior; educational metrics.

1. INTRODUCCIÓN

La evaluación educativa se ha apoyado tradicionalmente en productos observables como exámenes, tareas, calificaciones y registros de participación. La expansión de entornos digitales y tecnologías de sensorización ha permitido incorporar nuevas fuentes de evidencia sobre cómo ocurre el aprendizaje. La analítica multimodal del aprendizaje integra datos provenientes de texto, audio, video, movimiento, posición, mirada, interacción digital y señales fisiológicas para estudiar procesos complejos. Entre 2019 y 2023, el campo insistió en que estas tecnologías debían vincularse con teoría del aprendizaje, decisiones pedagógicas y criterios éticos, evitando que la disponibilidad de datos determinara por sí misma aquello que se considera relevante evaluar (1-4).

El desarrollo reciente de inteligencia artificial y aprendizaje automático ha aumentado la capacidad para procesar grandes volúmenes de información multimodal. Algoritmos de clasificación, reconocimiento, modelado temporal y fusión de datos pueden identificar patrones difíciles de observar directamente por docentes, especialmente en aprendizaje colaborativo, simulaciones y actividades presenciales complejas. Las revisiones recientes muestran un crecimiento sostenido de la investigación que combina IA y MMLA y señalan beneficios relacionados con extracción de características, personalización y retroalimentación. No obstante, la integración pedagógica de estos avances permanece menos desarrollada que su componente técnico, lo que limita su impacto real sobre evaluación y aprendizaje (5-8).

La evaluación inclusiva introduce una exigencia adicional: las métricas deben permitir que estudiantes diversos demuestren sus aprendizajes sin convertir diferencias de comportamiento en déficits. La analítica multimodal puede ampliar la evidencia disponible y reducir la dependencia de una única forma de participación, pero también puede producir el efecto contrario si interpreta contacto visual, movimiento, tiempo de respuesta, proximidad, expresión facial o participación verbal como indicadores universales de atención y competencia. La literatura sobre learning analytics e inclusión muestra que la representación de estudiantes con discapacidad continúa siendo limitada, por lo que la

validez de muchas inferencias para poblaciones diversas todavía necesita investigación específica (9-11).

Una contribución importante de la MMLA consiste en hacer visibles procesos que tradicionalmente resultaban difíciles de evaluar. En simulaciones clínicas, por ejemplo, los datos de posicionamiento permiten analizar comportamientos espaciales y colaboración; en aprendizaje cooperativo, audio, movimiento y proximidad pueden aportar información sobre dinámicas grupales; y en formación docente, la integración de diferentes señales puede apoyar evaluación formativa y reflexión. Estas posibilidades desplazan el interés desde el resultado final hacia el proceso mediante el cual se construye el aprendizaje. Sin embargo, transformar señales físicas o digitales en métricas educativas requiere demostrar que la medida representa realmente el constructo que pretende evaluar (12-16).

La validez constituye por ello uno de los problemas centrales. Un sensor puede medir con precisión una posición corporal, pero esa precisión técnica no demuestra que la posición represente colaboración, compromiso o aprendizaje. Del mismo modo, un algoritmo puede clasificar patrones con alta exactitud y continuar produciendo interpretaciones pedagógicamente débiles. Los marcos recientes sobre diseño, generalización y evaluación de MMLA subrayan la necesidad de reportar características de sensores, procedimientos analíticos, métricas de desempeño y condiciones de implementación. Una evaluación inclusiva requiere además comprobar que esas relaciones no cambian sistemáticamente entre grupos de estudiantes o contextos educativos diferentes (2,6,17-19).

Los riesgos éticos aumentan a medida que la analítica incorpora datos más íntimos. Cámaras, micrófonos, sensores portables y sistemas de posicionamiento pueden registrar información continua sobre estudiantes, generando preocupaciones relacionadas con privacidad, vigilancia, consentimiento, propiedad de los datos y asimetrías de poder. Una revisión sistemática de MMLA identificó una tensión persistente entre producir conocimiento detallado sobre el aprendizaje y evitar prácticas invasivas. Estudios centrados en fairness, accountability, transparency and ethics muestran además que los estudiantes valoran los beneficios de las visualizaciones multimodales, pero demandan

representaciones precisas, transparencia y formas de consentimiento más continuas y significativas (20-23).

En este escenario, la MMLA ofrece una oportunidad para reconsiderar qué cuenta como evidencia educativa, pero su valor inclusivo depende de cómo se diseñen e interpreten las nuevas métricas. Esta revisión narrativa tiene como objetivo analizar la evidencia publicada entre 2019 y 2023 sobre analítica multimodal del aprendizaje aplicada a evaluación, inteligencia artificial y comportamiento, identificando sus aportes para una evaluación inclusiva, las modalidades y métricas emergentes, los principales riesgos de sesgo y privacidad y las condiciones necesarias para una implementación pedagógicamente válida. El análisis busca diferenciar entre medir más aspectos del estudiante y comprender mejor el aprendizaje, dos propósitos que no necesariamente son equivalentes (1,7,9,20,24,25).

2. REVISIÓN DE LA LITERATURA

De la analítica digital a la analítica multimodal

La analítica convencional utiliza principalmente rastros generados dentro de plataformas digitales, como accesos, clics, entregas o resultados. La MMLA amplía este repertorio al incorporar señales del mundo físico y social. Worsley et al. propusieron doce compromisos para consolidar el campo, incluyendo conexión con teoría, consideración del contexto, transparencia metodológica y responsabilidad ética (1). Posteriormente, los estudios de mapeo confirmaron que la investigación se organiza alrededor de contextos de aprendizaje, procesos, sistemas, modalidades y tecnologías, evidenciando una transición desde prototipos aislados hacia arquitecturas capaces de integrar diferentes fuentes de datos (6). La fusión de datos constituye un componente central porque la multimodalidad no consiste simplemente en acumular sensores. La integración debe permitir relaciones interpretables entre señales heterogéneas y fenómenos educativos. Chango et al. revisaron métodos de data fusion utilizados en MMLA y educational data mining, mostrando diferentes niveles de integración y desafíos relacionados con sincronización, dimensionalidad y calidad de los datos (5). Estas dificultades adquieren importancia evaluativa: combinar señales de baja

calidad puede producir una métrica aparentemente sofisticada pero pedagógicamente inestable.

Inteligencia artificial y nuevas capacidades analíticas

La IA fortaleció especialmente las etapas de extracción de características, reconocimiento de patrones y modelado. El manual especializado de MMLA y el estudio de mapeo del campo documentaron métodos computacionales para integrar señales heterogéneas y mostraron que el desarrollo técnico debía conectarse con teoría, diseño pedagógico y evaluación responsable (6,24). Las aplicaciones empíricas disponibles también combinaron aprendizaje automático con datos procedentes de diferentes fuentes para examinar procesos como distracción, interacción entre pares y rendimiento, aunque sus resultados dependieron del contexto y de la calidad de las medidas utilizadas (25).

Estas capacidades permiten automatizar parte de la observación que antes requería codificación manual. Sin embargo, automatizar una descripción no equivale a automatizar una evaluación. El significado educativo de una característica depende del contexto, la tarea y la teoría que la relaciona con el aprendizaje. Por ello, los sistemas de IA deberían funcionar como instrumentos para aumentar la observación docente y no como sustitutos del juicio pedagógico. La interpretabilidad resulta especialmente importante cuando una métrica multimodal puede afectar retroalimentación, clasificación o decisiones de progreso.

Comportamiento, colaboración y evaluación formativa

Uno de los campos con mayor desarrollo es el aprendizaje colaborativo. Las señales de voz, proximidad, posición y movimiento permiten estudiar cómo los miembros de un equipo distribuyen tareas e interactúan. Yan et al. demostraron que datos de posicionamiento de 304 estudiantes de ciencias de la salud podían relacionarse con desempeño grupal y alcanzar un AUC de 0,74 al distinguir equipos con diferente desempeño colaborativo (12). Estos resultados muestran que las métricas espaciales pueden aportar evidencia complementaria, aunque no deben interpretarse como equivalentes directos del aprendizaje.

En formación sanitaria, Fernandez-Nieto et al. utilizaron narrativas construidas con datos multimodales de equipos para orientar la reflexión de estudiantes sobre su desempeño (16). De forma complementaria, Martinez-Maldonado et al. documentaron un despliegue de MMLA durante dos años con 399 estudiantes y 17 educadores, identificando lecciones técnicas, pedagógicas, participativas, éticas y de sostenibilidad en condiciones auténticas (15). Estos estudios son relevantes porque desplazan la MMLA desde la detección hacia el cierre del ciclo analítico: los datos adquieren valor cuando regresan a estudiantes y docentes en forma de evidencia comprensible que puede orientar reflexión y mejora.

Evaluación inclusiva y diversidad de evidencias

Una evaluación inclusiva requiere múltiples formas de demostrar aprendizaje. Desde esta perspectiva, la MMLA puede reducir la dependencia de una única señal y reconocer que participación, colaboración y resolución de problemas pueden expresarse de maneras distintas. La revisión de Khalil et al. sobre learning analytics e inclusión encontró potencial para reducir discriminación, mejorar retención y validar diseños para grupos marginados, pero también mostró que la investigación sobre discapacidad continúa siendo un área emergente (9). Este vacío obliga a evitar afirmaciones de inclusión que no hayan sido comprobadas con poblaciones diversas.

La multimodalidad puede favorecer inclusión cuando ofrece evidencia complementaria y permite contextualizar un comportamiento, pero puede generar exclusión si convierte una conducta mayoritaria en norma. Métricas basadas en mirada, postura, proximidad o cantidad de intervenciones verbales requieren especial cautela. Una persona puede aprender eficazmente sin mostrar el patrón conductual esperado por el modelo. Por ello, las métricas inclusivas deberían evaluar constructos educativos y no conformidad conductual, permitir múltiples trayectorias hacia un mismo resultado y conservar mecanismos para revisión humana.

Privacidad, vigilancia y consentimiento

La expansión de modalidades aumenta la sensibilidad de los datos. Prinsloo et al. mostraron que buena parte de la MMLA se desarrolla en espacios intermedios entre experimentación y práctica educativa y entre apoyo al aprendizaje e intrusión, debido al uso de sensores, dispositivos conectados e IA (20). La privacidad no puede tratarse únicamente mediante anonimización posterior; debe influir en qué datos se recogen, durante cuánto tiempo, quién puede acceder a ellos y qué inferencias se consideran legítimas.

Li et al. examinaron la disposición de estudiantes universitarios a consentir el uso de sus datos para learning analytics y encontraron diferencias de participación y consentimiento entre subpoblaciones, además de relaciones con la confianza institucional y la preocupación por la recopilación de datos (22). Estos resultados son esenciales para una evaluación inclusiva: un sistema que depende de datos cedidos de manera desigual puede representar peor a ciertos grupos y reforzar disparidades, por lo que la transparencia, la agencia estudiantil y la revisión continua de las condiciones de consentimiento deben formar parte del diseño.

Diseño, generalización y aplicación educativa

La transición desde prototipos hacia sistemas utilizables exige resolver problemas de escalabilidad y generalización. Yan et al. revisaron 96 artículos y recomendaron mejorar estándares de reporte, fortalecer colaboración interdisciplinaria y desarrollar lineamientos éticos frente a riesgos de sesgo, privacidad e igualdad (7). Chejara et al. propusieron un marco específico para evaluar y reportar generalización de modelos de aprendizaje automático en MMLA, destacando que un sistema útil en un laboratorio puede perder rendimiento cuando cambia la población, el entorno o la tecnología (2).

Las investigaciones del periodo confirman que la integración tecnológica debe estar subordinada al escenario de aprendizaje y sus restricciones. La revisión de Alwahaby et al. encontró evidencia débil de impacto de la MMLA en resultados de aprendizaje en contextos reales y una atención todavía escasa a los desafíos éticos (18). En conjunto, la literatura

sugiere que una MMLA inclusiva necesita ser válida, sostenible, interpretable y adaptable, además de técnicamente precisa.

3. METODOLOGÍA

Diseño de la revisión

Se realizó una revisión narrativa de literatura científica publicada entre 2019 y 2023. El diseño se seleccionó para integrar evidencia heterogénea procedente de estudios empíricos, revisiones sistemáticas, estudios de mapeo, trabajos de diseño y contribuciones metodológicas relacionadas con analítica multimodal del aprendizaje, inteligencia artificial, evaluación, comportamiento, inclusión y ética. La revisión no pretendió estimar un efecto agregado, sino identificar tendencias, convergencias, limitaciones y principios aplicables a la construcción de una evaluación multimodal inclusiva.

Estrategia de búsqueda

La búsqueda se orientó mediante combinaciones de términos como “multimodal learning analytics”, “artificial intelligence”, “assessment”, “formative assessment”, “collaborative learning”, “inclusive education”, “disability”, “fairness”, “privacy”, “ethics”, “sensors”, “positioning”, “video analytics” y “multimodal data fusion”. Se priorizaron publicaciones con DOI o URL editorial verificable y trabajos aparecidos en revistas y actas académicas especializadas en tecnología educativa, learning analytics, inteligencia artificial educativa y evaluación.

Criterios de inclusión

Se incluyeron publicaciones de 2019 a 2023 que abordaran directamente MMLA o aportaran evidencia necesaria para interpretar sus implicaciones en evaluación inclusiva. Se consideraron estudios sobre integración multimodal, sensores, inteligencia artificial, aprendizaje colaborativo, simulación, retroalimentación, generalización, privacidad, fairness y diseño. Se priorizaron trabajos con metodología explícita, población o corpus identificable y resultados relacionados con procesos de aprendizaje o evaluación.

Criterios de exclusión

Se excluyeron trabajos anteriores a 2019 y posteriores a 2023, documentos sin trazabilidad bibliográfica suficiente, aplicaciones de análisis multimodal sin propósito educativo y estudios donde la multimodalidad fuera únicamente una característica técnica sin relación con aprendizaje, evaluación o comportamiento. También se evitó interpretar como evidencia directa de inclusión cualquier estudio que no hubiera examinado diversidad, discapacidad o condiciones de equidad.

Síntesis de la evidencia

Las publicaciones fueron organizadas en seis dimensiones: modalidades y fusión de datos; inteligencia artificial; comportamiento y colaboración; evaluación formativa; inclusión y validez; y privacidad, fairness y diseño responsable. La síntesis comparó qué señales se utilizaron, qué constructos se pretendieron representar, qué métricas se generaron y qué riesgos fueron reportados. Se distinguió entre precisión técnica de sensores o modelos y validez pedagógica de las inferencias.

Rigor y trazabilidad

Las afirmaciones principales se vincularon con referencias numeradas en estilo Vancouver. Se priorizó la trazabilidad mediante DOI y fuentes editoriales verificables. Debido al carácter narrativo, no se aplicó metaanálisis ni protocolo PRISMA de revisión sistemática. La interpretación consideró la concentración de estudios en educación superior, el uso frecuente de muestras pequeñas y la limitada validación con poblaciones diversas como restricciones para generalizar conclusiones sobre evaluación inclusiva.

4. RESULTADOS

La síntesis de la literatura muestra una expansión de la MMLA entre 2019 y 2023. El campo avanzó desde aplicaciones de sensorización y marcos conceptuales hacia estudios de fusión, posicionamiento, voz, video, interacción y aprendizaje automático capaces de producir métricas sobre colaboración, comportamiento y procesos de aprendizaje. Sin embargo, el

crecimiento técnico no fue acompañado por una validación equivalente en diversidad de poblaciones y contextos. Las revisiones del periodo señalaron evidencia limitada de impacto en entornos auténticos y una representación todavía insuficiente de discapacidad e inclusión (9,18).

Modalidades y nuevas métricas educativas

Las modalidades más recurrentes incluyen registros digitales, audio, video, posición, movimiento, mirada y señales fisiológicas. Su combinación permite construir métricas que describen participación verbal, proximidad, distribución espacial, secuencias de interacción, postura y coordinación grupal. Estas métricas amplían la evaluación de procesos, pero su utilidad depende de la relación demostrable con el constructo educativo. Los estudios de simulación y colaboración muestran que posición y comunicación pueden aportar evidencia sobre desempeño de equipos, mientras los trabajos metodológicos insisten en que la fusión debe reportarse con suficiente detalle para ser reproducible y generalizable (2,5,7,12-16).

Tabla 1. Modalidades, métricas emergentes y precauciones para evaluación inclusiva

Modalidad	Métrica posible	Aporte evaluativo	Riesgo inclusivo	Salvaguarda
Registros digitales	Secuencias, tiempo, frecuencia	Proceso de interacción	Confundir actividad con aprendizaje	Triangulación
Audio	Turnos, duración, patrones de habla	Colaboración y comunicación	Penalizar estilos comunicativos diversos	Interpretación contextual
Video/postura	Gestos, orientación, postura	Dinámica no verbal	Normativizar comportamiento corporal	No usar como proxy único
Posicionamiento	Proximidad, movimiento, zonas	Coordinación espacial	Inferir participación desde cercanía	Validación con la tarea
Eye tracking	Fijaciones y trayectorias	Atención visual a recursos	Equiparar mirada con comprensión	Complementar con desempeño

Modalidad	Métrica posible	Aporte evaluativo	Riesgo inclusivo	Salvaguarda
Fisiología	Activación y variabilidad	Procesos afectivos/cognitivos	Inferencias sensibles o ambiguas	Minimización y consentimiento
Multimodal fusion	Patrones combinados	Visión integral del proceso	Amplificación de errores/sesgos	Validación por modalidad

Inteligencia artificial y automatización

La IA aparece principalmente en extracción de características, reconocimiento de patrones, clasificación y predicción. El manual de MMLA de 2022 reúne métodos y sistemas para analizar datos complejos del aprendizaje, al tiempo que subraya la necesidad de conectar las técnicas con la teoría y con decisiones pedagógicas responsables (24). En un estudio empírico, Liao y Wu combinaron datos de discurso, cuestionarios e indicadores derivados mediante aprendizaje automático para examinar distracción digital, aprendizaje entre pares y rendimiento; el trabajo ilustra el potencial de la automatización, pero también la dependencia contextual de las métricas resultantes (25).

Evaluación formativa, colaboración y simulación

La evidencia más convincente procede de escenarios donde las métricas regresan a docentes y estudiantes como retroalimentación. En simulación sanitaria, Yan et al. analizaron datos de 304 estudiantes organizados en 76 equipos y obtuvieron un AUC de 0,74 para diferenciar desempeño colaborativo mediante posicionamiento, además de relaciones entre comportamiento espacial y rendimiento grupal (12). Fernandez-Nieto et al. mostraron cómo la narración de datos multimodales podía orientar reflexión sobre el trabajo en equipo (16), mientras Martinez-Maldonado et al. documentaron los retos y aprendizajes de un despliegue auténtico de dos años (15). Estos hallazgos sugieren que el valor evaluativo aumenta cuando las métricas facilitan interpretación y acción en lugar de limitarse a clasificar.

Tabla 2. Estudios representativos 2019-2023 y contribución a la evaluación

Estudio	Diseño/corpus	Tecnología o enfoque	Resultado principal	Contribución
Worsley et al. (1)	Marco conceptual	MMLA	12 compromisos	Fundamentos responsables
Yan et al. (7)	Revisión, 96 trabajos	MMLA	Retos de escala, sostenibilidad y ética	Criterios de implementación
Yan et al. (12)	304 estudiantes/76 equipos	Posicionamiento	AUC 0,74 en desempeño colaborativo	Evaluación de simulación
Ouhaichi et al. (6)	Mapeo sistemático 399	MMLA + IA	14 tópicos en 4 temas	Estado del campo
Martinez-Maldonado et al. (15)	estudiantes/17 educadores; despliegue de 2 años	MMLA en contexto auténtico	Lecciones técnicas, pedagógicas y éticas	Condiciones de implementación
Li et al. (22)	4.000 estudiantes invitados; encuesta posterior	Consentimiento en learning analytics	Diferencias de participación, confianza y privacidad	Agencia y equidad de datos
Giannakos et al. (24)	Manual especializado	Métodos y sistemas MMLA	Síntesis de técnicas y fundamentos	Base metodológica del campo
Liao y Wu (25)	Estudio empírico	Datos multimodales + aprendizaje automático	Distracción, pares y rendimiento	Aplicación contextual de IA

Inclusión, fairness y validez

La evidencia específicamente inclusiva es menor que la evidencia técnica. La revisión de Khalil et al. encontró 26 estudios sobre learning analytics, inclusión y discapacidad y concluyó que existe potencial para reducir discriminación y apoyar grupos marginados, pero también vacíos importantes (9). En MMLA, este problema se intensifica porque las métricas pueden basarse en comportamientos corporales y comunicativos. Una métrica inclusiva

debe admitir que diferentes estudiantes pueden alcanzar el mismo aprendizaje mediante patrones de interacción distintos. Por ello, la comparación con un comportamiento promedio no debe utilizarse automáticamente como criterio de calidad.

Privacidad, transparencia y consentimiento

Los resultados éticos son consistentes: cuanto más detallada es la observación multimodal, mayor es el riesgo de intrusión. Prinsloo et al. describieron la tensión entre apoyo y vigilancia en MMLA, especialmente cuando se utilizan sensores, wearables y sistemas adaptativos (20). Li et al. mostraron que la confianza institucional, la preocupación por la recopilación y las diferencias entre subpoblaciones influyen en la disposición a consentir el uso de datos educativos (22). La revisión general de privacidad en learning analytics identificó 47 trabajos y confirmó que la protección de datos sigue siendo una preocupación estructural (19). La inclusión, por tanto, requiere también que los estudiantes puedan rechazar modalidades especialmente sensibles sin quedar académicamente perjudicados.

Tabla 3. Marco propuesto para evaluación multimodal inclusiva

Dimensión	Pregunta de control	Riesgo	Criterio recomendado
Constructo	¿Qué aprendizaje representa la señal?	Proxy sin validez	Justificación pedagógica
Multimodalidad	¿Cada modalidad aporta información necesaria?	Sobre-recolección	Minimización de datos
IA	¿Puede explicarse la inferencia?	Caja negra	Interpretabilidad
Diversidad	¿Funciona para perfiles distintos?	Sesgo de grupo	Validación desagregada
Comportamiento	¿Se confunde conducta con competencia?	Normalización	Múltiples evidencias
Privacidad	¿El dato es proporcional al propósito?	Vigilancia	Privacy by design
Consentimiento	¿Puede revisarse o retirarse?	Consentimiento nominal	Consentimiento continuo

Dimensión	Pregunta de control	Riesgo	Criterio recomendado
Retroalimentación	¿La métrica permite actuar?	Clasificación sin apoyo	Feedback formativo
Decisión	¿Existe revisión humana?	Automatización de alto impacto	Human-in-the-loop
Accesibilidad	¿Existe alternativa equivalente?	Exclusión por tecnología	Vías alternativas

5. DISCUSIÓN

Los resultados confirman que la MMLA puede transformar la evaluación al ampliar la evidencia disponible sobre procesos que los instrumentos convencionales observan de manera limitada. El análisis de posición, voz, movimiento y secuencias de interacción permite describir cómo se organiza una actividad y no únicamente cuál fue su resultado. Esta capacidad resulta especialmente valiosa en colaboración, simulación y tareas complejas, donde una calificación final oculta procesos relevantes. Sin embargo, la multimodalidad no garantiza por sí misma una evaluación mejor. La calidad depende de que cada señal tenga una relación teórica y empírica defendible con aquello que se pretende evaluar (1,6,12,15).

La inteligencia artificial acelera esta transformación al automatizar tareas de extracción y modelado que serían impracticables manualmente. El manual de MMLA muestra la diversidad de métodos disponibles para analizar señales complejas y convertirlas en representaciones útiles, pero también la necesidad de anclar esos métodos en teoría, contexto y propósitos educativos explícitos (24). Esta asimetría representa un riesgo: la capacidad técnica para detectar un patrón puede avanzar más rápido que la capacidad educativa para explicar su significado. En evaluación inclusiva, ninguna predicción debería adquirir valor normativo únicamente porque un algoritmo pueda producirla de forma consistente.

Los estudios de posicionamiento ilustran esta diferencia. La relación encontrada entre comportamientos espaciales y desempeño de equipos de simulación demuestra que el espacio contiene información educativa útil (12). No obstante, proximidad, movimiento o

permanencia en una zona no son competencias por sí mismos. Su interpretación depende de la tarea, roles y contexto. Aplicar una métrica espacial desarrollada en simulación sanitaria a otro entorno sin nueva validación podría producir conclusiones incorrectas. Los marcos de generalización y escalabilidad son, por ello, componentes centrales de una evaluación responsable (2,7).

La evaluación formativa parece constituir uno de los usos más defendibles de MMLA. Cuando las visualizaciones y métricas se emplean para provocar reflexión, el sistema no necesita convertir cada señal en una sentencia sobre capacidad. Fernandez-Nieto et al. mostraron que las narrativas de datos multimodales pueden guiar la reflexión sobre el trabajo en equipo (16). Martinez-Maldonado et al. documentaron, además, las condiciones y dificultades de incorporar estas soluciones en una experiencia educativa auténtica de larga duración (15). En estos usos, la analítica funciona como evidencia para diálogo y no como sustituto de la evaluación profesional.

Desde la perspectiva inclusiva, la principal advertencia consiste en no convertir el comportamiento promedio en estándar universal. Un estudiante puede participar menos verbalmente y contribuir de otras maneras; otro puede evitar contacto visual y mantener alto nivel de comprensión; diferentes patrones de movimiento pueden responder a estrategias personales o necesidades de accesibilidad. Si un sistema interpreta estas diferencias como falta de compromiso, la nueva métrica reproduce una forma de normalización. La escasez de estudios específicamente centrados en discapacidad e inclusión impide asumir que los modelos existentes funcionan de forma equivalente para toda la diversidad estudiantil (9).

Este problema afecta también a fairness. Una representación multimodal puede parecer más completa porque incorpora muchas variables, pero aumentar dimensiones no elimina sesgos. Una modalidad sesgada puede contaminar la fusión y varias señales correlacionadas pueden reforzar una inferencia incorrecta. La evaluación debería incluir análisis desagregados, sensibilidad a contextos y mecanismos de revisión. La multimodalidad debe

utilizarse para ofrecer diferentes evidencias de aprendizaje, no para multiplicar criterios mediante los cuales un estudiante pueda ser considerado atípico.

La privacidad constituye un límite estructural. MMLA puede registrar información que excede ampliamente lo que un estudiante produce conscientemente para ser evaluado. Voz, video, ubicación o señales fisiológicas pueden revelar aspectos sensibles y permitir inferencias secundarias. Prinsloo et al. advierten que el campo opera en una frontera entre apoyo y encroachment (20), mientras Jin et al. muestran que estudiantes reclaman mayor control y transparencia (22). Una evaluación inclusiva no debería obligar a aceptar vigilancia intensiva como condición para acceder a oportunidades educativas.

En conjunto, los hallazgos apoyan una concepción de MMLA como infraestructura de evidencia complementaria. Su finalidad debería ser enriquecer la comprensión del aprendizaje, identificar barreras, mejorar feedback y permitir múltiples formas de participación. La IA puede aumentar la capacidad analítica, pero las decisiones de alto impacto requieren juicio humano, explicabilidad y posibilidad de impugnación. El avance más importante no consiste en medir cada comportamiento, sino en seleccionar únicamente aquellas señales cuya utilidad pedagógica, proporcionalidad ética y validez para poblaciones diversas puedan justificarse.

6. CONCLUSIONES

La analítica multimodal del aprendizaje amplía las posibilidades de evaluación al integrar evidencias digitales, espaciales, comunicativas, visuales y fisiológicas. Entre 2019 y 2023, la literatura mostró avances importantes en fusión de datos, posicionamiento, análisis colaborativo, retroalimentación e incorporación de inteligencia artificial. Estas tecnologías permiten observar procesos educativos que las calificaciones y registros tradicionales representan de forma incompleta.

El principal aporte para una evaluación inclusiva no consiste en aumentar el número de variables observadas, sino en diversificar las formas mediante las cuales se interpreta el aprendizaje. La multimodalidad puede reducir la dependencia de una única conducta o

producto, siempre que las métricas se utilicen como evidencias complementarias y no como estándares rígidos de comportamiento.

La inteligencia artificial incrementa la capacidad para procesar datos multimodales, pero la precisión técnica no garantiza validez educativa. Cada inferencia debe justificar qué constructo representa, cómo fue validada, para qué población funciona y qué errores puede producir. Las métricas derivadas de postura, mirada, proximidad, voz o fisiología requieren especial prudencia porque pueden confundirse fácilmente con atención, compromiso o competencia.

La privacidad, fairness, transparencia y consentimiento son condiciones de validez y no elementos accesorios. Los sistemas inclusivos deben minimizar datos, ofrecer alternativas accesibles, permitir revisión de inferencias y conservar supervisión humana. Los estudiantes y docentes deben participar en el diseño de métricas y visualizaciones para asegurar que las representaciones sean comprensibles y pedagógicamente útiles.

En conclusión, una MMLA verdaderamente inclusiva debe pasar de la lógica de medir más a la lógica de comprender mejor. Las nuevas métricas educativas tienen valor cuando ayudan a identificar procesos y barreras, enriquecen la retroalimentación y reconocen diferentes formas de demostrar aprendizaje. El desafío para la investigación futura es desarrollar sistemas de IA multimodal interpretables, generalizables y evaluados con poblaciones diversas, capaces de ampliar la equidad sin transformar la observación educativa en vigilancia ni la diversidad conductual en desviación.

REFERENCIAS BIBLIOGRÁFICAS

1. Worsley M, Martinez-Maldonado R, D'Angelo C. A new era in multimodal learning analytics: twelve core commitments to ground and grow MMLA. J Learn Anal. 2021;8(3):10-27. <https://doi.org/10.18608/jla.2021.7361>.
2. Chejara P, Prieto LP, Ruiz-Calleja A, Rodríguez-Triana MJ, Shankar SK, Kasepalu R. EFAR-MMLA: an evaluation framework to assess and report generalizability of machine

- learning models in MMLA. Sensors. 2021;21(8):2863. <https://doi.org/10.3390/s21082863>.
3. Mangaroska K, Martinez-Maldonado R, Vesin B, Gašević D. Challenges and opportunities of multimodal data in human learning: the computer science students' perspective. J Comput Assist Learn. 2021;37(4):1030-1047. <https://doi.org/10.1111/jcal.12542>.
4. Dominguez F, Ochoa X, Zambrano D, Camacho K, Castells J. Scaling and adopting a multimodal learning analytics application in an institution-wide setting. IEEE Trans Learn Technol. 2021;14(3):400-414. <https://doi.org/10.1109/TLT.2021.3100778>.
5. Chango W, Lara JA, Cerezo R, Romero C. A review on data fusion in multimodal learning analytics and educational data mining. WIREs Data Min Knowl Discov. 2022;12(4):e1458. <https://doi.org/10.1002/widm.1458>.
6. Ouhaichi H, Spikol D, Vogel B. Research trends in multimodal learning analytics: a systematic mapping study. Comput Educ Artif Intell. 2023;4:100136. <https://doi.org/10.1016/j.caeai.2023.100136>.
7. Yan L, Zhao L, Gašević D, Martinez-Maldonado R. Scalability, sustainability, and ethicality of multimodal learning analytics. In: LAK22: 12th International Learning Analytics and Knowledge Conference. New York: ACM; 2022. p.13-23. <https://doi.org/10.1145/3506860.3506862>
8. Yan L, Martinez-Maldonado R, Gallo Cordoba B, Deppeler J, Corrigan D, Gašević D. Mapping from proximity traces to socio-spatial behaviours and student progression at the school. Br J Educ Technol. 2022;53(6):1645-1664. <https://doi.org/10.1111/bjet.13203>.
9. Khalil M, Slade S, Prinsloo P. Learning analytics in support of inclusiveness and disabled students: a systematic review. J Comput High Educ. Published online March 14, 2023. <https://doi.org/10.1007/s12528-023-09363-4>
10. Baker RS, Hawn A. Algorithmic bias in education. Int J Artif Intell Educ. 2022;32:1052-1092. <https://doi.org/10.1007/s40593-021-00285-9>.
11. Deho OB, Zhan C, Li J, Liu J. How do the existing fairness metrics and unfairness mitigation algorithms contribute to ethical learning analytics? Br J Educ Technol. 2022;53. <https://doi.org/10.1111/bjet.13217>

12. Yan L, Martinez-Maldonado R, Zhao L, Dix S, Jaggard H, Wotherspoon R, et al. The role of indoor positioning analytics in assessment of simulation-based learning. *Br J Educ Technol*. 2023;54(1):267-292. <https://doi.org/10.1111/bjet.13262>.
13. Zhao L, Yan L, Gašević D, Dix S, Jaggard H, Wotherspoon R, et al. Modelling co-located team communication from voice detection and positioning data in healthcare simulation. In: LAK22. New York: ACM; 2022. p.370-380. <https://doi.org/10.1145/3506860.3506935>
14. Yan L, Martinez-Maldonado R, Zhao L, et al. How do teachers use open learning spaces? Mapping from teachers' socio-spatial data to spatial pedagogy. In: LAK22. New York: ACM; 2022. p.87-97. <https://doi.org/10.1145/3506860.3506872>
15. Martinez-Maldonado R, Echeverria V, Fernandez-Nieto G, Yan L, Zhao L, Alfredo R, et al. Lessons learnt from a multimodal learning analytics deployment in-the-wild. *ACM Trans Comput-Hum Interact*. 2023;31(1):Article. 8. <https://doi.org/10.1145/3622784>.
16. Fernandez-Nieto GM, Echeverria V, Buckingham Shum S, Mangaroska K, Kitto K, Palominos E, et al. Storytelling with learner data: guiding student reflection on multimodal team data. *IEEE Trans Learn Technol*. 2021;14(5):695-708. <https://doi.org/10.1109/TLT.2021.3131842>.
17. Nasir J, Bruno B, Chetouani M, Dillenbourg P. What if social robots look for productive engagement? Automated assessment of goal-centric engagement in learning applications. *Int J Soc Robot*. 2022;14:55-71. <https://doi.org/10.1007/s12369-021-00766-w>.
18. Alwahaby H, Cukurova M, Papamitsiou Z, Giannakos M. The evidence of impact and ethical considerations of multimodal learning analytics: a systematic literature review. In: Giannakos M, Spikol D, Di Mitri D, Sharma K, Ochoa X, Hammad R, editors. *The Multimodal Learning Analytics Handbook*. Cham: Springer; 2022. p.289-325. https://doi.org/10.1007/978-3-031-08076-0_12
19. Liu Q, et al. Understanding privacy and data protection issues in learning analytics using a systematic review. *Br J Educ Technol*. 2023. <https://doi.org/10.1111/bjet.13388>
20. Prinsloo P, Slade S, Khalil M. Multimodal learning analytics—in-between student privacy and encroachment: a systematic review. *Br J Educ Technol*. 2023;54:1566-1586. <https://doi.org/10.1111/bjet.13373>.

21. Currier C. Unresolved privacy and ethics issues related to learning analytics in higher education and academic librarianship. *Emerg Libr Inf Perspect.* 2021;4(1). <https://doi.org/10.5206/elip.v4i1.13463>
22. Li W, Sun K, Schaub F, Brooks C. Disparities in students' propensity to consent to learning analytics. *Int J Artif Intell Educ.* 2022;32(3):564-608. <https://doi.org/10.1007/s40593-021-00254-2>.
23. Ouyang F, Dai X, Chen S. Applying multimodal learning analytics to examine the immediate and delayed effects of instructor scaffoldings on small groups' collaborative programming. *Int J STEM Educ.* 2022;9:45. <https://doi.org/10.1186/s40594-022-00361-z>.
24. Giannakos M, Spikol D, Di Mitri D, Sharma K, Ochoa X, Hammad R, editors. *The Multimodal Learning Analytics Handbook*. Cham: Springer; 2022. <https://doi.org/10.1007/978-3-031-08076-0>
25. Liao CH, Wu JY. Deploying multimodal learning analytics models to explore the impact of digital distraction and peer learning on student performance. *Comput Educ.* 2022;190:104599. <https://doi.org/10.1016/j.compedu.2022.104599>.

CONFLICTO DE INTERESES

Los autores afirman que no existen conflictos de intereses en este estudio y que se han seguido éticamente los procesos establecidos por esta revista. Además, aseguran que este trabajo no ha sido publicado parcial ni totalmente en ninguna otra revista.

FINANCIAMIENTO

Los autores no recibieron financiamiento para el desarrollo de esta investigación.

CONTRIBUCIÓN DE AUTORÍA

Autores e iniciales: Graciela Mamani Torres, Wilson Ángel Gutiérrez Rodríguez y Giovanni Juan Maldonado Aguirre

Planteamiento de la idea: GMT, GJMA

Estructuración metodológica: WAGR

Levantamiento y gestión de información: GJMA

Sistematización de datos: GMT, GJMA

Procesamiento e interpretación de resultados: WAGR

Verificación de la información: GJMA

Elaboración de recursos visuales: WAGR

Redacción inicial del manuscrito: GMT

Corrección académica y edición final: GJMA

Todos los autores revisaron y aprobaron la versión final del manuscrito. GMT, WAGR

© 2024; Los autores. Este es un artículo en acceso abierto, distribuido bajo los términos de una licencia Creative Commons (<https://creativecommons.org/licenses/by/4.0>) que permite el uso, distribución y reproducción en cualquier medio, siempre que la obra original sea correctamente citada.