



**Development and psychometric analysis of an instrument for
assessing group dynamics during collaborative problem-
solving in a university context**

**Desarrollo y análisis psicométrico de un instrumento
destinado a la evaluación de dinámicas grupales durante
la resolución colaborativa de problemas en el contexto
universitario**

Para citar este trabajo:

Villon Ortega, C. D. . (2025). Desarrollo y análisis psicométrico de un instrumento destinado a la evaluación de dinámicas grupales durante la resolución colaborativa de problemas en el contexto universitario. Imperium Académico Multidisciplinary Journal, 2(2), 1-17. https://estrellaediciones.com/index.php/imperium_academico/article/view/40

Autores:

Catherine Denisse Villon Ortega
Universidad de Guayaquil
Guayaquil – Ecuador
Catherine.villon@educacion.gob.ec
<https://orcid.org/0009-0002-5160-4931>

Autor de Correspondencia: Catherine Denisse Villon Ortega, Catherine.villon@educacion.gob.ec

RECIBIDO: 06-Marzo-2025

ACEPTADO: 20-Marzo-2025

PUBLICADO 31- Marzo -2025



Resumen

El presente estudio tuvo como objetivo diseñar y evaluar un instrumento para medir los procesos grupales involucrados en la resolución colaborativa de problemas en la educación superior. La Escala para la Evaluación de Procesos Grupales en la Resolución Colaborativa de Problemas (GROUPS) consta de 24 ítems de autoinforme distribuidos en cuatro dimensiones: Exploración y Comprensión, Representación y Formulación, Planificación y Ejecución, y Monitoreo y Reflexión. Esta investigación instrumental se realizó en Chile con una muestra de 939 estudiantes universitarios, quienes respondieron el instrumento tras participar en una tarea colaborativa de resolución de problemas de tres semanas, diseñada conjuntamente por el equipo investigador y docentes. Los análisis de validez estructural mostraron un ajuste adecuado al modelo de cuatro factores (con índices RMSEA y SRMR inferiores a .05, y CFI y TLI superiores a .95), además de altos niveles de confiabilidad en todas las dimensiones (alfa ordinal superior a .80). Se describen los procesos grupales identificados y se discuten las características del instrumento, sus principales hallazgos y las implicaciones para fortalecer las competencias en resolución colaborativa de problemas en la educación superior.

Palabras clave: Medición; Dinámicas grupales; Solución conjunta de problemas

Abstract

The present study aimed to design and evaluate an instrument to measure group processes involved in collaborative problem-solving in higher education. The Group Processes in Collaborative Problem-Solving Scale (GROUPS) consists of 24 self-report items distributed across four dimensions: Exploration and Comprehension, Representation and Formulation, Planning and Execution, and Monitoring and Reflection. This instrumental research was conducted in Chile with a sample of 939 university students, who completed the instrument after participating in a three-week collaborative problem-solving task, jointly designed by the research team and faculty. Structural validity analyses showed an adequate fit to the four-factor model (with RMSEA and SRMR indices below .05, and CFI and TLI above .95), in addition to high levels of reliability in all dimensions (ordinal alpha above .80). The identified group processes are described, and the characteristics of the instrument, its main findings, and the implications for strengthening collaborative problem-solving skills in higher education are discussed.

Keywords: Measurement; Group dynamics; Joint problem-solving

1. Introducción

La resolución colaborativa de problemas (RCP) se entiende como la habilidad para participar eficazmente en un proceso donde dos o más individuos trabajan conjuntamente para solucionar un problema, compartiendo tanto la comprensión como el esfuerzo necesario, y combinando sus conocimientos y capacidades para alcanzar una solución común OECD (2021). Esta competencia resulta esencial en un contexto global cada vez más complejo y dinámico, donde la tecnología digital y las herramientas de comunicación juegan un rol fundamental en la facilitación y el desarrollo de procesos colaborativos Kapur et al. (2020; Zheng et al. (2022).

En un contexto marcado por rápidos avances tecnológicos y la creciente complejidad interdisciplinaria de los desafíos contemporáneos, la resolución efectiva de problemas requiere la colaboración entre individuos con habilidades, formaciones y perspectivas diversas. Por ello, las competencias en resolución colaborativa de problemas (RCP) se consideran esenciales dentro del conjunto de habilidades del siglo XXI Voogt et al., (202); Saavedra et al. (2020). Aunque la resolución de problemas a nivel individual ha sido objeto de estudio durante décadas en disciplinas como la psicología, recientemente ha aumentado el interés en su dimensión



colaborativa. Investigaciones recientes han explorado factores que afectan el desempeño en tareas de RCP, incluyendo características de personalidad, diversidad cognitiva del equipo, conocimientos previos, percepciones mutuas y creencias sobre las capacidades de los compañeros Bielaczyc et al. (2020); Zhang et al. (2023).

La investigación sobre aprendizaje grupal y toma de decisiones colectivas ha sido extensa, incluyendo estudios sobre el aprendizaje colaborativo asistido por computadora Kirschner et al. (2021); Slavin (2020). Sin embargo, persiste la necesidad de profundizar en la comprensión de los procesos y dinámicas grupales específicos que ocurren durante tareas de resolución colaborativa de problemas (RCP), especialmente cuando estas involucran tecnologías de la información y la comunicación (TIC) Scoular et al. (2021); Figueroa et al. (2025). Además, aunque tradicionalmente la resolución de problemas se ha analizado en entornos controlados de laboratorio, en años recientes se ha incrementado el interés por estudiar estas habilidades en contextos reales, donde las interacciones sociales y tecnológicas influyen directamente en el desempeño colaborativo Graesser et al. (2020); Voogt et al. (2021).

El desarrollo de instrumentos para evaluar y monitorear habilidades como la resolución colaborativa de problemas es fundamental para diseñar evaluaciones efectivas, establecer indicadores precisos y planificar intervenciones pedagógicas adecuadas. Este proceso enfrenta el desafío de equilibrar criterios que a menudo resultan conflictivos, tales como la validez ecológica, la validez interna, la posibilidad de aplicación a gran escala y la eficiencia en el tiempo de análisis Azevedo et al., (2021); Graesser et al. (2020). Por ello, la búsqueda de metodologías que satisfagan estos múltiples requerimientos es clave para avanzar en la evaluación de competencias complejas en contextos educativos actuales Pellegrino et al. (2022).

Diversos autores coinciden en la importancia de generar conocimiento sobre las habilidades de resolución colaborativa de problemas (RCP), especialmente para fortalecer la capacidad de escuelas y docentes en el diseño curricular que facilite la adquisición y desarrollo de estas competencias mediante una enseñanza intencionada en contextos educativos Graesser et al. (2020); Voogt et al. (2021). En este sentido, desde 2015, el Programa para la Evaluación Internacional de Estudiantes (PISA) ha incorporado la evaluación de la RCP, aplicando pruebas digitales que involucran interacciones entre humanos y sistemas automatizados, evaluando a más de 540,000 estudiantes de 72 países OCDE (2017); OCDE (2024). Esta inclusión refleja la relevancia creciente de la RCP como competencia clave para enfrentar los retos educativos y sociales actuales, y la necesidad de contar con instrumentos y metodologías que permitan medirla con rigor y en contextos reales.

Diversos estudios basados en la evaluación PISA 2015 evidencian que los estudiantes no migrantes obtienen mejores resultados en tareas de resolución colaborativa de problemas (RCP) cuando asisten a escuelas con una alta proporción de estudiantes migrantes, lo que sugiere que la diversidad en el aula fomenta perspectivas variadas que enriquecen la planificación estratégica para resolver problemas OECD (2017); OECD (2021). Sin embargo, aunque esta evaluación a gran escala ha permitido avances significativos, presenta limitaciones importantes, como la falta de un entorno ecológicamente válido y la imposibilidad de analizar en profundidad los patrones discursivos durante la interacción colaborativa Graesser et al. (2020); Azevedo et al. (2021). Estas restricciones indican la necesidad de complementar las evaluaciones digitales masivas con metodologías que permitan un estudio más detallado y contextualizado de las dinámicas grupales en la resolución colaborativa

El proyecto Assessment and Teaching of 21st-Century Skills (ATC21S) se enfocó en desarrollar mediciones estandarizadas para evaluar diversas habilidades de resolución colaborativa de problemas (RCP) en interacciones entre humanos, con el propósito de facilitar comparaciones en



investigaciones futuras Griffin et al. (2020); Care et al. (2021). Según Graesser et al. (2020), este marco conceptual permite a los investigadores descomponer las habilidades de RCP analizando mecanismos discursivos clave, tales como la participación activa, la adopción de perspectivas y la regulación social durante la colaboración. Así, el ATC21S contribuye a una comprensión más detallada y sistemática de los procesos sociales y cognitivos implicados en la resolución colaborativa de problemas en contextos educativos contemporáneos.

El principal desafío del modelo ATC21S radica en que su aplicación requiere observadores capacitados y un tiempo prolongado para el análisis, lo que limita su utilidad para generar evaluaciones formativas ágiles en contextos educativos. Para superar esta limitación, Scoular et al. (2020) propusieron un enfoque basado en un modelo de inferencia que identifica habilidades clave del marco ATC21S a partir de patrones lingüísticos en interacciones humanas, tales como la negociación, la recopilación de información, la perseverancia, la memoria transaccional y la sistematicidad. Este método ofrece a los docentes indicadores concretos sobre las capacidades de sus estudiantes mediante el análisis de ítems registrados, facilitando así una evaluación más práctica y accesible en entornos educativos Scoular et al. (2020); Griffin et al. (2022).

En la literatura reciente, algunos estudios han abordado la evaluación de las habilidades de resolución colaborativa de problemas (RCP) en la educación superior. Por ejemplo, investigaciones como las de Pazos et al. (2010) desarrollaron un instrumento observacional para valorar el aprendizaje colaborativo en grupos dirigidos por compañeros, considerando dos dimensiones principales: el enfoque en la resolución de problemas y el estilo de interacción grupal. Este tipo de herramientas resulta fundamental para comprender y mejorar las dinámicas de trabajo en equipo en contextos universitarios López et al. (2022); Martínez et al. (2023). Además, la evaluación de estas competencias se ha vinculado con modelos educativos que promueven la formación integral y el desarrollo de habilidades blandas, esenciales para enfrentar los retos actuales en la educación superior Cáceres et al. (2021).

Para el diseño del instrumento, se realizó una observación previa de grupos de estudiantes de primer y segundo año a lo largo de un trimestre, durante una actividad semanal en la que enfrentaban problemas elaborados por los docentes; estos problemas se distinguían por su complejidad, estructura poco definida, la necesidad de integrar diversa información y por no tener una solución inmediata ni única.

El instrumento clasifica los grupos en cuatro categorías: instrucción simple, discusión apoyada, instrucción elaborada y discusión guiada. En este contexto, Visschers-Pleijers et al. (2021) desarrollaron un cuestionario de autoinforme para evaluar la calidad de las interacciones en el aprendizaje grupal dentro de tareas basadas en problemas en educación superior, considerando tres dimensiones clave: preguntas exploratorias, razonamiento acumulativo y gestión de conflictos. Basándose en el enfoque centrado en la calidad de las interacciones, su estudio se centró en grupos tutoriales donde los estudiantes colaboran para integrar y sintetizar información obtenida de manera individual Visschers-Pleijers et al. (2021); Martínez et al. (2023). Este tipo de evaluación resulta fundamental para comprender cómo las dinámicas grupales influyen en el aprendizaje colaborativo en entornos universitarios.

Diversos estudios recientes han concluido que las preguntas exploratorias constituyen el componente más determinante para la productividad en grupos tutoriales. En este sentido, Dindar et al. (2020) evaluaron 25 grupos de tres estudiantes universitarios mediante una tarea simulada por computadora, analizando la relación entre la experiencia metacognitiva y su impacto tanto percibido como objetivo en el desempeño durante una tarea de resolución colaborativa de problemas (RCP). Lo innovador de su propuesta radica en asignar un valor significativo al autoinforme, destacando el papel crucial de la metacognición en el rendimiento de la RCP Dindar



et al. (2020); López et al. (2023).

Evaluar las habilidades de resolución colaborativa de problemas (RCP) en la educación superior es crucial debido a su relevancia para el desempeño efectivo en contextos profesionales y laborales. No obstante, a pesar de su importancia, los niveles actuales de desarrollo de estas competencias son bajos y existe poca claridad sobre las estrategias curriculares que podrían promover su fortalecimiento en los entornos educativos OECD (2017); Zárate (2023). Por ello, resulta indispensable continuar con el desarrollo de instrumentos que permitan evaluar los procesos grupales en diversas tareas de resolución de problemas, fomentando así las interacciones significativas entre estudiantes y facilitando intervenciones pedagógicas más efectivas Pazos et al. (2022); Zambrano et al. (2023). Estas evaluaciones contribuyen a mejorar la planificación curricular y a potenciar las habilidades sociales y cognitivas necesarias para enfrentar los retos actuales en la educación superior y el mundo laboral Rodríguez et al. (2020); Ministerio de Educación Ecuador (2024).

Se requieren instrumentos que, basados en la experiencia y metacognición de los participantes, integren dimensiones clave para analizar la calidad de las interacciones y procesos grupales, tales como la construcción de una comprensión compartida, la planificación, el monitoreo y la ejecución tanto del desempeño individual como colectivo, la comunicación exploratoria y la regulación de tareas. Además, la literatura destaca la necesidad de desarrollar herramientas que ofrezcan información e indicadores sobre el funcionamiento grupal, facilitando a los docentes proporcionar retroalimentación oportuna a los estudiantes. Finalmente, estos instrumentos deben ser sencillos de aplicar y analizar, sin requerir necesariamente la intervención de evaluadores altamente especializados.

El presente estudio se enfocó en la creación de un instrumento de autoinforme destinado a evaluar los procesos grupales involucrados en tareas de resolución colaborativa de problemas (RCP) en el ámbito de la educación superior, proporcionando además evidencia sobre su validez y confiabilidad. La herramienta desarrollada tiene como finalidad medir dichos procesos grupales durante las actividades de RCP, basándose en una conceptualización específica que se detalla a continuación.

Tabla 1
Esquema de conceptualización

Procesos de Resolución de Problemas	Comprensión (Entender y mantener una comprensión compartida)	Acción (Tomar medidas para resolver el problema)	Organización (Establecer y mantener la organización grupal)
Explorar y Comprender	Identificar perspectivas y habilidades de los miembros del equipo.	Descubrir el tipo de interacción colaborativa necesaria según los objetivos.	Comprender los roles necesarios para resolver el problema.
Representar y Formular	Construir una representación compartida y negociar el significado del problema.	Identificar y describir las tareas a realizar.	Describir roles del equipo y protocolos de comunicación/reglas.
Planificar, Ejecutar y Actuar	Comunicar acciones necesarias con los miembros del equipo.	Ejecutar el plan acordado.	Seguir reglas de participación (ej., animar a otros a contribuir).



Monitorear y Reflexionar	Monitorear y corregir la comprensión compartida.	Evaluar resultados de las acciones tomadas.	Proporcionar retroalimentación y ajustar la organización grupal.
---------------------------------	--------------------------------------------------	---------------------------------------------	------------------------------------------------------------------

Fuente: Elaboración propia

El análisis reciente de las habilidades necesarias para la resolución colaborativa de problemas (CPS, por sus siglas en inglés) destaca que muchas de estas competencias requieren un alto nivel de metacognición, entendida como la capacidad de reconocer y regular los propios procesos cognitivos con un propósito específico Dindar et al. (2020); López et al. (2023). Los juicios subjetivos acerca de las tareas grupales y sus resultados pueden influir significativamente en el desarrollo efectivo de las actividades de CPS Dindar et al. (2020). Asimismo, la naturaleza colaborativa de la resolución de problemas implica que los participantes deben expresar, justificar y debatir sus puntos de vista y opiniones, lo que hace que muchos de estos procesos sean observables y analizables OECD (2017); Hesse et al. (2021). Gran parte de estas habilidades son comunicativas, es decir, involucran la capacidad de transmitir información de manera clara y reportar las acciones realizadas en momentos específicos, facilitando así la construcción de una comprensión compartida dentro del grupo OECD (2017); Martínez et al. (2022).

El proceso de Exploración y Comprensión (EU) consiste en analizar el problema a través de la observación y la interacción, recopilando información relevante, identificando limitaciones y obstáculos, y demostrando un entendimiento tanto de los datos iniciales como de la información nueva. Además, implica reconocer las perspectivas y habilidades de los integrantes del equipo, establecer el tipo de colaboración requerida para alcanzar los objetivos planteados y definir los roles necesarios para la resolución efectiva del problema.

El proceso de Representación y Formulación (RF) implica el uso de tablas, gráficos, símbolos o palabras para representar diferentes aspectos del problema y formular hipótesis sobre los factores relevantes y sus interrelaciones, con el objetivo de construir una representación mental clara y coherente del problema. Este proceso también incluye la creación de una representación compartida entre los miembros del equipo, la negociación del significado del problema, la identificación de la tarea a realizar y la definición de los roles y la organización del grupo.

El proceso de Planificación y Ejecución (PE) consiste en diseñar una estrategia o plan para abordar el problema, llevar a cabo dicho plan y establecer objetivos tanto generales como específicos. Este proceso abarca la comunicación constante entre los integrantes del equipo acerca de las acciones a realizar, la implementación coordinada del plan acordado y el respeto por las normas previamente establecidas para garantizar un trabajo efectivo.

El proceso de Monitoreo y Reflexión (MR) abarca la supervisión del avance, la respuesta a la retroalimentación recibida y la reflexión sobre la solución, la información disponible y la estrategia empleada. Este proceso implica mantener una comprensión compartida entre los miembros del equipo, seguir de cerca las acciones y resultados obtenidos, ofrecer retroalimentación continua y ajustar la organización y los roles dentro del grupo cuando sea necesario para mejorar el desempeño.

Este estudio describe la creación de la Escala para la Evaluación de los Procesos Grupales en la Resolución Colaborativa de Problemas (GROUPS), un instrumento diseñado para medir la presencia y el grado de procesos grupales específicos durante una tarea de resolución colaborativa de problemas (CPS). La herramienta se fundamenta en la conceptualización previamente expuesta y evalúa cuatro dimensiones clave de los procesos de resolución, considerando en cada una los tres tipos de habilidades propias de la CPS. A continuación, se



presentan las características principales del instrumento y el procedimiento seguido para su desarrollo.

2. Metodología

Este estudio corresponde a un diseño instrumental cuyo propósito fue crear, desarrollar y evaluar las propiedades psicométricas de un instrumento, siguiendo la metodología propuesta por Montero et al. (2007), que establece los lineamientos para el desarrollo y validación de pruebas psicológicas. Además, el diseño del estudio se basó en los pasos sugeridos por Fonseca et al. (2019) para la elaboración rigurosa de instrumentos psicométricos, asegurando así la validez y confiabilidad del instrumento desarrollado Montero et al. (2007); Fonseca et al. (2019).

Implementación de la prueba piloto

Se empleó una muestra no probabilística compuesta por 120 estudiantes de educación superior inscritos en un instituto profesional técnico en Santiago de Chile, que ofrece programas de duración entre dos y cuatro años. Estos estudiantes, provenientes de diversos programas académicos, estaban cursando una asignatura de educación general en distintas áreas. Antes de completar el instrumento de evaluación, los participantes realizaron una tarea de resolución colaborativa de problemas (CPS, por sus siglas en inglés). La edad promedio de los estudiantes fue de 23.7 años, con una distribución de género donde el 35% correspondía a mujeres (n = 42) y el 65% a hombres (n = 78).

Implementación del test corregido

Se empleó un muestreo no probabilístico por conveniencia para seleccionar a los participantes en la administración de la prueba revisada, dado que el instrumento se aplicó tras la realización de una tarea de resolución colaborativa de problemas (CPS) en clases ya establecidas. El único criterio de inclusión fue pertenecer a dichos cursos y haber completado la tarea CPS descrita en el procedimiento. De esta manera, 939 estudiantes completaron tanto la tarea CPS como la versión revisada del instrumento en su totalidad. Al igual que en la muestra piloto, todos los participantes eran estudiantes de educación superior en un instituto profesional, matriculados en diversos programas. La edad promedio de esta muestra fue de 24,09 años (DE = 12,80), con una distribución de género donde el 56% eran mujeres (n = 526), el 42,4% hombres (n = 402) y un 1,2% (n = 11) se identificaron con otro género.

Herramienta de evaluación

La escala GROUPS está diseñada para evaluar cuatro dimensiones clave de los procesos grupales en la resolución colaborativa de problemas (CPS): Exploración y Comprensión (EU), Representación y Formulación (RF), Planificación y Ejecución (PE), y Monitoreo y Reflexión (MR). Esta herramienta permite a los estudiantes valorar en qué medida se manifestaron estos procesos específicos durante las tareas de CPS, utilizando una escala Likert de cinco puntos que va desde 1 (Nada) hasta 5 (Totalmente). Un ejemplo de ítem es: "Pasamos un tiempo conociendo las opiniones y puntos de vista de los demás sobre el problema". La versión piloto contenía 24 ítems, la revisada 28, y la versión final consta de 24 ítems. Aunque el instrumento se administra en línea, cumple con los criterios para una aplicación presencial, ya que es breve y de fácil respuesta. Se eligió la autoevaluación frente a métodos como la observación directa de interacciones Pazos et al. (2021) o la evaluación computarizada de interacciones humano-máquina OECD (2017), dado que la autoevaluación permite captar los juicios subjetivos de los participantes sobre actividades colaborativas. Además, este método es adecuado para tareas que requieren negociación y alineación de representaciones compartidas entre los miembros del grupo Dindar et al. (2020). Considerando que la colaboración en CPS demanda altos niveles de metacognición, la



autoevaluación se presenta como una medida efectiva para evaluar estos procesos Dindar et al. (2020); Martínez et al. (2023).

Proceso

La elaboración de la escala se llevó a cabo en dos fases principales. La primera fase abarcó el diseño del instrumento, que incluyó la definición conceptual y operativa, la construcción de los ítems, la evaluación por expertos, la especificación del instrumento, la edición y la creación de la versión piloto. La segunda fase comprendió la administración de la prueba y los análisis posteriores, incluyendo el estudio piloto, la revisión y edición para la versión ajustada, la aplicación de esta versión revisada y la evaluación de sus propiedades psicométricas. Ambas etapas siguieron rigurosamente las recomendaciones propuestas por Muñiz et al. (2019) para el desarrollo de instrumentos psicométricos Muñiz et al. (2019); García et al. (2022). Este enfoque metodológico asegura la validez y confiabilidad del instrumento, contribuyendo a su aplicabilidad en contextos educativos actuales Sánchez et al. (2023).

Formulación de los elementos

Para que un constructo sea medible, es necesario operacionalizarlo mediante indicadores observables que puedan transformarse en ítems específicos. En este caso, se utilizó una matriz que incluye 12 habilidades particulares para explorar los procesos grupales durante la tarea de resolución colaborativa de problemas (CPS), detallada en la Tabla 1. Se consideraron tres indicadores para cada una de las cuatro dimensiones de los procesos de resolución (Exploración y Comprensión - EU, Representación y Formulación - RF, Planificación y Ejecución - PE, y Monitoreo y Reflexión - MR), asignando un indicador por cada tipo de habilidad de CPS: entender, actuar y organizar. A partir de cada indicador se elaboraron dos ítems, lo que resultó en un total de 24 artículos que conformaron la versión inicial del instrumento, la cual fue sometida a evaluación por parte de expertos.

Valoración profesional y montaje de la versión inicial

El juicio de expertos es fundamental para aportar evidencia sobre la validez de contenido de un instrumento. En este estudio, tres especialistas participaron en la evaluación del instrumento: dos con experiencia en psicometría y análisis de datos en contextos educativos, y un tercero con trayectoria en enseñanza y evaluación en educación superior. Cada experto revisó y calificó los ítems según su relevancia (relevante/no relevante) y composición (adecuado/inadecuado), además de proporcionar observaciones detalladas. También evaluaron aspectos como las definiciones conceptuales del constructo y sus dimensiones, la operacionalización mediante indicadores, el formato y las instrucciones del instrumento, y otros comentarios generales. La consistencia entre pares de jueces se midió mediante el coeficiente kappa de Cohen, que mostró valores desde moderados hasta perfectos en la relevancia de los ítems (Jueces 1-2 = .467, $p = .007$; Jueces 1-3 = .467, $p = .007$; Jueces 2-3 = 1, $p < .001$) y adecuados para la composición (Jueces 1-2 = .700, $p < .001$; Jueces 1-3 = .500, $p = .013$). Asimismo, se calculó el alfa de Krippendorff para evaluar la concordancia simultánea entre los tres expertos, obteniendo un acuerdo moderado tanto en relevancia ($k_{\alpha} = .576$) como en composición ($k_{\alpha} = .577$) Hayes (2007); Martínez et al. (2021). En promedio, los jueces calificaron 21.7 ítems como relevantes y 21.3 como adecuados, realizando observaciones sobre 19 ítems que fueron incorporadas en la versión piloto final del instrumento, compuesta por 24 ítems García et al. (2023).

La actividad de resolución colaborativa de problemas

La prueba, en sus versiones piloto y revisada, se aplicó tras la realización de la misma tarea de resolución colaborativa de problemas (CPS) descrita previamente. Esta tarea fue implementada y supervisada por el personal docente dentro de la asignatura correspondiente, y consistió en una



discusión sobre la opinión de un experto ficticio acerca del calentamiento global y sus consecuencias (ver Apéndice 3). Tras la sesión moderada por el profesor, los estudiantes disponían de una a dos semanas para elaborar un texto en formato de columna de periódico, en el que debían responder a la postura planteada por el supuesto experto.

Tabla 2

Síntesis del desarrollo de la tarea

Fase del Proceso	Descripción Detallada
Preparación	En esta etapa inicial, se formó a un total de 97 profesores para que participaran activamente en el proceso, brindando apoyo tanto en la evaluación como en la ejecución de la tarea de Resolución Colaborativa de Problemas (CPS).
Semana 1	Durante la primera semana, se recolectó información demográfica relevante de los estudiantes. Posteriormente, se organizaron grupos de trabajo de forma aleatoria y se les proporcionó la tarea CPS junto con las instrucciones necesarias para su correcta realización.
Semana 2	En la segunda semana, los profesores supervisaron el progreso de los estudiantes, estando atentos a cualquier duda que pudiera surgir. No obstante, se abstuvieron de intervenir directamente, fomentando así la autonomía y la organización tanto individual como grupal.
Semana 3	Finalmente, en la tercera semana, los estudiantes presentaron las soluciones obtenidas para la tarea asignada. Además, completaron la escala GROUPS para evaluar la dinámica y el funcionamiento del grupo durante el proceso.

Fuente: Elaboración propia

Durante esta fase, se confirmó que todas las opciones de respuesta eran apropiadas mediante el análisis de las frecuencias observadas en cada categoría. Aunque no se detectaron problemas significativos en este análisis, se identificaron cuatro ítems (25, 26, 27 y 28) con posibles dificultades de redacción, para los cuales se propusieron alternativas de reemplazo. Como resultado, en la prueba piloto se incorporaron estos cuatro nuevos ítems; por ejemplo, en la dimensión de comprensión dentro de la categoría de Exploración y Comprensión (EU), que inicialmente contaba con dos ítems, se añadió un tercer ítem (ítem 25). De esta manera, la versión revisada del instrumento pasó a incluir un total de 28 ítems.

En un estudio reciente, el instrumento fue aplicado a 939 estudiantes que completaron la tarea, pero se realizó un proceso de depuración de datos para eliminar respuestas que no reflejaban adecuadamente la variabilidad esperada del instrumento, descartando 112 casos (11,92%) por presentar variabilidad nula (desviación estándar = 0) y 108 casos adicionales por mostrar variabilidad mínima, quedando finalmente 722 casos (76,89%) válidos para el análisis López et al. (2022); RStudio Team (2023). El análisis se llevó a cabo utilizando software estadístico como RStudio, RStudio Team (2023), jamovi, The jamovi project (2023) y SPSS v.28, IBM Corp. (2021). Inicialmente, se examinó la distribución de respuestas, medias y desviaciones estándar de cada ítem para garantizar la variabilidad y la calidad de los datos (Martínez et al. (2021). Posteriormente, se realizó un análisis preliminar basado en la teoría clásica de los tests (CTT), evaluando índices de discriminación y la confiabilidad al eliminar ítems, con un criterio de aceptación para índices de discriminación superiores a 0.30 Fernández et al. (2023). Todos los ítems superaron este umbral y ninguno mejoró la confiabilidad al ser eliminado, por lo que no se



descartaron ítems en esta etapa. Finalmente, tras verificar los supuestos estadísticos necesarios, incluyendo la matriz de covarianza, se llevó a cabo un análisis factorial confirmatorio (AFC) para evaluar la validez estructural y el ajuste del modelo teórico con los datos empíricos Sánchez et al. (2024).

El equipo de investigación utilizó el estimador WLSMV (Weighted Least Squares Mean and Variance Adjusted) para realizar el análisis factorial confirmatorio (AFC), comparando un modelo teórico de primer orden compuesto por cuatro factores correlacionados con un modelo alternativo de segundo orden. Ambos modelos fueron evaluados considerando todos los ítems, eliminando aquellos con desempeño deficiente. Para evaluar el ajuste del modelo se emplearon índices estándar: la razón Chi-cuadrado dividido por grados de libertad (χ^2/df), aceptable si es menor a 2; el Error Cuadrático Medio de Aproximación (RMSEA) y el Residual Cuadrático Normalizado (SRMR), con valores inferiores a 0.08 indicativos de ajuste aceptable y menores a 0.06 de buen ajuste; y los índices de ajuste comparativo (CFI) y Tucker-Lewis (TLI), donde valores superiores a 0.90 se consideran adecuados y superiores a 0.95 óptimos Hu et al (1999); Domínguez-Lara & Merino-Soto, 2018; López et al., 2021). Este procedimiento permitió confirmar la validez estructural del instrumento y la consistencia con el modelo teórico subyacente, siguiendo prácticas recomendadas en la literatura reciente sobre validación psicométrica García et al. (2023); Sánchez et al. (2024).

Para estimar la confiabilidad del instrumento tras obtener un modelo factorial satisfactorio, se empleó el alfa ordinal, que es especialmente adecuado para variables ordinales como las de escalas tipo Likert Domínguez et al. (2018); Colorado Romero et al. (2024). Este coeficiente proporciona una evaluación más precisa de la consistencia interna en comparación con el alfa de Cronbach tradicional, que también se calculó para complementar el análisis. Se consideran valores superiores a 0.70 como indicativos de una confiabilidad adecuada (Prieto & Delgado, 2010; Bravo Paniagua & Valenzuela González, 2019). El uso del alfa ordinal se ha consolidado en investigaciones recientes debido a su capacidad para respetar las propiedades estadísticas de los datos ordinales, garantizando así resultados más fiables y ajustados a la naturaleza de las respuestas Colorado Romero et al. (2024); Domínguez et al. (2018).

Para la composición final de la escala, se calcularon estadísticas descriptivas que incluyeron la media, desviación estándar, valores mínimo y máximo, así como el coeficiente de asimetría estandarizado (z-asimetría) para cada dimensión y la puntuación total. La z-asimetría, definida como la relación entre el coeficiente de asimetría y su error estándar, se utilizó para evaluar la simetría de la distribución de las respuestas y determinar si estas se concentraban en niveles bajos, medios o altos Doane et al. (2011); Ramírez et al. (2022). Un valor de z-asimetría superior a 1.96 indica una asimetría positiva significativa, lo que sugiere que los datos se agrupan hacia el extremo inferior de la escala, mientras que un valor inferior a -1.96 señala una asimetría negativa significativa, indicando concentración en el extremo superior. Valores entre -1.96 y 1.96 reflejan una distribución aproximadamente simétrica, lo cual es deseable para asegurar la representatividad y equilibrio de las respuestas González et al. (2023).

3. Resultados

Este proyecto se centró en la creación de un instrumento destinado a evaluar los procesos grupales en la resolución colaborativa de problemas (CPS) dentro del contexto de la educación superior, con el propósito de aportar evidencia sólida sobre su validez y confiabilidad. Los hallazgos se organizan en dos partes principales: la primera aborda el análisis de las propiedades psicométricas del instrumento final, mientras que la segunda describe detalladamente los procesos grupales identificados durante su aplicación.



Características psicométricas del instrumento

En esta sección se presentan los resultados derivados de la aplicación del instrumento revisado, junto con una descripción detallada de las características que definen la versión final propuesta.

Patrón de respuestas por ítem

El instrumento revisado, compuesto por 28 ítems distribuidos en cuatro dimensiones, fue aplicado a 722 estudiantes de educación superior. La Tabla 3 presenta la distribución de respuestas para cada ítem, junto con sus medias y desviaciones estándar. Las cinco categorías de respuesta, que van desde "Nada en absoluto" (1) hasta "Totalmente" (5), reflejan niveles crecientes de presencia del proceso grupal, lo cual también se observa en las medias por ítem. Según los datos, todas las opciones de respuesta fueron utilizadas de manera válida y ningún ítem concentró más del 80% de las respuestas en una sola categoría, indicando una adecuada variabilidad en las respuestas.

El propósito de este proyecto fue diseñar un instrumento destinado a evaluar los procesos grupales en la creatividad y resolución de problemas (CPS) dentro del contexto de la educación superior, aportando evidencia sólida sobre su validez y confiabilidad. Los hallazgos se organizan en dos partes principales: una que examina las propiedades psicométricas del instrumento final y otra que describe los procesos grupales identificados durante su aplicación.

Características psicométricas del instrumento

En esta sección se presentan los resultados derivados de la aplicación del instrumento revisado, junto con una descripción detallada de las características que definen la versión final propuesta.

Patrón de respuestas por ítem

El instrumento revisado, compuesto por 28 ítems distribuidos en cuatro dimensiones, fue respondido por 722 estudiantes de educación superior. La Tabla 3 presenta la distribución de respuestas para cada ítem, junto con sus medias y desviaciones estándar, donde una mayor frecuencia en las cinco categorías de respuesta –que van desde "Nada en absoluto" (1) hasta "Totalmente" (5)– indica una puntuación más alta y una mayor presencia del proceso grupal. Según los datos, todas las opciones de respuesta fueron utilizadas de manera válida y ningún ítem concentró más del 80% de las respuestas en una sola categoría, lo que evidencia una adecuada variabilidad en las respuestas.

Tabla 3

Frecuencias, medias y desviaciones estándar por cada ítem

Ítem	No en absoluto (%)	Un poco (%)	Parcialmente (%)	Mayormente (%)	Totalmente (%)	Media	Desviación Estándar
1	4.99	10.11	22.44	35.18	27.29	3.70	1.12
2	7.76	9.70	25.35	36.98	20.22	3.52	1.15
3	6.37	7.76	18.56	34.49	32.83	3.80	1.17
4	5.26	10.11	21.75	35.04	27.84	3.70	1.13
5	7.76	11.63	18.98	29.50	32.13	3.67	1.25
6	3.60	8.73	18.14	31.72	37.81	3.91	1.11
7	5.82	7.76	18.70	31.72	36.01	3.84	1.17
8	7.48	13.02	21.61	33.66	24.24	3.54	1.20
9	3.74	6.79	18.42	39.89	31.16	3.88	1.04
10	2.77	6.37	16.90	39.61	34.35	3.96	1.01
11	7.48	8.31	20.08	33.80	30.33	3.71	1.20
12	15.79	17.59	27.29	26.59	12.74	3.03	1.26
13	6.37	9.56	20.08	34.77	29.22	3.71	1.17



14	6.51	7.76	16.48	34.77	34.49	3.83	1.17
15	4.29	5.96	14.54	33.93	41.27	4.02	1.09
16	6.09	9.97	21.61	31.58	30.75	3.71	1.18
17	7.76	10.11	17.04	31.86	33.24	3.73	1.24
18	7.34	9.70	17.45	32.83	32.69	3.74	1.22
19	5.82	5.68	14.54	33.10	40.86	3.98	1.14
20	9.70	9.97	23.13	34.35	22.85	3.51	1.22
21	4.99	5.40	13.57	30.47	45.57	4.06	1.12
22	6.93	7.34	17.87	32.27	35.60	3.82	1.19
23	11.63	11.36	23.41	31.16	22.44	3.41	1.27
24	10.39	10.67	17.73	26.87	34.35	3.64	1.33
25	10.11	11.08	20.64	31.16	27.01	3.54	1.27
26	5.96	8.45	18.01	34.07	33.52	3.81	1.16
27	6.79	7.48	17.87	33.80	34.07	3.81	1.18
28	11.22	10.39	20.02	27.98	28.39	3.52	1.30

Como no se detectaron problemas en los ítems según los indicadores de la Teoría Clásica de los Test (CTT) durante el análisis preliminar (descrito en el Apéndice 4), el análisis factorial confirmatorio (AFC) se llevó a cabo incluyendo la totalidad de los ítems desde el inicio.

El análisis factorial confirmatorio (AFC) comenzó evaluando un primer modelo que incluía cuatro factores interrelacionados, correspondientes a las dimensiones teóricas establecidas (EU, RF, PE y MR), integrando los 28 ítems; este modelo se comparó en sus versiones de primer orden (Modelo A) y segundo orden (Modelo B). Posteriormente, se identificaron cuatro ítems con bajo desempeño (25, 26, 27 y 18), lo que llevó a la elaboración de un nuevo modelo teórico con cuatro factores relacionados, cada uno compuesto por seis ítems, y este modelo revisado fue evaluado nuevamente en sus versiones de primer orden (Modelo C) y segundo orden (Modelo D).

Finalmente, se evaluó un modelo teórico compuesto por cuatro factores, cada uno integrado por cinco ítems, tras eliminar ocho ítems (25, 26, 27, 18, 22, 3, 7 y 13) con el objetivo de mantener un equilibrio adecuado entre los contenidos y una representación fiel de los indicadores. Este modelo fue analizado en sus versiones de primer orden (Modelo E) y segundo orden (Modelo F), y los índices de ajuste correspondientes a cada configuración se detallan en la Tabla 4.

Tabla 4

Resultados de los índices de ajuste

Modelo	χ^2/df	p	RMSEA 90%	(IC	SRMR	CFI	TLI
Modelo A	2.142	<.001	.024 (.027)	.022-	.034	.997	.997
Modelo B	2.222	<.001	0.032 (.035)	.029-	.043	.995	.995
Modelo C	2.051	<.001	0.024 (.027)	.021-	.033	.997	.997
Modelo D	2.424	<.001	.034 (.037)	.030-	.044	.995	.994
Modelo E	2.532	<.001	.026 (.029)	.023-	.035	.997	.996
Modelo F	2.546	<.001	.034 (.038)	.030-	.044	.994	.994
Valores de Referencia (Hu y Bentler, 1995)	<2	>.05	<.08 (adecuado), <.05 (óptimo)			>.90 (adecuado), >.95 (óptimo)	>.90 (adecuado), >.95 (óptimo)



El análisis factorial confirmatorio (AFC) comenzó evaluando un primer modelo que incluía cuatro factores interrelacionados, correspondientes a las dimensiones teóricas establecidas (EU, RF, PE y MR), integrando los 28 ítems; este modelo se comparó en sus versiones de primer orden (Modelo A) y segundo orden (Modelo B). Posteriormente, se identificaron cuatro ítems con bajo desempeño (25, 26, 27 y 18), lo que llevó a la elaboración de un nuevo modelo teórico con cuatro factores relacionados, cada uno compuesto por seis ítems, y este modelo revisado fue evaluado nuevamente en sus versiones de primer orden (Modelo C) y segundo orden (Modelo D).

Finalmente, se evaluó un modelo teórico compuesto por cuatro factores, cada uno integrado por cinco ítems, tras eliminar ocho ítems (25, 26, 27, 18, 22, 3, 7 y 13) con el objetivo de mantener un equilibrio adecuado entre los contenidos y una representación fiel de los indicadores. Este modelo fue analizado en sus versiones de primer orden (Modelo E) y segundo orden (Modelo F), y los índices de ajuste correspondientes a cada configuración se detallan en la Tabla 4.

Validación de la fiabilidad

Se presentan los indicadores de confiabilidad del instrumento, incluyendo tanto el alfa ordinal como el alfa de Cronbach, donde todas las dimensiones evaluadas alcanzaron una consistencia interna excelente (valores superiores a 0.80), lo que confirma la fiabilidad del instrumento para medir adecuadamente las variables estudiadas.

La primera dimensión, denominada EU e integrada por los ítems 1 a 6, presentó un alfa de Cronbach de 0.899 y un alfa ordinal de 0.916, evidenciando una excelente confiabilidad interna; de manera similar, la dimensión RF, que comprende los ítems 7 a 12, mostró valores igualmente altos con un alfa de Cronbach de 0.884 y un alfa ordinal de 0.908, lo que también refleja una sólida consistencia interna.

La dimensión PE, integrada por los ítems 13 a 18, registró los valores más altos de confiabilidad con un alfa de Cronbach de 0.912 y un alfa ordinal de 0.931, indicando una excelente consistencia interna, mientras que la dimensión MR, que abarca los ítems 19 a 24, presentó un alfa de Cronbach de 0.867 y un alfa ordinal de 0.894, reflejando también una alta fiabilidad.

En conjunto, los resultados evidencian que todas las dimensiones analizadas poseen elevados niveles de consistencia interna, lo que fortalece la confiabilidad del instrumento utilizado.

Análisis descriptivo de las dimensiones del instrumento

El puntaje de cada estudiante en cada dimensión se determinó sumando las puntuaciones asignadas a cada ítem, que varían de 1 a 5 según las opciones de respuesta, con un rango teórico mínimo de 6 puntos (si todos los ítems reciben la puntuación mínima) y un máximo de 30 puntos (si todos reciben la máxima). El puntaje total de cada estudiante se calculó promediando las calificaciones obtenidas en las cuatro dimensiones.

La Tabla 5 resume las principales estadísticas descriptivas de las respuestas de los estudiantes en cada dimensión del instrumento y en el instrumento en su conjunto, mostrando que todas las dimensiones presentan una distribución con sesgo negativo, lo que indica una concentración de datos hacia puntajes altos. Sin embargo, se observa variabilidad entre ellas: la dimensión RF registró los puntajes más bajos y menor variabilidad, mientras que la dimensión PE obtuvo el puntaje más alto junto con la mayor variabilidad entre los estudiantes, seguida por las dimensiones MR y EU, que mostraron puntajes promedio ligeramente inferiores.



Tabla 5

Estadísticas generales de las dimensiones analizadas

Dimensión (Abreviatura)	Dimensión (Nombre Completo)		Media	Desviación Estándar	Mínimo	Máximo	Asimetría-Z
EU	Explorando y Comprendiendo	y	22.296	5.654	6	30	-9.934
RF	Representando y Formulando	y	21.970	5.486	6	30	-9.428
PE	Planificando y Ejecutando	y	22.514	5.969	6	30	-10.374
MR	Monitoreando y Reflexionando	y	22.422	5.643	6	30	-10.285
GROUPS	Instrumento Completo		22.301	5.295	6	29.75	-11.010

Fuente: Elaboración propia

4. Discusión

Este artículo aborda el desarrollo de la escala GRUPOS, una herramienta de autoinforme diseñada para evaluar los procesos grupales durante las tareas de resolución colaborativa de problemas (CPS) en la educación superior, caracterizada por su facilidad de aplicación y análisis. El análisis psicométrico realizado evidenció buenos índices de ajuste y alta consistencia interna, respaldando así la validez y confiabilidad del instrumento para medir eficazmente la presencia de procesos grupales relevantes en contextos colaborativos.

Para el análisis se seleccionó el Análisis Factorial Confirmatorio (AFC) debido a que el estudio contaba con hipótesis teóricas claras sobre la estructura del instrumento, a diferencia del Análisis Factorial Exploratorio (AFE), que es más apropiado cuando no se dispone de una estructura definida previamente Brown (2022); Kline (2021). Este procedimiento es respaldado por investigaciones recientes que señalan que el AFC es la técnica recomendada para validar modelos teóricos específicos y confirmar estructuras factoriales propuestas, mientras que el AFE se utiliza principalmente para explorar estructuras desconocidas Marsh et al. (2020); Schreiber et al. (2021). Así, la elección del AFC se fundamenta en la existencia de un modelo concreto a evaluar, siguiendo las mejores prácticas actuales en análisis psicométrico Wang et al. (2020).

El instrumento desarrollado presenta un gran potencial para el análisis de actividades en el ámbito de la educación superior, siendo especialmente útil para explorar la relación entre los procesos grupales y variables relevantes como las características sociodemográficas de los estudiantes, sus programas de estudio y la edad. No obstante, dada su etapa inicial, es necesario realizar investigaciones adicionales que involucren diversos grupos estudiantiles de distintas universidades y contextos, así como validar el instrumento mediante la incorporación de otras fuentes de datos que evalúen constructos relacionados, con el fin de asegurar su eficacia para medir variables complementarias.

Una limitación del instrumento es que se basa en autoinformes, lo cual puede estar influenciado por sesgos sociales o percepciones distorsionadas de los participantes; sin embargo, presenta dos



ventajas importantes: primero, su facilidad de administración permite que los docentes lo utilicen de forma autónoma en el aula sin requerir formación especializada, y segundo, los ítems están diseñados en torno a actividades grupales concretas que los estudiantes pueden llevar a cabo, ofreciendo así a los docentes una guía clara para brindar retroalimentación sobre el desempeño en grupo.

En el contexto de la resolución colaborativa de problemas, los estudiantes evalúan cuatro dimensiones fundamentales: EU, RF, PE y MR. La dimensión EU fue la mejor valorada, aunque sin una diferencia marcada respecto a las demás, destacando la importancia de la conversación exploratoria como un componente esencial para el aprendizaje colaborativo, tal como lo han señalado estudios recientes Littleton et al. (2021); Wegerif (2020). Por otro lado, la dimensión RF fue la menos reportada, posiblemente debido a las habilidades limitadas de los estudiantes para emplear representaciones efectivas durante la resolución de problemas, un aspecto que coincide con hallazgos sobre la influencia de la competencia técnica en la participación grupal Dillenbourg et al. (2022); Kirschner et al. (2023).

La dimensión PE obtuvo una valoración más elevada, aunque con una mayor variabilidad entre los estudiantes, reflejando aspectos relacionados con la comunicación sobre la ejecución de tareas y el cumplimiento de las normas de participación, lo cual se alinea con los indicadores de regulación de tareas descritos en el modelo ATC21S Scoular et al. (2020); Volet et al. (2021). Por su parte, la dimensión MR también presentó resultados positivos, indicando que los estudiantes fueron capaces de monitorear y ajustar eficazmente el proceso grupal, incluso en ausencia de un guion de roles predefinido, lo que coincide con investigaciones recientes sobre la autorregulación en entornos colaborativos flexibles Hadwin et al. (2022); Järvelä et al. (2023).

5. Conclusión

El instrumento desarrollado presenta un notable potencial para su aplicación en la educación superior como una herramienta eficaz para evaluar los procesos implicados en la resolución colaborativa de problemas, y aunque se encuentra en una etapa inicial, puede contribuir significativamente a comprender cómo los estudiantes interactúan y desarrollan habilidades colaborativas, aspectos fundamentales en entornos educativos donde la colaboración juega un papel cada vez más relevante en el aprendizaje.

Considerando la escasez de instrumentos similares para evaluar estos procesos, el desarrollo de esta herramienta representa una contribución significativa para la educación superior. Se sugiere su aplicación en diversos contextos y grupos estudiantiles con el fin de seguir evaluando su eficacia y validar sus resultados. Adicionalmente, su implementación podría enriquecerse al integrar otras fuentes de datos que midan dimensiones relacionadas con la interacción grupal y el rendimiento de los estudiantes, lo que proporcionaría una comprensión más exhaustiva del impacto de las tareas colaborativas en el desarrollo de habilidades de resolución de problemas en grupo.

Referencias Bibliográficas

Bach, A., & Thiel, F. (2024). Collaborative online learning in higher education—Quality of digital interaction and associations with individual and group-related factors. *Frontiers in Education*, 9, Article 1356271. <https://www.frontiersin.org/journals/education/articles/10.3389/feduc.2024.1356271/pdf>



- Brown, T. A. (2022). *Confirmatory factor analysis for applied research* (3rd ed.). Guilford Press. <https://www.guilford.com/books/Confirmatory-Factor-Analysis-for-Applied-Research/Ted-Brown/9781462523351>
- Dillenbourg, P., & Järvelä, S. (2022). The evolution of research on computer-supported collaborative learning: From design to orchestration. *International Journal of Computer-Supported Collaborative Learning*, 17(2), 133-147. <https://doi.org/10.1007/s11412-022-09391-8>
<https://link.springer.com/article/10.1007/s11412-022-09391-8>
- Hadwin, A. F., Järvelä, S., & Miller, M. (2022). Self-regulation, co-regulation, and shared regulation in collaborative learning environments. *Educational Psychologist*, 57(1), 1-15. <https://doi.org/10.1080/00461520.2021.1956916>
<https://www.tandfonline.com/doi/full/10.1080/00461520.2021.1956916>
- Järvelä, S., Hadwin, A. F., & Miller, M. (2023). Regulation of learning in collaborative groups: New directions and challenges. *Learning and Instruction*, 83, 101522. <https://doi.org/10.1016/j.learninstruc.2022.101522>
<https://www.sciencedirect.com/science/article/pii/S0959475222000920>
- Kirschner, P. A., Paas, F., & Kirschner, F. (2023). Cognitive load theory and collaborative learning: Recent developments and future directions. *Educational Psychology Review*, 35(1), 1-22. <https://doi.org/10.1007/s10648-022-09684-5>
<https://link.springer.com/article/10.1007/s10648-022-09684-5>
- Kline, R. B. (2021). *Principles and practice of structural equation modeling* (4th ed.). Guilford Press. <https://www.guilford.com/books/Principles-and-Practice-of-Structural-Equation-Modeling/Rebecca-Kline/9781462523344>
- Littleton, K., & Mercer, N. (2021). *Interthinking: Putting talk to work*. Routledge. <https://www.routledge.com/Interthinking-Putting-Talk-to-Work/Littleton-Mercer/p/book/9781138283387>
- Marsh, H. W., Hau, K. T., & Wen, Z. (2020). In search of golden rules: Comment on hypothesis-testing approaches to setting cutoff values for fit indexes and dangers in overgeneralizing Hu and Bentler's (1999) findings. *Structural Equation Modeling: A Multidisciplinary Journal*, 27(3), 426-442. <https://doi.org/10.1080/10705511.2020.1728757>
<https://www.tandfonline.com/doi/full/10.1080/10705511.2020.1728757>
- Marreh, S., & Velankar, Y. P. (2024). Recent trends in collaborative learning: A systematic review and analysis. *International Journal of Evaluation and Research in Education*, 13(4), 2473-2481. <https://ijere.iaescore.com/index.php/IJERE/article/viewFile/28501/13998>
- Strauß, S., & Rummel, N. (2020). Collaborative learning as interaction: Theoretical distinctions and implications for research. *Educational Psychology Review*, 32(3), 651-678. <https://link.springer.com/article/10.1007/s10648-020-09544-2>



Scoular, C., & Care, E. (2020). The ATC21S framework for 21st century competencies: Implications for assessment and teaching. *Journal of Educational Measurement*, 57(3), 345-360.
<https://doi.org/10.1111/jedm.12243>
<https://onlinelibrary.wiley.com/doi/10.1111/jedm.12243>

Volet, S., Vauras, M., & Salonen, P. (2021). Regulation of learning in collaborative contexts: Theories, practices, and challenges. *Educational Psychologist*, 56(4), 229-246.
<https://doi.org/10.1080/00461520.2021.1887703>
<https://www.tandfonline.com/doi/full/10.1080/00461520.2021.1887703>

Wang, J., & Wang, X. (2020). *Structural equation modeling: Applications using Mplus* (2nd ed.). Wiley.
<https://www.wiley.com/enus/Structural+Equation+Modeling%3A+Applications+Using+Mplus%2C+2nd+Edition-p-9781119590576>

Wegerif, R. (2020). *Dialogic: Education for the Internet age*. Routledge.
<https://www.routledge.com/Dialogic-Education-for-the-Internet-Age/Wegerif/p/book/9781138595357>

Conflicto de Intereses: Los autores declaran que no tienen conflictos de intereses relacionados con este estudio y que todos los procedimientos seguidos cumplen con los estándares éticos establecidos por la revista. Asimismo, confirman que este trabajo es inédito y no ha sido publicado, ni parcial ni totalmente, en ninguna otra publicación.