



UNIVERSIDAD DE CONCEPCIÓN
FACULTAD DE CIENCIAS SOCIALES

**EFECTO DEL MENTEFACTO CONCEPTUAL SOBRE LA METACOGNICIÓN
Y LA COMPRENSIÓN CONCEPTUAL: UN ESTUDIO EXPERIMENTAL EN
EL CONTEXTO DE EVALUACIÓN FORMATIVA**

TESIS PRESENTADA PARA OPTAR AL GRADO ACADÉMICO DE DOCTOR EN
PSICOLOGÍA

POR: María Paula Baquero Vargas
PROFESORA GUÍA: Claudia Paz Pérez Salas

Concepción, Chile, 2024

Se autoriza la reproducción total o parcial, con fines académicos, por cualquier medio o procedimiento, incluyendo la cita bibliográfica del documento.

AGRADECIMIENTOS

Agradezco a mi profesora guía, Claudia Pérez, por su orientación permanente, sus enseñanzas y sus aportes precisos en todo momento. Gracias por concederme la posibilidad de disfrutar estos años de estudio a través de la confianza y respeto hacia mis ritmos de trabajo y necesidades personales.

Agradezco también a mi esposo Pablo Emilio De La Cruz por acompañarme en esta travesía al extranjero y a la presencia de mis hijos Tomás y Abril en todo momento.

Así mismo, agradezco el apoyo financiero otorgado por la Agencia Nacional de Investigación y Desarrollo (ANID) para la realización de esta tesis doctoral.

Tabla de contenido

6

Abstract8

9

18

*Aprendizaje Centrado en el Estudiante*18

*Evaluación Formativa*19

*Características De La Evaluación Formativa*20

21

*Taxonomía revisada de Bloom como base para comprender el impacto de las evaluaciones objetivas en línea sobre el aprendizaje*23

*El problema de las Evaluaciones Formativas Objetivas y su relación con las Dimensiones Cognitivas de Bloom*25

*Metacognición*28

*Metacognición y evaluación formativa*28

*Componentes de la Metacognición*29

*Monitoreo Metacognitivo*30

*Comprensión Conceptual*34

*Medición de la comprensión*36

*Conocimiento Estructural como Alternativa para la Evaluación Formativa*38

*Técnicas De Representación Externa Del Pensamiento Estructural*39

*Técnicas de conocimiento estructural y su impacto en la educación*44

*Los mentefactos conceptuales como actividades de evaluación formativa en línea*46

*Pregunta de Investigación*49

*Hipótesis de Investigación*50

*Objetivos*51

*Objetivos específicos*51

Métodos52

*Diseño*52

*Participantes*53

<i>Variables</i>	54
<i>Variable Independiente</i>	54
<i>Variables Dependientes</i>	54
<i>Instrumentos y Materiales</i>	56
<i>Procedimiento</i>	62
<i>Consideraciones éticas</i>	67
Resultados	70
Discusión	78
Conclusiones	84
Referencias	86
Índice de Tablas	
Tabla 1. <i>Dimensiones, significados y subcomponentes de la taxonomía de Bloom</i>	24
Tabla 2. <i>Dimensiones de conocimiento</i>	33
Índice de Figuras	
<i>Figura 1. Técnica de Clasificación de tarjetas abierta.</i>	40
<i>Figura 2. Mapa Conceptual sobre el tema de Mapas Conceptuales.</i>	41
<i>Figura 3. Representación de las redes asociativas “Pathfinder”.</i>	43
<i>Figura 4. Estructura del Mentefacto conceptual.</i>	47
<i>Figura 5. Actividad del grupo experimental: Mentefacto Conceptual.</i>	61
<i>Figura 6. Actividad de emparejamiento para el grupo control.</i>	62
<i>Figura 7. Duración del proceso experimental.</i>	67
<i>Figura 8. Desempeño en comprensión.</i>	66
<i>Figura 9. Número de estudiantes que incluyeron en sus explicaciones conceptos y características claves para definir el concepto de granulocitos.</i>	72
<i>Figura 10. Precisión Absoluta</i>	
<i>Figura 11. Medida de sesgo para cada grupo</i>	77
<i>Figura 12. Errores presentados durante la ejecución de la actividad experimental</i>	76

Resumen

Debido a la preferencia masiva por el uso de pruebas objetivas para valorar formativamente el aprendizaje de los estudiantes en entornos virtuales, y el hecho de que la mayoría de evaluaciones objetivas en línea limitan a los estudiantes a usar habilidades de pensamiento basadas principalmente en la memoria, aspecto que interfiere en la comprensión y el aprendizaje autorregulado de los estudiantes, surge la necesidad de diversificar e innovar en el diseño de actividades de evaluación formativa en línea que promuevan procesos de pensamiento más elaborados. El objetivo de la presente investigación es explorar el efecto de los mentefactos conceptuales automatizados (organizadores gráficos objetivos, fijos, estructurales y jerárquicos) sobre la comprensión conceptual y el monitoreo metacognitivo de estudiantes universitarios.

El presente estudio tiene un diseño experimental de grupos paralelos, donde cincuenta estudiantes fueron asignados aleatoriamente a una actividad de evaluación experimental (construcción de un mentefacto conceptual), o a una actividad de evaluación de control (emparejamiento), basada en un texto académico sobre "granulocitos". Se aplicó una rúbrica para calificar la comprensión conceptual de los estudiantes después de realizar la actividad de evaluación formativa asignada, y se utilizaron juicios retrospectivos para valorar el nivel de confianza de los estudiantes con respecto a su desempeño.

Los resultados demuestran que los estudiantes que resolvieron el mentefacto conceptual obtuvieron puntuaciones significativamente mejores en comprensión

conceptual y juicios de monitoreo metacognitivo más precisos, en comparación con los estudiantes que resolvieron la actividad de emparejamiento.

Abstract

Due to the massive preference for objective testing in online environments to formatively assess students' learning, and the fact that most objective testing limits students to use thinking skills based mainly on memory, thus interfering with comprehension and self-regulated learning, the need to diversify and innovate in online formative assessment activities arises. The aim of the research is to explore the effect of automated conceptual mindfacts (objective, fixed, structural, and hierarchical graphical organizers) on conceptual understanding and metacognitive monitoring.

The present study conducted a parallel-group design experiment with fifty undergraduate students randomly assigned to an experimental assessment activity (conceptual mindfacts) and a control assessment activity (matching) based on an academic text about "Granulocytes." A rubric was applied to rate students' conceptual comprehension after the assessment activity, and a judgment scale was used to rate students' confidence regarding their performance.

Results demonstrate that students who solved the conceptual mindfact obtained significantly better punctuations on conceptual understanding and more accurate metacognitive monitoring judgements than those who solved the matching activity.

Introducción

La evaluación formativa es una práctica educativa que potencia el aprendizaje de los estudiantes de forma consciente y constante, maximizando las oportunidades de aprendizaje significativo. De forma más precisa, la evaluación formativa busca rastrear, monitorear, potenciar y asegurar el aprendizaje de los estudiantes, promoviendo habilidades metacognitivas que puedan llevarlos a mejorar el conocimiento que tienen sobre sus propias capacidades y estrategias de aprendizaje. Al mismo tiempo, la evaluación formativa informa al docente sobre las prácticas de instrucción que debe cambiar, adaptar o mejorar durante el proceso de enseñanza, para asegurar el aprendizaje de todos los estudiantes (Al-Hattami, 2020; Cosi et al., 2020; Leenknecht et al., 2020; McCallum y Milner, 2020).

Debido a la creciente popularidad de las prácticas de evaluación formativa en la actualidad, y la necesidad de rastrear permanentemente el desempeño de los estudiantes en entornos de aprendizaje virtuales, se ha dedicado una especial atención al diseño e innovación de actividades de evaluación formativa basadas en computador; también denominadas en la literatura como evaluaciones electrónicas, automatizadas, o sistematizadas, para ser implementadas en ambientes en línea (Bälter et al., 2013; Buldu y Buldu, 2010; Chen y Chen, 2009; Costa et al., 2010; Faber et al., 2017; Faber y Visscher, 2018; Hacker et al., 2009).

Como consecuencia, los sistemas de gestión de aprendizaje, cursos online masivos y abiertos (MOOCS), herramientas de creación de contenidos en línea, aulas virtuales, entre otras, disponen progresivamente de más esquemas de evaluación formativa

automatizadas para rastrear la comprensión de los estudiantes durante el proceso de aprendizaje (Jaggars y Xu, 2016). Desafortunadamente, los proveedores de la mayoría de estos recursos, no incorporan un diseño pedagógico en sus productos, por lo que las herramientas disponibles no necesariamente aseguran la calidad ni eficacia del aprendizaje (Henrie et al., 2015; Hettiarachchi et al., 2015). En el caso de la evaluación formativa particularmente, se ha demostrado que los bajos tamaños del efecto reportados en la literatura en relación con la eficacia de las prácticas de evaluación formativa, se deben en parte a que las actividades de evaluación diseñadas poseen criterios de calidad deficientes (Appiah y van Tonder, 2018; Dunn y Mulvenon, 2009; Farrell y Rushby, 2016; Lane et al., 2019; Lee et al., 2020).

La calidad de los instrumentos que usa un docente para evaluar el conocimiento de los estudiantes es fundamental porque a través de ella se promueve la reflexión y el cuestionamiento en los estudiantes. En este sentido, el tipo de pregunta usada cobra relevancia como herramienta pedagógica porque determina el potencial para promover niveles desafiantes de reflexión y cuestionamiento en el aula. De acuerdo con Minte Münzenmayer et al., (2021), se ha demostrado que la calidad de las preguntas en los procedimientos de evaluación influye en la calidad del aprendizaje de los estudiantes.

En revisiones recientes, se ha demostrado que las actividades de evaluación formativa más usadas por los docentes en los ambientes de evaluación en línea, tanto de forma sincrónica como asincrónica, son actividades con esquemas de tipo objetivo (opción múltiple, verdadero-falso, emparejamiento o asociación, etc.) (Febriani y Irsyad Abdullah, 2018; Rolim y Isaias, 2018), que, a pesar de sus atributos prácticos, cuando se

diseñan sin tener en cuenta ciertos criterios de calidad, limitan a los estudiantes a usar procesos de pensamiento basados principalmente en el recuerdo (Lai y Bower, 2019; Pellegrino, 2005; Reimann y Schilcher, 2013). Este hecho ha sido confirmado en diversos estudios que han demostrado que las actividades objetivas pobremente diseñadas, usualmente promueven en los estudiantes solo el primer o segundo nivel de conocimiento en la taxonomía de Bloom (como por ejemplo el recuerdo), y no permite a los estudiantes demostrar otras capacidades o procesos de pensamiento tales como el análisis o la creación, lo que compromete el desempeño académico y el aprendizaje (Bell et al., 2001; Chiheb, Faizi, y El Afia, 2011; Darling-Hammond et al., 2017; Galizzi, 2010; Hettiarachchi et al., 2015; Lai y Bower, 2019; Minte Münzenmayer et al., 2021; Pellegrino, 2005; Reimann y Schilcher).

El hecho de que las evaluaciones formativas usadas en contextos virtuales usualmente promuevan el uso del recuerdo como principal proceso cognitivo, acarrea varios problemas. En primer lugar, al no ascender en la pirámide de dimensiones cognitivas de la taxonomía de Bloom, los objetivos de aprendizaje permanecen en un nivel muy básico, lo que implica que la comprensión de los estudiantes puede verse afectada, llevando a su vez a una transferencia de aprendizaje limitada o nula. Este problema se conoce como conocimiento inerte, que sucede cuando los estudiantes adquieren gran cantidad de conocimiento factual, pero no lo comprenden a profundidad y por tanto no lo integran y organizan sistemáticamente dentro de una disciplina (Al-Hattami, 2020; Chou et al., 2017; Cosi et al., 2020; McLaughlin y Yan, 2017; Orhan Göksün y Gürsoy, 2019; Zainuddin et al., 2020).

Por otro lado, teniendo en cuenta la importancia que tiene la metacognición en la efectividad de las prácticas de evaluación formativa para monitorear el aprendizaje, y reconociendo que la naturaleza del aprendizaje en línea hace un llamado a que los aprendices auto-dirijan y tomen responsabilidad de su aprendizaje, es primordial promover el desarrollo de habilidades metacognitivas de regulación, como el monitoreo metacognitivo, durante la ejecución de las actividades de evaluación formativa en línea (Zhao y Ye, 2020). Desafortunadamente, así como evaluaciones con estructuras cerradas se han relacionado con niveles básicos de comprensión, también se ha demostrado que estas se relacionan con bajos niveles de precisión en el monitoreo metacognitivo en los estudiantes (Akcaoğlu et al., 2023; Klein et al., 2018; Murray, 2014; Nietfeld et al., 2005; Pinchok y Brandt, 2009; Purnamawati y Saliruddin, 2017; Zohar y Barzilai, 2015). Esto sucede porque usualmente las estructuras objetivas no proveen a los estudiantes con oportunidades para reflexionar sobre su proceso de aprendizaje, no permiten a los estudiantes evidenciar su progreso y otorgan un sentido de exceso de confianza sobre el aprendizaje que en ocasiones puede incluso inhibir el monitoreo y aprendizaje regulado (Dunlosky y Lipko, 2007; Liu et al., 2023; Nederhand et al., 2021; Zarestky et al., 2022).

En consecuencia, los estudiantes presentan dificultad para discriminar de forma precisa si un desempeño ha sido bueno o malo y por tanto no emprenden las acciones necesarias para mejorar (Gutierrez, 2021; Zhao y Ye, 2020). Particularmente, esta dificultad o sesgo en el propio monitoreo del desempeño está relacionada con la “ilusión de saber” que se refiere a la creencia de haber alcanzado una buena comprensión cuando

en realidad no ha sido lograda. La ilusión de saber suele considerarse un exceso de confianza subjetiva en la comprensión de la información que objetivamente es incorrecta, o un exceso de confianza en la capacidad de recordar información que no se puede recordar (Avhustiuk et al., 2018; Hall et al., 2007).

En este sentido, para asegurar la efectividad de las prácticas de evaluación formativa en línea, las actividades diseñadas deben cumplir con varios criterios de calidad (Ibarra-Sáiz y Rodríguez-Gómez, 2020; Ragupathi, 2016; Tomanek et al., 2008): que promuevan la comprensión, que estimulen habilidades de pensamiento diversificadas (Lane et al., 2019; Ropohl y Rönnebeck, 2019) y que desarrollen procesos metacognitivos que conlleven a una mejor precisión en el monitoreo de su aprendizaje, otorgando un sentido de confianza adecuado, que permita a los estudiantes emprender las acciones necesarias para mejorar su desempeño (Avhustiuk et al., 2018; Dunlosky y Lipko, 2007; Hall et al., 2007).

Por estas razones, varios estudios y revisiones de la literatura señalan la necesidad de diseñar actividades de evaluación en línea que promuevan habilidades de pensamiento de mayor complejidad, que aseguren comprensión, que desarrollen habilidades de monitoreo metacognitivo precisas, y que tengan la propiedad de ser automatizadas para acceder a ellas por medio de sistemas de gestión de aprendizaje, o puedan ser usadas en contextos de aprendizaje asincrónico. Esto con el fin de que sean útiles y coherentes con la realidad actual y el redireccionamiento de la educación hacia modelos híbridos y virtuales de enseñanza (Kruger et al., 2015; Lopes, 2014; Nipper y

Holmberg, 2003; Oldfield, Broadfoot, Sutherland, y Timmis, 2012; Turnbull, Chugh, y Luck, 2019; Villarroel y Bruna, 2019).

La necesidad de innovar en la evaluación formativa en línea ha sido abordada por la comunidad científica. Varios estudios han desarrollado diversos tipos de actividades innovadoras que puedan ser automatizadas, e incluso han abordado la posibilidad de automatizar actividades que puntúan alto en la taxonomía de Bloom, como ensayos escritos. Sin embargo, aunque se han propuesto ideas interesantes, puede evidenciarse que las investigaciones suelen proponer actividades para disciplinas muy específicas (matemáticas, física, computación, idiomas, enfermería, medicina, literatura), o para desarrollar habilidades particulares (procedimientos dentales o médicos), lo cual dificulta el acceso a estas actividades de manera global o su difusión en sistemas de gestión de aprendizaje (Barana y Marchisio, 2016; Leary et al., 2018; H. S. Lee et al., 2019; Mao et al., 2018; Moreno y Pineda, 2020; Zhu et al., 2020).

Una alternativa para desarrollar actividades de evaluación formativa que promuevan habilidades de pensamiento más elaboradas en ambientes virtuales, es considerar el conocimiento estructural. El conocimiento estructural se refiere a la capacidad para comprender las relaciones entre conceptos de cierto dominio, donde tanto los conceptos como las relaciones se almacenan en la memoria semántica a largo plazo, y gracias a su organización en red, son más fáciles de recordar (Anohina-Naumecca, 2020; Beissner, Jonassen, y Grabowski, 1994). Para alcanzar su dominio, se debe enfatizar en la relación significativa entre conceptos de alta proximidad, más que en la memorización de hechos y conceptos (Ausubel, 1968). De forma práctica,

convierten conocimiento tácito en explícito, y permite a los estudiantes revisar, evaluar, y guiar su propio pensamiento en un acto de creación (Novak y Gowin, 1984).

El conocimiento estructural puede ser una alternativa para generar actividades de evaluación formativa más apropiadas, porque al relacionar conceptos de forma significativa, promueve procesos cognitivos más elaborados (análisis, evaluación y creación) que a su vez influyen en la comprensión y metacognición de los estudiantes (Beissner et al., 1994; Chiheb et al., 2011; Hayes y Heit, 2018; Joshi et al., 2022; Nesbit y Adesope, 2006; Novak, 1990; Redford et al., 2012; Solomon et al., 1999; Trumpower y Vanapalli, 2016). De hecho, construir organizadores gráficos por medio de categorizaciones taxonómicas, jerarquías, relaciones, exclusiones, etc., constituye en sí mismo un proceso metacognitivo, donde el estudiante debe planear, monitorear y evaluar el proceso de construcción constantemente (Chevron, 2014; Huang y Tsapali, 2022).

Teniendo en mente que el conocimiento estructural se refiere a la representación interna del conocimiento en la mente humana, se deben usar representaciones externas para su evaluación: representaciones gráficas que consisten en mapas espaciales que muestran conceptos que están relacionados y que de forma explícita o implícita demuestran la naturaleza de la relación entre los conceptos (Beissner et al., 1994). Algunos de estos métodos incluyen técnicas como la clasificación de tarjetas, el uso de mapas conceptuales, mentefactos conceptuales y redes asociativas tipo “pathfinder” (Clariana, 2010).

Entre las representaciones externas del conocimiento estructural, los mapas conceptuales son los más conocidos y han sido ampliamente estudiados en el ámbito educativo demostrando impactos positivos en el desempeño académico de los estudiantes, la comprensión conceptual y la metacognición (Gouli, Gogoulou, y Grigoriadou, 2003; Hattie, 2015; Huang y Tsapali, 2022; Joshi et al., 2022; Liu et al., 2023; Novak, 1990; Redford et al., 2012; Schroeder, Nesbit, Anguiano, y Adesope, 2018). Sin embargo, a pesar de su utilidad para representar la estructura de conocimiento de una persona, los mapas conceptuales presentan una gran dificultad para ser usados como actividad de evaluación formativa en línea; debido a su estructura abierta (cantidad indefinida de niveles), la corrección y retroalimentación automática de estos esquemas requiere de programas especializados que no son de acceso común para el ámbito educativo (Chang, Sung, y Chen, 2001; Harper et al., 2003). Adicionalmente, las técnicas de clasificación de tarjetas y redes asociativas se basan en las relaciones de proximidad que existen entre ciertos conceptos, sin profundizar en las características subyacentes de los mismos, razón por la cual se han usado principalmente como herramientas diagnósticas más que pedagógicas (Clariana, 2010; Curtis y Davis, 2003).

Por el contrario, los mentefactos conceptuales, debido a su estructura cerrada, tienen el potencial de ser programados como actividad de evaluación formativa en línea. Los mentefactos conceptuales son un tipo de organizador gráfico, que representa la relación que existe entre un concepto determinado y otro de alta proximidad conceptual, junto con las características principales que los diferencian. Al poseer una estructura jerárquica, permanente y cerrada que promueve distintas operaciones mentales

(supraordinar, infraordinar, excluir, deducir e inducir) (De Zubiria Samper, 2000), los mentefactos conceptuales tienen el potencial para ser programados como actividad de evaluación formativa automatizada en línea. Además, los estudios realizados a la fecha, demuestran que los mentefactos conceptuales propician el aprendizaje significativo, tienen efectos positivos en el desempeño académico y mejoran la apropiación conceptual de los contenidos enseñados en estudiantes universitarios (Del Carmen Gutierrez, Martinez, Reyes, y Mendez, 2021; Ichpas, 2019; Mendoza y Huarachi, 2008; Monsalve Flórez, 2018; Quevedo, 2018; Riveros Melo, 2017).

Por tanto, debido a la poca diversidad y efectividad de actividades objetivas automatizadas disponibles para hacer evaluación formativa en línea, que promuevan el uso de habilidades de pensamiento diferentes al recuerdo, que favorezcan el desarrollo de habilidades de monitoreo metacognitivo precisas y promuevan altos niveles de comprensión en los estudiantes; esta investigación explora si la construcción de mentefactos conceptuales en línea promueve mejor comprensión conceptual y mayores niveles de precisión en el monitoreo metacognitivo de estudiantes universitarios, en comparación con actividades objetivas de emparejamiento aplicadas en línea. En ese sentido, la pregunta de investigación de esta tesis es: ¿Cuál es el efecto de los mentefactos conceptuales, como actividad de evaluación formativa en línea, sobre las habilidades metacognitivas de monitoreo y la comprensión conceptual de estudiantes universitarios, en comparación con actividades objetivas de emparejamiento en línea?

Marco Teórico

Aprendizaje Centrado en el Estudiante

La presente investigación está enmarcada en una concepción del aprendizaje centrada en el estudiante, que tiene sus raíces en una perspectiva constructivista del aprendizaje. El aprendizaje centrado en el estudiante se caracteriza por un trabajo colaborativo donde docentes y estudiantes están activamente comprometidos con el contenido y el proceso de aprender, atendiendo de forma particular a las necesidades, habilidades e intereses del aprendiz. Como consecuencia, los estudiantes se apropian de su proceso de aprendizaje y desarrollan procesos profundos de reflexión, pensamiento crítico y habilidades de auto-regulación (Singhal, 2017).

Desde una perspectiva pedagógica, el aprendizaje centrado en el estudiante implica cambiar la concepción del docente como gestor único del aprendizaje y resignificar su rol teniendo en cuenta los siguientes aspectos (Hoidn, 2016):

- a) El docente se convierte en un facilitador que comparte responsabilidades con el estudiante a lo largo de todo el proceso de aprendizaje.
- b) El contexto social y cultural del estudiante, sus experiencias, estilos de aprendizaje y necesidades, son tomados en cuenta para dirigir el proceso de enseñanza-aprendizaje.
- c) Los estudiantes construyen sus propios significados por medio del trabajo proactivo, la reflexión y desarrollo del pensamiento crítico.
- d) Hay un énfasis en la interdisciplinariedad con el objetivo de consolidar mayores niveles de conocimiento y desarrollo de habilidades.

- e) El estudiante juega un rol importante en la construcción de objetivos de aprendizaje.
- f) El proceso de aprendizaje no prioriza únicamente la transferencia del conocimiento, por el contrario, pretende generar procesos de comprensión profundos, significativos y reflexivos.
- g) La evaluación que acompaña el aprendizaje centrado en el estudiante es generalmente formativa y con retroalimentación continua.

Evaluación Formativa

Enmarcada dentro de la perspectiva del aprendizaje centrada en el estudiante, la evaluación formativa se caracteriza por evaluar el proceso de aprendizaje del estudiante durante el transcurso de la instrucción, sin el propósito de adjudicarle calificaciones. Es un trabajo colaborativo donde docentes y estudiantes se comprometen con el objetivo de entender el proceso de aprendizaje de los aprendices, sus organizaciones mentales, fortalezas, debilidades, estrategias y puntos a mejorar. Para los docentes, la evaluación formativa sirve como fuente de información para mejorar, ajustar, adaptar o cambiar los contenidos, objetivos o didácticas de instrucción usados, con el fin de promover comprensión y mejores desempeños en los estudiantes. Los métodos usados en la evaluación formativa buscan monitorear el proceso de cada aprendiz, para tener un conocimiento diferenciado de su nivel de comprensión y desempeño, logrando identificar qué tan cerca se encuentran de los objetivos de aprendizaje establecidos. Adicionalmente, por medio de estrategias de retroalimentación, se busca mejorar las estrategias metacognitivas usadas por los estudiantes, con el fin de que logren

monitorear su propio proceso de aprendizaje (Al-Hattami, 2020; Cosi et al., 2020; Leenknecht et al., 2020; McCallum y Milner, 2020).

Como práctica educativa, la evaluación formativa es un proceso deliberado que requiere de la aplicación de una serie de pasos que se deben ejecutar de forma continua durante el proceso de enseñanza aprendizaje: Primero se deben establecer objetivos de aprendizaje claros y con criterios de éxito específicos, que permitan determinar cómo progresan los estudiantes hacia la consecución de las metas. Segundo, los estudiantes deben desarrollar diversas actividades de aprendizaje que sirvan para reforzar la comprensión de los contenidos y permitan valorar el avance hacia los objetivos de aprendizaje. Tercero, los docentes deben verificar el desempeño de los estudiantes para retroalimentar su progreso en relación con los objetivos de aprendizaje. Por último, los docentes deben establecer adaptaciones que permitan mejorar el desempeño de los estudiantes antes de que termine el periodo de enseñanza-aprendizaje (Antoniou y James, 2014; Leenknecht et al., 2020).

Características De La Evaluación Formativa

La evaluación formativa se caracteriza por (Andrade y Cizek, 2010; Conrad y Openo, 2018; Gedye, 2010; Gikandi et al., 2011; Ogange et al., 2018):

- a) Promover el desarrollo de habilidades de pensamiento.
- b) Comunicar objetivos de aprendizaje claros y específicos.
- c) Promover logros de aprendizaje que representan resultados con aplicaciones más allá del contexto de aprendizaje.

- d) Identificar la comprensión actual del estudiante y los pasos necesarios para alcanzar las metas deseadas.
- e) Motivar a los estudiantes para que auto-monitoreen su progreso hacia las metas de aprendizaje.
- f) Diseñar evaluaciones enmarcadas en actividades de aprendizaje auténticas.
- g) Entregar retroalimentación no evaluativa.
- h) Promover metacognición y reflexión de los estudiantes sobre su trabajo.

Evaluación Formativa En Línea

Debido al desarrollo tecnológico y su impacto en la educación, la convergencia entre evaluación formativa y evaluación asistida por computador, da vida al concepto de evaluación formativa en línea. Si bien la evaluación formativa en línea comparte las mismas características que la evaluación formativa presencial descritas anteriormente, la evaluación formativa en línea ofrece varias ventajas: a) el fácil acceso y disponibilidad, b) la utilización de características interactivas como imágenes y videos c) la provisión de retroalimentación inmediata e individualizada, y d) el acceso a información proporcionada por las analíticas de datos (Nagandla et al., 2018). Así mismo, varios estudios han investigado los efectos de la evaluación formativa en línea y se ha encontrado que los mecanismos propuestos se relacionan positivamente con el aumento del compromiso de los estudiantes, el aumento en el tiempo dedicado al estudio de material educativo y la identificación de deficiencias de aprendizaje a través de esquemas de retroalimentación formativa (Sudakova et al., 2022).

Tipos de evaluación formativa en línea. Gracias a la cantidad y variedad de opciones que ofrece la tecnología, la educación en línea ha adaptado y desarrollado diversas actividades, herramientas o tareas, que permiten a la evaluación formativa en línea valorar el proceso de aprendizaje de los estudiantes de forma rápida y continua. Si bien las posibilidades de evaluación basadas en la web son múltiples, se pueden categorizar en dos grupos principales: objetivas o subjetivas.

Las evaluaciones objetivas en línea se caracterizan por tener un esquema de corrección basado en solo una respuesta correcta y entregan una corrección y retroalimentación automática inmediata. Para su ejecución, los estudiantes responden las preguntas apuntando o presionando el mouse, moviendo objetos en la pantalla o escribiendo números o palabras de extensión limitada. A este grupo pertenecen tareas como los cuestionarios de opción múltiple, verdadero-falso y tareas de emparejamiento o asociación principalmente (Al-Hattami, 2020; Chou et al., 2017; Cosi et al., 2020; McLaughlin y Yan, 2017; Orhan Göksün y Gürsoy, 2019; Zainuddin et al., 2020). Las evaluaciones subjetivas en línea, se caracterizan por permitir diferentes respuestas como aceptables, y su corrección automática puede ser difícil. En esta categoría se pueden incluir los foros, los blogs, estudios de caso, los ensayos, portafolios electrónicos, reflexiones, juegos de rol, simulaciones, entre otras (Appiah y van Tonder, 2018; Farrell y Rushby, 2016; Villarroel y Bruna, 2019).

Evaluaciones objetivas en línea. A continuación, se presenta una breve descripción de cada una de las modalidades de evaluación objetiva en línea más comunes (Boitshwarelo et al., 2017; Chiheb et al., 2011):

- a) **Opción múltiple:** Es una actividad donde se presenta un enunciado que plantea un problema o pregunta redactado en sentido afirmativo o interrogativo, acompañado de un conjunto de opciones de respuestas. Los estudiantes deben seleccionar una o más opciones de respuesta de acuerdo con las instrucciones dadas.
- b) **Verdadero-Falso:** Esta forma de actividad requiere que los estudiantes lean una afirmación y decidan si esta es correcta o incorrecta.
- c) **Asociación/ Emparejamiento (Términos pareados):** Esta actividad consiste en presentar una serie de elementos en dos columnas paralelas, donde cada palabra, símbolo o frase de una columna puede asociarse con los elementos de la otra. Una de las columnas se denomina premisa y la otra respuesta.
- d) **Completar:** Las preguntas de completar proporcionan a los estudiantes una frase o pregunta parcial y luego les piden que escriban la palabra (o palabras) en el espacio en blanco que mejor complete el enunciado o la pregunta.
- e) **Respuesta corta:** Los estudiantes tipean una palabra o frase en respuesta a una pregunta concreta. La respuesta debe corresponder de forma precisa a alguna de las repuestas aceptables previamente programadas.

Taxonomía revisada de Bloom como base para comprender el impacto de las evaluaciones objetivas en línea sobre el aprendizaje

Para comprender el impacto que tiene el uso de determinado tipo de evaluación sobre el aprendizaje de los estudiantes en contextos de aprendizaje en línea, se tomará como referente la taxonomía revisada de Bloom. La taxonomía de Bloom se ha

convertido en un referente académico que permite categorizar de forma jerárquica los objetivos de aprendizaje en relación con la complejidad del proceso cognitivo que se requiere para su ejecución. En este sentido, su uso se ha extendido también para clasificar el nivel de compromiso cognitivo que puede alcanzar determinada evaluación o actividad de aprendizaje. Si bien esta herramienta es un referente teórico, las investigaciones suelen hacer alusión a este modelo como criterio práctico para clasificar el potencial de las actividades o de los logros de aprendizaje de los estudiantes (Anderson et al., 2001).

El trabajo original de la taxonomía fue desarrollado en 1956 por Bloom y otros colaboradores. En los años 90, antiguos estudiantes de Bloom, Lorin Anderson y David R. Krathwohl, revisaron la taxonomía de su maestro y la volvieron a publicar en 2001. Uno de los aspectos clave de esta revisión es el cambio de los sustantivos de la propuesta original a verbos para significar las acciones correspondientes a cada categoría. Otro aspecto fue modificar el proceso “síntesis” por un criterio más amplio como “crear”, además, se modificó la jerarquía en que se presentan las distintas categorías, la taxonomía de 1956 tenía como proceso más alto de la jerarquía “evaluar”, mientras que a partir de 2001 está se cambió por “crear”. A continuación, la Tabla 1 representa cada una de las dimensiones cognitivas propuestas en la taxonomía revisada de Bloom y los subcomponentes cognitivos respectivos que la caracterizan (Anderson et al., 2001).

Tabla 1

Dimensiones, significados y subcomponentes de la taxonomía revisada de Bloom 2001

Dimensiones	Significado	Subprocesos
Recordar	Recuperar conocimiento relevante de la memoria a largo plazo	Reconocer y recordar
Comprender	Construir significado sobre la información recibida	Interpretar, ejemplificar, clasificar, resumir, inferir, comparar, explicar
Aplicar	Utilizar un procedimiento en una situación determinada	Ejecutar, implementar
Analizar	Descomponer material en sus partes constitutivas y comprender cómo estas partes están relacionadas unas con otras y con la estructura u objetivo principal	Diferenciar, organizar, atribuir
Evaluar	Elaborar juicios basados en criterios o estándares	Verificar, criticar
Crear	Poner elementos juntos con el objetivo de formar un todo novedoso y coherente o hacer un producto original	Generar, planear, producir

Nota: Adaptada de: A taxonomy for learning teaching and assessing. A revision of Bloom`s Taxonomy of Educational Objectives, por Anderson et al., 2001.

El problema de las Evaluaciones Formativas Objetivas y su relación con las Dimensiones Cognitivas de Bloom

Entre las evaluaciones objetivas y subjetivas, la literatura refiere que las evaluaciones objetivas son las más usadas por docentes en contextos de evaluación virtual debido a los diversos atributos que poseen para facilitar la labor docente (Febriani y Irsyad Abdullah, 2018): permitir a los docentes evaluar gran cantidad de estudiantes en corto tiempo, proveer a los docentes con oportunidades para calibrar la comprensión de los estudiantes de forma rápida y eficiente, proveer corrección y retroalimentación automática y además acceder a ellas por medio de los sistemas de gestión de aprendizaje, aplicaciones o programas en línea, de forma masiva y diversificada (Appiah y van Tonder, 2018; Farrell y Rushby, 2016; Fernández-Serrano, Pérez-García, Perales, y Verdejo-García, 2010; Timmis, Broadfoot, Sutherland, y Oldfield, 2015; Winkley, 2010). Sin embargo, se ha evidenciado que cuando las actividades de evaluación objetivas son diseñadas pobremente, limitan la capacidad para promover habilidades de pensamiento superiores en estudiantes, y debido a su estructura cerrada, difícilmente proporcionan retroalimentación útil por lo que dificultan procesos de monitoreo adecuados (Al-Hattami, 2020; Chou et al., 2017; Cosi et al., 2020; McLaughlin y Yan, 2017; Orhan Göksün y Gürsoy, 2019; Zainuddin et al., 2020).

A continuación, se citarán algunos estudios que demuestran cómo actividades objetivas usadas en el contexto educativo, difícilmente ascienden en la taxonomía de habilidades de pensamiento de Bloom. Por ejemplo, Bell et al. (2001), encontraron en un estudio sobre habilidades superiores de pensamiento y herramientas de evaluación, que la mayoría de las tareas de evaluación construidas por docentes, correspondían a tareas objetivas de opción múltiple que, al no estar correctamente construidas, se enfocaban

principalmente en habilidades de pensamiento basadas en el recuerdo (primera dimensión de la taxonomía de Bloom) (Bell et al., 2001). Así mismo, Bloxham y Boyd (2007), concluyen en su libro sobre “Desarrollo de evaluaciones efectivas en la educación superior”, que el tipo de evaluación usada por el docente, influye en la calidad y cantidad de logro obtenido por los estudiantes. En este sentido, enfatizan que evaluaciones pobremente diseñadas pueden llevar a desarrollar una comprensión deficiente del material debido a las limitaciones del tipo de actividad o herramienta usada. Este hecho fue corroborado por los autores en un análisis de tareas objetivas de evaluación docente, donde se evidenció que todos los items valorados, promovían tan solo la primera dimensión cognitiva de la taxonomía de Bloom.

Por su parte, Abosalem, (2016), en su artículo sobre técnicas de evaluación y habilidades de pensamiento superiores en estudiantes, reporta que son varios los autores que han manifestado cómo las evaluaciones objetivas usualmente operan en el nivel más bajo de la taxonomía de Bloom (recordar). En su revisión, reporta un estudio realizado por Fleming y Chambers (1983, citado por Abosalem, 2016), donde de las 8800 actividades de evaluación analizadas, el 80% se encontraba en el nivel de recuerdo de la taxonomía de Bloom.

Recientemente, en un estudio realizado en Chile, Minte Münzenmayer et al., (2021), reportan que, de 2002 preguntas analizadas en instituciones de educación superior, 1059 correspondían a un nivel de pensamiento basado en el recuerdo.

Uno de los aspectos más importantes de esta taxonomía en el ámbito de la evaluación, es su claridad frente a la correspondencia entre objetivos de aprendizaje,

procesos cognitivos necesarios para alcanzar dichos objetivos y formas de evaluar para acceder a estos procesos. Esta correspondencia es importante debido a la naturaleza jerárquica de la taxonomía, donde procesos en un mayor nivel jerárquico implican mayores demandas al sistema cognitivo, por tanto, tareas de evaluación más complejas. De acuerdo con Anderson et al. (2001), la coherencia entre logros de aprendizaje, procesos cognitivos y tareas seleccionadas es clave para asegurar la comprensión, el pensamiento metacognitivo en los estudiantes y la transferencia del aprendizaje (Anderson et al., 2001; Bell et al., 2001; Darling-Hammond, 2017; Hess et al., 2009; Mohamed y Lebar, 2017).

Metacognición

Metacognición y evaluación formativa

La metacognición se puede definir como la conciencia del propio conocimiento acerca de una tarea, tema o pensamiento y el control ejecutivo involucrado para llevar a cabo tal proceso mental. Las actividades de evaluación formativa deben promover habilidades metacognitivas en los estudiantes porque por medio de la evaluación formativa, los estudiantes ven la evaluación como un proceso de auto mejoramiento, en oposición a un mecanismo netamente valorativo. Así mismo, por medio de la evaluación permanente que se realiza durante la evaluación formativa, los estudiantes desarrollan habilidades para hacer preguntas significativas, cuestionan los comentarios del docente, monitorean su progreso y reflexionan activamente sobre su propio aprendizaje (Murray, 2014; Pinchok y Brandt, 2009; Purnamawati y Saliruddin, 2017; Zohar y Barzilai, 2015).

Particularmente, las habilidades de monitoreo metacognitivo son importantes para los estudiantes en el aprendizaje, porque ayudan a hacer un seguimiento de los procesos cognitivos y de las estrategias de regulación para resolver problemas. De modo más preciso, ayudan a gestionar el proceso de aprendizaje mediante la asignación eficaz de la atención, la memoria y el tiempo cuando se estudia o se realiza un trabajo académico (Ellis y Zimmerman, 2001). La capacidad de controlar el rendimiento con precisión por parte de los estudiantes los alerta de su necesidad de adaptarse al entorno de aprendizaje y de tomar decisiones sobre su educación. Sin un seguimiento preciso, puede resultar imposible un control de la metacognición (Azevedo y Alevén, 2013; Hacker, Dunlosky, y Graesser, 2009; Ozturk, 2017; Rivers et al., 2020; Tobias y Everson, 2009).

Componentes de la Metacognición

Expertos en el aprendizaje clasifican los componentes de la metacognición en dos grupos principales: conocimiento de la cognición y regulación de la cognición. El conocimiento de la cognición involucra la conciencia sobre qué conocimientos, estrategias o habilidades necesita una persona para realizar una tarea de aprendizaje y se compone de conocimiento declarativo (repertorio de estrategias cognitivas disponibles), procedimental (conjunto de procesos para implementar estrategias cognitivas) y condicional (el dónde, el cuándo y por qué aplicar las estrategias de acuerdo a las demandas exigidas) (Azevedo y Alevén, 2013; Hacker et al., 2009; Ozturk, 2017; Rivers et al., 2020).

Por su parte, el segundo componente de la metacognición, la regulación de la cognición, se ha definido como el control sobre el proceso de aprendizaje que realiza

una persona. Por su carácter procedimental, también es referido en la literatura como estrategias o habilidades metacognitivas. La regulación de la cognición involucra procesos como la planificación (preparar lo necesario antes de hacer una tarea), el monitoreo de la comprensión (habilidades de seguimiento efectivo para la consecución de tareas y el seguimiento sobre la efectividad del proceso de aprendizaje) y, finalmente, la evaluación (juicio holístico y general sobre qué tan bien se cumplió la tarea) (Ellis y Zimmerman, 2001; Gutierrez de Blume, 2021).

Monitoreo Metacognitivo

El monitoreo sobre la comprensión o también referido en la literatura como monitoreo metacognitivo, calibración de la comprensión, juicios metacognitivos, metacomprensión, metamemoria o autoregulación para la comprensión, ha sido conceptualizado como el proceso de evaluar el estado actual de conocimiento. Se refiere al grado en que la percepción que tiene una persona sobre su desempeño corresponde con su desempeño actual, medido de manera objetiva (Azevedo y Alevén, 2013; Hacker, Dunlosky, y Graesser, 2009; Ozturk, 2017; Rivers et al., 2020).

Esta capacidad para predecir de forma precisa la propia comprensión es un componente esencial del aprendizaje autorregulado efectivo. En una era de pruebas de alto nivel, rendimiento y competencia, la habilidad para desempeñarse bien en un examen es fundamental. Los estudiantes que estudian para una prueba necesitan tener un monitoreo preciso de su capacidad de retención de conocimiento si desean controlar habilidades más complejas en su estudio. Cuando los estudiantes demuestran fuertes sesgos en sus juicios de calibración, es posible que no tomen las medidas correctivas

necesarias para mejorar o evaluar cuidadosamente sus respuestas durante o después de un examen (Olaya Torres et al., 2023). En este sentido, aquellos estudiantes que tienen la capacidad de distinguir entre lo que han aprendido y lo que les falta por aprender, tienen una ventaja porque pueden omitir el estudio del material más familiar y enfocarse en el que menos dominan, por su parte, los estudiantes con pobre monitoreo metacognitivo usualmente invierten mucho tiempo repasando material familiar, omitiendo el estudio de material nuevo o desconocido (Akcaoglu et al., 2023; Dunlosky y Lipko, 2007; Hacker et al., 2008; Tobias y Everson, 2009).

Existen diversos estudios que reportan la importancia de un adecuado monitoreo en los procesos de aprendizaje. Por ejemplo, Richardson et al., (2012) identificaron el monitoreo metacognitivo como una de las estrategias de aprendizaje más influyentes para los estudiantes en aulas tradicionales y señalan que aquellos estudiantes con capacidad de calibración metacognitiva alta (precisión en su monitoreo), tienen un desempeño académico más efectivo. Por su parte, el modelo conceptual de aprendizaje autorregulado de Butler y Winne, (1995) estipula que el monitoreo metacognitivo sirve como etapa crucial del mecanismo de auto retroalimentación en el sistema cognitivo de los estudiantes. Del mismo modo, Hacker et al., (2008) concluyeron que el monitoreo está positivamente asociado con el rendimiento en los exámenes a lo largo de un semestre. Por su parte, el estudio de Nietfeld et al., (2005) sobre calibración de la metacognición, concluye que los estudiantes universitarios con mayor precisión en la calibración metacognitiva tienen más probabilidades de ser estudiantes de alto rendimiento, que la capacidad de realizar una calibración precisa tiene un impacto

positivo en la autorregulación y que una autorregulación eficaz conduce a un mejor rendimiento en el aprendizaje.

Medición del monitoreo metacognitivo. La medición del monitoreo debe considerar varias variables: el tiempo en que se hace el juicio de valor sobre el desempeño, el tipo de contenido que se valorará y la medida de calibración a reportar. Con relación al tiempo, el monitoreo se puede medir en distintas fases del aprendizaje: durante la adquisición de material nuevo, antes de un examen (juicios prospectivos) o después de un examen (juicios retrospectivos). Respecto al contenido, el monitoreo puede ser medido con diferentes niveles de especificidad. A nivel global, los estudiantes juzgan su desempeño con base en todas las preguntas de un examen completo, a nivel de concepto, los estudiantes juzgan su desempeño en relación con una temática particular o concepto nuevo adquirido, y al nivel de ítem, los estudiantes juzgan su desempeño en relación con una pregunta particular (Schraw, 2009). Por último, respecto a la medida de calibración que se desea reportar, existen diversas medidas cuantitativas para medir el nivel de precisión en el monitoreo de los estudiantes. Las más conocidas son la precisión absoluta, la precisión relativa y el sesgo. Si un investigador interesado en saber si un individuo hace juicios consistentes a lo largo de un examen con distintos ítems, podría usar una medida de precisión relativa. Por su parte, si el investigador está interesado en los cambios en la precisión dados por una intervención, tratamiento, práctica etc., entonces una medida de precisión absoluta es más apropiada (Hacker et al., 2008; Nietfeld et al., 2005).

Precisión absoluta. Mide el nivel de discrepancia entre el desempeño del estudiante y su nivel de confianza sobre ese desempeño, ya sea en un examen o con relación a un concepto particular. En este escenario, el investigador pide al estudiante que valore su confianza (que tan seguro se siente) respecto al desempeño en una tarea, examen, pregunta etc., en una escala continua entre 1% y 100%. Un puntaje de 0 corresponde a una precisión perfecta y desviaciones más lejanas del 0 corresponden a una precisión menor en el monitoreo metacognitivo del individuo (Schraw, 2009; Schraw y Sperling, 1994).

Precisión relativa. Se refiere a la precisión en la predicción del desempeño al comparar un ítem respecto de otro u otros (Schraw, 2009; Schraw y Sperling, 1994).

Sesgo. Valora el grado en que un individuo demuestra exceso o falta de confianza en su desempeño. Dado que el índice de sesgo puede ser positivo o negativo provee información sobre la dirección y la magnitud de la discrepancia entre confianza y desempeño. La dirección del sesgo (positivo o negativo) indica si existe falta de confianza o exceso de confianza: valores negativos indican falta de confianza y valores positivos exceso de confianza. Por su parte, la magnitud de la discrepancia provee información sobre la severidad del error de juicio. Valores positivos y negativos más alejados del cero demuestran mayor error en el juicio de confianza sobre el desempeño (Schraw, 2009). El exceso de confianza es prevalente entre estudiantes de bajo desempeño y cuando las evaluaciones son más difíciles y es conocido en el ámbito educativo como “la ilusión de saber” (Hacker et al., 2000).

Comprensión Conceptual

Junto con la metacognición, verificar la comprensión de los estudiantes es uno de los objetivos principales de la evaluación formativa y constituye uno de los pasos en el proceso de realizar esta práctica educativa. De hecho, muchos autores definen la evaluación formativa como el proceso sistemático de recoger evidencia permanente sobre el nivel actual de comprensión de los estudiantes (Morris et al., 2021).

Antes de explicar qué es la comprensión conceptual, es importante reconocer que de acuerdo con Anderson et al., (2001) existen cuatro dimensiones del conocimiento: factual, conceptual, procedimental y metacognitivo (Tabla 2), y que cada dimensión tiene un conjunto de características particulares sobre la información que contiene. En este mismo sentido, Anderson et al., (2001), explican que cada tipo de conocimiento está asociado con un nivel de procesamiento cognitivo distinto que debe ser evaluado de forma independiente. Por tanto, cuando se habla de comprensión conceptual en este trabajo, se hace referencia a la valoración de la comprensión de conocimientos pertenecientes a la dimensión conceptual únicamente.

Tabla 2

Dimensiones de conocimiento

Dimensión	Características	Subtipos
Factual	Información discreta, aislada y muy específica sobre algún dominio o disciplina. Son los elementos básicos necesarios para conocer la disciplina. Requiere bajo nivel de abstracción.	Conocimiento sobre terminología y sobre detalles específicos como eventos, personas, fechas o fuentes de información.

Conceptual	Formas de conocimiento con organizaciones más complejas donde las diferentes partes de la información están interrelacionadas significativamente en un todo funcional.	Conocimiento de clasificaciones, categorías, principios, generalizaciones, teorías, modelos y estructuras.
Procedimental	Indicaciones o pasos sobre cómo hacer algo, y los criterios para usar habilidades, algoritmos, técnicas y métodos en momentos determinados.	Conocimiento de las habilidades, algoritmos, técnicas y métodos específicos en una disciplina. Conocimiento de criterios para determinar cuándo aplicar procedimientos apropiados.
Metacognitivo	Comprende la cognición en general y conciencia sobre el conocimiento de la propia cognición.	Conocimiento estratégico, sobre tareas cognitivas incluyendo conocimiento contextual y condicional apropiado y auto conocimiento.

La comprensión conceptual es definida como el producto de las relaciones que una persona establece entre los elementos que conforman un contenido o información asociada a un dominio determinado. Así mismo, puede ser entendida como el resultado de las generalizaciones que hace un aprendiz, sobre la naturaleza o las propiedades de un concepto (Bisson, Gilmore, Inglis, y Jones, 2016; Chadwick, 2009; De-Zubiria, 2015;

Rittle-Johnson y Schneider, 2014) Solomon et al. (1999) explican que el proceso de construir una comprensión acerca de algún concepto, implica la capacidad de inferir, interpretar, predecir, organizar, categorizar, explicar y razonar.

Medición de la comprensión

Para medir la comprensión es necesario tener claras las propiedades constitutivas de la “comprensión” como variable. De acuerdo con Flores et al., (2020), la comprensión tiene cuatro propiedades constituyentes. En primer lugar, la comprensión se refiere a un “*logro*” y no a una acción. Si bien para lograr una comprensión adecuada se requieren de distintas habilidades y procesos; la comprensión propiamente dicha es el resultado de estas acciones. Por esta razón, la comprensión se puede medir como un resultado.

En segundo lugar, la comprensión “*tiene grados*”, es decir que no es una variable dicotómica, sino que se puede calibrar el grado en que alguien ha adquirido una comprensión y esta puede variar de persona en persona. Esta propiedad permite que se le puedan asignar valores diferentes al logro en comprensión.

En tercer lugar, la comprensión es “*polimorfa*”, es decir que los criterios para identificar si se ha logrado una comprensión dependen de aquello que se quiere comprender. No es lo mismo una comprensión sobre el concepto de “perro” que una comprensión sobre “el proceso de siembra del café”; como no es lo mismo lo que se comprende por “familia” de una cultura a otra.

Por último, la comprensión es “*disposicional*”, lo que significa que es necesario valorar la adquisición de ciertas habilidades para saber si una persona alcanza la

comprensión deseada. En este sentido, las rúbricas para evaluar cobran sentido porque establecen aquello que la persona debe demostrar, de acuerdo con el contexto establecido, para determinar si ha alcanzado una comprensión deseada. Para Anderson et al. (2001), algunas de estas habilidades son:

- a) Ejemplificar: Ocurre cuando los estudiantes dan ejemplos nuevos y específicos de un concepto general o de un principio, identificando las características diferenciadoras de este.
- b) Interpretar: Ocurre cuando un estudiante puede convertir información de una forma representacional a otra. Esto implica la capacidad para traducir palabras en dibujos, números en palabras, palabras en significados, etc.
- c) Clasificar: Ocurre cuando el estudiante reconoce que algo pertenece a una categoría determinada. Clasificar involucra la capacidad para identificar características relevantes o patrones que encajan en categorías específicas de un concepto o principio.
- d) Resumir: Ocurre cuando un estudiante expresa en sus palabras, abstracciones del material general, que representa lo más importante de la información enseñada.
- e) Inferir: Ocurre cuando el estudiante puede abstraer un concepto o principio, describiendo características y relaciones relevantes.
- f) Comparar: Ocurre cuando un estudiante tiene la capacidad para detectar similitudes y diferencias entre dos o más objetos, ideas, problemas o

situaciones. Implica poder encontrar correspondencias uno a uno entre elementos y patrones.

Conocimiento Estructural como Alternativa para la Evaluación Formativa

El conocimiento estructural es definido como la forma en que se organizan e interrelacionan conceptos con el fin de lograr una mejor comprensión sobre un dominio determinado. Representa la forma en que las personas estructuran su conocimiento y se basa en la noción de que la organización realizada, refleja la profundidad de la comprensión que se tiene acerca de algo (Anohina-Naumeca, 2020; Davis et al., 2003; Trumpower y Vanapalli, 2016).

La idea del conocimiento estructural se basa en las observaciones de Ausubel (1968), quien aseguró que el aprendizaje significativo involucra la asimilación de nuevos conocimientos en relación con las estructuras de conocimiento existentes. Esto ha sido apoyado por investigaciones sobre redes semánticas que han confirmado que el conocimiento se almacena en forma relacional. Los seres humanos no adquieren conocimiento en forma aislada, en cambio, le dan sentido al conocimiento relacionándolo con el que ya tienen. En otras palabras, se ha demostrado que, para adquirir conocimiento de orden superior, debe haber una comprensión no solo de los hechos, procedimientos y conceptos que contiene una disciplina determinada, sino por sobre todo debe haber una comprensión de las relaciones entre estos hechos, procedimientos y conceptos (Anohina-Naumeca, 2020; Davis et al., 2003; Trumpower y Vanapalli, 2016).

Desde la psicología cognitiva, la memoria a largo plazo es considerada como el almacenamiento principal del conocimiento y está compuesta por diversos subsistemas diferenciados por el tipo de conocimiento que almacenan: a) la memoria episódica almacena recuerdos relacionados con la experiencia personal, b) la memoria semántica almacena conocimientos generales en forma de conceptos, c) la memoria perceptual almacena contenidos de datos sensoriales pasados, y d) la memoria procedimental almacena conocimientos sobre cómo hacer diferentes acciones o procesos (Goldstein, 2008).

En este sentido, el conocimiento estructural se basa en el argumento de que la memoria semántica está activamente involucrada en el proceso de aprendizaje porque contiene los conceptos que son parte integral de cualquier dominio de conocimiento. Este tipo de memoria ha recibido el nombre del sistema conceptual humano, donde los conceptos se almacenan en la mente como redes semánticas, seguidos por estructuras más complejas como los esquemas. Los esquemas son marcos conceptuales que contienen conjuntos de conceptos inter-relacionados en una organización significativa (Anohina-Naumeca, 2020).

Técnicas De Representación Externa Del Pensamiento Estructural

Teniendo en mente que el conocimiento estructural se refiere a la representación interna del conocimiento en la mente humana, se deben usar representaciones externas para su evaluación. Existen diversos métodos que permiten plasmar la representación externa del conocimiento estructural. Algunos de estos métodos incluyen técnicas de clasificación de tarjetas y técnicas de representaciones de redes como mapas

conceptuales, mentefactos conceptuales o redes asociativas “pathfinder” (Clariana, 2010; Davis et al., 2003).

Técnicas de clasificación de tarjetas. El método de clasificación de tarjetas se usa para recoger información sobre las asociaciones y clasificaciones que tiene una persona acerca de un conjunto de datos, asociados a una temática particular. La técnica consiste en un conjunto de cartas que tienen impresos conceptos o imágenes en un lado de su cara. Los participantes deben clasificar las tarjetas en montones y cada montón refleja la estructura de conocimiento del participante respecto a alguna temática o categoría de clasificación (Figura 1). La clasificación puede hacerse de forma abierta o cerrada. La clasificación abierta sucede cuando los participantes clasifican tarjetas en montones creados por sí mismos y la cerrada ocurre cuando los participantes clasifican tarjetas dentro de montones previamente etiquetados por un evaluador. Los resultados usualmente se analizan de forma cualitativa, pero existen instrumentos específicos basados en la clasificación de tarjetas que poseen sus propios parámetros de corrección (Harper et al., 2003).

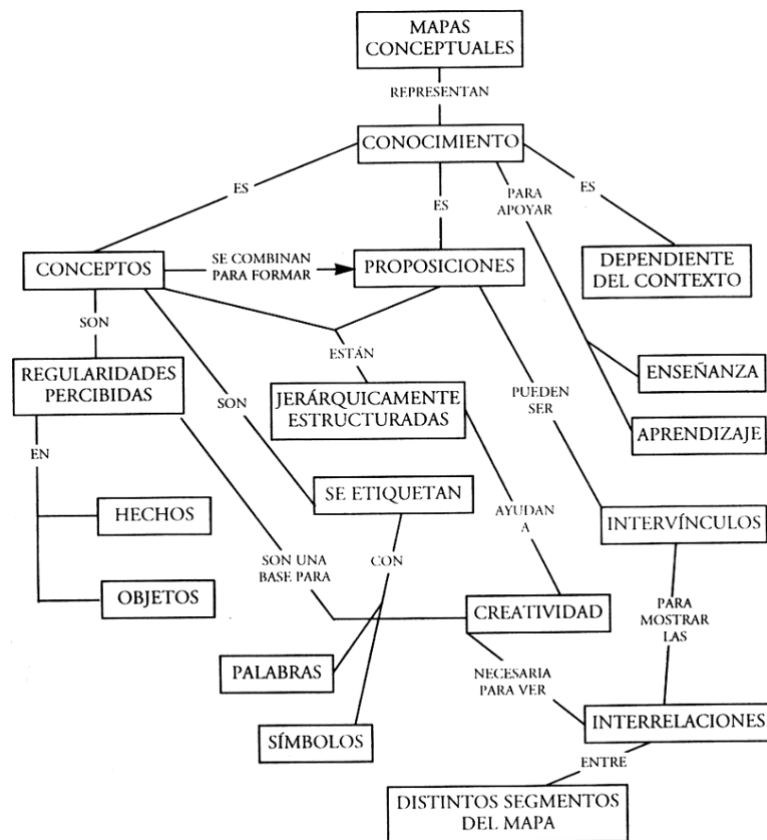
Figura 1. *Técnica de Clasificación de tarjetas abierta.*



Nota: Tomado de <https://www.possibleinc.com/blog/card-sorting-la-tecnica-que-mejora-la-estructura-de-tu-web/>

Mapas conceptuales. Los mapas conceptuales son instrumentos gráficos, de estructura abierta, que sirven para organizar y representar la relación entre conceptos; relación que es indicada por medio de una línea que conecta dos conceptos entre sí (Figura 2). Encima de cada línea se escribe una palabra conocida como conector, que especifica la relación que existe entre los dos conceptos asociados. La organización de un mapa conceptual es usualmente jerárquica, localizando conceptos y proposiciones de lo más general a los más específicos, o en otras palabras los conceptos más incluyentes y generales en la parte superior y los más específicos en la parte inferior (Novak, 1990; Novak y Cañas, 2007; Novak y Gowin, 1984).

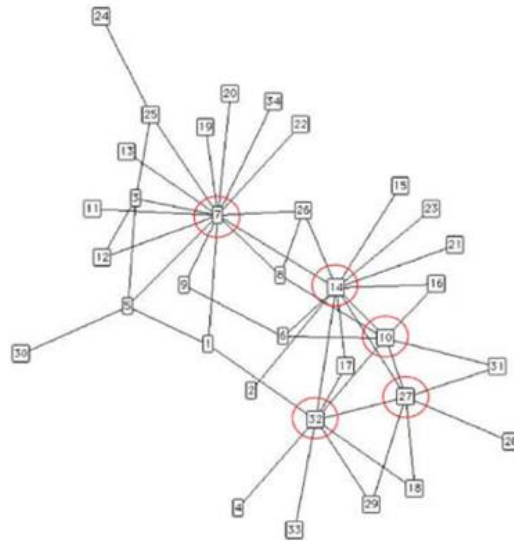
Figura 2. Mapa Conceptual sobre el tema de Mapas Conceptuales.



Nota: Tomado de: Conocimiento y aprendizaje. Los mapas conceptuales como herramientas facilitadoras para escuelas y empresas, por Joseph Novak, 1998, p. 54.

Redes asociativas “Pathfinder”. Las redes asociativas se usan para medir el nivel de proximidad entre diversos conceptos. La técnica inicia presentando al participante dos conceptos que debe valorar numéricamente según su nivel de proximidad. Este proceso continúa hasta que se haya recolectado suficiente información que permita calcular algorítmicamente la representación gráfica de la proximidad entre los conceptos abordados (Figura 3). Esta proximidad refleja la estructura relativa del conocimiento que tiene el participante, y su representación es comparada con gráficos de proximidad realizados por expertos sobre el mismo tema (Matzner, 2008).

Figura 3. Representación de las redes asociativas “Pathfinder”.



Nota: Tomado de: Representación del conocimiento y percepción subjetiva del proceso de aprendizaje profesional: estudio cualitativo en personal de enfermería, por Jiménez-Adán, Casas-García, Luengo-González, 2010.

Mentefactos conceptuales. Los mentefactos conceptuales son instrumentos de conocimiento que se usan para representar las relaciones más significativas que tiene un concepto con otros de alta proximidad y las características claves que los diferencian. Estas relaciones se representan en un diagrama que tiene una estructura predeterminada, fija, jerárquica y cerrada. Para su ejecución, se selecciona un concepto general que se ubica en el centro del esquema, y los participantes deben identificar los conceptos que representan relaciones de supraordinación, exclusión e infraordinación con el concepto seleccionado (De Zubiría, 2000). Los mentefactos constituyen potentes sintetizadores cognitivos que organizan, estructuran, relacionan y preservan los conocimientos,

favoreciendo la memorización de los mismos a largo plazo (De Zubiria Samper, 2000; Herrán y Linares Álvaro, 2013).

Técnicas de conocimiento estructural y su impacto en la educación

Si bien es posible indagar acerca de la estructura de conocimiento de una persona usando cualquiera de estas técnicas, los mapas conceptuales y los mentefactos conceptuales son las únicas técnicas que se han estudiado fuertemente en el ámbito educativo (Harper et al., 2003; Matzner, 2008).

Estudios empíricos y mapas conceptuales. La revisión de Trumpower y Vanapalli (2016), concluye que los mapas conceptuales mejoran el desempeño académico, aumentan la comprensión de los estudiantes con bajo desempeño académico, ayudan a organizar ensayos escritos en estudiantes de secundaria y promueven el uso de estrategias de aprendizaje, ayudando a los estudiantes a relacionar conceptos que luego se transfieren de forma implícita en aprendizajes posteriores. Por su parte, Buldu y Buldu, (2010), revelaron que los mapas conceptuales tienen el potencial para mejorar el aprendizaje de los estudiantes y ayudan a promover la conciencia sobre los procesos de aprendizaje en los alumnos. Así mismo, en un meta-análisis publicado por Nesbit y Adesope (2006), se concluye que construir mapas conceptuales es una actividad pedagógica más efectiva para la retención y transferencia de conocimiento, que otras alternativas como la lectura de textos, las clases magistrales o las discusiones de clase. De la misma manera, el meta-análisis de Hattie (2015), demostró que el tamaño del efecto promedio del uso de mapas conceptuales en ámbitos educativos es medio ($d=0.64$). Finalmente, el meta-análisis de Schroeder, Nesbit, Anguiano, y Adesope (2018),

reporta que aprender construyendo mapas conceptuales tiene un tamaño del efecto más grande ($g = 0.72$).

Estudios empíricos y Mentefactos Conceptuales. Los mentefactos conceptuales fueron desarrollados en Colombia por Miguel De Zubiría Samper en el año 1995. Aunque no son conocidos globalmente, existen varios estudios desarrollados en diversas Universidades de Colombia, Ecuador y Perú. Entre los estudios realizados con mentefactos conceptuales reportados en esta sección, cuatro corresponden a estudios desarrollados con estudiantes de educación básica, cuatro con estudiantes universitarios de pregrado y uno de tipo teórico-reflexivo. Entre estos estudios, dos reportan usar métodos cuasi-experimentales, uno exploratorio descriptivo, uno se reconoce como investigación acción, dos reportan ser experimentales y uno reporta ser pre-experimental. Metodológicamente ninguno de los estudios experimentales está realizado bajo condiciones de control de laboratorio y los análisis estadísticos reportados se basan en comparaciones de medidas de tendencia central. Entre los objetivos de investigación se analiza la efectividad de los mentefactos como estrategia de aprendizaje, como estrategia metacognitiva, como herramienta para el desarrollo del pensamiento histórico, como herramienta para mejorar el desempeño académico y como herramienta para mejorar la comprensión de estructuras biológicas y conocimientos estadísticos precisos (Alvarado Bustamante et al., 2022; Del Carmen Gutiérrez et al., 2021; Ibañez Pinilla, 2006; Mendoza y Huarachi, 2008; Mera Cabezas y Viñan Ajila, 2023; Mucha Hospinal et al., 2022; Quevedo, 2018; Riveros Melo, 2017; Torres Alférez, 2016).

Estos estudios concluyen que los mentefactos permiten a los estudiantes la apropiación conceptual de enseñanzas mejorando su nivel de aprendizaje en forma significativa y describen que los mentefactos favorecen el desarrollo de la creatividad, comprensión lectora, análisis, desarrollo de habilidades cognitivas y comprensión conceptual (Alvarado Bustamante et al., 2022; Del Carmen Gutiérrez et al., 2021; Ibañez Pinilla, 2006; Mendoza y Huarachi, 2008; Mera Cabezas y Viñan Ajila, 2023; Mucha Hospinal et al., 2022; Quevedo, 2018; Riveros Melo, 2017; Torres Alférez, 2016).

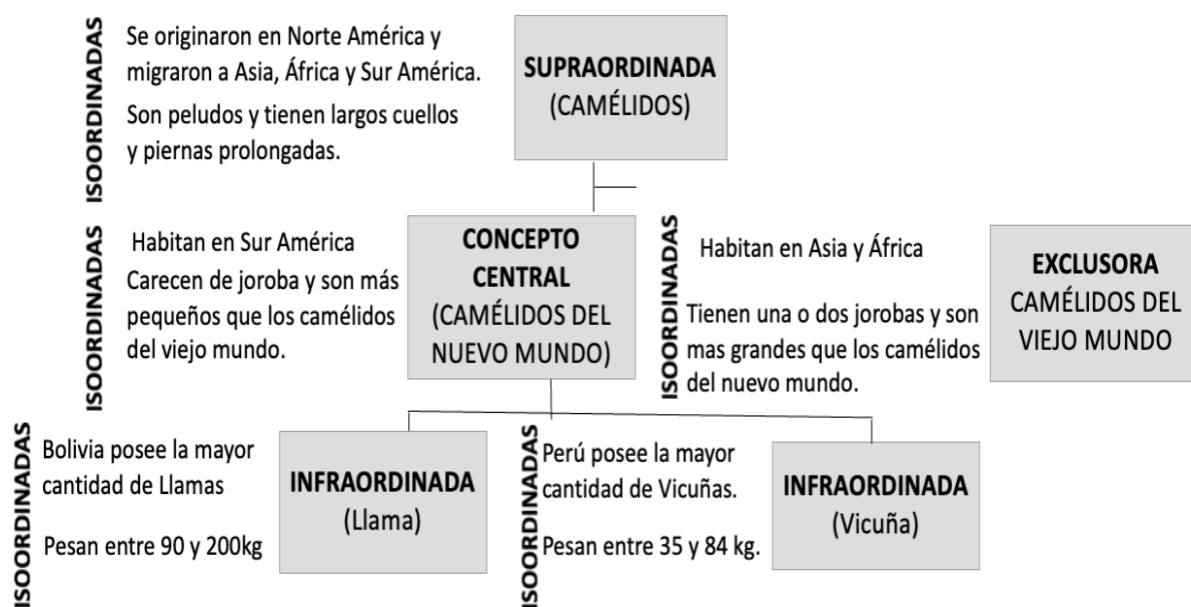
Los mentefactos conceptuales como actividades de evaluación formativa en línea

Si bien todas las técnicas de conocimiento estructural abordadas anteriormente tienen el objetivo de comprender la estructura de conocimiento que tiene un estudiante acerca de un tema, los mentefactos conceptuales son la única de estas técnicas que tiene una estructura cerrada y fija, que posibilita la corrección y retroalimentación inmediata del desempeño de un estudiante de forma automatizada simple. Esto se debe a que generar una corrección automática y proveer retroalimentación inmediata para técnicas de estructura abierta, implica procesos matemáticos muy complejos que dificultan el acceso a estas herramientas de forma libre en el ámbito educativo (Chang et al., 2001; Joseph et al., 2017). Por esta razón, los mentefactos conceptuales son una alternativa dentro del pensamiento estructural que puede ser implementada en ambientes en línea y permite entregar información inmediata sobre el desempeño del participante y guiarlo en el proceso.

La estructura del mentefacto conceptual y sus operaciones intelectuales. Un mentefacto conceptual es un instrumento de conocimiento gráfico que sirve para

comprender un concepto determinado por medio de las relaciones que tiene con otros conceptos de alta proximidad y sus características diferenciadoras. La elaboración de un mentefacto conceptual contempla cuatro elementos: *supraordinada*, *exclusora*, *infraordinada* e *isoordinada* (Figura 4), y por tanto cuatro operaciones intelectuales asociadas: *supraordinar*, *excluir*, *infraordinar* e *isoordinar* (De La Herrán y Linares Álvaro, 2013).

Figura 4. Estructura del Mentefacto conceptual aplicado al conocimiento de camélidos del nuevo mundo.



Nota: Elaboración Propia.

La construcción de un mentefacto conceptual inicia determinando el “concepto central”. Este se refiere al concepto que se quiere enseñar, comprender, conocer o conceptualizar. Por ejemplo, un concepto central puede ser “diseños de investigación empíricos en psicología”, acorde con la clasificación de Ato, López, y Benavente

(2013). Una vez se ha decidido este concepto, se procede a construir las demás relaciones del mentefacto.

En segundo lugar, se encuentra el elemento *supraordinada*, que implica el proceso mental de *supraordinar* (De Zubiria Samper, 2000). Este proceso corresponde a la capacidad para encontrar un concepto que contenga de la forma más cercana posible al concepto central (aquel que se quiere enseñar). Por ejemplo, el concepto que contiene de forma más cercana los diseños de investigación empíricos en psicología podría ser “diseños de investigación en psicología”.

En tercer lugar, la estructura de un mentefacto conceptual contempla un concepto *excluser*, esto es, un concepto que pertenece a la clase *supraordinada* (la clase más general) pero que se excluye, diferencia u opone al concepto central (De Zubiria Samper, 2000). En el presente ejemplo habría que indagar por diseños de investigación en psicología, diferentes a los empíricos. En este caso tendríamos diseños metodológicos, instrumentales y teóricos; todos estos representarían las *excluseras* en el mentefacto conceptual.

El cuarto elemento son las *infraordinadas*, que se refieren a conceptos que representan subtipos del concepto central (De Zubiria Samper, 2000). En el presente ejemplo, debemos indagar por tipos de diseños de investigación empíricos en psicología. Tendríamos como *infraordinadas* diseños empíricos manipulativos, asociativos y descriptivos. Cada uno de ellos representa una *infraordinada* diferente en el mentefacto conceptual.

Por último, todos los mentefactos conceptuales deben poseer *isoordinadas*, que hacen referencia a las características principales de cada uno de los conceptos encontrados, enfocándose en lo que diferencia a un concepto de otro (De Zubiria Samper, 2000). Por ejemplo, una *isoordinada* de los diseños de investigación empíricos en psicología de tipo manipulativo, es que son de naturaleza experimental. Como regla general, las *isoordinadas* deben enfocarse en criterios predefinidos que se especificarán en cada uno de los conceptos que componen el mentefacto. Por ejemplo, se podría definir como criterio, la función del tipo de investigación. Esto implicaría que a cada concepto del mentefacto se le adjudicaría una *isoordinada* que describa de la forma más precisa posible, la función principal que cumple en el ámbito de investigativo.

Debido a que la construcción de un mentefacto conceptual implica las cuatro operaciones intelectuales anteriormente ejemplificadas, es evidente que el proceso de elaboración del mismo promueve el uso de habilidades de pensamiento diferentes: clasificar, diferenciar, organizar, inferir y deducir (De Zubiria Samper, 2000); todas estas, dimensiones que ocupan niveles diferentes en la taxonomía de Bloom.

Pregunta de Investigación

Debido a la necesidad de innovar en el desarrollo de actividades en entornos virtuales, que puedan ser usadas como actividades de evaluación dentro del proceso de evaluación formativa, que promuevan procesos de pensamiento superiores y promuevan habilidades metacognitivas y de comprensión en estudiantes; que al mismo tiempo cumplan con las virtudes de las evaluaciones objetivas automatizadas. Y teniendo en cuenta que acorde con la revisión de literatura realizada no existen estudios rigurosos

sobre el uso de los mentefactos conceptuales, la presente investigación se planteó la siguiente pregunta de investigación:

¿Cuál es el efecto de los mentefactos conceptuales como actividad de evaluación formativa, sobre el monitoreo metacognitivo y la comprensión conceptual de estudiantes universitarios, en comparación con actividades objetivas de emparejamiento en ambientes en línea?

Se seleccionó la comprensión conceptual porque es el elemento de aprendizaje básico sobre el cual se construye una competencia de aprendizaje y porque es el filtro básico del cual depende la evaluación formativa para valorar si un estudiante ha avanzado en el aprendizaje de conceptos. Por otro lado, se seleccionó el monitoreo dentro de los procesos metacognitivos porque el monitoreo es el mecanismo por medio del cual un estudiante rastrea sus debilidades y habilidades para organizar su aprendizaje.

Hipótesis de Investigación

H1: El uso de mentefactos conceptuales como herramienta de evaluación formativa promueve mejores niveles de comprensión conceptual en estudiantes universitarios, en comparación al uso de actividades de emparejamiento.

H2: El uso de mentefactos conceptuales como herramienta de evaluación formativa, promueve mejores niveles de precisión absoluta durante el monitoreo metacognitivo, en comparación al uso de actividades de emparejamiento.

H3: El uso de mentefactos conceptuales como herramienta de evaluación formativa, promueve mejores niveles de confianza durante el monitoreo metacognitivo, en comparación al uso de actividades de emparejamiento.

Objetivos

Objetivo General: Evaluar el efecto de los mentefactos conceptuales sobre el monitoreo metacognitivo y la comprensión conceptual de estudiantes universitarios en comparación al uso de actividades de emparejamiento.

Objetivos específicos

1.1 Comparar el nivel de comprensión conceptual alcanzado por los estudiantes universitarios con el uso de mentefactos conceptuales como herramienta de evaluación formativa, en comparación al uso de actividades de emparejamiento.

1.2 Comparar el nivel de precisión absoluta alcanzada por los estudiantes durante el monitoreo metacognitivo con el uso de mentefactos conceptuales como herramienta de evaluación formativa, en comparación al uso de actividades de emparejamiento.

1.3 Comparar el nivel de sesgo alcanzada por los estudiantes durante el monitoreo metacognitivo con el uso de mentefactos conceptuales como herramienta de evaluación formativa, en comparación al uso de actividades de emparejamiento.

Método

Diseño

La presente investigación tuvo un diseño experimental que según Ato et al., (2013) es aquél que busca establecer relaciones causa-efecto entre variables seleccionadas, asegurando la manipulación de por lo menos una variable independiente. Dentro de los diseños experimentales, esta investigación corresponde a un diseño experimental de grupos paralelos con medidas post test (Shadish, Cook y Campbell, 2002).

En coherencia con lo anterior, este estudio exploró el efecto que tiene la variable independiente “tipo de actividad”, sobre las variables dependientes: “precisión absoluta”, “sesgo” y “comprensión conceptual”. Para tal fin, veinticinco estudiantes fueron asignados aleatoriamente a un grupo experimental que realizó la construcción de un mentefacto conceptual (organizador gráfico), y otros veinticinco fueron asignados a un grupo control que realizó una actividad de asociación/emparejamiento. Ambas actividades estuvieron basadas en el contenido de un texto de tipo académico sobre el tema de granulocitos.

La selección de los participantes se realizó por medio de un muestreo no probabilístico de voluntarios (Taherdoost, 2018), y la asignación de los participantes a los grupos se realizó por medio de un método de asignación aleatoria simple a través de la herramienta en línea: <https://www.randomizer.org/>. Esta herramienta permitió asignar a cada participante un número que a su vez correspondía a uno de los dos grupos. Todos los números fueron aleatorizados al inicio del experimento y se fueron asignando en

orden de llegada de los participantes. Se realizó un análisis de potencia para establecer el tamaño de la muestra que nos permitiera alcanzar la potencia deseada.

Las variables dependientes (precisión absoluta, sesgo y comprensión conceptual) se analizaron a través del estadístico t para muestras independientes con el fin de conocer si las diferencias entre los grupos eran estadísticamente significativas.

Participantes

Cincuenta estudiantes chilenos (64% mujeres), entre los 18 y 26 años ($M = 21.44$, $SD = 2.31$), de semestres primero a noveno ($M = 5.1$, $SD = 2.4$) fueron reclutados de cuatro facultades de la Universidad de Concepción (Ciencias Sociales 60%, Ingeniería 16%, Humanidades y Artes 14% y Educación 10%). Los participantes fueron reclutados a través de un muestreo no probabilístico de voluntarios por medio de volantes, colaboración de docentes para la difusión del estudio y acceso directo a los estudiantes en salas de clase y espacios de la Universidad. De estos 50 participantes, 25 fueron asignados aleatoriamente a un grupo experimental y 25 a un grupo control. El cálculo de la muestra se basó en los resultados de una prueba piloto y en un análisis de literatura con una aproximación actuarial. Como criterio de exclusión se determinó que no podrían participar estudiantes pertenecientes a la Facultad de Ciencias Biológicas, Ciencias Naturales y Oceanográficas, Ciencias Químicas, Ciencias Veterinarias, Enfermería, Farmacia, Medicina y Facultad de Odontología. Este criterio de exclusión permitió controlar el efecto de los conocimientos previos sobre la ejecución de la tarea porque eliminó la posibilidad de reclutar estudiantes con posibles conocimientos sobre granulocitos.

Variables

Variable Independiente

Tipo de actividad: Entendida como la actividad de aprendizaje que realizaron los estudiantes para reforzar el contenido de un texto académico.

Posee dos niveles:

Grupo experimental: Actividad basada en la construcción de un mentefacto conceptual.

Grupo control: Actividad basada en la asociación entre conceptos y definiciones (emparejamiento).

Se seleccionó la actividad de emparejamiento porque dentro de las actividades objetivas son las que se asemejan más al tipo de proceso mental que implica la construcción del mentefacto conceptual. Según Anderson et al. (2001), las mejores actividades para promover procesos de clasificación y ordenamiento son las actividades de asociación como el “emparejamiento”. Adicionalmente el proceso interactivo de arrastrar elementos de un lado de la pantalla a otro y la posibilidad de utilizar exactamente los mismos componentes en ambas pruebas es posible por medio de la prueba de emparejamiento.

Variables Dependientes

Precisión Absoluta

Definición Conceptual: Grado en que la percepción que tiene una persona sobre su desempeño se corresponde con su desempeño actual.

Definición operacional: Discrepancia entre el juicio estimado acerca del desempeño alcanzado (en una escala de 0 a 100) y el desempeño real del estudiante en la comprensión conceptual del tema asignado (granulocitos). La fórmula corresponde a la diferencia en el valor absoluto entre el desempeño real y el estimado.

Medida de Sesgo

Definición Conceptual: Valora el grado en que un individuo demuestra exceso o falta de confianza en su desempeño.

Definición operacional: Discrepancia entre el juicio estimado acerca del desempeño alcanzado (en una escala de 0 a 100) y el desempeño real del estudiante en la comprensión conceptual del tema asignado (granulocitos). La fórmula corresponde a la diferencia entre el desempeño real y el desempeño estimado. Dado que el índice de sesgo puede ser positivo o negativo provee información sobre la dirección y la magnitud de la discrepancia entre confianza y desempeño. La dirección del sesgo (positivo o negativo) indica si existe falta de confianza o exceso de confianza: valores negativos indican falta de confianza y valores positivos exceso de confianza. Por su parte, la magnitud de la discrepancia provee información sobre la severidad del error de juicio. Valores positivos y negativos más alejados del cero demuestran mayor error en el juicio de confianza sobre el desempeño (Schraw, 2009).

Nivel de Comprensión conceptual

Definición conceptual: Capacidad para construir significados de mensajes orales o escritos adquiridos en el proceso de instrucción y las relaciones que se establecen entre

los elementos que conforman un contenido o información (Bisson, Gilmore, Inglis, y Jones, 2016; Chadwick, 2009)

Definición operacional: Se entenderá como la calidad de la explicación construida por los estudiantes acerca del significado del concepto a aprender, en este caso de los granulocitos, al finalizar el experimento. La calidad de la explicación será valorada por medio de una rúbrica que consta de 10 ítems basados en indicadores de comprensión conceptual. El puntaje máximo que puede alcanzar un participante en la rúbrica es de 100 puntos.

Variables de control: En ambos grupos se registró automáticamente el tiempo que tardaron los estudiantes en resolver cada actividad y el número de errores que cometieron resolviendo la actividad. Así mismo, el nivel de complejidad de las actividades se midió en el pilotaje y no demostró diferencias significativas entre los dos grupos (Ver en resultados).

Instrumentos y Materiales

Medición de los juicios de confianza para el índice de Monitoreo Metacognitivo

Debido a que el monitoreo se refiere a la discrepancia entre el juicio de confianza y el desempeño real del estudiante, se midieron los juicios de confianza de los estudiantes posterior a la valoración de la comprensión conceptual. Los juicios de confianza se valoraron por medio de una escala continua de 0% a 100% donde cada estudiante debía indicar qué tan seguro o segura se sentía respecto a la calidad de la explicación que había entregado para el concepto de “granulocitos”. Esta escala fue

presentada en un formulario de Google y los estudiantes tenían la posibilidad de indicar el porcentaje de forma libre.

Medición de la Comprensión Conceptual

Una rúbrica fue diseñada para valorar la comprensión conceptual de los estudiantes después de realizar la tarea experimental, a través de la siguiente pregunta abierta sobre la comprensión del concepto de granulocitos: “A continuación, explica de la forma más completa, detallada y extensa posible, lo que has entendido acerca de los granulocitos.” Se usó una pregunta abierta porque de acuerdo con Anderson et al., (2001), las preguntas abiertas tienen la capacidad para capturar la estructura completa de conocimiento que una persona ha formado sobre un tema o concepto particular. Adicionalmente, en el meta análisis de Prinz et al. (2020) se reporta que los estudios relacionados con monitoreo metacognitivo que midieron comprensión por medio de preguntas abiertas, tuvieron tamaños de efecto muy grandes ($d = 0.72$) en comparación con otras formas de preguntas de tipo objetivo que tuvieron tamaños de efecto algo menores ($d = 0.5$).

La rúbrica midió el número de conceptos de alta proximidad que los estudiantes fueron capaces de relacionar con el concepto de granulocitos (agranulocitos, leucocitos, eosinófilos, neutrófilos y basófilos) y las características más representativas de cada uno de estos conceptos de alta proximidad, incluyendo las características principales de los granulocitos. Una rúbrica inicial de 14 ítems fue diseñada para puntuar la comprensión reportada por los estudiantes. Esta rúbrica tenía un puntaje máximo de 140 y se validó por medio del proceso del juicio de cinco jueces expertos en el tema de granulocitos

(cuatro bacteriólogas y una médica) usando el índice de validez de contenido de Lawshe para pertinencia y coherencia (Tristan-López, 2008). Para tal fin, cada juez tuvo a su disposición una matriz de planeación del instrumento que incluyó los ítems a valorar, sobre los cuales deberán emitir su opinión con respecto al cumplimiento de tres condiciones: 1) esencial, 2) útil pero no esencial y 3) no necesario. Una vez emitidas las opiniones de cada uno de los jueces y cada una de las condiciones, se establece el consenso entre las valoraciones de los jueces participantes, para lo cual se emplea la razón de validez de contenido (Cvr, por sus siglas en inglés). Se espera que el grado de acuerdo, principalmente en la categoría de esencial, supere el 50% del total de jueces participantes, para considerar que cada ítem es apto para formar parte del instrumento en cuestión. Para *coherencia* los 14 ítems obtuvieron un índice de validez de 1, sin embargo, para *pertinencia* diez ítems obtuvieron un índice de validez de 1, mientras que 4 obtuvieron un índice de validez de 0.6. Debido a que valores inferiores a 0.8 no son aceptable para juicios con 5 jueces o menos, estos cuatro ítems fueron eliminados de la rúbrica (Taherdoost, 2018b). Por tanto, la rúbrica final, constó de 10 ítems con un puntaje máximo de 100 puntos.

Texto de lectura sobre granulocitos.

Debido a que los estudiantes realizaron una actividad de evaluación formativa, el contenido de las mismas se basó en la lectura previa de un texto de tipo académico denominado “granulocitos” (Anexo 2). Se seleccionó esta temática por varias razones: a) Es un tema que no es común en áreas de ciencias sociales y esto evita que los participantes tengan conocimientos previos al respecto, b) es un tema de alta

especificidad conceptual, lo que asegura un nivel de dificultad alto para participantes, c) es un tema basado en evidencia científica y su taxonomía está definida de manera homogénea en la literatura. El texto sobre granulocitos describe la taxonomía y características principales de los granulocitos y es una adaptación de varios textos académicos sobre inmunología y sangre. El texto fue revisado por una bacterióloga y una médica quienes analizaron y verificaron que todo el contenido del texto fuese coherente con la literatura actual sobre granulocitos. El texto cuenta con 3488 caracteres, 517 palabras, 34 frases, 6 párrafos, y un tiempo aproximado de lectura de 2.6 minutos (Obtenido por medio de Legible: <https://legible.es/>).

Legibilidad lingüística del texto. La legibilidad lingüística de un texto, también denominada comprensibilidad, perspicuidad o lecturabilidad, se refiere a la facilidad para comprender un texto sin necesidad de realizar grandes esfuerzos. Se puede medir aplicando algoritmos sencillos, que son específicos de cada lengua y requieren una investigación científica previa para su validación. Algunas de las fórmulas válidas para el español son:

- a) Escala de lecturabilidad: La lecturabilidad se refiere a la legibilidad lingüística de un texto. José Fernández Huerta (1959), creó la segunda fórmula para medir legibilidad en textos en español, basada en la fórmula de Flesch para el inglés. El resultado de la fórmula obtenido por el analizador de legibilidad de textos en línea (legible.es) reporta una lecturabilidad de 50, lo que corresponde a un nivel “difícil” de lecturabilidad. Acorde con el autor, un nivel difícil de lecturabilidad es adecuado para estudiantes que se encuentren en cursos secundarios, esto es un

nivel escolar entre estudios preuniversitarios y universitarios de especialización (Fernández Huerta, 1959).

- b) La legibilidad es una fórmula para calcular la facilidad lectora de un texto. Fue desarrollada por Miguel Muñoz Baquedano y José Muñoz Urra en Chile en 2006. El resultado obtenido por el analizador arrojó un valor de legibilidad de 41, lo que lo ubica en difícil, que corresponde al sexto nivel de dificultad entre siete posibles (Baquedano, 2006).

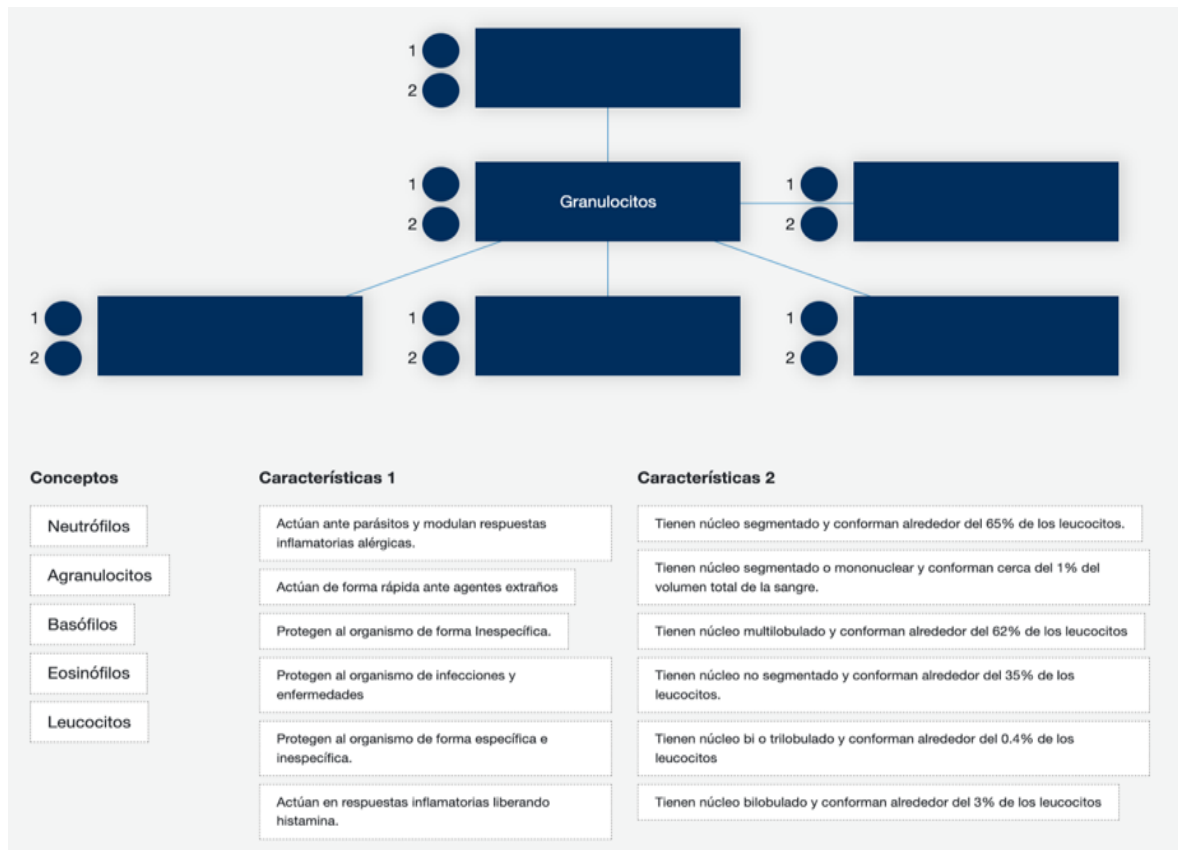
Actividades de evaluación formativa en línea

Las actividades para ambos grupos fueron sistematizadas por un especialista en programación y su acceso se realizó por medio de una página web donde a cada participante se le asignó un usuario y contraseña. El acceso a cada usuario dentro de la página fue bloqueado automáticamente una vez el participante resolvía la actividad, esto permitió controlar que cada participante no pudiera abrir la aplicación más de una vez. Ambas actividades contaron con componentes idénticos en sus interfaces, es decir, los mismos conceptos y características. Adicionalmente, ambas actividades fueron programadas para proveer retroalimentación inmediata sobre cada respuesta dada durante la resolución de la actividad y se resolvían arrastrando y soltando elementos desde una parte de la pantalla hasta otra.

Actividad del grupo Experimental: La actividad del grupo experimental consistió en organizar un mentefacto conceptual arrastrando elementos ubicados en la parte inferior de la pantalla sobre el organizador en la parte superior (Figura 5). Para esto, los estudiantes tuvieron que arrastrar los elementos que constituyen el mentefacto

desde tres listas diferentes. Los ítems de la lista uno que correspondían a conceptos debían ser arrastrados a los rectángulos disponibles en el diagrama. Las listas dos y tres que corresponden a características, debían ser arrastradas a los círculos con el número de la lista correspondiente. Es decir, todas las características de la lista 1 se debían arrastrar a algún círculo con el número 1 y todas las características de la lista 2 debían arrastrarse a algún círculo con el número 2 en el diagrama.

Figura 5. *Actividad del grupo experimental: Mentefacto Conceptual*



Nota: Imagen tomada de la página web diseñada para el experimento.

Actividad del grupo Control: El grupo control resolvió una actividad tradicional de emparejamiento de conceptos con sus características (Figura 6). Para resolver la

actividad, los participantes debían emparejar las características de la lista 1 a cualquier círculo 1 del concepto seleccionado y las características de la lista 2 a cualquier círculo 2 del concepto seleccionado.

Figura 6. *Actividad de emparejamiento para el grupo control*

Nota: Imagen tomada de la página web diseñada para el experimento.

Procedimiento

Estudio Piloto: El estudio piloto tuvo lugar en el primer semestre de 2022 y presentó un método de grupos paralelos donde 28 estudiantes (65% mujeres) de la Facultad de Ciencias Sociales de la Universidad de Concepción, entre 18 y 30 años ($M = 21.7$, $SD = 2.3$) fueron reclutados por medio de un muestreo no probabilístico por conveniencia. Todos los participantes fueron asignados aleatoriamente a las condiciones

experimentales: construcción de mentefactos (grupo experimental) y emparejamiento (grupo control).

Los resultados de este estudio indicaron que ninguno de los estudiantes tenía conocimientos previos sobre el concepto de granulocitos ($M = 0$), razón por la cual esta variable no se convirtió en covariante durante el estudio piloto. Así mismo, en el estudio piloto se midió el nivel de complejidad de las actividades experimentales por medio de una valoración realizada por cada participante después de resolver la actividad asignada. Los resultados de una prueba t de muestras independientes basados en las valoraciones de una escala Likert de 5 puntos, indicaron que no hubo diferencias estadísticamente significativas entre la dificultad que percibieron los participantes al realizar la tarea del grupo experimental ($M_{mentefacto} = 2.78$, $SD = 0.97$) y la dificultad experimentada al resolver la tarea del grupo de control ($M_{emparejamiento} = 3.21$, $SD = 1.05$), $t(26) = 1.119$, $p = .274$.

Contexto del experimento: El experimento tuvo lugar de manera individual y presencial en un escenario de laboratorio del Departamento de Psicología. Cada participante fue ubicado frente a la pantalla del computador, con acceso a Internet, a una distancia aproximada de 50 centímetros donde podía alcanzar el teclado y el mouse. Todas las aplicaciones fueron realizadas en horas de la tarde, en días entre lunes y viernes y fueron supervisadas por la misma investigadora, asegurando el mismo procedimiento a lo largo del tiempo. Si bien el experimento no fue ciego en relación con la supervisión del experimento, las instrucciones y la ejecución del experimento fueron administradas automáticamente por un computador, reduciendo así la influencia del

investigador sobre los participantes. El experimento tuvo una duración aproximada de 45 minutos.

Proceso del experimento

Los tiempos establecidos para cada uno de los pasos del procedimiento se encuentran en la figura 7.

a) Consentimiento informado: Todos los estudiantes leyeron y firmaron el consentimiento informado aprobado por el Comité de ética, bioética y bioseguridad de la Vicerrectoría de Investigación y Desarrollo de la Universidad de Concepción (Anexo 3).

b) Registro de datos demográficos: Los estudiantes registraron sus datos personales (nombres, apellidos, correo electrónico, carrera, edad, universidad y semestre) en la primera página de un formulario de Google en línea.

c) Valoración de conocimientos previos como medida control: Los estudiantes debían responder a la pregunta: ¿Qué son los granulocitos? o indicar “no sé” en caso de no tener conocimiento al respecto en la segunda página del formulario de Google.

d) Lectura del texto sobre granulocitos: La lectura se presentó por medio de una presentación de cinco diapositivas en PowerPoint que los estudiantes podían manipular para controlar el ritmo de lectura deseado. Las instrucciones sobre la lectura indicaron que la lectura solo se podía hacer una vez, aspecto que fue supervisado por la investigadora encargada.

e) Instrucciones sobre la actividad asignada: Las instrucciones de ambos grupos fueron presentadas por medio de una presentación en PowerPoint, una para el grupo experimental y otra para el grupo control. Las instrucciones indicaban a los participantes

en qué consistía la actividad a resolver, proveía un ejemplo de una actividad resuelta con el tipo de actividad correspondiente al grupo y una simulación de cómo resolver la actividad asignada paso a paso, por medio de una diapositiva interactiva.

f) Acceso a la página web del experimento: Una vez leídas las instrucciones en PowerPoint, cada estudiante tuvo acceso a la página web del experimento donde debía ingresar con un usuario y contraseña asignados.

g) Desarrollo de un ensayo de la tarea experimental en la página web programada con la actividad asignada: Para asegurar un adecuado manejo de la interface de la actividad asignada y comprensión de las instrucciones dadas previamente, cada estudiante tuvo oportunidad de resolver una actividad de ejemplo antes de realizar la tarea experimental propiamente tal. Este ensayo fue sobre el tema de “camélidos” (tipos de camellos). Durante este simulacro los estudiantes practicaron la forma de hacer la actividad correspondiente, resolviendo dudas sobre el procedimiento indicado para hacer la actividad y la forma práctica de arrastrar los objetos en el espacio con el mouse.

h) Actividad experimental asignada: Una vez que los participantes terminaban de resolver la actividad de simulacro en su totalidad, eran dirigidos a la actividad experimental (mentefacto conceptual) sobre los granulocitos o a la actividad control (actividad de emparejamiento) sobre los granulocitos descritas previamente. Esta actividad tenía que ser resuelta por completo para poder cerrar la actividad. Si bien se midió el tiempo y el número de errores por participante, esta información no estuvo disponible en la pantalla de los participantes durante la actividad.

i) Medición de la comprensión conceptual: Una vez finalizada la actividad experimental, indicado a través de un letrero personalizado en la pantalla de cada participante, los estudiantes fueron remitidos nuevamente al formulario de Google donde debían responder de forma abierta a la siguiente indicación: “A continuación, explica de la forma más completa, detallada y extensa posible, lo que has entendido acerca de los granulocitos”. Para esto, los estudiantes tuvieron el tiempo que desearon para responder la pregunta sin consultar el texto, la actividad realizada, o internet en sus celulares.

j) Medición del monitoreo metacognitivo: En seguida a la pregunta sobre comprensión en el formulario de Google, los estudiantes debían responder a la siguiente pregunta: En una escala de 0 a 100, ¿Qué tan seguro/a te sientes con la explicación que acabas de escribir? Escribe el número que consideres teniendo en cuenta que 0 corresponde a nada seguro/a y 100 a totalmente seguro/a.

k) Envío de formulario: Una vez terminado el juicio de confianza, los estudiantes enviaban el formulario y se daba por terminada la sesión.

Figura 7. *Duración del proceso experimental.*

Lectura y firma del consentimiento informado	• 5 minutos
Registro de datos demográficos	• 3 minutos
Valoración de conocimientos previos	• 2 minutos
Lectura del texto sobre granulocitos	• 8 minutos
Lectura de instrucciones	• 5 minutos
Actividad de simulacro	• 5 minutos
Aplicación de Actividades experimentales	• 5 minutos
Valoración de comprensión conceptual	• 6 minutos
Valoración del monitoreo metacognitivo	• 1 minuto
Tiempo total aproximado	• 40 minutos

Consideraciones éticas

La presente investigación se acogió a diversos códigos y principios de ética aplicados a la investigación en educación y en psicología. Entre ellos asumió los principios éticos del Informe Belmont (Belmont/Report, 1979), que se acogen a cualquier tipo y ámbito de investigación y se basan en los 3 aspectos:

- a) Respeto por las personas: Contempla los requisitos morales donde se debe reconocer la autonomía personal de los participantes y la protección de aquellas personas con una disminución de la misma.
- b) Beneficencia: Se refiere al compromiso de no dañar y posibilitar la maximización de beneficios para los participantes en el estudio.

c) Justicia: Se refiere a la obligación que tiene todo investigador de proveer un trato humanitario a los participantes.

De forma más específica, esta investigación también se acogió a los códigos en investigación en el ámbito de la educación. Entre ellos el código de la Asociación Americana para la Investigación Educativa (AERA, 2011) se centra en cinco principios generales: Competencia Profesional, Integridad, Responsabilidad Científica y Profesional, Respeto de los Derechos de Personas, su Dignidad y Diversidad y Responsabilidad Social. Estos principios, al igual que los principios éticos para psicólogos de la APA (APA, 2017) velan por los derechos y dignidad de los participantes en la investigación; haciendo una llamada a la protección de la privacidad de los participantes y de los datos, aludiendo de forma particular a la responsabilidad del investigador respecto a la transparencia en los procesos de investigación y rechazo absoluto del fraude y el plagio.

Por otro lado, se declara que la presente investigación no representó ningún tipo de riesgo mayor a las actividades habituales de los estudiantes, ni interfiere con la integridad física o emocional de los mismos; dando así cumplimiento a la Ley 20120 nacional. Por el contrario, los estudiantes que participaron en la investigación se pudieron ver beneficiados por la adquisición de nuevos conocimientos y herramientas para potenciar sus estrategias de estudio. Finalmente, esta investigación garantizó la privacidad y confidencialidad de los datos personales de cada una de las personas que participaron en el estudio. Si bien las actividades fueron realizadas en línea, el acceso a los datos solo se podía acceder por medio de una clave privada que migró los datos a un

archivo de Excel de acceso privado. De este modo, también se da cumplimiento a la Ley 19628 nacional.

Resultados

Resultados de la investigación principal

En primer lugar, no se registró ninguna pérdida de información durante el proceso de recolección y análisis de datos, por tal, los cincuenta participantes reclutados fueron incluidos en los resultados. Los datos obtenidos por medio de la aplicación web diseñada fueron registrados automáticamente en la nube para ser descargados en cualquier momento.

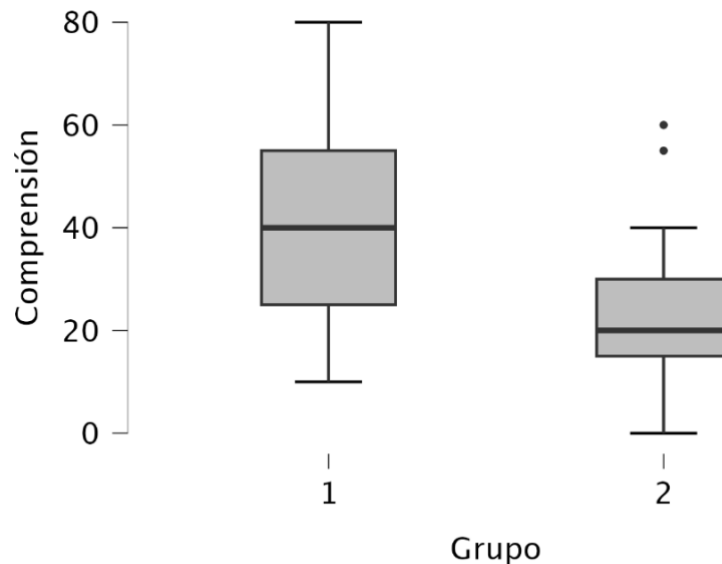
Se analizaron las variables de control conocimientos previos, tiempo y errores. Respecto a los conocimientos previos, se corroboró que ninguno de los participantes tenía conocimientos previos sobre el concepto de granulocitos, razón por la cual no fue necesario incluir “conocimientos previos” como covariable en el estudio ($M = 0$).

Respecto al tiempo (medido en minutos) empleado para resolver la actividad, se evidencia que, si bien los participantes del grupo experimental demoraron menos tiempo en promedio en resolver la actividad, no se registraron diferencias estadísticamente significativas entre los grupos: ($M = 3.8$, $SD = 1.5$) y ($M = 4.1$, $SD = 2.3$), $t(48) = .56$, $p = .28$, $d = 0.16$. Por su parte, respecto a la cantidad de errores, una prueba t de student para muestras independientes mostró que se encontraron diferencias estadísticamente significativas entre los grupos, donde el grupo experimental presentó menos errores resolviendo la tarea que el grupo control: ($M = 10.8$, $SD = 8.3$) y ($M = 15.9$, $SD = 11.2$), $t(48) = 1.82$, $p = .037$, $d = 0.5$.

Habiendo analizado los resultados de las variables de control, enseguida se reportarán los resultados para las variables dependientes del estudio. Con respecto a la

variable de comprensión, se realizó una prueba *t* de *Welch* de muestras independientes para comparar el desempeño de los estudiantes obtenido en la rúbrica de comprensión sobre granulocitos, rúbrica que se valoró en una escala de 0 a 100 puntos. Los resultados de la prueba *t* de *Welch* indican que hubo una diferencia estadísticamente significativa entre las puntuaciones de comprensión entre el grupo experimental ($M = 40.8$, $SD = 19.82$) y el grupo control ($M = 23.68$, $SD = 14.12$), $t(43.3) = 3.51$, $p < .001$, $d = 0.995$. La figura 8 representa el desempeño en comprensión obtenido en cada grupo. Los valores del eje *y* representan el puntaje de la rúbrica de comprensión.

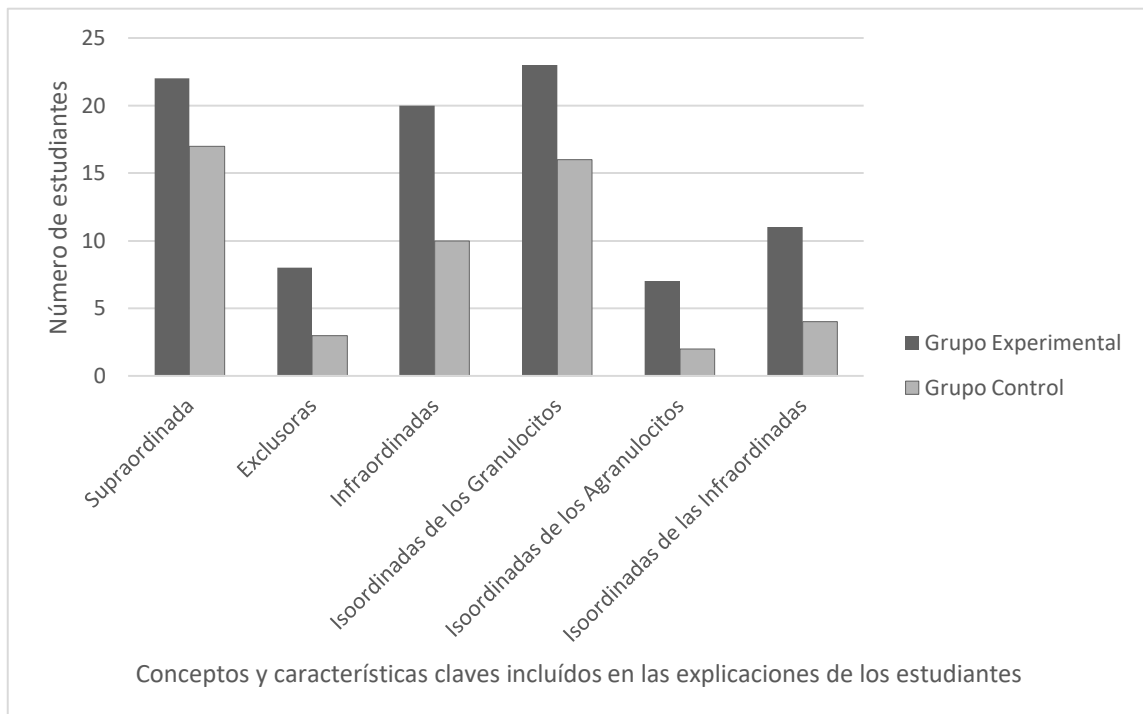
Figura 8. *Desempeño en comprensión.*



Para obtener un alto rendimiento en la rúbrica de comprensión, los estudiantes tenían que demostrar el conocimiento de diferentes conceptos de proximidad relacionados con los granulocitos. La Figura 9 representa la frecuencia de estudiantes que incluyeron en sus explicaciones los siguientes conceptos y características:

supraordinada (categoría general que incluye a los granulocitos: los leucocitos o glóbulos blancos), exclusoras (otros conceptos pertenecientes a la categoría general de los granulocitos: granulocitos), infraordinadas (tipos o clasificaciones de los granulocitos: basófilos, eosinófilos y neutrófilos), isoordinadas de granulocitos (características principales de los granulocitos relativas a su función y anatomía), isoordinadas de agranulocitos (características principales de los agranulocitos relativas a su función y anatomía) e isoordinadas de infraordinadas (características principales de los basófilos, eosinófilos y neutrófilos relativas a su función).

Figura 9. *Número de estudiantes que incluyeron en sus explicaciones conceptos y características claves para definir el concepto de granulocitos.*



Como se evidencia en la figura 9, una mayor cantidad de estudiantes del grupo experimental fueron capaces de incluir los elementos claves del significado de

granulocitos en sus explicaciones en comparación con los estudiantes del grupo control. Estos datos son muy importantes porque podrían indicar que, en el proceso de conceptualizar, lo que más se le dificultó a los estudiantes fue la capacidad para diferenciar un concepto de otros similares y caracterizarlo.

Además de las diferencias cuantitativas en la comprensión entre los grupos, se pudieron evidenciar diferencias de tipo cualitativo en relación a la estructura, organización y coherencia de las explicaciones elaboradas por los estudiantes. Por ejemplo, la mayoría de los estudiantes del grupo experimental explicó su comprensión sobre los granulocitos de forma jerárquica, describiendo primero la categoría supraordinada (Leucocitos), acercándose a las excluyentes en segundo lugar (Agranulocitos), demostrando las características generales de los granulocitos y sus diferencias con los agranulocitos en tercer lugar y, por último, definiendo las subclases de granulocitos (basófilos, neutrófilos y eosinófilos) y sus características. Así mismo, los participantes del grupo experimental fueron capaces de recordar más información y detalles específicos. Por ejemplo:

- a) *Los granulocitos son un tipo de leucocitos que se encargan de proteger nuestro cuerpo o sistema inmune de manera inespecífica. Componen el 65% de los leucocitos. Tienen un núcleo segmentado. Además, se dividen en 3 tipos, los cuales tienen funciones específicas. Actuar contra enfermedades inflamatorias, infecciones y agentes extraños (Participante # 22).*
- b) *Los granulocitos son una parte de los leucocitos, que forman parte del sistema inmune. Los leucocitos también se dividen en agranulocitos, cuyo núcleo es*

distinto al primero ya mencionado. Los granulocitos se dividen en 3 partes que se diferencian por sus funciones, pero que ahora no recuerdo sus nombres. Sin embargo, uno de ellos funcionaba para respuestas rápidas del sistema, otro para reaccionar ante parásitos y un tercero con función antihistamínica (Participante #4).

c) Los granulocitos, células pertenecientes a los glóbulos blancos, comprenden uno de los planes de acción para defender al cuerpo de infecciones, bacterias, hongos, etc. Estos comprenden el 65% de la totalidad de los linfocitos, y tienen, a su vez, tres sub divisiones. Por un lado se encuentran los neutrocitos (?), considerando el 62% de los granulocitos, y encargados de atacar las infecciones que pudiesen entrar al cuerpo. Luego se encuentran los bastoncitos (?), comprendiendo el 4% aproximadamente de los granulocitos, y responsables de la defensa contra las bacterias y hongos. Y por último, se encuentran los eostocitos (?), comprendiendo el 0.1% de los granulocitos, encargados de la respuesta sobre la inflamación, controlando así la histamina. (Participante #3).

d) Primeramente, los granulocitos y los agranulocitos son quienes conforman a los leucocitos, estos corresponden a los glóbulos blancos, encargados de proteger nuestro sistema inmune. Los granulocitos conforman el 65% de los leucocitos y, entre sus características principales está que presentan un núcleo segmentado. Estos tienen 3 subdivisiones, donde una de ellas correspondía al 62% de ese 65%, otro al 2,6% y el último al 0,4% (no recuerdo el nombre de cada subdivisión. (Participante #6).

Por el contrario, algunos participantes del grupo de control tendieron a expresar información segmentada en orden aleatorio, al igual que las proposiciones emparejadas durante la actividad en línea:

a) *Se encargan de proteger al organismo de agentes externos extraños potencialmente dañinos, como las bacterias, inflamaciones, etc.* (Participante #23).

b) *Son células del sistema inmune que combaten sustancias ajenas a las del cuerpo.* (Participante #37).

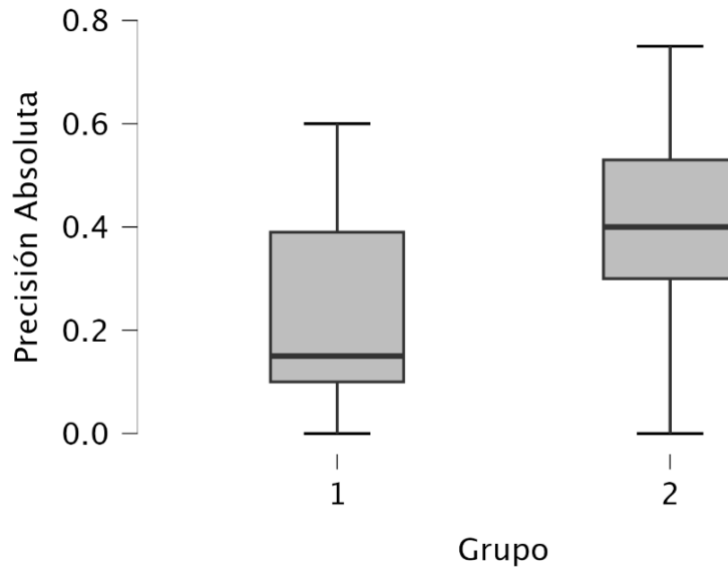
c) *Protegen al organismo de forma inespecífica y tienen núcleo segmentado.* (Participante #24).

d) *Los granulocitos son parte de lo que compone a los glóbulos blancos. Tienen distintas funciones dependiendo de lo que este atacando al cuerpo y ayudan a combatir agentes extraños en el sistema humano.* (Participante #49).

En cuanto a la segunda variable dependiente, precisión absoluta, se realizó una prueba *t* de *Student* de muestras independientes para comparar la precisión absoluta entre los grupos. La precisión absoluta corresponde a la medida de discrepancia entre el desempeño real en la prueba de comprensión y el juicio de valor de confianza sobre su desempeño en esta prueba (ambas medidas en una escala de 0 a 100). Los resultados indican que hubo una diferencia estadísticamente significativa entre las puntuaciones de precisión en el grupo experimental ($M = 0.23$, $SD = 0.19$) y el grupo control ($M = 0.38$, $SD = 0.20$), $t(48) = 2.72$, $p = .004$, $d = .77$. La figura 10 representa las puntuaciones

obtenidas para la precisión absoluta dentro de cada grupo. Es importante recordar que valores más cerca del 0 indican una mayor precisión.

Figura 10. *Precisión Absoluta.*

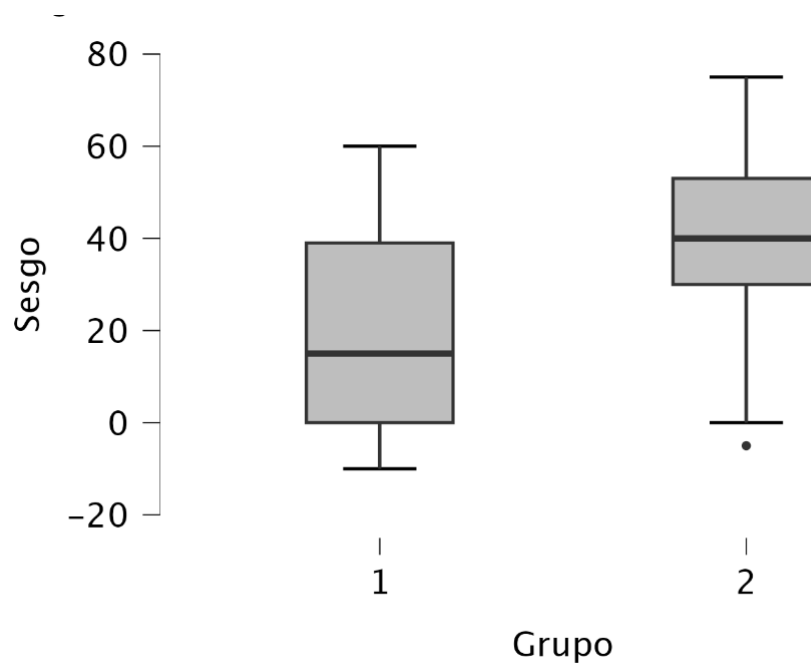


Los resultados acerca de la precisión absoluta significan que los estudiantes del grupo control fueron menos precisos que los estudiantes del grupo experimental estimando su desempeño en la comprensión del concepto de granulocitos.

La precisión absoluta previamente reportada está asociada con la tercera variable dependiente del estudio, la confianza sobre el desempeño, variable que se midió por medio de una medida de sesgo. La medida de sesgo es similar a la medida de precisión absoluta, es decir que representa la diferencia entre el desempeño y el nivel de confianza estimado sobre el desempeño, pero acepta valores negativos que permiten medir la dirección y magnitud de la confianza respecto al desempeño. Los resultados para la prueba *t* de *Student* indican que hubo una diferencia estadísticamente significativa entre

las puntuaciones de precisión en el grupo experimental ($M = 20.2$, $SD = 22.1$) y el grupo control ($M = 38.1$, $SD = 21.6$), $t(48) = 2.8$, $p = .003$, $d = .81$. La figura 11 representa la medida de sesgo obtenida en cada uno de los grupos. Valores negativos representan falta de confianza, valores iguales a cero representan una confianza adecuada y valores positivos representan exceso de confianza, donde a mayor magnitud, mayor exceso de confianza.

Figura 11. Medida de sesgo para cada grupo.



Discusión

Los resultados apoyan la hipótesis de que los mentefactos conceptuales promueven una mejor comprensión conceptual de textos académicos, una mayor precisión en el monitoreo metacognitivo y niveles de confianza mayores, mientras se cometen menos errores en la ejecución de la actividad, en comparación con actividades de emparejamiento. En este sentido es importante destacar que los estudiantes del grupo experimental no solo fueron capaces de incluir más conceptos de proximidad relacionados con los granulocitos, sino que sus explicaciones demostraron una comprensión más estructurada, organizada y coherente del concepto.

Uno de los argumentos que se defendieron en el problema de investigación como necesidad para generar nuevas formas de evaluación en línea, era la importancia de que las herramientas o actividades utilizadas promovieran procesos de pensamiento que ascendieran en la pirámide de Bloom. Al leer las explicaciones dadas por los estudiantes del grupo experimental y conocer el número de conceptos de alta proximidad y características incluidas en sus explicaciones, podríamos afirmar que efectivamente los estudiantes de este grupo fueron capaces de conceptualizar de manera más profunda el significado de “granulocitos”. Esto nos lleva a suponer que la construcción de mentefactos conceptuales involucró un conjunto diversificado de habilidades de pensamiento que ayudaron a los estudiantes a comprender el concepto de forma más organizada, robusta y coherente; habilidades que de acuerdo con Solomon et al. (1999) involucran: categorización, deducción, predicción, comparación, análisis, comunicación, planeación y monitoreo, entre otros.

En este sentido, al leer las explicaciones realizadas por los participantes del grupo experimental, fue notable ver cómo comprendían que un concepto no solo se define por sus características, sino por todos los conceptos de proximidad que se relacionan con él. Como lo describieron Anderson et al., (2001), la comprensión se refiere a las relaciones que una persona establece entre los elementos que constituyen el contenido en un dominio particular (Bisson, Gilmore, Inglis, y Jones, 2016; Chadwick, 2009). Aunque la comprensión conceptual es solo una parte del proceso de aprendizaje necesario para alcanzar la competencia en un contexto concreto, la comprensión es la base sobre la que se construyen el resto de habilidades. Cuando existe una buena comprensión de los conocimientos, los estudiantes pueden relacionar, organizar, discutir, juzgar, manipular y explorar otras destrezas que podrían conducir a un aprendizaje más práctico o procedimental (Anderson et al., 2001). Por lo tanto, disponer de una herramienta automatizada que sirva para verificar los conocimientos y que pueda ayudar a los estudiantes a adquirir una comprensión conceptual de los textos académicos es fundamental, sobre todo en la enseñanza superior en línea, donde se espera que los estudiantes tengan unas habilidades de comprensión lectora adecuadas y sean parcialmente autónomos en su aprendizaje.

Por otro lado, los resultados reportados apoyan las ideas obtenidas en investigaciones previas con mentefactos conceptuales donde se concluye que los mentefactos conceptuales ayudan a los estudiantes a adquirir conceptos en áreas temáticas específicas, favorecen la comprensión lectora, la comprensión conceptual y el desarrollo de habilidades cognitivas superiores (Alvarado Bustamante et al., 2022; Del

Carmen Gutiérrez et al., 2021; Ibañez Pinilla, 2006; Mendoza y Huarachi, 2008; Mera Cabezas y Viñan Ajila, 2023; Mucha Hospinal et al., 2022; Quevedo, 2018; Riveros Melo, 2017; Torres Alférez, 2016).

En cuanto al monitoreo metacognitivo, los mentefactos conceptuales ayudaron a los estudiantes a emitir juicios más precisos sobre su rendimiento en comparación con los estudiantes del grupo de control. La capacidad para predecir con exactitud el propio rendimiento es fundamental porque ayuda a los estudiantes a realizar un seguimiento de los procesos y estrategias cognitivas en curso durante la resolución de problemas y les alerta sobre su necesidad de adaptarse al entorno de aprendizaje tomando decisiones tácticas para mejorar su educación (Ellis y Zimmerman, 2001). Especialmente en las prácticas de evaluación formativa en línea, los procesos de seguimiento eficaces son cruciales para dirigir las acciones de los estudiantes con el fin de mejorar oportunamente los resultados del aprendizaje (Nietfeld et al., 2005).

Sobre la "ilusión de saber", descrita previamente como la creencia de haber logrado una buena comprensión cuando no se ha logrado (Avhustiuk et al., 2018; Hall et al., 2007), observamos que los estudiantes del grupo experimental mostraron niveles más bajos de exceso de confianza que los estudiantes del grupo de control. Esto podría explicarse por el hecho de que la construcción de mentefactos conceptuales ejercitaba diferentes habilidades de pensamiento que pueden ser más variadas que las que se ponen en juego durante las actividades de emparejamiento. Como afirma la literatura, las actividades que no pueden promover habilidades de pensamiento de alto orden pueden no estimular habilidades metacognitivas precisas (Avhustiuk et al., 2018). De hecho, se

ha demostrado que el uso permanente de evaluaciones objetivas genera un efecto de exceso de confianza, por lo que se sugiere variar el tipo de herramienta usada para evaluar y dar prioridad a actividades que promuevan más procesos cognitivos (Dutke et al., 2010).

Es importante mencionar que, además de que la actividad del grupo experimental permitió una mejor comprensión, precisión en el monitoreo y promovió niveles de pensamiento más complejos, los estudiantes la ejecutaron con mayor rapidez y menor cantidad de errores. Así mismo, se debe reconocer que estas diferencias se encontraron en un proceso experimental que no incluyó un tiempo de estudio por parte de los estudiantes, donde no tuvieron la oportunidad de hacer una doble lectura del texto académico, donde no había un interés particular por el tema, donde no hubo una instrucción sobre el tema y donde no tuvieron la oportunidad de resolver inquietudes, lo que significa que, en un proceso de aprendizaje formal, los mentefactos conceptuales podrían tener un potencial aún mayor.

Los resultados sugieren que las preguntas cerradas automatizadas utilizadas tradicionalmente en los entornos virtuales dan a los estudiantes una sensación de exceso de confianza que puede interferir en su capacidad para autorregular su aprendizaje y lograr una comprensión profunda. Aunque solo se probaron actividades de emparejamiento, otras preguntas cerradas estructuradas podrían tener los mismos efectos. Esto es crucial porque la mayoría de los cursos en línea síncronos y asíncronos verifican la comprensión utilizando preguntas de opción múltiple, actividades de emparejamiento, preguntas de rellenar lo que está en negro o cuestionarios de verdadero-

falso. En consecuencia, los proveedores de tecnología, los docentes que trabajan de manera virtual, los responsables de las políticas de educación en línea, los diseñadores, etc., deben tener en cuenta que la estructura de una evaluación puede afectar al potencial para lograr una comprensión profunda y un aprendizaje autorregulado adecuado (Avhustiuk et al., 2018; Hall et al., 2007). En el presente estudio, ambos grupos tenían el mismo contenido (conceptos y características idénticas), lectura e información. Sin embargo, la estructura de las tareas y los procesos cognitivos implicados durante su resolución llevaron a los estudiantes a alcanzar diferentes niveles de comprensión y monitoreo.

El diseño de la evaluación debe ser riguroso para que las prácticas de evaluación formativa en línea logren su objetivo; de lo contrario, es posible que no conduzcan a la conciencia metacognitiva durante el aprendizaje. Es fundamental tener en cuenta que no porque la comprensión se evalúe constantemente y provea retroalimentación permanente significa que se evalúe correctamente. Los instrumentos utilizados para evaluar el progreso son cruciales para que los estudiantes comprendan y controlen su aprendizaje. Si las evaluaciones proporcionadas no estimulan las habilidades de pensamiento de alto orden, los estudiantes pueden incurrir en una falsa sensación de exceso de confianza que les impida tomar medidas para mejorar su aprendizaje (Avhustiuk et al., 2018; Ellis y Zimmerman, 2001; Hall et al., 2007).

Por último, más allá de los resultados obtenidos, es importante dejar claro que los mentefactos tienen el poder de ejercitar muchas operaciones intelectuales que permitirán a los estudiantes comprender textos académicos complejos o conceptos teóricos y

convertirlos en estructuras conceptuales. De hecho, de acuerdo con De Zubiría (2000), existen muy pocos conceptos que nosotros como adultos dominamos. En este sentido, los mentefactos conceptuales imponen un reto importante a los docentes: organizar los conceptos en nuestra mente antes de tratar de comunicarlos, porque es imposible e incluso extraño enseñar algo que no se ha estructurado y esperar que otros lo organicen, comprendan y recuerden (De Zubiría, 2000).

Conclusiones

Concluimos que los mentefactos conceptuales promueven niveles más altos de comprensión conceptual y niveles más precisos de monitoreo metacognitivo en comparación con las actividades de emparejamiento, por lo que son una excelente alternativa para ser utilizados como tareas automatizadas de evaluación formativa en línea. Este estudio fue la primera investigación que exploró los mentefactos conceptuales automatizados en un entorno experimental riguroso con variables de comprensión y metacognición juntas.

Los resultados obtenidos en esta investigación soportan la idea de Minte Münzenmayer et al., (2021), de que los instrumentos que usa un docente para evaluar el conocimiento de los estudiantes son fundamentales como herramientas pedagógicas, porque ayudan a promover niveles desafiantes de reflexión y cuestionamiento en el aula, convirtiéndose en posibles predictores de aprendizaje (Minte Münzenmayer et al., 2021).

Este estudio tuvo limitaciones que es necesario mencionar. En primer lugar, los datos de esta investigación se obtuvieron en un entorno de laboratorio, por lo que los análisis representan únicamente los resultados obtenidos en este contexto. Así mismo, no se realizaron ciegos para la aplicación del experimento ni para el análisis de datos. Por otro lado, la aplicación desarrollada para realizar las actividades en línea quedó con letra pequeña por lo que algunos estudiantes con dificultad de visión tuvieron que agrandar la pantalla para poder ejecutar el experimento y se evidenció una incomodidad durante la ejecución que pudo haber afectado su desempeño. Esto se pudo haber evitado

disponiendo de una pantalla más grande para que todos los participantes tuvieran un tamaño más grande de letra durante la aplicación.

A pesar de las limitaciones nombradas anteriormente, este estudio demostró el potencial que tienen los mentefactos conceptuales como herramienta para conceptualizar información académica compleja, promoviendo el uso de procesos cognitivos superiores y asegurando procesos de regulación metacognitiva más precisos que otras medidas en línea típicamente usadas, regulando los niveles de confianza de los estudiantes.

Futuras investigaciones podrían explorar el uso de mentefactos conceptuales automatizados en entornos más ecológicos ya que podrían ser especialmente apropiados en procesos de instrucción formal. También podría ser interesante estudiar el proceso de construcción de mentefactos sin proveer a los estudiantes con las opciones de respuesta, sino, directamente desde la información de los textos académicos. Por último, comparar el potencial de los mentefactos conceptuales respecto a los mapas conceptuales es un área de estudio importante, ya que el uso de mapas conceptuales es ampliamente difundido y sería interesante conocer si los mentefactos conceptuales son más efectivos que los mapas conceptuales para la comprensión de conceptos.

Referencias

- Abosalem, Y. (2016). Assessment Techniques and Students ' Higher-Order Thinking Skills. *International Journal of Secondary Education*, 4(1), 1–11.
<https://doi.org/10.11648/j.ijsedu.20160401.11>
- AERA. (2011). Code of Ethics: American Educational Research Association. *American Educational Research Association*, 12. <https://doi.org/10.3102/0013189x11410403>
- Akcaoğlu, M. Ö. y Mor, E. y Külekçi, E. (2023). The mediating role of metacognitive awareness in the relationship between critical thinking and self-regulation. *Thinking Skills and Creativity*, 47. <https://doi.org/10.1016/j.tsc.2022.101187>
- Al-Hattami, A. A. (2020). E-Assessment of Students ' Performance During the E-Teaching and Learning. *International Journal of Advanced Science and Technology*, 29(8), 1537–1547.
- Alvarado Bustamante, M. y Esteban Huerto, Y. N. y Muñoz Carrillo, M. (2022). *Promoviendo el uso de los mentefactos para el aprendizaje en los estudiantes de educación básica alternativa de la institución educativa "Yarowilca" [Segunda Especialidad]*.
- Anderson, L. W. y Krathwohl, D. R. y Airasian, P. W. y Cruikshank, K. A. y Mayer, R. E. y Pintrich, P. R. y Raths, J. y Wittrock, M. c. (2001). *A taxonomy for learning teaching and assessing. A revision of Bloom`s Taxonomy of Educational Objectives*. Pearson Education.
- Andrade, H. L. y Cizek, G. J. (2010). *Handbook of Formative Assessment*. Routledge.
- Anohina-Naumeca, A. (2020). The Conceptual Model of Formative Assessment of Structural Knowledge. In M. J. Spector, B. B. Lockee, & M. D. Childress (Eds.), *Learning, Design, and Technology*. Springer International Publishing. <https://doi.org/10.1007/978-3-319-17727-4>

- Antoniou, Panayiotis. y James, M. (2014). Exploring formative assessment in primary school classrooms: Developing a framework of actions and strategies. *Educational Assessment, Evaluation and Accountability*, 26(2), 153–176. <https://doi.org/10.1007/s11092-013-9188-4>
- APA. (2017). *ETHICAL PRINCIPLES OF PSYCHOLOGISTS AND CODE OF CONDUCT* (pp. 1–16). file:///Users/Ricart/Downloads/APA 2010.pdf
- Appiah, M. y van Tonder, F. (2018). E-Assessment in Higher Education: A Review. *International Journal of Business Management and Economic Research*, 9(6), 1454–1460. www.ijbmer.com
- Ato, M. y López, J. y Benavente, A. (2013). Un sistema de clasificación de los diseños de investigación en psicología. *Anales de Psicología*, 29(3), 1038–1059. <https://www.redalyc.org/pdf/167/16728244043.pdf>
- Ausubel, D. (1968). Significado y aprendizaje significativo. In *Psicología Educativa* (pp. 1–23). Trillas.
- Avhustiuk, M. M. y Pasichnyk, I. D. y Kalamazh, R. V. (2018). The illusion of knowing in metacognitive monitoring: Effects of the type of information and of personal, cognitive, metacognitive, and individual psychological characteristics. *Europe's Journal of Psychology*, 14(2), 317–341. <https://doi.org/10.5964/EJOP.V14I2.1418>
- Azevedo, R. y Alevén, V. (2013). *International Handbook of Metacognition and Learning Technologies*. Springer.
- Bälter, O. y Enström, E. y Klingenberg, B. (2013). The effect of short formative diagnostic web quizzes with minimal feedback. *Computers and Education*, 60(1), 234–242. <https://doi.org/10.1016/j.compedu.2012.08.014>
- Baquedano, M. M. (2006). *Legibilidad y Variabilidad de los Textos*.

- Barana, A. y Marchisio, M. (2016). Ten good reasons to adopt an automated formative assessment model for learning and teaching Mathematics and scientific disciplines. *Procedia - Social and Behavioral Sciences*, 228, 608–613. <https://doi.org/10.1016/j.sbspro.2016.07.093>
- Beissner, K. L. y Jonassen, D. H. y Grabowski, B. L. (1994). Using and Selecting Graphic Techniques to Acquire Structural Knowledge. *Performance Improvement Quarterly*, 7(4), 20–38. <https://doi.org/10.1111/j.1937-8327.1994.tb00648.x>
- Bell, E. y Allen, R. y Brennan, P. (2001). *Assessment of higher order thinking skills: A discussion of the data from the 2001 random sampling exercise and a workshop for teachers*. https://www.qcaa.qld.edu.au/downloads/publications/research_qbssss_assess_hots_01.pdf
- Belmont/Report. (1979). The Belmont Report Ethical Principles and Guidelines for the Protection of Human Subjects of Research. In *National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research*.
- Bisson, M.-J. y Gilmore, C. y Inglis, M. y Jones, I. (2016). Measuring Conceptual Understanding Using Comparative Judgement. *International Journal of Research in Undergraduate Mathematics Education*, 2(2), 141–164. <https://doi.org/10.1007/s40753-016-0024-3>
- Bloxham, S. y Boyd, P. (2007). *Developing Effective Assessment in Higher Education: a Practical Guide*. Open Univeristy Press.
- Boitshwarelo, B. y Reedy, A. K. y Billany, T. (2017). Envisioning the use of online tests in assessing twenty-first century learning: a literature review. *Research and Practice in Technology Enhanced Learning*, 12(16), 1–16. <https://doi.org/10.1186/s41039-017-0055-7>

- Buldu, M. y Buldu, N. (2010). Concept mapping as a formative assessment in college classrooms: Measuring usefulness and student satisfaction. *Procedia - Social and Behavioral Sciences*, 2(2), 2099–2104. <https://doi.org/10.1016/j.sbspro.2010.03.288>
- Butler, D. L. y Winne, P. H. (1995). Feedback and Self-Regulated Learning: A Theoretical Synthesis. *Review of Educational Research*, 65(3), 245–281. <https://doi.org/10.3102/00346543065003245>
- Campbell, D. y Stanley, J. (1963). *Experimental and quasi-experimental designs for research*. Rand McNally.
- Chadwick, D. (2009). Building Conceptual Understandings in the Social Sciences. In *Learning Media Limited*. <https://ssol.tki.org.nz/Social-studies-1-10/Teaching-and-learning/Effective-teaching-in-social-studies/Building-conceptual-understandings>
- Chang, K. en y Sung, Y. T. y Chen, S. F. (2001). Learning through computer-based concept mapping with scaffolding aid. *Journal of Computer Assisted Learning*, 17(1), 21–33. <https://doi.org/10.1046/j.1365-2729.2001.00156.x>
- Chen, C. M. y Chen, M. C. (2009). Mobile formative assessment tool based on data mining techniques for supporting web-based learning. *Computers and Education*, 52(1), 256–273. <https://doi.org/10.1016/j.compedu.2008.08.005>
- Chen, G. y Wang, Y. y Guo, X. (2018). The impact of self-assessment on self-regulated learning and self-efficacy in e-learning environments: A meta-analysis. *Internet and Higher Education*, 36, 23–30.

- Chevron, M.-P. (2014). A metacognitive tool: Theoretical and operational analysis of skills exercised in structured concept maps. *Perspectives in Science*, 2(1–4), 46–54.
<https://doi.org/10.1016/j.pisc.2014.07.001>
- Chiheb, R. y Faizi, R. y el Afia, A. (2011). Using objective online testing tools to assess students' learning: Potentials and limitations. *Journal of Theoretical and Applied Information Technology*, 24(1), 69–72.
- Chou, P. y Chang, C. y Lin, C. (2017). BYOD or not : A comparison of two assessment strategies for student learning. *Computers in Human Behavior*, 74, 63–71.
<https://doi.org/10.1016/j.chb.2017.04.024>
- Clariana, R. (2010). Computer-Based Diagnostics and Systematic Analysis of Knowledge. In D. Ifenthaler, P. Pirnay-Dummer, & N. M. Seel (Eds.), *Computer-Based Diagnostics and Systematic Analysis of Knowledge*. Springer US. <https://doi.org/10.1007/978-1-4419-5662-0>
- Conrad, D. y Openo, J. (2018). *Assessment Strategies for On-line Learning: Engagement and Authenticity* (T. Anderson, Ed.; Issue October 10). AU PRESS.
<https://doi.org/10.15215/aupress/9781771992329.01>
- Cosi, S. y Voltas, N. y Lázaro-Cantabrana, J. L. y Morales, P. y Calvo, M. y Molina, S. y Quiroga, Á. (2020). Formative assessment at university using digital technology tools. *Revista de Currículum y Formación Del Profesorado*, 24(1), 164–183.
<https://doi.org/10.30827/profesorado.v24i1.9314>
- Costa, D. S. J. y Mullan, B. A. y Kothe, E. J. y Butow, P. (2010). A web-based formative assessment tool for Masters students: A pilot study. *Computers and Education*, 54(4), 1248–1253. <https://doi.org/10.1016/j.compedu.2009.11.011>

- Darling-Hammond, L. (2017). Developing and Measuring Higher Order Skills: Models for State Performance Assessment Systems. *The Council of Chief State School Officers (CCSSO)*, 3(1), 1–58. www.ccsso.org
- Darling-Hammond, L. y Schopp, M. y Minnich, C. (2017). *Developing and Measuring Higher Order Skills: Models for State Performance Assessment Systems THE COUNCIL OF CHIEF STATE SCHOOL OFFICERS.*
- Davis, M. A. y Curtis, M. B. y Tschetter, J. D. (2003). Evaluating cognitive training outcomes: Validity and utility of structural knowledge assessment. *Journal of Business and Psychology*, 18(2), 191–206. <https://doi.org/10.1023/A:1027397031207>
- De La Herrán, A. y Linares Álvaro, M. (2013). Mapas conceptuales y mentefactos: comparación y propuesta para favorecer aprendizajes significativos formativos. *Revista Iberoamericana de Educación*, 61(3), 1–16. <https://doi.org/10.35362/rie6131074>
- De Zubiria Samper, M. (2000). *Pedagogías del siglo XXI: Mentefactos I. El arte de pensar para enseñar y de enseñar para pensar* (Tercera). Fondo de Publicaciones Bernardo Herrera Merino.
- Del Carmen Gutierrez, L. y Martinez, P. S. B. y Reyes, A. M. F. y Mendez, V. N. P. (2021). The impact of the conceptual mind facts methodology on the academic performance of engineering students. *International Journal of Engineering Education*, 37(3), 841–852.
- De-Zubiria, M. (2015). Cómo enseñar promoviendo el pensamiento conceptual. In *Didácticas conceptuales 3*. Fundación Tecnológica Alberto Merani y Fundación Internacional de Pedagogía Conceptual Alberto Merani.

- Dijkstra, E. M. y van der Meijden, H. y van der Veen, J. T. (2014). Investigating the relationship between performance assessment, task complexity, and metacognitive monitoring. . *Assessment & Evaluation in Higher Education*, 39(4), 416–431.
- Dunlosky, J. y Lipko, A. R. (2007). Metacomprehension: A Brief History and How to Improve Its Accuracy. In *Psychological Science* (Vol. 16, Issue 4).
- Dunn, K. E. y Mulvenon, S. W. (2009). Dunn, K. E., & Mulvenon, S. W. (2009). A Critical Review of Research on Formative Assessment: The Limited Scientific Evidence of the Impact of Formative Assessment in Education, 14(7). Retrieved from <https://pareonline.net/pdf/v14n7.pdf> A Critical Review o. *Practical Assessment, Research & Evaluation*, 14(7). <https://pareonline.net/pdf/v14n7.pdf>
- Dutke, S. y Barenberg, J. y Leopold, C. (2010). Learning from text: Knowing the test format enhanced metacognitive monitoring. *Metacognition and Learning*, 5(2), 195–206. <https://doi.org/10.1007/s11409-010-9057-1>
- Ellis, D. y Zimmerman, B. (2001). Enhancing self-monitoring during self-regulated learning of speech role of self-monitoring during self-regulated learning. In H. J. Hartman (Ed.), *Metacognition in Learning and Instruction* (pp. 205–228). Kluwer Academic Publishers.
- Faber, J. M. y Luyten, H. y Visscher, A. J. (2017). The effects of a digital formative assessment tool on mathematics achievement and student motivation: Results of a randomized experiment. *Computers and Education*, 106, 83–96. <https://doi.org/10.1016/j.compedu.2016.12.001>

- Faber, J. M. y Visscher, A. J. (2018). The effects of a digital formative assessment tool on spelling achievement: Results of a randomized experiment. *Computers and Education*, 122(March), 1–8. <https://doi.org/10.1016/j.compedu.2018.03.008>
- Farrell, T. y Rushby, N. (2016). Assessment and learning technologies : An overview. *British Journal of Educational Technology*, 47(1), 106–120. <https://doi.org/10.1111/bjet.12348>
- Febriani, I. y Irsyad Abdullah, M. (2018). A systematic review of formative assessment tools in the blended learning environment. *International Journal of Engineering and Technology(UAE)*, 7(4), 33–39. <https://doi.org/10.14419/ijet.v7i4.11.20684>
- Fernández Huerta, J. (1959). Medidas sencillas de lecturabilidad. *Revista Pedagógica de La Sección Femenina de Falange ET y de Las JONS*, 214, 29–32.
- Fernández-Serrano, M. J. y Pérez-García, M. y Perales, J. C. y Verdejo-García, A. (2010). Prevalence of executive dysfunction in cocaine, heroin and alcohol users enrolled in therapeutic communities. *European Journal of Pharmacology*, 626(1), 104–112. <https://doi.org/10.1016/j.ejphar.2009.10.019>
- Flores, E. P. y Nolasco, A. y Da, B. y Rogoski, N. (2020). Narrative Comprehension: Concept Analysis and a Methodological Proposal. *Psicologia: Teoria e Pesquisa*, 2–30. <https://doi.org/10.31234/osf.io/6edtv>
- Gao, J. y Li, X. y Luo, L. (2018). Online quizzes as a tool for teaching and learning: A quantitative study of their effectiveness in higher education courses. *International Journal of Emerging Technologies in Learning (IJET)*, 13(8), 105–118.
- Gedye, S. (2010). Formative assessment and feedback: a review. *Planet*, 23(1), 40–45. <https://doi.org/10.11120/plan.2010.00230040>

- Gikandi, J. W. y Morrow, D. y Davis, N. E. (2011). Online formative assessment in higher education: A review of the literature. *Computers and Education*, 57, 2333–2351.
<https://doi.org/10.1016/j.compedu.2011.06.004>
- Goldstein, E. B. (2008). *Cognitive psychology: Connecting mind, research, and everyday experience* Wadsworth. Belmont.
- Gouli, E. y Gogoulou, A. y Grigoriadou, M. (2003). A Coherent and Integrated Framework Using Concept Maps for Various Educational Assessment Functions. *Journal of Information Technology Education*, 2, 26.
- Gutierrez, A. P. (2021). Calibrating Calibration: A Meta-Analysis of Learning Strategy Instruction Interventions to Improve Metacognitive Monitoring Accuracy. *Journal of Educational Psychology*, 114(4), 681–700. <https://doi.org/10.1037/edu0000674>
- Gutierrez de Blume, A. P. (2021). Calibrating Calibration: A Meta-Analysis of Learning Strategy Instruction Interventions to Improve Metacognitive Monitoring Accuracy. *Journal of Educational Psychology*, 114(4), 681–700. <https://doi.org/10.1037/edu0000674>
- Hacker, D. J. y Bol, L. y Bahbahani, K. (2008). Explaining calibration accuracy in classroom contexts: The effects of incentives, reflection, and explanatory style. *Metacognition and Learning*, 3(2), 101–121. <https://doi.org/10.1007/s11409-008-9021-5>
- Hacker, D. J. y Dunlosky, J. y Graesser, A. C. (2009). *Handbook of Metacognition in Education*. Routledge.
- Hall, C. C. y Ariss, L. y Todorov, A. (2007). The illusion of knowledge: When more information reduces accuracy and increases confidence. *Organizational Behavior and Human Decision Processes*, 103(2), 277–290. <https://doi.org/10.1016/j.obhdp.2007.01.003>

- Harper, M. E. y Jentsch, F. G. y Berry, D. y Cathy Lau, H. y Bowers, C. y Salas, E. (2003). TPL-KATS-card sort: A tool for assessing structural knowledge. *Behavior Research Methods, Instruments, and Computers*, 35(4), 577–584. <https://doi.org/10.3758/BF03195536>
- Hattie, J. (2015). The applicability of Visible Learning to higher education. *Scholarship of Teaching and Learning in Psychology*, 1(1), 79–91. <https://doi.org/10.1037/stl0000021>
- Henrie, C. R. y Halverson, L. R. y Graham, C. R. (2015). Measuring student engagement in technology-mediated learning: A review. *Computers and Education*, 90, 36–53. <https://doi.org/10.1016/j.compedu.2015.09.005>
- Hess, K. K. y Jones, B. S. y Carlock, D. y Walkup, J. R. (2009). Cognitive Rigor: Blending the Strengths of Bloom’s Taxonomy and Webb’s Depth of Knowledge to Enhance Classroom-level Processes. *Knowledge Creation Diffusion Utilization*, 1–8. http://www.standardsco.com/PDF/Cognitive_Rigor_Paper.pdf
- Hettiarachchi, E. y Huertas, M. A. y Guerrero, A.-E. (2015). Introducing a Formative E-Assessment System to Improve Online Learning Experience and Performance Hybrid Intensional Logic View project TeSLA (An Adaptive Trust-based e-assessment System for Learning) View project. In *Article in Journal of Universal Computer Science*. <https://www.researchgate.net/publication/281269991>
- Hettiarachchi, E. y Mor, E. y Huertas, M. A. y Guerrero-Roldán, A.-E. (2015). Introducing a Formative E-Assessment System to Improve Online Learning Experience and Performance. *Journal of Universal Computer Science*, 21(8), 2–22. <https://www.researchgate.net/publication/281269991>

- Hoidn, S. (2016). The Pedagogical Concept of Student-Centred Learning in the Context of European Higher Education Reforms. *European Scientific Journal, ESJ*, 12(28), 439. <https://doi.org/10.19044/esj.2016.v12n28p439>
- Hoon, J. y Lee, H. y Kwon, O. (2020). Effects of reflective writing on metacognition and learning outcomes in e-learning. . *Journal of Educational Computing Research*, 58(3), 497–519.
- Huang, Y. y Tsapali, M. (2022). The Relationship Between Students' Metacognition and Graphic Organisers: A Scoping Review. *Research Square*, 1–22. <https://doi.org/10.21203/rs.3.rs-2120141/v1>
- Ibañez Pinilla, M. (2006). *Mentefactos conceptuales como estrategia didáctico- pedagógica de los conceptos básicos de la teoría de muestreo aplicados en investigación en salud*. 4, 62–72.
- Ibarra-Sáiz, M. S. y Rodríguez-Gómez, G. (2020). Evaluating assessment. Validation with PLS-SEM of ATAE scale for the analysis of assessment tasks. *RELIEVE - Revista Electronica de Investigacion y Evaluacion Educativa*, 26(1), 1–18. <https://doi.org/10.7203/relieve.26.1.17403>
- Jaggars, S. S. y Xu, D. (2016). How do online course design features influence student performance? *Computers and Education*, 95, 270–284. <https://doi.org/10.1016/j.compedu.2016.01.014>
- Joseph, C. y Conradsson, D. y Nilsson Wikmar, L. y Rowe, M. (2017). Structured feedback on students' concept maps: The proverbial path to learning? *BMC Medical Education*, 17(1), 1–9. <https://doi.org/10.1186/s12909-017-0930-3>
- Kruger, D. y Inman, S. y Ding, Z. y Kang, Y. y Kuna, P. y Liu, Y. y Lu, X. y Oro, S. y Wang, Y. (2015). Improving Teacher Effectiveness: Designing Better Assessment Tools in Learning Management Systems. *Future Internet*, 7, 484–499. <https://doi.org/10.3390/fi7040484>

- Lai, J. W. M. y Bower, M. (2019). How is the use of technology in education evaluated? A systematic review. *Computers and Education*, 133, 27–42.
<https://doi.org/10.1016/j.compedu.2019.01.010>
- Lai, K. W. y Bower, M. (2019). Examining the potential of e-assessment for promoting higher order thinking: A review of the literature. . *Computers & Education*, 133, 1–15.
- Lane, R. y Parrila, R. y Bower, M. y Bull, R. y Cavanagh, M. y Forbes, A. y Jones, T. y Leaper, D. y Khosronejad, M. y Pellicano, L. y Powell, S. y Ryan, M. y Skrebneva, I. (2019). *Formative Assessment Evidence and Practice Literature Review*. (pp. 1–174). AITSL.
- Leary, M. O. y Scully, D. y Karakolidis, A. y Pitsia, V. (2018). The state-of-the-art in digital technology-based assessment. *European Journal of Education*, 53(2), 160–175.
<https://doi.org/https://doi.org/10.1111/ejed.12271>
- Lee, H. y Chung, H. Q. y Zhang, Y. y Abedi, J. y Warschauer, M. (2020). The Effectiveness and Features of Formative Assessment in US K-12 Education: A Systematic Review. *Applied Measurement in Education*, 33(2), 124–140. <https://doi.org/10.1080/08957347.2020.1732383>
- Lee, H. S. y Pallant, A. y Pryputniewicz, S. y Lord, T. y Mulholland, M. y Liu, O. L. (2019). Automated text scoring and real-time adjustable feedback: Supporting revision of scientific arguments involving uncertainty. *Science Education*, 103(3), 590–622.
<https://doi.org/10.1002/sce.21504>
- Leenknecht, M. y Wijnia, L. y Köhler, M. y Fryer, L. y Rikers, R. y Loyens, S. (2020). Formative assessment as practice: the role of students' motivation. *Assessment and Evaluation in Higher Education*, 1–20. <https://doi.org/10.1080/02602938.2020.1765228>

- Li, J. y Zhang, M. y Zhao, J. (2020). A review of formative assessment practices in higher education: Opportunities for technology enhancement. *Computers & Education*, 155(104002).
- Lopes, A. P. (2014). Learning Management Systems in Higher Education. *EDULEARN 14 Conference, July*, 5360–5365. https://doi.org/10.1007/978-3-030-49108-6_23
- Mao, L. y Liu, O. L. y Roohr, K. y Belur, V. y Mulholland, M. y Lee, H. S. y Pallant, A. (2018). Validation of Automated Scoring for a Formative Assessment that Employs Scientific Argumentation. *Educational Assessment*, 23(2), 121–138.
<https://doi.org/10.1080/10627197.2018.1427570>
- Matzner, M. (2008). Construction of a structural knowledge test. *University of Twente*, 1–19.
http://essay.utwente.nl/58833/%5Cnhttp://essay.utwente.nl/58833/1/scriptie_M_Matzner.pdf
- McCallum, S. y Milner, M. M. (2020). The effectiveness of formative assessment: student views and staff reflections. *Assessment and Evaluation in Higher Education*, 1–16.
<https://doi.org/10.1080/02602938.2020.1754761>
- McLaughlin, T. y Yan, Z. (2017). Diverse delivery methods and strong psychological benefits: A review of online formative assessment. *Journal of Computer Assisted Learning*, 33, 562–574.
<https://doi.org/10.1111/jcal.12200>
- Mendoza, L. y Huarachi, G. (2008). Los Mentefactos Conceptuales como Estrategia Metacognitiva en el logro de los Aprendizajes Significativos de los Alumnos en la Asignatura de Corrientes Pedagógicas Contemporáneas. *Ciencia y Desarrollo*, 17–22.
- Mera Cabezas, L. A. y Viñan Ajila, G. L. (2023). Las estrategias metodológicas: Mapa conceptual y mentefacto en el aprendizaje de Ciencias Naturales con estudiantes de décimo año de Educación General Básica de la Unidad Educativa Gonzol. In *Repositorio*

Digital Universidad Nacional de Chimborazo.

<http://dspace.unach.edu.ec/handle/51000/11211>

- Minte Münzenmayer, A. y Sepúlveda Obreque, H. y Jaramillo Alvarado, R. y Díaz-Levicoy, D. (2021). Evaluación en Educación Superior: características y demandas cognitivas de preguntas escritas. *Revista de Estudios y Experiencias En Educación*, 20(44), 43–52.
<https://doi.org/10.21703/0718-5162.v20.n43.2021.003>
- Mohamed, R. y Lebar, O. (2017). Authentic Assessment in Assessing Higher Order Thinking Skills. *International Journal of Academic Research in Business and Social Sciences*, 7(2), 466–476. <https://doi.org/10.6007/IJARBSS/v7-i2/2021>
- Moreno, J. y Pineda, A. F. (2020). A Framework for Automated Formative Assessment in Mathematics Courses. *IEEE Access*, 8, 30152–30159.
<https://doi.org/10.1109/ACCESS.2020.2973026>
- Morris, R. y Perry, T. y Wardle, L. (2021). Formative assessment and feedback for learning in higher education: A systematic review. In *Review of Education* (Vol. 9, Issue 3). John Wiley and Sons Inc. <https://doi.org/10.1002/rev3.3292>
- Mucha Hospinal, L. F. y Lora Loza, M. G. y Chamorro-Mejía, R. y Aliaga Salguerán, J. J. y Raiser Vásquez Ramírez, M. y Cifuentes, C. A. M. (2022). Mentefactos conceptuales como estrategia para el aprendizaje de Estadística Inferencial en estudiantes universitarios. *Editorial Ciencias Médicas*, 36(2), 1–12. <https://orcid.org/0000-0002-1973-7497>
- Murray, J. W. (2014). Higher-order Thinking and Metacognition in the First-year Core-education Classroom: A case study in the use of color-coded drafts. *Open Review of Educational Research*, 1(1), 56–69. <https://doi.org/10.1080/23265507.2014.964297>

- Nagandla, K. y Sulaiha, S. y Nalliah, S. V. (2018). Online formative assessments: exploring their educational value. In *Original Article Journal of Advances in Medical Education & Professionalism* (Vol. 6, Issue 2).
- Nesbit, J. C. y Adesope, O. O. (2006). Learning With Concept and Knowledge Maps : A Meta-Analysis. *Review of Educational Research*, 76(3), 413–448.
<https://doi.org/https://doi.org/10.3102/00346543076003413>
- Nietfeld, J. L. y Cao, L. y Osborne, J. W. (2005). Metacognitive monitoring accuracy and student performance in the postsecondary classroom. In *Article in The Journal of Experimental Education*. <https://www.researchgate.net/publication/284066045>
- Nipper, S. y Holmberg, C. (2003). *Online Education. Learning Management Systems* (1st ed.). NKI Forlaget.
- Novak, J. D. (1990). Concept maps and Vee diagrams: two metacognitive tools to facilitate meaningful learning. . *Instructional Science*, 19(1), 29–52.
- Novak, J. D. y Anderson, O. R. (n.d.). *Implications of Parallels in Ausubelian Ideas of Meaningful Learning, Concept Mapping, and Recent Studies in Neurobiology, Especially as Related to Learning in Science and Cognate Disciplines*. 1–14.
- Novak, J. D. y Cañas, A. J. (2007). Theoretical Origins of Concept Maps, How to Construct Them, and Uses in Education. *Reflecting Education*, 3(1), 29–42.
- Novak, J. D. y Gowin, D. B. (1984). *Learning how to learn*. (Cambridge University Press, Ed.).
- Ogange, B. O. y Agak, J. y Okelo, K. O. y Kiprotich, P. (2018). Student Perceptions of the Effectiveness of Formative Assessment in an Online Learning Environment. *Open Praxis*, 10(1), 29. <https://doi.org/10.5944/openpraxis.10.1.705>

- Olaya Torres, A. J. y Puente Ferreras, A. y Montoya Londoño, D. M. y Gutierrez de Blume, A. (2023). Metacognitive Judgments as an Emerging Research Trend. A Conceptual Review. *Ánfora*, 30(54), 254–281. <https://doi.org/10.30854/anf.v30.n54.2023.910>
- Oldfield, A. y Broadfoot, P. y Sutherland, R. y Timmis, S. (2012). Assessment in a Digital Age: A research review. In *Technology Enhanced Assessment: Review of the Literature* (pp. 1–40). University of Bristol. <http://www.bristol.ac.uk/media-library/sites/education/documents/researchreview.pdf>
- Orhan Göksün, D. y Gürsoy, G. (2019). Comparing success and engagement in gamified learning experiences via Kahoot and Quizizz. *Computers & Education*, 135(February), 15–29. <https://doi.org/10.1016/j.compedu.2019.02.015>
- Ozturk, N. (2017). Assessing Metacognition: Theory and Practices. *International Journal of Assessment Tools in Education*, 4(2), 134–148. <https://doi.org/10.21449/ijate.298299>
- Pellegrino, J. W. (2005). Assessment and accountability: An overview. . *Educational Technology Research and Development*, 53(4), 5–15.
- Pinchok, N. y Brandt, W. C. (2009). *Connecting Formative Assessment Research to Practice* (p. 22). Learning Point Associates. <https://doi.org/10.1002/9780470741276.ch14>
- Prinz, A. y Golke, S. y Wittwer, J. (2020). How accurately can learners discriminate their comprehension of texts? A comprehensive meta-analysis on relative metacomprehension accuracy and influencing factors. *Educational Research Review*, 31, 100358. <https://doi.org/10.1016/J.EDUREV.2020.100358>

- Purnamawati, P. y Saliruddin, S. (2017). The effectiveness of the use of metacognition-based industrial electronic learning tools in growing higher order thinking skills (HOTS). *Jurnal Pendidikan Vokasi*, 7(2), 139. <https://doi.org/10.21831/jpv.v7i2.13447>
- Quevedo, D. I. (2018). *Mapas y Mentefactos Conceptuales como Herramientas para el Aprendizaje Significativo en la Enseñanza de la Historia*. Universidad Distrital Francisco José de Cladas.
- Ragupathi, K. (2016). *Designing Effective Online Assessments*.
- Reimann, P. y Schilcher, A. (2013). Does computer-based assessment enhance learning? A comparison of computer-based and traditional assessment methods for undergraduate students. *British Journal of Educational Technology*, 44(4), 571–584.
- Richardson, M. y Abraham, C. y Bond, R. (2012). Psychological correlates of university students' academic performance: A systematic review and meta-analysis. *Psychological Bulletin*, 138(2), 353–387. <https://doi.org/10.1037/a0026838>
- Riveros Melo, N. A. (2017). *La estructuración conceptual de los niveles de organización en los sistemas vivos mediante Mentefactos* [Maestría]. Universidad Nacional de Colombia.
- Rivers, M. L. y Dunlosky, J. y Persky, A. M. (2020). Measuring metacognitive knowledge, monitoring, and control in the pharmacy classroom and experiential settings. *American Journal of Pharmaceutical Education*, 84(5), 549–560. <https://doi.org/10.5688/ajpe7730>
- Rolim, C. y Isaias, P. (2018). Examining the use of e-assessment in higher education: teachers and students' viewpoints. *British Journal of Educational Technology*. <https://doi.org/10.1111/bjet.12669>

- Ropohl, M. y Rönnebeck, S. (2019). Making learning effective—quantity and quality of pre-service teachers' feedback. *International Journal of Science Education*, 41(15), 2156–2176.
<https://doi.org/10.1080/09500693.2019.1663452>
- Schraw, G. (2009). A conceptual analysis of five measures of metacognitive monitoring. *Metacognition and Learning*, 4(1), 33–45. <https://doi.org/10.1007/s11409-008-9031-3>
- Schraw, G. y Sperling, R. (1994). Assessing metacognitive awareness. *Contemporary Educational Psychology*, 19, 460–475. <https://doi.org/https://doi.org/10.1006/ceps.1994.1033>
- Schroeder, N. L. y Nesbit, J. C. y Anguiano, C. J. y Adesope, O. O. (2018). Studying and Constructing Concept Maps: a Meta-Analysis. *Educational Psychology Review*, 30(2), 431–455. <https://doi.org/10.1007/s10648-017-9403-9>
- Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). Experimental and quasi-experimental designs for generalized causal inference. Houghton Mifflin Company.
- Singhal, Dr. D. (2017). Understanding Student- Centered Learning and Philosophies of Teaching Practices. *International Journal of Scientific Research and Management*, February.
<https://doi.org/10.18535/ijstrm/v5i2.02>
- Solomon, K. O. y Medin, D. L. y Lynch, E. B. (1999). Concepts Do More Than Categorize. *Trends in Cognitive Science*, 3(3), 11–105.
- Sudakova, N. E. y Savina, T. N. y Masalimova, A. R. y Mikhaylovsky, M. N. y Karandeeva, L. G. y Zhdanov, S. P. (2022). Online Formative Assessment in Higher Education: Bibliometric Analysis. In *Education Sciences* (Vol. 12, Issue 3). MDPI.
<https://doi.org/10.3390/educsci12030209>

- Taherdoost, H. (2018a). Sampling Methods in Research Methodology; How to Choose a Sampling Technique for Research. *SSRN Electronic Journal*, September.
<https://doi.org/10.2139/ssrn.3205035>
- Taherdoost, H. (2018b). Validity and Reliability of the Research Instrument; How to Test the Validation of a Questionnaire/Survey in a Research. *International Journal of Academic Research Management*, 5(3), 28–36. <https://doi.org/10.2139/ssrn.3205040>
- Tawfik, Andrew. A. (2020). Electronic formative assessment in higher education: A review of the literature. . *Journal of Educational Computing Research*, 58(1), 1–27.
- Timmis, S. y Broadfoot, P. y Sutherland, R. y Oldfield, A. (2015). Rethinking assessment in a digital age: opportunities, challenges and risks. *British Educational Research Journal*, 1–23.
<https://doi.org/10.1002/berj.3215>
- Tobias, S. y Everson, H. T. (2009). The importance of knowing what you know: A knowledge monitoring framework for studying metacognition in education. . In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Handbook of metacognition in education* (pp. 107–127). Taylor & Francis Group.
- Tomanek, D. y Talanquer, V. y Novodvorsky, I. (2008). What do science teachers consider when selecting formative assessment tasks? *Journal of Research in Science Teaching*, 45(10), 1113–1130. <https://doi.org/10.1002/tea.20247>
- Torres Alférez, B. M. (2016). *Aplicación de mentefactos conceptuales para mejorar los niveles de aprendizajes de los estudiantes en el área de ciencia, tecnología y ambiente del cuarto grado de secundaria en la institución educativa Champagnat - Tacna* [Maestría en Tecnología Educativa]. Universidad Nacional Jorge Basadre Grohmann.

- Trumpower, D. L. y Vanapalli, A. S. (2016). Structural Assessment of Knowledge as, of and for Learning. *Learning, Design, and Technology*, 1–22. <https://doi.org/10.1007/978-3-319-17727-4>
- Turnbull, D. y Chugh, R. y Luck, J. (2019). Learning Management Systems: An Overview. *Encyclopedia of Education and Information Technologies*, 0–7. <https://doi.org/10.1007/978-3-319-60013-0>
- Villarroel, V. y Bruna, D. (2019). ¿Evaluamos lo que realmente importa? El desafío de la evaluación auténtica en educación superior. *Calidad En La Educación*, 50, 492–509. <https://doi.org/https://doi.org/10.31619/caledu.n50.729>
- Wilson, M. y Korn, J. H. (2007). Attention during multiple-choice testing. *Applied Cognitive Psychology*, 21(5), 615–625.
- Winkley, J. (2010). E-assessment and innovation. *Becta*, 1–22.
- Zainuddin, Z. y Shujahat, M. y Haruna, H. y Chu, S. K. W. (2020). The role of gamified e-quizzes on student learning and engagement: An interactive gamification solution for a formative assessment system. *Computers and Education*, 145, 1–15. <https://doi.org/10.1016/j.compedu.2019.103729>
- Zarestky, J. y Bigler, M. y Brazile, M. y Lopes, T. y Bangerth, W. (2022). Reflective Writing Supports Metacognition and Self-regulation in Graduate Computational Science and Engineering. *Computers and Education Open*, 3, 100085. <https://doi.org/10.1016/j.caeo.2022.100085>

- Zhang, Y. y Li, L. y Liu, Y. y Guo, X. (2021). An investigation of the relationship between self-assessment and academic performance in undergraduate students: The mediating role of metacognitive monitoring. . *Frontiers in Psychology, 12*(766831).
- Zhao, L. y Ye, C. (2020). Time and Performance in Online Learning: Applying the Theoretical Perspective of Metacognition. *Decision Sciences Journal of Innovative Education, 18*.
<https://doi.org/10.1111/dsji.12216>
- Zhu, M. y Liu, O. L. y Lee, H. S. (2020). The effect of automated feedback on revision behavior and learning gains in formative assessment of scientific argument writing. *Computers and Education, 143*, 1–15. <https://doi.org/10.1016/j.compedu.2019.103668>
- Zohar, A. y Barzilai, S. (2015). Metacognition and teaching higher order thinking (HOT) in science education: Students' learning, teachers' knowledge and instructional practices. *The Routledge International Handbook of Research on Teaching Thinking, June, 229–242*.

Anexos

Anexo 1: Rúbrica de comprensión

No.	Indicador	10 puntos	5 puntos	0 puntos
1	Supraordinada	Nombra a los Leucocitos o Glóbulos blancos como categoría supraordinada de los granulocitos.	Reconoce que los granulocitos son un subtipo de célula sanguínea pero no nombra cuál.	No enuncia a qué categoría pertenecen los granulocitos.
2	Exclusora	Nombra que los agranulocitos también pertenecen a los leucocitos o glóbulos blancos.	Enuncia que existen distintos tipos de leucocitos distintos a los granulocitos, pero no los nombra.	No reconoce la existencia de tipos de leucocitos distintos a los agranulocitos.
3	Infraordinadas	Explicita que los granulocitos se dividen en 3 clases o grupos.	Explicita que los granulocitos se subdividen, pero no establece en cuántas categorías.	No explicita que los granulocitos tienen subtipos.
4	Diferencias entre concepto central y exclusoras	Establece que los agranulocitos se diferencian de los granulocitos en por lo menos una de dos: a) el tipo de defensa que tienen: específica o no específica. b) su estructura física: (núcleo o volumen en la sangre)	Explicita que los granulocitos y agranulocitos se diferencian en su defensa o estructura sin especificar el tipo de defensa o estructura.	No especifica que los granulocitos son distintos a los agranulocitos.
5	Isoordinadas de las exclusoras	Indica por lo menos una: 1. Los agranulocitos tienen núcleo no segmentado o monolobulado. 2. Conforman alrededor del 35% de los leucocitos.	Nombra otra característica distinta a las dos anteriores.	No adjudica ninguna característica a los agranulocitos.
6	Función del concepto central	Establece que los granulocitos defienden al organismo de forma inespecífica	Establece que los granulocitos defienden al organismo de agentes extraños, de	No establece ni la función de los granulocitos ni su forma de defensa.

			enfermedades, entre otras.	
7	Isoordinadas del concepto central	Indica una de dos: 1. Los granulocitos tienen núcleo multilobulado. 2. Conforman alrededor del 65% de leucocitos.	Resalta características sobre la presencia de gránulos en los granulocitos.	No resalta ninguna característica de los granulocitos.
8	Isoordinadas de las Infraordinadas.	Indica que hay un subtipo de granulocito que actúa en respuestas inflamatorias o usando histamina	Nombra algún dato morfológico o funcional sobre los basófilos.	No nombra ni describe las características de los basófilos.
9	Isoordinadas de las Infraordinadas.	Indica que hay un subtipo que actúan de forma rápida ante agentes extraños	Nombra algún dato morfológico o funcional sobre los neutrófilos.	No nombra ni describe las características de los neutrófilos.
10	Isoordinadas de las Infraordinadas.	Indica que hay un subtipo que se encarga de actuar ante parásitos o que modula respuestas alérgicas	Nombra algún dato morfológico o funcional sobre los eosinófilos.	No nombra ni describe las características de los eosinófilos.

Anexo 2: Lectura Granulocitos

Los leucocitos o glóbulos blancos, son células que hacen parte del sistema inmune del organismo y se encargan de defenderlo de microorganismos invasores y enfermedades infecciosas. Están presentes en la sangre y en el sistema linfático y conforman aproximadamente el 1% del volumen total de la sangre. Se clasifican en granulocitos o agranulocitos según su granularidad y la forma de su núcleo. Los granulocitos se caracterizan por la presencia de diversos tipos de gránulos en su citoplasma (primarios, secundarios y terciarios) y un núcleo segmentado. Se subdividen en neutrófilos, eosinófilos y basófilos y conforman aproximadamente el 65% del total de leucocitos. Los agranulocitos, son glóbulos blancos mononucleares que carecen de gránulos secundarios y se subdividen en linfocitos y monocitos. Conforman aproximadamente el 35% del total de leucocitos.

Los granulocitos y monocitos protegen al organismo frente a los microorganismos invasores de forma inespecífica (acciones defensivas innatas que no requieren sensibilización previa de los invasores), principalmente mediante la ingestión por fagocitosis. Adicionalmente, los monocitos se convierten en macrófagos encargándose de eliminar restos de células muertas. Los linfocitos son las células efectoras de las respuestas inmunitarias específicas (requieren de acciones particulares para la sustancia invasora), generando reacciones destinadas a neutralizarlo o destruirlo. Hay tres tipos principales de linfocitos: las células B, las células T y las células NK (células agresoras naturales).

Los neutrófilos conforman aproximadamente el 62% del total de leucocitos, lo que los convierte en el subtipo más abundante de glóbulos blancos y poseen un núcleo multilobulado que puede tener entre dos y cinco lóbulos. Los neutrófilos son la primera línea de defensa del organismo y su función más importante es la respuesta rápida ante la invasión de agentes extraños, donde su objetivo principal es combatir bacterias y hongos. Para esto, usan un mecanismo llamado fagocitosis, donde por medio del contenido de sus gránulos logran la destrucción de los microorganismos invasores.

Los basófilos componen aproximadamente el 0.4% del total de glóbulos blancos y tienen un núcleo que puede ser bilobulado o trilobulado. Los basófilos juegan un papel muy importante en las respuestas inflamatorias, donde la liberación de histamina causa vasodilatación, contracción del músculo liso (en el árbol bronquial) y aumento de la permeabilidad de los vasos sanguíneos. La liberación masiva del contenido de sus gránulos puede causar un shock anafiláctico que puede llegar hasta la muerte si no es controlado. Cuando la inflamación es producida por alérgenos, la sustancia que provoca la reacción alérgica interacciona con una inmunoglobulina de tipo E, que da lugar a que la célula libere el contenido de sus gránulos.

Los eosinófilos componen aproximadamente el 3% del total de glóbulos blancos y su núcleo es bilobulado. Una de las principales funciones de los eosinófilos es la defensa inespecífica frente a parásitos. Para esto, los eosinófilos desgranulan su proteína básica mayor o proteína catiónica del eosinófilo en la superficie de los gusanos parásitos y los destruyen formando poros en sus cutículas, lo que facilita el acceso de agentes como superóxidos y peróxido de hidrógeno al interior del parásito. Por otro lado, tienen una función importante en la modulación de respuestas inflamatorias alérgicas.

Anexo 3: Consentimiento Informado



Universidad de Concepción
Departamento de Psicología

Efecto del Mentefacto Conceptual sobre la Metacognición y Comprensión Conceptual:
un Estudio Experimental en el contexto de Evaluación Formativa
Investigador Responsable: María Paula Baquero Vargas

Estimado (a) participante:

Ud. ha sido invitado/a participar en el estudio **Efecto del Mentefacto Conceptual sobre la Metacognición y Comprensión Conceptual: un Estudio Experimental en el contexto de Evaluación Formativa**, a cargo de **María Paula Baquero Vargas**, estudiante de Doctorado en Psicología de la Universidad de Concepción.

- A. *Propósito de la investigación:* El objetivo de esta investigación, desarrollada por el/la estudiante María Paula Baquero Vargas (Profesora guía Claudia Paz Pérez Salas) es analizar el efecto que tienen distintas actividades de aprendizaje online sobre las habilidades metacognitivas y la comprensión de estudiantes universitarios.
- B. *Descripción de su participación:* Si usted decide participar del estudio, se le pedirá que acepte su participación a través del presente consentimiento informado. Su participación consistirá en leer un texto con un contenido académico específico, para después realizar una actividad de refuerzo sobre el texto que leyó. Una vez termine la actividad, se le pedirá que responda dos cuestionarios sobre la actividad que realizó. En conjunto el experimento tendrá una duración aproximada de 40 minutos.
- C. *Posibles riesgos:* Un eventual riesgo podría ser que se sintiera incómodo/a mientras contesta este cuestionario. No obstante, usted es libre de dejar el estudio en cualquier momento, sin necesidad de dar ningún tipo de explicación.

- D. *Beneficios*: La información que usted aporte será de gran valor para la investigación respecto a la relación existente entre las variables de estudio: Estrategias de aprendizaje, metacognición y comprensión.
- E. *Confidencialidad y resguardo de la información*: Toda la información derivada de su participación será manejada con estricta confidencialidad. Sólo el equipo de investigación tendrá acceso a los datos por usted proporcionados. La información será resguardada según todos los requerimientos que las leyes chilenas explicitan (ley 20.120). Asimismo, tanto en el análisis como en la publicación y difusión científica de los resultados, no se identificará la identidad de ninguno de los/as participantes ni su respectiva organización, para así resguardar el anonimato. La información que entregue mediante su participación sólo será utilizada con fines científicos y relativos a esta investigación y no será usada con fines ajenos a los explícitamente expresados en este documento.
- F. *Voluntariedad*: La participación en esta investigación es absolutamente voluntaria y usted puede retirarse en cualquier momento del estudio, sin que ello tenga ninguna consecuencia.
- G. *Derechos del/de la participante*: Cualquier pregunta que yo quisiera hacer con relación a mi participación en este estudio será contestada por el/la investigador/a responsable María Paula Baquero Vargas en el correo electrónico: mapaulis84@gmail.com. Para cualquier duda que no me haya sido satisfactoriamente resuelta por el investigador responsable me podré dirigir al Presidente del Comité Ético Científico del Departamento de Psicología de la Universidad de Concepción, Dr. Mauricio Garrido, al correo electrónico mauriciogarrido@udec.cl. Además, se podrá comunicar con la Dra. Andrea Rodríguez T. Presidenta del Comité de Ética, Bioética y Bioseguridad de la Universidad de Concepción. Teléfono: (41) 2204302.

Después de haber recibido y comprendido la información de este documento y de haber podido aclarar todas mis dudas, otorgo el consentimiento para participar en el estudio: **“Efecto del Mentefacto Conceptual sobre la Metacognición y Comprensión Conceptual: un Estudio Experimental en el contexto de Evaluación Formativa”**.

Comprendo y acepto la información que se entregó anteriormente y declaro conocer los objetivos del estudio.

En atención a estas consideraciones, libremente marque la que corresponda.

Yo ACEPTO

Yo NO ACEPTO