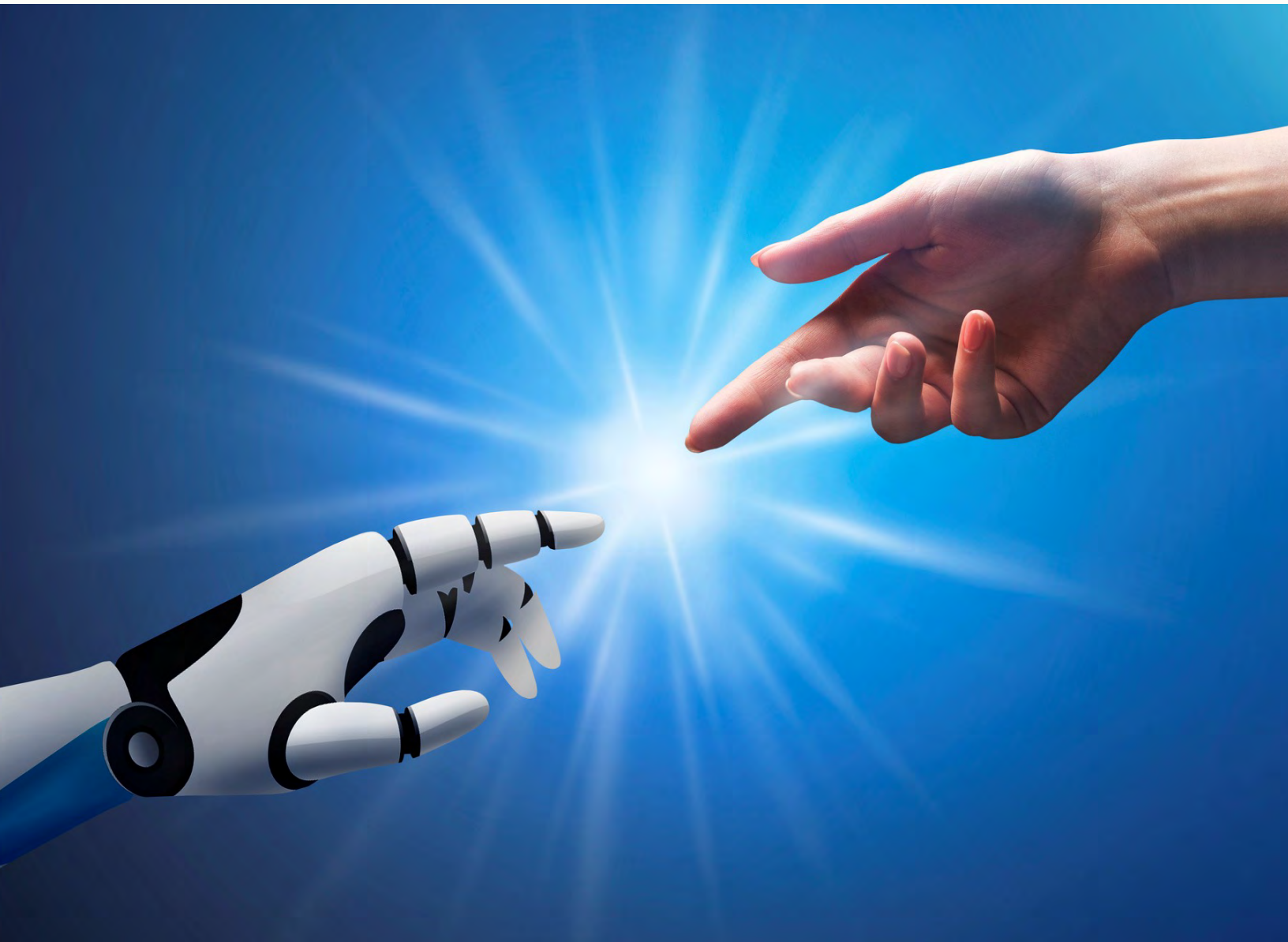
GLOBAL PSYCHOLOGY ALLIANCE

Psychology for Global Challenges

Top 10

Principles for Understanding Artificial Intelligence in Psychology



PRINCIPLES FOR UNDERSTANDING ARTIFICIAL INTELLIGENCE IN PSYCHOLOGY

GLOBAL PSYCHOLOGY ALLIANCE CONTRIBUTORS

AUTHORS

Amanda Clinton, PhD - American Psychological Association
Angel Elgier, PhD - Pontificia Universidad Católica del Perú
Djavila Ho, PsyD - Jamaican Psychological Society & KAHLE Journey
Kanika Jindal, MPhil - Harmony Therapy World, India & Middle East Psychological Association, Qatar
John Francis Leader, PhD - University College Dublin
Marijana Markovikj, PhD - University of Ss Cyril and Methodius, Skopje
Timo Ruusuvirta, PhD - The Finnish Psychological Association
Diwakar Singh Sikarwar, PhD - Counsellor Council of India

COORDINATION

Maria Elena Humphrey, PhD - Central American Union of Psychology Associations

REVIEWERS

Devin Fowlie, PhD - Canadian Psychological Association
Tatiana Manrique, PhD - Colombian College of Psychologists
Shukria Khan, PhD - Pakistan Psychological Association
Zoltan Vass, PhD - Károli Gáspár University of the Reformed Church, Budapest
Cheng-Ta Yang, PhD - Taiwanese Psychological Association
Peejay Bengwasan, PhD - Psychological Association of the Philippines
Jonathan B. Aganan, PhD - University of Santo Tomas, Manila, Philippines

LAYOUT AND GRAPHIC DESIGN

Jonathan B. Aganan, PhD - University of Santo Tomas, Manila, Philippines
Rafael Iñigo I. Marqueses - University of Santo Tomas, Manila, Philippines

PRINCIPLES FOR UNDERSTANDING ARTIFICIAL INTELLIGENCE IN PSYCHOLOGY

TABLE OF CONTENTS

TOP 10 PRINCIPLES FOR UNDERSTANDING ARTIFICIAL INTELLIGENCE (AI) IN PSYCHOLOGY....1

Ethics, Transparency, and AI....1

Professional Competence and AI....1

Access, Bias, and Inclusion with AI....2

INTRODUCTION....3

METHODOLOGY....4

CONCLUDING REMARKS....5

THE PRINCIPLES EXPLAINED....6

Ethics, Transparency, and AI....7

Professional Competence and AI....11

Access, Bias, and Inclusion with AI....14

GLOSSARY....17

PRINCIPLES FOR UNDERSTANDING ARTIFICIAL INTELLIGENCE IN PSYCHOLOGY

TOP 10 PRINCIPLES FOR UNDERSTANDING ARTIFICIAL INTELLIGENCE (AI) IN PSYCHOLOGY

ETHICS, TRANSPARENCY, AND AI

PRINCIPLE 1: ETHICAL RESPONSIBILITY

Psychologists should be informed by professionally recognized ethical principles relating to the use of AI.

PRINCIPLE 2: TRANSPARENCY IN SELECTION OF AI TOOLS

Psychologists should use only those AI tools that offer transparency regarding their development and provide clear, conspicuous, and persistent disclosure to users that they are interacting with a nonhuman entity. When full transparency is unavailable, psychologists should document known limitations and seek available technical documentation about data practices.

PRINCIPLE 3: INFORMED CONSENT

Psychologists who incorporate AI-based tools into their professional activities should clearly advise relevant parties about potential risks and benefits. Consent must be explicit regarding limitations of AI tools, must include opt-in for use of individual data, and must not use deceptive design patterns.

PRINCIPLE 4: ACCOUNTABILITY, GOVERNANCE, AND PROFESSIONAL RESPONSIBILITY

Psychologists who develop, deploy, or rely on AI systems share professional accountability for their appropriate use and outcomes alongside other stakeholders and should advocate for clear governance structures that distribute responsibility appropriately.

PROFESSIONAL COMPETENCE AND AI

PRINCIPLE 5: EVIDENCE BASE

Psychologists who integrate AI into professional activities should possess sufficient foundational understanding of how their selected AI tools function, including their general methodology, data requirements, and known limitations relevant to psychological applications

PRINCIPLE 6: HUMAN OVERSIGHT AND PROFESSIONAL JUDGMENT

Psychologists who incorporate AI tools into their professional activities must have sufficient expertise and discretion to ensure appropriate oversight of AI-supported outcomes.

PRINCIPLE 7: PSYCHOLOGICAL SAFETY AND HARM PREVENTION

Psychologists who employ AI systems should prioritize maximize benefit alongside psychological safety, and actively work to prevent potential harms.

ACCESS, BIAS, AND INCLUSION WITH AI

PRINCIPLE 8: VALIDATION AND SCIENTIFIC EVIDENCE

Psychologists who use AI tools should critically evaluate whether such systems are supported by rigorous, transparent, and contextually appropriate scientific evidence that demonstrates reliability, validity, and relevance for the populations and purposes for which they are used.

PRINCIPLE 9: ALGORITHMIC BIAS AND FAIRNESS

Psychologists who employ AI systems should proactively identify, evaluate, and mitigate algorithmic bias, ensuring that AI-supported outputs do not systematically disadvantage or discriminate against individuals or groups based on demographic, cultural, socioeconomic, or contextual factors.

PRINCIPLE 10: ACCESS, EQUITY, AND INCLUSION

Psychologists who integrate AI tools into their professional activities should promote equitable access and inclusive design, ensuring that AI-supported psychological services do not exacerbate existing disparities or exclude individuals and communities based on structural, economic, linguistic, or technological barriers.

INTRODUCTION

Since the COVID-19 pandemic, the combination of increased global attention to mental health issues and public adoption of digital tools has resulted in significant need for psychological science informing human factors associated with use of technology.

The increasing prevalence of technology and the widespread availability of digital platforms, tools, and applications addressing well-being and mental health highlight the need for psychological science to ensure both effectiveness and safety. This heightened digital exposure has amplified discussions about psychological well-being and raised new ethical challenges regarding the responsible integration of emerging technologies, notably artificial intelligence (AI).

The Top 10 Principles for Understanding Artificial Intelligence in Psychology is a response to the rapid advancement of AI and its profound influence on the practice of psychology across various specialties, reshaping how psychologists work. This guidance offers an overview of critical knowledge and professional obligations for psychologists committed to responsible and effective integration of AI into mental health service delivery. As an introductory endeavor, these principles are focused on recommendations in the areas of ethics, clinical service provision, and understanding bias as well as professional competence, human oversight, and the promotion of equitable access.

It is essential to acknowledge global disparities in the development, accessibility, and cultural relevance of AI technologies. Psychologists recognize that unequal digital infrastructure, economic constraints, and cultural contexts may limit the ability of many communities to benefit from AI-based psychological tools. Ensuring equitable access and culturally responsive design is therefore a central consideration in the ethical integration of AI.

METHODOLOGY

The methodological orientation of this publication is conceptual and normative rather than empirical. Instead of generating data, it synthesizes existing ethical codes (e.g., American Psychological Association, 2017; British Psychological Society, 2021; Canadian Code of Ethics for Psychologists, 2017), international human rights and data protection frameworks (e.g., Europe's General Data Protection Regulation, U.S. Health Insurance Portability and Accountability Act, World Health Organization guidance) and scholarly data on digital health and AI in psychology.

The determination of these 10 principles within the three categories of **ethics; professional competence; and access, bias, and inclusion**, resulted from the thematic analysis of contributions from psychologists across the Global Psychology Alliance, based on 20 survey responses from distinct regions of the world.
This approach ensured that the framework is grounded in existing global standards, yet critically responsive to the evolving landscape of digital technologies.

The implications of this framework are multidimensional: The guidance calls for rigorous, culturally sensitive validation of AI-based and digital psychological tools in an effort to prevent harm from untested or biased algorithms. Furthermore, these principles highlight the urgent need to embed digital ethics, AI literacy, and critical thinking into curricula for psychologists at undergraduate, graduate, and continuing professional education levels.

In terms of mental health service delivery, this framework stresses that psychologists not just engage with AI as consumers, but that they actively model the ethical use of technology, ensuring informed consent, confidentiality, and human dignity in all digital interactions. Beyond psychology, the framework has policy implications: It can inform international guidelines, national regulations, and platform design standards, ultimately contributing to a coherent governance architecture for digital mental health.

The Top 10 Principles for Understanding Artificial Intelligence in Psychology is particularly timely for the field at a global level. Although there has been an increase in research data concerning AI in mental health service delivery, a noticeable gap remained in terms of guidance for psychologists that addresses ethical implementation, potential bias, integration approaches, and professional competence. This document fulfills this need by offering carefully drafted principles co-created by Global Psychology Alliance psychologist participants from more than 80 countries.

CONCLUDING REMARKS

Best practice in the use of AI in psychology requires psychologists to act in accordance with applicable laws, professional regulations, codes of ethics, and established standards of good practice within their local professional contexts. This guidance seeks to support, rather than replace or supersede, these frameworks. The safe, ethical, and effective use of AI in psychology relies on individual practitioners as well as a range of stakeholders, including but not limited to developers, institutions, and regulators. It is inherently a multistakeholder endeavor, requiring coordinated responsibility among psychologists, professional bodies, educators, researchers, technology developers, employers, regulators, and policymakers, among others. Shared governance, interdisciplinary collaboration, and ongoing dialogue are essential to ensure that AI systems are designed, implemented, monitored, and updated in ways that protect clients, shield psychologists, uphold professional values, and promote public trust while allowing the responsible integration of innovation into psychological practice.

EDITORIAL NOTE:

The Top 10 Principles for Understanding Artificial Intelligence in Psychology are summarized at the beginning of this document. The following pages expand upon the principles in their overarching description, while also providing increased explanation and relevance

REFERENCES

American Psychological Association. (2017). Ethical principles of psychologists and code of conduct. https://www.apa.org/ethics/code

British Psychological Society. (2021). Code of ethics and conduct. https:..explore.bps.org.uk/content/report-guidelines/bpsrep.2021.inf94

Canadian Psychological Association. (2017). Canadian code of ethics for psychologists. https://cpa.ca/doc/File/Ethics/CPA_Code_2017_4thEd.pdf

PRINCIPLE

1

ETHICS, TRANSPARENCY, AND AI

ETHICAL RESPONSIBILITY

Psychologists should be informed by professionally recognized ethical principles relating to the use of AI.

EXPLANATION

The development and use of AI should be grounded in ethical principles that prioritize respect for the dignity of persons and peoples, moral rights, human well-being, fairness, and responsibility. Psychologists should be aware of, understand, and apply research that provides insight into human decision-making, cognitive biases, and social behavior as it relates to AI and whether its use is indicated or contraindicated. Psychologists should understand how AI systems can be designed to align with social values and avoid unintended harm.

RELEVANCE FOR PSYCHOLOGISTS

For psychologists, this principle is highly relevant because their expertise in human behavior and ethics positions them not only to recognize and utilize, but also to contribute to the creation of, AI systems that simultaneously respect human dignity and complement efforts to improve mental health. To achieve this, psychologists should first be aware of the ethical implications of AI in their own specific professional contexts. Psychologists can advise on potential psychological impacts of AI deployment, such as effects on user trust, dependency, or decision-making, and ensure that ethical safeguards reflect real world human concerns. The use of AI in health services settings should be underpinned by an understanding of ethical guidelines directly relevant to AI coupled with traditional psychological guidelines for practice. Psychologists have a responsibility to be aware of ethical gaps and limitations that may be present in ethical use of AI and take steps to augment these.

REFERENCES

American Psychological Association. (2024). Ethical guidance for AI in the professional practice of health service psychology. https://www.apa.org/topics/artificial-intelligence-machine-learning/ethicalguidance-professional-practice.pdf

European Federation of Psychologists' Associations. (n.d.) Digitalisation priorities. https://www.efpa.eu/digitalisation

Resnik, D.B., & Hosseini, M. (2025). The ethics of using artificial intelligence in scientific research: new guidance needed for a new tool. AI and Ethics, 5, 1499–1521. https://doi.org/10.1007/s43681-024-00493-8

Wu, J. (2024, November 4). Social and ethical impact of emotional AI advancement: The rise of pseudo-intimacy relationships and challenges in human interactions. Frontiers in Psychology, 15. https://doi.org/10.3389/fpsyg.2024.1410462

PRINCIPLE

2

TRANSPARENCY IN SELECTION OF AI TOOLS

Psychologists should use only those AI tools that offer transparency regarding their development and provide clear, conspicuous, and persistent disclosure to users that they are interacting with a nonhuman entity. When full transparency is unavailable, psychologists should document known limitations and seek available technical documentation about data practices.

EXPLANATION

Transparency in AI requires systems to provide clear explanations of the data they use, how that data was collected, and the processes underlying algorithmic decisions. Without such clarity, the use of AI tools risks making it difficult for users and professionals to assess accuracy, fairness, and accountability. AI transparency tools can bridge the gap between complex technical processes and user understanding.

RELEVANCE FOR PSYCHOLOGISTS

For psychologists, transparency is critical in both research and applied settings. Psychologists must be able to understand how AI tools generate outputs to evaluate their validity and reliability in clinical, organizational, or research contexts. Transparency also allows psychologists to explain AI outputs clearly to clients, thereby maintaining trust and professional integrity.

REFERENCES

Resnik, D. B., & Hosseini, M. (2026). Disclosing artificial intelligence use in scientific research and publication: When should disclosure be mandatory, optional, or unnecessary? Accountability in Research, 33(2).

PRINCIPLE

3

INFORMED CONSENT

Psychologists who incorporate AI-based tools into their professional activities should clearly advise relevant parties about potential risks and benefits. Consent must be explicit regarding limitations of AI tools, must include opt-in for use of individual data, and must not use deceptive design patterns.

EXPLANATION

Incorporating AI tools into research, training, educational settings, or psychological practice requires obtaining informed consent when data may be collected, stored, or analyzed. Additionally, any use of AI tools that informs decision-making, supports service delivery, or influences interactions with clients, students, participants, or other stakeholders must be clearly disclosed.

Informed consent in AI development requires ensuring that individuals whose data is used, as well as those who interact with AI systems, are aware of how their information is collected, processed, and applied. Additionally, when AI tools are incorporated into professional practice, disclosure to clients is necessary to maintain transparency and autonomy. This principle emphasizes respect for individuals' rights and the ethical imperative of choice.

RELEVANCE FOR PSYCHOLOGISTS

For psychologists, informed consent is already a cornerstone of ethical practice, and AI adds new dimensions to this responsibility. Psychologists must communicate clearly with clients about the role of AI tools in assessments, interventions, or data handling, ensuring clients can make informed decisions about their participation. This principle safeguards client trust and helps psychologists remain compliant with professional ethical standards.

REFERENCES

American Psychological Association. (2017). Ethical principles of psychologists and code of conduct. https://www.apa.org/ethics/code

European Federation of Psychologists' Associations AISBL. (2025). Meta-code of ethics (https://www.efpa.eu/meta-code-ethic

PRINCIPLE

4

ACCOUNTABILITY, GOVERNANCE, AND PROFESSIONAL RESPONSIBILITY

Psychologists who develop, deploy, or rely on AI systems share professional accountability for their appropriate use and outcomes alongside other stakeholders and should advocate for clear governance structures that distribute responsibility appropriately.

EXPLANATION

The use of AI in psychology does not transfer ethical or professional responsibility from human practitioners to technological systems. Regardless of the level of automation involved, psychologists retain accountability for decisions informed or supported by AI tools.

AI systems may involve multiple stakeholders, including developers, vendors, institutions, and policymakers, which can obscure responsibility when harms occur. Psychologists should therefore seek clarity regarding how AI systems function, who controls their deployment, and how decisions are made and reviewed. Accountability requires understanding not only what an AI system does, but also how and why it produces particular outputs.

Psychologists should support governance structures that promote transparency, documentation, auditability, and mechanisms for oversight and redress. This includes being prepared to explain AI-informed decisions to clients, participants, or stakeholders, and to intervene or override AI outputs when professional judgment indicates concern.

Establishing clear accountability frameworks at an individual and organizational professional level helps prevent over-reliance on automated systems and reinforces the central role of human ethical reasoning in psychological practice.

RELEVANCE FOR PSYCHOLOGISTS

Accountability is essential for maintaining public trust in psychological science and practice. When AI tools are introduced without clear governance, responsibility may become diffuse, increasing the risk of ethical lapses, unchallenged errors, or harm without recourse.

For psychologists operating in diverse institutional and regulatory environments, explicit attention to accountability ensures alignment with professional standards, legal obligations, and ethical codes. This principle reinforces that AI is a tool within psychological practice–not a substitute for professional responsibility, judgment, or ethical deliberation.

REFERENCES

American Psychological Association. (2017). Ethical principles of psychologists and code of conduct. https://www.apa.org/ethics/code
European Commission. (2019). Ethics guidelines for trustworthy AI.

European Commission. (2019). Ethics guidelines for trustworthy AI

Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., Luetge, C., Madelin, R., Pagallo, U., Rossi, F., Schafer, B., Valcke, P., & Vayena, E. (2018). AI4People–An ethical framework for a good AI society. Minds and Machines, 28(4), 689–707. https://doi.org/10.1007/s11023-018-9482-5

Organisation for Economic Co-operation and Development. (2019). OECD principles on artificial intelligence.

PRINCIPLE

5

PROFESSIONAL COMPETENCE AND AI

EVIDENCE BASE

Psychologists who integrate AI into professional activities should possess sufficient foundational understanding of how their selected AI tools function, including their general methodology, data requirements, and known limitations relevant to psychological applications

EXPLANATION

Professional competence in the use of AI requires that psychologists possess sufficient foundational knowledge to critically understand how AI tools function, including their underlying theoretical assumptions, data inputs, and methodological limitations. This includes familiarity with the psychological, computational, and statistical principles relevant to the AI systems they select and apply within their professional roles.

Competence in this context does not imply technical expertise in developing AI systems, but rather the ability to make informed, reflective judgments about when, how, and whether AI tools are appropriate for specific professional purposes. Psychologists should remain attentive to the current state of empirical research regarding the effectiveness, risks, and limitations of AI-supported practices, and should engage in ongoing education to ensure that their use of AI remains aligned with evolving scientific knowledge and professional standards.

RELEVANCE FOR PSYCHOLOGISTS

As students of the human experience, the primary aim of psychologist researchers and practitioners is to improve lives. AI can help psychologists better understand individual, group, or societal functioning, only when they themselves are professionally capable of using the technologies. Psychologists who are competent in the employment of AI tools as complementary to science and practice must understand their design, stated intent, and effectiveness in order to protect the public from potential harms as well as to maximize productive health outcomes.

REFERENCES

American Psychological Association. (2023, July 1). AI is changing every aspect of psychology. Here's what to watch for. Monitor on Psychology, 54(5). https://www.apa.org/monitor/2023/07/psychology-embracing-ai

American Psychological Association. (2024). Ethical guidance for AI in the professional practice of health service psychology. https://www.apa.org/topics/artificial-intelligence-machine-learning/ethicalguidance-ai-professional-practice

Cruz-Gonzalez, P., He, A. W.-J., Lam, E. P., Ng, I. M. C., Li, M. W., Hou, R., … Sánchez Vidaña, D. I. (2025). Artificial intelligence in mental health care: a systematic review of diagnosis, monitoring, and intervention applications. Psychological Medicine, 55, e18. doi:10.1017/S0033291724003295

PRINCIPLE

6

HUMAN OVERSIGHT AND PROFESSIONAL JUDGMENT

Psychologists who incorporate AI tools into their professional activities must have sufficient expertise and discretion to ensure appropriate oversight of AI-supported outcomes.

EXPLANATION

The use of AI in psychology should not replace professional judgment or human responsibility. When AI tools are employed to support psychological work– such as decision-making, assessment, intervention planning, evaluation, or training– psychologists remain accountable for interpreting, contextualizing, and acting upon AI-generated outputs.

Human oversight is essential to ensure that AI-supported processes are applied appropriately, ethically, and in ways that respect the complexity of human behavior and lived experience. This includes monitoring AI outputs for plausibility, relevance, and potential unintended consequences, as well as intervening when automated recommendations conflict with professional judgment, ethical standards, or contextual considerations.

This principle applies across professional settings and populations, recognizing that AI may be used not only in clinical care but also in educational assessment, organizational decision-making, research workflows, and other domains of psychological practice. Maintaining meaningful human oversight helps safeguard the quality, fairness, and integrity of psychological work while preventing over-reliance on automated systems.

RELEVANCE FOR PSYCHOLOGISTS

In psychological research and practice in applied settings, human supervision is essential to follow and apply ethical principles. This aspect is especially important for the analysis of research data, because without this oversight, conclusions may be drawn based on automated patterns that overlook the context in which the research was conducted. In practice, human oversight is necessary to mitigate and limit “automation bias” where disregard for nuanced client data may produce outputs that potentially result in biased recommendations and potential harm to clients. This ensures that the psychologist is the final decision maker, preventing an over-reliance on technology and ensuring that AI only assists with–but does not substitute for–professional judgment. This oversight allows psychologists to maintain accountability, ethical standards, and the integrity of the psychological process.

REFERENCES

American Psychological Association. (2024). Ethical guidance for AI in the professional practice of health service psychology. https://www.apa.org/topics/artificial-

Babu, A., & Joseph, A. P. (2024). Artificial intelligence in mental healthcare: transformative potential vs. the necessity of human interaction. Frontiers in Psychology, 15, 1378904. https://doi.org/10.3389/fpsyg.2024.1378904

European Commission. (n.d.) Article 14: Human oversight | EU Artificial Intelligence Act. https://artificialintelligenceact.eu/article/14/

Özer M. (2024). Is artificial intelligence hallucinating? Turkish Journal of Psychiatry, 35(4), 333–335. Advance online publication. https://doi.org/10.5080/u27587

PRINCIPLE

7

PSYCHOLOGICAL SAFETY AND HARM PREVENTION

Psychologists who employ AI systems should prioritize maximize benefit alongside psychological safety, and actively work to prevent potential harms.

EXPLANATION

The integration of AI into psychological work introduces new forms of risk that may affect individuals' well-being, autonomy, and sense of agency. Psychologists have a responsibility to anticipate, identify, and mitigate potential psychological harms associated with the use of AI systems, including emotional distress, confusion, stigmatization, or inappropriate influence on thoughts, behaviors, or decisions.

This principle applies across contexts in which AI may shape psychological experiences, such as assessments, recommendations, feedback systems, learning environments, organizational settings, or research participation. Psychologists should carefully consider how AI outputs are framed, communicated, and acted upon, particularly when interacting with vulnerable populations or when AI systems generate content that may be perceived as authoritative or personalized.

Preventing harm requires ongoing monitoring of AI-supported processes, responsiveness to unintended consequences, and readiness to modify or discontinue AI use when risks outweigh benefits. By centering psychological safety, psychologists help ensure that AI technologies support human well-being rather than undermine it.

RELEVANCE FOR PSYCHOLOGISTS

Psychologists are charged with addressing human behavior in its most sensitive and unique contexts. Achieving expertise in the science and practice of psychology is the result of years of rigorous training and supervised practice. Effective therapeutic intervention relies on evidence-based approaches and contextually oriented applications built on a foundation of care and trust. Psychologists integrating AI into their work with clients must be cognizant of the rapidly evolving research literature as well as clinical insights associated with both the potential and the limits of AI tools and therapeutics. The psychologist is ultimately responsible for AI indicators and outputs that are integrated into any process associated with health services interventions (e.g., hallucinations by note-taking tools, analytics for tone of voice/patient themes, flags indicating relapse or suicidal ideation). Psychologists possessing AI competence and an understanding of quality care that includes AI can safeguard users (and themselves) while maintaining the integrity of the profession by prioritizing patient well-being, promoting accuracy, sustaining trust, and taking care to avoid harm.

REFERENCES

Abrams, Z. (2025, June 26). Artificial intelligence is reshaping how psychologists work. American Psychological Association Services, Inc. https://www.apaservices.org/practice/news/artificial-intelligence-psychologists-work

American Psychological Association. (2025, November). Health advisory: Use of generative AI chatbots and wellness applications for mental health. https://www.apa.org/topics/artificial-intelligence-machinelearning/health-advisory-chatbots-wellness-apps

British Psychological Society. (2024, October 9). AI and the work of psychologists: Practical applications and ethical considerations. https://www.bps.org.uk/blog/ai-and-work-psychologists-practical-applicationsand-ethical-considerations

Putica, A., Khanna, R., Bosl, W., Saraf, S., & Edgcomb, J. (2025, July 24). Ethical decision-making for AI in mental health: The Integrated Ethical Approach for Computational Psychiatry (IEACP) framework. Psychological Medicine, 55, e213. https://doi.org/10.1017/S0033291725101311

PRINCIPLE

8

ACCESS, BIAS, AND INCLUSION WITH AI

VALIDATION AND SCIENTIFIC EVIDENCE

Psychologists who use AI tools should critically evaluate whether such systems are supported by rigorous, transparent, and contextually appropriate scientific evidence that demonstrates reliability, validity, and relevance for the populations and purposes for which they are used.

EXPLANATION

AI systems used in psychological contexts must be supported by robust scientific evidence consistent with established standards of psychological research and practice. Psychologists have an ethical and professional responsibility to ensure that AI tools are not treated as neutral or self-validating technologies, but as instruments whose outputs depend on underlying models, data, and assumptions that require systematic evaluation.

Validation involves assessing whether AI systems demonstrate adequate reliability, validity, and accuracy for their intended psychological purposes, as well as whether their performance generalizes across cultural, linguistic, demographic, and contextual variations. Evidence derived from one population, setting, or use case may not be transferable to others, particularly in global or low-resource contexts.

Psychologists should critically evaluate available documentation, peer-reviewed evidence, and validation studies, and remain cautious when evidence is limited, proprietary, or inaccessible. When empirical support is insufficient, AI tools should be used only in exploratory, supportive, or research-oriented roles, rather than as bases for decision-making that may affect individuals' rights, well-being, or opportunities.

By grounding AI use in scientific validation, psychologists help preserve the integrity of psychological science and protect individuals from untested, misleading, or overstated technological claims.

RELEVANCE FOR PSYCHOLOGISTS

The global expansion of AI in mental health, education, and organizational settings has outpaced the development of universally accepted validation standards. Many commercially available tools lack independent evaluation, rely on unrepresentative datasets, or provide insufficient documentation for users to assess their appropriateness.

Without rigorous validation, AI tools may produce inaccurate or unstable outputs, generate biased predictions that disproportionately affect certain groups, undermine clinical or educational decision-making, and erode public trust in psychological science and technology-supported care. Ensuring validation is therefore essential for protecting individuals and communities, promoting equitable outcomes, and maintaining the scientific integrity of psychological practice worldwide.

REFERENCES

Terry, J., Strait, G., Alsarraf, S., Weinmann, E., & Waychoff, A. (2025). Artificial intelligence in scale development: Evaluating AI-generated survey items against gold-standard measures. Current Psychology, 44, 16339–16350. https://doi.org/10.1007/s12144-025-08240-w

PRINCIPLE

9

ALGORITHMIC BIAS AND FAIRNESS

Psychologists who employ AI systems should proactively identify, evaluate, and mitigate algorithmic bias, ensuring that AI-supported outputs do not systematically disadvantage or discriminate against individuals or groups based on demographic, cultural, socioeconomic, or contextual factors.

EXPLANATION

AI systems used in psychological contexts may reproduce, amplify, or obscure existing social and structural inequalities embedded in training data, model design, or deployment contexts. Algorithmic bias can emerge at multiple stages of the AI lifecycle, including data collection, feature selection, model training, validation, and real-world implementation.

Psychologists have a responsibility to recognize that biased outputs are not merely technical errors, but may reflect historical inequities related to race, gender, ethnicity, language, disability, age, socioeconomic status, or geographic location. These biases can lead to unfair assessments, unequal access to services, or disproportionate harms for already marginalized populations.

To promote fairness, psychologists should critically examine whether AI systems perform equitably across diverse populations and contexts, and whether outcomes differ systematically between groups. This includes questioning whose data is represented, whose experiences are missing, and how algorithmic decisions may interact with existing power asymmetries in psychological practice, research, and policy.

When bias is identified or suspected, psychologists should advocate for mitigation strategies such as data diversification, model recalibration, transparency about limitations, or discontinuation of AI use. Fairness in AI is not a one-time achievement but an ongoing process requiring monitoring, reflection, and interdisciplinary collaboration.

RELEVANCE FOR PSYCHOLOGISTS

Algorithmic bias directly threatens core ethical principles of psychology, including justice, beneficence, and respect for persons and peoples. In applied settings–such as assessment, diagnosis, intervention planning, risk evaluation, or digital mental health–biased AI systems may reinforce disparities rather than reduce them.

For psychologists working in global, multicultural, or low-resource contexts, unexamined algorithmic bias can lead to inappropriate generalizations, cultural misinterpretations, or exclusion of populations whose data are underrepresented in dominant AI systems. Addressing bias is therefore essential to ensuring that AI-supported psychological practices remain equitable, culturally responsive, and aligned with professional ethical standards.

REFERENCES

American Psychological Association. (2017). Ethical principles of psychologists and code of conduct. https://www.apa.org/ethics/code

Barocas, S., Hardt, M., & Narayanan, A. (2019). Fairness and machine learning. https://fairmlbook.org/

Benjamin, R. (2019). Race after technology: Abolitionist tools for the new Jim Code. Polity Press.

Mehrabi, N., Morstatter, F., Saxena, N., Lerman, K., & Galstyan, A. (2021). A survey on bias and fairness in machine learning. ACM Computing Surveys, 54(6), 1-35. https://doi.org/10.1007/s12144-025-08240

PRINCIPLE

10

ACCESS, EQUITY, AND INCLUSION

Psychologists who integrate AI tools into their professional activities should promote equitable access and inclusive design, ensuring that AI-supported psychological services do not exacerbate existing disparities or exclude individuals and communities based on structural, economic, linguistic, or technological barriers.

EXPLANATION

The integration of AI into psychological practice has the potential to expand access to services, knowledge, and interventions. However, without deliberate attention to equity and inclusion, AI systems may also reinforce or deepen existing disparities across regions, populations, and social groups.

Barriers related to cost, infrastructure, digital literacy, language availability, disability, and connectivity can limit who benefits from AI-enabled psychological tools. In many global contexts–particularly in low- and middle-income countries, rural areas, or marginalized communities–these barriers may systematically exclude individuals who could otherwise benefit from psychological support.

Psychologists should critically assess whether AI tools are accessible, usable, and culturally appropriate for the populations they serve. This includes considering language diversity, disability access, contextual relevance, and the broader socio-technical conditions under which AI systems are deployed. Equity in AI is not achieved solely through technical performance, but through intentional design and implementation that accounts for unequal starting conditions.

Psychologists are encouraged to advocate for inclusive development practices, transparency about access limitations, and alternative non-AI pathways when technological solutions are not appropriate or feasible. Ensuring equitable access requires ongoing reflection on who is included, who is excluded, and why.

RELEVANCE FOR PSYCHOLOGISTS

Access and inclusion are foundational ethical concerns in psychology. AI systems that are inaccessible or poorly adapted to diverse contexts risk privileging already advantaged groups while marginalizing others. For psychologists working across cultural, national, and socioeconomic boundaries, attention to access is essential to uphold principles of justice and social responsibility.

In applied practice, research, education, and policy contexts, psychologists play a key role in evaluating whether AI-supported tools genuinely broaden access or merely shift services toward populations with greater technological resources. Addressing access and inclusion helps ensure that AI contributes to reducing–not reproducing–global and local inequalities in psychological care.

REFERENCES

American Psychological Association. (2017). Ethical principles of psychologists and code of conduct. https://www.apa.org/ethics/code

D'Ignazio, C., & Klein, L. F. (2020). Data feminism. MIT Press.

United Nations. (2015). Sustainable Development Goals. https://sdgs.un.org/goals

World Health Organization. (2021). Ethics and governance of artificial intelligence for health. https://www.who.int/publications/i/item/9789240029200

GLOSSARY

Algorithmic bias: Systematic and repeatable patterns in AI systems that produce unfair outcomes, typically disadvantaging underrepresented groups due to biased training data, design choices, or deployment contexts.

Artificial intelligence (AI): Technology that enables computers and machines to simulate human learning, comprehension, problem solving, decision-making, creativity, and autonomy by relying on neural networks and algorithms.

Automation bias: The tendency to over-rely on automated recommendations, sometimes disregarding contradictory human judgment or contextual information.
Black box AI: Systems whose decision-making processes are not interpretable.

Digital tools: Electronic devices, websites, software, and applications that rely on technology to complete tasks, organize information, streamline processes and/or effectively communicate. Digital tools rely on electronics and digital systems. Examples include hardware tools (smartphone, laptop) and software tools (programs such as Word or Excel and online services such as search engines).

Digitalization: Conversion and adaptation of a system or process to be operated with the use of computers and the internet.
Explainability/interpretability: The degree to which AI decision can be understood by humans.

Generative AI: AI systems that create new content (text, images, etc.).

Hallucination: A phenomenon where an AI model generates convincing but false or unfounded information.

Intervention: A defined technique or structure utilized to create positive change in human thought, behavior, emotion, relationships, or overall well-being and healthy living. A psychological intervention is based on theoretical and empirical science.

Large language model (LLM): A type of AI trained on vast amounts of text and image data to understand, summarize, and generate human-like language.

Machine learning: The subset of AI where systems improve through data exposure.

Neural network: Computational architecture inspired by biological neurons.

GLOSSARY, ctd.

Psychological services: The full range of interventions and activities undertaken by psychology professionals, spanning all fields of psychology across diverse settings. These services include but are not limited to approaches based on research, clinical/counseling/school service delivery, health, prevention, school, guidance counseling, and public health, aimed at improving understanding and working with individuals, couples, families, groups, societies, and technology.

Training data: The initial dataset used to teach a machine learning algorithm how to process information and make predictions.


Asociación Nacional de
Profesionales en Ciencias
Cognitivas y del Comportamiento


Association of Psychologists in Nepal
APN
2016


aysen
ASSOCIATION DES
PSYCHOLOGUES DU
SENEGAL


Cci
Counsellor Council of India


efpsa
European Federation of
Psychology Students' Associations


საქართველოს ფსიქოლოგთა
პროფესიული კავშირი


GPA


IARR
International
Association for
Relationship
Research


THE KOREAN PSYCHOLOGICAL ASSOCIATION
1946


SEPs
Sociedad
Ecuatoriana de
Psicoterapia


.UBA PSICOLOGÍA
Facultad de Psicología


ALAMOC
Asociacion Latinoamericana de Analisis,
Modificacion del Comportamiento y
Terapia Cognitiva Conductual
Intergrante de
WORLD CONFEDERATION
WCCBT


UNSA


NIGERIAN PSYCHOLOGICAL ASSOCIATION
NPA


СИМБОЛИКУМ


COLEGIO COLOMBIANO DE
PSICÓLOGOS
COL
PSIC


MEPA
MIDDLE EAST
PSYCHOLOGICAL ASSOCIATION


mental
health
awareness
initiative


COLLEGE PSYCHOLOGY ASSOCIATION OF VILLA


SRI LANKA
PSYCHOLOGICAL
ASSOCIATION


PSYCHOLOGICAL
SOCIETY OF KENYA
In Service to Humanity


POLISH PSYCHOLOGICAL ASSOCIATION


AMERICAN
PSYCHOLOGICAL
ASSOCIATION


Singapore
Psychological
Society
(Established 1979)


TAIWANESE
PSYCHOLOGICAL
ASSOCIATION


TANZANIAN
PSYCHOLOGICAL
ASSOCIATION


ZPA
Zimbabwe
Psychological
Association


الجمعية السعودية
لعلم النفس
Saudi Psychological Association


the british
psychological society
promoting excellence in psychology


ASOCIACIÓN MEXICANA
DE PSICOLOGÍA Y
DESARROLLO COMUNITARIO


IPN
INDIAN
PSYCHOLOGISTS
NETWORK


Czech Psychology Network for Global Changes


UCCAP
Unión Centroamericana de Colegios y
Asociaciones de Psicología
Guatemala | Honduras | El Salvador | Nicaragua | Costa Rica | Panamá



GLOBAL
PSYCHOLOGY
ALLIANCE

Psychology for Global Challenges