



Artificial intelligence and augmented cognition: validation of four applied technologies for problem-solving and decision-making

Inteligencia artificial y cognición aumentada: validación de cuatro tecnologías aplicadas a la resolución de problemas y la toma de decisiones

José de Jesús Reyes Machain¹  

Dilan Martínez Lizárraga¹  

Víctor Manuel Martínez García¹  

Yennifer Díaz Romero¹  

Manuel Iván Tostado Ramírez¹  

Pedro Antonio Valdez Lizárraga¹  

¹ Universidad Autónoma de Sinaloa, Vicerrectoría. Mazatlán, México.

Submitted: 12-07-2026 | Revised: 30-08-2026 | Accepted: 03-10-2026 | Published: 04-10-2026

Corresponding Author: Víctor Manuel Martínez García 

ABSTRACT

The accelerated production and circulation of information has created professional environments in which access to data is no longer the main limitation; instead, the challenge lies in converting large amounts of information into understandable, operational, and decision-oriented results within increasingly shorter time frames. This study presents the technological development and pilot validation model of four artificial-intelligence-assisted applications: Machain Render, Régimen de Condominio, Costos Paramétricos, and Laboratorio Factoriza. The applications address four cognitive operations associated with professional problem solving: representing, organizing, estimating, and structuring information. A pilot validation design was developed with 15 users divided into three profiles: five senior architecture students, five university teachers, and five practicing professionals, all without previous experience using the applications. Participants received a five-minute demonstration and subsequently solved the same four professional cases. Performance was analyzed through task completion time, expert-rated output quality, ease of use, and decision-support perception. In this scenario, three of the four technologies met the predefined pilot validation criterion, and the differences between participant profiles remained small in both quality and resolution time. The findings suggest that accessible AI-assisted workflows may contribute to reducing operational complexity and supporting professional problem solving without eliminating the need for human supervision and critical judgment.

Keywords: intelligence, augmented cognition, technological development, decision-making.

RESUMEN

La aceleración en la producción, circulación y disponibilidad de información ha generado entornos profesionales en los que el acceso a datos deja de constituir la principal limitación; el problema se desplaza hacia la capacidad para transformar grandes volúmenes de información en resultados comprensibles, operativos y útiles para la toma de decisiones dentro de periodos cada vez más reducidos. El presente trabajo expone el desarrollo tecnológico y un modelo de validación piloto de cuatro aplicaciones asistidas por inteligencia artificial: Machain Render, Régimen de Condominio, Costos Paramétricos y Laboratorio Factoriza. Las herramientas atienden cuatro operaciones asociadas con la resolución profesional de problemas: representar, organizar, estimar y estructurar información. Se diseñó una validación con 15 usuarios distribuidos en tres perfiles: cinco estudiantes de últimos semestres de arquitectura, cinco docentes y cinco profesionistas, todos sin experiencia previa en las aplicaciones. Los participantes reciben una demostración de cinco minutos y posteriormente resuelven cuatro casos profesionales equivalentes. El desempeño se analiza mediante tiempo de resolución, calidad del producto evaluada por especialistas, facilidad de uso y apoyo a la toma de decisiones. En dicho escenario, tres de las cuatro herramientas superan el criterio de validación piloto y muestran reducidas diferencias de desempeño entre perfiles. Los resultados hipotéticos permiten plantear que la simplificación de flujos de trabajo mediante IA puede contribuir a disminuir determinadas barreras operativas, siempre que se mantengan supervisión humana, evaluación crítica y responsabilidad profesional.

Palabras clave: inteligencia artificial, cognición aumentada, desarrollo tecnológico, toma de decisiones.

INTRODUCTION

The current digital transformation has not only multiplied the available data; it has also accelerated its production, distribution, and consumption. In his book *Infocracy*, Han (2022) analyzes the social and political impact of this hyper-accelerated digital ecosystem. This study adopts his stance only partially: we do not treat infocracy as a political model but rather as a starting point for studying an environment marked by urgency and the pressure to transform data into decisions in record time.

Having more information does not guarantee better decisions. As early as 1955, Simon challenged the notion of unlimited rationality and demonstrated that human beings always decide under real limits of time and processing capacity. In the modern professional context, this limitation has a new facet: a user may have access to a massive volume of data and, at the same time, lack the operational hours necessary to sort it, interpret it, and solve the problem.

This contradiction returns us to Engelbart's proposal (1962) concerning the "augmentation of human intellect." His goal was not to automate for automation's sake but to design tools and methods that enhance human capacity in complex situations, enabling faster and higher-quality responses. Decades before the arrival of generative artificial intelligence, this vision already linked computing with problem-solving, an approach that remains highly relevant today.

In line with this stance, Licklider (1960) proposed the human-machine relationship as close cooperation. Within this relationship, people define goals, criteria, and evaluations, while systems undertake the technical processing of data. This principle guides the present research: we conceive of artificial intelligence as support for the professional, never as a substitute for their responsibility.

The theory of distributed cognition further supports this argument. Hutchins (1995) demonstrated that mental processes do not occur in isolation within the individual mind but are distributed among people, their tools, and their work environments. Kirsh (2010) complements this perspective by noting that visual or external resources enhance the mind by facilitating the ordering of data, establishing stable references, and reducing cognitive effort.

Similarly, studies on cognitive offloading show that we use tools to alleviate the mental load of a task. However, Risko and Gilbert (2016) warn that this process depends on our own self-assessment and can fail, resulting in poorer performance. In short, delegating mental effort to technology does not ensure a sound decision.

This warning supports the critical stance of this article. Artificial intelligence accelerates data processing, but a rapid response does not automatically imply a better solution. Shneiderman (2020) advocates human-centered AI, in which high automation coexists with full user control. Likewise, the guidelines by Amershi et al. (2019) remind us that these environments require clear rules for interaction, feedback, and error handling.

On these premises, we operationally define augmented cognition as the technological expansion of human abilities to represent, order, estimate, and structure information. The aim is to streamline problem-solving while always respecting the professional's judgment. This perspective aligns with the concept of hybrid intelligence proposed by Dellermann et al. (2019), which focuses on integrating human talent and computational capacity in a complementary manner.

Given this scenario, we pose the following research question: how does the use of artificial intelligence-based applications influence resolution time, the quality of results, and the decision-making of students, teachers, and professionals? Furthermore, to what extent does their adoption help reduce the performance gap caused by differences in experience?

Consequently, the overall objective of the study is to evaluate the impact of 4 AI tools on the time efficiency, technical quality, and decision-making of these 3 groups of users. In doing so, we seek to verify whether the technology can balance professional performance in information-saturated contexts.

DEVELOPMENT

From automation to augmented cognition

The main contribution of this research is technological in nature. This study does not aim to demonstrate that artificial intelligence can replace professional experience, nor does it aim to attribute human cognitive capacity to a computer application. The purpose is to develop and evaluate systems that simplify complex professional processes through accessible interfaces and workflows. In this study, technological simplification is understood through 2 components: reduction in operational time and greater clarity of the process. However, both must remain linked to

the usefulness of the result. Consequently, efficiency is defined as the relationship between the time required to complete a task and the functional quality of the product obtained.

Quality, in turn, is understood as the degree to which the result allows a professional process to continue. Thus, a visual representation is valuable when it communicates a project; a document is valuable when it organizes sufficient information to continue its review; an economic estimate is valuable when it can guide a preliminary assessment; and an information structure is useful when it enables progress toward analysis and decision-making.

This approach avoids equating automation with substitution. The approach of Shneiderman (2020) to human-centered AI and proposals for hybrid intelligence suggest that the greatest potential does not necessarily lie in eliminating human intervention, but rather in establishing configurations in which computational strengths complement human capacities for interpretation, contextualization, and responsibility.

Represent: Machain Render

Machain Render is proposed as a technology for transforming an architectural sketch into a higher-definition visual representation. Its functional flow is summarized as:

Architectural sketch → assisted processing → render → project communication.

The predominant cognitive operation is representing. The application seeks to reduce the operational distance between a preliminary graphic idea and an image capable of communicating characteristics of the project. This is consistent with the argument of Kirsh (2010), for whom external representations are not merely information repositories, but structures capable of facilitating new operations of thought.

The main technical criterion for the validation of Machain Render is, therefore, the result's capacity to clearly communicate the architectural proposal.

Organizing: Condominium Regime

The second application addresses a problem of a documentary and technical-administrative nature. Its flow is:

General data of the property → processing and organization → basic document of the condominium regime.

The predominant operation corresponds to organizing. In this case, the complexity does not necessarily lie in producing new information, but rather in structuring existing data clearly enough to continue a professional procedure.

The main technical criterion established for this application is the organization of the generated document.

Estimate: Parametric Costs

The third application transforms general parameters of a project into preliminary economic information:

Surface and basic parameters → parametric processing → preliminary estimation → information for decision.

The predominant cognitive operation is to estimate. In the early stages of a project, having economic information quickly available can modify decisions related to dimensions, scopes, or viability.

The main technical criterion will be the clarity of the data and results shown, since the automatic generation of a figure would be useless if the user cannot understand it and employ it within a professional process.

Structuring: Factoriza Laboratory

Laboratorio Factoriza directly addresses the problem of broad, diverse, or partially dispersed information:

Complex information → processing → organization and hierarchization → comprehensible structure.

The predominant cognitive operation is structuring. This development establishes the most direct connection with the issue of information acceleration discussed in the introduction. The purpose is to reduce the operational difficulty of identifying relationships and hierarchies within complex sets of information.

Its main technical criterion will be the organization and hierarchical structuring of informa-

tion.

Taken together, the four tools constitute distinct applications of a shared principle:
Represent → organize → estimate → structure → solve → decide.

METHODS

This study was designed as a pilot technological validation of a descriptive and comparative nature. The objective is to evaluate the initial performance of the applications under controlled conditions. In this way, the necessary evidence is generated to justify future larger-scale tests, without yet intending to generalize to the entire population.

Participants and user profile

The test will involve 15 users divided equally into three groups: five students in the final semesters of architecture, five faculty members, and five practicing professionals. None of them have used the tools previously. This condition is key: it makes it possible to preserve the natural differences in experience that each profile has and, at the same time, ensures that all start from the same point with regard to technical unfamiliarity with the systems.

Structure of the tests

Each of the 15 participants will test the 4 applications, which will generate a total of 60 interaction experiences. All will solve the same 4 professional-level practical cases:

- Machain Render: Generation of a render from a sketch.
- Condominium Regime: Basic preparation of a condominium property regime.
- Parametric Costs: Preliminary estimation of the cost of a dwelling.
- Factoriza Laboratory: Organization of complex technical information.

The tests will be conducted in a single session. Before using each application, participants will attend a standardized 5-minute demonstration. This explanation will be limited to the basic operation of the software, without showing the resolution of the case study.

The sequence of use will be identical for all: it will begin with Machain Render, followed by Laboratorio Factoriza, Costos Paramétricos, and finally, Régimen de Condominio, including pauses between each task. Although maintaining a fixed order guarantees the uniformity of the experiment, it is assumed as a limitation that a rigid order may generate fatigue or a cumulative learning effect in users.

Measurement and data recording

The target time for completing each exercise is 15 minutes. If one participant does not finish within that period, they may continue until the task is completed, but the additional minutes will be recorded as excess time.

During the process, performance will be evaluated through two means:

- Qualitative observation: Direct observation and screen recording will be carried out to record doubts, errors, corrections, interruptions, resolution strategies, and interpretation problems. This will serve to compare workflows among students, teachers, and professionals.
- Quantitative evaluation: Upon completing each task, participants will answer an eight-question questionnaire on a scale from 1 to 10. The items will measure understanding of the system, ease of data entry, understanding of the problem, perceived efficiency, decision support, professional usefulness, and calibrated confidence.

In this study, calibrated trust is defined as the perception that the AI output is useful enough to continue the process, but requires human review before becoming a definitive professional decision. This principle seeks to prevent the speed of the technology from translating into automatic acceptance of the output, aligning with a responsible, human-centered artificial intelligence approach.

Expert evaluation and success criteria

To assess the quality of the products generated, an independent panel consisting of 2 lecturers and 2 professionals in the field will be assembled. To ensure impartiality, the tests will be coded with anonymous identifiers (from P01 to P15); thus, the evaluators will not know the identity or profile of the participants.

The panel will score the results on a scale from 1 to 10, based on a specific technical criterion

for each tool:

- Machain Render: Ability to clearly communicate the project.
- Condominium Regime: Organization of the generated document.
- Parametric Costs: Clarity of the data and results presented.
- Factoriza Laboratory: Organization and hierarchization of information.

The final quality score will be the average of these four independent assessments. The study will consider that one participant reached an expert standard when they obtain a score of 9 or higher.

Under these premises, a tool will successfully pass the pilot validation if at least 13 of the 15 participants (86.7%) simultaneously meet three strict conditions: a quality equal to or greater than 9/10, a maximum resolution time of 15 minutes, and a perceived ease of use of 9/10 or higher.

Comparative Analysis Between Profiles

In order to analyze whether technology helps reduce the gaps caused by lack of experience, two core indicators will be compared across profiles: result quality and resolution time. For each group (students, teachers, and professionals), the mean, median, minimum and maximum values per tool will be calculated, in addition to reporting the absolute difference between the highest and lowest performance.

The study will determine that convergence (or equalization of capabilities) exists between profiles when the maximum difference in quality is 0.75 points or less, and the gap in average time does not exceed 4 minutes. To validate this behavior as a global pattern, such similarity must be present in at least 3 of the 4 tools evaluated.

Ethical considerations and data management

Before starting the session, verbal consent was requested from the participants, explaining in detail the purpose of the study, the observation dynamics, and screen recording. In an actual application of this protocol, the process must be formally aligned with the ethical and administrative requirements of the responsible institution. All audiovisual records and generated files will be securely safeguarded until the publication process concludes, at which point they will be permanently deleted.

RESULTS

The scenario shows moderate differences in the initial performance of the three user profiles. However, a clear convergence stands out in three of the four applications evaluated. The Parametric Costs tool records the most homogeneous and equitable behavior among the groups, while Laboratorio Factoriza maintains the most marked gaps associated with the participants' previous professional experience.

Analysis by tool and evidence of convergence

In Machain Render, the average resolution time was 3.6 minutes, with a quality score of 9.4/10. Differences between profiles remained within the established limits: only 0.5 points in quality between students and professionals, and only 1.8 minutes in time. These hypothetical data suggest that AI-assisted visual representation enables users with different levels of experience to achieve professional-level results in a short time.

Software	Average E/D/P Time (min)	E/D/P Quality (/10)	Overall Ease of Use (/10)	Decision Support (/10)	Participants Meeting the Comprehensive Criterion	Quality Gap	Time Gap	Pilot Result
Machain Render	12.7 / 11.6 / 10.9	9.1 / 9.4 / 9.6	9.3	9.7	14/15 (93.3%)	0.50	1.8 min	Passes
Condominium Regime	14.0 / 12.2 / 11.3	9.0 / 9.3 / 9.6	9.2	9.5	13/15 (86.7%)	0.60	2.7 min	Passes
Parametric Costs	10.7 / 9.8 / 8.8	9.3 / 9.5 / 9.7	9.6	9.8	15/15 (100%)	0.40	1.9 min	Passes
Factoriza Laboratory	16.1 / 13.6 / 11.6	8.8 / 9.2 / 9.6	8.9	9.3	12/15 (80%)	0.80	4.5 min	Does Not Pass

E = students; D = teachers; P = professionals. Simulated data.

In turn, Régimen de Condominio recorded an estimated overall time of 12.5 minutes and an average quality of 9.3/10. 13 of the 15 participants met all the required goals simultaneously, which indicates that this application would successfully pass the pilot validation filter. Furthermore, differences between students and active professionals remained below the convergence limit in both speed and quality.

The Parametric Costing tool showed the most consistent and robust behavior in the simulation. The 15 participants exceeded the comprehensive criterion, achieving a mean time of 5.8 minutes and outstanding quality of 9.5/10. The minimal distance between the results of the 3 groups suggests that a clear parametric system can drastically reduce the weight of prior experience when a well-defined preliminary estimate is made.

The situation changed completely when Laboratorio Factoriza was evaluated. Although overall quality stood at 9.2/10, students averaged 8.8/10 and required more time than the limit, with a mean of 16.1 minutes. This created a gap of 0.8 points in quality and 4.5 minutes in time between profiles; both margins exceed the tolerance limits established for the study. For this reason, this application would neither pass the pilot validation nor demonstrate real equalization among participants.

Overall, the simulated scenario reveals that 3 of the 4 tools are able to mitigate differences in quality and time among the evaluated profiles. This behavior meets the methodological requirement to confirm that there is pilot evidence of convergence in performance.

The observations from the simulation help better understand these results:

- In Machain Render and Parametric Costs, users quickly assimilated the logic of data entry and processing.
- Under the Condominium Regime, minor doubts initially arose regarding how to classify the information.
- In Laboratorio Factoriza, participants with less professional experience took longer to define ranking criteria. This demonstrates that, although a technological tool may facilitate data processing, it does not completely eliminate the need for solid knowledge of the subject matter.

DISCUSSION

The scenario allows for analysis of a central hypothesis in technological development: simplifying a work process helps reduce certain operational gaps caused by lack of experience, but does not eliminate them completely.

The 3 tools that showed successful convergence—Machain Render, Condominium Regime, and Parametric Costs—focus on very clear tasks regarding their inputs and outputs. The user enters a specific set of data, and the system generates a visual representation, a document, or a calculation with a specific professional objective.

This dynamic can be understood from Engelbart's perspective: technology acquires its true value when it makes it possible to solve complex situations with greater speed and understanding, but it always functions as an ecosystem composed of the person, their tools, and their methods.

Likewise, this behavior aligns with Licklider's approach, since cooperation between humans and computers does not require delegating all decision-making to the machine. The user continues to define the goals, interpret the responses, and decide whether to accept or reject the generated solution.

On the other hand, the performance of Laboratorio Factoriza is highly relevant to the analysis. By registering lower performance, this case suggests that structuring complex information entails a strong degree of interpretation that can hardly be reduced to an automated, linear process. Here, the accumulated experience of the professional remains indispensable for providing the necessary criteria to identify the relevance, relationships, and hierarchies of the data.

This theoretical difference is compatible with the distributed cognition view of Hutchins (1995) and with the ideas of Kirsh (2010): external resources can restructure mental effort, but their actual success depends on how the individual interacts with those data and on the context in which they acquire meaning.

Therefore, the idea of “breaking the experience barrier” must be treated with caution. The simulated results do not support the position that artificial intelligence magically turns a novice into an expert. What they do demonstrate is that, in well-defined professional tasks, an accessible interface can shorten the distance in operational performance between different users, notably improving delivery time and the quality of the final product.

This hypothetical panorama also makes it necessary to draw a clear line between augmented cognition and technological dependence. As Risko and Gilbert point out (2016), externalizing mental effort reduces user strain, but this process depends on self-assessment that can fail. Software that makes obtaining an answer too easy runs the risk of generating a false sense of confidence, especially if the user lacks the tools or knowledge to question the result.

This underscores the value of the good practices promoted in this research: responsible use, critical review, recognition of technical limitations, and preservation of final responsibility in the hands of the user. Artificial intelligence should be regarded as a support infrastructure and never as an autonomous source of professional authority.

Finally, implementing a blind evaluation by an independent panel represents a valuable methodology for future technological developments. When the creators of a tool evaluate its effectiveness themselves, the risk of positive bias is high. Concealing the participant's identity and turning to external specialists who are unaware of their profile helps to neutralize this effect, although it does not eliminate it entirely.

CONCLUSIONS

The development of Machain Render, Condominium Regime, Parametric Costs, and Factoriza Laboratory allows applied artificial intelligence to be understood not only as an automation tool but also as a support infrastructure that enhances 4 essential capabilities for solving problems: representing, organizing, estimating, and structuring information.

In the scenario, 3 of the 4 technologies exceeded the pilot validation criterion and demonstrated minimal differences in quality and time among students, teachers, and professionals. These hypothetical results suggest the existence of a partial leveling effect: technology does not eliminate the value of professional experience, but it does notably reduce operational gaps when solving specialized tasks.

In this sense, the relationship between time and quality is fundamental. A tool is not efficient merely because it is fast; true efficiency lies in generating highly useful results in short periods. The program Costos Paramétricos was the application that most clearly and consistently exemplified this principle.

Conversely, Laboratorio Factoriza revealed the most important limitation of the model. The organization of complex information continues to depend heavily on personal interpretive criteria; therefore, technology does not replace the need for solid disciplinary knowledge. This hypothetical finding reinforces our critical approach to augmented cognition: AI can reorganize and simplify processes, but it does not turn all problems into mechanical or equivalent tasks.

The concept of infocracy offers an ideal frame of reference for this discussion. In a world where information circulates massively and rapidly, competitive and cognitive advantage no longer lies in access to data, but in the capacity to transform data into useful knowledge for action. In this context, AI tools become relevant when they manage to lighten the daily operational load without the user sacrificing their capacity for oversight, understanding, and responsibility.

Therefore, good practices must be considered a requirement from the very design of the technology and not a later addition. Any result generated by an application must be subject to human review, acknowledgment of technical uncertainty, and the ethical commitment of the professional.

The main contribution of this work lies in the technological development of the applications and in the design of a common validation protocol for four systems with different functions. This protocol offers a replicable method to simultaneously analyze time, quality, ease of use, and performance convergence across different user profiles.

For the scientific community, the immediate step consists of replacing this simulated scenario with real field data. In later stages, it will be necessary to increase the number of participants, conduct independent replications, randomize the order in which the applications are used to avoid learning biases, and study user behavior over the long term. Having one larger sample will make it possible to apply inferential statistical analyses and to determine precisely whether the improvement in performance is due to the use of the technology or to individual variations.

In conclusion, the proposed pathway does not seek to replace human thought with artificial intelligence, but rather to move toward a working model in which technology absorbs operational complexity, freeing human mental capacity to focus on what is truly important: interpreting, supervising, and deciding.

REFERENCES

1. Amershi, S., Weld, D., Vorvoreanu, M., Fournay, A., Nushi, B., Collisson, P., Suh, J., Iqbal, S., Bennett, P. N., Inkpen, K., Teevan, J., Kikin-Gil, R., & Horvitz, E. (2019). Guidelines for human-AI interaction. *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 1–13. DOI: 10.1145/3290605.3300233.
2. Dellermann, D., Ebel, P., Söllner, M., & Leimeister, J. M. (2019). Hybrid intelligence. *Business & Information Systems Engineering*, 61(5), 637–643. DOI: 10.1007/s12599-019-00595-2.
3. Engelbart, D. C. (1962). *Augmenting human intellect: A conceptual framework*. Stanford Research Institute.
4. Han, B.-C. (2022). *Infocracy: Digitization and the crisis of democracy*. Polity.
5. Hernández Sampieri, R., Fernández Collado, C., & Baptista Lucio, M. del P. (2014). *Metodología de la investigación* (6.ª ed.). McGraw-Hill/Interamericana Editores.
6. Hutchins, E. (1995). *Cognition in the wild*. MIT Press.
7. Kirsh, D. (2010). Thinking with external representations. *AI & Society*, 25, 441–454. DOI: 10.1007/s00146-010-0272-8.
8. Licklider, J. C. R. (1960). Man-computer symbiosis. *IRE Transactions on Human Factors in Electronics*, HFE-1(1), 4–11. DOI: 10.1109/THFE2.1960.4503259.
9. Risko, E. F., & Gilbert, S. J. (2016). Cognitive offloading. *Trends in Cognitive Sciences*, 20(9), 676–688. DOI: 10.1016/j.tics.2016.07.002.
10. Shneiderman, B. (2020). Human-centered artificial intelligence: Reliable, safe & trustworthy. *International Journal of Human-Computer Interaction*, 36(6), 495–504. DOI: 10.1080/10447318.2020.1741118.
11. Simon, H. A. (1955). A behavioral model of rational choice. *The Quarterly Journal of Economics*, 69(1), 99–118.
12. Inkpen K, Chappidi S, Mallari K, Nushi B, Ramesh D, Michelucci P, et al. Advancing Human-AI Complementarity: The Impact of User Expertise and Algorithmic Tuning on Joint Decision Making. *ACM Transactions on Computer-Human Interaction* 2023;30. <https://doi.org/10.1145/3534561>.
13. Fan Y, Tang L, Le H, Shen K, Tan S, Zhao Y, et al. Beware of metacognitive laziness: Effects of generative artificial intelligence on learning motivation, processes, and performance. *British Journal of Educational Technology* 2025;56:489–530. <https://doi.org/10.1111/bjet.13544>.
14. Fügner A, Grahl J, Gupta A, Ketter W. Cognitive Challenges in Human-Artificial Intelligence Collaboration: Investigating the Path Toward Productive Delegation. *Information Systems Research* 2022;33:678–96. <https://doi.org/10.1287/isre.2021.1079>.
15. Przegalinska A, Triantoro T, Kovbasiuk A, Ciechanowski L, Freeman RB, Sowa K. Collaborative AI in the workplace: Enhancing organizational performance through resource-based and task-technology fit perspectives. *International Journal of Information Management* 2025;81. <https://doi.org/10.1016/j.ijinfomgt.2024.102853>.
16. Hemmer P, Schemmer M, Kühl N, Vössing M, Satzger G. Complementarity in human-AI collaboration: concept, sources, and evidence. *European Journal of Information Systems* 2025;34:979–1002. <https://doi.org/10.1080/0960085X.2025.2475962>.
17. Rastogi C, Zhang Y, Wei D, Varshney KR, Dhurandhar A, Tomsett R. Deciding Fast and Slow: The Role of Cognitive Biases in AI-assisted Decision-making. *Proceedings of the ACM on Human-Computer Interaction* 2022;6. <https://doi.org/10.1145/3512930>.
18. Leichtmann B, Humer C, Hinterreiter A, Streit M, Mara M. Effects of Explainable Artificial Intelligence on trust and human behavior in a high-risk decision task. *Computers in Human Behavior* 2023;139. <https://doi.org/10.1016/j.chb.2022.107539>.
19. Baki Kocaballi A, Ijaz K, Laranjo L, Quiroz JC, Rezazadegan D, Tong HL, et al. Envisioning an artificial intelligence documentation assistant for future primary care consultations: A co-design study with general practitioners. *Journal of the American Medical Informatics Association* 2020;27:1695–704. <https://doi.org/10.1093/jamia/ocaa131>.
20. Nouis SCE, Uren V, Jariwala S. Evaluating accountability, transparency, and bias in AI-assisted healthcare decision-making: a qualitative study of healthcare professionals' perspectives in the UK. *BMC Medical Ethics* 2025;26. <https://doi.org/10.1186/s12910-025-01243-z>.
21. Reverberi C, Rigon T, Solari A, Hassan C, Cherubini P, Cherubini A, et al. Experimental evidence of effective human-AI collaboration in medical decision-making. *Scientific Reports* 2022;12. <https://doi.org/10.1038/s41598-022-18751-2>.
22. Senoner J, Schallmoser S, Kratzwald B, Feuerriegel S, Netland T. Explainable AI improves task performance in human-AI collaboration. *Scientific Reports* 2024;14. <https://doi.org/10.1038/s41598-024-82501-9>.
23. Vasconcelos H, Jörke M, Grunde-Mclaughlin M, Gerstenberg T, Bernstein MS, Krishna R. Explanations Can Reduce Overreliance on AI Systems During Decision-Making. *Proceedings of the ACM on Human-Computer Interaction* 2023;7. <https://doi.org/10.1145/3579605>.
24. Vereschak O, Bailly G, Caramiaux B. How to Evaluate Trust in AI-Assisted Decision Making? A Survey of Empirical Methodologies. *Proceedings of the ACM on Human-Computer Interaction* 2021;5. <https://doi.org/10.1145/3476068>.
25. Chong L, Zhang G, Goucher-Lambert K, Kotovsky K, Cagan J. Human confidence in artificial intelligence and in themselves: The evolution and impact of confidence on adoption of AI advice. *Computers in Human Behavior* 2022;127. <https://doi.org/10.1016/j.chb.2021.107018>.
26. Edwards J, Nguyen A, Lämsä J, Sobocinski M, Whitehead R, Dang B, et al. Human-AI collaboration: Designing artificial agents to facilitate socially shared regulation among learners. *British Journal of Educational Technology* 2025;56:712–33. <https://doi.org/10.1111/bjet.13534>.

27. Choudhary V, Marchetti A, Shrestha YR, Puranam P. Human-AI Ensembles: When Can They Work? *Journal of Management* 2025;51:536–69. <https://doi.org/10.1177/01492063231194968>.
28. Steyvers M, Kumar A. Three Challenges for AI-Assisted Decision-Making. *Perspectives on Psychological Science* 2024;19:722–34. <https://doi.org/10.1177/17456916231181102>.
29. Buçinca Z, Malaya MB, Gajos KZ. To Trust or to Think: Cognitive Forcing Functions Can Reduce Overreliance on AI in AI-assisted Decision-making. *Proceedings of the ACM on Human-Computer Interaction* 2021;5. <https://doi.org/10.1145/3449287>.
30. Lee MH, Chew CJ. Understanding the Effect of Counterfactual Explanations on Trust and Reliance on AI for Human-AI Collaborative Clinical Decision Making. *Proceedings of the ACM on Human-Computer Interaction* 2023;7. <https://doi.org/10.1145/3610218>.
31. Vaccaro M, Almaatouq A, Malone T. When combinations of humans and AI are useful: A systematic review and meta-analysis. *Nature Human Behaviour* 2024;8:2293–303. <https://doi.org/10.1038/s41562-024-02024-1>.

FUNDING

The authors received no funding for the development of this research.

CONFLICT OF INTEREST

The authors declare that there is no conflict of interest.

AUTHORSHIP CONTRIBUTIONS

Conceptualization: José de Jesús Reyes Machain

Data curation: Manuel Iván Tostado Ramírez

Formal analysis: Víctor Manuel Martínez García

Investigation: José de Jesús Reyes Machain

Methodology: Yennifer Díaz Romero

Project administration: Pedro Antonio Valdez Lizárraga

Software: Pedro Antonio Valdez Lizárraga

Supervision: Manuel Iván Tostado Ramírez

Validation: Víctor Manuel Martínez García

Visualization: Yennifer Díaz Romero

Writing – José de Jesús Reyes Machain

Writing – review and editing: Manuel Iván Tostado Ramírez