



**Te Poari Kaimātai
Hinengaro o Aotearoa**

NEW ZEALAND PSYCHOLOGISTS BOARD

Guidelines for the use of Artificial Intelligence in Psychology in Aotearoa New Zealand

Draft - April 2024

Acknowledgements

Drafting guidelines for a health profession takes a village approach. It is important that the guidelines are culturally, ethically, and legally sound, are understandable and most of all useful for the field. A range of practitioners and advisors contributed to the development of the Guidelines for the use of Artificial Intelligence (**AI**) in Psychology in Aotearoa New Zealand. Te Poari Kaimātai Hinengaro o Aotearoa / New Zealand Psychologists Board are very grateful to Vijaya Dharan, Rosie Dobson, Kobus Du Plooy, Desiree Grant, Jessica Gu, Thomas Jenkin, Claire Komatas, Tom Nesor, Chris Neuenfeldt, Briony Pentecost and Aroha Waipara-Panapa for their assistance in drafting these guidelines. Thank you to Dr Karaitiana Taiuru for his permission to include his Te Tiriti based ethical principles for artificial intelligence in the guidelines. Brian Emery (cultural advisor to Te Poari) was generous in his feedback regarding Māori data sovereignty. Tūmāia Kaiārahi were also critical in the feedback stage and provided valuable input, both collectively and individually, to produce the final version of this document. Claro Law provided helpful legal guidance and feedback regarding the information contained in these guidelines.

Beginning the Korero: Purpose of the Guidelines

1. The Health Practitioners Competence Assurance Act (**the HPCA Act**) mandates Te Poari Kaimatai Hinengaro o Aotearoa / New Zealand Psychologists Board (**the Board**) to assure the public that psychologists are fit to practise and that they provide competent, high quality and safe services. To meet these obligations, the Board has adopted the Code of Ethics for Psychologists Working in Aotearoa/New Zealand 2002 (**the Code of Ethics**; developed in conjunction with the New Zealand Psychological Society and the New Zealand College of Clinical Psychologists) as a guide to ethical practice. The Code of Ethics delineates the manner or the tikanga in which psychologists ought to carry out their practice. All other statements, including the current guidelines, of how psychologists should conduct their practice must be consistent with the Code of Ethics and its ethical principles of respect for the dignity of persons, responsible caring, integrity in relationships, and responsibility to society.
2. Guidelines adopted by the Board support psychologists in providing competent and ethical practice by translating or expanding on the Code of Ethics in relation to more specific aspects of their professional behaviour. Guidelines are not definitive, binding, or enforceable by themselves. They have the least authority of any of the regulatory documents. However, a disciplinary or review body may use the guidelines in evaluating a psychologist's knowledge and competency. Guidelines are recommendations rather than mandatory standards but supplement the Code of Ethics which is the highest and most aspirational regulatory document. In addition, the Board's intent in developing guidelines is to assist practitioners in delivering "best practice" both for the safety of the public, and to lessen the likelihood of practitioners facing complaints about their practice from those with whom they work.
3. The central purpose of the Guidelines for the use of AI in Psychology in Aotearoa New Zealand (**the Guidelines**) is to promote practice consistent with ethical principles. Every AI scenario must be taken on its own merits by the psychologist, with critical thought and judgement applied to each individual circumstance. It is not possible for the Guidelines to provide an exact answer of what to do in the infinite number of possible scenarios and contingencies that might exist regarding AI. However, it is intended that the Guidelines will assist psychologists in their critical thought and judgement that they will apply to AI scenarios which may arise in their practice, research, teaching, or supervision.

Firstly, what is Artificial Intelligence?

Before the Guidelines are presented, the following brief definitions of artificial intelligence (AI) are offered:

- Put simply, AI is an umbrella term for technology that enables computers and machines to simulate human intelligence and problem-solving capabilities. This allows such technologies to perform tasks previously requiring human intervention. Digital assistants, autonomous vehicles, and generative AI tools (e.g., ChatGPT) are some well-known examples of AI in our daily lives. AI uses machine learning, involving the development of algorithms that ‘learn’ from available data, arriving at increasingly more accurate predications or decisions over time.
- A more technical definition of AI from a High-Level Experts Group of the European Commission is: “systems designed by humans that, given a complex goal, act in the physical or digital world by perceiving their environment, interpreting the collected structured or unstructured data, reasoning on the knowledge derived from this data and deciding the best action(s) to take (according to pre-defined parameters) to achieve the given goal. AI systems can also be designed to learn to adapt their behaviour by analyzing how the environment is affected by their previous actions” (AIHLEG, 2018).
- AI has been around since the 1960s and has become increasingly sophisticated since 2019. Initially, AI systems focused on performing a specific task by following rules to analyse data and make predictions based on pattern recognition but did not technically create anything new (playing against a computer at chess would be an example of an early AI system). Generative AI is a more recent form of AI that can create something new (e.g., text, imagery, audio, video) from existing large data sets by learning underlying patterns to generate new pattern creation. ChatGPT is a well-known example of a large language model AI system where algorithms use enormous data sets to understand, summarise, and generate new content.
- To be clear, the current Guidelines use the term artificial intelligence (AI) to broadly cover any system where one or more of the following applies:
 - ✓ Machine learning systems – that process training data
 - ✓ Classifier systems – that are used to put information into categories
 - ✓ Interpreter systems – that turn input data into standardised outputs
 - ✓ Generative systems – that are used to create new text, images, audio, or video.
 - ✓ Automation – where computers take on tasks that people have hitherto done (e.g., driving)

What are the Risks and Benefits of Artificial Intelligence for Psychology?

AI can significantly augment the breadth of knowledge in fields pertinent to psychology, presenting practitioners with an expanded understanding previously unexplored in depth. This enhancement has the potential to notably increase a psychologist's expertise. However, the integration of AI in psychological practice is not without its risks and challenges. These include:

- **Variable Depth of Knowledge:** While AI offers commendable introductory insights, it often lacks a nuanced understanding of specific psychology-related fields. The effectiveness of AI is contingent on the quality and breadth of the data it processes.
- **Bias and Inaccuracy:** Identifying biases within AI algorithms can be challenging. Instances where AI inadvertently alters demographic characteristics of historical figures highlight the risk of incorporating unintentional inaccuracies, emphasizing the need for transparency and critical evaluation.
- **Subtle Biases:** The subtlety of biases in AI-generated content can complicate their identification, especially in areas unfamiliar to the practitioner. Awareness and critical scrutiny are essential in mitigating this issue.
- **Misinformation and Ethical Misuse:** The deliberate use of AI to fabricate or distort information, particularly with the intent of manipulating public opinion or for personal benefit, is ethically indefensible.
- **Plagiarism and Intellectual Honesty:** Employing AI to derive and summarize research, while ethically acceptable, contrasts sharply with presenting AI-generated content as one's original work. This distinction underscores the imperative for explicit ethical guidelines in the use of AI.
- **Accountability:** It is crucial that all academic users, including psychology students, employ AI responsibly, particularly in ensuring academic integrity when leveraging AI for coursework or research.

Conversely, when used responsibly, AI offers considerable potential benefits to a psychologist, including:

- Reviewing and synthesizing extensive information sets.
- Generating novel ideas and identifying gaps in research, alongside outlining potential research methodologies.

- Assisting in educational settings as a tutoring or coaching tool.
- Facilitating the clarification of diagnoses and the formulation of treatment plans.
- Assisting in administrative tasks such as handouts and letters.

The responsible employment of AI can significantly enhance the efficiency and effectiveness of psychological practices. Nonetheless, psychologists must remain vigilant regarding evolving ethical and moral challenges as AI technology progresses. The current Guidelines are intended to assist psychologists in navigating these challenges.

The Guidelines at a Glance

The Guidelines have been structured into ten overarching principles for the psychologist considering the use of AI in their practice of psychology. The principles are listed below, with more detail provided under each principle on the following pages.

1. Psychologists to consider the unique cultural context of Aotearoa New Zealand when using AI in their practice.
2. Psychologists are encouraged to view their obligations in understanding an AI tool as similar to those when using psychometric measures.
3. Psychologists to consider their obligations under Te Tiriti and the principles of Māori Data Sovereignty when using AI tools.
4. Psychologists to consider the potential biases in results from an AI tool and avoid perpetuating any form of discrimination based on biased data sets.
5. Psychological services and opinions that a registered psychologist offers should not be exclusively delegated to AI.
6. Psychologists should only use AI tools that they have assessed as being ethically robust and transparent about the parameters of data sharing.
7. Psychologists should consider the principles of the Privacy Act 2020 and a privacy impact assessment when using AI tools in their work.
8. Psychologists should be transparent about the use of AI and inform those with whom they work that they are using AI if the AI tool performs a significant part of the service, and/or if the person(s) would reasonably expect to know when consenting to a service.
9. The onus is on the psychologist to ensure the person(s) with whom they are working understands the role of any AI, to the extent the person requires to make an informed choice at the outset of a psychological service or research.
10. Person(s) working with a psychologist should not be disadvantaged if they do not wish to have AI tools used in their care or their data entered into an AI system.

1. Psychologists to consider the unique cultural context of Aotearoa New Zealand when using AI in their practice

- 1.1. In the rapidly evolving landscape of technology, artificial intelligence (AI) brings both unique opportunities and complex challenges. As psychologists in Aotearoa New Zealand embark on considering the use of AI in their practice, it has become important for the Board to publish this set of guidelines that not only aligns with the Code of Ethics and other relevant health legislation but is also embedded within the principles of the Ti Tiriti o Waitangi and considers the unique cultural context of the country. The cultural tapestry of Aotearoa New Zealand, with its rich history and diverse communities, demands a nuanced approach to the implementation of AI in psychological practice. Therefore, while existing guidelines from other parts of the world were considered in the development of the current guidelines, careful consideration was paid towards ensuring that it is appropriate for the unique Aotearoa New Zealand context.
- 1.2. As AI continues to weave itself into all spheres of our everyday lives, psychologists find themselves at the crossroads of innovation and ethical responsibility. This juncture necessitates careful consideration of how AI can be ethically and effectively integrated into psychological frameworks and services, ensuring that its benefits are maximized while minimizing potential risks.
- 1.3. At the heart of the current guidelines lies a commitment to Ti Tiriti o Waitangi, cultural competence, and responsiveness. Recognizing the rich diversity embedded in Aotearoa's cultural fabric, psychologists must ensure that AI systems are not only free from biases but also resonate with the values and perspectives of all the communities they serve. This requires engagement with Māori to hear their views regarding AI. The principles of transparency, informed consent, and data guardianship serve as foundational pillars, safeguarding the privacy and trust essential for effective therapeutic relationships.
- 1.4. The integration of AI, when approached with critical reflection and dedication to cultural sensitivity can become a catalyst for positive change, reflecting the enduring commitment of psychologists in Aotearoa to the holistic well-being of their clients and communities. As the future unfolds, these guidelines are a commitment to the profession's evolving responsibility in the face of continuing technological advancement.

2. Psychologists are encouraged to view their obligations in understanding an AI tool as similar to those when using psychometric measures.

2.1. Psychologists are well-versed in the selection, evaluation, and communication of psychometric measures and the data they produce. Psychometrics are a common source of data that psychologists in all scopes of practice use to assist in the outcomes for people we work with (e.g., in measuring baseline levels of a construct, assisting diagnostic opinion, evaluating progress, and assessing risk). Given the high frequency in which psychometrics are used in psychological practice, psychologists are socialised early in their professional training to only use those measures that are robust across technical and ethical domains.

2.2. The current guidelines propose that psychologists take a similar critical approach when considering the use of AI tools in their practice. While it is unrealistic (and unnecessary) for all psychologists to have an in-depth understanding of the technical development of an AI tool, they must have sufficient knowledge of the AI tool's:

- purpose and appropriateness in the specific context,
- use of data,
- accuracy,
- confidentiality and privacy mechanisms, and
- diversity and bias policies

and be satisfied that these are at a standard to both benefit and not cause harm to those with whom we work. Just as with psychometric measures, these pieces of information would also be the minimum a psychologist would need to know to potentially inform the people with whom they work during the consent process.

2.3. Psychologists have been taught to critically evaluate psychometric measures prior to use, and the current guidelines encourage psychologists to apply the same capacity for critical analysis to any AI tools available now or in the future, to safeguard against potential harms to people with whom we work.

2.4. The development of AI tools in all spheres of society is likely to significantly escalate in the coming years. It is not a defence for a health professional to say, in the event of a complaint regarding the use of an AI tool, that they were not aware that the tool had a particular use of data that they were not aware of. Just like with psychometrics, psychologists must only use those AI tools that publish sufficient information for the user to critically evaluate whether the tool meets ethical, legal and privacy standards and to communicate fully the purpose and limitations of the tool to the person(s) using their services.

3. Psychologists to consider their obligations under Te Tiriti and the principles of Māori Data Sovereignty when using AI tools.

3.1. Māori data is defined by the Waitangi Tribunal's WAI 2522 report as "Digital or digitizable information or knowledge that is about or from Māori people, language, culture, resources, or environments."

3.2. Māori data sovereignty refers to the inherent rights and interests of Māori, whānau, hapū, iwi and Māori organisations have in relation to the creation, collection, access, analysis, interpretation, management, dissemination, re-use and control of data relating to Māori, whānau, hapū, iwi and Māori organisations as guaranteed in He Whakaputanga and or Te Tiriti and the provided recognition of rights with the United Declaration of Rights of Indigenous Peoples (Taiuru, 2023). It can be further defined as Māori data governance. The principles, structures, accountability mechanisms, legal instruments, and policies through which Māori exercise control over Māori data.

3.3. Since 2016 there have been significant advancements for the protection and recognition of Māori data sovereignty, particularly with consideration of new technologies. This includes the WAI 2522 report decision that Māori data be recognised as taonga (treasure) imbued with Te Tiriti rights. This means that a psychologist should be particularly mindful when working with Māori in respecting the control and choice they have over their own Māori data.

3.4. Dr Karaitiana Taiuru is a leading authority and Māori technology ethicist. He offers the follow principles (Taiuru, 2023) for psychologists to consider when working with Māori and AI¹:

- Principle 1: Tino Rangatiratanga. AI systems to embed Māori leadership, decision-making and governance at all levels of the system's life cycle from inception, design, release to monitoring.
- Principle 2: Equity. AI systems to achieve equity outcomes for Māori.
- Principle 3: Active protection. Ensuring informed consent for the use of Māori data in AI development, with robust procedures in place to prevent biases or predictions that harm Māori.
- Principle 4: Mana whakahaere. Effective and appropriate stewardship or kaitiakitanga over AI systems is required. An understanding of the source and intended use of data is required, so that it is not repurposed without permission or in a way that will diminish the mana of Māori.

¹ Dr Taiuru has given his permission for these principles to be included in the Guidelines

- Principle 5: Mana Motuhake. This requires that tikanga are followed throughout the whole AI development and deployment cycles, with Māori deciding what data and data uses are controlled and allowed.
- Principle 6: Tapu/Noa – cultural safe practices. AI should not be culturally unsafe or break the rules of Tapu and Noa.

4. Psychologists to consider the potential biases in results from an AI tool and avoid perpetuating any form of discrimination based on biased data sets.

- 4.1. AI bias or machine learning bias refers to AI systems that produce results that reflect and perpetuate human biases within a society, including historical and current social inequalities.
- 4.2. There are two main reasons why an AI system may produce biased results. The first is the quality of the training or input data that the AI system has learned from. This data input can include skewed human decisions or reflect inherent inequities. The second source of potential inequality is the representation in the training data sets. For example, a data set may have groups of the population that are over or under-represented. An AI system and its results can only be as good as the quality of the data it was trained on. Given the data that an AI is trained on will always originate from a human mind or systems (with their range of biases), it is unclear if creating truly unbiased AI is technologically possible.
- 4.3. It is critical that psychologists understand that the potential for bias in results from an AI system exists, and that this can negatively impact society's trust in AI. Human judgement is also subject to the same potential for bias and, in the field of psychology at least, has gone through its own period of scrutiny and research into how to optimize both accuracy of prediction and minimise a range of biases. This can be seen in the rigorous development that goes into producing psychometric measures, for example.
- 4.4. Psychologists should approach the use of AI with the issue of potential bias in mind, just as they would look at normative data that a psychometric tool had been normed on. Doing this allows the psychologist to be informed about how applicable an AI system is for their needs. As with Principle 5 below, humans and machines need to work together for enhanced health outcomes in a "multidisciplinary approach" rather than machines replacing humans, or humans negating the positive impact machine learning can have for society.
- 4.5. Other practical strategies a psychologist can use to ensure that they are using AI systems that seek to minimise bias are:
 - Ensuring the AI tool has a bias/fairness policy and a commitment to investing in research to minimise bias in their AI tool.
 - The psychologist having an awareness of AI bias and ensuring they maintain oversight and critical analysis and review of any results produced by AI.
 - Where possible, reviewing the type of training data used by an AI tool to ensure the data is representative of diversity across domains.

5. Psychological services and opinions that a registered psychologist offers should not be exclusively delegated to AI.

- 5.1. Psychological services have relied heavily on human judgment, particularly in the infancy of the profession. As the profession evolved, a variety of objective tools (e.g., psychometrics) have been employed in the assistance of the psychologist offering services to a range of people. These tools have assisted psychologists, rather than replacing them in the overall services they provide.
- 5.2. It is clear there is a significant shortage of mental health professionals globally, and one of the potential benefits of AI is its capacity to increase access to mental health services and support. At the time the current guidelines were prepared there is already a growing market of AI tools that enable people to interact with an AI therapist via electronic devices. This is expected to increase significantly in the coming decade, with corresponding research into the efficacy of this mode of mental health intervention.
- 5.3. When a person interacts with an AI therapist tool it is reasonable to assume that the person understands that the service is with an AI, and that a human is not involved. However, when a person procures the services of a (human) psychologist, it is with the expectation that a human is leading the provision of that service. Psychologists are encouraged to consider the extent to which any AI tool is being used to replace their own critical thought and sure transparency around this. The use of AI should not exclusively replace a psychologist's service when that service is advertised or presented as a service offered by a human. In addition, particularly at this stage of the evolution of AI, there is the possibility of machine error and bias to occur, making the critical thought and review of a human an essential feature of psychological services.
- 5.4. In a report to OECD, Laukkonen et al (2019) point out that "AI is good at following rules, but many situations in life are embedded in specific contexts that change the rules. In such circumstances, the complex interactions between values, beliefs, and goals can provide axioms—and agency—for decision-making in thorny moral scenarios." Such decision making they emphasise will be unique to humans, as AI cannot be generalised nor is it wholly capable of dealing with *ambiguity*, *uncertainty*, and *complexity* of situations. Arguably, psychologists will need to exercise this sense of agency and professional knowledge in every step of their practice which involves complex human contexts and interactions.

6. Psychologists should only use AI tools that have been assessed as ethically robust and transparent about the parameters of data sharing.

Privacy is an important but non-absolute right in Aotearoa New Zealand

- 6.1. The emergence of AI highlights a tension between two competing interests: society's interest in the innovation of technology and individual interests related to privacy of information used in these technologies (Boniface, 2021). Too far in either direction could result in hindering the development of AI tools that could enhance health outcomes, or rendering the protections of privacy meaningless in action. What is the priority for our society? Research into the perspectives of health service users in Aotearoa New Zealand has shown support for AI in health but with clear conditions around intent, governance, privacy, security, transparency and restrictions on commercial gain (Dobson & Whittaker, 2023).
- 6.2. Privacy is not an absolute concept in law or ethics. In Aotearoa New Zealand, a health practitioners' code of ethics, in addition to the Code of Rights, the Privacy Act 2020 and the Health Information Privacy Code 2020, provide an individual's right to privacy and the execution of the parameters of this right. The legal protections related to privacy in Aotearoa New Zealand are designed to be flexible, and are presented as principles and ideals rather than rigid requirements (Boniface, 2021). However, as stated in the Office of the Privacy Commissioner's *Artificial intelligence and the Information Privacy Principles* (published in September 2023), the Privacy Act 2020 applies to the use of AI.

AI presents challenges to an individual's right to privacy

- 6.3. The nature of AI means that there are likely new and extensive possibilities for breaches of privacy. In order to make connections rapidly and accurately AI requires enormous datasets which could easily be shared between different systems. An individual's data may potentially be moved and manipulated in a variety of ways that cannot possibly be anticipated at the point of collection. AI makes it harder to see, understand and explain how personal information may be potentially used in the future (OPC Guidelines, 2023).
- 6.4. These features of AI present numerous challenges for privacy, including the impossibility to apply conventional methods of protection to large scale and global technological systems. Privacy protections rely on people and organisations who can understand the context and take responsibility for their actions (OPC Guidelines, 2023). The further an individual's data is from a human agent, the less likely a privacy breach could be anticipated or identified.

6.5. The Office of the Privacy Commissioner (**OPC**) has published a set of guidelines in September 2023 regarding the use of AI². The OPC guidelines remind the reader that in Aotearoa New Zealand privacy law applies to the use of AI tools. In our local context, this means that any use of AI needs to comply with the 13 information privacy principles provided in the Privacy Act 2020 and the parallel 13 health information rules in the Health Information Privacy Code 2020. These principles and rules govern the collection, storage, use and sharing of personal/health information. The OPC guidelines are clear that these principles must be upheld if you are building an AI tool, using an AI tool to support decision making or using AI in any aspect of your work with people.

6.6. The reader is encouraged to read the OPC Guidelines as an adjunct to the Board's Guidelines on the use of AI as they give guidance regarding adhering to each of the 13 information privacy principles. However, the following practical privacy points are specifically highlighted from the OPC Guidelines here:

- Have senior leadership approval based on full consideration of risks and mitigations of the AI tool
- Review whether a generative AI tool is necessary and proportionate given potential privacy impacts and consider whether you could take a different approach
- Conduct a privacy impact assessment before using these tools (discussed further in the next Principle 7)
- Be transparent, telling people how, when, and why the tool is being used and the potential implications of its use
- Engage with Māori about potential risks and impacts to the taonga of their information
- Develop procedures to ensure accuracy of information and access by individuals to their information
- Ensure human review prior to acting on AI outputs to reduce risks of inaccuracy and bias
- Ensure that personal information is not retained (in any form) or disclosed by the AI tool and that personal information is not used for secondary purposes.

6.7. Many psychologists work within organisations who will have their own AI guidelines and policies which psychologists should be familiar with.

² The OPC Guidelines indicate that updates to this document will occur as needed.

7. Psychologists should consider the principles of the Privacy Act 2020 and a privacy impact assessment when using AI tools in their work.

7.1. The OPC Guidelines make the useful point that there is much excitement and urgency around new AI tools, despite best practices for these tools yet to be fully developed. They urge people to consider privacy prior to using AI that may require the input of personal information. The OPC Guidelines state that they expect organisations to do a privacy impact assessment prior to the use of AI.

7.2. The OPC website offers guidance on writing and conducting a privacy impact assessment. A privacy impact assessment is a tool to help agencies identify and assess the privacy risks arising from their collection, use of and handling of personal information. Ways to mitigate or minimise these risks are also proposed in a privacy impact assessment. These assessments are particularly useful when an agency is considering introducing a new system (such as AI).

7.3. The OPC website has a resource: *How to do a Privacy Impact Assessment* that psychologists are directed to. In summary the main steps of a privacy impact statement include:

- Gather all the information you need and sketch out how and where the information you are collecting will go
- Check this information out against the Information Privacy Principles in the Privacy Act
- Identify any real privacy risks and how to mitigate them
- Produce a privacy impact assessment report
- Take action
- Review and adjust the privacy impact assessment as the project develops

8. Psychologists should be transparent about the use of AI and inform those with whom they work that they are using AI if the AI tool performs a significant part of the service, and/or if the person(s) would reasonably expect to know when consenting to a service.

Gaining Informed Consent is Essential for Whakawhanaungatanga

8.1. Engaging in the informed consent korero (conversation) is a clear sign of relationship building for those we work with and respect for their human dignity and autonomy/self-governance. It is an integral part of the initial engagement process and establishing a working partnership. Gaining informed consent means that the person has the right to choose whether to receive the psychological service (or to take part in research) based on the best information available, and to withdraw that consent should their opinion change. Informed consent should occur at the beginning of an engagement, should be documented so that there is a written record retained, and is often a dynamic process of partnership as the unfolding process requires revision of mutual understanding and agreement.

There is a Legal and Ethical requirement for Gaining Informed Consent in Aotearoa New Zealand

8.2. Gaining informed consent from people with whom you are working is both a legal and ethical requirement for all health practitioners working in Aotearoa New Zealand. The Health and Disability Patient Code of Rights³ (hereafter the Code of Rights) provides the legal basis for consent in Aotearoa New Zealand in Rights 6 and 7. These rights convey the two core pre-requisites for consent: whether the patient was given appropriate information to base their consent on and whether they were competent to understand the information when making a decision.

8.3. The Tikanga Mataatika/Code of Ethics for psychologists working in Aotearoa New Zealand⁴ recognises that obtaining informed consent is a fundamental expression of respect for the dignity of persons and peoples (Principle 1.7). Further relevant guidance is given in Principle 1.7.6 of the Code of Ethics: *In obtaining informed consent, psychologists provide as much information as a reasonable or prudent persons, family, whānau, or community would want to know before making a decision or consenting to an activity. This includes warning of any potential risks or consequences.* This reflects the test of materiality provided

³ Under Review as of December 2023

⁴ Under Review as of December 2023

for in the Code of Rights (Right 6) and in relevant case law (e.g., Rogers v Whitaker [1992] 67 ALJR 47 [High Court of Australia]).

8.4. AI presents a number of complicating features for the psychologist and the informed consent process, not least because of the emerging nature of AI. The following points are made to assist the psychologist in considering when and how to inform the people with whom they work about the use of AI in their work.

Does the psychologist always need to inform the client of their use of AI at the consent stage?

8.5. The answer to this will be partly contingent on the extent of the AI tool and its role in the psychological service provided. If the AI tool is used as an adjunct to the psychologist's reasoning, disclosure of its use at the consent stage may not be required. Psychologists, like all health practitioners, use many systems of information to inform their thinking (e.g., memories of past lectures, current readings, discussions with supervisors and colleagues). It would be unnecessary, in most cases, to disclose all of these to a client. In the case of AI being used as a partial tool in the overall service, the psychologist is still the main provider and "in charge" of the health service (Cohen, Attwood, & Williams, 2020).

8.6. However, if the AI tool is being used as a substitute for any of the significant roles the psychologist would reasonably be expected to perform, or is being used as the predominant or central method of reasoning or recommendation, or if results/advice from AI are automatically followed with limited input from the psychologist, then it would be likely that a person should be provided this information at the consent stage. Effectively, the AI tool could then be considered the main provider of the service, and this is likely to be a material point that most people would want to know when consenting to a service. Consider the example from Cohen, Attwood and Williams (2020): if you consented to a particular surgeon performing surgery on you and then awoke to find a completely different surgeon had completed the operation (without good reason), you would likely feel deceived and not informed about a significant part of the treatment.

8.7. A further point to consider is the emerging nature of AI and the significant place it currently holds in societal discourse and thought. As of 2024, AI and its various uses in health care does not yet appear to be commonplace, or business as usual, although this is likely to change rapidly in the near future. Because of this, and the likelihood that members of the public will have varying opinions and knowledge about the emerging use of AI, it may be wise to consider always mentioning when AI is used in any psychological service you provide at this stage in the evolution of AI. However, there may also be risks in 'over disclosure' of information at the consent stage, and this is discussed further in the following principle.

What information should be conveyed in the consent process?

- 8.8. As per the Code of Rights, the information conveyed in the consent process should be what is material to the person(s) you are working with. There is no legal or ethical expectation that competent individuals have to understand the actual scientific and technological processes involved with their diagnosis, treatment or any other health service they are engaged with. The expectation is that they are given the material information to make decisions regarding their self-determination and that they are given the opportunity to reach their own conclusion for their wellbeing without being misled or deceived (Boniface, 2021). This might include, why the AI is used and to what extent (e.g., if the output/decision from AI is reviewed by the psychologist).
- 8.9. In addition, the individual characteristics of a person (e.g., their level of cognitive functioning, their own expertise in the area of AI, their fears and anxieties) should all be considered in what is 'material' to them in making their decision about whether or not to engage in a psychological service that utilises AI (Boniface, 2021).
- 8.10. There does not appear to be an expectation that all possible risks regarding the use of AI are identified and communicated to a client. All aspects of healthcare carry potential unknown risks and AI is no different to this. Over disclosure of risks may make it difficult for clients to distinguish meaningful risks from trivial ones (Boniface, 2021). The risks that will be relevant to a particular client are the most important to communicate. A psychologist's experience with AI may also be a relevant factor to explore. A HDC case from 2009 (HDC decision 08HDC20258) found that a surgeon who used robotic assisted surgical technology on a 69 year old man (who later developed complications), had a duty to inform the patient that he had had limited experience with that technology.
- 8.11. When using AI tools, it is important that psychologists are always transparent during the consent process about what data will be collected and used by the tool. This includes being transparent about when a tool collects personal data and/or uses this for secondary purposes, even when data is de-identified first and the secondary purposes are not yet known. For example, it is common for generative AI tools to use the data entered into them to continue to train the tool and develop future tools. It is also important to consider that when it comes to sensitive data that de-identification can never be guaranteed, and future deep learning techniques may be able to identify individuals from previously de-identified data.

Psychologists need to be especially aware of the danger of using AI to help generate or revise articles intended for publication

8.12. The integration of AI into psychological research and journal publication presents both significant opportunities and ethical challenges. AI can inadvertently introduce errors or misinterpretations in research data or manuscript drafts, which could lead to flawed conclusions or misleading information. Ethical deployment of AI necessitates clear acknowledgment of its use in the final draft submitted for publication or other use, including the origins and limitations of the AI-generated content. Psychologists should ensure that all AI-assisted research undergoes rigorous peer review, maintaining high standards of accuracy and ethical integrity. Furthermore, informed consent protocols must evolve to cover the collection and use of data by AI systems, safeguarding confidentiality, and participant rights.

9. The onus is on the psychologist to ensure the person(s) understands the role of any AI to the extent they require to make an informed choice at the outset of a psychological service or research.

9.1. What if the AI is complex and the person with whom you are working is unlikely to be able to evaluate whether its use in their case furthers or stymies their healthcare interests? The decision maker should not be overwhelmed by technical information (Boniface, 2021). The onus is on the psychologist to ensure that the person understands the AI tool being used, without conveying complicating elements that may risk the person will not be able to adequately make a decision for themselves.

9.2. The law makes no expectation that patients are able to understand completely, or to the same level as a health practitioner. It would be unlikely a patient can explain the mechanisms of prescribed medication, and thus do not also need to explain the mechanism of an AI tool.

9.3. Although acquiring extensive knowledge of AI coding, programming, and functioning is unrealistic for most healthcare providers, those who plan to use these technologies in practice should be able to:

- Provide people with whom they are working an explanation of how the AI programme or system works, similar to how a psychologist explains how psychometrics are used in practice
- Explain the healthcare provider's experience using the AI program or system
- Describe to people the risks versus potential benefits of the AI technology (e.g., compared to human accuracy)
- Discuss the human versus machine roles and responsibilities in diagnosis, treatment, and procedures
- Describe any safeguards that have been put in place, such as cross-checking results between clinicians and AI programmes, and on-going monitoring or training of the tools and programmes
- Explain issues related to confidentiality of person's information and any data privacy risks

9.4. Taking the time to provide people with these additional details during the informed consent process and to answer any questions can help ensure that they have the appropriate information to make informed decisions about their treatment. Following the informed consent process, providers should document these discussions in health records and include copies of any related consent forms.

- 9.5. Additional care should be taken when a person has reduced capacity to consent. The Code of Rights outlines the required steps in the consent process for people with reduced capacity. The process of being transparent about AI needs to be adapted and flexible to people's differing levels of capacity.

10. Person(s) working with a psychologist should not be disadvantaged if they do not wish to have AI used in their care or their data entered into an AI system.

- 10.1. One of the overarching and significant ethical principles in healthcare is the respect for the autonomy of a person and their decision making. Given the concerns of some people regarding the emerging nature of AI and the sensitive nature of their data, it is likely that psychologists will encounter people who wish to have psychological services, but do not consent to their data being used in an AI tool for any purpose. The psychologist will need to respect this decision and ensure the person can still access the psychological services that they require. If it is not possible to offer a particular psychological service without the use of an AI tool, the psychologist must make reasonable efforts to refer the person on to another equivalent service that does not rely on AI.
- 10.2. People may have a range of ideas and beliefs about AI that may or may not be accurate. The psychologist can offer the relevant information necessary for an informed choice to be made but should avoid trying to unduly influence or convince a person that a particular AI tool is safe. The person with whom the psychologist is working can be presumed competent (unless there is good reason to suggest otherwise) to make their own decision regarding whether they wish to engage with a service that uses, in part, an AI tool. This decision must be respected by the psychologist.

Appendix

As an interesting experiment, the psychologists tasked with writing the current Guidelines asked ChatGPT to produce AI Guidelines for a psychologist working in Aotearoa New Zealand. Here is the result:

Guidelines for Psychologists Using AI in New Zealand

Introduction: Artificial Intelligence (AI) is increasingly being integrated into various fields, including psychology, to enhance research, diagnosis, treatment, and therapy processes. As psychologists in New Zealand embrace AI technologies, it's essential to establish guidelines to ensure ethical, safe, and effective usage. These guidelines aim to provide a framework for psychologists to navigate the integration of AI responsibly within their practice.

1. Ethical Considerations: a. Respect for Autonomy: Ensure that individuals' autonomy and decision-making capacities are respected throughout the AI-driven processes. b. Privacy and Confidentiality: Safeguard sensitive data and ensure compliance with relevant privacy laws such as the New Zealand Privacy Act to protect clients' confidentiality. c. Transparency: Be transparent about the use of AI technologies in psychological practice, including informing clients about how their data will be used and the limitations of AI systems. d. Equity and Diversity: Mitigate bias in AI algorithms to ensure fair and equitable treatment for all individuals, regardless of factors such as race, gender, or socio-economic status. e. Informed Consent: Obtain informed consent from clients before using AI tools or systems in their assessment, diagnosis, or treatment.
2. Competence and Training: a. Continuous Education: Stay informed about developments in AI technologies relevant to psychological practice through ongoing education and training. b. Technical Proficiency: Develop the necessary technical skills to use AI tools effectively and responsibly within the field of psychology. c. Collaboration: Collaborate with experts in AI and related fields to ensure the appropriate selection, implementation, and interpretation of AI-driven solutions. d. Supervision: Seek supervision or consultation from experienced professionals when integrating AI into psychological practice, especially in complex or novel applications.
3. Assessment and Diagnosis: a. Validation: Use AI-driven assessment tools that have been validated for reliability, validity, and cultural appropriateness in the New Zealand context. b. Integration: Integrate AI-generated insights with clinical judgment and consideration of individual client factors to inform accurate diagnoses and treatment planning. c. Interpretation: Exercise caution when interpreting AI-generated results, considering the limitations and potential biases inherent in the algorithms and datasets used.
4. Treatment and Intervention: a. Personalization: Tailor AI-driven interventions to the specific needs and preferences of each client, considering cultural, linguistic, and contextual factors. b. Monitoring and Evaluation: Continuously monitor the effectiveness and impact of AI-driven interventions on client outcomes and adjust treatment plans accordingly. c. Human Oversight: Maintain a human-centered approach to therapy, ensuring that AI technologies supplement rather than replace the therapeutic relationship between psychologists and clients.
5. Data Security and Governance: a. Data Protection: Implement robust security measures to protect client data from unauthorized access, breaches, or misuse. b. Data Ownership: Clarify

ownership and usage rights of data collected through AI-driven tools, ensuring that clients retain control over their personal information. c. Compliance: Adhere to relevant regulations and standards governing data protection and ethical conduct, including the New Zealand Health Information Privacy Code and professional codes of ethics.

6. Professional Responsibility:
 - a. Accountability: Take responsibility for the ethical implications and outcomes of using AI technologies in psychological practice, including any potential risks or unintended consequences.
 - b. Advocacy: Advocate for ethical guidelines and regulatory frameworks that promote responsible use of AI in psychology at the organizational and policy levels.
 - c. Continuous Reflection: Reflect on the ethical, social, and professional implications of integrating AI into psychological practice, engaging in ongoing dialogue and self-assessment.

Conclusion: As psychologists in New Zealand harness the potential of AI to enhance their practice, adherence to ethical guidelines is paramount to ensure the responsible and beneficial use of these technologies. By prioritizing ethical considerations, maintaining competence, and upholding professional standards, psychologists can leverage AI to improve client outcomes while safeguarding their well-being and rights.

References

- AIHLEG. (2018). *A definition of AI: Main capabilities and scientific disciplines*. Bruxelles: European Commission.
- Boniface, C. (2021). *The legal impact of artificial intelligence on the New Zealand health system*. A thesis completed for a PhD at the University of Canterbury
- Cohen, G., Attwood, J.A. & Williams, L. (2020). Informed consent and medical artificial intelligence: What to tell the patient? *Georgetown Law Journal*, 108, 1425 – 1469
- Dobson, R. & Whittaker, R. (2023). What Do Health Service Users Think About the Use of Their Data for AI Development? In: Bichel-Findlay, J. Otero, P. Scott, P. Huesing E. eds., 2023. MEDINFO 2023: The future is accessible. Proceedings of the 19th World Congress on Medical and Health Informatics. IOS Press.
- Laukkonen, R., Biddell, H., & Gallagher, R. (2019). Preparing humanity for change and artificial intelligence: Learning to learn as a safeguard against volatility, uncertainty, complexity, and ambiguity. *OECD report*.
- Office of the Privacy Commissioner (2023). *Artificial intelligence and the information privacy principles*.
- Taiuru, K. (2023). Te Tiriti Based Artificial Intelligence Ethical Principles. Available from: <http://www.taiuru.co.nz/AI-Principles>
- Waitangi Tribunal's WAI 2522 Report on the comprehensive and progressive agreement for trans-pacific partnership (2023)