



IMPACTO DEL ENTORNO DE RESOLUCIÓN DE PROBLEMAS DE FERMI EN LA COMPLEJIDAD Y FLEXIBILIDAD DE LAS ESTRATEGIAS DE MAESTROS EN FORMACIÓN

César Gallart 

Universitat de València
cesar.gallart@uv.es

Carlos Segura 

Universitat de València
carlos.segura@uv.es

Lluís Albarracín* 

Universitat Autònoma de Barcelona
lluis.albarracin@uab.cat

RESUMEN: Las competencias en resolución de problemas y de modelización son centrales en el currículo de Matemáticas. Los problemas de Fermi son idóneos para que los estudiantes desarrollen estas competencias, pues presentan situaciones reales y abiertas que promueven, por un lado, la inclusión de suposición realistas que aumentan la complejidad de las soluciones, y por otro, la flexibilidad en el uso de varias estrategias. Sin embargo, los docentes de Primaria no introducen estos problemas en sus aulas, y tienen dificultades para resolverlos satisfactoriamente. Nuestro propósito es conocer qué entorno de resolución es más eficaz para que los futuros maestros/as desarrollen estrategias complejas y las usen de manera flexible a lo largo de una secuencia de problemas de Fermi. Para ello, presentamos un estudio cuasi experimental con tres entornos de resolución: 55 futuros maestros resuelven problemas de Fermi en el aula, evocando la situación; 41 futuros maestros los resuelven experimentando en el lugar real de los problemas; y 41 futuros maestros los resuelven interrogando a ChatGPT. Desde un enfoque de métodos mixtos, analizamos las estrategias completadas, las suposiciones realistas, los tipos de estrategias utilizadas y la flexibilidad de los participantes en cada entorno. Una combinación de análisis de dependencia de variables y de su varianza nos permite

comparar la influencia del entorno de resolución en la tasa de estrategias completadas y su nivel de complejidad, así como en cuáles son más utilizadas y en el nivel de flexibilidad de los futuros maestros. Los resultados muestran que, aunque completar una estrategia es accesible para los futuros maestros con independencia del entorno de resolución, sí hay diferencias significativas en la complejidad de esas estrategias según el entorno y también en la flexibilidad. El entorno de experimentación *in situ* es el más eficaz para desarrollar estas habilidades.

PALABRAS CLAVE: resolución de problemas, modelización matemática, problemas de Fermi, entornos de resolución, ChatGPT.

IMPACT OF FERMI'S PROBLEM-SOLVING ENVIRONMENT ON THE COMPLEXITY AND FLEXIBILITY OF PRE-SERVICE TEACHERS' STRATEGIES

ABSTRACT: Problem-solving and modelling competencies are central to the Mathematics curriculum. Fermi problems are ideal for students to develop these competencies, as they present real and open situations that promote, on the one hand, the inclusion of realistic assumptions that increase the complexity of the solutions, and on the other hand, flexibility in the use of multiple strategies. However, primary school teachers do not introduce these problems in their classrooms and have difficulties solving them satisfactorily. We aim to find out which solving environment is most effective for pre-service teachers to develop complex strategies and use them flexibly throughout a sequence of Fermi problems. To this end, we present a quasi-experimental study with three solving environments: 55 pre-service teachers solve Fermi problems in the classroom, evoking the situation; 41 pre-service teachers solve them by experimenting at the real problem site; and 41 pre-service teachers solve them by questioning ChatGPT. Using a mixed methods approach, we analyse the strategies completed, the realistic assumptions, the types of strategy used and the participants' flexibility for each solving environment. A combination of variable dependence and variance analysis allows us to compare the influence of the solving environment on the rate of completed strategies and their level of complexity, as well as on which strategies are most used and the pre-service teachers' level of flexibility. The results show that, although completing a strategy is accessible to pre-service teachers regardless of the solving environment, there are significant differences in the complexity of these strategies according to the environment and also in the flexibility demonstrated. The on-site experimentation environment is the most effective for developing these skills.

KEYWORDS: Problem solving, mathematical modelling, Fermi problems, solving environments, ChatGPT.

Recibido: 19/12/2024

Aceptado: 19/03/2025

EXTENDED ABSTRACT

There is a consensus in mathematics education on the central role of problem solving in developing mathematical competence. In this sense, many educational programmes around the world emphasise the importance of solving modelling problems, as they make it possible to connect mathematics with the real world. In fact, in the Spanish mathematics curriculum, both problem solving and mathematical modelling are specific competencies.

Fermi problems pose a real situation, without data, in which the estimation of a quantity that cannot be obtained directly is required. These problems, especially when they require the estimation of large quantities of elements in a delimited area, are ideal for introducing mathematical modelling in primary school classrooms, as they can be solved by developing simple models in which to apply elementary knowledge of measurement and estimation of lengths and areas, as well as basic arithmetic calculations.

Moreover, since Fermi problems are open-ended, they can be solved using multiple strategies. Specifically, previous work has categorised three solution strategies for this type of problem: Linearisation (arranging the elements to be estimated in two dimensions, width and length), Base Unit (establishing the total area where the elements must be located, then dividing this area by the area occupied by a single element) and Density (estimating the elements within a specified portion of the given area). These strategies can be adapted to the contextual characteristics of the problem by introducing realistic assumptions (such as subtracting the space lost by obstacles or considering the heterogeneity of the sizes or distribution of the elements to be estimated). Consequently, using Fermi problem sequences that consist of estimating a large number of elements in a delimited area allows for promoting flexibility in students at this stage, namely, their ability to use multiple strategies, changing from one problem to another to adapt to their contextual characteristics.

Despite the potential of Fermi problems to develop modelling problem-solving skills in primary education, it is well known that primary school teachers hardly introduce these problems in their classrooms. In fact, several studies have found that there is a high proportion of pre-service teachers who have difficulties in solving them satisfactorily and who do not flexibly use multiple strategies. Teacher training needs to incorporate effective strategies to facilitate the successful introduction of Fermi problems into the classroom.

In this sense, certain problem-solving settings could facilitate the development of pre-service teachers' competencies, promoting their flexibility and the incorporation of realistic assumptions that enrich their strategies. Hence, several studies have explored the effect of outdoor (experimenting in the real place where the problems are located) and indoor (in the classroom, evocation from images) settings on the solution of modelling problems. In recent years, the emergence of AI in education suggests that we add a third problem-solving setting: indoor AI support (in this study, questioning ChatGPT).

Accordingly, this study aims to determine which of these three solving environments (outdoor experimentation, indoor ChatGPT interrogation, or indoor evocation) is most effective in enabling prospective teachers to develop complex strategies (incorporating realistic assumptions for greater estimation accuracy) and use them flexibly (switching strategies from one problem to another) across a sequence of Fermi problems.

To achieve this objective, an exploratory quasi-experimental study was carried out with a sample of pre-service teachers (N=137). For the experimentation, a pencil and paper questionnaire were used with a sequence of four Fermi problems located in the environment of the Faculty of Education (University of Valencia) and therefore familiar to all participants. These four problems consist of estimating a large number of elements in a rectangular area and have different contextual characteristics (differences in the size, distribution and shape of the elements to be estimated): number of students that fit in the porch of the faculty; number of tiles in a certain area; number of blades of grass in a small garden; and number of cars that can be parked in a lot. The participants were distributed into three groups according to the problem-solving setting: 55 students in the indoor evocation setting, 41 students in the outdoor experimentation setting, and 41 students in the ChatGPT interrogation setting. The dependent variables analysed were the rate of completed solutions, the strategies used, the inclusion of realistic assumptions in the strategy, and the flexibility in the use of these strategies throughout the problem sequence. The analysis was conducted using a mixed methods approach, combining analysis of variable dependence and analysis of variance.

The findings of the study show that Fermi problems are accessible to pre-service teachers regardless of the problem-solving setting, as in all of them most participants can develop complete solutions. However, this does not mean that these solution strategies are complex and best adapted to the contextual characteristics. Thus, we found that there are significant differences in terms of the inclusion of realistic assumptions and flexibility depending on the problem-solving setting. The outdoor experimental setting promotes more complex strategies (with the inclusion of realistic assumptions that help to obtain a more accurate estimate, especially those that consist of removing obstacles from the space to refine the data collection) than the other two settings. This is because, in the outdoor setting, the solver directly perceives relevant features of the real context –obstacles, irregularities– that impact on the demanded estimate and that may go unnoticed in the indoor settings. Although the images included in the problem statements show these obstacles, this has not been sufficient for them to be taken into account by the participants who have solved the problems indoors, with or without AI support.

On the other hand, although the three strategies (Density, Base Unit and Linearisation) are used in all three problem-solving settings, the results show that their use does depend on the setting. Thus, the use of the Base Unit strategy is significantly lower in the outdoor experimental environment, with the Density strategy being the most used in this environment. In fact, in the outdoor setting there is a frequent use of non-conventional measurement referents that allows participants to ‘concentrate’ the elements to be estimated within the referent, which leads naturally to the Density strategy. In contrast, in the other two settings, Base Unit was the most frequently used strategy. In the case of the ChatGPT setting, pre-experimental tests showed that the AI only uses Base Unit and Density, and more frequently the former, which could explain the results. However, participants in this environment also use the Linearisation strategy, suggesting that they are not always conditioned or constrained by the AI, but combine its use with their own experience.

Finally, we found significant differences in the flexibility demonstrated by pre-service teachers in the three problem-solving settings. Both the experimental and ChatGPT settings promote greater flexibility in the use of multiple strategies than the indoor evocation setting.

However, although participants in the outdoor experimental setting showed slightly greater flexibility than those in the ChatGPT setting, this difference was not statistically significant.

These findings highlight the importance of incorporating contextualised problem-solving experiences into teacher training, as well as the potential of digital tools such as ChatGPT to complement and diversify mathematical problem-solving strategies. Both the use of outdoor experimentation settings and AI-supported indoor settings can be used as an effective scaffold for primary school teachers to demonstrate flexibility in solving modelling problems and to develop more complex and accurate problem-solving strategies. Future research could further explore the impact of these settings in pre-service teacher training, for example, on their problem-solving performance or the self-efficacy they experience. Our results also provide an opportunity to further research on how pre-service teachers interact with ChatGPT in order to characterise usage profiles in solving modelling problems.

1. INTRODUCCIÓN

La resolución de problemas en contextos reales constituye un enfoque fundamental para promover el desarrollo de la competencia matemática en la educación primaria (Verschaffel et al., 2000). En los últimos años, se ha enfatizado la necesidad de conectar los contenidos matemáticos con situaciones del mundo real, no solo para aumentar la relevancia de la disciplina, sino también para desarrollar habilidades transferibles como son el razonamiento matemático y la toma de decisiones en situaciones complejas (Maaß et al., 2019). Estos problemas requieren que los estudiantes movilicen habilidades y conocimientos como la estimación de medidas, la identificación de variables relevantes, la gestión de múltiples estrategias y la toma de decisiones fundamentadas.

A pesar de estos beneficios, el uso de problemas en contextos reales también plantea retos significativos. Por ejemplo, la dificultad para conectar los conocimientos extramatemáticos con estrategias de resolución y para gestionar la incertidumbre asociada a la falta de datos (Galbraith y Stillman, 2001). Otro reto es garantizar que los docentes reciban la formación adecuada para gestionar situaciones en las que se empleen múltiples estrategias de resolución (Heinze et al., 2009).

Por tratarse de problemas matemáticos del mundo real accesibles para educación primaria, nos centraremos en el uso de problemas de Fermi, en los que se deben estimar cantidades a partir de romper el problema inicial en varios subproblemas (Ärlebäck, 2009). Dada su formulación abierta, sin datos, estos problemas permiten desarrollar múltiples estrategias (Segura y Ferrando, 2023). Sin embargo, estudios previos ponen de manifiesto que los futuros maestros tienen dificultades para resolver este tipo de problemas hasta el punto de no poder completar estrategias exitosas (Ferrando et al., 2020), y menos aún enriquecerlas para hacerlas más complejas con suposiciones realistas adecuadas (Segura et al., 2021). Por lo tanto,

es necesario investigar sobre entornos de resolución que mejoren la competencia de los futuros maestros para asegurar su implementación exitosa en las aulas.

En este sentido, Segura et al (2023) mostraron que la resolución de secuencias de problemas de Fermi por futuros maestros experimentando en el lugar real de los problemas es un entorno de resolución que favorece la adaptación de las estrategias a las características contextuales. Esto podría promover mayor flexibilidad en el uso de múltiples resoluciones que en un entorno de evocación en el aula (Segura y Ferrando, 2023). Además, la adaptación al contexto que favorece el entorno experimental podría tener influencia en el desarrollo de estrategias más complejas (Segura et al., 2021). Por otro lado, la irrupción de la Inteligencia Artificial (IA) en la Educación Matemática introduce un nuevo entorno de resolución de problemas que podría ser también efectivo en la mejora de estos aspectos (Pepin et al., 2025). En consecuencia, el propósito de este trabajo es analizar si estos entornos son eficaces para impulsar que los futuros maestros desarrollen estrategias complejas y las usen flexiblemente en una secuencia de problemas de Fermi.

2. MARCO TEÓRICO

2.1. Problemas de Fermi

Según Ärlebäck (2009), los problemas de Fermi son problemas abiertos que requieren que los estudiantes hagan suposiciones sobre la situación planteada y estimen las cantidades relevantes necesarias (p. 331). Una de sus particularidades radica en su formulación: estos problemas parecen difusos en su planteamiento y ofrecen poca información sobre cómo abordar su resolución (Efthimiou y Llewellyn, 2007). Según Sriraman y Knott (2009), los problemas de Fermi son problemas de estimación cuyo objetivo es que los estudiantes realicen conjeturas fundamentadas. Solo un análisis detallado de la situación permite descomponer el problema en subproblemas más simples y llegar a la solución (Carlson, 1997). Sin embargo, la forma en que se utilizan estos problemas en el aula va más allá de identificar cantidades relevantes y realizar estimaciones basadas en conocimientos previos. En su implementación se fomenta que los estudiantes desarrollen sus propias estrategias de resolución (Ferrando y Albarracín, 2021), obteniendo información del mundo real por sus propios métodos, ya sea a través de mediciones in situ o empleando herramientas digitales (Keune y Henning, 2003; Tangney y Bray, 2013).

El desafío que plantean los problemas de Fermi varía según la edad de los estudiantes, ya que utilizan su conocimiento de manera diferente en cada nivel educativo (Ferrando y Albarracín, 2021). Algunas investigaciones se han centrado en los procesos de resolución en estudiantes de primaria (Haberzettel et al., 2018; Toalongo et al., 2024), concluyendo que, en este nivel, ya pueden generar sus propias soluciones de diversas maneras y desarrollar nuevos conocimientos

matemáticos. En estudios previos (Ferrando y Albarracín, 2021; Montejo-Gámez et al., 2024) se analizaron y clasificaron las estrategias de resolución desarrolladas para resolver problemas de Fermi que requerían la estimación de grandes cantidades en áreas delimitadas, y son las siguientes:

- **Linealización:** es la disposición de los elementos a estimar en una cuadrícula, con las cantidades de estos elementos estimadas en función de cada dimensión (por ejemplo, altura y anchura).
- **Densidad:** el cálculo del número de elementos dentro de una porción especificada del área dada, identificada por el estudiante.
- **Unidad base:** consiste en establecer el área total donde deben ubicarse los elementos, para luego dividir esta área por la ocupada por un único elemento.

2.2. Complejidad de estrategias y flexibilidad de futuros maestros

Aunque los problemas de Fermi son potencialmente accesibles para estudiantes de distintos niveles educativos (Årlebäck, 2009), su resolución puede suponer un reto para los docentes de educación Primaria. Así, se ha encontrado que los futuros maestros/as tienen dificultades para completar las estrategias de resolución de problemas de Fermi (Ferrando et al., 2020) y cometen errores durante su desarrollo (Segura y Ferrando, 2023).

Además, se ha observado que las estrategias generadas por los futuros maestros son mayoritariamente simples y no se adaptan a características contextuales que puedan dificultar la estimación (Segura et al., 2021), como la presencia de obstáculos o de irregularidades. Cuando el resolutor percibe estas características contextuales, puede tratar de superarlas introduciendo adaptaciones a su estrategia. Estas adaptaciones, que denominamos factores de complejidad o suposiciones realistas (Segura et al., 2021), como restar el espacio perdido por obstáculos o considerar la heterogeneidad de los tamaños o distribución de los elementos a estimar, sirven para enriquecer la estrategia desarrollada a partir de hacerla más compleja y con la intención de obtener una estimación más precisa y rigurosa.

Una manera efectiva de fomentar el desarrollo y adaptación de estrategias en la formación de maestros es el uso de secuencias de problemas de Fermi (Ferrando et al., 2020; Segura y Ferrando, 2023). Su objetivo final es que los futuros maestros adquieran flexibilidad en el uso de múltiples estrategias, entendida como la capacidad de los individuos para elegir entre diferentes estrategias de solución al enfrentarse a una actividad matemática (Heinze et al., 2009). En el marco de la resolución de problemas se considera que el uso flexible de múltiples estrategias es esencial para construir un

conocimiento profundo, conectado y adaptado a las características del problema o del contexto (Heinze et al., 2009; Star y Rittle-Johnson, 2008).

Segura y Ferrando (2023) definen flexibilidad a lo largo de una secuencia de problemas de Fermi como la capacidad de cambiar de estrategia de un problema a otro según las características del contexto. En este trabajo se distinguen tres niveles: la alta flexibilidad supone usar varias estrategias en varios problemas, la flexibilidad moderada consiste en cambiar de estrategia sólo en un problema, mientras que la falta de flexibilidad es utilizar la misma estrategia durante toda la secuencia. Segura y Ferrando (2023) encontraron que los futuros maestros con alta flexibilidad tenían mayor competencia en resolución de problemas de Fermi, siendo capaces de completar sus estrategias sin errores. La falta de flexibilidad, por el contrario, se relaciona con dificultades para desarrollar las estrategias y con errores en su implementación, incluyendo carencias en medida y estimación de magnitudes (Segura et al., 2025). Por todo ello, se deduce que fomentar la flexibilidad en los futuros docentes podría ser una forma eficiente de mejorar su competencia en la resolución de problemas en contextos reales.

2.3. Entornos de resolución de problemas

En el desarrollo de estrategias más complejas y su uso flexible a través de secuencias de Fermi podrían tener impacto los entornos de resolución. Así, Jablonsky (2023) ha explorado el impacto en el proceso de resolución de problemas de estimación de tamaños de objetos reales en distintos entornos, in situ (outdoors), experimentando directamente sobre el propio objeto (en línea con las rutas matemáticas de Buchholtz, 2021), o evocando dicho objeto en el aula (indoors) con ayuda de imágenes o modelos 3D, obteniendo diferencias en la identificación de las características relevantes del objeto (simplificación y estructuración) y recogida de datos (matematización) entre ambos entornos. Así, Buchholtz (2021) rediseña el ciclo de modelización en las rutas matemáticas para poner el énfasis en el trabajo relacionado directamente con el objeto real (matematización contextualizada). En Segura et al. (2023), los futuros maestros diseñaron planes para resolver problemas de Fermi, en contextos cercanos y familiares, mediante su evocación en el aula. Se observó que, posteriormente, cuando estos mismos futuros maestros implementaban sus planes en un entorno de resolución experimentando en el lugar real de los problemas (outdoors), adaptaron sus estrategias a las características contextuales percibidas in situ, cambiando de estrategia si era necesario o incorporando suposiciones realistas. Hartmann y Schukajlow (2021) analizan el efecto de trabajar en estos entornos en aspectos afectivos, llegando a la conclusión de que es el problema y no el entorno lo que es importante para el desarrollo del interés y las emociones.

En los últimos años, a estos dos entornos de resolución (evocación en el aula, experimentación in situ), podemos añadirles un tercero: resolver problemas con la asistencia de la IA, una herramienta de enorme potencial educativo (Lo, 2023; Zhang y Aslam, 2021). Aunque la investigación sobre el uso de la IA en educación matemática es todavía escasa (Kuhail et al., 2023), se ha comprobado que es capaz de resolver algunos problemas matemáticos (Noster et al., 2024) y en particular, problemas de modelización matemática (Spreitzer et al., 2024) y de Fermi (López-Simó y Rezende, 2024). La interrogación a una IA, como ChatGPT, podría dar soporte a estimaciones para hacerlas verosímiles o ideas para suposiciones realistas, aunque este entorno de resolución en los problemas de Fermi apenas se ha investigado.

3. OBJETIVOS DEL ESTUDIO

El objetivo general del presente estudio es conocer qué entorno de resolución, de entre los que proponemos (experimental in situ, interrogativo con ChatGPT y evocación en el aula) es más eficaz para que los futuros maestros desarrollen estrategias complejas y las usen de manera flexible a lo largo de una secuencia de problemas de Fermi. Tras lo expuesto en los antecedentes, planteamos las siguientes preguntas de investigación:

- PI1. ¿Depende la capacidad de los futuros maestros para completar una estrategia del entorno de resolución?
- PI2. ¿Cuál es el efecto del entorno de resolución en la complejidad de las estrategias desarrolladas por los futuros maestros?
- PI3. ¿Dependen las estrategias utilizadas por los futuros maestros del entorno de resolución?
- PI4. ¿Cuál es el efecto del entorno de resolución en la flexibilidad de los futuros maestros?

4. METODOLOGÍA

Para abordar estas preguntas, hemos realizado un estudio cuasi-experimental, con carácter exploratorio, sobre una muestra de conveniencia formada por futuros maestros de educación primaria. La intervención consistió en fijar tres entornos de resolución (variable independiente) para estudiar su influencia en el desarrollo de las estrategias, su complejidad y la flexibilidad en su uso (variables dependientes).

4.1. Participantes

El estudio se ha realizado sobre una muestra de N=137 futuros maestros en formación (a partir de ahora participantes), todos ellos pertenecientes al tercer curso del Grado de Educación Primaria de la Facultad de Formación del Profesorado de la Universidad de Valencia (España). La edad media de los participantes es de 22,2 años, 45 son hombres y 92 mujeres. Los participantes estaban finalizando el segundo semestre del curso 2023-24, habiendo recibido formación en aritmética, geometría, medida, estadística y probabilidad. Antes de comenzar la experiencia se les explicó que los datos recogidos se utilizarían en el desarrollo de esta investigación, con su consentimiento informado.

4.2. Entornos de resolución

Se proponen tres entornos de resolución, experimental in situ; interrogación con ChatGPT y evocación en el aula, para estudiar su impacto en la capacidad de los participantes de completar una estrategia de resolución, su complejidad y la flexibilidad mostrada a lo largo de la secuencia de problemas propuestos.

Los participantes del entorno experimental resolvieron los problemas en el lugar donde estos se sitúan, pudiendo percibir sus características de manera directa, pero sin la ayuda de instrumentos de medida.

Los participantes del entorno con ChatGPT trabajaron en su aula, accediendo a la IA desde sus ordenadores (versión gratuita 3.5). Se les indicó expresamente que no podían utilizar motores de búsqueda como Google ni otras tecnologías. No recibieron ninguna instrucción previa sobre el uso de ChatGPT, más allá de su propia experiencia personal. Las interacciones con la IA, recogidas por los investigadores, no son objeto de análisis en este estudio, aunque permitieron comprobar cómo estos participantes obtenían sus datos.

Por último, los participantes del entorno evocación en el aula resolvieron los problemas sin ningún tipo de apoyo, reconstruyendo mentalmente los lugares de los problemas.

Los 137 participantes se asignaron aleatoriamente a uno de los tres entornos, de manera que estos quedaron con un número similar de alumnos. Así, se asignaron 41 participantes al entorno in situ; 41 participantes al entorno con ChatGPT; y 55 participantes al entorno en el aula. Los participantes trabajaron individualmente en su entorno, supervisados por dos de los investigadores, cuya única interacción con ellos era recordarles las condiciones de la experiencia.

4.3. Materiales

Para la experimentación, se prepara un cuestionario de lápiz y papel con una secuencia de cuatro problemas de Fermi situados en el entorno de la Facultad y por tanto accesible y familiar para todos los participantes. Estos cuatro problemas consisten en estimar una cantidad grande de elementos en una superficie rectangular. Su enunciado y características contextuales puede verse en la Tabla 1.

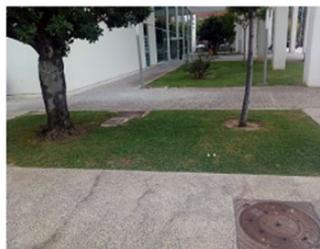
Tabla 1. *Secuencia de problemas de Fermi*

	Enunciado del problema	Características contextuales
P1	¿Cuánta gente se puede resguardar debajo del porche de entrada a la facultad si llueve?	Tamaño elementos: mediano Disposición elementos: desordenada Forma elementos: irregular
P2	¿Cuántas baldosas hay en el espacio entre el edificio de magisterio y el gimnasio?	Tamaño elementos: pequeño Disposición elementos: muy ordenada Forma elementos: muy regular
P3	¿Cuántas briznas de césped hay en este espacio?	Tamaño elementos: muy pequeño Disposición elementos: muy desordenada Forma elementos: irregular
P4	¿Cuántos coches caben en el parking de Magisterio?	Tamaño elementos: grande Disposición elementos: ordenada Forma elementos: regular

El enunciado de los problemas se acompaña de una imagen de su ubicación real y un espacio en blanco donde escribir su respuesta (ver Figura 1). El cuestionario se encabeza con la siguiente instrucción, “Explica, paso a paso, cómo los resolverías, los datos y procedimientos que utilizarías e intenta dar una solución numérica con este plan que describes”, instrucción que se repite verbalmente por los investigadores al comienzo de la experiencia.

Figura 1. *Enunciado y fotografía del problema P3 en el cuestionario*

Problema 3 ¿Cuántas briznas de hierba hay en el jardín de tu instituto?



Esta secuencia ha sido utilizada en investigaciones previas (Segura y Ferrando, 2023) y su objetivo principal es promover la flexibilidad: se busca que los participantes reconozcan que, según las características contextuales del problema (como el tamaño de los elementos, su regularidad o el orden en su distribución), puede ser más apropiado emplear una estrategia específica, mientras que en otros problemas, con características distintas, será más conveniente utilizar una diferente, a pesar de que todos pueden resolverse de la misma manera. Por ejemplo, en P3 una estrategia adecuada puede ser la Densidad, dado el tamaño de los elementos (muy pequeños) con distribución irregular, mientras que en P2, donde los elementos son de mayor tamaño y con una distribución regular, puede ser más adecuada la estrategia de Linealización (Ferrando et al., 2020). Conviene señalar que en las resoluciones previas realizadas por los investigadores con ChatGPT, la IA empleó únicamente dos estrategias: Unidad Base y Densidad.

4.4. Análisis de datos

La experiencia se llevó a cabo en una sesión de 60 minutos. Todos los participantes pudieron completar el cuestionario en el tiempo previsto, sin que en ningún entorno fuera necesario añadir tiempo adicional. Se utilizó un enfoque de métodos mixtos para el análisis de los datos. Así, las 548 resoluciones (137 participantes que resuelven, cada uno, cuatro problemas) fueron analizadas cualitativamente por dos de los investigadores, categorizándolas en incompletas y completas. Estas últimas se analizaron cualitativamente para clasificarlas en estrategias de resolución y para identificar la inclusión de suposiciones realistas. Como se resume en la Figura 8, estas categorías se cuantificaron en las variables nivel de complejidad y flexibilidad, para realizar un análisis cuantitativo de relación entre estas variables y la variable independiente (entorno de resolución). A continuación, detallamos las distintas fases del proceso de análisis de datos.

4.4.1. Categorización de las estrategias

En la primera fase del análisis, dos investigadores clasificaron las soluciones en dos categorías: completas e incompletas. Se consideraron completas aquellas que describen con suficiente claridad el proceso necesario para poder alcanzar la estimación demandada (sin tener en cuenta si se trata de una estimación precisa o adecuada, pues el rendimiento no es objeto de esta investigación). Las soluciones completas se codificaron, a su vez, según los tres tipos de estrategias descritos: Linealización, Unidad Base y Densidad. Los casos dudosos se acordaron entre los dos investigadores sin dificultades, ya que la fiabilidad y consistencia del proceso se ha constatado en estudios previos (Segura y Ferrando 2023; Segura et al., 2025).

En la Figura 2 podemos ver un ejemplo resolución completa mediante la estrategia de Unidad Base en P1; en la Figura 3, mediante la estrategia de Linealización en P2; y en la Figura 4, de Densidad en P3.

Figura 2. Estrategia completa de Unidad Base en P1

Si suponemos que el parche tiene una longitud de ~~50m~~ 40 m y una anchura de 5 m, el área del parche será de 50 m^2 . Sabiendo que el espacio medio de una persona es $0,5\text{ m}^2$ dividimos 50 m^2 entre $0,5\text{ m}^2$ y son 100 personas. Por tanto, se podrán resguardar 100 personas.

Figura 3. Estrategia completa de Linealización en P2

$$7 \text{ pasos} \times 6 \text{ baldosas (□)}$$

$$6 \text{ pasos} \times 5 \text{ baldosas (□)}$$

$$+ \frac{1}{2} \text{ paso baldosera} \rightarrow 2,5 \text{ baldosas (□)}$$

ancho
 $5,5 \times \text{paso aprox.}$

$$5 \text{ pasos} \times 5,5 \text{ baldosas} \times 10 = 27,5 \times 10 = 275$$

$$\begin{array}{r} 5,5 \\ \times 5 \\ \hline 27,5 \end{array}$$

$$\text{Total} \rightarrow 275 \times 74,5 = 20487,5 \text{ baldosas}$$

Largo
 $7 \times 6 = 42$
 $6 \times 5 = 30$
 $\text{Total} = 42 + 30 + 2,5 = 74,5 \text{ baldosas}$



Figura 4. Estrategia completa de Densidad en P3

Chat GPT: Hay ≈ 750.000 briznas/ m^2
 1 paso humano $\rightarrow 0,8\text{ m}$

Supongo: En ese espacio hay 3-4 pasos de largo y ≈ 9 pasos de ancho
 $\rightarrow 0,8 \cdot 4 = 3,2\text{ m}$ / $0,8 \cdot 9 = 7,2\text{ m}$

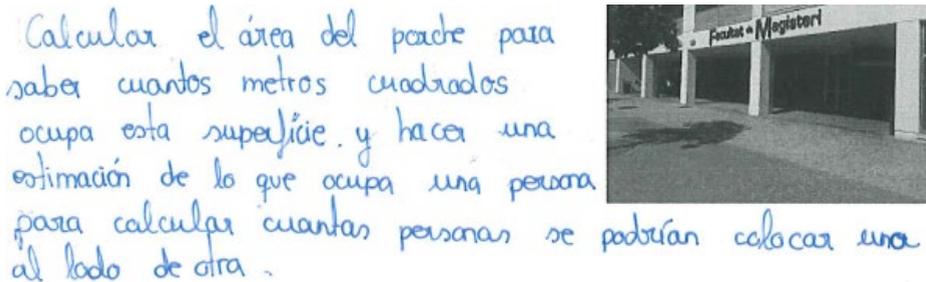
$$3,2 \times 7,2 \rightarrow A = 23,04\text{ m}^2$$

$$23,04 \cdot 750.000 = 17.280.000 \text{ briznas de césped hay en ese espacio.}$$



Las soluciones incompletas son aquellas en blanco o donde no se desarrolla suficientemente una estrategia, pues falta identificar variables o las relaciones entre ellas, no incluyendo todos los razonamientos necesarios para llegar a la estimación demandada. Por ejemplo, en la Figura 5 el resolutor identifica una variable (área del porche), pero no define correctamente la otra variable (“lo que ocupa una persona”, no especifica la magnitud) y no establece una relación entre ellas (no explica cómo obtener cuántas personas caben en el porche a partir de estas dos variables).

Figura 5. Estrategia incompleta en P1



4.4.2. Nivel de complejidad de las estrategias

Paralelamente a la clasificación de las soluciones según el tipo de estrategia utilizada, los dos codificadores identificaron si las estrategias incluían suposiciones realistas, utilizando la categorización validada en Segura et al. (2021). Se observaron dos tipos de suposiciones realistas:

- (i) La eliminación de obstáculos que restan espacio disponible a los elementos a estimar. Por ejemplo, el área que ocupan los árboles en P3, como se observa en la Figura 6.
- (ii) La heterogeneidad de la densidad en la distribución de los elementos, suponiendo que pueden darse distintas densidades según la configuración de la distribución. En la Figura 7 se observa esta suposición en P1, ya que el participante primero ofrece varias posibilidades según que las personas que vayan a ocupar el porche se encuentren más o menos “cómodas”, y luego toma una decisión ajustada a una suposición (“si llueve lo más lógico no es apretujarse”).

Figura 6. Suposición realista incluida en P3

En el espacio hay: $150 \text{ bitinos/m}^2 \times 4,5 \text{ m}^2 = 675 \text{ bitinos}$
Habrá menos ya que tendremos que calcular el espacio que ocupan los árboles y quitarse ese espacio a la cantidad total de bitinos.

Figura 7. Suposición realista incluida en P1

$8 \uparrow$
 $\quad \quad \quad \rightarrow$
 $\quad \quad \quad 40$

$40 \times 8 = 320 \text{ baldosas}$
En cada baldosa caben $\frac{4}{3}$ personas apretujadas.
 $\frac{2}{3}$ personas cómodamente.

Si llueve, lo más lógico no es apretujarse, con lo cual ponemos 3 personas por baldosa. Bajo el porche podrían resguardarse 960 personas

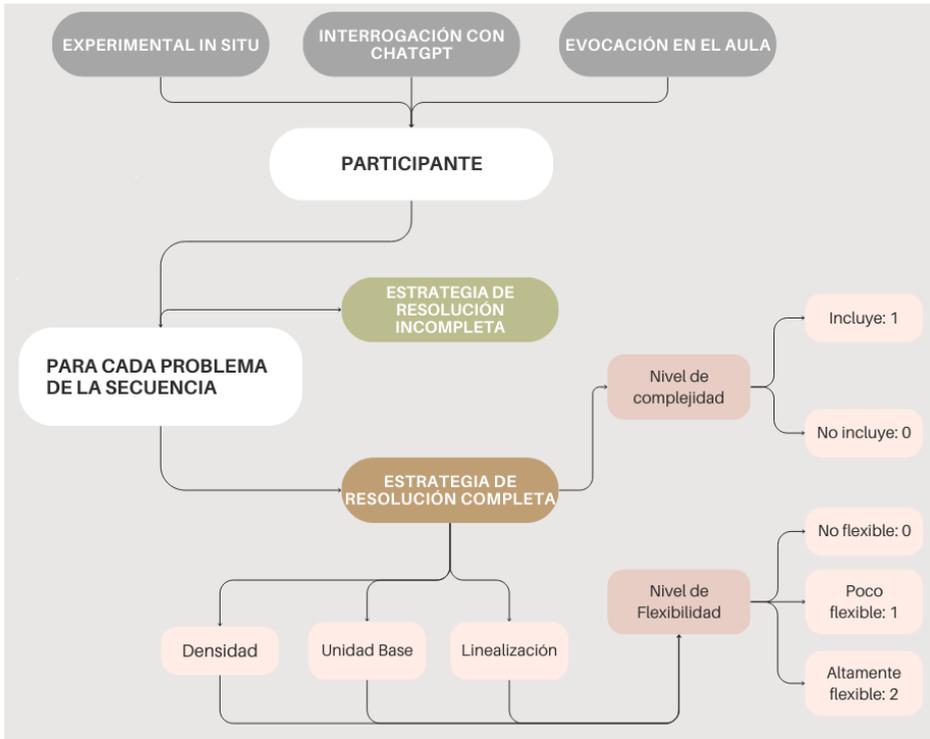


Finalmente, se codificaron todas las resoluciones completas asignándoles un nivel de complejidad: aquellas que no incluyen suposiciones realistas se codifican con 0 y las que incluyen suposiciones realistas se codifican con 1.

4.4.3. Nivel de flexibilidad de los participantes

En la segunda fase del análisis, se siguió el proceso de codificación descrito por Segura y Ferrando (2023) para evaluar el nivel de flexibilidad de cada participante: un participante se consideró *no flexible* (se codifica con 0) si utilizó la misma estrategia en todos los problemas de la secuencia; *moderadamente flexible* (se codifica con 1) si cambió de estrategia únicamente en uno de los problemas; y *altamente flexible* (se codifica con 2) si cambió de estrategia en dos o más problemas. Por ejemplo, un participante que utiliza la estrategia de Unidad Base en P1, la estrategia de Linealidad en P2, la estrategia de Densidad en P3 y de nuevo la estrategia de Unidad Base en P4, es categorizado como altamente flexible, mientras que uno que utiliza la estrategia de Densidad en P1, P2 y P3 y Unidad Base en P4 se categoriza como moderadamente flexible.

Figura 8. Esquema de análisis y codificación



De este modo, para cada participante hemos considerado las siguientes variables: entorno de resolución (experimental in situ, interrogación con ChatGPT o evocación en el aula), estrategia de resolución (completa o incompleta) en cada uno de los problemas, nivel de complejidad (0 o 1) y tipo de estrategia (Densidad, Unidad Base o Linealización) de cada resolución completa y nivel de flexibilidad (0, 1 o 2) mostrado a lo largo de la secuencia de problemas. El esquema de análisis y codificación puede verse en la Figura 8.

5. RESULTADOS

5.1. Estrategias completas e incompletas según el entorno de resolución

La Tabla 2 muestra los resultados del análisis que clasifica en completas e incompletas las 164 estrategias de los 41 participantes que resolvieron la secuencia de problemas experimentando in situ, las 164 estrategias de los 41 participantes que resolvieron los problemas en el aula con ChatGPT; y las 220 estrategias de los 55 participantes que resolvieron los problemas evocando la realidad en el aula.

Tabla 2. Frecuencia (y porcentaje) de estrategias completas e incompletas según el entorno

Entorno	Estrategias completas	Estrategias incompletas
Experimental in situ	135 (82,3%)	29 (17,7%)
Interrogación con ChatGPT	142 (86,6%)	22 (13,4%)
Evocación en aula	177 (80,5%)	43 (19,5%)

Se observa que la proporción de estrategias incompletas es similar en los tres entornos, variando entre el 13% y el 20%. Para comprobar si la proporción de estrategias incompletas depende del entorno de resolución se ha efectuado un análisis χ^2 , obteniendo que no existe relación de dependencia ($\chi^2 (2, 548) = 2.53, p = .28$).

Además, cuando comparamos el porcentaje de participantes que completó con éxito toda la secuencia de problemas en cada entorno, encontramos que en el entorno experimental son un 43,9%, mientras que en los entornos con ChatGPT y de evocación en el aula son un 56,1% y un 60%, respectivamente. Estas diferencias tampoco son significativas ($\chi^2 (2, 137) = 2.55, p = .28$).

5.2. Nivel de complejidad de las estrategias según en el entorno de resolución

La Tabla 3 recoge el nivel de complejidad medio de las estrategias desarrolladas por los participantes en cada entorno. Un valor cercano a 0 indica que la mayoría de las estrategias en ese entorno no han incluido suposiciones realistas, mientras que un valor cercano a 1 indica que la mayoría sí lo han hecho.

Tabla 3. Nivel de complejidad medio de las estrategias en cada entorno

Entorno	Nivel de complejidad medio de las estrategias
Experimental in situ	0.40
Interrogación con ChatGPT	0.11
Evocación en aula	0.06

El nivel de complejidad medio de las estrategias en el entorno experimental es mucho mayor que en los otros dos. Un análisis ANOVA confirma que existen diferencias significativas en la complejidad de las estrategias según el entorno ($F (2, 545) = 47.19, p < .001$, con tamaño de efecto grande $\eta^2 = 0.15$). El análisis post-hoc de Games-Howell para comparaciones múltiples muestra que hay significativamente más estrategias complejas en el entorno experimental que en los otros dos ($p < .001$). Aunque el nivel de complejidad medio de las estrategias en el entorno con ChatGPT casi duplica el del entorno de evocación en el aula, esta diferencia no es significativa ($p = .19$), por ser ambos muy bajos.

5.3. Tipos de estrategia según el entorno de resolución

La Tabla 4 recoge los resultados del análisis de la frecuencia de los tipos de estrategia completa usados por los participantes en cada entorno. Se observa que aparecen los tres tipos de estrategia en los tres entornos y que la proporción en el uso de cada tipo es similar en el entorno con ChatGPT y en el entorno de evocación en el aula. Sin embargo, estas proporciones son distintas en el entorno experimental.

Tabla 4. *Frecuencia de uso de cada estrategia (y porcentaje respecto al total de las estrategias en ese entorno) en los entornos*

Entorno	Linealización	Unidad Base	Densidad
Experimental in situ	53 (39%)	14 (10%)	68 (51%)
Interrogación con ChatGPT	33 (23%)	68 (48%)	41 (29%)
Evocación en aula	34 (19%)	88 (50%)	55 (31%)

En efecto, un análisis χ^2 confirma que la proporción de uso de las estrategias depende del entorno ($\chi^2(4, 454) = 61.03, p < .001$ con tamaño de efecto mediano $V = 0.26$). En el entorno experimental el uso de la estrategia Unidad Base está significativamente por debajo de su uso en los otros 2 entornos, mientras que ocurre lo contrario con el uso de la estrategia Densidad.

5.4. Nivel de flexibilidad de los participantes según el entorno de resolución

Por último, la Tabla 5 recoge los resultados del análisis de los niveles de flexibilidad de los participantes en cada entorno. Un valor del nivel de flexibilidad medio cercano a 0 significa que en ese entorno la mayoría de los participantes no son flexibles (han utilizado la misma estrategia para resolver la secuencia de problemas), mientras que un nivel cercano a 2 significa que la mayoría de los participantes son muy flexibles (han cambiado de estrategia en dos o más problemas de la secuencia).

Tabla 5. *Nivel de flexibilidad medio de los participantes en cada entorno*

Entorno	Nivel de flexibilidad medio
Experimental in situ	1.34
Interrogación con ChatGPT	1.15
Evocación en aula	0.75

Se observan diferencias en el nivel de flexibilidad medio en cada entorno y confirmamos que son significativas con un análisis ANOVA, obteniendo $F(2, 134) =$

10.21, $p < .001$ con un tamaño de efecto mediano cercano a alto ($\eta^2 = 0.13$). Las pruebas post-hoc HSD Tukey de comparaciones múltiples muestran que los participantes del entorno evocación en el aula son significativamente menos flexibles que aquellos que trabajaron en el entorno experimental ($p < .001$) y también que aquellos que lo hicieron con ChatGPT ($p < .05$). Aunque la flexibilidad media en el entorno experimental es mayor que la demostrada en el entorno con ChatGPT, esta diferencia no es significativa ($p = .38$).

6. DISCUSIÓN

6.1. ¿Depende la capacidad de los participantes para completar una estrategia del entorno de resolución?

Aunque el porcentaje de estrategias incompletas en el entorno evocación en el aula es ligeramente superior al observado en el entorno experimental, y este, a su vez, lo es respecto al entorno con ChatGPT, estas diferencias no resultan ser significativas. Tampoco hay diferencias significativas cuando comparamos la proporción de participantes que completó la secuencia en cada entorno, aunque la proporción de participantes del entorno experimental que no completó la secuencia es ligeramente mayor. Aunque, globalmente, en el entorno experimental tienen menos estrategias incompletas que en el evocativo, sus dificultades se concentran en el problema P3, pues en el lugar real del problema perciben la irregularidad de la distribución del césped (Segura et al., 2023) y esto afecta al número de secuencias completadas.

En cualquier caso, la cercanía y familiaridad con la localización de los cuatro problemas parece ser clave en la ausencia de diferencias significativas, lo que confirman estudios previos sobre la accesibilidad de los contextos cercanos (Albarracín et al., 2022). Además, los problemas propuestos no requieren de modelos complejos (Ärlebäck, 2009), lo que permite que puedan ser abordados adecuadamente desde cualquier entorno. En consecuencia, el entorno no parece ser un factor determinante que influya en la capacidad para plantear una estrategia completa.

6.2. ¿Cuál es el efecto del entorno de resolución en la complejidad de las estrategias desarrolladas por los participantes?

Sin embargo, sí que se aprecian diferencias estadísticamente significativas entre los tres entornos en cuanto a la inclusión de suposiciones realistas. Los participantes del entorno experimental presentan soluciones significativamente más complejas que los participantes de los otros entornos. Un entorno de resolución in situ promueve la percepción directa de características relevantes del contexto real –

mayoritariamente la eliminación de áreas con obstáculos– que impactan en la estimación demandada y esto facilita que los participantes introduzcan suposiciones realistas para refinar la obtención de datos (Buchholtz, 2021; Jablonsky, 2023) y lograr estimaciones más precisas (en la línea de Segura et al., 2021, 2023). Aunque las imágenes que se incluyen en los enunciados muestran estos obstáculos (columnas en el porche de la facultad en P1, alcantarillado en P2 y árboles en P3), esto no ha sido suficiente para que sean tomados en cuenta por los participantes que han resuelto los problemas en el aula (con o sin apoyo de la IA).

La otra suposición realista detectada, la consideración de distintas densidades, ha sido muy residual: solo en 3 de las 96 estrategias categorizadas como complejas, una en el entorno in situ y las otras dos en el entorno con ChatGPT. Esto podría explicarse porque no se “percibe” directamente en la realidad (en el caso de P1 y P4, los elementos a estimar no están todos “visibles”).

6.3. ¿Dependen las estrategias utilizadas por los participantes del entorno de resolución?

Aunque las tres categorías de estrategias (Densidad, Unidad Base y Linealización) son utilizadas en los tres entornos, los resultados muestran que su uso sí que depende del entorno. Así, el uso de la estrategia de Unidad Base es significativamente menor en el entorno experimental, siendo mayoritaria en este entorno la estrategia de Densidad. Este resultado está en la línea con Segura et al. (2023), pues la posibilidad de experimentar en el lugar del problema promueve que el resolutor determine cuántos elementos caben en un área pequeña y luego aplique el razonamiento proporcional. En este entorno, se observa un uso frecuente de referentes de medida no convencionales –pasos, como en la Figura 3, o baldosas que recubren el suelo, como en la Figura 7– que les permiten “concentrar” los elementos a estimar dentro del referente, lo que conduce de manera natural a la estrategia de Densidad (Figura 7). Esta estrategia es la que Ferrando y Albarracín (2021) encontraron que requiere mayor madurez matemática.

Por el contrario, en los otros dos entornos fue Unidad Base la estrategia más utilizada. En Ferrando et al. (2020) ya se había encontrado que en los planes de resolución diseñados en el aula era esta estrategia la más empleada, lo que confirma el resultado. En el caso del entorno con ChatGPT, como hemos comentado anteriormente, la IA utiliza Unidad Base y Densidad, y con mayor frecuencia la primera, lo que podría explicar los resultados. Cabe señalar que los participantes de este entorno también utilizan la estrategia de Linealización, lo que sugiere que no siempre se encuentran condicionados ni limitados por la IA, sino que combinan su uso con su propia experiencia para desarrollar una estrategia de resolución completa. Aunque estas intuiciones deben confirmarse en un estudio cualitativo más profundo de la interacción entre resolutor y ChatGPT, sugieren que la IA tiene

potencial como mediador en la resolución de problemas (Tangney y Bray, 2013; Zhang y Aslan, 2021). En este sentido, un análisis más detallado de las resoluciones apunta a que algunos participantes incorporan estimaciones propias basadas en la evocación del lugar a las pautas de resolución que les proporciona la IA (por ejemplo, en la Figura 2, el resolutor siguió las pautas de ChatGPT pero da una estimación personal de las dimensiones del porche, porque conoce ese espacio). Por el contrario, en la Figura 4, observamos que el resolutor recurre a ChatGPT para obtener datos concretos (la longitud de un paso y la densidad de briznas de césped por metro cuadrado). Como mostramos en la Figura 9, en ambos entornos los participantes han interiorizado como referentes de estimación (Segura et al., 2025) algunos elementos presentes en las imágenes, como tapas de alcantarilla o columnas.

Figura 9. Estimaciones basadas en elementos presentes en las imágenes en P3

La alcantarilla es un metro cuadrado
y a ojo veo que caben 8 en fila
y 4 en columna. Entonces esta infor-
mación se la doy al chatgpt y que
lo calcule.

$$100 \cdot 32 = 3200 \text{ briznas}$$

6.4. ¿Cuál es el efecto del entorno de resolución en la flexibilidad de los participantes?

Por último, hemos observado diferencias en los niveles de flexibilidad mostrados por los participantes en los tres entornos. El nivel de flexibilidad es significativamente menor en el entorno de evocación en el aula que en los otros dos. Además, las pruebas previas realizadas con ChatGPT parecían sugerir que los participantes con acceso a esta IA podrían mostrar menor flexibilidad que los participantes en el entorno in situ. Esta hipótesis también era respaldada por los resultados de Segura et al. (2023) sobre adaptabilidad de las estrategias a las características del contexto en el entorno experimental. Sin embargo, pese a que los participantes del entorno experimental son ligeramente más flexibles que los del entorno con ChatGPT, esta diferencia no resulta ser significativa. Estos resultados amplían los obtenidos por Segura y Ferrando (2023) sobre flexibilidad a otros entornos de resolución, y apuntan a que las relaciones entre flexibilidad, adaptabilidad y complejidad de las estrategias

(Segura et al., 2021, 2025) no son directas y deben ser estudiadas con mayor profundidad, en la línea con otros estudios sobre problemas no-rutinarios (Elia et al., 2009; Heinze et al., 2009).

7. CONCLUSIONES

Nuestro objetivo era estudiar el efecto del entorno de resolución en el desarrollo de las estrategias, su complejidad y la flexibilidad en su uso por futuros maestros, con la finalidad de aportar conocimiento que permita caracterizar su competencia resolviendo problemas de contexto real. La accesibilidad de los problemas de Fermi facilita el desarrollo de estrategias completas con independencia del entorno. Sin embargo, el entorno experimental resulta más eficaz para fomentar el uso de estrategias complejas, ya que permite percibir con mayor claridad características del contexto que pueden influir en la estimación requerida. Esto impacta en la matematización de la situación real, pues los futuros maestros tratan de obtener datos con la mayor precisión posible, incorporando suposiciones realistas, incluso sin disponer de instrumentos de medición convencionales. Finalmente, tanto el entorno experimental como el entorno con ChatGPT promueven de manera más eficaz que el entorno de evocación en el aula la flexibilidad de los futuros maestros, es decir, que usen múltiples estrategias a lo largo de la secuencia de problemas. En consecuencia, los futuros maestros que resuelven problemas de Fermi en el aula sin apoyo de tecnología desarrollan menos variedad de estrategias y éstas son más simples.

Estos resultados apuntan a la importancia de fomentar en la formación de maestros actividades de resolución de problemas fuera del aula, pues permiten desarrollar aspectos de su competencia como la riqueza de las resoluciones o el uso y comparación de múltiples estrategias. Por otro lado, nuestros resultados abren la puerta a investigar en detalle la forma en la que los futuros maestros interactúan con ChatGPT desde el punto de vista de explorar y caracterizar su potencial educativo.

Existen limitaciones en este estudio que conviene señalarse, pues abordarlas en el futuro será esencial para comprender mejor cómo estos entornos han influido en las estrategias de los futuros maestros. Así, no se han analizado los prompts –interrogaciones e instrucciones– utilizados por los participantes que interactuaron con ChatGPT, ni se ha analizado cómo integraron en sus estrategias las respuestas de la IA o el efecto de su experiencia previa con esta herramienta. Además, no hemos considerado el impacto de estos entornos en el rendimiento, en el sentido de la adecuación de las respuestas numéricas proporcionadas a la realidad, ni tampoco el efecto que puede tener el entorno en aspectos como el interés, la autoeficacia o el potencial para trabajar colaborativamente en grupo. En este sentido, aunque se insistió en que los participantes resolvieran individualmente los problemas, en el

entorno experimental esta condición no se puede controlar completamente, no pudiéndose evitar interacciones debidas a gestos y señales visibles por todos.

Este trabajo muestra que la interacción con entornos de resolución del mundo real o tecnológicos mediados por la IA puede enriquecer la experiencia de aprendizaje de los estudiantes. Esto implica que los docentes deberían considerar la variabilidad de los entornos en la planificación de problemas matemáticos del mundo real. En este sentido, como muestra este estudio, los formadores de docentes pueden promover que los futuros maestros sean competentes introduciendo entornos variados, resoluciones complejas y tratamiento flexible de las estrategias mediante la inclusión de este tipo de actividades en las asignaturas del grado de Maestro/a en Educación Primaria.

AGRADECIMIENTOS

Los autores agradecen el apoyo del proyecto PID2021-126707NB-I00 financiado por MCIU/AEI/10.13039/501100011033 y por «ERDF A way of Making Europe», y también del Proyecto de Acciones Especiales de Investigación UV-INV-AE-3668116 y el soporte del grupo de investigación GIPEAM (2021 SGR 00159).

REFERENCIAS BIBLIOGRÁFICAS

- Albarracín, L., Segura, C., Ferrando, I. y Gorgorió, N. (2022). Supporting mathematical modelling by upscaling real context in a sequence of tasks. *Teaching Mathematics and its Applications: An International Journal of the IMA*, 41(3), 183-197. <https://doi.org/10.1093/teamat/hrab027>
- Ärlebäck, J. B. (2009). On the use of realistic Fermi problems for introducing mathematical modelling in school, *The Mathematics Enthusiast*, 6(3), 331-364. <https://doi.org/10.54870/1551-3440.1157>
- Buchholtz, N. (2021). Processos de modelação dos alunos envolvidos em trilhos matemáticos. *Quadrante*, 30(1), 140-157. <https://doi.org/10.48489/quadrante.23699>
- Carlson, J. E. (1997). Fermi problems on gasoline consumption. *The Physics Teacher*, 35, 308-309. <https://doi.org/10.1119/1.2344696>
- Efthimiou, C. J. y Llewellyn, R. A. (2007). Cinema, Fermi problems and general education. *Physics education*, 42(3), 253.
- Elia, I., van den Heuvel-Panhuizen, M. y Koliopoulos, D. (2009). An investigation of young children's flexibility in solving addition tasks through the use of multiple solutions. *Educational Studies in Mathematics*, 71(1), 15-37. <https://doi.org/10.1007/s10649-008-9156-3>

- Ferrando, I. y Albarracín, L. (2021). Students from grade 2 to grade 10 solving a Fermi problem: analysis of emerging models. *Mathematics Education Research Journal*, 33(1), 61-78. <https://doi.org/10.1007/s13394-019-00292-z>
- Ferrando, I., Segura, C. y Pla-Castells, M. (2020). Relations entre contexte, situation et schéma de résolution dans les problèmes d'estimation. *Canadian Journal of Science, Mathematics and Technology Education*, 20(3), 557-573. <https://doi.org/10.1007/s42330-020-00102-w>
- Galbraith, P. y Stillman, G. (2001). Assumptions and context: Pursuing their role in modelling activity. En J. F. Matos, W. Blum, S. K. Houston y S. P. Carreira (Eds.), *Modelling and mathematics education: ICTMA 9: Applications in science and technology* (pp. 300-310). Horwood Publishing.
- Haberzettl, N., Klett, S. y Schukajlow, S. (2018). Mathematik rund um die Schule – Modellieren mit Fermi-Aufgaben [Mathematics around school – Modelling with Fermi tasks]. En K. Eilerts y K. Skutella (Eds.), *Neue Materialien für einen realitätsbezogenen Mathematikunterricht 5. Ein ISTRON-Band für die Grundschule* (pp. 31-41). Springer Spectrum.
- Hartmann, L. M. y Schukajlow, S. (2021). Interest and Emotions While Solving Real-World Problems Inside and Outside the Classroom. En Leung, F. K. S., Stillman, G. A., Kaiser, G., Wong, K. L. (Eds.), *Mathematical Modelling Education in East and West*. Springer. https://doi.org/10.1007/978-3-030-66996-6_13
- Heinze, A., Star, J. R. y Verschaffel, L. (2009). Flexible and adaptive use of strategies and representations in mathematics education. *ZDM Mathematics Education*, 41(5), 535-540. <https://doi.org/10.1007/s11858-009-0214-4>
- Jablonski, S. (2023). Indoors vs. Outdoors: Student Perception of Different Modelling Settings. *Research in Integrated STEM Education*, 1(3), 403-421. <https://doi.org/10.1163/27726673-bja00016>
- Keune, M. y Henning, H. (2003). Modelling and spreadsheet calculation. En Q. X. Ye, W. Blum, S.-K. Houston y Q. Y. Yian (Eds.), *Mathematical modelling in education and culture* (pp. 101-110). Horwood Publishing.
- Kuhail, M., Alturki, N., Alramlawi, S. y Alhejori, K. (2023). Interacting with educational chatbots: A systematic review. *Education and Information Technologies*, 28(1), 973-1018. <https://doi.org/10.1007/s10639-022-11177-3>
- Lo, C. K. (2023). What Is the Impact of ChatGPT on Education? A Rapid Review of the Literature. *Education Sciences*, 13, 410. <https://doi.org/10.3390/educsci13040410>
- López-Simó, V. y Rezende, M. F. (2024). Challenging ChatGPT with Different Types of Physics Education Questions. *The Physics Teacher*, 62, 290-294. <https://doi.org/10.1119/5.0160160>

- Maass, K., Geiger, V., Ariza, M. R. y Goos, M. (2019). The Role of Mathematics in interdisciplinary STEM education. *ZDM Mathematics Education*, 51, 869-884. <https://doi.org/10.1007/s11858-019-01100-5>
- Montejo-Gómez, J., López-Centella, E. y Fernández-Ahumada, E. (2024). Solving Estimation Tasks: Novel Features of the Emerging Models When Three-Dimensional Geometry Becomes Relevant. En H. S. Siller, V. Geiger, G. Kaiser, (Eds.), *Researching Mathematical Modelling Education in Disruptive Times* (pp. 607-618). Springer.
- Noster, N., Gerber, S. y Siller, H. S. (2024). Pre-Service Teachers' Approaches in Solving Mathematics Tasks with ChatGPT. *Digital Experiences in Mathematics Education*, 1-25. <https://doi.org/10.1007/s40751-024-00155-8>
- Pepin, B., Buchholtz, N. y Salinas-Hernández, U. (2025). A Scoping Survey of ChatGPT in Mathematics Education. *Digital Experiences in Mathematics Education*. <https://doi.org/10.1007/s40751-025-00172-1>
- Segura, C. y Ferrando, I. (2023). Pre-service teachers' flexibility and performance in solving Fermi problems. *Educational Studies in Mathematics*, 113(2), 207-227. <https://doi.org/10.1007/s10649-023-10220-5>
- Segura, C., Ferrando, I., y Albarracín, L. (2021). Análisis de los factores de complejidad en planes de resolución individuales y resoluciones grupales de problemas de estimación de contexto real. *Quadrante*, 30(1), 31-51. <https://doi.org/10.48489/quadrante.23592>
- Segura, C., Ferrando, I., y Albarracín, L. (2023). Does collaborative and experiential work influence the solution of real-context estimation problems? A study with prospective teachers. *The Journal of Mathematical Behavior*, 70, 101040. <https://doi.org/10.1016/j.jmathb.2023.101040>
- Segura, C., Gallart, C. y Ferrando, I. (2025). Influence of pre-service primary school teachers' prior knowledge of measurement and measurement estimation in solving modelling problems. *Journal of Mathematics Teacher Education*. <https://doi.org/10.1007/s10857-025-09685-3>
- Spreitzer et al (2024) Mathematical Modelling Abilities of Artificial Intelligence Tools: The Case of ChatGPT. *Education Sciences*, 14, 698. <https://doi.org/10.3390/educsci14070698>
- Sriraman, B., y Knott, L. (2009). The mathematics of estimation: Possibilities for interdisciplinary pedagogy and social consciousness. *Interchange*, 40(2), 205-223.
- Star, J. R., y Rittle-Johnson, B. (2008). Flexible problem solving: The roles of strategy choice and strategy flexibility in knowing when and how to use procedures. *The Journal of Problem Solving*, 2(2), 71-91. <https://doi.org/10.7771/1932-6246.1042>

- Tangney, B., y Bray, A. (2013). Mobile technology, maths education & 21C learning. In *Proceedings of the 12th world conference on mobile and contextual learning* (pp. 20–27). College of the North Atlantic–Qatar.
- Toalongo, X., Trelles, C. y Alsina, Ángel. (2024). Los Problemas de Fermi y las Modelling Eliciting Activities como un recurso para fomentar la Modelización Matemática entre el alumnado de Educación Primaria. *Edma 0-6: Educación Matemática En La Infancia*, 13(1), 58-92. <https://doi.org/10.24197/edmain.1.2024.58-92>
- Verschaffel, L., Greer, B. y de Corte, E. (2000). *Making sense of word problems*. Swets & Zeitlinger.
- Zhang, K. y Aslan, A.B. (2021). AI technologies for education: Recent research & future directions. *Computer and Education: Artificial Intelligence 2*, 100025. <https://doi.org/10.1016/j.caeai.2021.100025>