Psicothema

Volumen 38, no. 3
ISSN 0214–9915 • eISSN 1886-144X
Colegio Oficial de Psicología del Principado de Asturias

Psicothema is indexed by Social Sciences Citation Index (WOS), Scopus, Google Scholar, SciELO, Dialnet, EBSCO Essentials Academic Search Premier, PubMed, PsycINFO, IBECS, Redinet, Psicodoc, Pubpsych, Fuente Académica Plus, IBZ Online, Periodicals Index Online, MEDLINE, EMBASE, ERIH PLUS, Latindex, MIAR, CARHUS Plus+ 2018, Rebiun, DOAJ, & Crossref.

D.L. AS 3779-1989 ISSN: 0214 - 9915 CODEN PSOTEG

This paper meets the requirements of ISO 9706:1994, Information & documentation – Paper for documents – Requirements for permanence, effective with Volume 7, Issue 2, 1995.

Publisher address:
Colegio Oficial de Psicología del Principado de Asturias
Ildefonso Sánchez del Río, 4 - 1º B
33001 Oviedo (Spain)
Tel.: +34 985 28 57 78 • Fax: +34 985 28 13 74
E-mail: psicothema@cop.es • http://www.psicothema.com

Psicothema

PUBLICATION GUIDELINES

Psicothema publishes empirical work in English which is done with methodological rigor and which contributes to the progress of any field of scientific psychology. As an exception, the Editorial Board may accept publication of work in Spanish if the content justifies such a decision. Theoretical work may also be accepted, if requested by the Editorial Board, with preference given to articles that engage with critical research issues or which discuss controversial approaches.

Submission of articles

1. Articles should be submitted via the journal's web page: www.psicothema.com (Authors section – submission of articles): http://www.psicothema.es/submit

2. Submissions must comply with the rules for preparation and publication of articles, as well as the ethical standards specified below.

3. Studies must be unpublished. Articles which have been fully or partially published elsewhere will not be accepted, nor will articles that are in the process of publication or which have been submitted to other journals for review. It will be assumed that all those who appear as authors have agreed to do so, and all those cited for personal correspondence have consented.

4. The activities described in the published articles will comply with generally accepted ethical standards and criteria, both in terms of work with human beings and animal experimentation, as well as all aspects of professional and publishing ethics.

5. The original work may be submitted in Spanish initially and receipt will be acknowledged immediately. If so, and if it is accepted, the authors will be responsible for translating it into English for publication.

6. Authors may only submit one article for consideration by Psicothema per year.

7. Names and surnames should be entered on the platform in the form they will be cited (a single surname, two separate surnames, hyphenated surnames, etc.). The affiliation of all authors must be indicated. **A maximum of two affiliations per author may be indicated. Affiliations must follow the format "entity or university (country, in English)"**. Do not include information about research groups or departments. Only one person may appear as corresponding author, who will be responsible for ensuring that the author names, order, and affiliations are correct.

8. Authors should suggest three people who they believe would be suitable reviewers for the article, clearly indicating their institutional affiliation and email address. Authors may also indicate people who, for whatever reason, they do not wish to be involved in the review process for their work. Pleas bear in mind the recommendations from the Committee on Publication Ethics (COPE) when suggesting the three reviewers https://publicationethics.org/files/Ethical_guidelines_for_peer_reviewers_0.pdf

9. Manuscripts are screened by the Editorial Board to assess relevance and interest for the journal and whether it follows the rules. Articles must faithfully conform to the editorial rules and fall within the editorial scope of the journal. It is a necessary, though not sufficient, condition that articles must comply with the rules for publication. Articles which do not follow Psicothema's rules will be rejected. In general, within around 10 days the Editorial Board will communicate a decision of interest to begin the review process.

10. Psicothema is only able to publish about 10% of the manuscripts it receives, which is why we apply a very rigorous screening and selection system. Many submissions are considered **non-priorities** by the Editorial Board without being sent for review.

11. If an article passes the Editorial Board screening, it will be sent to a minimum of two reviewers to evaluate its scientific quality. The journal has a **policy of "double blind" reviews**, meaning that both authors and reviewers are anonymous during the review process. To that end, manuscripts must not contain information that would allow the authors to be identified. Most reviewers report back within the agreed three week period. The review process, from receiving an article to the decision to modify it or reject it, usually takes around two months.

12. If, after receiving the reviewers' reports, the Editorial Board decides that the article needs "modifications" to be published, the authors should send the modifications in the requested format together with a point-by-point response to all the comments made by the reviewers and the Editorial Board. Failure to respond in the required format within the set timescale will lead to the article being rejected and removed from the management platform, with no possibility of re-submission.

13. The Editorial Board is responsible for the final decision to accept the article for publication or not. The editors usually make their decisions as quickly as possible once they have received all the necessary reports.

14. After an article has been accepted, and before publication, the authors must sign a copyright agreement. Printing rights and rights of reproduction in any format or medium belong to Psicothema, who will not reject any reasonable request from authors for permission to reproduce their contributions.

15. It is the authors' responsibility to obtain relevant permissions to reproduce copyright-protected material. They are also responsible for disclosing possible conflicts of interest, declaring sources of funding and their participation in the research, and providing access, where necessary, to databases, procedure manuals, scores, and other experimental material that may be relevant. These aspects must be declared in the articles, as described below.

For any questions or clarifications, the journal can be contacted via the email address psicothema@cop.es

Manuscript preparation

1. File format: Articles must be sent in DOC or DOCX format. Microsoft Word documents must not be locked or password-protected, they should not have comments in the margins or information that might reveal the authors' identities. The file should be anonymised in "file properties" so that author information does not appear.

2. Length: The maximum length for articles is **6,000 words** (including the title, abstracts, key words, in-text references, acknowledgements, figures, and tables). The 6,000 word limit **does not include the list of references**. If authors wish to provide supplementary material, the article should include a unique, persistent web link (see point 18 about supplementary material).

3. Format: The articles must be in Microsoft Word format, using **12-point Times New Roman**, in a single column with 3 cm margins, paragraphs left-aligned and double spaced (except for tables and figures which may use single spacing). Page numbers must be included in the lower right corner. Limit sections and subsections to three levels of headings and follow the recommendations in the APA 7th edition about "Sentence case" in the list of references. Psicothema does not allow footnotes, annexes, or appendices. Any such content should be incorporated appropriately into the text (see point 18 about supplementary material).

4. Language: Although articles may be submitted and reviewed in Spanish, accepted articles are usually published in English. Once articles are accepted, the authors must provide an English translation of the reviewed article, within the indicated timeframe, for publication. Psicothema accepts American and British English, but not a mix of the two. Any text in English must be of appropriate professional quality, which will be reviewed by a professional native-speaking translator. Following that review, Psicothema may suggest changes, or if necessary, request a new translation or revision of the translation, the costs of which will be borne by the article's authors.

5. Title page: The first page of the article contains the article title in English and in Spanish, the running title (in English), the total number of words

in the article (not counting references) and a **declaration of authorship, originality and the fact that the work is previously unpublished**. This declaration is obligatory as one of the measures the journal takes to avoid plagiarism. The submitted text must be anonymized, avoiding use of the authors names or anonymizing other possible references that may identify them. Follow the APA 7th edition rules for capitalization of titles and subtitles (i.e., "Title case"). Use upper case for the first letter of all nouns, verbs, adjectives, adverbs, pronouns, and any word longer than three letters.

6. **Title:** The title should be short, descriptive, clear, accurate, and easy to read. It should engage the reader's interest and name variables or topics addressed. Ensure that the main key phrase of the topic is in the article title and avoid superfluous words. Remember that searches normally use key phrases rather than individual words (for example, "mental health in people with disability" not just "health"). Try to include the topic at the start of the title. If the title is "creative", add a more descriptive subtitle after a colon. A descriptive title will help the article to be found in databases. The Editorial Board reserves the right to change titles and abstracts of articles accepted for publication in order to follow the above rules and enhance the article's impact and dissemination.

7. **Abstracts and key words:** the second page of the article contains the abstracts (in Spanish and English) and 3-5 key words or terms. Abstracts must be no more than 200 words and **structured** in four sections: Background, Method, Results, and Conclusions. The abstract should be a single paragraph with these titles in bold, followed by colons and upper case. The key words cover essential elements of the paper such as the research topic, population, method, or application of the results. Avoid general terms and empty words (pronouns, adverbs etc.), or redundant words such as analysis, description, research, etc. Nouns are preferred. Pay particular attention to selection of key words as they are used to index the article.

8. **Article:** The article introduction begins on the third page. The introductory section should not include the article title, or the subtitle "Introduction", or subsections. Following that, the "Method" section should contain the following subsections "Participants", "Instruments", "Procedure", and "Data Analysis", and no others, in no other order, and with no other titles. Where appropriate, in the procedure section it is obligatory to provide information about ethical aspects of the study, the ethics committee that approved the study and the reference code (anonymized during the review process). For research with children, express mention must be made about obtaining informed consent. Pay particular attention to the APA rules about the presentation of statistical and mathematical results in the text, as well as tables and figures. At the end, there should be a single "Discussion" section which should include both discussion along with limitations and conclusions of the study. The discussion section should not have any subsections.

9. **Declaration of author contributions:** Where there is more than one author, there must be a declaration of responsibilities at the end of the article, before the references, specifying what contribution each of the authors made. To specify each author's contribution, use the criteria established by the CRediT taxonomy (Contributor Roles Taxonomy; https://credit.niso.org). Please use the full name of each author as it appears in the manuscript to declare their contributions, followed by the CRediT roles performed. Follow this example: **John White:** Conceptualization, Methodology, Software. **Nuria García-Fernández:** Data curation, Writing - Original draft. **Lucinda Jackson:** Visualization, Investigation. **Laura Gayo:** Supervision, Software, Validation. **Michael Gutiérrez:** Writing - Review and Editing.

   If a group of authors made equal contributions, please also use the CRediT taxonomy to specify their contributions: **John White:** Conceptualization, Writing – Original draft, Writing - Review and Editing. **Lucinda Jackson:** Conceptualization, Writing – Original draft, Writing review and Editing.

   Psicothema does not permit the use of other formulas to indicate equal contributions, such as 'contributed equally to this work ', co-first authors, co-last authors, or co-senior authors.

10. **Corresponding author:** Psicothema allows only **one corresponding author**, who will take primary responsibility for communication with the journal during the manuscript submission, peer review, and publication process, as well as for ensuring providing correct details of authorship (including the names of co-authors, addresses and affiliations), ethics, acknowledgements, sources of funding, conflict of interests, and declarations. The corresponding author is responsible for having ensured that all authors have agreed to be so listed, and have approved the manuscript submission to the journal. After publication, the corresponding author is the point of contact for queries about the published paper. It is their responsibility to inform all co-authors of any matters arising in relation to the published paper and to ensure such matters are dealt with promptly.

11. **Acknowledgements:** any acknowledgements should be included at the end of the text, before the references, in a separate section titled "Acknowledgements".

12. **Sources of Funding:** Priority will be given to work supported by competitive national and international projects. A section titled "Funding" must be included following the "Acknowledgements" section (if one is included) and before the list of references. The "Funding" section must clearly specify the funding body with the assigned code in brackets. It must also be clearly indicated whether the source of funding had any kind of participation in the study. If there was no participation, include the following sentence, "The source of funding did not participate in the design of the study, the data collection, analysis, or interpretation, the writing of the article, or in the decision to submit it for publication". If no funding was received, add the following, "This study did not receive any specific assistance from the public sector, the commercial sector, or non-profit organizations".

13. **Conflict of interests:** Authors must report any economic or personal relationship with other people or organizations that may inappropriately influence their work. If there are none, following the funding section, in a section titled "Conflict of Interest", authors should state: "The author(s) declare(s) that there are no conflicts of interest".

14. **Declaration of availability of data:** The authors should state, in a section titled "Data Availability Statement", whether the research data associated with the article is available and where or under what conditions it may be accessed. They may also include links (where appropriate) to the dataset.

15. **Reference style:** Articles must be written following the guidelines in the 7th edition of the Publication Manual of the American Psychological Association. Articles that do not comply with these rules will be rejected. Some of the requirements are summarized below.

Bibliographical references in the text should include the author's surname and year of publication (in brackets, separated by a comma). If the author's name forms part of the narrative, it should be followed by the year in brackets. If there are more than two authors, only the first author's surname is given, followed by "et al." and the year; if there is confusion, add subsequent authors until the work is clearly identified. In every case, the references in the bibliography must be complete (up to 20 authors). When citing different articles in the same brackets, order them alphabetically. To cite more than one study from the same author or authors from the same year, add the letters a, b, c, as necessary, repeating the year (e.g., 2021a, 2021b).

The list of references at the end of the article must be alphabetical and comply with the following rules:

a) Books: Author (surname, comma, initials of first name(s) and a full stop); if there are various authors, separate them with a comma; before the final author use a comma and "&"; year (in brackets) and full stop, The full title in italics and full stop; finally, the publisher. For example:

Lezak, M., Howieson, D. B., & Loring, D. W. (2004). *Neuropsychological assessment* (4th ed.). Oxford University Press.

b) Chapters of books with various authors, reports from conferences or similar: Author(s); year; title of the work being cited, followed by "In", the director(s), editor(s), or compiler(s) and in brackets Ed., adding an s if plural; the title of the book in italics and in brackets the page numbers of the cited chapter; the publisher. For example:

de Wit, H., & Mitchell, S. H. (2009). Drug effects on delay discounting. In G. J. Madden & W. K. Bickel (Eds.), *Impulsivity: The behavioral and neurological science of discounting* (pp. 213-241). American Psychological Association.

c) Journal articles: Author(s); year; article title; full name of the journal in italics; volume number in italics; issue number in brackets with no space between it and the volume number; first and last page number. The doi should be included in URL format. For example:

Muñiz, J., & Fonseca-Pedrero, E. (2019). Diez pasos para la construcción de un test. *Psicothema, 31*(1), 7-16. https://doi.org/10.7334/psicothema2018.291

For documents that do not have a doi, it is no longer necessary to use "Retrieved from", instead give the URL directly. For example:

Walker, A. (2019, November 14). *Germany avoids recession but growth remains weak*. BBC News. https://www.bbc.com/news/business-50419127

d) Pay particular attention to the rules in the 7th edition of the APA manual for citing work presented in **conferences, doctoral theses, and software,** as well as the rules for the **use of acronyms in text** and in the references section.

e) When the original version of the cited work (book, chapter, or article) **is not in English**, cite the original title and give the English translation in square brackets (with no separation from the original, without using italics).

For further information and other cases, consult the 7th edition of the APA publication manual or the following page: https://apastyle.apa.org/style-grammar-guidelines/references/examples

16. Figures and tables should be included at the end of the manuscript, one per page. They should also follow the APA 7th edition guidelines, be appropriately numbered and cited in the text, indicating approximately where they should be placed. They must have a short, descriptive title that helps understand the content, and follow the APA recommendations about title case, with no full stop. They should be 7 or 14 cm wide and have clear, legible lettering and symbols. Avoid wasted space and make best use of the space available. Figures must be submitted in editable formats, consistent with the format of the rest of the article. If that is not possible, they must have a minimum resolution of 300ppp.

17. Pre-registration of studies and plans of analysis: as a general rule, Psicothema recommends pre-registering submitted studies. If authors have pre-registered studies or plans of analysis, links to that pre-registration should be provided in the article.

18. Supplementary material. Psicothema recommends sharing the data that has been used in the research and supplementary material in institutional or thematic open-access repositories, federated in the European Open Science Cloud (EOSC). Provide a web link if access is to be provided to databases or any other supplementary material, using unique, persistent identifiers.

19. We encourage authors to consult the following standard guidelines when preparing their manuscripts (although due to the multidisciplinary nature of the journal, this is not obligatory):

Case Reports - **CARE**

Diagnostic accuracy - **STARD**

Observational studies - **STROBE** (von Elm et al., 2008), **MQCOM** (Chacón et al., 2019) o **GREOM** (Portell et al., 2015)

Randomized controlled trial - **CONSORT** and **SPIRIT** (Hopewell et al., 2022)

Systematic reviews, meta-analyses – **PRISMA** (Page et al., 2020)

Test adaptation - **International Test Commission Guidelines** (Hernández et al., 2020)

Test development - **Ten steps for test development** (Muñiz & Fonseca, 2019)

Publication of articles

1. Publication rates: Psicothema is an **"open access"** journal. All of the articles will always be free to those who want to read or download them. In order to provide this open access, Psicothema charges a publication fee which the authors or their funders must pay. The price depends on the length of the manuscript. In general, the average price per article is between €180 and €210, based on a mean of 6-7 pages per article, at €30 per laid-out page.

2. Print Proofs: Once an article has been accepted for publication, the contact person will receive an email with the print proofs in PDF format to check and correct spelling-typographical errors. Only minimal corrections can be made to the content of the article once it has been accepted. **Substantial modifications and changes will not be accepted** other that correcting printing or translation errors, possible errors detected during the review process, or incorporating suggestions made by the Editorial Board. No changes will be accepted in this phase to authorship, addition of new affiliations, or details such as including research groups or departments. Galley proofs should be checked carefully, following the instructions provided with them, to confirm that they match the accepted original. Corrected proofs should be returned within the requested timeframe (48-72 hours). Corrections must be made in the PDF file itself, no other means of correction will be accepted. It is vital to check that names, surnames, ORCID codes, and affiliations are all correct in this stage. The corresponding author is responsible for gaining approval from all co-authors for the corrected print proofs. If the proof article is not reviewed within the timeframe or manner specified, that version of the article will be published and subsequent changes or corrections will not be possible.

3. Published version: Once the edition of Psicothema containing the article is published, the author will receive a copy of their article in PDF format. The final version typeset by Psicothema will be available online via DOI. We strongly recommend sharing the final version published by Psicothema on social networks, (Facebook, Twitter, LinkedIn…), university and public repositories (Mendeley, Cosis…), scientific social networks (ResearchGate, Academia.edu, Kudos ...), personal and institutional websites, blogs, Google Scholar, ORCID, Web of Science ResarcherID, ScopusID...

Ethical standards

Psicothema is committed to the scientific community to ensure the ethical and quality standards of published articles. Its references are the "Core practices" defined by the **Committee on Publication Ethics (COPE)** for journal editors, the **American Psychological Association (APA)** Code of Conduct, and the Code of Ethics for Psychology from the **Spanish General Council of Psychology**.

Use of inclusive, non-sexist language. At Psicothema, we are firmly committed to equality and respect for all, recognizing and appreciating diversity. For this reason, authors should ensure that they use bias-free language, avoid stereotypes, and engage with inclusive, non-sexist language, albeit prioritizing grammatical correctness, economy of language, and accuracy, given the limitations of space. Pay particular attention to the presentation of data, so that participants' characteristics are described and analysed properly, without presenting information that is irrelevant to testing hypotheses, achieving objectives, or presenting results of the study. Avoid condescending, obsolete, or inappropriate language, as well as the use of labels related to stereotypes. We recommend reporting where potential gender differences are found in the results.

Responsible authorship. Psicothema promotes transparency via the declaration of authors' contributions. All signatories must have made substantial contributions in each of the following aspects: (1) conception and design of the study, or data acquisition, or analysis and interpretation of data, (2) drafting the article or critical review of the intellectual content, and (3) final approval of the submitted version. The list and order of authors should be carefully reviewed before the initial submission of the article. Any addition, removal, or re-ordering must be done before the article is accepted, with the approval of the Psicothema Editorial Board and the consent of all named authors. A form for this is available on request.

Open science. To facilitate the reproducibility of research and reuse of data, code, types of software, models, algorithms, protocols, methods, and any other useful material related to the project should be shared.

We recommend that authors publish the original study data in public open-access repositories online, such as FigShare (http://figshare.com), Mendeley Data (https://data.mendeley.com/), Zenodo (http://zenodo.org/), DataHub (http://datahub.io) and DANS (http://www.dans.knaw.nl/). Where data or supplementary material is shared, a corresponding reference should be included in the manuscript and the list of references, using unique, persistent identifiers.

Funding sources. In the acknowledgements section, authors should include data on the organizations that provided economic funding for the study or preparation of the article, and briefly describe the role any funding body played in designing the study, data collection, analysis, and interpretation, writing the article, or the decision to submit it for publication. If there was no participation from the funding body, this should be indicated as suggested in the "Preparation of Articles" section. The author responsible for submitting the article should include this metadata at the time of submission in the corresponding section.

San Francisco declaration on research assessment (DORA). As part of its commitment to open knowledge, Psicothema follows this initiative because it shares the need to address the quality assessment of scientific articles (not only the journals in which they are published), to consider the value and impact of all research outputs (including data and software), and to consider the societal impact of research from a broader perspective (including qualitative indicators, such as the influence on scientific policies and practices, together with a responsible use of quantitative indicators). To this end, it is committed to remove restrictions on the number of references that can be included in the bibliography, not counting them as part of the maximum number of words, to encourage responsible authorship practices and to provide information about the specific contributions of each author (CRediT), to mandate the citation of primary literature in favor of reviews in order to give credit to the group(s) who first reported a finding, and to make available a variety of journal-based metrics and article-level metrics (PlumX).

Good publishing practice in gender equality. Psicothema is committed to gender policies that lead to real equality between men and women in society through various actions: (1) pursuing equal proportions of women and men in the editorial team, as well as in those who review the articles; (2) recommending the use of inclusive language in scientific articles; (3) recommending that articles report whether the original study data considered sex or gender in order to identify possible differences; and (4) including the full names of the authors of published articles. To that end, authors must include their full names (not just first initials) in the metadata, which will appear in the published articles.

Authors' rights

Acknowledgement of receipt. Receipt of the article will be immediately communicated to the authors by email.

Screening. Articles will be reviewed by the Editor-in-Chief, Executive Editor, Managing Editor, and the Associate Editors. The editorial team may directly reject studies if, in their opinion, they do not follow the journal's publication rules, do not meet the minimum requirements, or do not fit the journal's objectives or priorities.

Review. Once past the Editorial Board screening, the articles will be reviewed by external reviewers and by the Associate Editor responsible for managing the article. The Associate Editor and the Managing Editor will consider the external reviewers' reports and will make the final decision on publication.

Reasoned reply. Except in cases of articles considered to be "non-priority" in the initial screening phase, authors will be given a reasoned response about the Editorial Board's final decision when that involves rejection (i.e., articles rejected after the peer review phase).

Confidentiality. Authorship of articles received will be kept anonymous and the evaluation process will be confidential, we commit to not disseminating the article more than necessary for the evaluation process and until the article is accepted for publication.

Use of data. The members of the Editorial Board will not use the results of unpublished work without the express consent of the authors.

Declaration of privacy. The names and email addresses provided to Psicothema will only be used for purposes established in the journal, they will not be given to third parties or used for commercial purposes.

Complaints and claims. Efforts will be made to respond to and resolve complaints and claims quickly and constructively. Complaints or claims should be send by email to psicothema@cop.es, clearly and accurately specifying the nature of the complaint, the contact details of the person making it, and sufficient data to demonstrate any possible violation of the journal's declaration of ethics. Complaints about published content must be made as soon as possible after publication, and after having first contacted the corresponding authors to try and find a direct resolution. Psicothema may be contacted where it is not appropriate to contact the authors, if the authors do not respond, or if they do not resolve the issue. If possible, documentation must be included as evidence of the situation. Psicothema will acknowledge receipt of the complaint by email, and may request additional information or documentation for clarification. Depending on the nature or complexity of the issue, if the content is reviewed and sufficiently documented, the Editorial Board will study the case and make any decision in accordance with the directives of the Committee of Publishing Ethics (COPE). The Editor-in-chief will make the final decision and a response will be sent by email. Other people and institutions will be consulted as necessary, including university authorities or subject-matter experts, and legal advice may be sought if the complaint has legal implications. Complainants will have to expressly request that a complaint be treated confidentially and the Editorial Board will do so as far as appropriate and in line with our management processes. It is possible that complainants will not receive any information about the state of any investigation until a final decision is reached, and it is important to bear in mind that investigations may take some time. Complaints that are outside Psicothema scope or that are presented in an offensive, threatening, or defamatory manner will be dismissed. Personal criticism or comments are not acceptable. Communication will be terminated if it is not cordial and respectful, or if there is persistent vague or unfounded complaint. Psicothema reserves the right to take appropriate legal measures if a complainant insists on a complaint that is unfounded, false or malicious.

Authors' responsibilities

Editorial rules. Authors should read and accept the editorial rules and journal's instructions before sending a manuscript. While the article is undergoing the evaluation process at Psicothema, it must not be in any evaluation process at other journals.

Ethical rules. Authors must comply with the ethical standards specified in the Psicothema rules for authors.

License for public communication. The authors cede to Psicothema the public communication rights of their article for free dissemination through the internet, portals and electronic devices, through its free provision to users for online consultation, printing, download and archive, guaranteeing free, open access to the publication.

Publication licence. The authors accept the Psicothema copyright policy and cede it the right of publication. Psicothema publishes its articles under CC-BY-NC-ND license.

Reviewers' responsibilities

Editorial rules. Reviewers must read and accept the journal's editorial rules and instructions before reviewing an article. They must also follow the COPE ethical directives for reviewers.

Professional responsibility. Reviewers must only accept articles for review for which they have sufficient knowledge to perform a proper review.

Conflict of interests. Reviewers will constructively and impartially review articles for which they consider themselves qualified, abstaining from reviewing articles in which there might be a conflict of interest.

Confidentiality. Reviewers will respect the confidentiality of the review process and will not use information obtained during the peer review process for personal gain or to others' advantage, or to discredit or disadvantage others. They will not involve other people in the review process without the authorization of Psicothema.

Suspicion of a breach of ethics. Reviewers will inform the Editorial Board if they detect poor practice, fraud, plagiarism, or self-plagiarism, as well as any other irregularity related to research or publication ethics.

Deadline for reviews. Reviewers will commit to meeting the review timeframes set by Psicothema, informing the Editorial Board if they need additional time or are unable to send a report after having accepted a review request.

Preparation of the report. The format of Psicothema's review report is open, but reviewers must use a short scoring rubric. Reviewers must be objective and constructive in their reviews, offering feedback that will help authors improve their articles. Reviewers must make fair, impartial, constructive assessments of the article's strengths and weaknesses, and avoid disparaging personal comments or baseless accusations. They must not suggest that authors add references to the reviewer's own work (or that of colleagues) just to increase the number of citations or raise the visibility of their work or the work of associated; suggestions must only be based on valid academic reasons.

The journal's responsibilities

The Editorial Board is not responsible for the ideas or opinions expressed by the authors in the journal articles or the reviewers in their reports. The opinions and facts expressed in the articles are solely and exclusively the authors' responsibility and do not represent the journal's opinions or scientific policies. The editorial organization is not responsible in any case for the credibility or authenticity of the articles.

Psicothema will strive to avoid scientific fraud, which includes fabrication, falsification, or omission of data; plagiarism; duplicate publications; and authorial conflicts. Particular attention in plagiarism is paid to avoiding passing others' work off as one's own, co-opting others' ideas without recognition, giving incorrect information about the source of a reference, and paraphrasing a source without mentioning it. Detection of **fraud or plagiarism** will lead to the rejection of the submitted or published article.

The Psicothema Editorial Board undertakes to ensure that everyone involved (authors, reviewers, editors, and journal management) comply with the expected ethical standards in every phase of the publishing process, from reception to publication of an article, basing this on the recommendations from **The Committee on Publication Ethics (COPE)** to resolve possible conflicts.

The readers' rights

Readers have the right to read all articles published in Psicothema for free immediately after their publication.

Updating published articles

Psicothema is committed to correcting important scientific errors or ethical issues in published articles. In order to be transparent about any change, the following criteria and procedures have been established for updating our published articles.

Minor errors. Minor errors that do not affect the readability or meaning, such as spelling, grammar, or layout mistakes are not sufficient for and do not justify an update, regardless of the source of the error.

Metadata errors. Requests to correct errors in an article's metadata (for example, title, author name, abstract) must be made during the galley correction process. Once an article has been published, corrections can only be made if the Editorial Board believes the request to be reasonable and important. Once approved, the article will be updated and republished on the Psicothema website, with notification to relevant databases.

Author name and affiliation. Authors must make any desired changes to author names, surnames and affiliations during the galley correction process. Once the article has been published, no changes will be made without valid, convincing reasons, especially if the ORCID code has been supplied correctly. Changes to names after publication will only be in exceptional cases where the authors adopt a new name (such as for marriage or after gender transition) and want it updated. In such cases the Committee on Publication Ethics (COPE) recommendations will be followed.

Corrections. Requests may be submitted to correct errors that affect scientific interpretation. Once a request is approved, the article will be updated and re-published on the Psicothema website, together with a notice of correction. This notice will be a separate publication with a link to the updated article in the most recent edition of the journal, in order to notify readers that there has been a significant chance to the article and that the revised version is available on the website. Relevant databases will then be notified about the update.

Retractions. If an article needs to be retracted from the research literature due to inadvertent errors during the review process, serious ethical violations, fabrication of data, plagiarism, or other reasons that threaten the integrity of the publication, Psicothema will follow the recommendations from the Committee on Publication Ethics (COPE) for retractions. In this case, the original publication will be amended with a "RETRACTED" mark but will remain available on the Psicothema website for future reference. Retractions will be published with the same authorship and affiliation as the retracted article, so that the notice and the original retracted article may be properly found in indexing databases. The retraction notice will be published in the most current edition of the journal. Partial retractions may be published in cases where results are partly incorrect. An article will only be removed completely from the Psicothema website and indexing databases in very exceptional circumstances, where leaving it online would constitute an illegal act or could cause significant harm.

Expression of concern. Psicothema may publish such a concern if the investigation about supposed bad conduct in research is inconclusive, complex, or very prolonged. In this case, a Psicothema editor may choose this option, detailing their concerns point by point, and any actions ongoing.

Comments and responses. Psicothema will only exceptionally publish comments on or responses to articles published in the journal, when the comments (i.e., short letters to the editors from readers who wish to publicly question a specific article) affect the editorial content of an article published in the journal, contain evidence of a claim, and the result of an investigation by the editorial team does not lead to rejecting the criticism, or correction or retraction of the article. In these exceptional cases, once the comment has been approved for peer review, the editorial team will contact the authors of the article in question and invite them to reply to the comment. The reply will allow the authors to publicly respond to the concerns raised. If the authors do not provide a reply within the set timescale or decide not to reply, the comment will be published along with a note explaining the absence of a response. Both comments and replies will be reviewed to ensure that the comment addresses significant aspects of the original article without becoming essentially a new article, the response directly addresses the concerns without evasion, and the tone of both is appropriate for a scientific journal. Only one round of comments and responses will be facilitated about any single article. Nonetheless, Psicothema recommends that readers direct their comments directly to the authors involved and use alternative forums for additional public discussion.

Archive and Digital Preservation Policies

Conditions for self-archiving preprints. Preprint versions of articles (the version initially submitted to the journal) may be shared at any time anywhere. Sharing an article (without review) on a preprint server, for example, is not considered prior publication. Because of that, before final publication, we recommend that authors self-archive the preprint

version on personal or institutional websites, scientific social networks, repositories, reference managers… Once published, if the preprint remains available in a repository, it must be specified that, "This is an electronic version of an article published on Psicothema (year). The final version is available on the official web page", also highlighting the full reference to the published article, including its DOI.

Psicothema's preservation policy. Psicothema publications are available on the journal website and in international open-access repositories online such as SciELO and Dialnet. Psicothema focuses on disseminating content and making it accessible through indexing services. Online access is free, while the printed version is subscription access. In addition, Psicothema allows self-archiving of preprint, post-print, and editorial version (immediately after publication). Exploitation rights and Psicothema's self-archiving permissions may be found at https://dulcinea.opensciencespain.org/ficha1034.

Conditions for self-archiving postprint articles. Authors are encouraged to share the final version of their accepted articles (the authors' version including changes suggested by reviewers and editors) on social networks and repositories until the editorial version is published in an edition of the journal. It must be expressly indicated that this is an article "in press" in the Psicothema journal. Once the editorial version is published, if the postprint is still available in a repository, it must be specified that, "This is an electronic version of an article published on Psicothema (year). The final version is available on the official web page", also highlighting the full reference to the published article, including its DOI.

Archive. The journal undertakes to provide XML metadata or in other specific formats, immediately after its publication and within three months in order to promote its dissemination in databases. Psicothema uses various national and international repositories: Clarivate Analytics, Scopus, Google Scholar, S2ciELO, Dialnet, EBSCO Essentials Academic Search Premier, PubMed, PsycINFO, IBECS, Redinet, Psicodoc, Pubpsych, Fuente Académica Plus, IBZ Online, Periodicals Index Online, MEDLINE, EMBASE, ERIH PLUS, Latindex, MIAR, CARHUS Plus+ 2018, Rebiun, DOAJ, & Crossref.

Digital preservation. In order to preserve permanent access to digital objects hosted on its own servers, Psicothema makes backups, monitors the technological environment to foresee possible migrations of obsolete formats or software, preserves digital metadata, uses DOI Digital Object Identifier) and ORCID. All articles published in Psicothema are also hosted and available in the institutional repository of the Universidad de Oviedo (REUNIDO). The articles published on Psicothema's website are available in easily reproducible format (PDF).

Anti-Plagiarism Policy

In compliance with our code of ethics and in order to guarantee the originality of the manuscripts submitted for evaluation, Psicothema applies anti-plagiarism software to all manuscripts that meet the minimum criteria of the preliminary review and are to be subjected to the review process.

The submission will be rejected in case of detecting practices of plagiarism or scientific fraud, either during the preliminary review by the editorial committee or once the peer review has started.

Articles

Correction

Psicothema

Volumen 38, no. 3

PUBLISHED BY

COLEGIO OFICIAL DE **PSICOLOGÍA** DEL PRINCIPADO DE ASTURIAS

MEMBER OF

Consejo General de la Psicología ESPAÑA

SPONSORED BY

CAJA RURAL DE ASTURIAS

SEPCyS
Sociedad Española para el avance de la Psicología Clínica y de la Salud. Siglo XXI

GRUPO EDITORIAL DE PSICOFUNDACIÓN

Psicothema (2026) 38(3) 113-122

Psicothema

https://www.psicothema.com/es • eISSN 1886-144X

Colegio Oficial de Psicología del Principado de Asturias

Psicothema

Article

A Multilevel Framework for the Safe and Ethical Use of AI in Mental Health Practice

Ignacio Pedrosa

International University of La Rioja (UNIR) (Spain)

ARTICLE INFO

Received: 04/11/2025
Accepted: 05/03/2026

Keywords:
Digital mental health
Psychological ethics
Artificial intelligence
Psychological regulation
Digital wellbeing

ABSTRACT

Background: Incorporating Artificial Intelligence (AI) into mental health practice represents a profound transformation of psychological care, enabling novel forms of assessment, intervention, and support. However, this shift also introduces substantial ethical and psychological risks, including erosion of autonomy, algorithmic bias, threats to digital rights, and strain on clinical judgment and therapeutic relationships. This article proposes PRISM, a multilevel framework designed to guide the ethical integration of AI in mental health across individual, technological, and institutional dimensions. **Method:** An integrative conceptual review was conducted drawing on Self-Determination Theory, information ethics, data justice, and international guidelines from the APA, WHO, EU, and UNESCO. The framework introduces the concept of psychological regulation of AI, extending ethical oversight to the subjective and relational impacts of algorithmic mediation. **Results:** PRISM articulates four operational principles (digital autonomy, affective transparency, algorithmic justice, and relational care) along with practical tools, including psychological impact indicators and a Psychological Model Card, to support ethical auditing and clinical digital wellbeing. **Conclusions:** A psychology-based regulatory approach is essential to ensure dignity, autonomy, and ethical sustainability in digital mental health. PRISM provides an operational bridge between technological governance and psychological practice, supporting responsible, person-centred AI use in clinical contexts.

Marco Multinivel para el Uso Seguro y Ético de la Inteligencia Artificial en la Práctica de la Salud Mental

Palabras clave:
Salud mental digital
Ética psicológica
Inteligencia artificial
Regulación psicológica
Bienestar digital

RESUMEN

Antecedentes: La integración de la Inteligencia Artificial (IA) en la práctica de la salud mental representa una transformación de la atención psicológica, ofreciendo nuevas formas de evaluación, intervención y apoyo. Pero también introduce riesgos éticos y psicológicos sustanciales como la erosión de la autonomía, el sesgo algorítmico, amenazas a los derechos digitales y la tensión en el juicio clínico y las relaciones terapéuticas. Se propone un marco multinivel (PRISM) diseñado para guiar la integración ética de la IA en la salud mental mediante dimensiones individuales, tecnológicas e institucionales. **Método**: Se realizó una revisión conceptual integradora basada en la Teoría de la Autodeterminación, ética de la información, justicia de datos y directrices internacionales (APA, OMS, UE y UNESCO). **Resultado**: PRISM articula cuatro principios (autonomía digital, transparencia afectiva, justicia algorítmica y cuidado relacional) junto con herramientas prácticas, incluyendo indicadores de impacto psicológico y una ficha modelo, para apoyar la auditoría ética y el bienestar digital clínico. **Conclusiones**: Un enfoque regulador psicológico es fundamental para garantizar la dignidad, autonomía y sostenibilidad ética en la salud mental digital. PRISM proporciona un puente operativo entre la gobernanza tecnológica y la práctica psicológica, apoyando el uso responsable y centrado en la persona en contextos clínicos.

Cite as: Pedrosa, I. (2026). A multilevel framework for the safe and ethical use of AI in mental health practice. *Psicothema, 38*(3), 113-122. https://doi.org/10.70478/psicothema.2026.38.11
Corresponding author: Ignacio Pedrosa, ignacio.pedrosa@unir.net

Over the past two decades, mental health has experienced accelerated digitalization. The advent of Artificial Intelligence (AI) has not only represented a technological revolution but also an epistemological and ethical reconfiguration of applied psychology (Dehbozorgi et al., 2025). The increasing deployment of platforms, mobile applications, machine learning approaches, simulated patients, and therapeutic chatbots has transformed psychological literacy, assessment, prevention, and treatment (Sanz et al., 2025; Torous et al., 2025; Wang et al., 2025).

Recent studies have introduced the concept of digital compassion, defined as the capacity of AI systems to modulate empathetic responses and emotional support in mental health contexts without replacing human interaction (Abou Hashish, 2025; Wiljer et al., 2019). Unlike earlier digital tools, such as traditional smartphone apps, which functioned primarily as passive instruments, AI operates as a mediating entity capable of simulating cognitive and affective processes, thereby reshaping core psychological practices (Wiljer et al., 2019). This shift calls for a specifically psychological form of regulation, as existing biomedical or legal frameworks (such as the EU AI Act) prioritize systemic safety while neglecting the subjective impact of algorithmic mediation and its effects on the therapeutic relationship.

Current AI applications in mental health include test construction, assessment, therapeutic support, and clinical decision assistance, ranging from risk detection models to conversational agents and documentation systems (Cheng, 2024; Dehbozorgi et al., 2025; Suárez-Álvarez et al., 2026).

Although these tools can enhance accuracy and accessibility, their opacity may undermine patient autonomy and the therapeutic alliance when limitations are not adequately communicated (Theunissen & Browning, 2022).

At the same time, AI offers substantial benefits, including improved diagnostic accuracy, increased access for underserved populations, and reduced administrative burden, allowing greater clinical focus on patient needs (Milasan & Scott-Purdy, 2025; Yıldız, 2025a).

Nevertheless, these advances introduce unresolved risks, such as algorithmic opacity, automated bias, sensitive data governance challenges, and the potential dehumanization of therapeutic relationships (Adler et al., 2024; WHO, 2023).

Gap Between Technological Development and Psychological Reflection

The development of AI has significantly outpaced the capacity of existing ethical and regulatory frameworks to address its psychological implications. Current regulatory efforts, including the EU AI Act (Regulation EU 2024/1689) and WHO guidelines (2023, 2025), remain largely shaped by a technical–biomedical bias that relegates ethical-psychological concerns to a secondary role. The EU AI Act's risk-based approach prioritizes systemic safety but lacks specific criteria for assessing the mental impact of AI systems on user subjectivity or the integrity of the therapeutic relationship.

Similarly, although the APA (2025) and WHO (2025) stress the importance of human oversight, they do not resolve what has been described as the oversight paradox: the inherent contradiction of requiring meaningful human control over agentic AI systems explicitly designed for autonomous goal achievement (Kyrychenko et al., 2026).

This limitation is compounded by the fragmented interaction between the GDPR, the Digital Services Act, and the AI Act, which produces an accountability paradox in clinical contexts (Kyrychenko et al., 2026). A single biased algorithmic output may trigger multiple regulatory responses without providing a coherent mechanism for reporting or mitigating complex psychological harm. By privileging technical transparency over affective transparency current frameworks risk generating moral fatigue among practitioners and mistrust among patients (Theunissen & Browning, 2022).

This regulatory gap underscores the need for a specialized psychological approach capable of integrating disparate legal requirements into a person-centred clinical reality.

Emerging Psychological Effects

Recent studies indicate that sustained interaction with AI-based therapeutic agents produces mixed psychological outcomes (Torous et al., 2025; Wang et al., 2025). While these systems enhance accessibility, skill acquisition, and anonymity, they are also associated with emotional dependency, reduced perceived control, digital moral fatigue, and confusion between human and artificial relationships.

Within this context, algorithmic anxiety (i.e., the distrust and uncertainty surrounding automated decision-making) has emerged as a relevant construct in social and clinical psychology (Li & Liu, 2025). Such subjective effects, often invisible to conventional efficacy metrics, directly affect psychological wellbeing, self-efficacy, and therapeutic adherence, with meaningful consequences for mental health outcomes.

These emerging effects highlight the need for evaluative approaches that extend beyond performance and safety indicators. A psychological regulatory perspective is required to assess AI-mediated mental health interventions through the lenses of subjectivity, relational dynamics, and ethics of care. From this standpoint, four core questions become central: how patient autonomy is preserved in automated environments; what limits should govern algorithmic supervision of behaviour; how explainability and algorithmic justice can be ensured in clinical contexts; and which ethical–digital competencies are required of mental health professionals. Addressing these questions requires an integrative framework capable of translating ethical principles into operational criteria for clinical practice, professional training, and institutional governance in digital mental health.

This work pursues a dual objective. First, it develops an integrative theoretical synthesis of the main ethical-psychological dilemmas associated with the application of AI in mental health. Second, it proposes a model articulated around three interdependent axes (i.e. individual, technological, and institutional) to guide policy design, professional training, and ethical best practices in the digital era. Unlike existing bioethical, medical, or legal frameworks (Tang et al., 2025), no equivalent concept has yet been defined in recent systematic reviews of digital health from a psychological perspective.

The study adopts an integrative conceptual review methodology, designed to synthesize diverse theoretical perspectives and empirical findings into a unified regulatory framework. The search was conducted across three primary high-impact databases: APA PsycInfo, PubMed, and Scopus, supplemented by grey literature from international regulatory bodies (EU, WHO, APA, UNESCO). The search strategy utilized combinations of Boolean operators and keywords, including: ('*Artificial Intelligence*' OR '*Generative AI*') AND ('*Mental Health*' OR '*Psychology*') AND ('*Ethics*' OR

'*Regulation*' OR '*Digital Rights*') AND ('*Algorithmic Bias*' OR '*Autonomy*'). The scope was primarily limited to the period from 2020 to 2025 to ensure the inclusion of the most recent developments in large multi-modal models and international AI governance.

Inclusion criteria focused on a) peer-reviewed articles discussing the psychological impact of AI on the therapeutic relationship or user subjectivity, b) official ethical guidelines from psychological associations, and c) international legislative frameworks for AI health technologies. Exclusion criteria were applied to purely technical or biomedical reports that did not address the human-AI interaction or the ethical-psychological dimension (Hoose & Kralikova, 2024).

Theoretical Framework: Psychology Facing Artificial Intelligence

The development of AI in mental health represents a historical inflection point comparable to the emergence of psychoanalysis or cognitive-behavioural therapy in their respective eras. Unlike previous paradigmatic shifts, psychology now shares agency with non-human systems capable of simulating cognitive and affective functions (Sanz et al., 2025). This unprecedented condition compels a reconsideration of the discipline's epistemological foundations and its role as guarantor of psychic integrity within increasingly digitalized clinical contexts.

Several psychological paradigms provide conceptual tools for understanding human-AI interaction. Self-Determination Theory (Deci & Ryan, 2000) posits that wellbeing depends on the satisfaction of autonomy, competence, and relatedness that AI systems may either support or undermine depending on their design and implementation. This perspective underscores the need to distinguish psychological regulation from broader approaches such as applied ethics or digital bioethics, which primarily prioritize systemic safety and regulatory compliance (Dencik & Sanchez-Monedero, 2022; EU, 2024/1689; Torous et al., 2025). In contrast, psychological regulation evaluates AI-mediated care through its impact on subjectivity, relational dynamics, and basic psychological needs, reframing regulation from external rule compliance to the preservation of psychological integrity in digital clinical environments.

While traditional ethical frameworks focus on systemic risk mitigation, psychological regulation targets specific constructs such as algorithmic anxiety (Yang & Sundar, 2025), the distress arising from unpredictable automated processes, and digital moral fatigue among practitioners (Epstein et al., 2019). By prioritizing understandability as an outcome of transparency and interpretability, this approach seeks to prevent the depersonalization of care (Padalkar et al., 2024) and to safeguard the therapeutic bond in technologically mediated interactions.

From a humanistic perspective, authenticity and empathy remain central to the therapeutic process. The partial delegation of these dimensions to algorithms raises critical questions about relational authenticity, perceived agency, and the limits of technological mediation in clinical care.

At a structural level, Social Cognitive Theory (Bandura, 2001) introduces the notion of digital self-efficacy, understood as perceived control over automated processes. Similarly, Information Ethics (Floridi, 2019) emphasizes the protection of autonomy, dignity, and agency as intrinsic values of the informational ecosystem, while the Data Justice perspective (Dencik & Sanchez-Monedero, 2022) and international guidelines (WHO, 2023) extend these concerns to equity and democratic participation. Together, these frameworks provide an essential ethical–psychological foundation for embedding human values into technological design.

These perspectives converge on a central premise: AI systems are not neutral. Their architectures embody implicit values that shape identity, perceived control, emotional experience, and professional responsibility. Consequently, AI transforms not only diagnostic and therapeutic procedures but also the structure of the clinical relationship itself. Before proposing a regulatory framework, it is therefore essential to analyse the ethical and psychological dilemmas that emerge when psychological practice becomes technologically mediated.

Emerging Ethical and Psychological Dilemmas

Building on these theoretical foundations, the following sections identify the primary ethical and psychological challenges that justify a dedicated regulatory framework.

Algorithmic Bias and Psychological Justice

A central ethical concern is algorithmic bias, which may replicate historical discrimination embedded in clinical and population datasets. Adler et al. (2024) showed that AI models in psychiatric diagnostics often underrepresent women and racial minorities, reducing diagnostic accuracy and undermining patients' epistemic trust.

Psychological regulation should include bias audits addressing both statistical and experiential dimensions, posing questions such as: *How does the patient perceive the system's fairness?* and *How does this perception influence therapeutic adherence and the alliance?*

Autonomy and Digital Consent

Autonomy, a cornerstone of psychological ethics, is challenged by the technical complexity of AI. Many users consent to data processing without a full understanding of algorithmic inference mechanisms, highlighting the need for cognitively and emotionally intelligible consent protocols (APA, 2025).

Therapeutic Relationship and Dehumanization

Studies on therapeutic chatbots (e.g., Replika) indicate that while accessibility improves, users may develop emotional dependency and experience reduced relational authenticity (Laestadius et al., 2024). Misalignment between AI responses and the user's affective-cognitive context exacerbates this disconnect (Loechner et al., 2025).

This emotional disconnect affects not only patients but also mental health professionals, who must manage interactions mediated by systems incapable of genuine affective responsiveness. This not only affects patients but also contributes to digital moral fatigue in practitioners, who must mediate interactions with systems incapable of genuine empathy. Ethical frameworks should therefore integrate the intersubjective experience of both patients and clinicians, combining the protection of patient wellbeing with the ethical supervision of algorithmic systems.

Privacy, Surveillance, and Psychological Wellbeing

Privacy research demonstrates that perceived constant surveillance negatively impacts emotional wellbeing and cognitive performance (Solove, 2021). Mental health monitoring tools analyzing text or voice can provoke exposure anxiety, heightening self-consciousness and behavioural inhibition.

Jointly, these dilemmas reveal that regulatory challenges are not merely technical, they are profoundly psychological. They underscore the necessity of a framework that safeguards autonomy, equity, and wellbeing while accounting for the relational dynamics inherent in AI-mediated care.

The PRISM Framework: A Psychological Regulatory Model for AI in Mental Health

The regulation of AI in mental health cannot be restricted to legal or technical mechanisms alone. It requires the integration of psychological principles that guide the design, implementation, and evaluation of intelligent systems in ways that uphold human dignity and wellbeing.

The proposed framework, hereafter referred to as the PRISM Framework -Psychological Regulation and Integrity in Smart Mental Health-, is grounded in four integrative principles, derived from both classical psychological ethics and contemporary research on AI:

- **Digital autonomy:** the individual's capacity to understand and make informed decisions regarding technologies that influence their mind, emotions, or behaviour.
- **Affective transparency:** the requirement for AI systems to clearly communicate their limitations in emotional understanding, to prevent false attributions of empathy or intentionality.
- **Algorithmic justice:** the ethical obligation to prevent bias through systematic auditing and the participatory inclusion of vulnerable or underrepresented groups in the design and validation process.
- **Relational care:** the preservation of authentic human presence and the quality of the therapeutic relationship as central components of psychological intervention.

This framework is not merely a normative proposal, but a conceptually grounded model based on the intersection of SDT and aspects linked to transparency and interpretability for understandability (Deci & Ryan, 2000; Joyce et al., 2023). The theoretical rigor of PRISM resides in the premise that psychological well-being in digital environments depends on the satisfaction of innate needs for autonomy, competence, and relatedness (Deci & Ryan, 2000). This framework primarily addresses the professional stakeholder, focusing on the ethical and psychological challenges clinicians face when interacting with AI.

Together, these principles constitute the foundation of the PRISM Framework, which is structured around three interdependent axes (i.e. individual, technological, and institutional) forming an ethical-psychological architecture designed to protect digital rights and promote mental health wellbeing (Table 1). Empirically, the Algorithmic Justice dimension is validated by findings demonstrating that smartphone-sensed behaviours are unreliable predictors across diverse populations. For instance, phone usage patterns that signify depression in younger cohorts can represent healthy engagement in older adults (Adler et al., 2024). This reliability gap requires the Individual Axis of the framework, which integrates trait-based assessments to calibrate trust and mitigate digital moral fatigue (Yang & Sundar, 2025). Furthermore, the Technological Axis adopts the operational definition of understandability. It posits that for a system to be clinically valid, its computational structures must stand in close correspondence with human clinical heuristics, such as abductive and inductive inference (Joyce et al., 2023). By shifting from post-hoc explanations to intrinsic interpretability, PRISM ensures that the mental impact of algorithmic mediation remains traceable and ethically auditable. Finally, the Institutional Axis is supported by implementation science models that emphasize the role of digital navigators in maintaining the human connection, a factor shown to significantly enhance the efficacy of digital interventions compared to unguided tools (Torous et al., 2025).

Individual Axis: Wellbeing, Digital Autonomy, and Agency

This axis represents the core of the proposed framework, as mental health is conceptualized as a dynamic balance among self-determination, perceived competence, and human connection (Deci & Ryan, 2000). The integration of AI introduces novel psychological determinants of digital wellbeing, including the following constructs:

- **Perceived digital autonomy:** individual's sense of control over the intelligent systems with which they interact. A deficit in this domain is associated with algorithmic anxiety and feelings of helplessness (Yang & Lu, 2026).
- **Algorithmic anxiety:** emotional response to the opacity or unpredictability of algorithmic processes, frequently observed in users who delegate clinical decisions to automated systems (Yang & Sundar, 2025).
- **Digital epistemic trust:** individual's willingness to regard system inferences as valid or benevolent. Excessive trust may foster dependency, whereas insufficient trust can result in technological rejection. Recent empirical studies indicate that psychological trust in human-AI collaboration depends more on moral attribution mechanisms and the perceived sense of control than on mere technical accuracy (Gupta et al., 2025).

Table 1
Conceptual Scheme of the PRISM Framework

Level/Axis	Guiding Principle	Key Psychological Components	Regulatory Mechanisms	Expected Outcomes
Individual	Digital Autonomy and Agency	Perceived autonomy, epistemic trust, algorithmic anxiety, moral fatigue	Digital wellbeing assessments, interactive informed consent	Greater perceived control, reduced technological anxiety
Technological	Ethical Design and Explainability	Transparency, accountability, inclusive design, interdisciplinary co-design	Ethical audits, model cards, psychological review of AI systems	Reduction of bias and increased public trust
Institutional	Governance and Digital Justice	Ethical training, shared responsibility, continuous oversight	Ethical-psychological observatories, interdisciplinary committees	Sustainability, social legitimacy, regulatory alignment

Note. Grounded in the socio-technical model of trust, which links user-related, system-related, and contextual factors (Gupta et al., 2025).

- **Digital moral fatigue:** emotional exhaustion experienced by professionals facing tensions between technical efficacy and the ethics of care (Wang et al., 2025).

Within this axis, psychological interventions should incorporate ethical-digital literacy programs in clinical contexts, enabling both patients and professionals to recognize the potential limits and risks of AI. It is also recommended that periodic assessments of digital wellbeing and perceived autonomy be implemented as psychological indicators of care quality.

To support the evaluation of these constructs, Table 2 presents a proposed set of psychological impact indicators for assessing AI in mental health. This adapted metric system operationalizes dimensions such as digital autonomy, algorithmic anxiety, and epistemic trust, providing an empirical foundation for the future validation of assessment instruments.

At the operational level, the APA (2025) recommends including comprehensible digital consent modules, enabling users to visualize, gradually and adaptively, how their data are processed and which decisions are automated.

Technological Axis: Responsible Design and Computational Psychology

The second axis addresses the ethical and psychological requirements that should guide the design, training, and validation of algorithmic models. Artificial intelligence systems involved in psychological processes must adhere to explainability, auditability, and human sensitivity as fundamental principles. Those are operationalized through the following components:

- **Psychological explainability:** Beyond providing technical justifications, systems must communicate their inferences in language that is comprehensible to users (IAPP, 2023; Joyce et al., 2023), aligning with their cognitive and emotional frameworks.
- **Ethical co-design models:** Following Gebru et al. (2021), the use of Datasheets for Datasets and Model Cards (Mitchell et al., 2018; OECD, 2022) ensures transparency regarding data provenance, limitations, and biases, while maintaining a person-centred approach (Yıldız, 2025a, 2025b).
- **Collaborative ethical audits:** Model review processes should involve multidisciplinary teams composed of psychologists, engineers, and legal experts, like biomedical ethics committees.
- **Responsible psychological simulation:** Instead of attempting to mimic human empathy, systems should explicitly convey their artificial nature, reducing emotional confusion and anthropomorphic projection by users (Floridi, 2019).

Operationally, this dimension should incorporate psychological impact metrics, such as anxiety, trust, and sense of agency, into usability testing protocols. At the same time, algorithms that process language or emotion should rely on culturally representative corpora to prevent epistemic exclusion. Additionally, it is proposed that psychological associations establish an ethical quality seal for digital mental health applications, analogous to medical certifications (Singh et al., 2025), accompanied by practical guidelines for responsible tool selection like those issued by the APA (2024).

To ensure that the principles of transparency, explainability, and accountability are operationally implemented, the adoption of a Psychological Model Card is proposed (Table 3). This instrument, inspired by the technical format introduced by Mitchell et al. (2019) and aligned with OECD (2022) recommendations, has been adapted to the clinical and ethical specificities of psychology. It provides a structured framework for documenting the characteristics, limitations, and psychological risks of AI systems applied in mental health contexts.

The Psychological Model Card functions as both an ethical assurance tool and a mechanism of accountability, linking technological regulation with the psychological wellbeing of users and practitioners. In the card, the left column specifies the required fields, while the right column provides a simulated example to illustrate comprehension and facilitate practical implementation.

Institutional Axis: Governance, Training, and Responsibility

This axis encompasses the macro level structures guiding policies, regulations, oversight and professional training necessary to ensure that AI ethics in mental health depend on permanent governance systems rather than the goodwill of individual developers. The PRISM framework adheres to a human-in-the-loop perspective, where AI is conceived as a support system rather than an autonomous agent. Building on the cyborg ontology in healthcare, mental health nurses and practitioners are envisioned as technologically augmented practitioners who integrate AI insights with human judgement, empathy, and compassion (Yıldız, 2025a). This axis includes the following key components:

- **Digital ethics committees:** Healthcare institutions and universities should incorporate technology specialized

Table 2
Proposal of Psychological Indicators for the Impact of AI in Mental Health

Dimension	Indicator	Reference instrument	Suggested adaption
Digital autonomy	Perceived control over algorithmic decisions	Perceived Autonomy Scale (Liang & Reiss, 2025)	Adapted version including items related to clinical AI use
Algorithmic anxiety	Concern about the automation of clinical judgment	Technological Anxiety Scale (Meuter et al., 2003)	Adaptation for therapeutic settings
Digital epistemic trust	Level of credibility granted to automated inferences	Trust in Automation Scale (McGrath et al., 2025)	Inclusion of ethical and empathic factors
Digital moral fatigue	Exhaustion arising from ethical conflicts generated by AI	Moral Fatigue Scale (Epstein et al., 2019)	Adaption for digital scenarios
General psychological wellbeing	Satisfaction, self-efficacy, and sense of digital purpose	Ryff Psychological Wellbeing Scale (Ryff, 1989)	Incorporate items related to interaction with AI

Table 3
Simplified Psychological Model Card for AI Systems in Mental Health

General system identification			
System name	"PsIA-Assist v1.0" (generic example for a clinical support model)		
Developer / Institution	Laboratory of Computational Psychology and Digital Ethics, University X		
Version / Review date	1.0 – January 2026		
Application domain	Support for psychological assessment and recommendation of mental health resources		
Target population	Adults (18-65 years), without severe diagnoses, in clinical or psycho educational contexts		
Purpose	To facilitate clinical decision making and therapeutic follow-up through natural language and emotional pattern analysis		
Psychological and ethical foundations			
Dimension	**Applied Principle**		**Source / Reference**
Autonomy and control	Explicit digital consent, explanatory interface, option to reject automation		APA (2025); Deci & Ryan (2000)
Psychological non-Maleficence	Mandatory human supervision for emotional interpretations		WHO (2023)
Emotional Transparency	Visible indicators of model confidence and textual justification of inferences		Gebru et al. (2021)
Sensitive data Protection	Encryption of psychological records and irreversible anonymization		EU AI Act (2024)
Data and sources			
Item	**Description**		
Data origin	Anonymized clinical corpus (European and Latin-American institutions, 2020-2024)		
Data type	Interview transcripts, self-reports, and emotional diaries		
Representativeness	Gender balance and linguistic/socioeconomic diversity		
Limitations	Under representation of adults > 55 years old and patients with severe disorders		
Detected biases	Higher emotional positivity in training corpus; tendency to undervalue mild distress		
Corrective measures	Annual retraining, weighted class adjustment, and qualitative psychological sample review		
Performance and wellbeing metrics			
Metric Type	**Indicator**	**Value/Status**	**Validation Source**
Technical	Emotional classification accuracy	0.87 (cross-validation, 2024)	Internal dataset
Psychological	User satisfaction (digital wellbeing scale)	4.3/5 (n=312)	Pilot study (2024)
Ethical	Cognitive understandability	82% of users understand explanations	Observational study (2024)
Emotional risk	Algorithmic anxiety index	0.12 (low)	X Questionnaire (2024)
Oversight and governance			
Item		**Description**	
Responsible ethics board		Psychological Committee for Digital Ethics (PCDE-2025)	
External review		Annual report from an independent psychological audit team	
Compliance officer		Coordinator for ethics and digital wellbeing at the institution	
Reporting channels		Anonymous form for reporting errors or perceived bias	
Anticipated psychological impact			
Item	**Positive Effect**	**Potential Risk**	**Mitigation Strategy**
Digital Autonomy	Greater perceived control through shared decision-making	Risk of excessive delegation	Retraining in ethical literacy
Epistemic Trust	Increased credibility in support systems	Overconfidence or blind trust	Explicit messaging about model limitations
Technological Anxiety	Reduction with guided use	Heightened anxiety in vulnerable users	Mandatory clinical supervision
Professional Moral Fatigue	Decrease through automation of repetitive tasks	Risk of empathic detachment	Group reflective supervision
Social responsibility and sustainability			
Item		**Description**	
Accessibility		Interface adapted for functional diversity	
Ethical license		Use limited to assistive, non-commercial purposes, protecting sensitive data from non-therapeutic secondary exploitations	
Public transparency		Open documentation of the model and audit outcomes	
Periodic review		Every 12 months, with public report on bias and digital wellbeing metrics	
Narrative Summary (for publication or institutional communication)			
PsIA-Assist is a mental health support system designed under psychological and ethical principles. It guarantees user autonomy, transparency of algorithmic decisions, and constant human supervision. Its development complies with the standards of the AI Act (2024) and the ethical guidelines of the APA (2025), prioritizing patient wellbeing and trust. Each system version is reviewed by an interdisciplinary committee of psychologists, engineers, and legal experts to ensure social responsibility and minimize the risk of emotional harm.			

Note. Simulated information is provided in each section to facilitate standardized applicability.

psychologists within their ethical review boards as well as promote "ethical quality seals" to ensure the responsible design, evaluation, and deployment of AI systems in clinical contexts.

- **Continuous professional training:** Academic programs should integrate digital psychology, algorithmic ethics, and human-technology rights into their curricula (APA, 2025).
- **Shared responsibility:** Regulatory frameworks must acknowledge **co-responsibility** among designers, users, and service providers, recognizing that psychological risk cannot rest solely on patients or therapists.
- **International collaboration:** National guidelines should be harmonized with the European AI Act (2024) and WHO (2023) standards to promote global coherence in ethical oversight.

Regulation must also include independent monitoring bodies tasked with documenting and disseminating best practices, acting as intermediaries between regulators and professional communities. A comprehensive operational guide should support mental health organizations, universities, and technology platforms in auditing, supervising, and sustaining ethical AI practices, thereby strengthening regulatory sustainability (Table 4).

Additionally, the potential implementation of a certification or labelling system for AI in mental health will provide clinicians with immediate, standardized insights into a system's understandability. Its practical implications include ensuring that practitioners have access to traceable audit trails and ensuring that mental impact becomes a measurable criterion of normative validity, allowing for a calibrated trust that is corresponding with the AI's actual capabilities and clinical risks.

Table 4

Operational Guide for Institutional Implementation

Objective: To assist mental health institutions, universities, and digital platforms in implementing the proposed framework.
Initial diagnosis
● Audit existing AI systems regarding transparency, bias, and supervision.
● Assess digital wellbeing among staff and users.
Training and awareness
● Conduct interdisciplinary workshops on algorithmic ethics.
● Implement programs on emotional digital literacy.
Ongoing tracking
● Establish an internal Psychological Digital Ethics Committee.
● Publish semi-annual reports on the psychological impact of AI.
Transparency and public communication
● Develop traceability panels and explanations of automated decisions.
● Include formal protocols for bias and error correction.

Note. Simulated information is provided in each section to facilitate standardized applicability.

To maximize the applied impact of the PRISM framework, the theoretical pillars have been translated into actionable clinical guidelines for practitioners and health organizations. First, to address the risk of therapeutic depersonalization, institutions should implement the role of digital navigators, specialized staff trained to bridge the gap between algorithmic outputs and patient understanding, thereby ensuring that technology supplements rather than supplants humanistic care (Torous et al., 2025). Second, clinicians should adopt psychological model cards as a standard pre-deployment audit tool to evaluate a system's understandability and its alignment with human clinical heuristics (Mitchell et al., 2018). Furthermore, regarding information integrity, practitioners are advised to use affective transparency cues, explicitly informing patients about the AI's lack of emotional sentience to prevent the revisited ELIZA effect and emotional dependency (Wang et al., 2025). Finally, to ensure algorithmic justice, health services must establish participatory co-design committees that involve patients from disregarded groups in the validation of diagnostic tools, mitigating the risk of structural biases. These directives transform the PRISM model from a normative proposal into a practical toolkit for the responsible democratization of mental health AI.

Discussion

The proposed PRISM framework is designed not to replace legal norms but to complement them from a person-centred, psychological perspective. As a discipline focused on behaviour and experience, psychology is uniquely positioned to anticipate the subjective consequences of healthcare digitalization. Unlike existing international frameworks which provide high-level moral imperatives, PRISM offers operational granularity, equipping practitioners with tools such as the Psychological Model Card to navigate fragmented enforcement architectures including GDPR, DSA, APA, and the AI Act.

By integrating ethical regulation from the standpoint of subjectivity, PRISM supports preventing emerging pathologies such as algorithmic dependence, digital moral fatigue, and surveillance anxiety, recently documented in clinical literature (Torous et al., 2025; Wang et al., 2025). Standardized operational tools, such as the Psychological Model Card for AI systems in mental health (Table 3), translate ethical and psychological principles into measurable structures, enabling ethical traceability and assessment of AI's psychological impact. Unlike traditional Model Cards focused exclusively on technical performance metrics, this adaptation incorporates variables such as digital autonomy, epistemic trust, algorithmic anxiety, and professional moral fatigue, promoting a holistic conception of digital wellbeing. It thus functions as a bridge between algorithmic ethics and applied psychology, with potential as a standard reference for ethical audits and institutional validation in digital mental health contexts.

The ethical sustainability of digital mental health ecosystems depends on cultivating moral resilience among practitioners and embedding human dignity in technological governance. Consistent with WHO (2025) recommendations, this requires robust frameworks ensuring transparency, multidisciplinary oversight, and the preservation of human interaction as the core of the therapeutic relationship. Digital environments should operate as hybrid ecosystems where technology extends, but does not replace, empathy and qualified care, thereby strengthening patient safety and ethical integrity.

Internationally, PRISM represents an epistemological shift. It positions psychology as a normative discipline capable of anticipating, measuring, and mitigating cognitive and emotional harm from AI, while introducing novel contributions. By integrating classical psychological principles such as autonomy, agency, and relationality, with current demands of algorithmic ethics and digital regulation, PRISM extends regulation beyond legal compliance to include the psychological dimension of user experience. Mental impact becomes a criterion of normative validity, transforming regulation into a form of preventive clinical intervention. The framework also provides operational guidelines for design, professional practice, and training, laying the foundation for empirical validation in future research.

Operationally, PRISM addresses growing public and professional concerns about the psychological impact of intensive AI use. It aims to influence policy and social relevance by offering a conceptual and practical toolkit for digital mental health. The framework supports interdisciplinary observatories and international codes of psychological practice regarding AI. Its broad scope facilitates practical implementation and replicable impact evaluation across clinical, institutional, and policy contexts, enhancing transferability and scalability. Success depends on collaboration among law, philosophy, data science, and applied psychology to resolve the human oversight paradox and ensure automated decision-making aligns with human dignity, autonomy, and vulnerability (Kyrychenko et al., 2026; Theunissen & Browning, 2022).

Future research should consolidate psychological regulation of AI through an interdisciplinary agenda integrating empirical, ethical, and technological approaches. Six priority areas are proposed: 1) Framework validation through participatory co-design inclusive development cycles involving clinicians, professional associations, industry developers, and, most importantly, patients from vulnerable or underrepresented groups; 2) assessment of the psychological impact of AI systems on users and professionals (algorithmic anxiety, moral fatigue, perceived control); 3) development of measurement instruments for digital autonomy, epistemic trust, and technological wellbeing; 4) design and validation of educational programs on ethical and emotional digital literacy; 5) longitudinal studies on the effects of continued interaction with conversational agents on therapeutic relationships and self-concept; and 6) promotion of international psychological standards for algorithmic ethics.

The tools proposed in this paper provide instrumental and methodological foundations for future research, offering impact indicators, governance instruments, and application guidelines to orient studies at the intersection of psychology and ethical technology. By integrating ethical principles with technological innovation, psychology can lead the design of digital mental health policies. Professionals must be trained in AI ethics, understand automation fundamentals, and actively contribute to regulatory frameworks. Institutions, in turn, must recognize that psychological wellbeing depends not only on clinical interventions but also on just, intelligible, and transparent digital environments. Instruments such as the Psychological Model Card exemplify how abstract ethical principles can be translated into operational assurance and professional accountability. Systematic implementation would allow institutions to integrate algorithmic ethics into daily practice, offering a tangible response to emerging regulatory challenges.

In conclusion, the ethical sustainability of digital mental health hinges on a robust multilevel framework that protects autonomy and human dignity in the face of increasingly intelligent systems. PRISM offers a structured path for ethically sustainable digital mental health, establishing psychology as a leading discipline in shaping human-centred AI governance.

Author Contribution

Ignacio Pedrosa: Conceptualization, Formal Analysis, Investigation, Writing - Original Draft, Writing - Review & Editing.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors

Declaration of Interests

The author declares that there is no conflict of interest.

Data Availability Statement

Data sharing is not applicable to this article as no new data were created or analysed in this study.

References

Abou Hashish, E. A. (2025). Compassion through technology: Digital empathy concept analysis and implications in nursing. *Digital Health*, *11*, Article 20552076251326221. https://doi.org/10.1177/20552076251326221

Adler, D. A., Stamatis, C. A., Meyerhoff, J., Mohr, D. C., Wang, F., Aranovich, G. J., Sen, S., & Choudhury, T. (2024). Measuring algorithmic bias to analyze the reliability of AI tools that predict depression risk using smartphone sensed-behavioral data. *npj Mental Health Research*, *3*(1), Article 17. https://doi.org/10.1038/s44184-024-00057-y

American Psychological Association [APA] (2024). *APA's AI tool guide for practitioners*. https://www.apaservices.org/practice/business/technology/tech-101/evaluating-artificial-intelligence-tool

APA (2025). *Ethical guidance for artificial intelligence in professional practice of health service psychology.* https://www.apa.org/topics/artificial-intelligence-machine-learning/ethical-guidance-professional-practice.pdf

Bandura, A. (2001). Social cognitive theory: An agentic perspective. *Annual Review of Psychology*, *52*(1), 1–26. https://doi.org/10.1146/annurev.psych.52.1.1

Cheng, X. (2024). Research on depression recognition based on university students' facial expressions and actions with the assistance of artificial intelligence. *Journal of Advanced Computational Intelligence and Intelligent Informatics*, *28*(5), 1126–1131. https://doi.org/10.20965/jaciii.2024.p1126

Deci, E. L., & Ryan, R. M. (2000). The 'what' and 'why' of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry*, *11*(4), 227–268. https://doi.org/10.1207/S15327965PLI1104_01

Dehbozorgi, R., Zangeneh, S., Khooshab, E., Nia, D. H., Hanif, H. R., Samian, P., Yousefi, M., Hashemi, F. H., Vakili, M., Jamalimoghadam, N., & Lohrasebi, F. (2025). The application of artificial intelligence in the field of mental health: A systematic review. *BMC Psychiatry*, *25*(1), Article 132. https://doi.org/10.1186/s12888-025-06483-2

Dencik, L., & Sanchez-Monedero, J. (2022). Data justice. *Internet Policy Review*, *11*(1). https://doi.org/10.14763/2022.1.1615

Epstein, E. G., Whitehead, P. B., Prompahakul, C., Thacker, L. R., & Hamric, A. B. (2019). Enhancing understanding of moral distress: The measure of moral distress for health care professionals. *AJOB Empirical Bioethics*, *10*(2), 113–124. https://doi.org/10.1080/23294515.2019.1586008

European Union (2024). *Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act). Official Journal of the European Union.* https://eur-lex.europa.eu/eli/reg/2024/1689/oj/eng

Floridi, L. (2019). *The logic of information: a theory of philosophy as conceptual design.* Oxford University Press.

Gebru, T., Morgenstern, J., Vecchione, B., Vaughan, J. W., Wallach, H., Daumé, H., & Crawford, K. (2021). *Datasheets for datasets* (arXiv:1803.09010). arXiv. https://doi.org/10.48550/arXiv.1803.09010

Gupta, A., Mund, A., Roy, S., Garg, P., & Yadav, D. K. (2025). Trust in AI systems: A social-psychological investigation of human–AI collaboration. *TPM – Testing, Psychometrics, Methodology in Applied Psychology, 32*(S7), 428-446. *https://tpmap.org/submission/index.php/tpm/article/view/2133*

Hoose, S., & Kralikova, K. (2024). Artificial intelligence in mental health care: Management implications, ethical challenges, and policy considerations. *Administrative Sciences*, *14*(9), Article 227. https://doi.org/10.3390/admsci14090227

IAPP (2023). *5 things to know about model cards*. https://iapp.org/news/a/5-things-to-know-about-ai-model-cards

Joyce, D. W., Kormilitzin, A., Smith, K. A., & Cipriani, A. (2023). Explainable artificial intelligence for mental health through transparency and interpretability for understandability. *npj Digital Medicine*, *6*(1), Article 6. https://doi.org/10.1038/s41746-023-00751-9

Kyrychenko, M., Mudryi, M., & Chaklosh, M. (2026). *Global AI governance overview: Understanding regulatory requirements across global jurisdictions* (arXiv:2512.02046). arXiv. https://doi.org/10.48550/arXiv.2512.02046

Laestadius, L., Bishop, A., Gonzalez, M., Illenčík, D., & Campos-Castillo, C. (2024). Too human and not human enough: A grounded theory analysis of mental health harms from emotional dependence on the social chatbot Replika. *New Media & Society*, *26*(10), 5923–5941. https://doi.org/10.1177/14614448221142007

Li, W., & Liu, X. (2025). Anxiety about artificial intelligence from patient and doctor-physician. *Patient Education and Counseling*, *133*, Article 108619. https://doi.org/10.1016/j.pec.2024.108619

Liang, H. & Reiss, M.J. (2025). The associations between students' attitudes toward AI and learning engagement: serial mediating roles of perceived autonomy and learning enjoyment. *Frontiers in Psychology 16*, Article 1681635. https://doi.org/10.3389/fpsyg.2025.1681635

Loechner, J., Carlbring, P., Schuller, B., Torous, J., & Sander, L. B. (2025). Digital interventions in mental health: An overview and future perspectives. *Internet Interventions-the Application of Information Technology in Mental and Behavioural Health*, *40*, Article 100824. https://doi.org/10.1016/j.invent.2025.100824

McGrath, M.J, Lack, O., Tisch, J. & Duenser, A. (2025). Measuring trust in artificial intelligence: validation of an established scale and its short form. *Frontiers in Artificial Intelligence, 8*, Article 1582880. https://doi.org/10.3389/frai.2025.1582880

Meuter, M. L., Ostrom, A. L., Bitner, M. J., & Roundtree, R. (2003). The influence of technology anxiety on consumer use and experiences with self-service technologies. *Journal of Business Research, 56*(11), 899–906.

Milasan, L. H., & Scott-Purdy, D. (2025). The future of artificial intelligence in mental health nursing practice: An integrative review. *International Journal of Mental Health Nursing*, *34*(1), Article e70003. https://doi.org/10.1111/inm.70003

Mitchell, M., Wu, S., Zaldivar, A., Barnes, P., Vasserman, L., Hutchinson, B., Spitzer, E., Raji, I. D., & Gebru, T. (2019). *Model cards for model reporting*. (arXiv:1810.03993). arXiv. https://doi.org/10.48550/ARXIV.1810.03993

Organisation for Economic Co-operation and Development [OECD] (2022). *Model Cards*. In OECD.AI – Catalogue of Tools & Metrics for Trustworthy AI. https://oecd.ai/en/catalogue/tools/model-cards

Padalkar, T., Henderson, N., Rocque, G. B., Hildreth, K., Whitlow, O., Ingram, S. A., Eltoum, N., Herring, L., Igunma, O., Hardy, C. M., & Chu, D. I. (2024). "We want that personal touch": Dilemma of digitalization of healthcare based on community health worker experiences of depersonalization in healthcare. *JCO Oncology Practice*, *20*(10_suppl), 261–261. https://doi.org/10.1200/OP.2024.20.10_suppl.261

Ryff, C.D. (1989). Psychological Well-Being Scale [Database record]. APA PsycTests. https://doi.org/10.1037/t04262-000

Sanz, A., Tapia, J. L., Garcia-Carpintero, E., Rocabado, J. F., & Pedrajas, L. M. (2025). ChatGPT simulated patient: Use in clinical training in psychology. *Psicothema*, *37*(3), 23–32. https://doi.org/10.70478/psicothema.2025.37.21

Singh, R., Bapna, M., Diab, A. R., Ruiz, E. S., & Lotter, W. (2025). How AI is used in FDA-authorized medical devices: A taxonomy across 1,016 authorizations. *npj Digital Medicine*, *8*(1), Article 388. https://doi.org/10.1038/s41746-025-01800-1

Solove, D. J. (2021). *Understanding privacy* (2nd ed.). Harvard University Press.

Suárez-Álvarez, J., He, Q., Guenole, N., & D'Urso, D. (2026). Using artificial intelligence in test construction: A practical guide. *Psicothema*, *38*(1), 1–12. https://doi.org/10.70478/psicothema.2026.38.01

Tang, D., Xi, X., Li, Y., & Hu, M. (2025). Regulatory approaches towards AI medical devices: A comparative study of the United States, the European Union and China. *Health Policy*, *153*, Article 105260. https://doi.org/10.1016/j.healthpol.2025.105260

Theunissen, M., & Browning, J. (2022). Putting explainable AI in context: Institutional explanations for medical AI. *Ethics and Information Technology*, *24*(2), Article 23. https://doi.org/10.1007/s10676-022-09649-8

Torous, J., Linardon, J., Goldberg, S. B., Sun, S., Bell, I., Nicholas, J., Hassan, L., Hua, Y., Milton, A., & Firth, J. (2025). The evolving field of digital mental health: Current evidence and implementation issues for smartphone apps, generative artificial intelligence, and virtual reality. *World Psychiatry*, *24*(2), 156–174. https://doi.org/10.1002/wps.21299

Wang, X., Zhou, Y., & Zhou, G. (2025). The application and ethical implication of generative ai in mental health: Systematic review. *JMIR Mental Health*, *12*, Article e70610. https://doi.org/10.2196/70610

World Health Organization [WHO] (2023). *Ethics and governance of artificial intelligence for health.* https://www.who.int/publications/i/item/9789240029200

WHO (2025). *Ethics and governance of artificial intelligence for health: Guidance on large multi-modal models*. https://www.who.int/publications/i/item/9789240084759

Wiljer, D., Charow, R., Costin, H., Sequeira, L., Anderson, M., Strudwick, G., Tripp, T., & Crawford, A. (2019). Defining compassion in the digital health age: Protocol for a scoping review. *BMJ Open*, *9*(2), Article e026338. https://doi.org/10.1136/bmjopen-2018-026338

Yang, C., & Lu, X. (2026). Educational pathways for enhancing algorithmic transparency: A discussion based on the phenomenological reduction method. *AI & Society*, *41*(1), 151–164. https://doi.org/10.1007/s00146-025-02475-8

Yang, H., & Sundar, S. S. (2025). AI anxiety: Explication and exploration of effect on state anxiety when interacting with AI doctors. *Computers in Human Behavior: Artificial Humans*, *3*, Article 100128. https://doi.org/10.1016/j.chbah.2025.100128

Yıldız, E. (2025a). Artificial intelligence in mental health nursing: Balancing clinical efficiency and the human touch: A quest for a new synthesis. *Journal of Psychiatric and Mental Health Nursing*, Article 13173. https://doi.org/10.1111/jpm.13173

Yıldız, E. (2025b). Ethical big data for personalised mental health nursing: A P4 and systems view. *Journal of Psychiatric and Mental Health Nursing*, *32*(6), 1404–1411. https://doi.org/10.1111/jpm.70038

GRUPO EDITORIAL DE PSICOFUNDACIÓN

Psicothema (2026) 38(3) 123-137

Psicothema

https://www.psicothema.com/es • eISSN 1886-144X

Colegio Oficial de Psicología del Principado de Asturias

Psicothema

Article

Acceptance and Commitment Therapy (ACT) in End-of-Life Care: A Meta-Analysis of Efficacy and Process of Change

Juanjo Macías[1], Lucía Lara-Merín[2] and Carlos López-Pinar[2]

[1] University of Burgos (Spain)
[2] European University of Valencia (Spain)

ARTICLE INFO

Received: 17/11/2025
Accepted: 22/01/2026

Keywords:
Acceptance and Commitment Therapy
Quality of life
Terminal illness
Palliative care
Systematic review
Meta-analysis

ABSTRACT

Background: Acceptance and Commitment Therapy (ACT) is an evidence-based, process-oriented psychotherapy that provides palliative care through compassionate, person-centered care for terminal illness. This study evaluates the effects of ACT on quality of life (QoL) and emotional distress and examines the role of process variables (e.g., psychological flexibility). **Method**: Systematic review with meta-analysis following searches in PubMed, SCOPUS and PsycInfo for studies measuring QoL, anxiety/depression symptoms, and process variables in ACT interventions. Random-effects meta-analyses and meta-regressions were performed. Risk of bias was assessed. **Results**: 13 studies were included (1,109 participants; M_{age} = 61.5; 55.7% women; 6 RCTs). ACT showed significant, moderate effects for QoL, anxiety and depression symptoms versus active controls (SMD = 0.44 to 0.62), sustained at 1-9 months follow-up (SMD = 0.35 to 0.56). Psychological flexibility emerged as the strongest predictor of QoL improvement. Face-to-face interventions demonstrated better results than remote delivery. **Conclusions**: These findings support the clinical relevance of ACT as an effective approach for addressing psychological distress and improving QoL in end-of-life care. Key limitations include self-report bias, high heterogeneity, potential overlap between measures, publication bias for QoL, and limited sample sizes in some analyses. High-quality trials are warranted, particularly in non-oncological populations, caregivers, and healthcare professionals.

Terapia de Aceptación y Compromiso (ACT) en Cuidados al Final de la Vida: Un Metaanálisis de Eficacia y Procesos de Cambio

Palabras clave:
Terapia de Aceptación y Compromiso
Calidad de vida
Enfermedad terminal
Cuidados paliativos
Revisión sistemática
Meta-análisis

RESUMEN

Introducción: La Terapia de Aceptación y Compromiso (ACT) es una psicoterapia basada en procesos que proporciona atención compasiva y centrada en la persona en cuidados paliativos. Este estudio evalúa la eficacia de ACT en calidad de vida (CdV) y malestar emocional, y examina variables de proceso (p.ej., flexibilidad psicológica). **Método**: Se buscó en PubMed, Scopus y PsycInfo estudios que midieran CdV, síntomas de ansiedad/depresión, y variables de proceso en intervenciones ACT. Se realizaron metaanálisis de efectos aleatorios y metarregresiones. Se evaluó el riesgo de sesgo. **Resultados**: Se incluyeron 13 estudios (1,109 participantes; M_{edad} = 61,5; 55,7% mujeres; 6 ECA). ACT mostró efectos moderados significativos en CdV, ansiedad y depresión frente a grupos de control activos (SMD = 0,44 to 0,62), mantenidos en el seguimiento de 1-9 meses (SMD = 0,35 to 0,56). Las intervenciones presenciales mostraron resultados superiores a las remotas. **Conclusiones**: Estos hallazgos apoyan la relevancia clínica de ACT para abordar malestar psicológico y mejorar CdV en cuidados al final de la vida. Las limitaciones incluyen sesgo de autoinforme, alta heterogeneidad, solapamiento entre medidas, sesgo de publicación para CdV y tamaños muestrales limitados. Se requieren ensayos de alta calidad en poblaciones no oncológicas, cuidadores y profesionales sanitarios.

Cite as: Macías, J., Lara-Merín, L., & López-Pinar, C. (2026). Acceptance and Commitment Therapy (ACT) in end-of-life care: A meta-analysis of efficacy and process of change. *Psicothema, 38*(3), 123-137. https://doi.org/10.70478/psicothema.2026.38.12
Corresponding author: Carlos López-Pinar, carloslopez.pinar@universidadeuropea.es

The global burden of terminal illness is substantial and is expected to increase significantly in the coming decades, driven in part by sustained population aging (Lee & Zhou, 2017). It is estimated that between 69% and 82% of individuals who die each year could have required palliative care for chronic or terminal conditions (Murtagh et al., 2014). Despite the crucial role of palliative care in health systems and universal health coverage, global access remains severely limited. The Lancet Commission estimates that over 61 million people experience serious health-related suffering annually, with at least 80% lacking basic palliative care, including essential pain medication (Knaul et al., 2018; Radbruch et al., 2020). These figures illustrate the substantial mortality burden associated with terminal illnesses and highlight the urgent need for early detection, effective treatment, and integrated palliative care strategies (Yang et al., 2024; Zhao et al., 2021). Tackling this challenge requires a robust, sustainable framework that ensures equitable access to supportive services across the illness trajectory. Given the evolving, multifaceted nature of terminal conditions, updated definitions are needed to capture their medical complexity and existential dimensions, as will be further addressed below.

While major international organizations (World Health Organization, European Association for Palliative Care) have not established specific operational criteria for end-of-life status (Surges et al., 2024; World Health Organization, 2020), consensus definitions describe it as an advanced, progressive, and incurable disease with no reasonable response to curative treatment, accompanied by multiple intense symptoms and significant emotional impact on patient and family (Organización Médica Colegial & Sociedad Española de Cuidados Paliativos, 2015). The lack of universal consensus on defining end-of-life leads to heterogeneity in study populations across palliative care research, with varying operational criteria including prognosis thresholds, disease staging, or enrollment in palliative care programs.

Palliative care addresses both physical suffering and existential challenges through comprehensive, multidisciplinary approaches that emphasize symptom management while acknowledging patients' personal narratives and sense of meaning (Teoli et al., 2025). The World Health Organization highlights that palliative care should be aimed at improving the QoL of patients and their families facing life-threatening illnesses (World Health Organization, 2022). Psychotherapy interventions based on contextual models may improve care by addressing the disease and the patient's psychological and relational context (Quevedo et al., 2022; Schiele et al., 2020). These approaches can be framed within contextual behavioral science (CBS), which emphasizes understanding and influencing behavior as embedded in its historical and situational context.

QoL has increasingly been recognized as a central outcome in end-of-life care, encompassing physical comfort, emotional well-being, and social connectedness beyond mere survival (Khan et al., 2005; World Health Organization, 2022). Approaching the final stage of life due to a terminal illness evokes not only the fear of death, but also a complex array of emotional and existential challenges affecting identity, meaning, and connection (Lamont, 2005). The most common manifestations are symptoms of anxiety and depression (Secinti et al., 2019). These conditions significantly impact QoL, treatment adherence, and symptom burden. In addition, a diagnosis of terminal illness often triggers what has been termed an "existential plight" (Weisman & Worden, 1977), a heightened awareness of mortality that encompasses profound concerns about meaning, identity, and impending death (Yalom, 2008). The medical community increasingly recognizes addressing these existential dimensions as integral to patient care (Tarbi & Meghani, 2019).

While cognitive-behavioral therapy (CBT) has demonstrated moderate efficacy in reducing anxiety and depressive symptoms in patients with advanced cancer (Wilaras, 2024; Xia et al., 2024), only a minority of CBT trials have explicitly addressed existential distress, and treatment effects in this population have often been non-significant (Campbell & Campbell, 2012). This limitation is notable given that existential concerns (such as death anxiety, loss of meaning, and spiritual suffering) are inherent to the palliative care context (Lee, 2020). Acceptance and Commitment Therapy (ACT), as both a philosophy and clinical approach (Carlsson, 2012; Kissane, 2012), may offer a more comprehensive framework for addressing these concerns. ACT targets psychological flexibility, the ability to remain open to difficult inner experiences while engaging in value-based action (Hayes, Strosahl, et al., 2006), through explicit emphasis on values clarification, acceptance of death-related anxiety, and commitment to meaningful action in the face of existential threat. Evidence suggests that psychological flexibility is a key mechanism of change, associated with decreased anxiety and depression in patients with terminal illnesses (Zhao et al., 2021), making it a central outcome in assessing ACT's efficacy in end-of-life care as a valuable complement to biomedical interventions (Kredentser & Chochinov, 2020; Sauer et al., 2024; Secinti et al., 2019).

ACT is a transdiagnostic, process-based therapy rooted in functional contextualism that emphasizes psychological flexibility, which enables patients to engage meaningfully with suffering while reorienting actions toward personal values (Hayes et al., 2013, 2022; Hayes, Strosahl, et al., 2006). The Extended Evolutionary Meta-Model identifies processes such as acceptance and values-based action as essential for addressing existential dilemmas and clinical symptoms (Ciarrochi et al., 2022; Hayes et al., 2020; Menzies & Menzies, 2024). There is evidence supporting the efficacy of ACT in managing chronic illnesses and other health conditions (Gloster et al., 2020; González-Fernández et al., 2018; Jiang et al., 2024).

Recently, ACT has expanded to include terminal illness. In their systematic review, Li et al. (2021) identified 6 studies (n = 261) examining the effect of ACT for patients with advanced cancer. They found that these interventions had a positive effect on emotional distress, health-related QoL and somatic issues such as sleep problems. Gibson Watt et al. (2023) conducted a systematic scoping review that also included caregivers. With a similar sample size (k = 7 intervention studies for patients) they found results consistent with Li et al. (2021), and expanded them with findings suggesting benefits for psychological distress in caregivers too. A recent meta-analysis (Fang et al., 2023) found 8 studies (n = 488) examining the effects of ACT for advanced cancer patients. They found large effects on QoL, anxiety and depression, while the effects on psychological flexibility were non-significant.

However, these reviews have important limitations that warrant an updated and more comprehensive meta-analysis. First, the only existing meta-analysis (Fang et al., 2023) excluded non-randomized controlled designs, potentially limiting the comprehensiveness of the evidence base. Second, this meta-analysis only examined immediate post-treatment effects without analyzing longer-term follow-up outcomes, leaving questions about the durability of ACT's

benefits. Third, none of the existing reviews, whether systematic reviews (Gibson Watt et al., 2023; Li et al., 2021) or meta-analysis (Fang et al., 2023), explored mechanisms of change, despite ACT theory emphasizing psychological flexibility as the key therapeutic process. Fourth, all existing reviews focused exclusively on patients with advanced cancer, whereas terminal illness encompasses a broader range of conditions including amyotrophic lateral sclerosis (ALS) and other progressive diseases that would benefit from palliative psychological interventions. Fifth, a substantial number of studies have been published in recent years, necessitating an updated synthesis. Although ACT does not pursue symptom reduction as its primary therapeutic objective, changes in process such as psychological flexibility and valued action are consistently associated with improvements in QoL and reductions in emotional distress (Gloster et al., 2020; Hayes, Luoma, et al., 2006).

The present study (a) systematically reviews and meta-analyzes the existing literature on the short- and longer-term effects of ACT on QoL and emotional distress (e.g., symptoms of anxiety and depression) in terminally ill patients, and (b) explores the mechanism of change by examining how changes in ACT process variables, particularly values-based living and psychological flexibility, may explain the potential improvements in these clinical outcomes.

Method

This systematic review with meta-analysis followed the PRISMA guidelines for systematic reviews and meta-analyses (Page et al., 2021). In addition, this review has been pre-registered in PROSPERO (CRD42024562107). The results of the selection process, with the reasons for excluding, the coding process (with the full codebook) and the fully extracted dataset can be consulted at this OSF link.

Information Sources and Search Strategy

MedLine (via PubMed), Scopus, and PsycInfo were consulted using the following search terms in the title, abstract, and keywords (for the population and the intervention): (*palliativ** OR *advanced progressive illness* OR *advanced care* OR *last year* OR *terminally* OR *terminal care* OR *terminal* OR *hospice* OR *advanced cancer* OR *amyotrophic lateral sclerosis)* AND *(acceptance and commitment)*

The full search strings are available at this OSF link. Where available, the search was filtered to show only studies published in peer-reviewed journals. Reference lists of identified articles were reviewed to identify additional studies. Previous reviews on the field of ACT research on patients with cancer and other potentially terminal conditions were also consulted (Fang et al., 2023; Gibson Watt et al., 2023; Li et al., 2021; Okuyama et al., 2017; Sauer et al., 2024). In addition, a forward reference search of articles citing the identified studies was conducted to potentially identify additional eligible studies. The search included articles published up to May 26, 2025. These search terms were piloted by the authors to ensure that they would cover all relevant studies comprehensively.

Eligibility Criteria

According to the PICOS strategy (Eriksen & Frandsen, 2018), studies meeting the following characteristics were included: (A) Population: Studies with a sample of adults (+18) who are terminally ill, including patients explicitly described as having advanced cancer or equivalent terminology (minimum stage III or IV with low survival rates, typically <50% five-year survival), advanced ALS, or patients receiving palliative care, both men and women; (B) Intervention: Studies in which at least one of the experimental groups received ACT. Combining ACT with other interventions was permitted, as long as the primary focus of the treatment remained addressing the ACT processes; (C) Comparison: For controlled trials, at least one of the arms did not receive ACT, allowing to isolate the specific effects of this intervention. Both active (e.g., treatment as usual, other psychological interventions, support groups) and inactive (e.g., waitlist, no interventions) control groups are accepted; (D) Outcomes: In order to address the main objective, studies had to include quantitative measures of depression symptoms, anxiety symptoms, and/or QoL as outcome variables. To address the question about the mechanism of change, they also had to include a measure of the main process of change variable proposed by ACT (e.g., psychological flexibility) or one of the six core process variables (e.g., valued action, cognitive fusion); (E) Study Designs: Both RCTs, nonrandomized controlled trials, and uncontrolled single-arm trials were included. Sensitivity analyses were conducted in order to control the randomization of the controlled trials. All study designs were included to provide a comprehensive synthesis of the literature. Between-group and within-subject analyses were conducted and reported separately to distinguish controlled from uncontrolled treatment effects; (F) Additional criteria: Priority was given to peer-reviewed articles in scientific journals, although doctoral theses were also included. Articles had to be written in English or Spanish, as these were the languages understood by the authors.

Selection Process

The initial title/abstract screening was conducted by a single reviewer using a conservative criterion that retained any potentially relevant study for full-text review, thereby minimizing the risk of premature exclusion. The final selection was made independently by two independent researchers by applying the selection criteria to the full texts. Agreement between the researchers was considered to be substantial, with a Cohen's kappa coefficient of .78. Disagreements were resolved by consensus.

Data Collection and Data Items

After the final selection of studies, the second and the last author independently coded the following variables: (A) study design (e.g., RCT, nonrandomized controlled trials, uncontrolled single-arm trials) and study characteristics (author, year, country); (B) population diagnosis (e.g., advanced cancer, amyotrophic lateral sclerosis); (C) sample mean age and percentage of women; (D) intervention type (e.g., ACT, ACT combined with other intervention), delivery (e.g., delivered by a therapist or self-delivered; face-to-face or remote), format (e.g., individual or group), and number of sessions; and (E) comparison (e.g., active control, inactive control). Discrepancies between the two researchers were resolved by consensus. The final author then extracted the sample size, means and standard deviations for each study arm. In cases where standard deviation values were not provided, they were calculated using the formulas recommended

by Higgins and Green (2011). For studies with insufficient data to calculate the ES, or in cases of uncertainty, the corresponding authors were contacted.

Methodological Quality Assessment

The Cochrane Risk of Bias Tool 2 (Sterne et al., 2019) was used to assess the risk of bias in RCTs, while risk of bias in nonrandomized controlled trials and uncontrolled single-arm trials was assessed with ROBINS-I (Sterne et al., 2016). RoB 2 evaluates the following five domains: (1) bias arising from the randomization process, (2) bias due to deviations from intended interventions, (3) bias due to missing outcome data, (4) bias in the measurement of the outcome, and (5) bias in the selection of the reported result. Each domain in RoB 2 was judged as having low, high, or some concerns risk of bias. ROBINS-I assesses seven domains: (1) bias due to confounding, (2) bias in the selection of participants into the study, (3) bias in the classification of interventions, (4) bias due to deviations from intended interventions, (5) bias due to missing data, (6) bias in the measurement of outcomes, and (7) bias in the selection of the reported result. Each domain in ROBINS-I was judged as having low, moderate, serious, or critical risk of bias, or no information. Two researchers assessed each study independently. Discrepancies were resolved by consensus.

One of the researchers evaluated publication bias by visually inspecting the symmetry of funnel plots. This was also assessed using the trim-and-fill method (Duval & Tweedie, 2000), to estimate the number of potentially missing studies and adjust effect size estimates accordingly, and the Egger test (Egger et al., 1997), to statistically assess funnel plot asymmetry, with $p < .05$ indicating potential publication bias.

Data Analysis

First, effect sizes (ESs) and their associated variances were estimated using Morris's proposal for within-subject (Morris, 2000) and between-group outcomes (Morris, 2008), adjusted with a small sample factor correction (Morris, 2000, 2008). The standard error was derived from the square root of the estimated variance. ES values of 0.2, 0.5, and 0.8 were interpreted as indicating small, moderate, and large effects, respectively (Cohen, 1988). Outcomes were extracted at post-treatment (first assessment immediately after treatment) and at follow-up. When multiple follow-up assessments were reported for the same outcome, we extracted the longest available follow-up time point.

A random effects model meta-analysis was then performed for each outcome by weighting the ESs by their inverse variance. This random effects model was chosen because of the expected high heterogeneity between studies and the assumption that the true effect size may vary across studies (Borenstein et al., 2010). To measure this heterogeneity among studies, the I^2 index was estimated for each outcome (Higgins et al., 2003). I^2 values of 25%, 50%, and 75% were interpreted as indicating low, moderate, and high heterogeneity, respectively.

Additional Analyses

Finally, to assess the mediating role of process variables of ACT (e.g., psychological flexibility model), meta-regressions were conducted for each outcome. In these meta-regression models, process variables were set as predictors, while QoL and depression and anxiety symptoms were selected as dependent variables. Meta-regressions were not conducted for outcomes with fewer than three studies because this number is insufficient to estimate between-study variance or fit a meaningful model (Thompson & Higgins, 2002). At least three studies are needed for minimal interpretability; however, higher numbers are preferred for reliable results. To assess other potential sources of heterogeneity, subgroup analyses were performed using $\chi 2$ tests. Studies were grouped according to the coded variables listed above. Significance was set at $p = .05$. Sensitivity analyses excluding studies mixing stage III and IV cancer patients were performed. Meta-analyses and meta-regressions were performed using the *meta* package (Balduzzi et al., 2019) in R (R Core Team, 2025). The R code and database are available at this OSF link.

Results

Five corresponding authors were contacted for further information. Two studies were excluded because not all individuals in their sample were in the terminal stage of cancer (Ksiksou et al., 2024; Lantheaume et al., 2020). The other three studies did not respond and were therefore excluded because we could not be sure whether they met the eligibility criteria (Han et al., 2019; Páez et al., 2007; Yusufov et al., 2023).

Finally, six RCTs, two nonrandomized controlled trials, and five uncontrolled single-arm trials, with a total of 1,109 participants, met the eligibility criteria and were included (see Table 1 and Figure 1). The mean age of the participants was 61.5 years ($SD = 3.9$), with the majority having been diagnosed with advanced cancer ($n = 9$). On average, 55.7% of the participants were women. Regarding the intervention, most studies employed standalone ACT ($n = 11$), which was predominantly delivered by a therapist ($n = 10$) and conducted in a face-to-face ($n = 6$) and individual ($n = 9$) setting. All control groups were active ($n = 9$). The mean number of sessions in the intervention was 6.8 ($SD = 2.3$). Eight studies included at least one follow-up assessment ($M = 4.1$ months, range: 1–9 months, $SD = 2.64$).

Risk of Bias Assessment Results

For all studies, there were some concerns about the measurement of the outcome, as all were self-reported (see Figure 2a and 2b). Overall, the remaining domains for controlled trials were considered to be at low risk of bias. For uncontrolled single-arm trials, several studies did not report enough information to assess the risk of bias due to deviations from the intended interventions, and there were some concerns about the management of missing data in some studies.

Regarding publication bias, the Egger test was not significant for all outcomes. This suggests no evidence of publication bias; however, tests may be underpowered and these results should be interpreted cautiously. Trim-and-fill adjusted ES for several outcomes, but ES remained statistically significant for all outcomes. Finally, visual inspection of funnel plots with 10 or more studies suggests some asymmetry only for within-subject depression at post-treatment. Trim-and-fill imputed two studies, but the adjusted ES remained significant. Therefore, the combined results for publication bias risk assessment suggest no overall evidence of unpublished studies with negative results.

Table 1

Main Characteristics of the Included Studies

Study	Country	Design	Allocated participants (n)		Sample characteristics			Intervention characteristics					Control group
			Treated	Controls	Diagnosis	Mean age (*SD*)	% of females	Type	Sessions planned	Mean sessions received (*SD*)	Applied by	Delivery	Format	
(Arch et al., 2019)	United States of America	Uncontrolled trial	35	-	Metastatic gastrointestinal, breast, gynecologic, prostate, myeloma and lung cancer	61.8 (11.9)	62.9	ACT	7	6.2 (0.8)	Combined	Combined	Combined	
(Gould et al., 2023)	United Kingdom	Uncontrolled trial	30	-	Amyotrophic lateral sclerosis	58.4 (13.8)	48.0	ACT	8	5.5 (3.4)	Combined	Combined	Individual	
(Gould et al., 2024)	United Kingdom	RCT	97	94	Amyotrophic lateral sclerosis	63.1 (11.0)	58.0	ACT	8	N/A	Therapist	Combined	Individual	Usual care: medication, symptom-management treatment, and hospital/ community services
(Hu et al., 2024)	China	Nonrandomized controlled trial	76	75	Advanced gastric cancer	57.0 (N/A)	33.8	ACT	N/A	N/A	Therapist	Face-to-face	N/A	Usual care (routine inpatient care, brief emotional support, patient education, and management of chemotherapy side effects) + personalized nutrition and usual care only
(Hulbert-Williams et al., 2021)	United Kingdom	Uncontrolled trial	10	-	Non-curative breast, esophageal, prostate, cervical, lung, pancreatic, bladder, colon with liver metastases, digestive/ peritoneal cancers	65.7 (11.5)	40.0	ACT	5	N/A	Therapist	Face-to-face	Individual	
(Li et al., 2024)	China	RCT	80	80	Advanced lung cancer	59.5 (7.6)	20.6	ACT	4	3.4 (N/A)	Therapist	Face-to-face	Individual	Usual care: brief lung cancer/treatment education and daily care, medication & side-effect instructions, routine physical checks, plus discharge advice on diet/exercise and retesting
(Martin & Pakenham, 2024)	Australia	Nonrandomized controlled trial	52	54	Palliative care patients: lung, breast, colorectal, gynecological, pancreatic, head & neck, hematological, bone/ soft tissue, prostate, skin, urological cancers; cardiovascular disease, respiratory failure, liver failure, renal failure, motor neuron disease	69.1 (9.4)	52.6	ACT	4 (modules)	N/A	Combined	Remote	Individual	Usual care: usual specialist multidisciplinary palliative care from the same service

Study	Country	Design	Allocated participants (n)		Sample characteristics			Intervention characteristics					Control group
			Treated	Controls	Diagnosis	Mean age (*SD*)	% of females	Type	Sessions planned	Mean sessions received (*SD*)	Applied by	Delivery	Format	
(Mosher et al., 2022)	United States of America	RCT	20	20	Advanced gastrointestinal cancer, including colorectal, pancreatic, stomach and liver	58.5 (10.0)	55.0	ACT	6	4.8 (2)	Therapist	Remote	Combined	Education support: 6 weekly ~50-min telephone sessions with *supportive listening* plus practical/health information and referrals (contact info for support services)
(Mosher et al., 2024)	United States of America	RCT	116	120	Metastatic breast cancer	59.0 (11.5)	100	ACT	6	5.4 (1.5)	Therapist	Remote	Individual	Education support: 6 weekly 50-min telephone sessions delivered by therapists with supportive-counseling experience
(Plumb Vilardaga et al., 2020)	United States of America	Uncontrolled trial	24	-	Advanced prostate, colorectal, lung, renal, breast, appendiceal, gallbladder cancer	66.0 (10.8)	29.0	ACT + Coping skills training	4	3.9 (0.6)	Therapist	Remote	Individual	
(Ramos et al., 2018)	United States of America	Uncontrolled trial	39	-	Palliative care patients with a life-limiting illness: advanced prostate and lung cancer; end-stage renal disease, advanced neurological or autoimmune disorders	63.1 (8.6)	N/A	ACT + CBT	6-8	N/A	Therapist	Face-to-face	Group	
(Rost et al., 2012)	United States of America	RCT	25	22	Late-stage ovarian cancer	56.0 (N/A)	100.0	ACT	12	N/A	Therapist	Face-to-face	Individual	Active CBT-style protocol (12 face-to-face sessions) including relaxation training, problem-solving, and a "five-column" cognitive restructuring
(Serfaty et al., 2019)	United Kingdom	RCT	20	20	Advance breast, colon, myeloma, prostate, bowel, non-Hodgkin lymphoma cancer	62.0 (11.5)	74.0	ACT	8	N/A	Therapist	Face-to-face	Individual	Talking Control: up to 8 sessions of structured empathic/ neutral supportive conversation, explicitly avoiding active therapeutic techniques (e.g., mindfulness, cognitive restructuring, problem-solving)

Note. RCT = Randomized controlled trial; ALS = Amyotrophic lateral sclerosis; N/A = Not available; ACT = acceptance and commitment therapy; CBT = Cognitive behavioral therapy.

Figure 1
PRISMA Flow Diagram

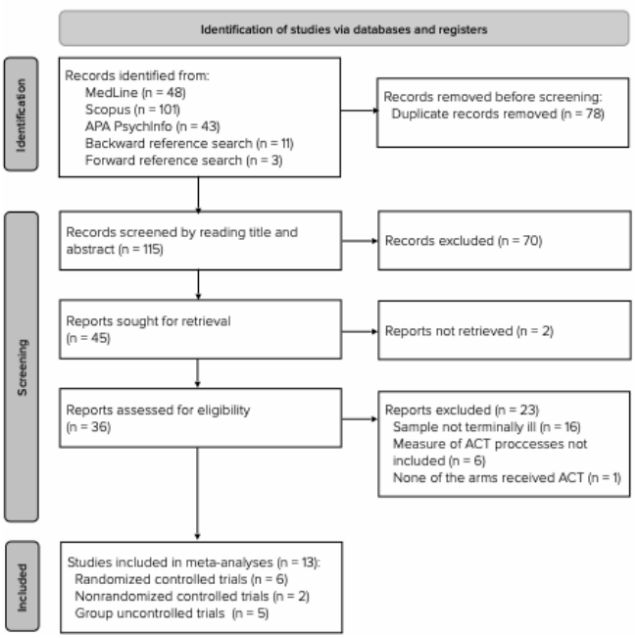


Figure 2a
Risk of Bias Summary: RoB 2 (Upper Panel)

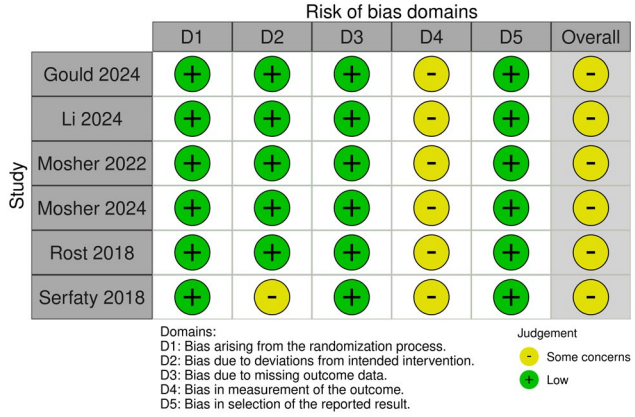
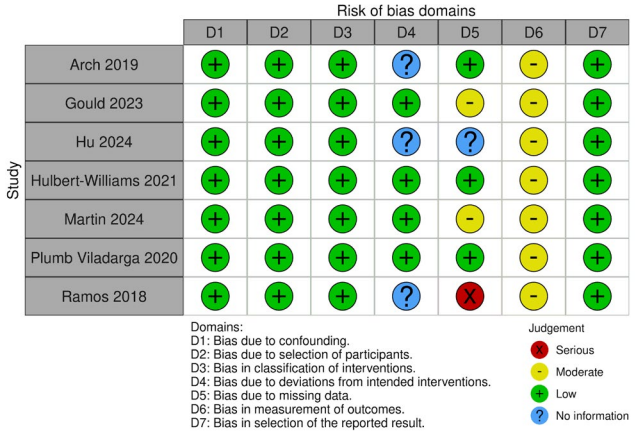


Figure 2b
Risk of Bias Summary: ROBINS-I (Bottom Panel)



Meta-Analyses Results

The full results of the meta-analysis are shown in Table 2, while Figures 3 and 4 show forest-plots for post-treatment QoL. See the rest of the forest-plots in Supplementary Material. Between-group analyses showed that, compared with participants in active control groups, those who received ACT improved significantly more on all outcome variables, with approximately moderate ESs (SMDs = 0.44 to 0.62). These improvements remained significant at follow-up, increasing for QoL (SMD = 0.56), while for anxiety and depression symptoms decreased (SMDs = 0.35 to 0.40). In terms of within-subject change, ES estimates were similar for QoL (post-treatment SMD = 0.46; follow-up SMD = 0.52) but smaller for anxiety and depression symptoms (SMDs = 0.19 to 0.45). However, the effect was non-significant for within-subject QoL at follow-up (CI 95% [-0.02 to 1.07], $p = .06$). Heterogeneity was high for QoL (I^2 = 89% to 94%) and low-to-moderate for anxiety and depression (I^2 = 22% to 60%, except between-group anxiety: I^2 = 93%).

Table 2
Standardized Mean Differences (SMD), 95% Confidence Intervals, Heterogeneity Analyses (I^2), and Risk of Bias for Between-Groups and Within-Subject Outcomes

Variable			Between-groups outcomes					Within-subject outcomes				
		Time point	*k*	*n*	SMD	95% CI	I^2	*k*	*n*	SMD	95% CI	I^2
Outcome variables	Quality of Life	PT	9	1051	0.44	0.10 to 0.79	90%	10	477	0.46	0.09 to 0.84	92.%
		FU	5	615	0.56	0.17 to 0.95	89%	6	314	0.52	-0.02 to 1.07	94%
	Anxiety symptoms	PT	4	442	0.62	0.06 to 1.19	93%	8	305	0.41	0.26 to 0.56	60%
		FU	2	305	0.40	0.21 to 0.60	22%	3	180	0.45	0.25 to 0.64	56%
	Depression symptoms	PT	4	442	0.58	0.33 to 0.82	59%	8	305	0.25	0.15 to 0.35	47%
		FU	2	305	0.35	0.13 to 0.58	36%	3	180	0.19	0.03 to 0.36	43%
Process variables	Psychological flexibility	PT	6	515	0.49	0.08 to 0.90	80%	10	332	0.51	0.02 to 1.01	92%
		FU	4	372	0.55	0.04 to 1.07	89%	5	209	0.73	-0.27 to 1.74	96%
	Valued action	PT	4	318	0.62	-0.17 to 1.41	94%	6	199	0.57	-0.32 to 1.46	92%
		FU	3	214	0.69	-0.50 to 1.89	95%	5	149	0.60	-0.45 to 1.66	97%

Note. PT = post-treatment; FU = follow-up; *k* = number of studies; *n* = number of participants; SMD = standardized mean differences; CI = confidence intervals

Figure 3
Forest Plot for Between-Group QoL at Post-Treatment

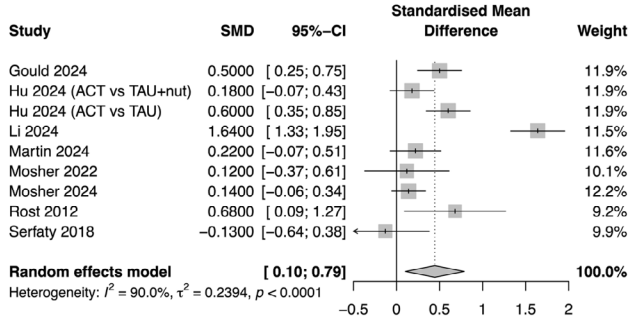
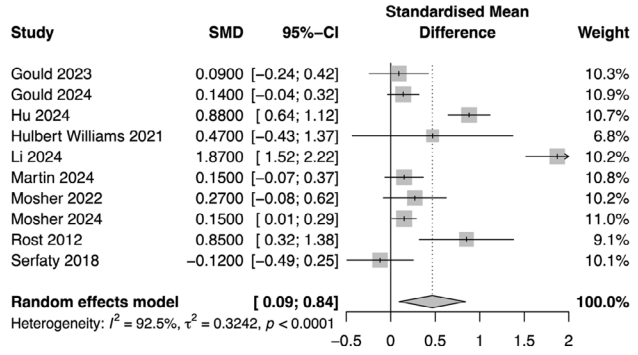


Figure 4
Forest Plot for Within-Subject QoL at Post-Treatment



Regarding process variables, those who received ACT improved significantly more than active controls on psychological flexibility, with moderate ESs (SMD = 0.49) that were maintained at follow-up (SMD = 0.55). The ES for within-subject change was similarly moderate (SMD = 0.51) but became non-significant at follow-up for this variable (CI 95% [-0.27 to 1.74], p = .15). However, ES for the valued action were moderate to large (SMDs 0.57 to 0.69), but confidence intervals were very wide, mainly due to the small sample size, and results were not statistically significant for both between-group and within-subject outcomes. Heterogeneity was high for all process variable outcomes (I^2 > 80%).

Examining the Role of the Process Variables in QoL and Anxiety and Depression Symptom Improvement

Valued action and psychological flexibility were the best predictors of QoL improvements (see Table 3, and Figures 5 and 6). These process variables significantly predicted all QoL outcomes, explaining between 75% and 100% of the variance. However, meta-regression using changes in between-group psychological flexibility as predictor after the treatment was not significant (p = .18) and explained only 18% of the variance.

Figure 5
Scatterplot for Psychological Flexibility Predicting Post-Treatment QoL

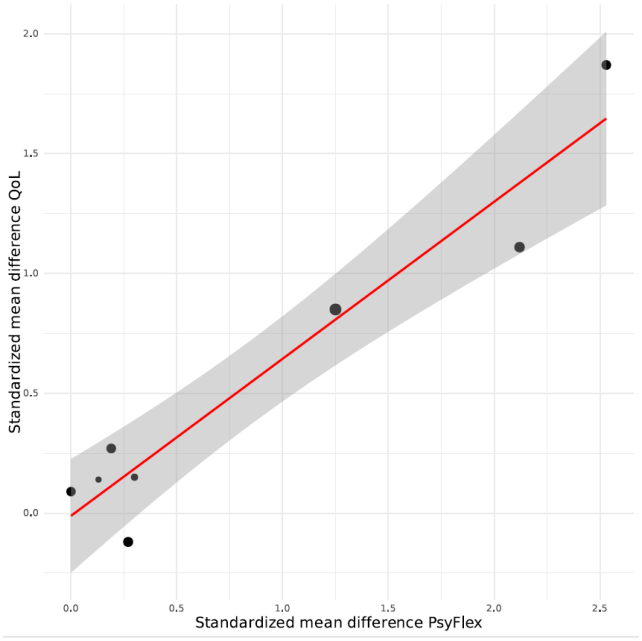


Figure 6
Scatterplot for Values Predicting Post-Treatment QoL

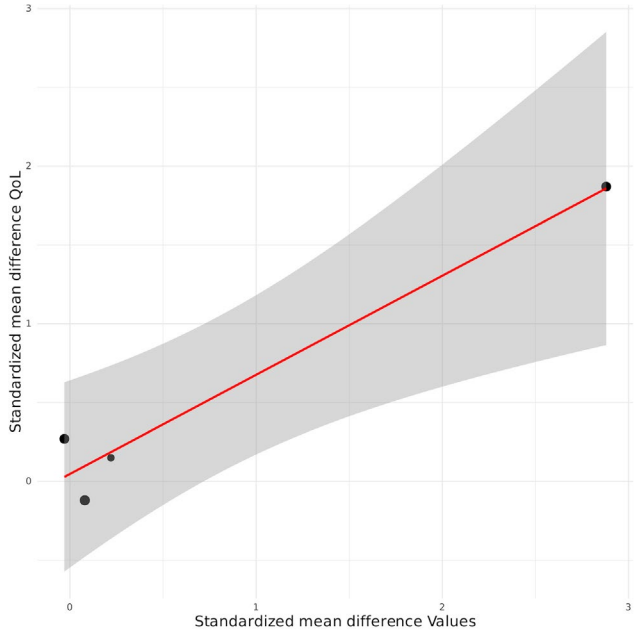


Table 3
Meta-Regression Results

Outcome	Covariates																			
	Percentage of women				Age of participants				Number of sessions				Psychological flexibility				Valued action			
	ß	SE	*p*	R^2	ß	SE	*p*	R^2	ß	SE	*p*	R^2	ß	SE	*p*	R^2	ß	SE	*p*	R^2
BG QoL (post-treatment)	-0.01	0.01	.17	.14	-0.02	0.05	.63	0	-0.04	0.09	.67	0	0.61	0.45	.18	.18	**0.98**	**0.12**	**<.001**	**1**
BG QoL (follow-up)	**-0.01**	**0.01**	**.01**	**.73**	0.03	0.10	.78	0	-0.09	0.14	.51	0	**0.71**	**0.19**	**<.001**	**1**	**0.44**	**0.23**	**.04**	**.75**
WS QoL (post-treatment)	-0.01	0.01	.13	.16	-0.05	0.05	.29	.02	-0.05	0.09	.60	0	**0.72**	**0.08**	**<.001**	**1**	**0.63**	**0.08**	**<.001**	**.98**
WS QoL (follow-up)	-0.01	0.01	.09	.31	-0.06	0.11	.60	0	-0.25	0.16	.13	.26	**0.63**	**0.08**	**<.001**	**0.98**	**0.53**	**0.16**	**<.001**	**.87**
BG anxiety (post-treatment)	-0.01	0.01	.29	.09	-0.05	0.06	.35	.06	-0.05	0.10	.61	0	0.32	0.59	.59	0	Not enough studies			
BG anxiety (follow-up)	Not enough studies for all the covariates																			
WS anxiety (post-treatment)	0.01	0.01	.19	0	-0.03	0.02	.07	.18	0.05	0.04	.28	0	**0.13**	**0.05**	**<.01**	**.99**	0.06	0.05	.19	1
WS anxiety (follow-up)	-0.01	0.01	.88	0	-0.05	0.05	.29	.11	-0.04	0.07	.59	0	0.03	0.11	.77	0	Not enough studies			
BG depression (post-treatment)	-0.01	0.01	.35	.28	-0.03	0.03	.24	.28	-0.02	0.05	.66	0	0.32	0.24	.18	.54	Not enough studies			
BG depression (follow-up)	Not enough studies for all the covariates																			
WS depression (post-treatment)	0.01	0.01	.38	0	-0.01	0.02	.54	0	0.05	0.03	.14	0	-0.01	0.07	.98	0	-0.05	0.05	.29	0
WS depression (follow-up)	-0.04	0.02	.11	.28	-0.01	0.21	.46	0	-0.37	0.31	.22	.11	Not enough studies				Not enough studies			

Note. BG = between-group, WS = within-subject; not enough studies (n = 2)

Meta-Regression Results

On the other hand, for the remaining meta-regressions, the other covariates (e.g., percentage of women in the sample, mean age of participants, number of sessions) did not significantly predict any of the outcomes, with the only exception of between-group QoL at follow-up (p = .01, R^2 = .73).

Subgroup Analyses Results

These analyses must be interpreted with caution due to the small sample sizes in most subgroups. The clearer subgroup differences were found in favor of face-to-face intervention delivery (compared to remote or combined deliveries) in improving QoL outcomes, both after treatment ($\chi^2(2) = 10.36$; $p < .01$ for between-group and $\chi^2(2) = 7.81$; $p = .02$ for within-subject outcomes) and at follow-up ($\chi^2(2) = 35.43$; $p < .01$ for between-group and $\chi^2(2) = 29.04$; $p < .01$ for within-subject outcomes). In addition, advanced cancer patients improved significantly more than ALS patients in depression outcomes at post-treatment ($\chi^2(1) = 12.28$; $p < .01$). Subgroup analyses were not statistically significant for the remaining outcomes.

Sensitivity Analyses Results

Sensitivity analyses revealed no statistically significant differences between randomized and non-randomized controlled trials (all $p > .05$). In addition, sensitivity analyses showed that effects on within-subject depression at follow-up became non-significant (SMD = 0.26, 95% CI [-0.13 to 0.65], $p = .19$) after removing studies that included mixed stage III and IV cancer patients (see Supplementary Table 1). All the other effects on outcome and process variables remained significant after removing those studies, and the heterogeneity was overall decreased (I^2 = 0% to 75% versus I^2 = 22% to 97%).

Discussion

This review had two primary objectives. First, it aimed to quantitatively synthesize evidence on the effects of ACT on QoL and depression and anxiety symptoms in people with advanced progressive illness. Second, it examined the role of process variables, such as psychological flexibility and values-based action, in their association with these clinical improvements. This meta-analysis found that ACT showed moderate effects in improving QoL and emotional symptoms in terminally ill patients when compared to active controls and in pre-post comparisons, though heterogeneity was high across outcomes. These effects were generally maintained at follow-up; psychological flexibility and valued action were associated with QoL improvements and emerged as candidate process-related correlates.

Our study showed that ACT was associated with statistically significant improvements in QoL in both controlled comparisons (e.g., usual care or educational support) and within-subject analyses. In some cases, these gains were maintained over time. These findings are consistent with Fang et al. (2023), who also reported a large ES for QoL improvement in advanced cancer patients. This may be because they included fewer studies and participants. These findings have been confirmed by other systematic reviews in similar populations (Li et al., 2021). However, our review uniquely expands the scope to include a broader range of terminally ill patients, such as those in palliative care and those with ALS.

The findings showed that increases in psychological flexibility were associated with QoL improvements, consistent with ACT's theoretical model, which proposes that improvements in clinical outcomes occur as consequences of changes in underlying psychological processes rather than as direct therapeutic targets (Hayes, 2020). Future research should examine whether ACT's emphasis on acceptance and values may help address existential concerns often present in end-of-life care, such as fear of death or uncertainty, which fall outside traditional diagnostic frameworks (Lee, 2020). Importantly, values work should be individualized and evaluated in terms of contextual fit and behavioral function, as rigid rule-following may also contribute to distress in some cases. However, the meta-regression analyses examining valued action as a predictor should be interpreted with considerable caution, as the R^2 values of 1.00 for several outcomes (QoL at post-treatment, and within-subject anxiety at post-treatment) suggest potential overfitting, likely due to the limited number of studies (k = 2-8 across models) relative to the complexity of the model, which may have resulted in perfect separation or multicollinearity issues rather than genuine moderation effects.

Research has highlighted the health benefits of having supportive relationships, which may promote physical and mental well-being (Cacioppo & Patrick, 2008). Conversely, loneliness and social isolation have been associated with increased mortality risk, with effect sizes comparable to well-established risk factors (Holt-Lunstad et al., 2015). These findings suggest ways to develop therapeutic interventions that can alleviate the adverse effects of social stress and improve overall QoL during critical life stages.

Although our meta-regression analyses for valued action were limited by small sample sizes and should be interpreted with caution, the theoretical framework of ACT suggests that value-based interventions may be clinically useful when they are workable in context and guided by the functional analysis of the individual's behavior (Menzies & Menzies, 2024; Wersebe et al., 2017). While previous evidence suggests that values clarification and valued action often precede a reduction in symptoms (Gloster et al., 2020), our findings for valued action as a predictor were inconclusive due to limited data and potential overfitting.

According to our results, psychological flexibility was an important predictor of long-term QoL improvement, particularly in the context of terminal illness. Improvements in psychological flexibility, which encompasses acceptance of difficult internal experiences, may have a cumulative positive impact on QoL over time. Our findings showed near-large effect sizes for psychological flexibility, both compared to active controls and within subjects, with slight improvements noted at follow-up. These results are consistent with a potentially important role of psychological flexibility in improving QoL.

Our findings differ from previous reviews by Fang et al. (2023), which reported no significant effect of psychological flexibility in the context of cancer and advanced cancer. The absence of clear findings on psychological flexibility in these studies could reflect methodological limitations. Many trials focus primarily on clinical outcomes such as anxiety, depression, or QoL, rather than directly assessing flexibility through validated tools, such as the AAQ-II. Furthermore, the heterogeneity of populations and interventions, coupled with limited process-focused measures, constrains the ability to detect significant effects. In our paper, we acknowledge these gaps and emphasize the importance of future research to systematically include flexibility assessments to capture ACT's core processes of change in end-of-life contexts. In contrast, Zhao et al. (2021) found a moderate ES for this variable. Our estimates for valued action were moderate but not statistically significant across all outcomes. This is likely due to the limited sample size in studies assessing these variables and the high heterogeneity of the results. The analysis shows that improvements may co-occur with, or be associated with, increases in psychological flexibility and, to a lesser extent, valued action. However, valued action should not be assumed to be uniformly beneficial, as its impact depends on the function of the behavior and the contextual contingencies in which it occurs. These findings suggest that psychological flexibility may be an important therapeutic process in end-of-life care, though further research with larger samples is needed to confirm this relationship.

Some argue that ACT promotes psychological flexibility by encouraging acceptance of thoughts and emotions rather than attempting to suppress them (Hayes & Hofmann, 2021). As early as the 1980s, research showed that attempting to control or suppress thoughts can paradoxically intensify them (Wegner et al., 1987). Conversely, accepting the symptoms often leads to their reduction (Hayes, Strosahl, et al., 2006). Together, these findings and the long-term effect on QoL highlight the need to consider delayed response patterns when evaluating the efficacy of ACT. This may help clinicians set realistic expectations that focus on improving QoL rather than immediate symptom relief. Our results are consistent with previous studies that have emphasized the importance of acceptance when coping with distress associated with terminal illness (Davis et al., 2017; Secinti et al., 2019). Our consistent results across between-group and within-subject outcomes align with previous reviews in cancer patients (Fang et al., 2023; Li et al., 2021) and advanced illness patients (Gibson Watt et al., 2023), though the ESs were smaller. Unlike QoL, ACT process variables did not consistently predict improvements in these emotional symptoms. Anxiety and depression in terminal illness may arise from specific fears, such as mortality and loss of autonomy (Butow et al., 2021; Hong et al., 2022), which are not fully addressed by reducing experiential avoidance. Overall, while ACT moderately reduces anxiety and depression in terminally ill patients, these symptoms are multifaceted and may be influenced by factors beyond psychological flexibility.

First, it is important to note that the results suggest that none of the other continuous variables (e.g., age, percentage of women, number of sessions) significantly predicted treatment outcome. These null findings may suggest that effects do not vary markedly by age or gender, although tests may be underpowered**.** Furthermore, even shorter interventions (e.g., four sessions) may yield similar ES to longer treatments (e.g., 12 sessions). Subgroup tests suggested larger effects for face-to-face delivery, though these results require cautious interpretation. This finding may suggest that face-to-face interactions may play an important role for terminally ill patients, perhaps by strengthening the therapeutic relationship, reducing distractions, and providing richer nonverbal communication (Berle et al., 2015). In contrast, therapy format (e.g., individual versus group therapy) does not appear to affect treatment outcomes. Finally, ACT's versatility is further supported by its potential benefits in treating terminally ill patients with different diagnoses (e.g., advanced cancer and ALS).

Our results suggest that ACT was associated with statistically significant improvements in QoL among people with advanced

progressive illnesses, with potential benefits beyond advanced cancer, though data for non-oncological diagnoses, such as ALS, remain limited. These findings suggest that ACT may be a promising approach for addressing some of the complex challenges of end-of-life care. Clinicians may consider ACT as one therapeutic option for terminally ill patients, particularly when the goal is to enhance psychological flexibility and QoL. Furthermore, ACT has shown utility for caregivers, addressing a critical gap in care systems (Gibson Watt et al., 2023). However, improving the QoL of people in palliative care does not necessarily improve family well-being, underscoring the necessity for tailored interventions for caregivers (Dionne-Odom et al., 2015).

There are several limitations to the present study that should be noted. First, a potential source of bias was identified primarily due to the self-reported nature of the measurement of the variables. However, it has been argued that in the context of psychotherapy research, there is limited evidence of an overestimation of effects due to self-reported outcomes (Munder & Barth, 2018). Future studies should incorporate multi-method assessment of outcomes and intervention-related change (e.g., behavioral/observational indicators of change process, or informant reports) to strengthen inferences beyond self-reported change. Second, we observed a high heterogeneity across studies for most outcomes, which is consistent with prior reviews in studies on ACT for cancer patients (Fang et al., 2023; González-Fernández & Fernández-Rodríguez, 2019; Sauer et al., 2024). This may complicate the interpretation of the true overall effect and limit the generalizability of the findings. Third, some evidence suggests that one of the main measures of psychological flexibility is highly correlated with psychological distress (Langer et al., 2024). This could limit the interpretation of meta-regression analyses examining the relationship between psychological flexibility and symptoms of anxiety and depression. In this sense, the meta-regressions using valued action as a predictor with R^2 values of 1 should be interpreted with caution, because of the risk of potential overfitting and multicollinearity issues. Fourth, the trim-and-fill test revealed some evidence of publication bias for QoL at the follow-up. Conversely, we did not detect evidence of publication bias for the other outcomes. Fifth, the small sample size in some meta-regressions and some groups in the subgroup analyses may affect the validity of these analyses and should be interpreted with caution. Sixth, the inclusion of stage III cancer patients in some studies may have introduced heterogeneity, though sensitivity analyses confirmed the robustness of the findings. Seventh, the inclusion of single-arm trials in within-subject analyses limits causal inferences for these estimates, though our primary conclusions are based on controlled between-group comparisons from controlled studies. Eighth, the inclusion of different study designs (RCTs, quasi-experimental, and single-arm studies) may limit direct comparability and requires cautious interpretation. However, we addressed this by separating controlled (between-group) from uncontrolled (within-subject) analyses and conducting sensitivity analyses comparing randomized versus non-randomized studies. Finally, although RCTs can reduce several threats to internal validity and provide estimates of average effects, they may have limited ecological validity for individualized, function-guided clinical practice (Pérez Álvarez, 2020); therefore, future research should complement group trials with idiographic, repeated-measures designs (e.g., single-case experimental designs and intensive time-series assessment) to evaluate clinically meaningful change at the individual level. Relatedly, these findings should not be interpreted as supporting a one-size-fits-all protocol; in end-of-life care, intervention components, particularly values work, should be individualized and evaluated with respect to behavioral function and contextual constraints.

In conclusion, integrating ACT interventions and addressing key psychological processes may improve our understanding of psychotherapy's potential to mitigate psychological distress in palliative care. ACT's emphasis on addressing existential dilemmas through values-based interventions provides a distinctive approach to preparing patients for the end-of-life. Although it is often difficult to feel fully prepared for the ethical and existential challenges of terminal illness, an existential perspective may offer a framework through which individuals can be guided by their values toward an authentic, dignified, and humane farewell. This, in turn, may be associated with improved QoL, even in the final moments. Our findings are broadly consistent with evidence suggesting that ACT can be a promising approach in palliative care, highlighting links between process-based interventions and process-related correlates of change. Taken together, these findings suggest that ACT may be associated with improvements in QoL and psychological distress, potentially by addressing existential dilemmas and fostering adaptive behaviors. This integration paves the way for future research at the intersection of psychotherapy and change processes, particularly values-based action, in end-of-life care. In this context, values work should be understood as clinically relevant when it is workable in the person's situation and informed by functional assessment, rather than assumed to be inherently beneficial. By promoting values-driven behavior and psychological flexibility, ACT may offer a compassionate, evidence-based approach that could support terminally ill patients, although more rigorous and adequately powered trials are needed to confirm effects and clarify mechanisms across settings and diagnoses.

Author Contributions

Juanjo Macías: Conceptualization, Writing - Original Draft. **Lucía Lara-Merín**: Writing - Review & Editing, Investigation. **Carlos López-Pinar**: Conceptualization, Writing - Original Draft, Investigation, Methodology, Formal Analysis, Visualization, Project Administration.

Acknowledgements

We wish to express our gratitude to Steven C. Hayes and Carmen Luciano for reviewing the manuscript and providing insightful comments that enhanced the final version of this article.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Conflict of Interest

In recent years, Carlos López-Pinar has received financial compensation for conducting training in behavioral therapies. This background is acknowledged as a potential source of bias when

interpreting some of the topics discussed in this article. The rest of the authors have no conflict of interests to declare.

Data Availability Statement

Both the database and R code used for data analysis are available at Open Science Framework in this link: https://osf.io/3jwty/?view_only=9b988d4f6be14fa8b4999360747a3dee

References

Arch, J. J., Fishbein, J. N., Ferris, M. C., Mitchell, J. L., Levin, M. E., Slivjak, E. T., Andorsky, D. J., & Kutner, J. S. (2019). Acceptability, feasibility, and efficacy potential of a multimodal acceptance and commitment therapy intervention to address psychosocial and advance care planning needs among anxious and depressed adults with metastatic cancer. *Journal of Palliative Medicine*, *23*(10), 1380–1385. https://doi.org/10.1089/jpm.2019.0398

Balduzzi, S., Rücker, G., & Schwarzer, G. (2019). How to perform a meta-analysis with R: A practical tutorial. *Evidence-Based Mental Health*, *22*(4), 153–160. https://doi.org/10.1136/ebmental-2019-300117

Berle, D., Starcevic, V., Milicevic, D., Hannan, A., Dale, E., Brakoulias, V., & Viswasam, K. (2015). Do patients prefer face-to-face or internet-based therapy? *Psychotherapy and Psychosomatics*, *84*(1), 61–62. https://doi.org/10.1159/000367944

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2010). A basic introduction to fixed-effect and random-effects models for meta-analysis. *Research Synthesis Methods*, *1*(2), 97–111. https://doi.org/10.1002/jrsm.12

Butow, P., Müller, F., Napier, C. E., Bartley, N., Ballinger, M. L., Biesecker, B., Juraskova, I., Meiser, B., Schlub, T. E., Thomas, D. M., Goldstein, D., Best, M. C., Newson, A., Tucker, K., Vines, R., Vines, K., Kirk, J., Young, M. A., Savard, J., & Jacobs, C. (2021). Longitudinal patterns in fear of cancer progression in patients with rare, advanced cancers undergoing comprehensive tumour genomic profiling. *Psycho-Oncology*, *30*(11), 1920–1929. https://doi.org/10.1002/pon.5764

Cacioppo, J. T., & Patrick, W. (2008). *Loneliness: Human nature and the need for social connection*. W. W. Norton & Company.

Campbell, C. L., & Campbell, L. C. (2012). A systematic review of cognitive behavioral interventions in advanced cancer. *Patient Education and Counseling*, *89*(1), 15–24. https://doi.org/10.1016/j.pec.2012.06.019

Carlsson, M. E. (2012). Sleep disturbance in relatives of palliative patients cared for at home. *Palliative and Supportive Care*, *10*(3), 165–170. https://doi.org/10.1017/S1478951511000836

Ciarrochi, J., Hayes, L., Quinlen, G., Sahdra, B., Ferrari, M., & Yap, K. (2022). Letting go, creating meaning: The role of Acceptance and Commitment Therapy in helping people confront existential concerns and lead a vital life. In R. G. Menzies, R. E. Menzies, & G. Dingle (Eds.), *Existential concerns and cognitive-behavioral procedures. An integrative approach*. Springer Nature.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates.

Davis, E. L., Deane, F. P., Lyons, G. C. B., & Barclay, G. D. (2017). Is higher acceptance associated with less anticipatory grief among patients in palliative care? *Journal of Pain and Symptom Management*, *54*(1), 120–125. https://doi.org/10.1016/j.jpainsymman.2017.03.012

Dionne-Odom, J. N., Azuero, A., Lyons, K. D., Hull, J. G., Tosteson, T., Li, Z., Li, Z., Frost, J., Dragnev, K. H., Akyar, I., Hegel, M. T., & Bakitas, M. A. (2015). Benefits of early versus delayed palliative care to informal family caregivers of patients with advanced cancer: Outcomes from the ENABLE III randomized controlled trial. *Journal of Clinical Oncology*, *33*(13), 1446–1452. https://doi.org/10.1200/JCO.2014.58.7824

Duval, S., & Tweedie, R. (2000). Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics*, *56*(2), 455–463. https://doi.org/10.1111/j.0006-341X.2000.00455.x

Egger, M., Davey Smith, G., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *BMJ*, *315*(7109), 629–634. https://doi.org/10.1136/bmj.315.7109.629

Eriksen, M. B., & Frandsen, T. F. (2018). The impact of patient, intervention, comparison, outcome (PICO) as a search strategy tool on literature search quality: A systematic review. *Journal of the Medical Library Association*, *106*(4), 420–431. https://doi.org/10.5195/jmla.2018.345

Fang, P., Tan, L., Cui, J., & Yu, L. (2023). Effectiveness of acceptance and commitment therapy for people with advanced cancer: A systematic review and meta-analysis of randomized controlled trials. *Journal of Advanced Nursing*, *79*(2), 519–538. https://doi.org/10.1111/jan.15543

Gibson Watt, T., Gillanders, D., Spiller, J. A., & Finucane, A. M. (2023). Acceptance and commitment therapy (ACT) for people with advanced progressive illness, their caregivers and staff involved in their care: A scoping review. *Palliative Medicine*, *37*(8), 1100–1128. https://doi.org/10.1177/02692163231183101

Gloster, A. T., Walder, N., Levin, M. E., Twohig, M. P., & Karekla, M. (2020). The empirical status of acceptance and commitment therapy: A review of meta-analyses. *Journal of Contextual Behavioral Science*, *18*, 181–192. https://doi.org/10.1016/j.jcbs.2020.09.009

González-Fernández, S., & Fernández-Rodríguez, C. (2019). Acceptance and commitment therapy in cancer: Review of applications and findings. *Behavioral Medicine*, *45*(3), 255–269. https://doi.org/10.1080/08964289.2018.1452713

González-Fernández, S., Fernández-Rodríguez, C., Paz-Caballero, M., & Pérez-Álvarez, M. (2018). Treating anxiety and depression of cancer survivors: Behavioral activation versus acceptance and commitment therapy. *Psicothema*, *30*(1), 14–20. https://doi.org/10.7334/psicothema2017.396

Gould, R. L., McDermott, C. J., Thompson, B. J., Rawlinson, C. V., Bursnall, M., Bradburn, M., Kumar, P., Turton, E. J., White, D. A., Serfaty, M. A., Graham, C. D., McCracken, L. M., Goldstein, L. H., Al-Chalabi, A., Orrell, R. W., Williams, T., Noad, R., Baker, I., Faull, C., … Waterhouse, S. (2024). Acceptance and commitment therapy plus usual care for improving quality of life in people with motor neuron disease (COMMEND): A multicentre, parallel, randomised controlled trial in the UK. *Lancet*, *403*(10442), 2381–2394. https://doi.org/10.1016/S0140-6736(24)00533-6

Gould, R. L., Rawlinson, C., Thompson, B., Weeks, K., Gossage-Worrall, R., Cantrill, H., Serfaty, M. A., Graham, C. D., McCracken, L. M., White, D., Howard, R. J., Bursnall, M., Bradburn, M., Al-Chalabi, A., Orrell, R., Chhetri, S. K., Noad, R., Radunovic, A., Williams, T., … Tallentire, L. (2023). Acceptance and commitment therapy for people living with motor neuron disease: An uncontrolled feasibility study. *Pilot and Feasibility Studies*, *9*(1), Article 116. https://doi.org/10.1186/s40814-023-01354-7

Han, J., Liu, J. E., Su, Y. L., & Qiu, H. (2019). Effect of a group-based acceptance and commitment therapy (ACT) intervention on illness cognition in breast cancer patients. *Journal of Contextual Behavioral Science*, *14*, 73–81. https://doi.org/10.1016/j.jcbs.2019.09.003

Hayes, S. C. (2020). *A liberated mind: How to pivot toward what matters*. Avery Publishing Group.

Hayes, S. C., Ciarrochi, J., Hofmann, S. G., Chin, F., & Sahdra, B. (2022). Evolving an idionomic approach to processes of change: Towards a unified personalized science of human improvement. *Behaviour Research and Therapy*, *156*, Article 104155. https://doi.org/10.1016/j.brat.2022.104155

Hayes, S. C., & Hofmann, S. G. (2021). "Third-wave" cognitive and behavioral therapies and the emergence of a process-based approach to intervention in psychiatry. *World Psychiatry*, *20*(3), 363–375. https://doi.org/10.1002/wps.20884

Hayes, S. C., Hofmann, S. G., & Ciarrochi, J. (2020). A process-based approach to psychological diagnosis and treatment: The conceptual and treatment utility of an extended evolutionary meta model. *Clinical Psychology Review*, *82*, Article 101908. https://doi.org/10.1016/j.cpr.2020.101908

Hayes, S. C., Levin, M. E., Plumb-Vilardaga, J., Villatte, J. L., & Pistorello, J. (2013). Acceptance and commitment therapy and contextual behavioral science: Examining the progress of a distinctive model of behavioral and cognitive therapy. *Behavior Therapy*, *44*(2), 180–198. https://doi.org/10.1016/j.beth.2009.08.002

Hayes, S. C., Luoma, J. B., Bond, F. W., Masuda, A., & Lillis, J. (2006). Acceptance and commitment therapy: Model, processes and outcomes. *Behaviour Research and Therapy*, *44*(1), 1–25. https://doi.org/10.1016/j.brat.2005.06.006

Hayes, S. C., Strosahl, K. D., & Wilson, K. G. (2006). *Acceptance and commitment therapy: The process and practice of mindful change*. Guilford Press.

Higgins, J. P. T., & Green, S. (2011). *Cochrane handbook for systematic reviews of interventions.* John Wiley & Sons.

Higgins, J. P. T., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *BMJ*, *327*(7414), 557–560. https://doi.org/10.1136/bmj.327.7414.557

Holt-Lunstad, J., Smith, T. B., Baker, M., Harris, T., & Stephenson, D. (2015). Loneliness and social isolation as risk factors for mortality. *Perspectives on Psychological Science*, *10*(2), 227–237. https://doi.org/10.1177/1745691614568352

Hong, Y., Yuhan, L., Youhui, G., Zhanying, W., Shili, Z., Xiaoting, H., & Wenhua, Y. (2022). Death anxiety among advanced cancer patients: A cross-sectional survey. *Supportive Care in Cancer*, *30*(4), 3531–3539. https://doi.org/10.1007/s00520-022-06795-z

Hu, H., Li, H., Xu, H., Liu, J., & Qi, Y. (2024). Effect of personalized nutrition combined with acceptance and commitment therapy on psychological resilience, quality of life and side effects of chemotherapy in patients with advanced gastric cancer. *British Journal of Hospital Medicine*, *85*(9), 1–16. https://doi.org/10.12968/hmed.2024.0326

Hulbert-Williams, N. J., Norwood, S. F., Gillanders, D., Finucane, A. M., Spiller, J., Strachan, J., Millington, S., Kreft, J., & Swash, B. (2021). Brief engagement and acceptance coaching for hospice settings (the BEACHeS study): Results from a Phase I study of acceptability and initial effectiveness in people with non-curative cancer. *BMC Palliative Care*, *20*(1), Article 96. https://doi.org/10.1186/s12904-021-00801-7

Jiang, X., Sun, J., Song, R., Wang, Y., Li, J., & Shi, R. (2024). Acceptance and commitment therapy reduces psychological distress in patients with cancer: A systematic review and meta-analysis of randomized controlled trials. *Frontiers in Psychology*, *14*, Article 1253266. https://doi.org/10.3389/fpsyg.2023.1253266

Khan, F. A., Akhtar, S. S., & Seikh, M. K. (2005). Cancer treatment - objectives and quality of life issues. *Malaysian Journal of Medical Sciences*, *12*(1), 3–5.

Kissane, D. W. (2012). The relief of existential suffering. *Archives of Internal Medicine*, *172*(19), 1501–1505. https://doi.org/10.1001/archinternmed.2012.3633

Knaul, F. M., Farmer, P. E., Krakauer, E. L., De Lima, L., Bhadelia, A., Jiang Kwete, X., Arreola-Ornelas, H., Gómez-Dantés, O., Rodriguez, N. M., Alleyne, G. A. O., Connor, S. R., Hunter, D. J., Lohman, D., Radbruch, L., del Rocío Sáenz Madrigal, M., Atun, R., Foley, K. M., Frenk, J., Jamison, D. T., … Zimmerman, C. (2018). Alleviating the access abyss in palliative care and pain relief—an imperative of universal health coverage: The Lancet commission report. *Lancet*, *391*(10128), 1391–1454. https://doi.org/10.1016/S0140-6736(17)32513-8

Kredentser, M. S., & Chochinov, H. M. (2020). Psychotherapeutic considerations for patients with terminal illness. *American Journal of Psychotherapy*, *73*(4), 137–143. https://doi.org/10.1176/appi.psychotherapy.20190048

Ksiksou, J., Maskour, L., & Alaoui, S. (2024). Effect of acceptance and commitment therapy on depression, anxiety, stress, and psychological flexibility in women with breast cancer. *Kontakt*, *26*(2), 104–111. https://doi.org/10.32725/kont.2024.023

Lamont, E. B. (2005). A demographic and prognostic approach to defining the end of life. *Journal of Palliative Medicine*, *8*(Suppl. 1), s-12-s-21. https://doi.org/10.1089/jpm.2005.8.s-12

Langer, Á. I., Ponce, F. P., Ordóñez-Carrasco, J. L., Fuentes-Ferrada, R., Mac-Ginty, S., Gaete, J., & Núñez, D. (2024). Psychometric evidence of the acceptance and action questionnaire-II (AAQ-II): An item response theory analysis in university students from Chile. *BMC Psychology*, *12*(1), Article 111. https://doi.org/10.1186/s40359-024-01608-w

Lantheaume, S., Montagne, M., & Shankland, R. (2020). Intervention focused on resources to reduce anxiety and depression disorders in cancer patients: A pilot study. *Encephale*, *46*(1), 13–22. https://doi.org/10.1016/j.encep.2019.07.005

Lee, J. (2020). Existential issues and psychosocial interventions in palliative care. *The Korean Journal of Hospice and Palliative Care*, *23*(4), 167–171. https://doi.org/10.14475/kjhpc.2020.23.4.167

Lee, R., & Zhou, Y. (2017). Does fertility or mortality drive contemporary population aging? The revisionist view revisited. *Population and Development Review*, *43*(2), 285–301. https://doi.org/10.1111/padr.12062

Li, H., Wong, C. L., Jin, X., Chen, J., Chong, Y. Y., & Bai, Y. (2021). Effects of acceptance and commitment therapy on health-related outcomes for patients with advanced cancer: A systematic review. *International Journal of Nursing Studies*, *115*, Article 103876. https://doi.org/10.1016/j.ijnurstu.2021.103876

Li, H., Wong, C. L., Jin, X., Chong, Y. Y., & Ng, M. S. N. (2024). Effects of acceptance and commitment therapy-based intervention on fatigue interference and health-related quality of life in patients with advanced lung cancer: A randomised controlled trial. *Journal of Contextual Behavioral Science*, *32*, Article 100758. https://doi.org/10.1016/j.jcbs.2024.100758

Martin, C. L., & Pakenham, K. I. (2024). Pilot evaluation of a self-help ACT intervention for palliative care patients. *Journal of Contextual Behavioral Science*, *33*, Article 100797. https://doi.org/10.1016/j.jcbs.2024.100797

Menzies, R. G., & Menzies, R. E. (2024). Existential therapies and the extended evolutionary meta-model: Turning existential philosophy into process-based therapy. *Journal of Contextual Behavioral Science*, *34*, Article 100840. https://doi.org/10.1016/j.jcbs.2024.100840

Morris, S. B. (2000). Distribution of the standardized mean change effect size for meta-analysis on repeated measures. *British Journal of Mathematical and Statistical Psychology*, *53*(1), 17–29. https://doi.org/10.1348/000711000159150

Morris, S. B. (2008). Estimating effect sizes from pretest-posttest-control group designs. *Organizational Research Methods*, *11*(2), 364–386. https://doi.org/10.1177/1094428106291059

Mosher, C. E., Lee, S., Addington, E. L., Park, S., Lewson, A. B., Snyder, S., Hirsh, A. T., Bricker, J. B., Miller, K. D., Ballinger, T. J., Schneider, B. P., Storniolo, A. M., Newton, E. V., Champion, V. L., & Johns, S. A. (2024). Randomized controlled trial of acceptance and commitment therapy for fatigue interference with functioning in metastatic breast cancer. *Journal of Clinical Oncology*, *43*(6), 662–671. https://doi.org/10.1200/JCO.24.00965

Mosher, C. E., Secinti, E., Wu, W., Kashy, D. A., Kroenke, K., Bricker, J. B., Helft, P. R., Turk, A. A., Loehrer, P. J., Sehdev, A., Al-Hader, A. A., Champion, V. L., & Johns, S. A. (2022). Acceptance and commitment therapy for patient fatigue interference and caregiver burden in advanced gastrointestinal cancer: Results of a pilot randomized trial. *Palliative Medicine*, *36*(7), 1104–1117. https://doi.org/10.1177/02692163221099610

Munder, T., & Barth, J. (2018). Cochrane's risk of bias tool in the context of psychotherapy outcome research. *Psychotherapy Research*, *28*(3), 347–355. https://doi.org/10.1080/10503307.2017.1411628

Murtagh, F. E., Bausewein, C., Verne, J., Groeneveld, E. I., Kaloki, Y. E., & Higginson, I. J. (2014). How many people need palliative care? A study developing and comparing methods for population-based estimates. *Palliative Medicine*, *28*(1), 49–58. https://doi.org/10.1177/0269216313489367

Okuyama, T., Akechi, T., Mackenzie, L., & Furukawa, T. A. (2017). Psychotherapy for depression among advanced, incurable cancer patients: A systematic review and meta-analysis. *Cancer Treatment Reviews*, *56*, 16–27. https://doi.org/10.1016/j.ctrv.2017.03.012

Organización Médica Colegial & Sociedad Española de Cuidados Paliativos. (2015). *Atención médica al final de la vida: conceptos y definiciones.* [End-of-life healthcare: Concepts and definitions]. OMC-SECPAL.

Páez, M. B., Luciano, C., & Gutiérrez, O. (2007). Tratamiento psicológico para el afrontamiento del cáncer de mama. Estudio comparativo entre estrategias de aceptación y de control cognitivo [Psychological treatment for coping with breast cancer. Comparative study between acceptance and cognitive control strategies]. *Psicooncología*, *4*(1), 75–95.

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., … Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, Article n71. https://doi.org/10.1136/bmj.n71

Pérez Álvarez, M. (2020). El embrollo científico de la psicoterapia: Cómo salir [The scientific mess of psychotherapy: How to get out of it]. *Papeles Del Psicólogo*, *41*(2), 174-183. https://doi.org/10.23923/pap.psicol2020.2944

Plumb Vilardaga, J. C., Winger, J. G., Teo, I., Owen, L., Sutton, L. M., Keefe, F. J., & Somers, T. J. (2020). Coping skills training and acceptance and commitment therapy for symptom management: Feasibility and acceptability of a brief telephone-delivered protocol for patients with advanced cancer. *Journal of Pain and Symptom Management*, *59*(2), 270–278. https://doi.org/10.1016/j.jpainsymman.2019.09.005

Quevedo, Y., Booij, L., Herrera, L., Hernández, C., & Jiménez, J. P. (2022). Potential epigenetic mechanisms in psychotherapy: A pilot study on DNA methylation and mentalization change in borderline personality disorder. *Frontiers in Human Neuroscience*, *16*, Article 955005. https://doi.org/10.3389/fnhum.2022.955005

R Core Team. (2025). *R: A language and environment for statistical computing (Version 4.5.1) [Computer software]*. R Foundation for Statistical Computing. https://www.R-project.org/

Radbruch, L., De Lima, L., Knaul, F., Wenk, R., Ali, Z., Bhatnaghar, S., Blanchard, C., Bruera, E., Buitrago, R., Burla, C., Callaway, M., Munyoro, E. C., Centeno, C., Cleary, J., Connor, S., Davaasuren, O., Downing, J., Foley, K., Goh, C., … Pastrana, T. (2020). Redefining palliative care—A new consensus-based definition. *Journal of Pain and Symptom Management*, *60*(4), 754–764. https://doi.org/10.1016/j.jpainsymman.2020.04.027

Ramos, K., Hastings, S. N., Bosworth, H. B., & Fulton, J. J. (2018). Life program: Pilot testing a palliative psychology group intervention. *Journal of Palliative Medicine*, *21*(11), 1641–1645. https://doi.org/10.1089/jpm.2017.0454

Rost, A. D., Wilson, K., Buchanan, E., Hildebrandt, M. J., & Mutch, D. (2012). Improving psychological adjustment among late-stage ovarian cancer patients: Examining the role of avoidance in treatment. *Cognitive and Behavioral Practice*, *19*(4), 508–517. https://doi.org/10.1016/j.cbpra.2012.01.003

Sauer, C., Haussmann, A., & Weissflog, G. (2024). The effects of acceptance and commitment therapy (ACT) on psychological and physical outcomes among cancer patients and survivors: An umbrella review. *Journal of Contextual Behavioral Science*, *33*, Article 100810. https://doi.org/10.1016/j.jcbs.2024.100810

Schiele, M. A., Gottschalk, M. G., & Domschke, K. (2020). The applied implications of epigenetics in anxiety, affective and stress-related disorders - A review and synthesis on psychosocial stress, psychotherapy and prevention. *Clinical Psychology Review*, *77*, Article 101830. https://doi.org/10.1016/j.cpr.2020.101830

Secinti, E., Tometich, D. B., Johns, S. A., & Mosher, C. E. (2019). The relationship between acceptance of cancer and distress: A meta-analytic review. *Clinical Psychology Review*, *71*, 27–38. https://doi.org/10.1016/j.cpr.2019.05.001

Serfaty, M., Armstrong, M., Vickerstaff, V., Davis, S., Gola, A., McNamee, P., Omar, R. Z., King, M., Tookman, A., Jones, L., & Low, J. T. S. (2019). Acceptance and commitment therapy for adults with advanced cancer (CanACT): A feasibility randomised controlled trial. *Psycho-Oncology*, *28*(3), 488–496. https://doi.org/10.1002/pon.4960

Sterne, J. A. C., Hernán, M. A., Reeves, B. C., Savović, J., Berkman, N. D., Viswanathan, M., Henry, D., Altman, D. G., Ansari, M. T., Boutron, I., Carpenter, J. R., Chan, A. W., Churchill, R., Deeks, J. J., Hróbjartsson, A., Kirkham, J., Jüni, P., Loke, Y. K., Pigott, T. D., … Higgins, J. P. (2016). ROBINS-I: A tool for assessing risk of bias in non-randomised studies of interventions. *BMJ*, *355,* i4919. https://doi.org/10.1136/bmj.i4919

Sterne, J., Savović, J., Page, M. J., Elbers, R. G., Blencowe, N. S., Boutron, I., Cates, C. J., Cheng, H.-Y., Corbett, M. S., Eldridge, S. M., Emberson, J. R., Hernán, M. A., Hopewell, S., Hróbjartsson, A., Junqueira, D. R., Jüni, P., Kirkham, J. J., Lasserson, T., Li, T., … Higgins, J. P. T. (2019). RoB 2: A revised tool for assessing risk of bias in randomised trials. *BMJ*, *366*, Article l4898. https://doi.org/10.1136/bmj.l4898

Surges, S. M., Brunsch, H., Jaspers, B., Apostolidis, K., Cardone, A., Centeno, C., Cherny, N., Csikós, À., Fainsinger, R., Garralda, E., Ling,

J., Menten, J., Mercadante, S., Mosoiu, D., Payne, S., Preston, N., Van den Block, L., Hasselaar, J., & Radbruch, L. (2024). Revised European Association for Palliative Care (EAPC) recommended framework on palliative sedation: An international Delphi study. *Palliative Medicine*, *38*(2), 213–228. https://doi.org/10.1177/02692163231220225

Tarbi, E. C., & Meghani, S. H. (2019). A concept analysis of the existential experience of adults with advanced cancer. *Nursing Outlook*, *67*(5), 540–557. https://doi.org/10.1016/j.outlook.2019.03.006

Teoli, D., Schoo, C., & Kalish, V. B. (2025). Palliative care. In *StatPearls*. StatPearls Publishing.

Thompson, S. G., & Higgins, J. P. T. (2002). How should meta-regression analyses be undertaken and interpreted? *Statistics in Medicine*, *21*(11), 1559–1573. https://doi.org/10.1002/sim.1187

Wegner, D. M., Schneider, D. J., Carter, S. R., & White, T. L. (1987). Paradoxical effects of thought suppression. *Journal of Personality and Social Psychology*, *53*(1), 5–13. https://doi.org/10.1037/0022-3514.53.1.5

Weisman, A. D., & Worden, J. W. (1977). The existential plight in cancer: Significance of the first 100 days. *The International Journal of Psychiatry in Medicine*, *7*(1), 1–15. https://doi.org/10.2190/UQ2G-UGV1-3PPC-6387

Wersebe, H., Lieb, R., Meyer, A. H., Hoyer, J., Wittchen, H.-U., & Gloster, A. T. (2017). Changes of valued behaviors and functioning during an acceptance and commitment therapy intervention. *Journal of Contextual Behavioral Science*, *6*(1), 63–70. https://doi.org/10.1016/j.jcbs.2016.11.005

Wilaras, A. (2024). Cognitive behavioural interventions for palliative care in adult patients: A scoping review. *Progress in Palliative Care*, *32*(6), 395–437. https://doi.org/10.1080/09699260.2024.2408127

World Health Organization. (2020). *Palliative care [Fact sheet]*. https://www.Who.Int/News-Room/Fact-Sheets/Detail/Palliative-Care

World Health Organization. (2022). *World mental health report: Transforming mental health for all*. World Health Organization.

Xia, W., Zheng, Y., Guo, D., Zhu, Y., & Tian, L. (2024). Effects of cognitive behavioral therapy on anxiety and depressive symptoms in advanced cancer patients: A meta-analysis. *General Hospital Psychiatry*, *87*, 20–32. https://doi.org/10.1016/j.genhosppsych.2024.01.006

Yalom, I. D. (2008). *Staring at the sun: Overcoming the terror of death*. Jossey-Bass.

Yang, Y., Yi, Y., Shi, X., & Yang, X. (2024). Comparative efficacy of psychological interventions on anxiety and depression in patients with cancer: A systematic review and network meta-analysis. *Medicine (Baltimore)*, *103*(21), Article e38155. https://doi.org/10.1097/MD.0000000000038155

Yusufov, M., McHugh, R. K., Greer, J. A., Dalrymple, K. L., Sannes, T. S., Braun, I. M., Tulsky, J. A., & Pirl, W. F. (2023). An acceptance and commitment therapy-based intervention for opioid use disorder risk in individuals with cancer: A treatment development study. *Journal of Contextual Behavioral Science*, *28*, 226–234. https://doi.org/10.1016/j.jcbs.2023.04.006

Zhao, C., Lai, L., Zhang, L., Cai, Z., Ren, Z., Shi, C., Luo, W., & Yan, Y. (2021). The effects of acceptance and commitment therapy on the psychological and physical outcomes among cancer patients: A meta-analysis with trial sequential analysis. *Journal of Psychosomatic Research*, *140*, Article 110304. https://doi.org/10.1016/j.jpsychores.2020.110304

GRUPO EDITORIAL DE PSICOFUNDACIÓN

Psicothema (2026) 38(3) 138-148

Psicothema

https://www.psicothema.com/es • eISSN 1886-144X

Colegio Oficial de Psicología del Principado de Asturias

Psicothema

Systematic Review

Effects of High-Intensity Interval Training on Mental Health Outcomes in Adolescents and Young Adults: A Systematic Review and Meta-Analysis

Miguel Heres[1], Susana Pulgar[1], Marcos Quintana-Cepedal[2], Maria Fernandez-del-Valle[2] and Hugo Olmedillas[2]

[1] European University of Atlantic (Spain)
[2] University of Oviedo (Spain)

ARTICLE INFO

Received: 29/10/2025
Accepted: 19/01/2026

Keywords:
Exercise
High-intensity interval training
Mental health
Youth

ABSTRACT

Background: Given rising physical inactivity and mental health challenges, time-efficient interventions like High-Intensity Interval Training (HIIT) require rigorous evaluation. This systematic review and meta-analysis synthesized the evidence for HIIT's effects on mental health in adolescents and young adults. **Method:** Six electronic databases were systematically searched for Randomized Controlled Trials (RCTs) up to September 2025, comparing HIIT (defined as ≥85% HRmax) against a passive control group in non-clinical populations aged 11-29 years. **Results:** 11 RCTs comprising 1,401 participants were included. The random-effects meta-analysis showed a modest reduction in perceived stress associated with HIIT (SMD = −0.29, 95% CI −0.59 to 0.004, p = .053), although the effect did not reach conventional statistical significance and substantial uncertainty remained. No statistically significant effects were observed for depressive symptoms, anxiety, or psychological wellbeing. **Conclusions:** HIIT may be associated with modest reductions in perceived stress; however, the certainty of the evidence was low to very low across all outcomes. These findings should be interpreted with caution, and further high-quality research is needed to clarify the effects of HIIT on mental health outcomes.

Efectos del Entrenamiento Interválico de Alta Intensidad en la Salud Mental de Adolescentes y Adultos Jóvenes: una Revisión Sistemática y Metaanálisis

Palabras clave:
Ejercicio
Entrenamiento interválico de alta intensidad
Salud mental
Jóvenes

RESUMEN

Antecedentes: Ante la creciente inactividad física y los problemas de salud mental, las intervenciones eficientes en tiempo, como el entrenamiento interválico de alta intensidad (HIIT), requieren una evaluación rigurosa. Esta revisión sistemática y metaanálisis sintetizó la evidencia sobre los efectos del HIIT en la salud mental de adolescentes y adultos jóvenes. **Método:** Se realizaron búsquedas sistemáticas en seis bases de datos electrónicas hasta septiembre de 2025 para identificar ensayos controlados aleatorizados que compararon HIIT (≥85% de la frecuencia cardíaca máxima) con un grupo control pasivo en poblaciones no clínicas de entre 11 y 29 años. **Resultados:** Se incluyeron 11 estudios con 1,401 participantes. El metaanálisis con modelos de efectos aleatorios mostró una reducción modesta del estrés percibido a favor del HIIT (DME = −0.29; IC del 95%: −0.59 a 0.004; p = .053), sin alcanzar la significación estadística convencional y con elevada incertidumbre. No se observaron efectos significativos sobre síntomas depresivos, ansiedad ni bienestar psicológico. **Conclusiones:** HIIT podría asociarse con reducciones modestas del estrés percibido; sin embargo, la certeza de la evidencia fue baja a muy baja. Se requieren estudios de mayor calidad para clarificar sus efectos sobre la salud mental.

Cite as: Heres, M., Pulgar, S., Quintana-Cepedal, M., Fernandez-del-Valle, M., & Olmedillas, H. (2026). Effects of High-Intensity Interval Training on mental health outcomes in adolescents and young adults: A systematic review and meta-analysis. *Psicothema, 38*(3), 136-146. https://doi.org/10.70478/psicothema.2026.38.13
Corresponding author: Hugo Olmedillas, olmedillashugo@uniovi.es

Depression and anxiety are the most prevalent mental health challenges among youth, affecting up to 30% of adolescents and young adults (Silva et al., 2020). These developmental periods are marked by increased vulnerability to mental health difficulties, driven by academic stress, social pressures, and developmental challenges (Patton et al., 2016; Twenge et al., 2019). This burden has intensified following the COVID-19 pandemic, with documented increases in psychological distress (Madigan et al., 2023; Racine et al., 2021). While antidepressants are effective, concerns regarding side effects in minors (Cipriani et al., 2016) underscore the need for accessible non-pharmacological alternatives such as exercise (Howlett et al., 2021). Studies have reported an inverse relationship between physical activity levels and symptoms of anxiety and depression (Bélair et al., 2018; Carter et al., 2021), reinforcing the potential of exercise as a complementary intervention. However, adherence to global physical activity guidelines remains low (Guthold et al., 2020), notwithstanding the benefits of exercise, largely due to barriers such as lack of time (Hoare et al., 2017).

Addressing these issues is essential, as individuals with elevated anxiety sensitivity may benefit significantly from physical activity (Mason & Asmundson, 2018; Schmidt et al., 2010). Although moderate-intensity exercise is widely recommended for managing depression (Pilling et al., 2009) and anxiety (Paluska & Schwenk, 2000), evidence in youth remains less conclusive than in adults (Carter et al., 2021; Schuch et al., 2016). The variability in these findings suggests that specific exercise parameters may play a crucial role in therapeutic efficacy. In this regard, Noetel et al. (2024) suggest that exercise intensity is a key determinant of mental health outcomes, with vigorous-intensity exercise potentially producing larger and more clinically meaningful effects than moderate-intensity activity. High-Intensity Interval Training (HIIT) appears to engage key neurobiological mechanisms more robustly, contributing to improved stress regulation (Heijnen et al., 2016; Miller & Raison, 2016). These adaptations may be particularly relevant for youth, whose brains are highly responsive to neuroplasticity enhancements promoted by vigorous exercise (Lubans et al., 2016a). Unlike moderate continuous training, HIIT imposes a rapidly fluctuating physiological demand that may trigger distinct neurobiological adaptations. Specifically, vigorous exercise bouts have been linked to greater acute increases in brain-derived neurotrophic factor (BDNF), catecholamines, and cortisol, which collectively mediate stress regulation and hippocampal plasticity (Heijnen et al., 2016; Miller & Raison, 2016). Furthermore, HIIT has shown superior efficacy in improving cardiorespiratory fitness compared to moderate exercise (Costigan et al., 2015). Given that low cardiorespiratory fitness is a strong predictor of depressive symptoms and psychological distress (Ruggero et al., 2015), the robust physiological stimulus provided by HIIT may offer a potent pathway for mental health promotion.

Despite this theoretical promise, the specific application of HIIT for mental health in young populations requires clarification. While HIIT acts as a time-efficient strategy addressing common barriers (Guthold et al., 2020), existing research remains fragmented. Previous reviews have been limited by substantial heterogeneity regarding intervention protocols, ranging from cycle ergometers to bodyweight circuits, and the frequent mixing of clinical and non-clinical samples (Leahy et al., 2020; Rodriguez-Ayllon et al., 2019). Consequently, it remains unclear whether the benefits observed in clinical settings translate to the prevention of symptoms in healthy youth, or if the variability in protocols (e.g., work-rest ratios, duration) dilutes the potential efficacy. Based on evidence linking higher exercise intensity to greater benefits, it is expected that HIIT may be associated with improvements in mental health outcomes, although substantial heterogeneity across studies is acknowledged.

This systematic review aims to evaluate the effects of HIIT on mental health outcomes in adolescents and young adults from a preventive perspective. Distinct from clinical trials focusing on symptom remission, this study targets non-clinical populations to determine whether HIIT can serve as a proactive strategy to manage subclinical symptomatology and buffer against stress. By addressing methodological limitations of previous research, this review offers new insights into the potential of HIIT as a time-efficient intervention for youth mental health.

Method

This systematic review and meta-analysis were conducted in accordance with the recommendations of the Cochrane Handbook for Systematic Reviews of Interventions (Higgins et al., 2024) and was reported following the guidelines established by the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) (Page et al., 2021). The study protocol was pre-registered in the International Prospective Register of Systematic Reviews (PROSPERO) under the registration number CRD42024567984.

Eligibility Criteria

The formulation of the research question, the process of evidence searching, and the definition of eligibility criteria were based on the components of the patient, intervention, comparator, outcomes and type of study (PICOS strategy) (Methley et al., 2014; Schardt et al., 2007). The titles and abstracts of all identified studies were screened, followed by a full-text review of relevant articles to assess their eligibility. Eligibility was determined by two independent researchers (MH and HO) according to the following predefined criteria: (1/P) healthy adolescents and young adults with no physical disease nor severe mental illness or mental disorder and aged between 11 and 29 years. The selection of this age range was guided by a public health and prevention-oriented perspective. Inclusion was restricted to passive control groups to specifically isolate the effects of HIIT relative to no structured exercise; (2/I) studies that follow a HIIT protocol at or above 85% of maximum heart rate (HR_{max}), which corresponds to over 75% of maximum oxygen uptake (VO_2max) (American College of Sports Medicine, 2018), being eligible any form of exercise of any duration and frequency. The HIIT intervention group may not perform any other interventions; (3/C) HIIT intervention compared to a passive control group that does not perform any intervention; (4/O) measurement of any mental health outcome; (5/S) studies should be RCTs.

Search and Study Selection

EMBASE, PsycInfo, PubMed, Scopus, SportDiscus and Web of Science were searched by two authors (MH and HO) from inception up to September 2025 for RCTs investigating HIIT in healthy adolescents and young adults.

In addition to the database search, a manual search was conducted by reviewing the reference lists of recent meta-analyses and all included studies to identify potentially relevant articles not captured

by the initial search. Studies identified through this manual process were subjected to the same independent screening procedures and were required to meet the same predefined inclusion and exclusion criteria as studies retrieved from the database search.

MH and HO independently conducted the screening and selection stages of the studies. Any divergence between investigators was solved in a consensus meeting with a third investigator (MF). The search terms used are available on the Open Science Framework (OSF): https://osf.io/vrs4u/files/bhfu6

Data Extraction

Data coding from each selected study was conducted by HO and independently cross-checked for accuracy by MH. Data extracted included: first author's last name and publication year, sample characteristics (baseline physical activity levels, sample size, mean age and standard deviation (or range if not available) and gender proportion), mental health assessments (outcome, instrument), intervention details (duration, frequency, time and type). Outcome data were reported as mean and standard deviation.

Assessment of Risk of Bias

The risk of bias in the included studies was assessed using the Cochrane Risk of Bias Tool for Randomized Clinical Trials (RoB 2.0), across its five standard domains. Two researchers (MQ and HO) independently applied the assessment tool to each included study. The risk of bias for each outcome was categorized as "high," "low," or "some concerns." and visualized using the robvis tool (McGuinness & Higgins, 2021). Any divergence between investigators was solved in a consensus meeting with a third investigator (MH). Formal inter-rater agreement statistics (e.g., kappa coefficients) were not calculated.

Assessment of Evidence Certainty

The Grading of Recommendations Assessment Development and Evaluation (GRADE) was used to evaluate the certainty of the evidence. Two researchers (MQ and HO) independently rated each meta-analysis, following GRADE guidelines which initially address RCTs to be rated high quality, and could be downgraded to moderate, low or very low based on the following domains: risk of bias, imprecision of results and inconsistency (Guyatt et al., 2011; Higgins, 2003). Any divergence between investigators was solved in a consensus meeting with a third investigator (MH).

Data Analysis

All analyses were conducted using R (The R Foundation for Statistical Computing), employing the *meta* and *metafor* packages (Schwarzer, 2025; Viechtbauer, 2010). Meta-analyses were performed when at least three studies were available for a given outcome (perceived stress, depressive symptoms, anxiety, wellbeing, and positive and negative affect).

Effect sizes for between-group comparisons accounting for change over time were calculated as standardized mean changes between groups using the escalc function in *metafor*, following the procedures described by Skvarc and Fuller-Tyszkiewicz (2024). This approach is consistent with the standardized mean change originally proposed by Becker (1988). When pre–post correlation coefficients were not reported, a conservative correlation of $r = 0.5$ was assumed for variance estimation, in accordance with recommendations from the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins et al., 2024).

Effect sizes were expressed as Cohen's *d*. To further improve interpretability, pooled effect sizes were additionally transformed into the percentage of superiority, representing the probability that a randomly selected participant from the intervention group shows a better outcome than a randomly selected participant from the control group. Effect size magnitude was interpreted as a continuous measure, without relying on conventional categorical benchmarks.

Pooled estimates were obtained using random-effects models with inverse-variance weighting, reflecting expected heterogeneity in exercise protocols, follow-up durations, and study populations across trials. Statistical heterogeneity was assessed using the I^2 statistic, with values above 50% and 75% indicating moderate and high heterogeneity, respectively.

Meta-analytic results were presented using forest plots. No moderator analyses were pre-specified in the PROSPERO protocol, as fewer than ten effect sizes were anticipated per outcome. Similarly, formal assessments of small-study effects or publication bias were not conducted due to the limited number of available studies. Statistical significance was set at $p < .05$.

Assessment of Reporting Biases

Assessment of publication bias using funnel plot inspection and Egger's regression test was planned for any meta-analysis that included at least 10 studies. However, as no meta-analysis met this minimum number of studies, these analyses could not be formally conducted. According to the Cochrane Handbook for Systematic Reviews of Interventions, tests for funnel plot asymmetry are underpowered and unreliable with fewer than 10 studies, making their results potentially misleading (Higgins et al., 2024).

Results

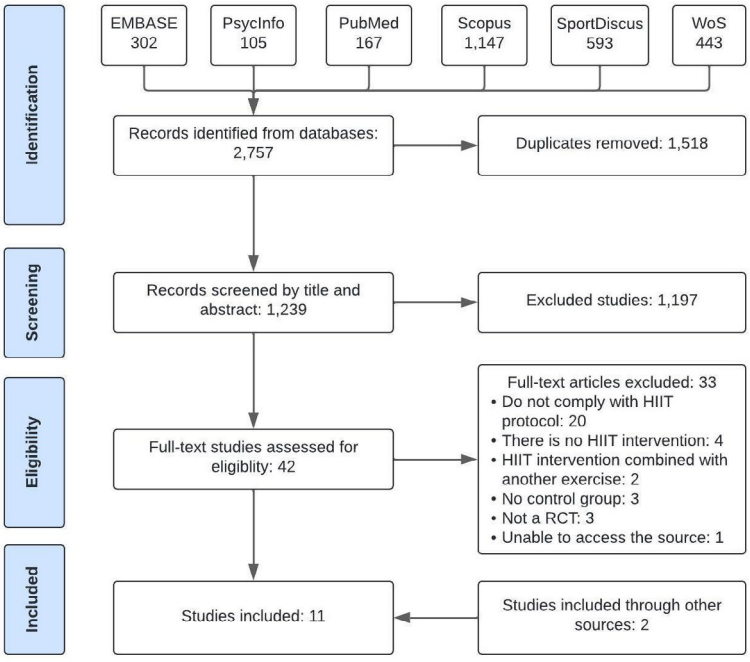
Search Results

The literature searches identified a total of 2,757 studies. Once 1,518 duplicates were removed, 1,239 titles and abstracts were examined. At full text review stage, 42 studies were reviewed to evaluate eligibility, of which nine met the inclusion criteria. Two studies were added from other sources for the final 11 studies included in this systematic review (Figure 1).

Study Characteristics

A total of 1,401 participants who completed treatments were included in the review: 709 in HIIT intervention groups and 692 in control groups. The duration of the interventions ranged from 2 to 16 weeks, except for one study that lasted 12 months (Lubans et al., 2020). Eight studies performed 3 sessions per week (Costigan et al., 2016; Eather et al., 2019; Leahy et al., 2018; Liu et al., 2022; Lubans et al., 2020; Lucibello et al., 2020; May et al., 2018; Shandu et al., 2023), one study 2 sessions per week (Harris et al., 2022), one study 1 session per week (Allen et al., 2018) and one study 10 sessions in 11 days (Ormsbee et al., 2013). Five studies used progressively longer sessions (Costigan et al., 2016; Eather et al., 2019; Leahy et

Figure 1
Prisma Flow Diagram



Note: Flowchart illustrating the study selection process, including number of records identified, screened, assessed for eligibility and finally included in the systematic review.

al., 2018; Lubans et al., 2020; Ormsbee et al., 2013), with session durations ranging from 8 to 40 minutes. One study included only male participants (Shandu et al., 2023); the remaining ten studies have female representation ranging from 31% to 82%, with a mean of 59%. The mean age of the participants was not reported in one study (Shandu et al., 2023), while in the remaining studies, ranged from 12 to 31 years. Mental health outcomes assessed across studies included salivary and hair cortisol, depressive symptoms, perceived stress, positive and negative affect, wellbeing and anxiety. Detailed study characteristics are presented in Table 1.

Risk of Bias

Risk of bias was independently assessed by two reviewers (MQ and HO) using the Cochrane Risk of Bias Tool (RoB 2.0). Based on this assessment, six studies were rated as low risk of bias, two were rated as having some concerns and three were rated as high risk of bias (Figure 2).

Meta-Analysis

Forest plots for all mental health outcomes derived from the meta-analyses are presented in Figure 3.

Figure 2
Risk of Bias Using Cochrane Risk of Bias 2.0 Tool

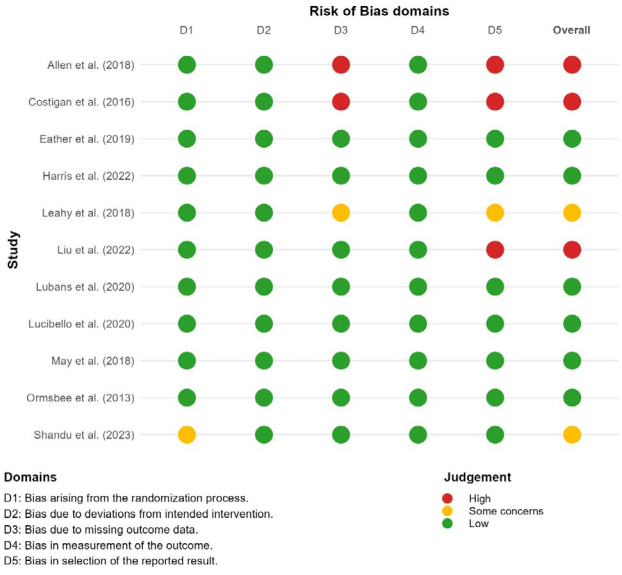


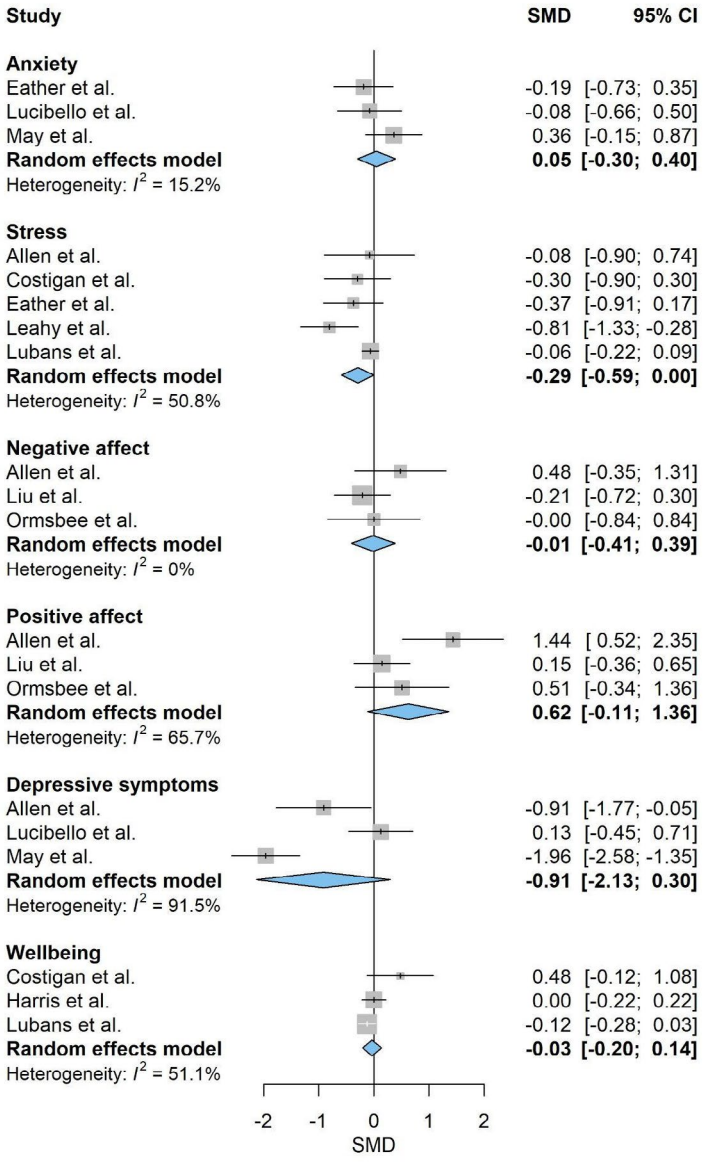
Table 1
Characteristics of the Studies Included in the Study.

Study	Participants	Measures	HIIT intervention	Results
Allen et al. (2018)	PA levels: moderate Smokers IG: *n* = 12, 30,4 ± 0,5 years, 33% women; CG: *n* = 11, 31 ± 2 years, 82% women	CES-D - baseline, week 4, week 8 and post intervention PANAS - baseline, week 4, week 8 and post intervention PSS - baseline, week 4, week 8 and post intervention	Stationary bike 12 weeks 1 session/week 20 minutes/session	Significant improvements at week 8 ($p < .015$) and week 12 ($p < .019$) in positive affect between groups Rest of mental health outcomes: no significant changes
Costigan et al. (2016)	PA levels: NR England high school students (15,8 ± 0,6 years) 31% women IG1: *n* = 21, IG2: *n* = 22; CG: *n* = 21	K10 - baseline and post intervention Flourishing Scale - baseline and post intervention Feelings State - pre and post every session	IG1: gross motor cardiorespiratory exercises; IG2: combination of cardiorespiratory and body weight resistance training exercises (PE classes) 8 weeks 3 session/week Progressive (from 8 to 10 min/ session)	Feelings State: significant differences ($p < .001$) between pre and post for IG1; improvements for IG2 ($p < .06$) Well-being (Flourishing Scale): small intervention effect for both IG1 and IG2 K10: No intervention effect
Eather et al. (2019)	PA levels: NR Australia university students (20,38 ± 1,88 years) IG: *n* = 22, 64% women; CG: *n* = 31, 68% women	STAI - baseline and post intervention PSS - baseline and post intervention Feelings State - pre and post every session	Combination of aerobic and core resistance using body weight or basic equipment 8 weeks 3 session/week Progressive (from 8 to 10 min/ session)	No significant changes
Harris et al. (2022)	PA levels: NR New Zealand students IG: *n* = 160, 12,1 ± 0,6 years, 55% women; CG: *n* = 163, 12,2 ± 0,6 years, 45% women	SDQ - baseline and post intervention Warwick-Edinburgh Mental Wellbeing Scale - baseline and post intervention	Major muscle groups (PE classes) 16 weeks (2 of rest) 2 session/week 15 minutes/session	No intervention effect
Leahy et al. (2018)	PA levels: NR Australia high school students (16,2 ± 0,4 years), 54% women IG: *n* = 38; CG: *n* = 30	SDQ - baseline and post intervention PSS - baseline and post intervention	Aerobic-based and resistance-based exercises 14 weeks 3 session/week From 12 to 20 minutes/session	Moderate improvement for IG for total difficulties ($p = .02$) and reduction in subscales emotional problems ($p = .02$) and peer problems ($p = .02$)
Liu et al. (2022)	PA levels: low University students (26,26 ± 2,21 years) IG: *n* = 33, 56% women; CG: *n* = 28, 57% women	SPANE - baseline and post intervention	Treadmill 12 weeks 3 session/week 40 minutes/session	Small differences in favour of the IG: trends upward for satisfaction, positive affect and well-being; trend downward for negative affect
Lubans et al. (2020)	PA levels: NR Australia high school students IG: *n* = 337, 16 ± 0,4 years, 50% women; CG: *n* = 333, 16 ± 0,5 years, 39% women	PSS - baseline, middle and post intervention SDQ - baseline, middle and post intervention Warwick-Edinburgh Mental Wellbeing Scale - baseline, middle and post intervention Hair cortisol (baseline-6 months) - baseline and middle intervention	Aerobic and body weight resistance exercises 12 months 3 session/week From 8 to 20 minutes/session	Hair cortisol: small difference in favour of the IG; moderation effects for mental health status: stronger effects on poor mental health at baseline Other mental health outcomes: no differences between groups
Lucibello et al. (2020)	PA levels: low University students (19,8 ± 2,2 years) 63% women IG: *n* = 25; CG: *n* = 21	BAI - baseline and post intervention BDI-II - baseline and post intervention	Cycle ergometer 9 weeks 3 session/week 20 minutes/session	No intervention effect
May et al. (2018)	PA levels: low USA university students (18,55 ± 0,99 years) 82% women IG: *n* = 30; CG: *n* = 30	SBI - baseline and post intervention CES-D - baseline and post intervention STAI - baseline and post intervention	Cyclo ergometer 4 weeks 3 session/week 20 minutes/session	No intervention effect

Study	Participants	Measures	HIIT intervention	Results
Ormsbee et al. (2013)	PA levels: moderate-high Active university students IG: *n* = 12, 19,3 ± 0,4 years, 58% women; CG: *n* = 10, 19,5 ± 0,4, 80% women	PANAS - baseline and post intervention Salivary cortisol - baseline, pre and post intervention	Cycle ergometer 10 sessions (within 2 weeks); 1 session/day Progressive (from 7,5 to 20 min/ session)	No significant changes
Shandu et al. (2023)	PA levels: low Male smoker students (18-30 years) IG: *n* = 18; CG: *n* = 14	WHOQOL-BREF - baseline and post intervention	Cycle ergometer 8 weeks 3 session/week 33 minutes/session	Significant improvements ($p < .05$) for IG

Note: BAI, Beck Anxiety Inventory; BDI-II, Beck Depression Inventory II; CES-D, Center for Epidemiological Studies - Depression; CG, Control Group; IG, Intervention Group; K10, Kessler Psychological Distress Scale; NR, Not Reported; PANAS, Positive And Negative Affect Scale; PA, Physical Activity; PSS, Perceived Stress Scale; SBI, School Burnout Inventory; SDQ, Strength and Difficulties Questionnaire; SPANE, Scale of Positive And Negative Experience; STAI, State-Trait Anxiety Scale; WHOQOL-BREF, World Health Organization Quality of Life.

Figure 3
Forest Plot



Note: CI, Confidence Intervals; SMD, Standardized Mean Difference.

Depressive Symptoms

The quantitative synthesis showed a non-significant effect of HIIT on depressive symptoms compared to no intervention (SMD = −0.91, 95% CI −2.13 to 0.30, p = .14), with substantial heterogeneity (I^2 = 91.5%; quality of evidence: very low). The pooled effect corresponded to a percentage of superiority of 74%. The prediction interval ranged from -6.03 to 4.20.

Anxiety

There was no statistically significant effect in favour of the control group for anxiety compared to the HIIT group (SMD = 0.05, 95% CI −0.30 to 0.40, p = .79), with low heterogeneity (I^2 = 15.2%; quality of evidence: very low). The pooled effect corresponded to a percentage of superiority of 51.4%. The prediction interval ranged from −0.91 to 1.00.

Wellbeing

There was no significant effect of HIIT on psychological wellbeing compared to no intervention (SMD = −0.03, 95% CI −0.20 to 0.14, p = .71), with moderate heterogeneity (I^2 = 51.1%; quality of evidence: low). The pooled effect corresponded to a percentage of superiority of 49.1%. The prediction interval ranged from -0.56 to 0.50.

Stress

The inclusion of HIIT protocols was associated with a modest reduction in perceived stress compared to no intervention; however, this effect did not reach conventional statistical significance (SMD = −0.29, 95% CI −0.59 to 0.005, p = .053), with moderate heterogeneity (I^2 = 50.8%; quality of evidence: very low). The pooled effect corresponded to a percentage of superiority of 58.2%. The prediction interval ranged from −1.06 to 0.47, indicating substantial uncertainty in the expected effects across future studies.

Positive and Negative Affect

There was no statistically significant effect of HIIT on positive affect compared to no intervention (SMD = 0.63, 95% CI −0.11 to 1.36, p = .09), with substantial heterogeneity (I^2 = 65.7%; quality of evidence: very low). The pooled effect corresponded to a percentage of superiority of 67.2%. The prediction interval ranged from −2.17 to 3.42. Conversely, there was no statistically significant effect of HIIT on negative affect compared to the control group (SMD = −0.01, 95% CI −0.41 to 0.39, p = .96), with no observed heterogeneity (I^2 = 0.0%; quality of evidence: very low). The pooled effect corresponded to a percentage of superiority of 49.7%. The prediction interval ranged from −0.94 to 0.92.

Discussion

This systematic review and meta-analysis examined the effects of HIIT on mental health outcomes in adolescents and young adults. The findings suggest a potential modest reduction in perceived stress following HIIT. Although this effect did not reach conventional statistical significance and the certainty of evidence was low, the direction of the effect is noteworthy. Regarding depressive symptoms and positive affect, while non-significant effects were observed, the estimates were generally directionally favourable to HIIT. No significant effects were found for anxiety, psychological well-being or negative affect, indicating variability in the impact of HIIT across different mental health domains. The certainty of the evidence for all analysed outcomes ranged from low to very low, underscoring the need for more rigorous, high-quality studies in this area.

To aid interpretation of the pooled effects, SMDs were complemented with prediction intervals and the percentage of superiority. Prediction intervals were wide for most outcomes and frequently crossed the null value, highlighting the substantial variability in expected effects across future studies. The percentage of superiority provided a clinically interpretable estimate of the probability that an individual receiving HIIT would show better outcomes than one in the control group; however, these values should be interpreted cautiously given the low certainty of evidence and the wide prediction intervals observed.

We did not observe a statistically significant effect of HIIT on depressive symptoms, although the estimated effects tended to favour HIIT. This pattern is broadly consistent with previous research suggesting that exercise, including HIIT, may influence mood and emotional regulation. Notably, some interventions included participants with baseline symptoms of depression or anxiety, and HIIT may have yielded greater mental health benefits in those individuals with poorer mental health status (Lubans et al., 2020). However, the interpretation of these findings is limited by the small number and low quality of the included RCTs (Schuch et al., 2016; Stubbs et al., 2017).

Consistent with these findings, the pooled analysis of depressive symptoms was marked by substantial heterogeneity (I^2 = 91.5%). This high level of statistical inconsistency is likely a direct reflection of the considerable clinical and methodological diversity across the included trials, as detailed in Table 1. For instance, intervention durations varied dramatically, from a short-term 2-week protocol by Ormsbee et al. (2013) to a 12-month intervention by Lubans et al. (2020). Similarly, the HIIT dose was inconsistent, with frequencies ranging from a single weekly session (Allen et al., 2018) to 3 sessions per week in the majority of studies. Furthermore, the nature of the HIIT protocols themselves was diverse, including cycle ergometer-based sessions (e.g., Lucibello et al., 2020) and bodyweight-based resistance exercises (e.g., Costigan et al., 2016). Such profound variations in intervention design, duration, and modality prevent a clear conclusion and underscore the difficulty in determining an optimal HIIT prescription for mental health outcomes based on the current evidence.

Perceived stress was the only outcome showing a consistent tendency towards improvement, representing a medium effect size with potential clinical relevance as a complementary management strategy. These results align with previous findings (Borrega-Mouquinho et al., 2021); however, the high heterogeneity suggests that physiological adaptations alone (e.g., anti-inflammatory pathways) do not guarantee benefits. Crucially, recent evidence indicates that individual traits, such as emotional impulsivity, may blunt the psychoneuroendocrine response to stress despite fitness gains (Javelle et al., 2024). Consequently, while HIIT shows promise, it should be prescribed cautiously, accounting for baseline psychological characteristics rather than as a uniform solution.

The lack of significant effects on anxiety and psychological well-being observed in this review may be partially explained by the short duration of many interventions and their emphasis on more transient affective states (Biddle et al., 2019; Harris et al., 2022; Schuch et al., 2016). Longer intervention periods may be necessary to detect measurable changes in these outcomes, particularly in individuals with more severe baseline symptoms (Eather et al., 2019). However, the current evidence on HIIT remains inconclusive, as studies vary widely in their methods, including differences in intervention duration, intensity, and outcome assessment (Hood et al., 2011; Little et al., 2010). Furthermore, some studies suggest that factors such as intensity, duration, and frequency of HIIT sessions may not significantly influence mental health outcomes (Gaia et al., 2024; Lucibello et al., 2020).

The low heterogeneity observed for anxiety (I^2 = 15.2%) and the moderate heterogeneity for psychological well-being (I^2 = 51.1%) indicate relatively consistent findings across studies, particularly for anxiety. However, the absence of statistically significant effects and prediction intervals centred around the null value suggest that HIIT may have limited or variable influence on anxiety and psychological well-being across different populations and intervention contexts.

It is possible that acute features of HIIT, such as high exercise intensity, may temporarily divert attention from negative thoughts or stressors, while the structured and goal-oriented nature of HIIT could contribute to feelings of accomplishment or self-efficacy. Nevertheless, these potential psychological responses may be transient and insufficient to produce sustained changes in more stable mental health constructs such as anxiety or psychological well-being.

No statistically significant effects were observed for either positive or negative affect following HIIT interventions. For positive affect, the pooled effect was non-significant and characterized by substantial heterogeneity (I^2 = 65.7%), suggesting considerable variability in responses across studies. In contrast, findings for negative affect were highly consistent (I^2 = 0%), with effect estimates centred around the null value, indicating no systematic effect of HIIT on this outcome.

The lack of significant effects on positive affect and the high heterogeneity (I^2 = 65.7%) align with the notion that emotional responses to HIIT are highly individualized. Lubans et al. (2016a; 2016b) suggest that these responses depend on baseline fitness and psychological mediators like self-esteem. Consequently, individuals with lower fitness may not experience the immediate physiological adaptations required to elicit emotional improvements, reinforcing the need to account for personal characteristics when prescribing HIIT.

Despite mixed findings, HIIT may be considered a time-efficient exercise modality with potential relevance for mental health outcomes, although current evidence remains limited. School-based programs show potential given the rising stress trends among youth (Patalay & Gage, 2019). However, outcomes vary considerably, possibly due to short intervention durations and implementation challenges common in educational settings (Naylor et al., 2015). Evidence suggests that interventions exceeding three months may be necessary to yield consistent mental health improvements (Brown et al., 2013), contrasting with the shorter protocols typical of experimental HIIT research.

From a clinical perspective, the observed reduction in perceived stress represents a medium effect size. While this suggests that HIIT can be a valuable component of stress management strategies for young people, the substantial heterogeneity indicates that individual responses vary significantly. Consequently, HIIT should be viewed as a complementary tool rather than a standalone first-line treatment for mental health difficulties in this population.

Despite the directionally favourable estimates of some findings, several methodological limitations of the included studies contribute to the low to very low certainty of the evidence. Due to the small number of studies included in each meta-analysis ($N < 10$), standard methods to assess small-study effects, such as funnel plots and Egger's test, lacked sufficient statistical power and were therefore not performed. Furthermore, sensitivity analyses excluding studies rated as high risk of bias were not performed due to the limited number of studies, preventing a formal assessment of how study quality influenced the pooled estimates. As a result, the observed reduction in perceived stress, characterized by marginal statistical significance and wide confidence and prediction intervals, may represent an imprecise estimate of the true effect and should be interpreted with caution until supported by a larger and more robust body of evidence. Second, substantial variability in study designs, participant characteristics, and intervention protocols limits the generalizability of the findings. We acknowledge that the broad age range (11–29 years) encompasses distinct developmental stages, but, age-related subgroup analyses were not feasible due to the limited number of studies. Differences in session duration, intensity prescription, frequency, and adherence rates likely contributed to the observed heterogeneity across mental health outcomes. Moreover, the predominant reliance on self-reported measures introduces the potential for reporting bias, influenced by factors such as social desirability and transient mood states. In addition, the use of diverse validated instruments raises concerns regarding measurement equivalence. Differences in developmental stage, symptom expression, and interpretation of self-report measures across the broad age range may limit comparability and contribute to heterogeneity.

It is important to note that, as this review focused on non-clinical populations, the initial burden of depressive and anxiety symptoms was generally low. This likely created a 'floor effect,' making it statistically difficult to detect large magnitude reductions in these specific outcomes compared to therapeutic interventions in clinical samples. Some studies also assessed physiological stress markers; however, these could not be included in the meta-analyses due to the limited number of available studies and were therefore considered only in the narrative synthesis, further highlighting challenges in outcome comparability. Additionally, the restriction of our search to peer-reviewed literature constitutes a limitation that may have introduced publication bias. However, this decision was made to prioritize the inclusion of studies with a verified standard of methodological reporting, given the high variability in protocols typically found in non-peer-reviewed sources. Finally, many interventions were of relatively short duration, which may be insufficient to detect changes in more stable mental health constructs such as anxiety or depressive symptoms. Future research should prioritize larger, well-powered randomized controlled trials with standardized HIIT protocols, comprehensive reporting of participant characteristics, longer intervention and follow-up periods, and the inclusion of age-appropriate subjective and objective mental health outcomes to strengthen the quality and interpretability of the evidence.

In conclusion, HIIT may be associated with modest reductions in perceived stress, although substantial uncertainty remains.

No statistically significant effects were observed for depressive symptoms, anxiety, or psychological wellbeing. Overall, the certainty of the evidence ranged from low to very low, and the findings should be interpreted with caution.

Author Contributions

Miguel Heres: Writing – Review & Editing, Writing – Original Draft, Validation, Methodology, Investigation, Formal Analysis, Data Curation, Conceptualization. **Susana Pulgar:** Writing – Review & Editing, Validation, Methodology. **Marcos Quintana-Cepedal:** Writing – Review & Editing, Validation, Methodology, Conceptualization, Formal Analysis. **Maria Fernandez-del-Valle:** Writing – Review & Editing, Supervision, Conceptualization. **Hugo Olmedillas:** Writing – Review & Editing, Writing – Original Draft, Validation, Methodology, Investigation, Formal Analysis, Data Curation, Conceptualization, Supervision.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Conflict of Interest Statement

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data Availability Statement

All data generated or analyzed during this study are available from the corresponding author upon request.

References

Allen, A., Carlson, S. C., Bosch, T. A., Eberly, L. E., Okuyemi, K., Nair, U., & Gordon, J. S. (2018). High-intensity interval training and continuous aerobic exercise interventions to promote self-initiated quit attempts in young adults who smoke: Feasibility, acceptability, and lessons learned from a randomized pilot trial. *Journal of Addiction Medicine, 12*(5), 373–380. https://doi.org/10.1097/ADM.0000000000000414

American College of Sports Medicine. (2018). *ACSM's guidelines for exercise testing and prescription* (10th ed.). Wolters Kluwer.

Becker, B. J. (1988). Synthesizing standardized mean-change measures. *British Journal of Mathematical and Statistical Psychology, 41*(2), 257–278. https://doi.org/10.1111/j.2044-8317.1988.tb00901.x

Bélair, M. A., Kohen, D. E., Kingsbury, M., & Colman, I. (2018). Relationship between leisure time physical activity, sedentary behaviour and symptoms of depression and anxiety: Evidence from a population-based sample of Canadian adolescents. *BMJ Open, 8*(10), Article e021119. https://doi.org/10.1136/bmjopen-2017-021119

Biddle, S. J. H., Ciaccioni, S., Thomas, G., & Vergeer, I. (2019). Physical activity and mental health in children and adolescents: An updated review of reviews and an analysis of causality. *Psychology of Sport and Exercise, 42*, 146–155. https://doi.org/10.1016/j.psychsport.2018.08.011

Borrega-Mouquinho, Y., Sánchez-Gómez, J., Fuentes-García, J. P., Collado-Mateo, D., & Villafaina, S. (2021). Effects of high-intensity interval training and moderate-intensity training on stress, depression, anxiety, and resilience in healthy adults during coronavirus disease 2019 confinement: A randomized controlled trial. *Frontiers in Psychology, 12*, Article 643069. https://doi.org/10.3389/fpsyg.2021.643069

Brown, H. E., Pearson, N., Braithwaite, R. E., Brown, W. J., & Biddle, S. J. (2013). Physical activity interventions and depression in children and adolescents: A systematic review and meta-analysis. *Sports Medicine, 43*(3), 195–206. https://doi.org/10.1007/s40279-012-0015-8

Carter, T., Pascoe, M., Bastounis, A., Morres, I. D., Callaghan, P., & Parker, A. G. (2021). The effect of physical activity on anxiety in children and young people: A systematic review and meta-analysis. *Journal of Affective Disorders, 285*, 10–21. https://doi.org/10.1016/j.jad.2021.02.026

Cipriani, A., Zhou, X., Del Giovane, C., Hetrick, S. E., Qin, B., Whittington, C., Furukawa, T. A., Salanti, G., & Geddes, J. R. (2016). Comparative efficacy and tolerability of antidepressants for major depressive disorder in children and adolescents: A network meta-analysis. *The Lancet, 388*, 881–890. https://doi.org/10.1016/S0140-6736(16)30385-3

Costigan, S. A., Eather, N., Plotnikoff, R. C., Hillman, C. H., & Lubans, D. R. (2016). High-intensity interval training for cognitive and mental health in adolescents. *Medicine & Science in Sports & Exercise, 48*(10), 1985–1993. https://doi.org/10.1249/MSS.0000000000000993

Costigan, S. A., Eather, N., Plotnikoff, R. C., Taaffe, D. R., & Lubans, D. R. (2015). High-intensity interval training for improving health-related fitness in adolescents: A systematic review and meta-analysis. *British Journal of Sports Medicine, 49*(19), 1253–1261. https://doi.org/10.1136/bjsports-2014-094490

Eather, N., Riley, N., Miller, A., Smith, V., Poole, A., Vincze, L., Morgan, P. J., & Lubans, D. R. (2019). Efficacy and feasibility of HIIT training for university students: The Uni-HIIT RCT. *Journal of Science and Medicine in Sport, 22*(5), 596–601. https://doi.org/10.1016/j.jsams.2018.11.016

Gaia, J. W. P., Schuch, F. B., Ferreira, R. W., Souza, E. L., Ferreira, V. M. S., & Pires, D. A. (2024). Effects of high-intensity interval training on depressive and anxiety symptoms in healthy individuals: A systematic review and meta-analysis of randomized clinical trials. *Scandinavian Journal of Medicine & Science in Sports, 34*(4), Article e14618. https://doi.org/10.1111/sms.14618

Guthold, R., Stevens, G. A., Riley, L. M., & Bull, F. C. (2020). Global trends in insufficient physical activity among adolescents: A pooled analysis of 298 population-based surveys with 1.6 million participants. *The Lancet Child & Adolescent Health, 4*(1), 23–35. https://doi.org/10.1016/S2352-4642(19)30323-2

Guyatt, G. H., Oxman, A. D., Sultan, S., Glasziou, P., Akl, E. A., Alonso-Coello, P., Atallah, Á., Brozek, J., Dahm, P., D'Amico, R., Furlan, A., Harmsen, R., Hempel, S., Helfand, M., Hróbjartsson, A., Jüni, P., Liberati, A., Marshall, J., … Schünemann, P. (2011). GRADE guidelines: 9. Rating up the quality of evidence. *Journal of Clinical Epidemiology, 64*(12), 1311–1316. https://doi.org/10.1016/j.jclinepi.2011.06.004

Harris, N., Warbrick, I., Fleming, T., Borotkanics, R., Atkins, D., & Lubans, D. (2022). Impact of high-intensity interval training including Indigenous narratives on adolescents' mental health: A cluster-randomised controlled trial. *Australian and New Zealand Journal of Public Health, 46*(6), 794–799. https://doi.org/10.1111/1753-6405.13312

Heijnen, S., Hommel, B., Kibele, A., & Colzato, L. S. (2016). Neuromodulation of aerobic exercise: A review. *Frontiers in Psychology, 6*, Article 1890. https://doi.org/10.3389/fpsyg.2015.01890

Higgins, J. P. T. (2003). Measuring inconsistency in meta-analyses. *BMJ, 327*(7414), 557–560. https://doi.org/10.1136/bmj.327.7414.557

Higgins, J. P. T., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M. J., & Welch, V. A. (Eds.). (2024). *Cochrane handbook for systematic reviews of interventions* (Version 6.5, updated August 2024). Cochrane Collaboration. https://training.cochrane.org/handbook

Hoare, E., Stavreski, B., Jennings, G. L., & Kingwell, B. A. (2017). Exploring motivation and barriers to physical activity among active and inactive Australian adults. *Sports, 5*(3), Article 47. https://doi.org/10.3390/sports5030047

Hood, M. S., Little, J. P., Tarnopolsky, M. A., Myslik, F., & Gibala, M. J. (2011). Low-volume interval training improves muscle oxidative capacity in sedentary adults. *Medicine & Science in Sports & Exercise, 43*(10), 1849–1856. https://doi.org/10.1249/MSS.0b013e3182199834

Howlett, N., Bottoms, L., Chater, A., Clark, A. B., Clarke, T., David, L., Irvine, K., Jones, A., Jones, J., Mengoni, S. E., Murdoch, J., Pond, M., Sharma, S., Sims, E. J., Turner, D. A., Wellsted, D., Wilson, J., Wyatt, S., & Trivedi, D. (2021). A randomised controlled trial of energetic activity for depression in young people (READY): A multi-site feasibility trial protocol. *Pilot and Feasibility Studies, 7*(1), Article 6. https://doi.org/10.1186/s40814-020-00734-7

Javelle, F., Bloch, W., Borges, U., Burberg, T., Collins, B., Gunasekara, N., Hosang, T., Jacobsen, T., Laborde, S., Löw, A., Schenk, A., Schlagheck, M. L., Schoser, D., Vogel, A., Walzik, D., & Zimmer, P. (2024). Eight weeks of high-intensity interval training versus stretching do not change the psychoneuroendocrine response to a social stress test in emotionally impulsive humans. *European Journal of Applied Physiology, 124*, 2893–2908. https://doi.org/10.1007/s00421-024-05471-w

Leahy, A. A., Eather, N., Smith, J. J., Hillman, C. H., Morgan, P. J., Plotnikoff, R. C., Nilsson, M., Costigan, S. A., Noetel, M., & Lubans, D. R. (2018). Feasibility and preliminary efficacy of a teacher-facilitated high-intensity interval training intervention for older adolescents. *Pediatric Exercise Science, 31*(1), 107–117. https://doi.org/10.1123/pes.2018-0039

Leahy, A. A., Mavilidi, M. F., Smith, J. J., Hillman, C. H., Eather, N., Barker, D., & Lubans, D. R. (2020). Review of high-intensity interval training for cognitive and mental health in youth. *Medicine & Science in Sports & Exercise, 52*(10), 2224–2234. https://doi.org/10.1249/MSS.0000000000002359

Little, J. P., Safdar, A., Wilkin, G. P., Tarnopolsky, M. A., & Gibala, M. J. (2010). A practical model of low-volume high-intensity interval training induces mitochondrial biogenesis in human skeletal muscle: Potential mechanisms. *The Journal of Physiology, 588*(6), 1011–1022. https://doi.org/10.1113/jphysiol.2009.181743

Liu, J., Zhang, Y., Li, X., Wang, D., Shi, B., You, Y., Min, L., Luo, B., Li, Y., Di, Q., & Ma, X. (2022). Exercise improves mental health status of young adults via attenuating inflammation factors but modalities matter. *Frontiers in Psychiatry, 13*, Article 1067890. https://doi.org/10.3389/fpsyt.2022.1067890

Lubans, D. R., Smith, J. J., Eather, N., Leahy, A. A., Morgan, P. J., Lonsdale, C., Plotnikoff, R. C., Nilsson, M., Kennedy, S. G., Holliday, E. G., Weaver, N., Noetel, M., Shigeta, T. T., Mavilidi, M. F., Valkenborghs, S. R., Gyawali, P., Walker, F. R., Costigan, S. A., & Hillman, C. H. (2020). Time-efficient intervention to improve older adolescents' cardiorespiratory fitness: Findings from the "Burn 2 Learn" cluster randomised controlled trial. *British Journal of Sports Medicine, 55*(13), 751–758. https://doi.org/10.1136/bjsports-2020-103277

Lubans, D. R., Smith, J. J., Morgan, P. J., Beauchamp, M. R., Miller, A., Lonsdale, C., Parker, P., & Dally, K. (2016b). Mediators of psychological well-being in adolescent boys. *Journal of Adolescent Health, 58*(2), 230–236. https://doi.org/10.1016/j.jadohealth.2015.10.010

Lubans, D., Richards, J., Hillman, C., Faulkner, G., Beauchamp, M., Nilsson, M., Kelly, P., Smith, J., Raine, L., & Biddle, S. (2016a). Physical activity for cognitive and mental health in youth: A systematic review of mechanisms. *Pediatrics, 138*(3), Article e20161642. https://doi.org/10.1542/peds.2016-1642

Lucibello, K. M., Paolucci, E. M., Graham, J. D., & Heisz, J. J. (2020). A randomized control trial investigating high-intensity interval training and mental health: A novel non-responder phenotype related to anxiety in young adults. *Mental Health and Physical Activity, 18*, Article 100327. https://doi.org/10.1016/j.mhpa.2020.100327

Madigan, S., Racine, N., Vaillancourt, T., Korczak, D. J., Hewitt, J. M. A., Pador, P., Park, J. L., McArthur, B. A., Holy, C., & Neville, R. D. (2023). Changes in Depression and Anxiety Among Children and Adolescents From Before to During the COVID-19 Pandemic: A Systematic Review and Meta-analysis. *JAMA pediatrics, 177*(6), 567–581. https://doi.org/10.1001/jamapediatrics.2023.0846

Mason, J. E., & Asmundson, G. J. (2018). A single bout of either sprint interval training or moderate intensity continuous training reduces anxiety sensitivity: A randomized controlled trial. *Mental Health and Physical Activity, 14*, 103–112. https://doi.org/10.1016/j.mhpa.2018.02.006

May, R. W., Seibert, G. S., Sanchez-Gonzalez, M. A., & Fincham, F. D. (2018). Self-regulatory biofeedback training: An intervention to reduce school burnout and improve cardiac functioning in college students. *Stress, 22*(1), 1–8. https://doi.org/10.1080/10253890.2018.1501021

McGuinness, L. A., & Higgins, J. P. T. (2021). Risk-of-bias VISualization (robvis): An R package and Shiny web app for visualizing risk-of-bias assessments. *Research Synthesis Methods, 12*(1), 55–61. https://doi.org/10.1002/jrsm.1411

Methley, A. M., Campbell, S., Chew-Graham, C., McNally, R., & Cheraghi-Sohi, S. (2014). PICO, PICOS and SPIDER: A comparison study of specificity and sensitivity in three search tools for qualitative systematic reviews. *BMC Health Services Research, 14*, Article 579. https://doi.org/10.1186/s12913-014-0579-0

Miller, A. H., & Raison, C. L. (2016). The role of inflammation in depression: From evolutionary imperative to modern treatment target. *Nature Reviews Immunology, 16*(1), 22–34. https://doi.org/10.1038/nri.2015.5

Naylor, P. J., Nettlefold, L., Race, D., Hoy, C., Ashe, M. C., Wharf Higgins, J., & McKay, H. A. (2015). Implementation of school-based physical activity interventions: A systematic review. *Preventive Medicine, 72*, 95–115. https://doi.org/10.1016/j.ypmed.2014.12.034

Noetel, M., Sanders, T., Gallardo-Gómez, D., Taylor, P., del Pozo Cruz, B., van den Hoek, D., Smith, J. J., Mahoney, J., Spathis, J., Moresi, M., Pagano, R., Pagano, L., Vasconcellos, R., Arnott, H., Varley, B., Parker, P., Biddle, S., & Lonsdale, C. (2024). Effect of exercise for depression: Systematic review and network meta-analysis of randomised controlled trials. *BMJ, 384*, Article e075847. https://doi.org/10.1136/bmj-2023-075847

Ormsbee, M. J., Kinsey, A. W., Chong, M., Friedman, H. S., & Panton, L. B. (2013). The influence of high-intensity interval training on the salivary cortisol response to a psychological stressor and mood state in

non-sedentary college students. *Journal of Exercise Physiology Online, 16*(1), 105–116. https://www.sjsu.edu/people/karen.moreno/courses/c1/s2/JEPonlineFEBRUARY2013_M_Ormsbee.pdf

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., … Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ, 372*, n71. https://doi.org/10.1136/bmj.n71

Paluska, S. A., & Schwenk, T. L. (2000). Physical activity and mental health: Current concepts. *Sports Medicine, 29*(3), 167–180. https://doi.org/10.2165/00007256-200029030-00003

Patalay, P., & Gage, S. H. (2019). Changes in millennial adolescent mental health and health-related behaviours over 10 years: A population cohort comparison study. *International Journal of Epidemiology, 48*(5), 1650–1664. https://doi.org/10.1093/ije/dyz006

Patton, G. C., Sawyer, S. M., Santelli, J. S., Ross, D. A., Afifi, R., Allen, N. B., Arora, M., Azzopardi, P., Baldwin, W., Bonell, C., Kakuma, R., Kennedy, E., Mahon, J., McGovern, T., Mokdad, A. H., Patel, V., Petroni, S., Reavley, N., Taiwo, K., Waldfogel, J., … Viner, R. M. (2016). Our future: a Lancet commission on adolescent health and wellbeing. *Lancet (London, England), 387*(10036), 2423–2478. https://doi.org/10.1016/S0140-6736(16)00579-1

Pilling, S., Anderson, I., Goldberg, D., Meader, N., & Taylor, C. (2009). Depression in adults, including those with a chronic physical health problem: Summary of NICE guidance. *BMJ, 339*, Article b4108. https://doi.org/10.1136/bmj.b4108

Racine, N., McArthur, B. A., Cooke, J. E., Eirich, R., Zhu, J., & Madigan, S. (2021). Global prevalence of depressive and anxiety symptoms in children and adolescents during COVID-19: A meta-analysis. *JAMA pediatrics, 175*(11), 1142–1150. https://doi.org/10.1001/jamapediatrics.2021.2482

Rodriguez-Ayllon, M., Cadenas-Sánchez, C., Estévez-López, F., Muñoz, N. E., Mora-Gonzalez, J., Migueles, J. H., Molina-Garcia, P., Henriksson, H., Mena-Molina, A., Martinez-Vizcaino, V., Catena, A., & Esteban-Cornejo, I. (2019). Role of physical activity and sedentary behavior in the mental health of preschoolers, children and adolescents: A systematic review and meta-analysis. *Sports Medicine, 49*(9), 1383–1410. https://doi.org/10.1007/s40279-019-01187-0

Ruggero, C. J., Petrie, T., Sheinbein, S., Greenleaf, C., & Martin, S. (2015). Cardiorespiratory fitness may help in protecting against depression among middle school adolescents. *The Journal of adolescent health: official publication of the Society for Adolescent Medicine, 57*(1), 60–65. https://doi.org/10.1016/j.jadohealth.2015.03.016

Schardt, C., Adams, M. B., Owens, T., Keitz, S., & Fontelo, P. (2007). Utilization of the PICO framework to improve searching PubMed for clinical questions. *BMC medical informatics and decision making, 7*, 16. https://doi.org/10.1186/1472-6947-7-16

Schmidt, N. B., Keough, M. E., Mitchell, M. A., Reynolds, E. K., Macpherson, L., Zvolensky, M. J., & Lejuez, C. W. (2010). Anxiety sensitivity: Prospective prediction of anxiety among early adolescents. *Journal of Anxiety Disorders, 24*(5), 503–508. https://doi.org/10.1016/j.janxdis.2010.03.007

Schuch, F. B., Vancampfort, D., Richards, J., Rosenbaum, S., Ward, P. B., & Stubbs, B. (2016). Exercise as a treatment for depression: A meta-analysis adjusting for publication bias. *Journal of Psychiatric Research, 77*, 42–51. https://doi.org/10.1016/j.jpsychires.2016.02.023

Schwarzer, G. (2025). meta: General package for meta-analysis (Version 8.2-1) [R package]. Comprehensive R Archive Network (CRAN). https://CRAN.R-project.org/package=meta

Shandu, N. M., Mathunjwa, M. L., Shaw, B. S., & Shaw, I. (2023). Effects of high-intensity interval training and continuous aerobic training on health-fitness, health-related quality of life, and psychological measures in college-aged smokers. *International Journal of Environmental Research and Public Health, 20*(1), Article 653. https://doi.org/10.3390/ijerph20010653

Silva, S. A., Silva, S. U., Ronca, D. B., Gonçalves, V. S. S., Dutra, E. S., & Carvalho, K. M. B. (2020). Common mental disorders prevalence in adolescents: A systematic review and meta-analyses. *PLOS ONE, 15*(4), Article e0232007. https://doi.org/10.1371/journal.pone.0232007

Skvarc, D. R., & Fuller-Tyszkiewicz, M. (2024). Calculating repeated-measures meta-analytic effects for continuous outcomes: A tutorial on pretest–posttest–controlled designs. *Advances in Methods and Practices in Psychological Science, 7*(1), 1–13. https://doi.org/10.1177/25152459231217238

Stubbs, B., Vancampfort, D., Rosenbaum, S., Firth, J., Cosco, T., Veronese, N., Gaughran, F., & Schuch, F. B. (2017). An examination of the anxiolytic effects of exercise for people with anxiety and stress-related disorders: A meta-analysis. *Psychiatry Research, 249*, 102–108. https://doi.org/10.1016/j.psychres.2016.12.020

Twenge, J. M., Cooper, A. B., Joiner, T. E., Duffy, M. E., & Binau, S. G. (2019). Age, period, and cohort trends in mood disorder indicators and suicide-related outcomes in a nationally representative dataset, 2005-2017. *Journal of abnormal psychology, 128*(3), 185–199. https://doi.org/10.1037/abn0000410

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software, 36*(3), 1–48. https://doi.org/10.18637/jss.v036.i03

GRUPO EDITORIAL DE PSICOFUNDACIÓN

Psicothema (2026) 38(3) 149-160

Psicothema

https://www.psicothema.com/es • eISSN 1886-144X

Colegio Oficial de Psicología del Principado de Asturias

Psicothema

Article

The Illness Management and Recovery (IMR) Scale: A Network Analysis of User and Professional Perspectives

Nuria Martín-Ordiales[1], Albert Sesé[2], Mª Pilar Martín-Chaparro[1], Maite Barrios[3] and Mª Dolores Hidalgo Montesinos[1]

[1] University of Murcia (Spain)
[2] University of the Balearic Islands (Spain)
[3] University of Barcelona (Spain)

ARTICLE INFO

Received: 08/11/2025
Accepted: 09/02/2026

Keywords:
Network analysis
IMR scale
Mental health recovery
Psychometrics
Measurement invariance

ABSTRACT

Background: Recovery is a multidimensional, subjective process that emphasizes personal strategies and empowerment beyond symptom remission. The *Illness Management and Recovery* (IMR) scale incorporates the perspectives of both users and professionals, although its internal structure has had only limited examination through advanced psychometric methods. Network analysis was applied to the IMR scale with the aims of (1) exploring interrelationships among items, (2) assessing network stability, and (3) analyzing measurement invariance between users and professionals. **Method:** Data were collected from 172 users and 49 mental health professionals using the IMR scale. Network analysis used the *Triangulated Maximally Filtered Graph* (TMFG) method and the *walktrap* algorithm, with stability assessed through *bootstrapping*. **Results:** The analysis indicated 39 significant connections among 15 items, highlighting coping strategies as the central node, along with symptom distress, functional difficulties, and symptom recovery. *Exploratory Graph Analysis* identified three factors: (1) Personal Growth and Planning, (2) Social and Health Behavior Management, and (3) Symptom Management and Resilience. **Conclusions:** Network analysis underscores the importance of coping strategies, suggesting that interventions should focus on strengthening these skills to support recovery.

La Escala de Gestión de la Enfermedad y Recuperación: un Análisis de Redes Desde las Perspectivas de Usuarios y Profesionales

Palabras clave:
Análisis de redes
Escala IMR
Recuperación en salud mental
Psicometría
Invariancia de medida

RESUMEN

Antecedentes: La recuperación es un proceso multidimensional y subjetivo que enfatiza las estrategias personales y el empoderamiento más allá de la remisión sintomática. La escala *Illness Management and Recovery* (IMR) integra las perspectivas de usuarios y profesionales, aunque su estructura interna ha sido escasamente examinada mediante métodos psicométricos avanzados. Se aplicó análisis de redes a la escala IMR con el objetivo de: (1) explorar las interrelaciones entre los ítems, (2) evaluar la estabilidad de la red y (3) analizar la invariancia de medida entre usuarios y profesionales. **Método**: Se recopilaron datos de 172 usuarios y 49 profesionales de salud mental aplicando la escala IMR. El análisis de redes empleó el método *Triangulated Maximally Filtered Graph* (TMFG) y el algoritmo *walktrap*, evaluándose la estabilidad mediante *bootstrapping*. **Resultados**: El análisis reveló 39 conexiones significativas entre 15 ítems, destacando las estrategias de afrontamiento como nodo central, junto con malestar sintomático, dificultades funcionales y recuperación sintomática. El *Exploratory Graph Analysis* identificó tres factores: (1) Crecimiento personal y planificación, (2) Gestión social y de conductas de salud, y (3) Manejo de síntomas y resiliencia. **Conclusiones**: El análisis de redes resalta la importancia de las estrategias de afrontamiento, sugiriendo orientar las intervenciones hacia su fortalecimiento.

Cite as: Martín-Ordiales, N., Sesé, A., Martín-Chaparro, M. P., Barrios, M., & Hidalgo Montesinos, M. D. (2026). The Illness Management and Recovery (IMR) Scale: A network analysis of user and professional perspectives. *Psicothema, 38*(3), 149-160. https://doi.org/10.70478/psicothema.2026.38.14
Corresponding author: Maite Barrios, mbarrios@ub.edu

Individuals diagnosed with severe mental disorders, such as schizophrenia, schizoaffective disorder, bipolar disorder or major depressive disorder with psychotic features, have often been grouped together due to overlapping symptom profiles and treatment needs (WHO, 2022). Historically, mental health classifications, such as those found in the ICD-11, have adopted a categorical approach based on symptom clusters and diagnostic thresholds. While clinically useful, these approaches may inadequately reflect the heterogeneity and complexity of individual experiences (Marder & Galderisi, 2017). A related issue here is that dominant conceptualizations of mental illness have long been framed within biomedical models, emphasizing neurobiological and genetic determinants of behavior (Insel & Quirion, 2005), although some authors have critically questioned the limitations of purely diagnostic or categorical frameworks (Johnstone & Boyle, 2023). By contrast, contemporary frameworks increasingly recognize the role of psychosocial factors, trauma histories, and environmental stressors in the development and course of mental health conditions (Read et al., 2005). Moreover, recent work highlights the importance of integrative and transdiagnostic models that combine genetic, psychological, and environmental factors, moving beyond purely biomedical approaches to better capture the complexity of mental health and psychopathology (Eaton et al., 2023). From a network perspective, some authors propose that dynamically interacting symptoms and elements constitute the disorder itself, forming personalized, transdiagnostic systems that may improve diagnosis and intervention (Roefs et al., 2022). Furthermore, network models challenge reductionist explanations by showing that mental disorders emerge from symptom interactions rather than a single biological cause, underscoring the need for holistic research strategies (Borsboom et al., 2018). This broader understanding has facilitated the emergence of more person-centered, holistic approaches in research and practice (Davidson & Roe, 2009; Slade, 2009).

One of the most significant conceptual advances was the emergence of the recovery model (Anthony, 1993), although it is only in the last decade or so that this model began to be adopted in national policies, translated into practical instruments, and subjected to scientific evaluation (Ahmed et al., 2012; Frost et al., 2017; Canacott et al., 2019). Unlike traditional perspectives that primarily focused on symptom remission and functional restoration, the recovery model emphasizes the possibility of living a meaningful and fulfilling life despite ongoing symptoms. Recovery is thus understood as a multidimensional and deeply subjective process (Slade, 2009), encompassing not only clinical stability but also hope, empowerment, subjective well-being, and personal agency. More recently, a consensus definition of the recovery process has been sought, with researchers recognizing the need to incorporate the perspectives of health professionals, users of mental health services, family members, and other significant individuals (Sampietro et al., 2022; Guerrero et al., 2024, 2025).

Being able to quantitatively assess the recovery process is also essential for establishing appropriate interventions for each user. To operationalize this construct, several measurement instruments have been developed, including the Recovery Assessment Scale (Giffort et al., 1995), the Mental Health Recovery Measure (Young et al., 2003), and the Recovery Self-Assessment (O´Connell et al., 2005), among other more recent ones such as the Self-Identify Stage of Recovery (SIRS) scale (Sampietro et al., 2024), the Individual Recovery Outcomes Counter (I.ROC) (Garrido-Cervera et al., 2025) and the MARS-12 (Balluerka et al., 2024). However, many of these tools primarily capture the perspective of users, while failing to incorporate the clinical insights of professionals (Felix et al., 2024). Furthermore, low use of assessment scales has been observed in clinical and intervention settings due to a lack of time and interest on the part of professionals (Gilbody et al., 2002).

Although recent research acknowledges the importance of collaboration between users, professionals, and family members in the recovery process (Burger et al., 2024), this is rarely reflected in existing scales. A notable exception is the Illness Management and Recovery (IMR) scale (Mueser et al., 2005), a unique instrument in that it captures both user and professional perspectives, thus providing a more comprehensive view of the recovery process. Importantly, the IMR is designed to measure multiple interrelated domains that are central to personal recovery, including goal setting, symptom management, coping strategies, knowledge about mental illness, social support, and functional outcomes, reflecting the core components of recovery-oriented practice and aligning with theoretical frameworks that emphasize empowerment, hope, and self-directed recovery (Hasson-Ohayon et al., 2008). A recent systematic review further supports that the IMR demonstrates relationships with recovery-relevant constructs such as hope, control over life, identity, symptom management, and adaptive coping, highlighting its utility not only as a clinical outcome measure but as an instrument capturing domains that actively contribute to the recovery process (Martín-Ordiales et al., 2024).

Previous studies have validated the psychometric properties of the IMR scale in diverse populations (Hasson-Ohayon et al., 2008; Sklar et al., 2012; Jensen et al., 2019). However, no study to date has analysed the internal structure of the scale using more advanced psychometric techniques such as network analysis. Beyond its methodological novelty, network analysis is particularly well suited for the study of complex and multidimensional constructs such as mental health recovery, as assessed by the IMR. Network analysis represents a paradigm shift in how psychological phenomena are conceptualized (Borsboom, 2017), insofar as it treats psychological elements as interconnected nodes within a complex system rather than as simple manifestations of an underlying latent cause. This perspective is especially relevant for recovery-oriented constructs, in which processes such as hope, empowerment, symptom management, and social functioning are theorized to be functionally interdependent rather than hierarchically organized. This approach allows for the identification of the most central items, bridge elements, and key domains within the network, offering insights that may not be fully captured by traditional latent variable models such as Structural Equation Models (SEM) or Item Response Models (IRT). Moreover, it paves the way for more personalized and collaborative mental health care, enabling professionals and users to jointly explore which network elements may maintain distress or support recovery, and thus design more targeted interventions (Epskamp et al., 2018). However, most current applications of this approach have focused primarily on symptom networks related to psychopathology (Buchwald et al., 2024), with limited exploration of positive processes or recovery in mental health (Fried et al., 2016).

Based on the conceptual framework of recovery-oriented processes, IMR items are expected to form distinct clusters corresponding to key recovery domains, including hope and empowerment, symptom

management, and social functioning (Mueser et al., 2005; Hasson-Ohayon et al., 2008). From a network perspective, items related to self-management and coping are anticipated to occupy central positions, reflecting their influence across multiple recovery processes. These expected clusters and central items provide a network-based theoretical rationale for examining the internal structure of the IMR, linking broad domains to potential item communities and guiding the interpretation of central and bridging nodes within the scale. Accordingly, we formulate the following hypotheses: (1) IMR items will cluster into communities corresponding to the theorized recovery domains, and (2) self-management and coping items will emerge as central nodes in the network.

To address these expectations, the present study applies network analysis to the IMR scale with the aim of identifying the central aspects of recovery from the perspectives of both users and professionals. The specific aims were as follows: (1) to examine the interrelationships among IMR items from a network perspective, (2) to assess the stability and structure of the estimated network, and (3) to analyze measurement invariance across the client (user self-report) and clinician versions of the scale.

Method

Participants

Users of mental health services: A total of 172 people completed the IMR client scale. Of these, 123 were men (71.5 %) and 45 were women (26.2 %), ranging in age from 19 to 68 years (M = 47.2, SD = 9.4). The majority of participants were single (75.6 %). The primary diagnosis in the sample was schizophrenia (61 %). A slight majority (56.4 %) had a recognized work disability.

Mental health professionals: A total of 49 mental health professionals completed the IMR clinician scale, a process that yielded 166 individual assessments, as each professional could report on multiple service users. Professionals were required to have detailed knowledge of the users they were providing information about, consistent with their being a lead professional in the patient's care and recovery process. The sample comprised 35 women (71.4 %) and 14 men (28.6 %), with the majority being either psychologists (53.1 %) or social workers (18.4 %). They reported an average of 11.1 years of clinical experience (SD = 7.9), including an average of 8.9 years (SD = 6.7) experience working specifically with individuals diagnosed with severe mental disorders (Supplementary material 1).

Instruments

The IMR scale (Mueser et al., 2005) is a 15-item scale derived from the IMR program that assesses key domains including progress toward personal goals, functional difficulties, relapse prevention, coping strategies, and use of medication (see Table 1). There are both client and clinician versions of the scale, with higher scores reflecting better recovery and illness management. Items are rated on a 5-point Likert-type response scale. In the current study, mental health professionals and service users completed their corresponding version of the scale. The Spanish translation of the IMR-client and IMR-professional scales (Martin-Ordiales et al., 2026) followed the International Test Commission's 2017 guidelines (Hernández et al., 2020). Participants completed the IMR with a trained professional present to provide standardized instructions and ensure understanding without altering item content.

Previous research has reported varying reliability coefficients for the test scores for the two versions: from .49 (Fortuna et al., 2020) to .85 (Tan et al., 2017) for the client version, and from .69 (Dalum et al., 2016) to .85 (Tan et al., 2017) for the clinician version.

Table 1
Key Domains of the IMR Scale

ITEM	DOMAIN
IMR1	Progress towards personal goals
IMR2	Knowledge about symptoms, treatment, etc.
IMR3	Involvement of family and friends in treatment
IMR4	Relationship with non-family members
IMR5	Time spent on work, volunteering, etc.
IMR6	Symptom distress
IMR7	Functional difficulties
IMR8	Relapse prevention plan
IMR9	Symptom recovery
IMR10	Psychiatric hospitalization
IMR11	Coping
IMR12	Participation in self-help groups
IMR13	Proper use of medication
IMR14	Impact of alcohol consumption
IMR15	Impact of drug use

Procedure

The study followed the ethical guidelines outlined in the Declaration of Helsinki and received approval from the Bioethics Committee of the University of Barcelona (CBUB; IRB No.: IRB00003099).

Participants were recruited through email and telephone outreach from mental health associations and private centers across Spain, using a convenience sampling method. Professionals were informed of the study aims and were asked to contact users who met the inclusion criteria. All participants were subsequently informed individually about the study, provided written informed consent, and completed the data collection protocol confidentially. Data collection took place between December 2022 and October 2023.

Inclusion criteria comprised adults (≥18 years) with a confirmed diagnosis of a severe mental disorder according to ICD-10 or DSM criteria (e.g., schizophrenia spectrum disorders, bipolar disorder, or severe depressive disorder with psychotic features), established at least two years prior to inclusion. All participants were fully informed about the study and provided written informed consent. Exclusion criteria comprised individuals with primary diagnoses other than severe mental illness (e.g., anxiety or substance use disorders), acute psychiatric episodes, significant cognitive impairment, major medical or neurological conditions, or insufficient language proficiency.

Data Analysis

Network Model

We estimated a network model by applying the Triangulated Maximally Filtered Graph (TMFG) method to the 15 items of the IMR scale in a sample of people with severe mental disorder. The TMFG is a network-filtering method designed to maximize the retention of relevant information while minimizing noise, iteratively

adding nodes based on their correlations (Previde Massara et al., 2017). To assess the stability of the network structure, we performed bootstrap resampling to examine variability in edge strength, applying a case-drop bootstrap procedure to evaluate sample dependence. These techniques ensured the robustness and reliability of the estimated network structure. We conducted a set of sensitivity analyses comparing the TMFG approach with an alternative regularized estimator (EBICglasso). All network estimations (both TMFG and GLASSO) were based on polychoric correlation matrices (Supplementary Material 2).

To further evaluate the reliability of the network, a correlation stability analysis was conducted for the centrality measures of the IMR network (Epskamp et al., 2018). This analysis involved repeatedly estimating the network while systematically dropping a proportion of cases from the original sample. The correlation stability coefficient (CS-coefficient) was computed for each centrality index, representing the maximum proportion of cases that can be dropped while maintaining a correlation of at least 0.7 with the original centrality values in 95% of the resamples. According to established guidelines (Epskamp et al., 2018), CS-coefficients below 0.25 indicate unstable estimates, values between 0.25 and 0.5 reflect moderate stability, and values above 0.5 indicate stable estimates.

Exploratory Graph Analysis

Next, we conducted an exploratory graph analysis (EGA) using the *EGAnet* package (Golino et al., 2024). The TMFG method and walktrap algorithm were applied within a bootstrap framework to identify the most stable and representative latent structure underlying the data. The walktrap algorithm, a community detection method based on random walks, identifies densely connected subgraphs by merging communities that random walkers are more likely to traverse together. This method enabled the detection of communities (or factors) among IMR items.

Configural Invariance Analysis

Finally, configural and metric invariance analyses were performed to evaluate whether the factor structure was consistent across the samples of users and professionals. The invariance function of the *EGAnet* package estimates configural invariance by comparing the factor structures of the two groups without imposing parameter constraints. Separate EGA models were estimated for each group to determine whether they shared the same number of factors and a similar pattern of relationships among the 15 IMR items. This procedure provided evidence of concordance between the responses of users and professionals and enabled a nuanced interpretation of the factor structure by highlighting items with stable versus variable associations with specific factors. To account for multiple comparisons and control the false discovery rate (FDR), *p*-values were adjusted using the Benjamini-Hochberg (BH) procedure.

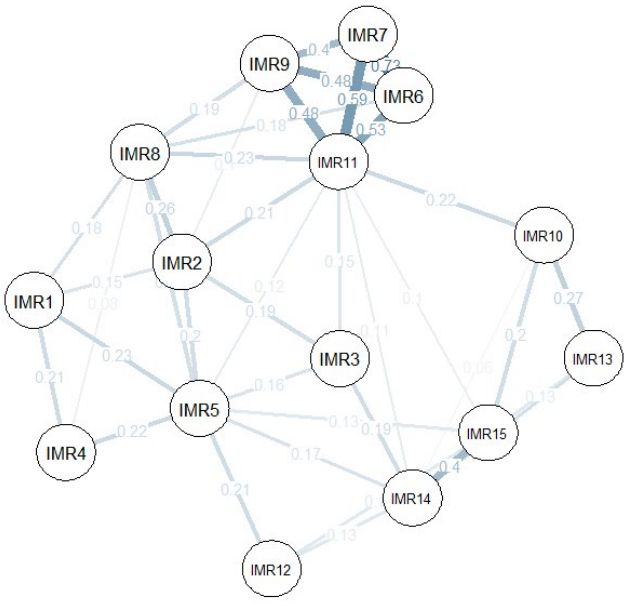
All analyses were conducted using R, with the following packages: *qgraph* (Epskamp et al., 2012), *ggplot2* (Wickham, 2016), *NetworkToolbox* (Christensen, 2018), *bootnet* (Epskamp et al., 2018), and *EGAnet* (Golino et al., 2024).

Results

Network Model

The network analysis offers a comprehensive view of the intricate relationships (edges) between various factors influencing mental health recovery. Figure 1 displays the sample network solution when considering user responses to the IMR items. There are 39 non-zero edges from a total of 105. Coping strategies (IMR11) emerge as a central hub, underscoring their pivotal role in managing symptoms and promoting well-being; note also the close community formed with IMR6, IMR7, and IMR9. The clustering of items related to social support (IMR3, IMR4, and IMR5) and substance use (IMR14, IMR15) reveals potential distinct and interpretable subgroups within the network. Items such as knowledge about symptoms and treatment (IMR2) and participation in self-help groups (IMR12) appear less interconnected, suggesting that while they may be important factors in recovery, their influence may be more specific to individual experiences. The strong connection between coping strategies (IMR11), symptom distress (IMR6), functional difficulties (IMR7), and symptom recovery (IMR9) emphasizes the importance of equipping users with robust coping mechanisms to prevent relapse.

Figure 1
Network Model With Strength Values Between the 15 IMR Items for Users



The centrality statistics (strength, closeness, betweenness, and expected influence) for user responses to the IMR scale are shown in Figure 2.

The item with the highest expected influence in the network is IMR11 (Coping), followed by IMR6 (Symptom distress), IMR7 (Functional difficulties), IMR5 (Time spent on work, volunteering, etc.), and IMR9 (Symptom recovery). The items with the least central role in the model are, in descending order, IMR4 (Relationships with

non-family members), IMR12 (Participation in self-help groups), IMR13 (Proper use of medication), and IMR3 (Involvement of family and friends in treatment). Regarding betweenness, items IMR11 (Coping) and IMR5 (Time spent on work, volunteering, etc.) yield the highest values, meaning that they have a greater number of edges in relation to their nearby nodes, as well as in relation to global closeness. Table 2 shows the values for the network statistics. Strength and expected and expected influence values are the same because the network is undirected and all edges are positive, reflecting the predominantly positive zero-order correlations among IMR items, which assess related aspects of recovery being consistent with the clinical content of the items.

Figure 2

Centrality Plot of the IMR Network for User Responses

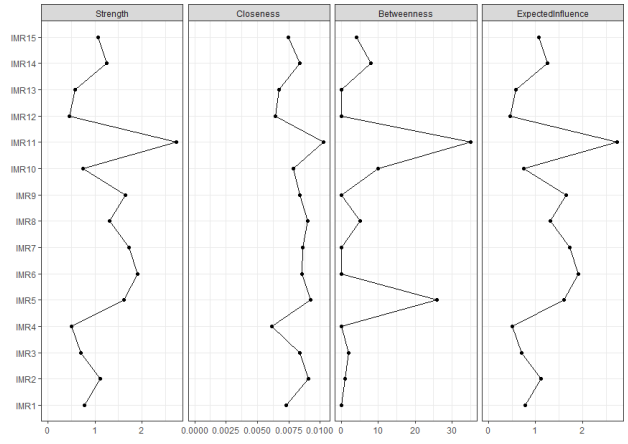


Table 2

Strength, Betweenness, and Closeness Values of the IMR Network Nodes

NODES	STRENGTH	BETWEENNESS	CLOSENESS
IMR11	2.73	70	0.010
IMR6	1.91	0	0.009
IMR7	1.72	0	0.009
IMR9	1.64	0	0.008
IMR5	1.61	52	0.009
IMR8	1.31	10	0.009
IMR14	1.25	16	0.008
IMR2	1.11	2	0.009
IMR15	1.07	8	0.007
IMR1	0.77	0	0.007
IMR10	0.75	20	0.008
IMR3	0.70	4	0.008
IMR13	0.58	0	0.007
IMR4	0.51	0	0.006
IMR12	0.46	0	0.006

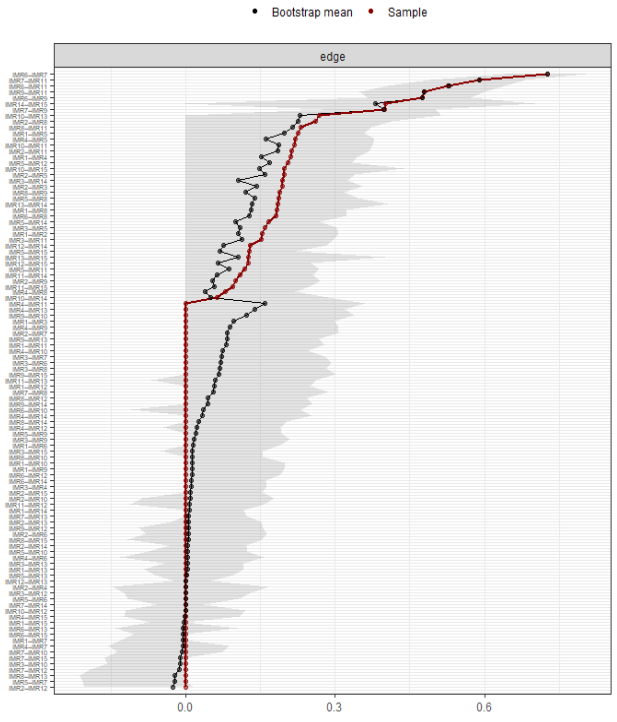
Note: Expected influence values are omitted because they are identical to strength values in an undirected network with exclusively positive edges.

A bootstrap analysis of all edges in the network provides crucial insights into the stability and reliability of the network structure. Figure 3 illustrates the variability in edge weights across 1000 resampled networks, offering a visual representation of the confidence intervals for each connection. The black line represents the average values of the bootstrap samples, offering a smoothed representation of the underlying distribution, while the red line depicts the observed values of the sample data. The gray shaded area represents the 95% confidence intervals for the bootstrap samples. Edges that appear consistently across bootstrap samples, indicated by narrow confidence intervals, suggest robust and reliable relationships between nodes.

Overall, the discrepancies between the bootstrap values and the sample values for the network edges are minimal and generally not significant. Most of the network structure observed in the original sample is consistent with the bootstrap results. While there are a few instances where the bootstrap solution yields non-zero values for edges that were initially zero in the sample network, these differences are characterized by very low strength values. This suggests that even where discrepancies exist, they represent weak connections that likely have little practical impact on the overall network structure. The consistency between the bootstrap and sample results largely supports the reliability of the observed network, indicating that the relationships between IMR scale items are stable across resamples. This pattern of stability is particularly observed for the edges connecting IMR6 ↔ IMR7, IMR7 ↔ IMR11, IMR6 ↔ IMR11, IMR9 ↔ IMR11, IMR6 ↔ IMR9, and IMR7 ↔ IMR9, suggesting that these associations are robust and reliable within the current sample.

Figure 3

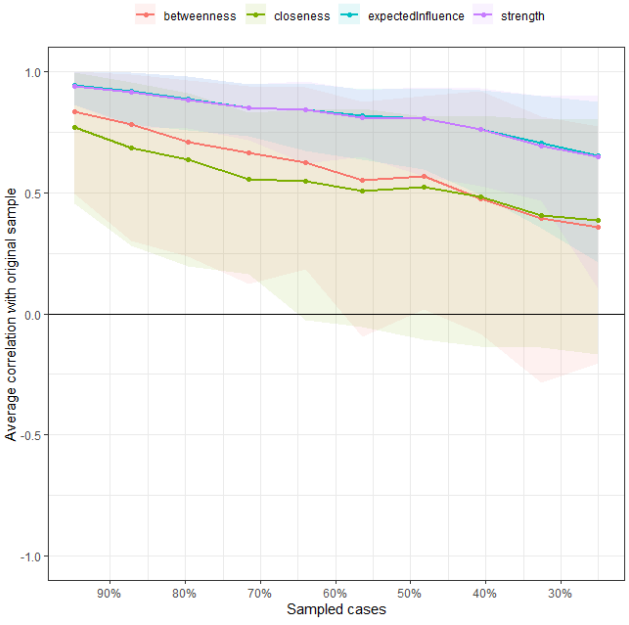
Bootstrap Edge Stability Analysis: Confidence Intervals and Network Connection Reliability



Results revealed varying levels of stability across centrality measures. Betweenness and closeness centrality demonstrated CS-coefficients of 0, indicating extremely poor stability. By contrast, expected influence and strength centrality exhibited moderate stability, with CS-coefficients of .36. This suggests that up to 36 % of the sample can be dropped while maintaining a 95 % probability of a correlation of at least .70 with the original centrality values.

Consequently, only expected influence and strength centrality measures can be cautiously interpreted in the current network structure. Figure 4 displays a plot with the results of a case-dropping bootstrap analysis for centrality measures in the IMR network. This plot illustrates the stability of centrality indices as cases are progressively removed from the sample. The plot shows four lines, each representing a different centrality measure (strength, expected influence, betweenness, and closeness). Lines that remain high and stable as the proportion of dropped cases increases indicate more stable centrality measures. Conversely, lines that drop quickly or show high variability suggest fewer stable measures. Generally, correlations above .70 are considered to indicate good stability.

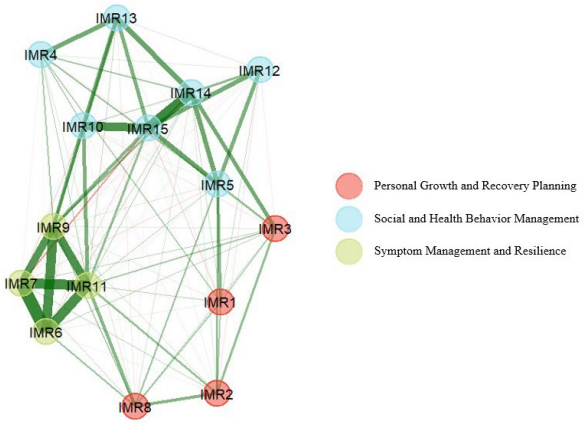
Figure 4
Centrality Stability in IMR Network: Case-dropping Bootstrap Analysis



Exploratory Graph Analysis

We estimated the EGA network using the 15 items of the IMR scale with a bootstrap approach involving 1,000 resamples. The most prevalent latent structure identified was a 3-factor model (59.5 %), followed by a 2-factor model (31.3 %), a 4-factor model (8.9 %), and an almost residual 5-factor model (0.3 %). Figure 5 displays the EGA network representing the 3-factor latent structure derived from the bootstrap analysis.

Figure 5
EGA Network for IMR Items After Bootstrap Analysis (n = 1000)



The model reveals three communities or factors: *Personal Growth and Recovery Planning*, which highlights personal goals, knowledge acquisition, social support, and relapse prevention, all essential elements for long-term recovery and personal empowerment; *Social and Health Behavior Management*, encompassing the influence of social relationships, activities, and health-related behaviors, such as medication adherence and the impact of substance use, including alcohol and drugs; and *Symptom Management and Resilience*, reflecting the focus on coping strategies, symptom recovery, and the challenges related to functioning and the distress caused by symptoms.

The Personal Growth and Recovery Planning factor is composed of items IMR1 (Progress toward personal goals), IMR2 (Knowledge about symptoms, treatments, etc.), IMR3 (Involvement of family and friends), and IMR8 (Relapse prevention plan). The Social and Health Behavior Management factor comprises items IMR4 (Relationship with non-family members), IMR5 (Time spent on work, volunteering, etc.), IMR10 (Psychiatric hospitalization), IMR12 (Participation in self-help groups), IMR13 (Proper use of medication), IMR14 (Impact of alcohol consumption), and IMR15 (Impact of drug use). Of these, the items with the strongest interrelationship are IMR10 (Psychiatric hospitalization), IMR14 (Impact of alcohol consumption), and IMR15 (Impact of drug use). Finally, the Symptom Management and Resilience factor is particularly notable for its consistency. This community comprises items IMR11 (Coping), IMR9 (Symptom recovery), IMR7 (Functional difficulties), and IMR6 (Symptom distress).

However, the robustness of this 3-factor structure may be compromised by volatility in the assignment of certain items. Table 3 shows the percentage of assignment of each item to the five potential

factors identified through the 1,000 bootstrap resamples. This helps to provide a clearer understanding of the stability and consistency of item-factor associations across resamples.

The table reveals that some items exhibit less stable factor assignments, with notable overlaps across multiple factors. Specifically, IMR1, IMR2, and IMR3 show significant distributions across different factors. IMR1 is assigned to factor 2 (36.4 %) and factor 3 (50.1%), IMR2 is mainly assigned to factor 3 (50.6 %) but also has a considerable overlap with factor 1 (33.8 %), while IMR3 is distributed across factor 2 (36.9 %) and factor 3 (39.3 %). IMR8, which is predominantly associated with factor 2 (51 %), also shows some overlap with factor 3 (39.6 %). These overlapping assignments indicate that these items have ambiguous associations, contributing to a less stable factor composition.

By contrast, the items in factor 1 (IMR6, IMR7, IMR9, and IMR11) show very high stability, with all assignments exceeding 90 %. Specifically, IMR6 and IMR7 are almost exclusively assigned to factor 1 (99.8 % each), while IMR9 (93.8 %) and IMR11 (98.7 %) are also predominantly assigned to this factor. This strong and consistent assignment of items to factor 1 indicates that it is the most stable and clearly defined factor in the analysis. Similarly, factor 2 shows strong and stable associations with IMR4, IMR10, IMR12, IMR13, IMR14, and IMR15, with the latter two items (IMR14 and IMR15) demonstrating particularly high assignments (88.8 % and 88.3 %, respectively). However, item IMR5, initially related to factor 2 (.584), also shows some overlap with factor 3 (.402). These stable assignments suggest that factors 1 and 2 have clearer and more consistent structures in comparison with factor 3

Table 3

Percentage of Assignment of Each Item to the Four Latent Structures with 2, 3, 4 and 5 Factors

	1	2	3	4	5
IMR6	**.998**	.000	.001	.001	.000
IMR7	**.998**	.000	.001	.001	.000
IMR11	**.987**	.006	.006	.001	.000
IMR9	**.938**	.020	.039	.003	.000
IMR8	*.510*	.047	*.396*	.047	.000
IMR14	.008	**.888**	.096	.007	.000
IMR15	.019	**.883**	.092	.005	.000
IMR13	.045	**.801**	.134	.018	.002
IMR12	.014	**.792**	.184	.009	.001
IMR10	.197	**.679**	.116	.008	.000
IMR4	.084	**.635**	.262	.017	.002
IMR5	.002	*.584*	*.402*	.011	.001
IMR2	*.338*	.099	*.506*	.057	.000
IMR1	.110	*.364*	*.501*	.025	.000
IMR3	.194	*.369*	*.393*	.043	.000

Note: Factors identified across 1,000 bootstrap resamples (high % in bold and overlaps in italics)

Configural Invariance

Configural invariance was only achieved with 13 items forming a 3-factor solution, as IMR4 and IMR5 showed different factor adscription. For the interpretation of metric invariance, only p-values adjusted using the Benjamini–Hochberg procedure were considered as the primary inferential criterion, while uncorrected p-values are reported for descriptive purposes.

Once the items with consistent factor adscription were identified, we proceeded to analyze metric invariance, that is, the equality of the factor loadings of the items across both groups: users and professionals. Table 4 presents the results of the analysis, including the 13 configurally invariant items, their factor loadings in the two groups.

Table 4

Factor Loadings of EGA Configural Solutions for Users and Professionals

	FACTOR 1		FACTOR 2		FACTOR 3	
ITEMS	USERS	PROF.	USERS	PROF.	USERS	PROF.
IMR1	.73	.79	.00	.00	.00	.09
IMR2	.80	.99	.08	.32	.00	.10
IMR3	.75	.65	.05	.10	.00	.00
IMR6	.08	.04	1.07	.92	.00	.00
IMR7	.00	.15	1.05	1.00	-.11	.17
IMR8	.77	.91	.19	.10	.00	.10
IMR9	.11	.00	.96	.55	.24	.28
IMR10	.00	.00	.13	.22	.83	.85
IMR11	.23	.34	1.01	1.00	.10	.20
IMR12	.00	.00	.00	.00	.62	.73
IMR13	.00	.06	.00	.05	.84	.57
IMR14	.00	.23	.06	.19	.97	1.16
IMR15	.00	.00	.01	.16	1.12	1.01

The results of the metric invariance analysis, as shown in Table 5, suggest that the metric invariance assumption holds for almost all items, with the exception of IMR6 and IMR9. After adjustment for multiple comparisons using the Benjamini–Hochberg procedure, only IMR9 showed statistically significant differences in factor loadings between users and professionals. Although IMR6 yielded a significant uncorrected p-value, this effect did not remain significant after correction and was therefore not interpreted as evidence of non-invariance. For all other items (IMR1, IMR2, IMR3, IMR8, IMR7, IMR11, IMR10, IMR12, IMR13, IMR14, and IMR15), no significant differences in factor loadings were found between the two groups, as both their p and p_BH values exceed .05.

Table 5

Metric Invariance Analysis Across Users and Professionals

ITEMS	FACTOR	DIFFERENCE IN LOADINGS	*p*	*p*_BH	SIG	DIRECTION
IMR1	1	-.059	.603	.713	ns	
IMR2	1	-.198	.067	.290	ns	
IMR3	1	.102	.601	.713	ns	
IMR8	1	-.141	.204	.580	ns	
IMR6	2	.153	.008	.052	**, ns	
IMR7	2	.054	.223	.580	ns	
IMR9	2	.409	.002	.026	**,*	1 > 2
IMR11	2	.004	.943	.946	ns	
IMR10	3	-.015	.946	.946	ns	
IMR12	3	-.109	.441	.657	ns	
IMR13	3	.267	.393	.657	ns	
IMR14	3	-.192	.314	.657	ns	
IMR15	3	.111	.455	.657	ns	

Note: Comparison: 1. Users vs. 2. Professionals

* $p < .05$, ** $p < .01$, ns = non-significant

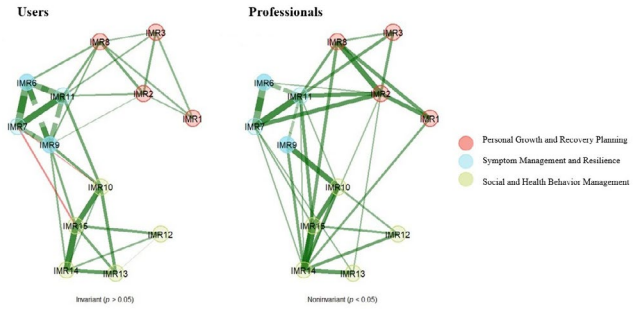
In summary, the metric invariance assumption appears generally supported for most items, although IMR9 shows some variation in factor loadings between groups. Due to the limited sample size, confirmatory factor analysis (CFA/SEM) could not be conducted; therefore, the findings should be interpreted as preliminary rather than definitive. Specifically, the factor loading for IMR9 is higher in

the users' group compared with the professionals' group, suggesting that the relationship between this item and the latent factors differs. For IMR6, the factor loadings are equivalent after adjusting for multiple comparisons, while the factor loadings for the remaining items are also consistent. Figure 6 shows the two 3-factor EGA configural models (13 IMR items) for users and professionals, and also indicates metric invariant and non-invariant items (IMR6 and IMR9 according to the uncorrected p-value).

The metric invariance analysis of the IMR scale across the two groups provided preliminary exploratory evidence that 12 of the 15 items may exhibit equivalent factor loadings. These findings suggest that the IMR instrument captures similar constructs in both groups, but they should be interpreted cautiously due to the exploratory nature of the analysis and the limited sample size.

Figure 6

EGA 3-Factor Configural Models and Metric Invariant and Non-Invariant Items Across the Two Groups



Discussion

The present study represents the first application of network analysis to the IMR scale, offering a novel approach to understanding the interconnectedness of recovery-related dimensions in individuals with severe mental disorders. Through the use of the TMFG method, EGA, and measurement invariance techniques, we were able to identify key nodes, stable factor structures, and the extent to which the IMR scale functions equivalently across users and health professionals.

In a first step, we developed a network model based on the service user version of the IMR scale. Coping (IMR11) appears to be a central aspect of the recovery concept, and accordingly the estimated network shows high values in strength, betweenness, closeness, and expected influence. This reinforces the idea that having personal coping strategies is key to recovery. Coping strategies such as problem-solving and search for social support have been shown to play a fundamental role in managing mental health challenges, particularly among individuals diagnosed with schizophrenia (Mitikhin et al., 2024). By applying their own coping techniques, people can gain control over their thoughts, emotions, and behaviors, which often leads to improved mental well-being and even personal growth (Gautam et al., 2024). Mental health interventions should therefore prioritize the development and reinforcement of adaptive coping tools, rather than focusing predominantly on symptom management. Coping strategies such as underestimation (i.e., minimizing perceived threats), diversion (i.e., redirecting attention), compensatory satisfaction (i.e., seeking fulfillment in alternative domains), reaction control (i.e., regulating emotional expression), and positive self-instruction (i.e., engaging in constructive internal dialogue) represent valuable techniques that can support individuals in dealing with psychological stressors more effectively (Holubova et al., 2015). By understanding these relationships, professionals can tailor interventions to address specific needs and improve treatment outcomes.

Another item with a high value for betweenness was IMR5, which refers to time spent in structured roles. This indicates that engagement in meaningful activities can improve the recovery process, as has been observed in recent studies (Guerrero et al., 2024; Bjørkedal et al., 2023; Chai et al., 2024). It is likewise consistent with the view that having a purpose in life and being able to focus attention on aspects of the person unrelated to mental illness may improve prospects for recovery (Johansson et al., 2025). Nevertheless, the lack of configural invariance for IMR4 and IMR5 suggests that social contact outside the family and engagement in structured roles may be conceptualized differently by users and professionals, highlighting domains where recovery-related meanings may diverge between experiential and clinical perspectives.

We also conducted an exploratory graph analysis, identifying a 3-factor solution as a plausible model. However, the presence of multiple alternative solutions across bootstrap samples indicates a degree of structural variability, therefore this 3-factor solution should be considered exploratory. The factor structures identified through EGA should be interpreted as tentative and hypothesis-generating, rather than as definitive representations of the underlying construct, particularly in light of the limited stability of some centrality measures and the absence of clustering control. Based on their composition, we labeled these factors: *Personal Growth and Recovery Planning*, *Social and Health Behavior Management*, and *Symptom Management and Resilience*. The items included in the first of these factors underscore the importance of setting and achieving personal goals, acquiring knowledge about one's mental health condition, engaging support from family and friends, and planning to prevent relapses. Together, these components form a model grounded in empowerment and strategic self-management. This approach emphasizes the active role of the individual in building resilience and maintaining long-term mental stability through informed, goal-directed actions. This interpretation is supported by existing literature: for example, Guerrero et al. (2024) found that participants reported progress in their recovery when they felt empowered and able to make autonomous decisions.

The composition of the second factor reflects a domain focused on social functioning and health-related behaviors. Here, the inclusion of items related to social relationships, structured activities, and community involvement suggests an emphasis on external interactions and lifestyle organization. By contrast, the presence of items addressing medication adherence and substance use highlights the importance of health maintenance and its impact on psychosocial well-being. This factor may represent a dimension encompassing both social integration and health management. Empirical studies have shown that participation in standardized intervention programs targeting these goals, such as the IMR program, leads to improvements in social functioning and reductions in substance use among individuals with severe mental illness (McGuire et al., 2014; Salyers et al., 2014). Similarly, recent studies support the importance of having meaningful support networks and engaging in personally meaningful social activities to enhance the recovery process (Guerrero et al., 2024; Høgh Egmose et al., 2023; Lee & Yu, 2024).

Although it is a less relevant aspect for users' perception of recovery (Guerrero et al., 2024), adherence to pharmacological treatment has been shown to enhance important dimensions of personal recovery in individuals with schizophrenia, particularly by fostering greater engagement in help-seeking behaviors and active participation in the recovery process (Tapia et al., 2021).

Finally, the structure of the third factor relates to an individual's ability to manage and recover from mental health problems. Specifically, it encompasses the practical and emotional aspects of coping, the process of symptom recovery, difficulties in daily functioning, and symptom distress, reflecting a cohesive domain of personal resilience and symptom management. Coping strategies refer to voluntary acts aimed at reducing stress, such as seeking help, talking about feelings, establishing self-care activities, and taking time for oneself (Gautam et al., 2024). In recovery-focused studies, individuals who employ positive coping strategies—such as planning and acceptance—tend to demonstrate higher levels of resilience (Holubova et al., 2015). This, in turn, results in lower levels of psychological distress and a stronger recovery orientation.

This 3-factor structure aligns with findings from previous psychometric studies that utilized factor analysis of responses to the IMR scale (Hasson-Ohayon et al., 2008; Sklar et al., 2012; Jensen et al., 2019). However, the novel contribution of the present study lies in elucidating the relationships and the strength of association among the constituent items, with the aim of informing and optimizing intervention strategies. From an applied perspective, such exploratory patterns can still offer valuable insights for understanding recovery-related processes in individuals with severe mental health disorders, while underscoring the need for replication in larger, independent samples.

Lastly, the analysis was completed by studying the concordance between the networks obtained from the responses of service users and professionals. In general, the networks obtained appear to be similar, although one significant difference should be highlighted. IMR9 (relating to symptom recovery/relapse) emerged as the node with the highest loading in the users' network, and its value was significantly higher than the corresponding value in the professionals' network, which suggests that this item may hold more significance or be more strongly associated with the latent construct of recovery for users than for professionals. This could indicate that users perceive symptom relapse more directly in relation to their therapeutic progress, while professionals may conceptualize it differently. This item (IMR9) is part of the Symptom Management and Resilience factor. These results are consistent with the idea that users perceive changes in symptoms and their impact on functioning in a way that is distinct from the perception of professionals (Polat et al., 2020). Previous studies have shown that people with mental health problems focus their perception of the recovery process more on issues such as empowerment, creating a life project, having a satisfactory social network, and hope. The stronger loading of IMR9 in the user group should also be interpreted in light of the marked imbalance between user and professional samples, as well as the non-independence of professional ratings, which may have contributed to the observed group difference. Accordingly, this difference in perception must be taken into account so as to offer interventions that address all connected nodes, avoiding focusing efforts on controlling positive and negative symptoms of the disease. This finding highlights the robustness of the IMR scale and confirms that it can provide a consistent assessment of user and professional perceptions of the therapeutic process and recovery.

This study has several limitations that should be acknowledged. First, the sample size is relatively small, which may limit the statistical power, the estimation of predictability values and the stability of network estimations across the diverse recovery-related indicators. Although bootstrap procedures were used to examine the precision of edge weights and the stability of centrality indices, these methods provide only indirect evidence of power and do not replace a formal simulation-based assessment. Thus, these findings should be interpreted as exploratory, and future research should employ larger samples in order to evaluate the robustness of the observed networks. However, it is important to note that the study successfully recruited a clinical sample of 172 individuals with severe mental disorders, predominantly schizophrenia—a population for which recruitment is well-documented to be methodologically challenging due to cognitive impairments, symptom severity, engagement barriers, and other practical difficulties in clinical research settings (Bucci et al., 2015; Deckler et al., 2022). Beyond sample size considerations, the clinical relevance of the sample and the paired responses obtained from the participants' reference professionals provide a rich, multi-perspective dataset that enhances the interpretability and ecological validity of the findings.

Furthermore, the inclusion criteria focused exclusively on individuals with severe mental disorders, which means that other clinical populations, such as those with substance use disorders or less severe mental health conditions, were not represented. This selective sampling may limit the generalizability of the network findings to broader mental health populations. Nevertheless, the detailed insights gathered from both service users and their professionals offer valuable information about recovery processes in severe mental illness that goes beyond the limitations imposed by sample size alone.

Finally, while configural invariance was achieved for most items, two items—IMR4 (Contact with people outside of the family) and IMR5 (Time in structured roles)—did not meet this criterion, indicating potential differences in how these items are interpreted or evaluated by service users and professionals. Future research particularly longitudinal studies, should further explore these items to better understand their behavior, strengthen the relationships between recovery domains, and improve measurement equivalence across respondent types.

Despite these limitations, network analysis has proven to be a useful approach for understanding the structure of the IMR scale and for identifying the most relevant concepts for recovery from both user and professional perspectives. The results provide further support for a 3-factor structure for this scale, although the current study provides additional insight into the central concept of each factor. Importantly, the findings may inform therapeutic interventions in real-world settings by highlighting specific targets for clinical practice. For instance, the identification of *Coping* (IMR11) and *Functional difficulties* (IMR7) as central nodes suggests that clinical teams should prioritize personalized interventions aimed at enhancing coping strategies and functional skills. This could be effectively achieved through psychosocial rehabilitation programs or social skills training, which address these core areas of recovery. The prominence of nodes such as *Relapse prevention plan* (IMR8) and *Progress towards personal goals* (IMR1) indicates that clinical training programs should integrate content focused on relapse planning and patient empowerment. These elements are essential

for supporting sustained recovery and should be emphasized in professional development curricula. Additionally, the observed differences in centrality between service users and professionals—such as the higher centrality of *Symptom recovery* (IMR9) among users—underscore the importance of fostering shared decision-making. Addressing these divergences could help to strengthen the therapeutic alliance by aligning treatment expectations and enhancing collaborative care. In conclusion, based on the findings presented, identifying the recovery aspects that are most relevant to service users, as well as the similarities and differences between users' and professionals' perspectives, may help inform more personalized and person-centered approaches to mental health care. These approaches align with recovery-oriented models that emphasize respect, collaboration, and individual recovery goals (Boardman & Dave, 2020).

In summary, this study adds to the literature on mental health recovery and highlights the utility of the IMR scale in shaping a line of research aligned with the needs of individuals with severe mental disorders.

Author Contributions

Nuria Martín-Ordiales: Conceptualization, Data Curation, Investigation, Resources, Visualization, Writing – Original Draft, Writing – Review and Editing. **Albert Sesé**: Conceptualization, Data Curation, Formal Analysis, Methodology, Writing – Original Draft, Writing – Review and Editing. **Mª Pilar Martín-Chaparro**: Conceptualization, Investigation, Resources, Supervision, Visualization, Writing – Original Draft, Writing – Review and Editing. **Maite Barrios**: Conceptualization, Data Curation, Formal Analysis, Funding Acquisition, Investigation, Methodology, Project Administration, Supervision, Validation, Visualization, Writing – Original Draft, Writing – Review and Editing. **Mª Dolores Hidalgo**: Conceptualization, Data Curation, Formal Analysis, Investigation, Methodology, Project Administration, Resources, Supervision, Validation, Visualization, Writing – Original Draft, Writing – Review and Editing.

Acknowledgements

This study was possible thanks to the collaboration of the following institutions: Fundación Murciana de Salud Mental CURAE MENTIS; ActivaMent Catalunya Associació; APICES-Salud Mental Cartagena; AFES- Salud Mental Murcia y Comarcas; AFESMO- Asociación Salud Mental Molina y Comarca; Fundación Familia, Albacete; AFPEM Horta Sud; FAEMA- Salud Mental Ávila; AFEMAC- Asociación Salud Mental Águilas; AFEMAR- Salud Mental y TEA de la Comarca del Mar Menor; ISOL- Asociación Murcia de Rehabilitación Psicosocial; AFEMTO- Asociación de Familias y Personas con Enfermedad Mental de Totana y Aledo; ASOFEM- Asociación de Salud Mental Lorca y Comarca; AFEMY-Asociación de Familias y Personas con Enfermedad Mental de Yecla; Centro de Día ACOFEM; Federación de Salud Mental Región de Murcia.

Funding

This study was supported by the Ministry of Science and Innovation of Spain-CIN/AEI/10.13039/501100011033 [grant PID2019-109887GB-I00], the Agency for the Management of University and Research Grants of the Government of Catalonia [grant 2021SGR01071].

Conflict of Interests

The authors declare that there are no conflicts of interest.

Data Availability Statement

R code availability statement: available from the corresponding author on reasonable request.

The datasets during and/or analysed during the current study available from the corresponding author on reasonable request.

References

Ahmed, A. O., Buckley, P. F., & Mabe, P. A. (2012). Recovery: International efforts at implementing and advancing the recovery model. *International Psychiatry, 9*(1), 4–6. https://doi.org/10.1192/S1749367600006140

Anthony, W. A. (1993). Recovery from mental illness: The guiding vision of the mental health service system in the 1990s. *Psychosocial Rehabilitation Journal, 16*(4), 11–23. https://doi.org/10.1037/h0095655

Balluerka, N., Gorostiaga, A., Sampietro, H. M., González-Pinto, A., & Aliri, J. (2024). Cross-cultural adaptation and psychometric validation of a Spanish version of the Maryland Assessment of Recovery Scale (MARS-12). *PLoS ONE, 19*(2), Article e0298554. https://doi.org/10.1371/journal.pone.0298554

Bjørkedal, S. B., Bejerholm, U., Hjorthøj, C., Møller, T., & Eplov, L. F. (2023). Meaningful Activities and Recovery (MA&R): A co-led peer occupational therapy intervention for people with psychiatric disabilities. Results from a randomized controlled trial. *BMC Psychiatry, 23*(1), Article 406. https://doi.org/10.1186/s12888-023-04875-w

Boardman, J., & Dave, S. (2020). Person centred care and psychiatry: Some key perspectives. *BJPsych International, 17*(3), 65–68. https://doi.org/10.1192/bji.2020.21

Borsboom, D. (2017). A network theory of mental disorders. *World Psychiatry, 16*(1), 5–13. https://doi.org/10.1002/wps.20375

Borsboom, D., Cramer, A. O. J., & Kalis, A. (2018). Brain disorders? Not really… Why network structures block reductionism in psychopathology research. *Behavioral and Brain Sciences, 1–54*. https://doi.org/10.1017/S0140525X17002266

Bucci, S., Butcher, I., Hartley, S., Neil, S. T., Mulligan, J., & Haddock, G. (2015). Barriers and facilitators to recruitment in mental health services: Care coordinators' expectations and experience of referring to a psychosis research trial. *Psychology and Psychotherapy: Theory, Research and Practice, 88*(3), 335-350. https://doi.org/10.1111/papt.12042

Buchwald, K., Narayanan, A., Siegert, R. J., Vignes, M., Arrowsmith, K., & Sandham, M. (2024). Centrality statistics of symptom networks of schizophrenia: A systematic review. *Psychological Medicine, 54*(6), 1061–1073. https://doi.org/10.1017/S003329172300363X

Burger, T. J., van Eck, R. M., Lachmeijer, M., de Wilde-Schutten, K. R., Lansen, M., van Alphen, C., ... & de Koning, M. B. (2024). Perspective matters in recovery: The views of persons with severe mental illness, family and mental health professionals on collaboration during recovery, a qualitative study. *BMC Psychiatry, 24*(1), Article 802. https://doi.org/10.1186/s12888-024-06198-w

Canacott, L., Moghaddam, N., & Tickle, A. (2019). Is the Wellness Recovery Action Plan (WRAP) efficacious for improving personal and

clinical recovery outcomes? A systematic review and meta-analysis. *Psychiatric Rehabilitation Journal, 42*(4), 372–381. https://doi.org/10.1037/prj0000368

Chai, C. A., Barrios, M., Gómez-Benito, J., Campoverde, K., & Guilera, G. (2024). Validation of the Comprehensive and Brief International Classification of Functioning, Disability and Health Core Sets for schizophrenia from the perspective of individuals diagnosed with the disorder: A worldwide study using focus groups. *Behavioral Sciences, 14*(11), Article 1032. https://doi.org/10.3390/bs14111032

Christensen, A. P. (2018). NetworkToolbox: Methods and measures for brain, cognitive, and psychometric network analysis in R. *The R Journal, 10*(2), 422–439. https://doi.org/10.32614/RJ-2018-065

Dalum, H. S., Waldemar, A. K., Korsbek, L., Hjorthøj, C., Mikkelsen, J. H., Thomsen, K., ... & Eplov, L. F. (2016). Participants' and staffs' evaluation of the Illness Management and Recovery program: A randomized clinical trial. *Journal of Mental Health, 27*, 30–37. https://doi.org/10.1080/09638237.2016.1244716

Davidson, L., & Roe, D. (2009). Recovery from versus recovery in serious mental illness: One strategy for lessening confusion plaguing recovery. *Journal of Mental Health, 16*(4), 459–470. https://doi.org/10.1080/09638230701482394

Deckler, E., Ferland, M., Brazis, S., Mayer, M. R., Carlson, M., & Kantrowitz, J. T. (2022). Challenges and strategies for the recruitment of patients with schizophrenia in a research setting. *International Journal of Neuropsychopharmacology, 25*(11), 924–932. https://doi.org/10.1093/ijnp/pyac058

Eaton, N. R., Bringmann, L. F., Elmer, T., Fried, E. I., Forbes, M. K., Greene, A. L., et al. (2023). A review of approaches and models in psychopathology conceptualization research. *Nature Reviews Psychology, 2*(10), 622–636. https://doi.org/10.1038/s44159-023-00218-4

Epskamp, S., Borsboom, D., & Fried, E. I. (2018). Estimating psychological networks and their accuracy: A tutorial paper. *Behavior Research Methods, 50*(1), 195–212. https://doi.org/10.3758/s13428-017-0862-1

Epskamp, S., Cramer, A. O., Waldorp, L. J., Schmittmann, V. D., & Borsboom, D. (2012). qgraph: Network visualizations of relationships in psychometric data. *Journal of Statistical Software, 48*(4), 1–18. https://doi.org/10.18637/jss.v048.i04

Felix, S., Valery, K. M., Caiada, M., Guionnet, S., Bonilla-Guerrero, J., Destaillats, J. M., & Prouteau, A. (2024). Personal recovery self-report outcome measures in serious mental illness: A systematic review of measurement properties. *Clinical Psychology Review, 112*, Article 102459. https://doi.org/10.1016/j.cpr.2024.102459

Fortuna, L. R., Falgas-Bague, I., Ramos, Z., Porche, M. V., & Alegría, M. (2020). Development of a cognitive-behavioral therapy with integrated mindfulness for Latinx immigrants with co-occurring disorders: Analysis of intermediary outcomes. *Psychological Trauma, 12*, 825–835. https://doi.org/10.1037/tra0000949

Fried, E. I., Epskamp, S., Nesse, R. M., Tuerlinckx, F., & Borsboom, D. (2016). What are "good" depression symptoms? Comparing the centrality of DSM and non-DSM symptoms of depression in a network analysis. *Journal of Affective Disorders, 189*, 314–320. https://doi.org/10.1016/j.jad.2015.09.005

Frost, B. G., Tirupati, S., Johnston, S., Turrell, M., Lewin, T. J., Sly, K. A., & Conrad, A. M. (2017). An Integrated Recovery-oriented Model (IRM) for mental health services: Evolution and challenges. *BMC Psychiatry, 17*(1), Article 22. https://doi.org/10.1186/s12888-016-1164-3

Garrido-Cervera, J. A., Ruiz-Granados, M. I., Cuesta-Vargas, A. I., & Sánchez-Guarnido, A. J. (2025). Validation study for the individual recovery outcomes counter (I.ROC) in a Spanish population. *European Journal of Psychiatry, 39*(2), Article 100287. https://doi.org/10.1016/j.ejpsy.2024.100287

Gautam, S., Jain, A., Chaudhary, J., Gautam, M., Gaur, M., & Grover, S. (2024). Concept of mental health and mental well-being, its determinants and coping strategies. *Indian Journal of Psychiatry, 66*(Suppl. 2), S231–S244. https://doi.org/10.4103/indianjpsychiatry.indianjpsychiatry_707_23

Giffort, D., Schmook, A., Woody, C., Vollendorf, C., & Gervain, M. (1995). Construction of a scale to measure consumer recovery. *Psychiatric Rehabilitation Journal, 18*(3), 131-139. https://db.arabpsychology.com/scales/recovery-assessment-scale-ras/

Gilbody, S. M., House, A. O., & Sheldon, T. A. (2002). Psychiatrists in the UK do not use outcome measures: National survey. *British Journal of Psychiatry, 180*(2), 101–103. https://doi.org/10.1192/bjp.180.2.101

Golino, H., Christensen, C., Moulder, R., Garrido, L. E., Jamison, L., & Shi, D. (2024). EGAnet: Exploratory Graph Analysis – A framework for estimating the number of dimensions in multivariate data using network psychometrics (Version 2.0.8) [Computer software]. https://doi.org/10.32614/CRAN.package.EGAnet

Guerrero, E., Barrios, M., Sampietro, H. M., Aza, A., Gómez-Benito, J., & Guilera, G. (2024). Let's talk about recovery in mental health: An international Delphi study of experts by experience. *Epidemiology and Psychiatric Sciences, 33*, e41:1–9. https://doi.org/10.1017/S2045796024000490

Guerrero, E., Barrios, M., Sampietro, H. M., Aza, A., Guilera, G., & Gómez-Benito, J. (2025). Recovery in mental health: An international Delphi study from a recovery-oriented professional perspective. *Social Science & Medicine, 381*, Article 118302. https://doi.org/10.1016/j.socscimed.2025.118302

Hasson-Ohayon, I., Roe, D., & Kravetz, S. (2008). The psychometric properties of the Illness Management and Recovery Scale: Client and clinician versions. *Psychiatry Research, 160*(2), 228–235. https://doi.org/10.1016/j.psychres.2007.06.013

Hernández, A., Hidalgo Montesinos, M. D., Hambleton, R. K., & Gómez Benito, J. (2020). International Test Commission guidelines for test adaptation: A criterion checklist. *Psicothema, 32*(3), 390–398. https://doi.org/10.7334/psicothema2019.306

Høgh Egmose, C., Heinsvig Poulsen, C., Hjorthøj, C., Skriver Mundy, S., Hellström, L., Nørgaard Nielsen, M., ... & Falgaard Eplov, L. (2023). The effectiveness of peer support in personal and clinical recovery: Systematic review and meta-analysis. *Psychiatric Services, 74*(8), 847–858. https://doi.org/10.1176/appi.ps.202100138

Holubova, M., Prasko, J., Hruby, R., Kamaradova, D., Ociskova, M., Latalova, K., & Grambal, A. (2015). Coping strategies and quality of life in schizophrenia: Cross-sectional study. *Neuropsychiatric Disease and Treatment, 11*, 3041–3048. https://doi.org/10.2147/NDT.S96559

Insel, T. R., & Quirion, R. (2005). Psychiatry as a clinical neuroscience discipline. *JAMA, 294*(17), 2221–2224. https://doi.org/10.1001/jama.294.17.2221

Jensen, S. B., Dalum, H. S., Korsbek, L., Hjorthøj, C., Mikkelsen, J. H., Thomsen, K., ... & Eplov, L. F. (2019). Illness management and recovery: One-year follow-up of a randomized controlled trial in Danish community mental health centers: Long-term effects on clinical and personal recovery. *BMC Psychiatry, 19*, Article 204. https://doi.org/10.1186/s12888-019-2048-0

Johansson, D., Skillmark, M., & Allgurin, M. (2025). Effects of psychosocial interventions on wellbeing in individuals with severe mental illness: A

systematic review. *Frontiers in Psychology, 16*, Article 1574303. https://doi.org/10.3389/fpsyg.2025.1574303

Johnstone, L., & Boyle, M. (2023). *Una introducción directa al Marco de Poder, Amenaza y Significado: Una alternativa al diagnóstico psiquiátrico.* Asociación Española de Neuropsiquiatría.

Lee, S. N., & Yu, H. J. (2024). Effectiveness of peer support programs for severe mental illness: A systematic review and meta-analysis. *Healthcare (Basel), 12*(12), Article 1179. https://doi.org/10.3390/healthcare12121179

Marder, S., & Galderisi, S. (2017). The current conceptualization of negative symptoms in schizophrenia. *World Psychiatry, 16*, 14–24. https://doi.org/10.1192/j.eurpsy.2022.12

Martín Ordiales, N., Hidalgo, M. D., Martín Chaparro, M. P., Ballester Plané, J., & Barrios, M. (2024). Assessing the psychometric properties of the Illness Management and Recovery Scale: A systematic review using the Consensus-based Standards for the selection of health Measurement Instruments (COSMIN). *Behavioral Sciences, 14*(4), Article 340. https://doi.org/10.3390/bs14040340

Martín-Ordiales, N., Hidalgo, M.D., Martín, M.P., & Barrios, M. (2026). The Illness Management and Recovery Scale: Adaptation and validation study of the Spanish version. *Evaluation & the Health Professions*. Online first https://doi.org/10.1177/01632787261441894

McGuire, A. B., Kukla, M., Green, A., & Salyers, M. P. (2014). Illness management and recovery: A review of the literature. *Psychiatric Services, 65*(2), 171–179. https://doi.org/10.1176/appi.ps.201200274

Mitikhin, V., Solokhina, T., Burygina, L., & Alieva, L. (2024). The nurses' assessment of the psychiatric care quality and the development of measures to improve it. *European Psychiatry, 67*(S1), S283–S284. https://doi.org/10.1192/j.eurpsy.2024.592

Mueser, K. T., Gingerich, S., Salyers, M. P., McGuire, A. B., Reyes, R. U., & Cunningham, H. (2005). Illness Management and Recovery (IMR) scales. In T. Campbell-Orde, J. Chamberlin, J. Carpenter, & H. S. Leff (Eds.), *Measuring the promise: A compendium of recovery measures* (Vol. II, pp. 32–35). Evaluation Center and Human Services Research Institute. https://facesandvoicesofrecovery.org/wp-content/uploads/Documents/Measuring-the-Promise-A-Compendium-of-Recovery-Measures-Volume-II.pdf

O'Connell, M., Tondora, J., Croog, G., Evans, A., & Davidson, L. (2005). From rhetoric to routine: Assessing perceptions of recovery-oriented practices in a state mental health and addiction system. *Psychiatric Rehabilitation Journal, 28*(4), 378–386. https://doi.org/10.2975/28.2005.378.386

Polat, S., Kutlu, Y., & Gültekin, B. (2020). Psychometric properties of the Turkish version of Illness Management and Recovery Scale–Patient form. *Perspectives in Psychiatric Care, 57*(1), 279–286. https://doi.org/10.1111/ppc.12559

Previde Massara, G., Di Matteo, T., & Aste, T. (2017). Network filtering for big data: Triangulated maximally filtered graph. *Journal of Complex Networks, 5*(2), 161–178. https://doi.org/10.48550/arXiv.1505.02445

Read, J., van Os, J., Morrison, A. P., & Ross, C. A. (2005). Childhood trauma, psychosis and schizophrenia: A literature review with theoretical and clinical implications. *Acta Psychiatrica Scandinavica, 112*(5), 330–350. https://doi.org/10.1111/j.1600-0447.2005.00634.x

Roefs, A., Fried, E. I., Kindt, M., Martijn, C., Elzinga, B., Evers, A., Wiers, R. W., Borsboom, D., & Jansen, A. (2022). A new science of mental disorders: Using personalised, transdiagnostic, dynamical systems to understand, model, diagnose and treat psychopathology. *Behaviour Research and Therapy, 153*, Article 104096. https://doi.org/10.1016/j.brat.2022.104096

Salyers, M. P., Fukui, S., Gearhart, T., Noll, J. P., & Williams, J. (2014). Clinical outcomes of the Illness Management and Recovery program in community mental health centers. *Psychiatric Services, 65*(8), 1101–1107. https://doi.org/10.1176/appi.ps.201300250

Sampietro, H. M., Barrios, M., Berrío, Á. I., Rojo, J. E., Guilera, G., & Gómez-Benito, J. (2024). *Evidence of validity and reliability for the Spanish version of the self-identified stage of recovery*. *Psicothema, 36*(2), 165–173. https://doi.org/10.7334/psicothema2023.213

Sampietro, H. M., Carmona, V. R., Sicilia Matas, L., Gavaldà Castet, C., Rojo, J. E., & Gómez Benito, J. (2022). Qué significa la recuperación para las personas usuarias de servicios de salud mental en Cataluña. *Quaderns De Psicologia*, *24*(2), Article e1880. https://doi.org/10.5565/rev/qpsicologia.1880

Sklar, M., Sarkin, A., Gilmer, T., & Groessl, E. (2012). The psychometric properties of the Illness Management and Recovery scale in a large American public mental health system. *Psychiatry Research, 199*(3), 220–227. https://doi.org/10.1016/j.psychres.2012.03.013

Slade, M. (2009). *Personal recovery and mental illness: A guide for mental health professionals.* Cambridge University Press. https://doi.org/10.1017/CBO9780511581649

Tan, C. H. S., Ishak, R. B., Lim, T. X. G., Marimuthusamy, P., Kaurss, K., & Leong, J. Y. J. (2017). Illness management and recovery program for mental health problems: Reducing symptoms and increasing social functioning. *Journal of Clinical Nursing, 26*, (21-22), 3471–3485. https://doi.org/10.1111/jocn.13712

Tapia, A. A., Soto, G. I., García, A. A., Pavez, F. A., & Alvarado, R. (2021). Effects of adherence to pharmacological treatment on the recovery of patients with schizophrenia. *BMC Psychiatry, 21*, Article 487. https://doi.org/10.1186/s12888-021-03547-2

Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis.* Springer-Verlag. https://doi.org/10.1007/978-3-319-24277-4_9

World Health Organization. (2022). *ICD-11: International classification of diseases* (11th rev.). WHO. https://icd.who.int/en/

Young, S. L., Ensing, D. S., & Egeland, B. (2003). Development of the mental health recovery measure. *Psychiatric Rehabilitation Journal, 27*(3), 231–234. https://doi.org/10.2975/27.2003.231.234

GRUPO EDITORIAL DE PSICOFUNDACIÓN

Psicothema (2026) 38(3) 161-172

Psicothema

https://www.psicothema.com/es • eISSN 1886-144X

Colegio Oficial de Psicología del Principado de Asturias

Psicothema

Article

Psychometric Properties of the Montreal Cognitive Assessment (MoCA) in Spanish Adults

David Arjol[1], Sara Mora-Simón[1,2], Emiliano Rodríguez-Sánchez[1,3], Manuel A. Gómez-Marcos[1,3], Luis García-Ortiz[1,3], José A. Maderuelo-Fernández[1,3], Cristina Lugones-Sánchez[1,3] and Jaime Unzueta-Arce[1,2]

[1] Instituto de Investigación Biomédica de Salamanca (Spain)
[2] Universidad de Salamanca (Spain)
[3] Gerencia de Atención Primaria de Salamanca, Gerencia Regional de Salud de Castilla y León (Spain)

ARTICLE INFO

Received: 04/12/2025
Accepted: 16/02/2026

Keywords:
MoCA
Construct validity
MIRT
Cognitive screening
Psychometrics

ABSTRACT

Background: The Montreal Cognitive Assessment (MoCA) is widely used for cognitive screening, but its dimensionality and domain-specific accuracy require validation with methods for ordinal items. The objective was to analyze the internal structure of the MoCA in a cognitive impairment-free Spanish community sample and examine its relationship with variables potentially related to the risk of cognitive impairment. **Method:** A total of 1,767 Spanish adults (aged 18-89 years) were assessed in primary care centers using the MoCA, and sociodemographic and health variables were recorded. Internal structure and metric properties were analyzed using factor models and item response theory, as well as their association with external variables and the concordance of the resulting solution. **Results:** The theoretical models tested did not demonstrate adequate fit. Cross-validation identified a 12-item solution in two correlated factors (Memory and Multiple), with excellent fit (CFI = .991; RMSEA = .019) and latent correlation r = .63. The abbreviated version showed high concordance with the total score (ICC = .96) and substantial agreement in dichotomous classification (k = .84). **Conclusions:** The MoCA exhibited a two-factor structure and consistent associations with external variables such as age, educational level, and vascular risk; a dimensional interpretation is recommended.

Propiedades Psicométricas del Montreal Cognitive Assessment (MoCA) en Adultos Españoles

Palabras clave:
Validez de constructo
MIRT
Rastreo cognitivo
Psicometría

RESUMEN

Antecedentes: El Montreal Cognitive Assessment (MoCA) se emplea ampliamente para el rastreo cognitivo, pero su dimensionalidad y precisión por dominio requieren validación con métodos para ítems ordinales. El objetivo fue analizar la estructura interna del MoCA en una muestra comunitaria española sin deterioro cognitivo (DC) y examinar su relación con variables potencialmente relacionadas con el riesgo de DC. **Método:** Se evaluó a 1,767 adultos (18-89 años) españoles procedentes de centros de atención primaria mediante el MoCA y se registraron variables sociodemográficas y de salud. Se analizaron la estructura interna y propiedades métricas mediante modelos factoriales y teoría de respuesta al ítem, así como su asociación con variables externas y la concordancia de la solución resultante. **Resultados:** Los modelos teóricos contrastados no mostraron ajuste adecuado. La validación cruzada identificó una solución de 12 ítems en dos factores correlacionados (Mnésico y Múltiple), con ajuste excelente (CFI = .991; RMSEA = .019) y correlación latente r = .63. La versión abreviada mostró alta concordancia con la puntuación total (CCI = .96) y acuerdo sustancial en la clasificación dicotómica (k = .84). **Conclusiones:** El MoCA mostró una estructura de dos factores y asociaciones coherentes con variables externas como la edad, el nivel educativo y el riesgo vascular; se recomienda priorizar una interpretación dimensional.

Cite as: Arjol, D., Mora-Simón, S., Rodríguez-Sánchez, E., Gómez-Marcos, M. A., García-Ortiz, L., Maderuelo-Fernández, J. A., Lugones-Sánchez, C., & Unzueta-Arce, J. (2026). Psychometric properties of the Montreal Cognitive Assessment (MoCA) in Spanish adults. *Psicothema, 38*(3), 161-172. https://doi.org/10.70478/psicothema.2026.38.15
Corresponding author: David Arjol, darjol@ibsal.es

Cognitive screening instruments enable quick, cost-effective assessments in the general population and facilitate referral to more comprehensive neuropsychological evaluations when appropriate (Chithiramohan et al., 2024). Their usefulness, however, depends on robust psychometric evidence to support structural validity and proper item functioning (Mokkink et al., 2024; Wen et al., 2025).

Since its introduction, and especially in the last decade, the Montreal Cognitive Assessment (MoCA; Nasreddine et al., 2005) has become established as a widely used test in screening for mild cognitive impairment (MCI) (O'Caoimh et al., 2025; Wen et al., 2025). Recent meta-analyses support its diagnostic accuracy (with cut-off points between < 23 and < 25 maximizing the sensitivity-specificity balance) and suggest greater sensitivity than the Mini-Mental State Examination (MMSE) in detecting MCI (Ciesielska et al., 2016; Islam et al., 2023). It is organized into seven domains, and while this guides practice, it does not in itself guarantee dimensional validity and requires empirical verification with appropriate methods (Sala et al., 2020; Wen et al., 2025). Although the MoCA was originally designed to sample multiple cognitive domains, it remains unclear whether its items primarily reflect a single general cognitive factor, multiple correlated dimensions, or a hierarchical structure, as prior research has reported inconsistent latent solutions across populations and analytic approaches (Sergi et al., 2025). Structural validity studies in community populations in Portugal (Duro et al., 2010; Freitas et al., 2012, 2015), Ireland (Coen et al., 2016), Chile (Cancino et al., 2020) and China (Luo et al., 2022) have, for example, reported divergent solutions ranging from six-factor models and hierarchical structures to more parsimonious models, thereby reinforcing the need to verify dimensionality in each population. Most of these studies have approached the question from the perspective of classical test theory (CTT) using confirmatory factor analysis (CFA) within the framework of structural equation modelling (SEM) (e.g., Coen et al., 2016; Duro et al., 2010; Freitas et al., 2012). While this is a valuable approach, it does not provide information on item-by-item performance or conditional accuracy across the latent trait (θ). At the same time, studies based on item response theory (IRT) have characterized MoCA at the item level, quantifying discrimination and information across θ and, therefore, conditional accuracy (Embretson & Reise, 2013). Rasch analysis (partial credit model [PCM]) and IRT models for ordinal items (graded response model [GRM]) present information profiles differentiated by domain: memory items and some executive function items provide greater discrimination and information, while orientation and naming items show a ceiling effect, supporting the use of multidimensional approaches (Aiello et al., 2022; Freitas et al., 2015; Tsai et al., 2012). A recent review reports substantial cross-population heterogeneity in MoCA factor solutions and item functioning (Sergi et al., 2025). Accordingly, its dimensional structure should be empirically verified in each population in which it is applied.

In Spain, the MoCA has shown adequate reliability (Pereiro et al., 2017), high test-retest stability (Lozano Gallego et al., 2009) and good-to-excellent inter- and intra-rater agreement (Porto, 2015), as well as satisfactory convergent validity (Pereiro et al, 2015, 2017; Vallés-Salgado et al., 2024) and criterion validity (Lozano Gallego et al., 2009; Matías-Guiu et al., 2017; Ramírez-Moreno et al., 2022; Vallés-Salgado et al., 2024), including community-based normative data (Ojeda et al., 2016; Pereiro et al., 2017) and recent clinical evidence on diagnostic performance in memory-clinic settings (Vallés-Salgado et al., 2024). However, evidence remains comparatively limited for structural validity in large community samples spanning a broad adult age range, particularly using methods appropriate for ordinal items and complemented with item-level modeling to document conditional measurement precision (Bugallo-Carrera et al., 2024; Vallés-Salgado et al., 2024). Consistent with best-practice test-validation frameworks, defensible score use requires evidence of a replicable dimensional structure and appropriate item functioning across key subgroups; thus, beyond Spanish studies focused on norms and diagnostic accuracy, the present study examines the MoCA's latent structure and conditional precision using modern psychometric approaches in a large, cognitive-impairment–free community dwelling sample (Badham et al., 2025; Ferrando et al., 2025). Likewise, evidence remains scarce on the relationship between MoCA and vascular risk and lifestyle biomarkers with prognostic relevance (e.g., composite indices such as CAIDE or ANU-ADRI), associations which have been found in other populations (Guo et al., 2024; Yin et al., 2024). This evidence is directly relevant for community screening and referral, where decisions depend on measurement precision in the range where early impairment is expected.

In view of the above, this study aimed to validate the MoCA in a large cognitive impairment-free Spanish community dwelling sample (aged 18-89 years) by assessing: (a) its internal structure; (b) its measurement properties by factor; and (c) its relationship with vascular risk indicators and lifestyle, in a framework integrating CTT and IRT.

Method

Participants

The sample comprised 1,767 adults aged 18-89 years (M = 51.5, SD = 16.2), of whom 58.5% were women. Participants were drawn from four independent community-based studies (n = 539, 223, 626, and 379) conducted in primary care settings in Castilla y León (Spain). Inclusion criteria were being 18 years of age or older, having Spanish as a native language, agreeing to participate voluntarily, and providing informed consent. Exclusion criteria included a history or diagnosis of neurological conditions with potential impact on cognition (e.g., dementia, MCI), stroke or traumatic brain injury, systemic diseases with cognitive impact (e.g., diabetes), significant sensory or motor limitations, illiteracy, incomplete MoCA, or prior professional familiarity with MoCA. Detailed sample characteristics are presented in Table 1.

Instruments

Montreal Cognitive Assessment (MoCA)

The MoCA is a cognitive screening test with a maximum score of 30 points that assesses seven domains: visuospatial/executive functions, naming, memory, attention, language, abstraction, and orientation. We administered the official Spanish version 8.1 of the MoCA (Nasreddine et al., 2005). This version has been previously validated linguistically (Acquadro et al., 2012). For an expanded description of the items, domains and score ranges, see Table S1.

Table 1
Demographic Characteristics of the Sample (N = 1,767)

Variable	Levels	Summary
Sociodemographic variables		
Age (years)		54 [39-65]
Sex	Women / Men	1,034 (58.5%) / 733 (41.5%)
Education level	Primary / Secondary / University	290 (16.4%) / 681 (38.5%) / 796 (45.0%)
Employment	Employed/ Unemployed	1,391 (78.7%) / 376 (21.3%)
Occupational field (CNO-11)	Natural Sciences / Art and Culture / Health / Social Sciences and Education / Processing and Manufacturing / Primary Resources / Finance and Administration / Transportation / Sales and Services	465 (33.4%) / 299 (21.5%) / 176 (12.7%) / 165 (11.9%) / 114 (8.2%) / 83 (6.0%) / 56 (4.0%) / 29 (2.1%) / 4 (0.3%)
Level of professional qualifications	A / B / C / D	83 (6.0%) / 452 (32.5%) / 375 (27.0%) / 481 (34.6%)
Clinical variables		
SBP / DBP (mmHg)		124 [112-137] / 78 [71-86]
Hypertension	Yes / No	594 (33.6%) / 1,173 (66.4%)
BMI	Underweight / Normal weight / Overweight / Obesity	29 (1.6%) / 749 (42.4%) / 670 (37.9%) / 319 (18.1%)
BMI (kg/m²)		25.7 [22.9-28.8]
Total cholesterol (mg/dL)		193 [170-218]
HDL / LDL (mg/dL)		57 [49-69] / 114 [94-135]
Triglycerides (mg/dL)		88 [67-120]
Glucose (mg/dL)		86 [80-93]
Antidepressants / Anxiolytics	Use	100 (5.7%) / 122 (6.9%)
Lifestyle		
Alcohol	Non-drinker / Low-risk drinker / At-risk drinker	564 (31.9%) / 945 (53.5%) / 258 (14.6%)
Alcohol (SDU/wk)		2.2 [0.0-6.3]
Physical activity	Low/ Moderate / High	242 (13.7%) / 929 (52.6%) / 596 (33.7%)
Physical activity (MET-min/wk)		2,970 [1,485-4,348]
Variable	**Levels**	**Summary**
Smoking	never smoked / ex-smoker > 5 years / ex-smoker 1-5 years / ex-smoker 0-1 year / yes, regularly	857 (48.5%) / 512 (29.0%) / 75 (4.2%) / 32 (1.8%) / 291 (16.5%)
PYI		13.6 [5.5-28.9]
CAIDE Index		
CAIDE index	Very low risk / Low risk / Moderate risk / High risk / Very high risk	1,073 (60.7%) / 389 (22.0%) / 184 (10.4%) / 114 (6.5%) / 7 (0.4%)

Note. Continuous variables are reported as median [*IQR*]; categorical variables as *n* (%). Rows for Occupational field (CNO-11) and Professional qualification levels were computed within the Employed subgroup (*n* = 1,391). Alcohol categories follow Spain's Ministry of Health thresholds (≤1 SDU/day for women; ≤2 SDU/day for men; reported here as SDU/wk). CNO-11 = Clasificación Nacional de Ocupaciones 2011 [National Classification of Occupations 2011 (CNO-11)]; SBP: systolic blood pressure; DBP: diastolic blood pressure; BMI = body mass index; HDL = high-density lipoprotein; LDL = low-density lipoprotein; MET = metabolic equivalent of task; PYI = pack-year index; CAIDE = Cardiovascular Risk Factors, Aging and Dementia.

Sociodemographic Variables

We recorded age, sex (assigned at birth), educational level, coded according to the Spanish system as illiterate, primary education (≤ 12 years of schooling), secondary education (13–15 years) or university studies (≥ 16 years), employment status, occupational field according to the Clasificación Nacional de Ocupaciones de España (Spanish National Classification of Occupations) (CNO-11) and level of professional qualification (categories A–D) according to the criteria of the Instituto Nacional de Estadística (National Institute of Statistics [INE]) (2011).

Clinical Variables

The use of antidepressants and/or anxiolytics was obtained by interview and verification in medical records. Blood pressure was measured with an OMRON M10-IT sphygmomanometer (Omron Healthcare, Kyoto, Japan) by taking three readings and averaging the last two, on the dominant arm, with the participant seated and after ≥ 5 minutes of rest, according to the European Society of Hypertension recommendations (Williams et al., 2018). Hypertension was defined as systolic blood pressure ≥ 140 mmHg or diastolic blood pressure ≥ 90 mmHg, a prior diagnosis, or antihypertensive treatment.

Body mass index (BMI = kg/m²) was calculated and classified, following WHO criteria, as obese (BMI ≥ 30 kg/m²), overweight (25-29.9 kg/m²), and normal weight (< 25 kg/m²). Fasting blood tests were carried out (total cholesterol, HDL, triglycerides, and glucose), and LDL was estimated with the Friedewald formula.

Lifestyle

Alcohol use was recorded through a structured questionnaire about the quantity and type of alcoholic drinks consumed in the previous seven days, expressed as standard drink units (SDUs) per week (1 SDU = 10 g of pure alcohol). Participants were classified in three categories, following the recommendation of the Spanish Ministry of Health (2020): non-drinker, low-risk drinker, and at-risk drinker. Physical activity was assessed using the short version of the IPAQ (Craig et al., 2003), deriving MET-min/week and the categories low/moderate/high. Smoking status was recorded to calculate the pack-year index (PKI).

CAIDE Index

The cardiovascular risk factors, aging and incidence of dementia (CAIDE) score was estimated as the weighted sum of age, sex, educational level, blood pressure, total cholesterol, BMI, and physical activity. The range is 0-15, and an ordinal classification with five levels—very low risk (0-5), low (6-7), moderate (8-9), high (10-11), and very high (12-15)—was used given its greater informative capacity and clinical utility (Kivipelto et al., 2006; Sindi et al., 2015).

Procedure

A multi-sample, instrumental study based on secondary data analysis was carried out. Data were drawn from four independent primary care–based studies implemented in community primary care settings in Castilla y León (Spain), with in-person assessments conducted under a shared standardized fieldwork framework. Trained professionals administered the MoCA and collected the remaining measures, and data quality was monitored throughout the fieldwork. Data collection occurred between October 2021 and September 2024. The databases of the four studies were integrated into an initial repository (N = 2,255) using variable harmonization and pre-specified cleaning rules. The studies initially contributed n = 610 (Martí-Lluch et al., 2023), 273 (Llamas-Ramos et al., 2024), 965 (Lugones-Sánchez et al., 2023), and 407 (Vicente-Gabriel et al., 2024) participants, respectively. For the present secondary analysis, range, consistency, and plausibility checks were applied, and a fully anonymized analytical dataset including only complete records for the variables used in the analyses was created prior to psychometric modeling. All studies were approved by the corresponding Ethics Committee and complied with the Declaration of Helsinki (World Medical Association, 2025). After cleaning the integrated database, 299 participants were excluded with type 2 diabetes, 23 with stroke prior to the MoCA, one with a neurological disorder, and 165 due to incomplete MoCA, producing a final sample of N = 1,767.

Data Analysis

To characterize the sample, we calculated descriptive statistics, with absolute frequencies and percentages for categorical variables and median and interquartile range (*IQR*) for continuous variables, as indicated by the Shapiro–Wilk test and inspection of histograms and Q–Q plots. Floor and ceiling effects were analyzed for the total MoCA score and for each resulting subscale (defined a priori as ≥20% of observations at the minimum or maximum value, respectively). Internal consistency was estimated using Cronbach's alpha (α) and McDonald's omega (ω) calculated on polychoric matrices; ω was included as a reliability estimate that is less dependent on restrictive assumptions for ordinal items. Hierarchical omega (ωh) was reported only when multidimensional or bifactor representations were evaluated.

Validation of the internal structure was performed in several phases. First, CFA was performed on the entire sample, with five theoretical models contrasted: 1) a unidimensional baseline model; 2) a two-factor correlated model (Verbal Memory and Executive Attention), as proposed by Duro et al. (2010); 3) a six-factor correlated model representing the theoretical domains of the test, following Freitas et al. (2012); 4) a second-order hierarchical model in which a global cognition factor accounts for covariance among first-order domains (Kline, 2023); and 5) a bifactor model with a general factor and domain-specific group factors, included as a diagnostic specification consistent with prior dimensionality evaluations in MoCA research (Luo et al., 2019) and broader structural syntheses (Sergi et al., 2025). The bifactor model was tested to quantify the proportion of common variance attributable to a general factor relative to specific factors and to inform interpretation of the total score and the incremental reliability of subscale scores (Kline, 2023; Luo et al., 2019). All CFAs were estimated for ordinal items using weighted least squares mean and variance adjusted (WLSMV, theta parameterization, χ^2 scaled). The fit was assessed using the comparative fit index (CFI), the Tucker-Lewis index (TLI), the root mean square error of approximation (RMSEA, 90% CI) and the standardized root mean square residual (SRMR) according to a priori criteria (CFI/TLI ≥ .95; RMSEA ≤ .06; SRMR ≤ .08) (Hu & Bentler, 1999). For bifactor specifications, additional diagnostics were computed, including explained common variance (ECV) and bifactor reliability indices, and local dependence was evaluated using *Q3* statistics.

In addition, we used an exploratory approach with cross-validation. The sample was randomly divided into two equivalent subsamples (exploratory, n = 884; confirmatory, n = 883) using simple random assignment with a fixed seed. Comparability between subsamples was examined with homogeneity checks for age, sex, and educational level (Table S2). In the exploratory subsample, an exploratory factor analysis (EFA) was applied to a polychoric matrix using ULS (minres) and oblimin rotation. The number of factors was determined by parallel analysis (500 replications), and KMO and Bartlett's test of sphericity were reported. Item retention followed prespecified psychometric criteria aimed at obtaining a stable and interpretable solution. Items were retained when they showed salient primary loadings on a single factor and adequate communalities. Items were considered for exclusion when they showed extreme floor or ceiling effects with limited variability, very low communalities, or weak and unstable contributions to the common factor structure. The resulting structure was validated in the confirmatory subsample using CFA, where its fit was compared to an alternative bifactor model.

Following factorial validation, multidimensional item response theory (MIRT) models were fitted, including a bifactor (general and specific factors) to evaluate essential unidimensionality and potential local dependence, and a two-factor correlated model.

Agreement between the full score and the abbreviated version was estimated using the intraclass correlation coefficient (ICC, 95% CI). Dichotomous agreement (normal performance vs. possible cognitive impairment) was assessed using Cohen's κ (95% CI). For the full MoCA, classification was based on the conventional cutoff of a total score < 26 originally proposed by Nasreddine et al. (2005) and widely used in cognitive screening research (e.g., Islam et al., 2023; Tsoi et al., 2017). An equivalent cutoff for the 12-item form was derived using equipercentile linking between the full and short-form scores: the percentile corresponding to the < 26 classification in the full-score distribution was identified, and the short-form integer threshold matching that percentile was selected. The linking function was logit-smoothed, and uncertainty was quantified via bootstrap resampling to obtain confidence intervals. Finally, associations between MIRT-derived latent scores and

sociodemographic, clinical, and lifestyle variables were examined using Spearman correlations. Additionally, a MIMIC-type structural equation model was specified for the two-factor CFA, in which predictors regressed directly on the latent factors. This model was estimated using robust maximum likelihood (MLR), and standardized effects were reported. As an additional analysis, we fitted regressions on the latent scores obtained through EAP.

All analyses were performed in R 4.3.1 (R Core Team, 2023) using *lavaan* (Rossel, 2012) and *semTools* (Jorgensen et al., 2016) for models with ordinal items and *mirt* (Chalmers, 2012) for IRT analyses. The significance level was set at α = .05 (two-tailed). Reproducible code and analysis scripts are available.

Results

Evidence of Validity Based on Internal Structure

None of the five theoretical models tested using CFA met the full set of prespecified fit criteria (CFI/TLI ≥ .95; RMSEA ≤ .06; SRMR ≤ .08). Complete fit indices are provided in Table S3. Parallel analysis suggested up to four factors; however, eigenvalues attenuated sharply after the second factor (F1 = 3.771; F2 = 0.929; F3 = 0.265; F4 = 0.078). In cross-validation, exploratory factor analysis (EFA) of the exploratory subsample (n = 884) identified a 12-item solution grouped into two correlated factors. The first factor, Memory, consists of the five delayed recall items, and the second, the Multiple factor, comprises following items: the category switch, cube copying, clock drawing, naming lion and rhinoceros, subtraction, and train-bicycle similarity. Accordingly, the two-factor solution was retained and explained 41.1% of the common variance (22.5% and 18.6% per factor) (see Tables S4–S6). This structure showed excellent fit in the confirmatory subsample (n = 883): CFI = .991; TLI = .988; RMSEA = .019 [90% CI; .012, .025]; SRMR = .046. A bifactor (general-specific) model did not provide a substantial improvement over the two-factor model (ΔCFI = .005; ΔRMSEA = .005), so the latter was retained for parsimony. Additional bifactor diagnostics suggested only moderate general-factor saturation (ECV = .58; ωh = .54) and limited specific-factor reliability (ωs Memory = .27; ωs Multiple = .39), with no evidence of meaningful local dependence (maximum positive $Q3$ = .029; 0 item pairs with $Q3$ > .20) (see Table S7). The retained structure explained 39.9% of the common variance (item R^2 values are presented in Table S8).

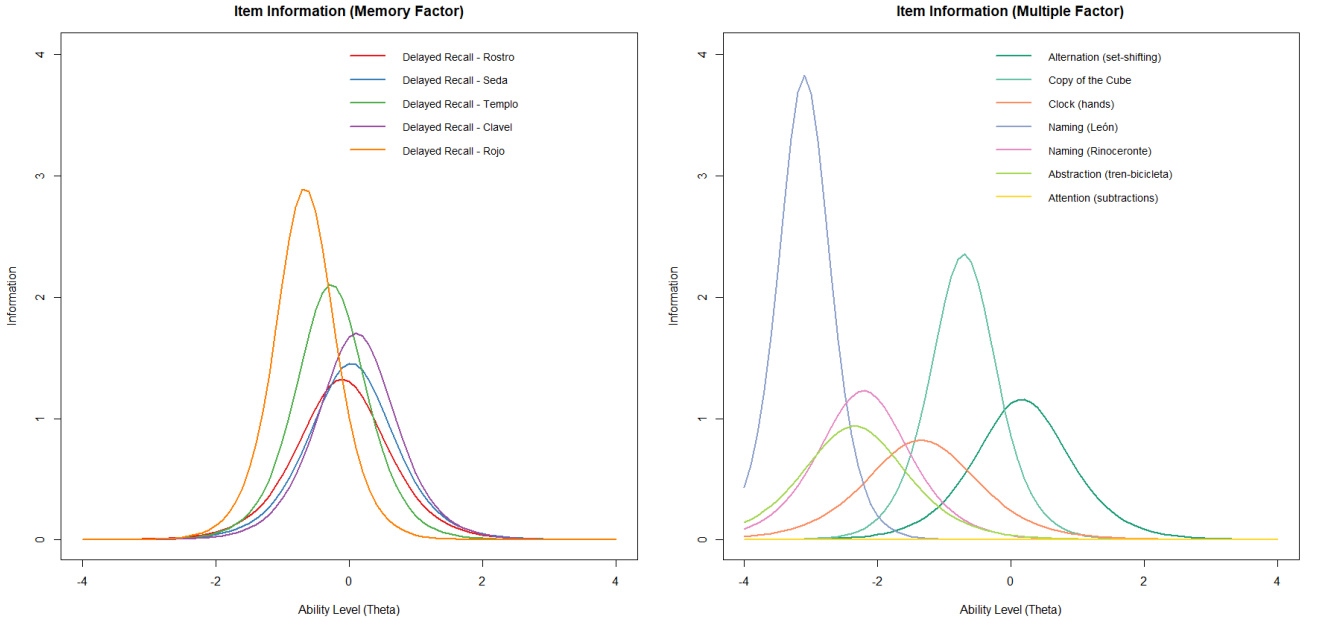
MIRT analyses of the general–specific (bifactor) specification showed estimation anomalies and lack of stable convergence, so this specification was not interpreted further. Instead, the two-correlated-factor model converged to a stable and consistent solution ($M2(_{51})$ = 92.45, p < .001; RMSEA = 0.020 [90% CI; .013, .026]; SRMR = 0.023; CFI = 0.990; TLI = 0.987; AIC = 25431.12; and BIC = 25582.8), with information profiles differentiated by factor and latent correlation r = .625.

The S-X² item fit was acceptable, and there was no relevant local dependence (|$Q3$| > .20) (see Table S7). Discrimination scores ranged from 1.35 to 2.00 for the Memory factor and 1.07 to 2.30 for the Multiple factor (Table S9). The item information curves (IIC) and test information curves (TIC) (Figure 1) showed distinct accuracy profiles: the Multiple factor reached its maximum information (information = 2.606) at a low ability level (θ = -2.9, SEM = 0.619), with information decreasing as θ increased; while the Memory factor reached its maximum (information = 3.071) across a broader ability range (θ = 0.3, SEM = 0.571).

Table 2 summarizes the reliability analysis of the MoCA, its official domains, and final subscales. The mean MoCA score was 24.10 (SD = 3.77; skewness = -0.81). There were marked

Figure 1

Item Information Curves From the Two-Factor MIRT Model



Note: Item Information Curves from the two-factor MIRT model. Memory Factor items peak in the low–mid θ range; Multiple Factor items peak mostly below the mean and lose precision at higher θ.

ceiling effects across several theoretical domains according to the prespecified criterion, including Visuospatial/Executive (29.2%), Naming (90.8%), Attention (46.0%), Language (46.0%), Abstraction (67.8%), and Orientation (93.0%) (see Table S10). No domain-level floor effects reached the prespecified threshold (≥20% at the minimum score). In terms of internal consistency, most domains had $\omega < .60$, with the exception of Delayed Recall ($\omega = .683$). Internal consistency was: ω_total = .717; ω_hierarchical = .613. The Memory factor showed $\omega = .683$ (skewness = -0.26), while the Multiple factor had $\omega = .585$ (skewness = -1.10).

Evidence of Validity Based on Relationships With Other Variables

The Memory factor was moderately associated with age (ρ = -0.46), educational level (ρ = 0.40), and the CAIDE index (ρ = -0.47), and more weakly associated with hypertension (ρ = -0.29) and BMI (ρ = -0.24). The Multiple factor showed moderate correlations with age (ρ = -0.55), educational level (ρ = 0.48), and CAIDE (ρ = -0.48), as well as a significant association with sex (ρ = -0.09) and hypertension (ρ = -0.32). All coefficients reached $p < .001$. The remaining variables showed no or small associations with the latent scores and did not contribute incremental variance to the MIMIC models ($|\rho| < .10$ and/or $q > .05$).

In the adjusted MIMIC model, the explained variance was $R^2 = .378$ for the Memory factor and $R^2 = .545$ for the Multiple factor. The results indicated that age (β = -0.43), education (β = 0.26), sex (β = 0.16), and BMI (β = -0.06) were significant predictors of performance on the Memory factor, while the significant predictors for the Multiple factor were age (β = -0.45), education (β = 0.36), and sex (β = -0.23) (see Table S11).

Finally, the agreement between the total MoCA score and the 12-item reduced version for classifying subjects' cognitive status (normal score or suspected cognitive impairment) was assessed. For continuous scores, the ICC (3, k) was 0.96 (95% CI [.956, .961]). For dichotomous classification (normal performance vs. possible impairment), the short-form cut-off was <11 (Table S12). Categorical agreement was $\kappa = 0.84$ (95% CI [0.819, 0.860]), with 92.2% overall agreement (95% CI [91.1%, 93.1%]). In terms of classification, 764/804 subjects with normal performance and 971/1000 with possible impairment were correctly classified. We found 40 false negatives and 29 false positives.

Discussion

In a large cognitive impairment-free Spanish community dwelling sample (N = 1,767) spanning early adulthood to old age (18-89 years), the MoCA yielded a correlated two-factor solution (Memory and Multiple) comprising 12 items. This study extends previous Spanish validation efforts— which have provided valuable evidence on normative adjustments and screening cut-offs (e.g., Pereiro et al., 2017; Vallés-Salgado et al., 2024)—by providing large-sample evidence on internal structure and item-level measurement precision in a cognitive impairment–free community sample using ordinal modeling, internal cross-validation, and IRT information functions. This solution offers a more parsimonious and interpretable reading of cognitive performance than alternatives based on discrete domains (e.g., Coen et al., 2016; Duro et al., 2010). Given the heterogeneity in MoCA factor solutions and item functioning reported across populations, these results further support the need for population-specific verification and careful score interpretation based on locally supported measurement evidence (Badham et al., 2025; Sergi et al., 2025). The choice of a correlated two-factor model over a general–specific bifactor model is congruent with current recommendations to prioritize simple models with substantive validity and robust empirical performance (Kline, 2023; Marsh et al., 2004; Rodríguez et al., 2016).

From a neuropsychological and neuroscientific perspective, the two-factor solution can be explained by current models of brain and cognitive function based on neural networks (Dragomir & Omurtag, 2021). These models align with the concept of interconnected yet distinct dynamic networks, including the episodic memory network

Table 2
Descriptive and Reliability Analysis of MoCA Items (N = 1,767): Total Scale, Theoretical Domains and Final Subscales

Scale / Domain	No. of Items	Range	Mean (*SD*)	Floor (%)	Ceiling (%)	Cronbach's α	McDonald's ω	Skewness	Kurtosis
MoCA (Total Score)	28	0-30	24.1 (3.77)	0.0%	3.5%	.736	.717	-0.81	0.51
MoCA Domains									
Visuospatial/Executive	5	0-5	3.73 (1.2)	0.3%	29.2%	.450	.434	-0.70	-0.11
Naming	3	0-3	2.90 (0.4)	0.2%	90.8%	.570	.575	-3.87	17.42
Attention	4	0-6	5.04 (1.1)	0.1%	46.0%	.510	.387	-1.41	1.63
Language	3	0-3	2.26 (0.8)	2.8%	46.0%	.410	.344	-0.83	-0.06
Abstraction	2	0-2	1.63 (0.6)	5.1%	67.8%	.460	NE	-1.29	0.65
Delayed Recall	5	0-5	2.90 (1.6)	11.0%	19.2%	.710	.673	-0.37	-0.95
Orientation	6	0-6	5.92 (0.4)	0.0%	93.0%	.640	.271	-5.02	34.26
12-item version of MoCA (Total Score)	12	0-14	9.99 (2.76)	0.00%	8.7%	.703	0.712	-0.63	0.09
Final Subscales (FA Model)									
Memory Subscale	5	0-5	2.90 (1.60)	11.0%	19.2%	.673	.673	-0.37	-0.95
Multiple Subscale	7	0-9	7.22 (1.69)	0.0%	25.4%	.551	.585	-1.10	0.90

Note. Scale reliability is reported using Cronbach's alpha (α) and McDonald's omega (ω). Alpha is provided to facilitate comparability with prior MoCA studies, whereas omega (computed from polychoric correlations) is reported as a less assumption-dependent estimate for ordinal items. Total score values for the abbreviated version and for each factor are based on the sum of their component items. NE = not estimable (e.g., omega cannot be computed or is unstable for very short domains, such as two-item components); FA = factor analysis.

(Qu et al., 2025) and the prefrontal cortex network and its connectivity (Menon & D'Esposito, 2021). First, the Memory factor comprises the five delayed free recall items on the verbal episodic memory task. This set of items explores the ability to retrieve previously learned and consolidated verbal information as part of the memory process. Studies on dynamic cellular and molecular mechanisms involved in memory have identified specific neuronal engrams for the encoding, consolidation, and retrieval of information, primarily involving the hippocampus and the medial and ventromedial prefrontal cortex, among others (Tallman et al., 2024; Teng et al., 2025). While being a very specific factor, it is also a prominent one; it has been observed that worsening performance on free recall tasks can be considered a predictor of the preclinical and asymptomatic phase of Alzheimer's disease, and is even associated with plasma biomarkers (Aschenbrenner et al., 2024; Grober et al., 2022, 2023).

Secondly, the Multiple factor includes seven items exploring cognitive abilities that depend on multiple neural networks with a markedly integrative character. These include visuoconstructive and visuospatial skills, as well as the executive components of cognitive flexibility and planning (tested with the category switch, cube copying and clock drawing items), the ability to identify and name visual stimuli (lion and rhinoceros items), and the capacity for conceptual reasoning, semantic categorization, and cognitive flexibility (train-bicycle similarity item). Sustained attention, working memory, calculation, and executive control (serial subtraction items) are skills whose performance has been linked in prior neurocognitive models to distributed cortico-subcortical systems and frontoparietal control resources, rather than to a single localized network (Dragomir & Omurtag, 2021; Hobeika et al., 2016; Imai et al., 2022). All of these networks support the functioning involved in various cognitive skills that require the integration of cortical and subcortical information and share a common frontal executive component. Lesions in any area or connection involved in these networks may result in poorer performance on this second factor.

Regarding validity evidence on the basis of relationships with other variables, the associations with age and educational level were moderate and in the expected direction, supporting construct validity (Aiello et al., 2022; Pereiro et al., 2017; Sachs et al., 2022). Additionally, the associations with cardiometabolic indicators—hypertension and BMI—and with the CAIDE index suggest that MoCA scores and their factors can capture gradients of vascular and lifestyle risk with prognostic relevance (Lentoor & Myburgh, 2022). This pattern is in line with previous evidence documenting how MoCA performance worsens along the cardiovascular disease continuum (Gagnon et al., 2022) and with reviews linking cardiovascular risk indices to cognitive impairment and dementia (Jia et al., 2023). Furthermore, associations between risk scores (e.g., CAIDE/ANU-ADRI), tau pathology, and cognition have been described in community cohorts (Guo et al., 2024; Yin et al., 2024). Taken together, these findings offer validity evidence based on relationships with other score variables versus clinical markers and composite indices, although caution is required when interpreting the above, given the lack of widespread validation for this index in Spain (Geethadevi et al., 2023; Kivipelto et al., 2006; Nicotra et al., 2024). Accordingly, these associations should be interpreted as supportive correlational evidence in a community screening context, rather than as diagnostic or causal evidence.

From a psychometric perspective, the findings are consistent with studies that question domain-based dimensionality and support more parsimonious models for their stability and interpretability (Coen et al., 2016; Duro et al., 2010; Roalf et al., 2016). Consistent with this literature, none of the prespecified theoretical CFA models achieved acceptable fit. This likely reflects structural features of the MoCA in community samples: brief ordinal items with heterogeneous content and pronounced ceiling effects can attenuate the inter-item covariance needed to support complex, domain-based or hierarchical specifications. Consequently, empirically derived, parsimonious structures may show greater stability than theoretically imposed ones, in line with the cross-cultural heterogeneity reported in recent structural reviews (Sergi et al., 2025). Regarding reliability, the moderate internal consistency observed—particularly for the Multiple factor—is expected given the brevity of the subscales, the heterogeneity of the cognitive processes involved, and the marked skewness typical of non-clinical samples. In this context, classical internal consistency indices may underestimate measurement quality; therefore, interpretation benefits from a dimensional approach supported by item response theory and conditional measurement precision. Our findings align with several psychometric studies in which a relatively distinct memory factor emerges alongside factors capturing other cognitive functions of varying complexity (Cancino et al., 2020; Coen et al., 2016; Duro et al., 2010; Freitas et al., 2012, 2015). In Portugal, two well-fitting factors have been described: a memory factor and another that integrates visuospatial, attentional, and executive components (Duro et al., 2010; Freitas et al., 2012, 2015). This pattern is similar to the findings of our study, although the exact composition of the two resulting factors is not identical. Importantly, our structure was derived at the item level, whereas Duro et al. (2010) analyzed aggregated subtests; thus, our Memory factor is strictly operationalized by delayed recall, and the Multiple factor aggregates heterogeneous non-mnemonic tasks, so these factors are not isomorphic to the Verbal Memory and Executive Attention domains previously described. Similarly, comparisons with Luo et al. (2019) require caution, as their bifactor modeling was used to evaluate essential unidimensionality rather than to propose a comparable two-factor content structure. Furthermore, IRT evidence shows inconsistent information per item and ceiling effects, which encourages interpretations by profile and questionnaire abridgement based on measurement properties for screening purposes (Aiello et al., 2022; Coen et al., 2016; Roalf et al., 2016). Accordingly, we emphasize factor-level interpretation and conditional precision, together with local normative data, to support use in community screening contexts (Pereiro et al., 2017; Vallés-Salgado et al., 2024).

The results also have implications for cognitive impairment screening in community settings. The Memory factor is particularly useful in measuring delayed free recall, an episodic memory index with recognized diagnostic value (Islam et al., 2023; Nasreddine et al., 2005; Tsoi et al., 2017). Overall, the relationships between the Multiple factor and cardiometabolic risk are consistent with the hypothesis of dopaminergic decline and prefrontal compensation, changes which are characteristic of aging (Bäckman et al., 2006; Cabeza, 2002). This pattern is consistent with known modifiable targets (e.g., blood pressure and metabolic control) and suggests that monitoring by factors can provide additional information in community follow-up. At the same time, the high concordance

observed between the 12-item form and the full MoCA total score supports continuity with standard screening practice based on total scores, facilitating practical interpretation without implying structural equivalence.

Information curves indicate that the Memory factor offers greater measurement precision in the clinically relevant range for early detection, while the Multiple factor concentrates information at low levels and displays a ceiling effect at mid-to-high performance levels; a pattern consistent with the IRT literature on delayed recall, calculation, attention, and executive functions, compared to tasks with a marked ceiling effect such as orientation or naming (Luo et al., 2019; Roalf et al., 2016; Sachs et al., 2022). More specifically, peak information for the Multiple factor at very low trait levels (θ far below the mean) reflects targeting toward poorer performance (i.e., greater discrimination among lower-performing individuals), rather than a floor effect (accumulation of minimum observed scores); in a high-functioning community sample, this targeting yields limited resolution at average-to-high ability levels and is expressed as reduced differentiation near the upper end of observed performance (a ceiling-type limitation). In community settings where prevention is a priority, a factor-based interpretation can help guide referral decisions, optimize resources, and avoid unnecessary testing (Coen et al., 2016; Islam et al., 2023; Pereiro et al., 2017). Clinically, this pattern supports using the Multiple factor primarily for screening and referral when poorer performance is suspected, whereas monitoring subtle change within the normal range should be interpreted with caution and, when needed, complemented with more demanding measures and the Memory factor (Luo et al., 2019; Roalf et al., 2016; Sachs et al., 2022). In practice, this factor-level information may add value over the total score when interpreting borderline performances, monitoring change, or deciding referral in primary care and community settings. Furthermore, in individuals with high cardiovascular risk (hypertension, dyslipidemia, hyperglycaemia, or obesity), visuo-attentional markers may be sensitive to change, allowing assessment of intervention effectiveness and their impact on cognitive performance (Ngandu et al., 2015; Williamson et al., 2019). Because item selection and measurement precision depend on the intended screening purpose and the trait range where maximal precision is needed, future research should develop and cross-validate purpose-specific short forms optimized for different regions of θ using IRT information-targeting criteria and, when available, external clinical anchors (Badham et al., 2025; Mokkink et al., 2024; Roalf et al., 2016).

This study has limitations that should be considered. First, although the sample is community-based, it was assembled from primary care studies conducted in Castilla y León; therefore, generalization to other Spanish regions and sociocultural contexts should be made with caution and confirmed through external replication. The sample's sociodemographic profile was broadly comparable to official population statistics for Castilla y León, with a similar age distribution and only a modest overrepresentation of women (approximately +7 percentage points); the higher educational attainment observed is consistent with the comparatively higher educational profile of this autonomous community (INE, 2025). The cross-sectional design precludes the inference of causality, and the 12-item solution, while improving parsimony, may affect content validity; these aspects should be assessed depending on the intended use (Mokkink et al., 2024; Roalf et al., 2016). Because the 12-item form was derived by prioritizing structural replicability and measurement precision under ordinal modeling, it does not preserve the full breadth of MoCA content and should be interpreted as a dimensional screening index rather than as a substitute for domain-level clinical profiling; when domain-specific information is required, use of the full MoCA remains preferable. Furthermore, measurement invariance was not formally evaluated; therefore, comparisons between subgroups are beyond the scope of this study and should be confirmed in future research. In addition, the present study focused on a cognitive impairment-free community sample; although clinical groups (e.g., MCI or dementia) were not included, validation in these populations is an important direction for future research. Excluding participants with diabetes may also limit generalizability to populations with metabolic comorbidities, and future studies should examine whether the measurement structure identified here remains stable when these individuals are included. Finally, validity evidence against external clinical criteria (e.g., standardized neuropsychological batteries, clinical diagnoses, or biomarker-informed outcomes) was not available; therefore, any short-form cut-off should be regarded as a linkage to the conventional MoCA threshold rather than an independently validated clinical decision rule. Likewise, because no neurobiological measures (e.g., neuroimaging or biomarkers) were collected, neurocognitive interpretations should be understood as theoretical contextualization rather than empirical evidence of neural substrates or causal mechanisms. However, this does not affect the internal validity of the structure and the estimated measurement properties. Strengths include the large sample size and age range, a psychometric approach aligned with current standards, and high agreement between the total score of the original format and the abbreviated version, all of which support the relevance of this study (American Educational Research Association et al., 2014; Kline, 2023; Roalf et al., 2016).

In conclusion, the results support a MoCA structure composed of 12 items grouped into two correlated factors in a large cognitive impairment-free Spanish community sample. This structure is parsimonious, interpretable, and consistent with hierarchical cognitive models. The Memory factor offers greater accuracy in the range relevant for early screening, while the Multiple factor provides more information at low levels and shows a ceiling effect at medium-to-high performance levels; a pattern consistent with the IRT-based literature. Therefore, supplementing the total score with factor profiles and demographic adjustments is a recommended strategy for detection and referral in community settings. These findings support the 12-item form as a brief screening option anchored to the standard MoCA metric, whereas clinical profiling requiring broader content coverage remains better served by the full MoCA (Mokkink et al., 2024; Roalf et al., 2016). Future priorities include conducting longitudinal studies, using adjusted normative data, studying measurement invariance, benchmarking the 12-item form and factor scores against external clinical criteria in independent samples, and developing clinical decision support tools, such as IRT-based calculators, to facilitate translation into clinical practice.

Author Contributions

David Arjol: Conceptualization, Methodology, Data Curation, Software, Formal Analysis, Validation, Visualization, Writing – Original Draft, Writing – Review & Editing. **Sara Mora-Simón:** Conceptualization, Methodology, Supervision, Writing – Original Draft, Writing – Review & Editing. **Emiliano Rodríguez-Sánchez:**

Data Collection, Investigation, Project Administration, Writing – Review & Editing. **Manuel A. Gómez-Marcos:** Data Collection, Investigation, Project Administration, Writing – Review & Editing. **Luis García-Ortiz:** Data Collection, Investigation, Project Administration, Writing – Review & Editing. **José A. Maderuelo-Fernández:** Data Collection, Investigation, Project Administration, Writing – Review & Editing. **Cristina Lugones-Sánchez:** Conceptualization, Methodology, Data Collection, Investigation, Project Administration, Validation, Supervision, Writing – Review & Editing. **Jaime Unzueta-Arce:** Conceptualization, Methodology, Formal Analysis, Validation, Supervision, Writing – Original Draft, Writing – Review & Editing..

Acknowledgements

We would like to express our sincere gratitude to all participants for generously giving up time to cooperate in this study. We also extend our gratitude to all the professionals and researchers who collaborated in the various phases of the project, and whose dedication and effort were essential to the completion of this research.

Funding

This study has been funded by the Spanish Ministry of Science and Innovation, Instituto de Salud Carlos III (ISCIII) through the RD24/0005/0018 (Network for Research on Chronicity, Primary Care, and Health Promotion (RICAPPS) funded by the European Union-Next Generation EU, Facility for Recovery and Resilience (MRR), and the projects PI19/01140, PI20/00321, PI21/00454 co-financed by the European Union. The government of Castilla y León also collaborated with funding this study through the research projects (GRS 1820/B/18, GRS 1944/B/19, GRS 2148/B/2020, GRS 2305/B/21, GRS 2500/B/22, GRS 2502/B/22). They played no role in the study design, data analysis, reporting results, nor the decision to submit the manuscript for publication.

Declaration of Interests

The authors declare that there are no conflicts of interest.

Data Availability Statement

The dataset, analysis scripts, and supplementary materials generated for this study have been deposited in the Zenodo repository.

References

Acquadro, C., Conway, K., Giroudet, C., & Mear, I. (2012). *Linguistic validation manual for health outcome assessments*. Mapi Institute.

Aiello, E. N., Solca, F., Torre, S., Carelli, L., Ferrucci, R., Priori, A., Verde, F., Silani, V., Ticozzi, N., & Poletti, B. (2022). Diagnostics and clinical usability of the Montreal Cognitive Assessment (MoCA) in amyotrophic lateral sclerosis. *Frontiers in Psychology*, *13*, Article 1012632. https://doi.org/10.3389/fpsyg.2022.1012632

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education (2014). *Standards for Educational and Psychological Testing*. American Educational Research Association.

Aschenbrenner, A. J., Hassenstab, J. J., Schindler, S. E., Janelidze, S., Hansson, O., Morris, J. C., & Grober, E. (2024). Free recall outperforms story recall in associations with plasma biomarkers in preclinical Alzheimer's disease. *The Journal of Prevention of Alzheimer Disease, 11*(6), 1696-1702. https://doi.org/10.14283/jpad.2024.130

Bäckman, L., Nyberg, L., Lindenberger, U., Li, S. C., & Farde, L. (2006). The correlative triad among aging, dopamine, and cognition: current status and future prospects. *Neuroscience & Biobehavioral Reviews*, *30*(6), 791-807. https://doi.org/10.1016/j.neubiorev.2006.06.005

Badham, L., Oliveri, M. E., & Sireci, S. G. (2025). Navigating multi-language assessments: best practices for test development, linking, and evaluation. *Psicothema*, *37*(3), 1-11. https://doi.org/10.70478/psicothema.2025.37.19

Bugallo-Carrera, C., Dosil-Díaz, C., Pereiro, A. X., Anido-Rifón, L., & Gandoy-Crego, M. (2024). Factors that indicate performance on the MoCA 7.3 in healthy adults over 50 years old. *BMC Geriatrics*, *24*(1), Article 482. https://doi.org/10.1186/s12877-024-05102-1

Cabeza, R. (2002). Hemispheric asymmetry reduction in older adults: the HAROLD model. *Psychology and Aging*, *17*(1), 85-100. https://doi.org/10.1037/0882-7974.17.1.85

Cancino, M., Rehbein, L., Gómez-Pérez, D., & Ortiz, M. S. (2020). Evaluación de funcionamiento cognitivo en el adulto mayor: Análisis y contrastación de los tres instrumentos más usados en Chile [Psychometric properties of three instruments to detect dementia]. *Revista Médica de Chile*, *148*(4), 452-458. https://doi.org/10.4067/s0034-98872020000400452

Chalmers, R. P. (2012). mirt: A multidimensional item response theory package for the R environment. *Journal of Statistical Software*, *48*(6), 1-29. https://doi.org/10.18637/jss.v048.i06

Chithiramohan, T., Santhosh, S., Threlfall, G., Hull, L., Mukaetova-Ladinska, E. B., Subramaniam, H., & Beishon, L. (2024). Culture-fair cognitive screening tools for assessment of cognitive impairment: A systematic review. *Journal of Alzheimer's Disease Reports*, *8*(1), 289-306. https://doi.org/10.3233/ADR-230194

Ciesielska, N., Sokołowski, R., Mazur, E., Podhorecka, M., Polak-Szabela, A., & Kędziora-Kornatowska, K. (2016). Is the Montreal Cognitive Assessment (MoCA) test better suited than the Mini-Mental State Examination (MMSE) in mild cognitive impairment (MCI) detection among people aged over 60? Meta-analysis. *Psychiatria Polska*, *50*(5), 1039-1052. https://doi.org/10.12740/PP/45368

Coen, R. F., Robertson, D. A., Kenny, R. A., & King-Kallimanis, B. L. (2016). Strengths and limitations of the MoCA for assessing cognitive functioning: findings from a large representative sample of Irish older adults. *Journal of Geriatric Psychiatry and Neurology*, *29*(1), 18-24. https://doi.org/10.1177/0891988715598236

Craig, C. L., Marshall, A. L., Sjöström, M., Bauman, A. E., Booth, M. L., Ainsworth, B. E., Pratt, M., Ekelund, U., Yngve, A., Sallis, J. F., & Oja, P. (2003). International physical activity questionnaire: 12-country reliability and validity. *Medicine and Science in Sports and Exercise, 35*(8), 1381–1395. https://doi.org/10.1249/01.MSS.0000078924.61453.FB

Dragomir, A., & Omurtag, A. (2021). Brain's networks and their functional significance in cognition. In N. V. Thakor (Eds), *Handbook of Neuroengineering* (pp. 1-30). Springer Singapore. https://doi.org/10.1007/978-981-16-5540-1_76

Duro, D., Simões, M. R., Ponciano, E., & Santana, I. (2010). Validation studies of the Portuguese experimental version of the Montreal Cognitive Assessment (MoCA): Confirmatory factor analysis. *Journal of Neurology, 257*(5), 728–734. https://doi.org/10.1007/s00415-009-5399-5

Embretson, S. E., & Reise, S. P. (2013). *Item response theory for psychologists. Psychology Press*. https://doi.org/10.4324/9781410605269

Ferrando, P. J., Morales-Vives, F., Casas, J. M., & Muñiz, J. (2025). Likert scales: A practical guide to design, construction and use. *Psicothema*, *37*(4), 1-15. https://doi.org/10.70478/psicothema.2025.37.24

Freitas, S., Prieto, G., Simões, M. R., & Santana, I. (2015). Scaling cognitive domains of the Montreal Cognitive Assessment: An analysis using the partial credit model. *Archives of Clinical Neuropsychology*, *30*(5), 435-447. https://doi.org/10.1093/arclin/acv027

Freitas, S., Simões, M. R., Marôco, J., Alves, L., & Santana, I. (2012). Construct validity of the Montreal Cognitive Assessment (MoCA). *Journal of the International Neuropsychological Society, 18*(2), 242–250. https://doi.org/10.1017/S1355617711001573

Gagnon, C., Saillant, K., Olmand, M., Gayda, M., Nigam, A., Bouabdallaoui, N., Rouleau, J.L., Desjardins-Crépeau, L., & Bherer, L. (2022). Performances on the Montreal Cognitive Assessment along the cardiovascular disease continuum. *Archives of Clinical Neuropsychology*, *37*(1), 117-124. https://doi.org/10.1093/arclin/acab029

Geethadevi, G. M., Quinn, T. J., George, J., Anstey, K. J., Bell, J. S., Sarwar, M. R., & Cross, A. J. (2023). Multi-domain prognostic models used in middle-aged adults without known cognitive impairment for predicting subsequent dementia. *Cochrane Database of Systematic Reviews*, (6), CD014885. https://doi.org/10.1002/14651858.CD014885.pub2

Grober, E., Lipton, R. B., Sperling, R. A., Papp, K. V., Johnson, K. A., Rentz, D. M., Veroff, A. E., Aisen, P. S., & Ezzati, A. (2022). Associations of stages of objective memory impairment with amyloid PET and structural MRI. *Neurology*, *98*(13), e1327-e1336. https://doi.org/10.1212/wnl.0000000000200046

Grober, E., Petersen, K. K., Lipton, R. B., Hassenstab, J., Morris, J. C., Gordon, B. A., & Ezzati, A. (2023). Association of stages of objective memory impairment with incident symptomatic cognitive impairment in cognitively normal individuals. *Neurology*, *100*(22), e2279-e2289. https://doi.org/10.1212/wnl.0000000000207276

Guo, Z. X., Liu, F., Wang, F. Y., Ou, Y. N., Huang, L. Y., Hu, H., Wang, Z., Fu, Y., Gao, P., Tan, L., & Yu, J. T. (2024). CAIDE score, Alzheimer's disease pathology, and cognition in cognitively normal adults: The CABLE study. *Journal of Alzheimer's Disease, 99*(4), 1273-1283. https://doi.org/10.3233/JAD-240005

Hobeika, L., Diard-Detoeuf, C., Garcin, B., Levy, R., & Volle, E. (2016). General and specialized brain correlates for analogical reasoning: A meta-analysis of functional imaging studies. *Human Brain Mapping*, *37*(5), 1953-1969. https://doi.org/10.1002/hbm.23149

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal, 6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Imai, A., Matsuoka, T., Kato, Y., & Narumoto, J. (2022). Diagnostic performance and neural basis of the combination of free- and pre-drawn Clock Drawing Test. *International Journal of Geriatric Psychiatry*, *37*(4), 1-9. https://doi.org/10.1002/gps.5699

Instituto Nacional de Estadística. (2011). *Clasificación Nacional de Ocupaciones 2011 (CNO-11)* [National Classification of Occupations 2011 (CON-11)]. https://www.ine.es/daco/daco42/clasificaciones/cno11_notas.pdf

Instituto Nacional de Estadística. (2025). Censo anual de población 2021–2025 [Annual Population Census 2021–2025]. https://www.ine.es/uc/dxF29hro

Islam, M. A., Hashem, R., Gad, M., Brown, A., Levis, B., Renoux, C., Thombs, B. D., & McInnes, M. D. (2023). Accuracy of the Montreal Cognitive Assessment tool for detecting mild cognitive impairment: A systematic review and meta-analysis. *Journal of the Alzheimer's Association, 71*(8), 2567-2577. https://doi.org/10.1002/alz.13040

Jia, R., Wang, Q., Huang, H., Yang, Y., Chung, Y. F., & Liang, T. (2023). Cardiovascular disease risk models and dementia or cognitive decline: A systematic review. *Frontiers in Aging Neuroscience*, *15*, Article 1257367. https://doi.org/10.3389/fnagi.2023.1257367

Jorgensen, T. D., Pornprasertmanit, S., Schoemann, A. M., Rosseel, Y., Miller, P., Quick, C., Garnier-Villarreal, M., Selig, J., Boulton, A., Preacher, K., Coffman, D., Rhemtulla, M., Robitzsch, A., Enders, C., Arslan, R., Clinton, B., Panko, P., Merkle, E., Chesnut, S., …Rhemtulla, M. (2016). *semTools: Useful tools for structural equation modeling* (R package version 0.5-7) [Computer software]. Comprehensive R Archive Network (CRAN). https://cran. r-project.org/web/packages/semTools/semTools. pdf

Kivipelto, M., Ngandu, T., Laatikainen, T., Winblad, B., Soininen, H., & Tuomilehto, J. (2006). Risk score for the prediction of dementia risk in 20 years among middle aged people: A longitudinal, population-based study. *The Lancet Neurology, 5*(9), 735–741. https://doi.org/10.1016/S1474-4422(06)70537-3

Kline, R. B. (2023). *Principles and practice of structural equation modeling*. Guilford Publications.

Lentoor, A. G., & Myburgh, L. (2022). Correlation between body mass index (BMI) and performance on the Montreal Cognitive Assessment (MoCA) in a cohort of adult women in South Africa. *Behavioural Neurology*, *2022*(1), Article 8994793. https://doi.org/10.1155/2022/8994793

Llamas-Ramos, I., Llamas-Ramos, R., Lugones-Sánchez, C., González-García, S., Tamayo-Morales, O., Alvarado-Omenat, J. J., ... & Rodríguez-Sánchez, E. (2024). Effect of a lifestyle-integrated functional exercise (LiFE) group intervention (sLiFE) to falls prevention in non-institutionalized older adults. Protocol of a randomised clinical trial. *Frontiers in Public Health*, *11*, Article 1304982. https://doi.org/10.3389/fpubh.2023.1304982

Lozano Gallego, M., Hernández Ferrándiz, M., Turró Garriga, O., Pericot Nierga, I., López Pousa, S., & Vilalta Franch, J. (2009). Validación del Montreal Cognitive Assessment (MoCA): Test de cribado para el deterioro cognitivo leve. Datos preliminares [Validation of the Montreal Cognitive Assessment (MoCA): Screening test for mild cognitive impairment. Preliminary data]. *Alzheimer. Realidades e Investigación en Demencia,* 43, 4-11. https://www.researchgate.net/publication/228740373

Lugones-Sánchez, C., Santos-Mínguez, S., Salvado, R., González-Sánchez, S., Tamayo-Morales, O., Hoya-González, A., Ramírez-Manent, J.I., Magallón-Botaya, R., Quesada-Rico, J.A., García-Cubillas, M. D., Rodríguez-Sánchez, E., Gómez-Marcos, M.A., Benito-Sanchez, R., Mira, A., Hernandez-Rivas, J.M., García-Ortiz, L., & MIVAS

III Researchers Group. (2023). Lifestyles, arterial aging, and its relationship with the intestinal and oral microbiota (MIVAS III study): a research protocol for a cross-sectional multicenter study. *Frontiers in Public Health*, *11*, Article 1164453. https://doi.org/10.3389/fpubh.2023.1164453

Luo, H., Andersson, B., Tang, J. Y., & Wong, G. H. (2019). Applying item response theory analysis to the Montreal Cognitive Assessment in a low-education older population. *Assessment*, *27*(7),1416-1428. https://doi.org/10.1177/1073191118821733

Luo, H., Andersson, B., Wong, G. H. Y., & Lum, T. Y. S. (2022). Longitudinal measurement properties of the Montreal Cognitive Assessment. *Journal of Clinical and Experimental Neuropsychology, 44*(9), 627-639. https://doi.org/10.1080/13803395.2022.2148634

Marsh, H. W., Hau, K. T., & Wen, Z. (2004). In search of golden rules: Comment on hypothesis-testing approaches to setting cutoff values for fit indexes and dangers in overgeneralizing Hu and Bentler's (1999) findings. *Structural Equation Modeling: A Multidisciplinary Journal*, *11*(3), 320-341. https://doi.org/10.1207/s15328007sem1103_2

Martí-Lluch, R., Bolíbar, B., Llobera, J., Maderuelo-Fernández, J. A., Magallón-Botaya, R., Sánchez-Pérez, Á., Fernández-Domínguez, M.J., Motrico, E., Vicens-Pons, E., Notario-Pacheco, B., Alves-Cabratosa, L., Ramos, R. & DESVELA Cohort Investigators. (2023). Role of personal aptitudes as determinants of incident morbidity, lifestyles, quality of life, use of health services, and mortality (DESVELA cohort): quantitative study protocol for a prospective cohort study in a hybrid analysis. *Frontiers in Public Health*, *11*, Article 1067249. https://doi.org/10.3389/fpubh.2023.1067249

Matías-Guiu, J. A., Valles-Salgado, M., Rognoni, T., Hamre-Gil, F., Moreno-Ramos, T., & Matías-Guiu, J. (2017). Comparative diagnostic accuracy of the ACE-III, MIS, MMSE, MoCA, and RUDAS for screening of Alzheimer disease. *Dementia and Geriatric Cognitive Disorders*, *43*(5-6), 237–246. https://doi.org/10.1159/000469658

Menon, V., & D'Esposito, M. (2021). The role of PFC networks in cognitive control and executive function. *Neuropsychopharmacology, 47*, 90-103. https://doi.org/10.1038/s41386-021-01152-w

Ministerio de Sanidad. (2020). *Límites de consumo de bajo riesgo de alcohol* [Low-risk alcohol consumption limits]. https://www.sanidad.gob.es/areas/promocionPrevencion/alcohol/documentosTecnicos/docs/Limites_Consumo_Bajo_Riesgo_Alcohol_Actualizacion.pdf

Mokkink, L. B., Elsman, E. B., & Terwee, C. B. (2024). COSMIN guideline for systematic reviews of patient-reported outcome measures version 2.0. *Quality of Life Research*, *33*, 2929-2939. https://doi.org/10.1007/s11136-024-03761-6

Nasreddine, Z. S., Phillips, N. A., Bédirian, V., Charbonneau, S., Whitehead, V., Collin, I., Cummings, J. L., & Chertkow, H. (2005). The Montreal Cognitive Assessment, MoCA: A brief screening tool for mild cognitive impairment. *Journal of the American Geriatrics Society*, *53*(4), 695–699. https://doi.org/10.1111/j.1532-5415.2005.53221.x

Ngandu, T., Lehtisalo, J., Solomon, A., Levälahti, E., Ahtiluoto, S., Antikainen, R., Bäckman, L., Hänninen, T., Jula, A., Laatikainen, T., Lindström, J., Mangialasche, F., Paajanen, T., Pajala, S., Peltonen, M., Rauramaa, R., Stigsdotter-Neely, A., Strandberg, T., Tuomilehto, J., … Kivipelto, M. (2015). A 2 year multidomain intervention of diet, exercise, cognitive training, and vascular risk monitoring versus control to prevent cognitive decline in at-risk elderly people (FINGER): A randomised controlled trial. *The Lancet*, *385*(9984), 2255-2263. https://doi.org/10.1016/S0140-6736(15)60461-5

Nicotra, A., Maestri, G., Salvadori, E., & Pantoni, L. (2024). Brain health assessment. An exploratory review of tools related to its cognitive dimension. *Cerebral Circulation-Cognition and Behavior*, *6*, Article 100188. https://doi.org/10.1016/j.cccb.2023.100188

O'Caoimh, R., Cadoo, S., Daly, B., & Molloy, D. W. (2025). Comparing short cognitive screening instruments in an outreach memory clinic in primary care. *International journal of Environmental Research and Public Health, 22*(3), Article 410. https://doi.org/10.3390/ijerph22030410

Ojeda, N., del Pino, R., Ibarretxe-Bilbao, N., Schretlen, D. J, & Peña, J. (2016). Montreal Cognitive Assessment Test: Normalization and standardization for Spanish population. *Revista de Neurología, 63*(11), 488–496. https://doi.org/10.33588/rn.6311.2016241

Pereiro, A. X., Ramos-Lema, S., Juncos-Rabadán, O., Facal, D., & Lojo-Seoane, C. (2015). Normative scores of the Cambridge Cognitive Examination-Revised in healthy Spanish population. *Psicothema, 27*(1), 32-39. https://doi.org/10.7334/psicothema2014.169

Pereiro, A. X., Ramos-Lema, S., Lojo-Seoane, C., Guàrdia-Olmos, J., Facal-Mayo, D., & Juncos-Rabadán, O. (2017). Normative data for the Montreal Cognitive Assessment (MOCA) in a Spanish sample of community-dweller adults. *European Geriatric Medicine*, *8*(3), 240–244. https://doi.org/10.1016/j.eurger.2017.04.003

Porto, J. M. (2015). *Validez diagnóstica de la evaluación cognitiva Montreal en el deterioro cognitivo posictus* [Diagnostic validity of the Montreal Cognitive Assessment in post-stroke cognitive impairment][Doctoral dissertation, University of Salamanca]. GREDOS USAL Repository. https://hdl.handle.net/10366/128548

Qu, J., Fu, J., Liu, X., Chen, M., Yang, T., & Shen, Z. (2025). Episodic memory neural mechanisms: Patterns, connectivity, and developmental dynamics. *Brain and Cognition, 191*, Article 106374. https://doi.org/10.1016/j.bandc.2025.106374

R Development Core Team. (2023). *R: A language and environment for statistical computing* [Computer software]. R Foundation for Statistical Computing. http://www.R-project.org/

Ramírez-Moreno, J. M., Bartolomé Alberca, S., Muñoz Vega, P., & Guerrero Barona, E. J. (2022). Screening for cognitive impairment with the Montreal Cognitive Assessment in Spanish patients with minor stroke or transient ischaemic attack. *Neurología*, *37*(1), 38–44. https://doi.org/10.1016/j.nrleng.2018.11.008

Roalf, D. R., Moore, T. M., Wolk, D. A., Arnold, S. E., Mechanic-Hamilton, D., Rick, J., Kabadi, S., Ruparel, K., Chen-Plotkin, A.S., Chahine, L.M., Dahodwala, N.A., Duda, J.E., Weintraub, D.A., & Moberg, P. J. (2016). Defining and validating a short form Montreal Cognitive Assessment (s-MoCA) for use in neurodegenerative disease. *Journal of Neurology, Neurosurgery & Psychiatry*, *87*(12), 1303-1310. https://doi.org/10.1136/jnnp-2015-312723

Rodríguez, A., Reise, S. P., & Haviland, M. G. (2016). Evaluating bifactor models: Calculating and interpreting statistical indices. *Psychological Methods, 21*(2), 137–150. https://doi.org/10.1037/met0000045

Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, *48*(2), 1-36.

Sachs, B. C., Chelune, G. J., Rapp, S. R., Couto, A. M., Willard, J. J., Williamson, J. D., Sink K. M., Coker, L. H., Gaussoin, S. A., Gure, T.

R., Lerner, A. J., Nichols, L. O., Still, C. H., Wadley, V. G., & Pajewski, N. M. (2022). Robust demographically-adjusted normative data for the Montreal Cognitive Assessment (MoCA): Results from the systolic blood pressure intervention trial. *The Clinical Neuropsychologist*, *36*(8), 2237-2259. https://doi.org/10.1080/13854046.2021.1967450

Sala, G., Inagaki, H., Ishioka, Y., Masui, Y., Nakagawa, T., Ishizaki, T., Arai, Y., Ikebe, K., Kamide, K., & Gondo, Y. (2020). The psychometric properties of the Montreal Cognitive Assessment (MoCA). *Swiss Journal of Psychology, 79*(3-4), 155-161. https://doi.org/10.1024/1421-0185/a000242

Sergi, M. R., Balsamo, M., D'Ignazio, G., Terrei, M., Palumbo, R., Salvatore, G., & Carlucci, L. (2025). Investigating the factor structure of the Montreal Cognitive Assessment: A qualitative review. *Frontiers in Psychology*, *16*, Article 1727202. https://doi.org/10.3389/fpsyg.2025.1727202

Sindi, S., Calov, E., Fokkens, J., Ngandu, T., Soininen, H., Tuomilehto, J., & Kivipelto, M. (2015). The CAIDE Dementia Risk Score App: The development of an evidence-based mobile application to predict the risk of dementia. *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring, 1*(3), 328-333. https://doi.org/10.1016/j.dadm.2015.06.005

Tallman, C. W., Luo, Z., & Smith, C. N. (2024). Human brain activity and functional connectivity associated with verbal long-term memory consolidation across 1 month. *Frontiers in Human Neuroscience*, *18*, Article 1342552. https://doi.org/10.3389/fnhum.2024.1342552

Teng, S., Chen, X., & Chen, Z. (2025). Dynamic memory enqrams: Unveiling the celular mechanisms of memory encoding, consolidation, generalizaton, and updating in the brain. *Brain Medicine, 1*(4), 50-61. https://doi.org/10.61373/bm025i.0044

Tsai, C. F., Lee, W. J., Wang, S. J., Shia, B. C., Nasreddine, Z., & Fuh, J. L. (2012). Psychometrics of the Montreal Cognitive Assessment (MoCA) and its subscales: Validation of the Taiwanese version of the MoCA and an item response theory analysis. *International Psychogeriatrics*, *24*(4), 651-658. https://doi.org/10.1017/S1041610211002298

Tsoi, K. K. F., Chan, J. Y. C., Hirai, H. W., Wong, A., Mok, V. C. T., Lam, L. C. W., Kwok, T. C. Y., & Wong, S. Y. S. (2017). Recall tests are effective to detect mild cognitive impairment: A systematic review and meta-analysis of 108 diagnostic studies. *Journal of the American Medical Directors Association*, *18*(9), 807.e17-807.e29. https://doi.org/10.1016/j.jamda.2017.05.016

Vallés-Salgado, M., Matias-Guiu, J. A., Delgado-Álvarez, A., Delgado-Alonso, C., Gil-Moreno, M. J., Valiente-Gordillo, E., López-Carbonero, J. I., Fernández-Romero, L., Peña-DeDiego, L., Oliver-Mas, S., Matías-Guiu, J., & Diez-Cirarda, M. (2024). Comparison of the diagnostic accuracy of five cognitive screening tests for diagnosing mild cognitive impairment in patients consulting for memory loss. *Journal of Clinical Medicine*, *13*(16), Article 4695. https://doi.org/10.3390/jcm13164695

Vicente-Gabriel, S., Lugones-Sánchez, C., Tamayo-Morales, O., Vicente Prieto, A., González-Sánchez, S., Conde, S., Gómez-Sánchez, M., Rodríguez-Sánchez, E., García-Ortiz, L., Gómez-Sánchez, L., Gómez-Marcos, M., & Fernandez-Matas, L. (2024). Relationship between addictions and obesity, physical activity and vascular aging in young adults (EVA-Adic study): a research protocol of a cross-sectional study. *Frontiers in Public Health*, *12*, Article 1322437. https://doi.org/10.3389/fpubh.2024.1322437

Wen, S., Cheng, D., Zhao, N., Chen, X., Lu, X., Li, Y., Liu, H., Gao, J., Hou, C., & Xu, R. (2025). Psychometric properties of screening tools for mild cognitive impairment in older adults based on COSMIN guidelines: A systematic review. *BMC Geriatrics*, *25*(1), Article 401. https://doi.org/10.1186/s12877-025-06030-4

Williams, B., Mancia, G., Spiering, W., Agabiti Rosei, E., Azizi, M., Burnier, M., Clement, D., Coca, A., De Simone, G., Dominiczak, A., Kahan, T., Mahfoud, F., Redon, J., Ruilope, L., Zanchetti, A., Kerins, M., Kjeldsen, S., Kreutz, R., Laurent, S., …Desormais, I. (2018). 2018 Practice Guidelines for the management of arterial hypertension of the European Society of Cardiology and the European Society of Hypertension. *Blood Pressure*, *27*(6), 314-340. https://doi.org/10.1080/08037051.2018.1527177

Williamson, J. D., Pajewski, N. M., Auchus, A. P., Bryan, R. N., Chelune, G., Cheung, A. K., Cleveland, M. L., Coker, L. H., Crowe, M. G., Cushman, W. C., Cutler, J. A., Davatzikos, C., Desiderio, L., Erus, G., Fine, L. J., Gaussoin, S. A., Harris, D., Hsieh, M., Johnson, K. C., …The SPRINT MIND investigators for the SPRINT Research Group. (2019). Effect of intensive vs standard blood pressure control on probable dementia: A randomized clinical trial. *JAMA*, *321*(6), 553-561. https://doi.org/10.1001/jama.2018.21442

World Medical Association. (2025). World Medical Association Declaration of Helsinki: ethical principles for medical research involving human participants. *Jama*, *333*(1), 71-74. https://doi.org/10.1001/jama.2024.21972

Yin, S., Gao, P. Y., Ou, Y. N., Fu, Y., Liu, Y., Wang, Z. T., Han, B., & Tan, L. (2024). ANU-ADRI scores, tau pathology, and cognition in non-demented adults: the CABLE study. *Alzheimer's Research & Therapy*, *16*(1), Article 65. https://doi.org/10.1186/s13195-024-01427-6

GRUPO EDITORIAL DE PSICOFUNDACIÓN

Psicothema (2026) 38(3) 173

Psicothema

https://www.psicothema.com/es • eISSN 1886-144X

Colegio Oficial de Psicología del Principado de Asturias

Psicothema

Correction

Corrigendum to "Psychometric properties of the Teachers´ Responses to Bullying Questionnaire (TRBQ) in Spanish students" [Psicothema 38(1), 46-57]

Laura Rodríguez-Pérez[1], Rosario Del Rey[1], Noemí García-Sanjuán[2], & Noelia Muñoz-Fernández[1]

[1] Universidad de Sevilla (Spain)

[2] Universidad Internacional de La Rioja (UNIR) (Spain)

This article was originally published with errors, which have now been corrected in the online version:

An error was identified in the Results section. Specifically, a paragraph corresponding to the measurement invariance analysis was inadvertently omitted during the publication process.

The missing paragraph should appear under the section "Measurement Invariance", immediately before the subsection "Differences in the Perception of Teacher Responses", and reads as follows:

"Table 3 presents the results of the measurement invariance analysis across educational level, gender, and bullying role. The results indicate that the TRBQ demonstrated full measurement invariance across all groups. The establishment of measurement invariance in the TRBQ allowed comparisons to be made between students' perceptions of teachers' responses to bullying strategies—non-intervention, restorative psychoeducational strategies, and disciplinary methods—based on educational level, gender, and bullying role involvement."

DOI of original article: https://doi.org/10.70478/psicothema.2026.38.05