IEEE SA
STANDARDS ASSOCIATION

IEEE SA ICAP WHITE PAPER

RECONCEPTUALIZING THE ETHICAL GUIDELINES FOR MENTAL HEALTH APPS:

VALUES FROM FEMINISM, DISABILITY STUDIES, AND INTERCULTURAL ETHICS

Executive Authors:
Dr. Matthew J. Dennis (TU Eindhoven) and Dr. Lily E. Frank (TU Eindhoven)

Academic Contributors: Arthur Bran Herbener (Aarhus University), Michał Klincewicz (Tilburg University), Malene Flensborg Damholdt (Aarhus University), Anna Puzio (Twente University), and Katherine Bassil (University Utrecht Medical Center)

NGO & Industry Contributors:
Jessica Stone (Stone Executive Coaching), Philipp Schneidenbach (Materna SE), Shriya Das (Compass Pathways), Ella Thomas (Relate), and Mat Rawsthorne (Virtual Health Labs)

IEEE

TRADEMARKS AND DISCLAIMERS

IEEE believes the information in this publication is accurate as of its publication date; such information is subject to change without notice. IEEE is not responsible for any inadvertent errors.

The ideas and proposals in this specification are the respective author's views and do not represent the views of the affiliated organization.

ACKNOWLEDGMENTS

The executive authors would like to thank the Philosophy & Ethics Capacity Group (TU Eindhoven) and the Ethics of Socially Disruptive Technologies Research Consortium for their generous support, which has taken the form of research time to complete this project.

Special thanks are given to Becky Inkster (University of Cambridge), Maria Palombini (IEEE), Christopher Burr (Alan Turing Institute), and Shana Pepin (IEEE).

The Institute of Electrical and Electronics Engineers, Inc. 3 Park Avenue, New York, NY 10016-5997, USA

PDF: STDVA27117 979-8-8557-0965-0

Find IEEE standards and standards-related product listings at: http://standards.ieee.org.

NOTICE AND DISCLAIMER OF LIABILITY CONCERNING THE USE OF IEEE CONFORMITY ASSESSMENT PROGRAM (ICAP) DOCUMENTS

This IEEE Conformity Assessment Program (ICAP) publication ("Document") is not a consensus standard. Specifically, this Document is NOT AN IEEE STANDARD. IEEE expressly disclaims all warranties (express, implied, and statutory) related to this Document, including, but not limited to, the warranties of: merchantability; fitness for a particular purpose; non-infringement; quality, accuracy, effectiveness, or currency. In addition, IEEE disclaims any and all conditions relating to results. This ICAP document is supplied "AS IS" and "WITH ALL FAULTS."

IN NO EVENT SHALL IEEE BE LIABLE FOR ANY ERRORS OR OMISSIONS OR DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO: PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE USE OF THIS WORK, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE AND REGARDLESS OF WHETHER SUCH DAMAGE WAS FORESEEABLE.

TABLE OF CONTENTS

RECONCEPTUALIZING THE ETHICAL GUIDELINES FOR MENTAL HEALTH APPS:

VALUES AND CONCEPTS FROM FEMINISM, DISABILITY STUDIES, AND INTERCULTURAL ETHICS

EXECUTIVE SUMMARY

Existing ethical guidelines that aim to guide the development of mental health apps tend to overemphasize the role of Western conceptual frameworks. While such frameworks have proved to be a useful first step in introducing ethics to a previously unregulated industry, the rapid global uptake of mental health apps requires thinking more deeply about the diverse populations these apps seek to serve. One way to do this is to introduce more intercultural ethical perspectives into app design and the guidelines that aim to encourage best practices. In addition to this, existing ethical guidelines can also benefit from the ethical scholarship from the feminist and disability traditions, both of which highlight specific ethical considerations for vulnerable users. Rethinking the ethical responsibilities of mental health app designers through the prisms of feminism, disability studies, and intercultural philosophy leads us to a more global and inclusive set of ethical considerations in app design. This white paper explores what existing guidelines for the regulation of mental health apps are missing. It also explores how these guidelines could be improved for users who inhabit an increasingly diverse and globalized world.

1. CHALLENGE: ETHICAL GUIDELINES FOR A DIVERSE AND GLOBALIZED WORLD

Mental health apps are increasingly being proposed by governments, industry, and medicine as a digital solution to the growing need for mental health care across the world. The advantages of mental health apps have inspired app designers to create a vast number of products that are only partially governmentally regulated and often only in the country where the app provider does business. There have been some attempts at regulation by the European Commission's Digital Services Act (DSA) and optional regulation by major industrial players. Nevertheless, the creators of these initiatives concede that they are responding to a dynamic and fluid range of technical solutions to the clinical needs of app users and are poorly equipped to guide app developers from an ethical point of view.

The current regulatory mandates that attempt to guide the development of mental health apps focus on legal issues, with a heavy emphasis on efficacy, privacy concerns, and data safeguarding, which, this report argues, can be usefully supplemented with greater attention to the nuanced ethical value conflicts that regularly arise in this space. Where ethical guidelines have been drafted (Burr and Powell [18]), it has been acknowledged that they draw from predominantly mainstream European or Northern American rights-based ethical frameworks. This white paper attempts to remedy this by broadening the scope of normative frameworks that inform regulations. The so-called "next billion internet users" (Arora [4]) will predominately come from the Global South, with mental health apps showing exponential growth in non-Western regions. To better serve this growing number of users, a new set of ethical considerations drawing on the conceptual resources of what have been often underrepresented traditions are introduced (Singh and Sagar [75]). A broad representation of ethical traditions is important for many reasons, but in this case, the traditions on which this paper focuses stand to directly benefit the growing number of users of mental health apps.

The first set of traditions drawn from are intercultural ethical ones. Given that many users of mental health apps will come from non-Western countries, it is vital to ensure that Western ethical notions do not exclusively inform the regulation of these products. This means that the ethical regulation of these apps should explore normative frameworks that are non-Western, those that have developed in traditions outside the Western world. This white paper certainly lays no claim to be exhaustive here. Instead, it takes a selection of non-Western ethical concepts to illustrate how mental health app developers can improve their products by thinking beyond the Western ethical canon. The hope is that the usefulness of these concepts will inspire developers and regulators

to think more widely about the ethical frameworks they draw from. Moreover, this aims to make the ethical guidelines that will inform the design of mental health apps more familiar to the ethical interests and concerns of non-Western users, while also providing ethical direction in this space.

The second set of traditions this paper examines has also been neglected and historically marginalized. This is the ethical theorizing of feminist and disability studies scholars, both of which have transformed the landscape of moral philosophy over the last four decades. Feminism has emphasized the role and importance of vulnerability, social relationships, interconnectedness, and relational autonomy. For example, it has rehabilitated the notion of care ethics, which emphasizes context in moral deliberation and the relational connections people have toward vulnerable members of the human community.

Disability studies also offer a challenging perspective on traditional ethical conceptions of human well-being and morality. Like feminism, this tradition emphasizes that humans are embodied beings but points to the diversity and variety of human embodiment, which is often neglected or judged critically.

As explored in the first section of this white paper, "The Regulatory Status Quo" (1.1), regulatory guidelines for mental health apps are currently highly limited from a geographical and cultural point of view, as there are no ethical codes for the designers of these products that draw on non-Western ethical traditions. Like the objection to the overarching use of Western ethical perspectives in the design of these apps (instead of intercultural ones), the same point can be made to plead for the increased use of feminist and disability studies in this domain. The use of apps is unequally divided along gender lines, and some of the most vulnerable users of the apps are non-able-bodied. This white paper proposes that listening to these traditions and actively incorporating their ethical insights into future guidelines for mental health apps will improve the design of these apps for tomorrow's global audience.

1.1. REGULATORY STATUS QUO

Summarizing the regulatory status quo in a concise yet useful way is challenging because of the breadth of existing regulations as well as their inconsistent level of depth and detail. Nevertheless, it is possible to map the key contours of the regulatory landscape in an informative way to show what is missing. To do this, this paper will make some broad observations before doing a deep dive into three approaches to regulation across specific regions (Europe, Germany, and the USA). The significant voids in regulatory frameworks necessitate caution, particularly in areas where app developers operate in a regulatory "wild west," encompassing much of the Global

South. It is also important to note that current regulatory approaches prioritize approaches with a medical background and a strong preference for apps that are based on cognitive behavioral therapy (CBT). Various reasons can be attributed to this, but one key factor is the relative ease with which CBT approaches lend themselves to the online format as well as how they can be tested for effectiveness easily.

1.1.1. MENTAL HEALTH APP REGULATION IN THE EUROPEAN UNION

In the European Union, mental health apps are primarily regulated by the Medical Devices Regulation (MDR) and the General Data Protection Regulation (GDPR). The MDR, which came into effect in May 2021, classifies mental health apps as medical devices if they diagnose or treat mental health conditions. These apps must have a CE mark to comply with EU safety standards.[1]

The GDPR, effective since 2018, regulates how health data are processed. Mental health apps that collect, store, or handle personal information must adhere to strict data protection, security, and user consent rules. Health data are considered sensitive, requiring explicit consent and additional.

The European Data Protection Board (EDPB) has provided specific guidelines that stress transparency, user control, and data minimization, ensuring mental health apps align with GDPR principles.[2].

Furthermore, the European Commission has created mHealth assessment guidelines to ensure the safety, quality, and reliability of health apps. They offer a framework for assessing the clinical effectiveness, data protection, and usability of these apps. Additionally, individual member states have their own regulations. For instance, Germany's Digital Health Applications (DiGA) program allows certain apps to be prescribed by doctors and reimbursed by health insurance.[3]

1.1.2. MENTAL HEALTH APP REGULATION IN GERMANY

In addition to the European-wide regulatory guidelines discussed above, it is useful to focus on Germany to see how a national government within the European Union is responding to the challenge of regulation. Mental health apps in Germany are specifically regulated by the Digital Health Apps Directive [Digitale Gesundheitsanwendungen-Verordnung (DiGAV)], which is the leading directive for any manufacturer of digital

[1] Available at https://eur-lex.europa.eu/eli/reg/2016/679/oj
[2] Available at https://edpb.europa.eu/our-work-tools/our-documents/guidelines/guidelines-032020-processing-data-concerning-health-purpose_en
[3] Available at https://www.bfarm.de/EN/Medical-devices/DiGA-and-DiPA/DiGA/_node.html

health apps. The main aspect of this regulation is the set of prerequisites for the apps to be accepted in health insurance policies and cost redemptions. Specifically, DiGAV governs apps to do the following things:

- Follow the EU MDR (Medical Device Regulation 2017/745, with a strong focus on quality management, unique device identification, and risk
- Process data according to GDPR (EU General Data Protection Regulation)
- Prove effectiveness (proof of positive impact)
- Be committed to being listed in DiGA (directory of registered digital health apps)

Along with the DiGAV, mental health apps made and used in Germany must follow the Digital Supply and Care Modernization Act (Digitale-Versorgung-und-Pflege-Modernisierungs-Gesetz [DVPMG]), which is a regulation that applies to any kind of digital care apps (such as medication support apps, alarm systems, and guidance apps for fulfilling requests). The DVPMG requires the providers of these products to follow the following mandates:

- Provide electronic documentation according to the Electronic Patients Archive (ePA)
- Facilitate digital communication with governing interfaces
- Conduct and provide Data Protection Impact Assessment (DPIA) according to GDPR

DVPMG will ultimately impact almost any existing medical directive in Germany as a “digital addendum.”

1.1.3. MENTAL HEALTH APP REGULATION IN THE UNITED STATES

In the USA, the regulatory framework for mental health apps predominantly involves oversight by the Food and Drug Administration (FDA), Federal Trade Commission (FTC), and Health Insurance Portability and Accountability Act (HIPAA) compliance requirements that vary based on the app’s specific functionalities and how it handles user data. FDA generally divides mental health apps into the following three categories:

- General wellness products (GWPs)
- Medical products
- FDA-approved apps

The FDA has outlined examples of "low-risk" apps eligible for regulatory discretion, particularly focusing on mental health apps. Regulatory discretion for low-risk GWP apps was emphasized during the COVID-19 pandemic following the Public Health Emergency (PHE) declaration.

The FDA oversees apps that function as medical devices. If an app solely provides educational information on mental health, it is not eligible for FDA approval. However, if an app offers therapeutic interventions, its creators may pursue FDA approval to certify it. To be considered a medical device, the software must be "intended for one or more medical purposes and achieve these purposes independent of a hardware medical device."[4]

The FDA enforces the FD&C Act that regulates the safety and effectiveness of medical devices, including some mobile medical apps. The agency focuses on digital health devices that could pose risks to consumers if they malfunction or if there are any breaches. Software functions that meet the definition of a device under section 201(h) of the FD&C Act are regulated as medical devices.

1.2. W.E.I.R.D DESIGN: CONNECTING INTERCULTURAL PERSPECTIVES WITH THE NEEDS OF MARGINALIZED USERS

As noted in Section 1 (Challenge: Ethical Guidelines for a Diverse and Globalized World), by far, the bulk of regulatory structures—mandatory guidelines, voluntary codes of conduct, and so on—are underpinned by Western ethical frameworks. The authors of this report wish to emphasize that such frameworks have been an effective way to begin the process of regulating mental health apps, although our aim here is to show why they are not comprehensive enough to accommodate the needs of the increasingly diverse users of these products. While we focus on the ethical frameworks that have informed the creation of apps themselves and their regulatory structures, it is vital to note that much of the empirical research that has informed their design is modeled on a subset—globally speaking—of users. Recently, the limits of research on this subset of users have become increasingly hard to ignore. Joseph Henrich's pivotal 2020 publication, "The WEIRDest People in the World: How the West Became Psychologically Peculiar and Particularly Prosperous," ignited discussions around the Western-centric bias prevalent in psychological research. This critique is centered on the fact that the majority of data is derived from Western, educated, industrialized, rich, and democratic (WEIRD) participants, highlighting a significant skew in the demographic profile of research subjects. Those who apply Henrich's

[4] See https://www.imdrf.org and consult "Software as a Medical Device (SaMD): Key Definitions."

findings to existing normative frameworks (Dennis & Clancy, 2022 [27]) point out that WEIRD users are in the minority compared to the diversity of users who use digital technologies.

It is also important to note that mental health apps that employ Western ethical frameworks consistently demonstrate a more progressive ethical approach than apps that have been designed without an ethical framework. Apps with no ethical framework typically focus on complying with the rules that allow for inclusion into Apple's App Store or Google's Android Marketplace. Apps that prioritize ethical considerations more comprehensively may add extra provisions regarding privacy (minimizing data capture by cookies, for example), autonomy (reducing the use of persuasive or e-nudging technologies, for instance), or fairness (minimizing or slowing down user personalization). These ethical considerations have broad appeal (both across cultures and geographic locations), but as subsequent sections of this white paper will show (especially Section 1.3), additional (or alternative) ethical considerations are important in non-Western cultures. In addition to sketching how these ethical considerations are important in Section 1.3, this white paper will examine them in detail in Sections 2.1–2.4. In these sections, we will discuss how the design of mental health apps affects access to mental health services in unpredictable ways for marginalized groups (2.1 Autonomy versus Relational Autonomy), how non-Western users often have different attitudes toward privacy and the use of their data (2.2 Social Justice and Social Harmony), how many non-Western traditions endorse a notion of autonomy that is highly relational rather than individualistic (2.3 Accessibility), and how a strong commitment to social justice is often mitigated by a robust commitment to social harmony (2.4 Efficacy).

1.3. WHICH FRAMEWORKS PROVIDE BETTER CONCEPTUAL RESOURCES FOR ETHICAL MENTAL HEALTH CARE APPS?

In the ethics of technology, there is a trend toward relational approaches, such as actor-network theory, mediation theory, postphenomenological approaches, new materialism, technofeminism, and much more. Hagendorff [40] refers to these as "blind spots" in AI ethics and argues that they need to be explored and implemented more extensively. Many of these approaches or their adaptations are characterized by their focus on relationality (e.g., relating to human, nonhuman, and technology), body, and embodiment and are open to intercultural approaches. This is particularly important in the discourse on mental health, as mental health is relational and not only related to the psyche and brain but also fundamentally connected to the body (Fuchs

[36]). Both aspects tend to be overlooked in the current debate on AI and mental health, which is very much focused on the brain.

Feminist ethics is a broad set of approaches and theories that aim to include the historically overlooked or even ridiculed moral insights of women and insights that come from traditional women's experiences, such as mothering, care work, emotional labor, the centrality of social relationships, and so on. Originally, the field focused on trying to understand whether there are differences in strategies of moral thinking between men and women (Gilligan [43]). The field has progressed considerably to focus on overlooked concepts such as vulnerability, relational autonomy, and new kinds of injustice (e.g., testimonial injustice) (Fricker [33]) and to include an understanding of power relations in ethics and political philosophy that takes into account intersection forms of oppression and how they interact, such as race, gender, disability, socio-economic status, and sexual orientation (cf., Garry, Khader, and Stone, 2017 [38]). This paper argues that the value of care and the centrality of vulnerability, as well as insights from feminist ethics, should be given greater attention in the development of mental health apps.

It can be observed that people, depending on gender, race, or disability, face very specific challenges in mental health that deviate from the norm typically assumed in apps and, therefore, require special consideration (Hendl, Jansky, and Wild [50]). As Perez [77] has pointed out in her seminal book "Invisible Women", medical technologies always use the "white man" as the prototype, resulting in a "gender data gap" where data from women are not collected or sufficiently considered—this also applies to many minority groups. This has a significant impact on the health of these groups and, as Perez shows using the example of heart attacks, whose symptoms in women are very different from those in men and therefore go unrecognized, can also have life-threatening consequences.

In recent years, queer feminism, which expands the feminist perspective to include queer individuals, has gained importance. Members of the LGBTQ+ community experience significant discrimination in healthcare, have specific needs, and are overlooked in digital healthcare (Austin and R. Goodman [6], Coleman, et al. [22], Chhabra and Kapadia [21]). Notably, among transpersons, "high rates of depression, substance abuse, self-harm, and suicide are well documented in the literature" (Talbert [93]). Wilson and Cariola [98] have identified "five core themes in their literature review on "LGBTQI+ Youth and Mental Health." They are as follows: (1) isolation, rejection, phobia, need for support; (2) marginalization; (3) depression, self-harm, and suicidality; (4) policy and environment; and (5) connectedness."

One strain of feminist ethics emphasizes the importance of the value of care. The ethics of care emphasizes the role that relationships play in defining our moral obligations and in making moral decisions. In contrast to ethical theories that focus on single-abstract moral principles, such as “maximize utility.”

Jacobs [55], [56] and Oosterlaken [74] have made the case for attention to the complexity of human vulnerability when designing technologies for health care, particularly digital technologies designed for health-related behavioral changes. Emphasis on vulnerability, which draws on feminist and care ethics (Mackenzie, Rogers, and Dodds [67]), also has a role to play when constructing guidelines, standards, or regulations for these technologies. Paying attention to human vulnerability in the design of mental health apps reveals ethical issues, value conflicts, and barriers to effectiveness that may have gone otherwise unnoticed. Jacobs illustrates this with her analysis of a technology that is meant to increase medication adherence for elderly patients, “Glow Cap,” part of a “smart” pill bottle, which lights up when it is time for someone to take their pills and is connected to a mobile app that allows for monitoring of medication usage and adherence (signaled by opening and closing the bottle, according to the schedule) (Jacobs [56]). Jacobs points out that a patient with mild dementia who strives to live as independently as possible may feel surveilled by the introduction of this technology, especially since the app allows for sharing her medication use data with caregivers or physicians. Because of such an individual’s vulnerability and dependence on caregivers, they may feel they have little control over whether or not they use such technology or even whether or not they share the information it produces.

Another source of inspiration and insights comes from disability studies and the recently emerging field of “crip-technoscience.” They are particularly relevant to design and regulatory frameworks for digital mental health technologies. According to the Crip Technoscience Manifesto (Hamraie and Fritsch [49]), “crip technoscience braids together two provocative concepts: ‘crip,’ the noncompliant, antiassimilationist position that disability is a desirable part of the world, and ‘technoscience,’ the coproduction of science, technology, and political life” (p. 2 [49]). This approach is contrasted with “mainstream ‘disability technoscience’ as a field of traditional expert relations and practices concerned with designing for disabled people rather than with or by disabled people” (p. 4 [49]).

One of the emerging patterns that has been recognized by scholars working in this field is “technoableism,” a term coined by Ashley Shew that describes:

> “…a rhetoric of disability that at once talks about empowering disabled people through technologies while at the same time reinforcing ableist tropes about what body-minds are good to have and who counts as

> worthy…Technoableists usually think they have the good of disabled people in mind. They do not see how their work reinscribes ableist tropes and ideas on disabled bodies and minds." [73]

In certain instances, mental health issues and disorders can be considered a form of disability, especially when they hinder an individual's capacity to engage in desired activities or to embody their valued identities. As such, it is appropriate to use lenses from disability studies to inform the design of technologies for diagnostics, monitoring, and treatment of mental health conditions.

There are several movements in the field of technology design that promise to center considerations of inclusivity, justice, representations, and values in their approaches. Some of these methods have become widely accepted, such as value-sensitive design VSD (Friedman et al. [34]; Friedman and Hendry [35]). Related methods and approaches include the Design Justice movement (Costanza-Chock [23]), inclusive design (Imrie and Hall [54]), universal design (Goldsmith [45]), participatory design (Bannon and Ehn [9]), and human-centered design (Kurosu [63]).

VSD is a structured approach to design that emerged from the field of software design and human-computer interaction. It is a tripartite, iterative process that aims to incorporate stakeholder values into the design process from the very beginning. Friedman, Hendry, and Borning explain: "Value-sensitive design is a theoretically grounded approach to the design of technology that accounts for human values in a principled and comprehensive manner throughout the design process." [34] Comprised of conceptual, empirical, and technical phases, VSD takes as a starting point the claim that not only can designers shape technologies but that technologies also shape us; that is, they influence the way we see the world, ourselves, and our interactions with each other. Technology is, therefore, not value-neutral and embodies values in that it creates affordances for users and other stakeholders to behave and think in ways that may or may not be the target of the technology. Conceptual research in VSD involves identifying stakeholders, direct and indirect; identifying relevant values and value conflicts; using analogous case studies; and analyzing design briefs, laws, standards, and ethical literature.

The empirical phase involves conducting stakeholder interviews and surveys using methods from social sciences. The technical phase involves investigating how current technical mechanisms and design systems either support or hinder/undermine certain values. The products of the technical investigation can include mock-ups or prototypes of improved technology/design and value-based reviews of existing technologies or the creation of design guidelines. The results of the technical phase are then meant to be fed back into the conceptual phase for further evaluation. Consider something like this: In discussing the translation of stakeholder values into tangible design requirements, Jacobs [55], [56] utilizes the mental health chatbot Woebot as a prime example.

She illustrates how the value of meaningful social affiliation can be embodied in design specifics. This principle asserts that in the interactions with Woebot, there exists a fundamental norm: “A person should be able to express their thoughts and emotions and receive attentive listening.”

Value-sensitive design (VSD) faces its share of criticism, but it serves as a valuable foundation and framework for investigating how values are integrated into technological solutions. This is particularly pertinent in the field of mental health technology, where the intersection of values, vulnerability, and disability demands careful consideration. As we will see in what follows, in addition to feminist approaches and those taken from disability studies, intercultural ethical frameworks have a role to play, which goes some way to recognizing the fact that the users of mental health apps come from diverse cultural backgrounds that have been informed by various non-Western ethical traditions.

2. KEY ETHICAL ISSUES FOR ETHICAL MENTAL HEALTH APPS

2.1. AUTONOMY VS. RELATIONAL AUTONOMY

Mental health apps are often presented as tools to support individuals seeking mental health support and care. These apps are advertised as being concerned with the user's behavior, nudging them in healthy directions through the collection and analysis of their data. While this can enhance a user's personal autonomy over their health, it can also undermine their autonomy in certain contexts. For instance, when health and economic interests are intertwined in mental health apps, it becomes rather challenging for users to distinguish between what is health advice and what is commercial intent (Sax, Helberger, and Bol [81]). This, in turn, raises questions about not only the transparency of such practices but also the autonomy of users. A large number of existing mental health apps are commercial with commercial motives, targeting users with advertisement through the (continuous) collection of personal and sensitive data [81]. Personalized strategies are increasingly becoming the norm, with big data promising to offer information that is distinctive to each individual and hence being more persuasive in nudging individuals to change their behavior (Alqahtani, Meier, and Orji [2]). The autonomy of users in mental health apps revolves around the alignment between recommended advice or nudges and the genuine desires of users. When app recommendations align with the user's authentic wishes, autonomy is enhanced. However, if the advice contradicts the user's authentic desires, this can potentially undermine autonomy. This is typically seen when mental health apps are designed in a way that prioritizes clinical utility, for example, leaving out personal utility and personal meaning. Moreover, improved personalized strategies (i.e., personalized communication and personalized response language) will increase the potential to influence user behavior using manipulative tactics (Shumanov and Johnson [84]). This is particularly pertinent in AI-driven tools, including large language models (LLMs), due to their algorithmic nature (Bassil [10]). They do not take into consideration an individual's user preferences, values, and boundaries. The impact of digital mental health tools on the therapeutic relationship has often been discussed. Trust is an important requirement in this context, especially in the case of mental health, where users are in a vulnerable position and invited to share very sensitive information about themselves. When economic incentives are integrated into mental health apps, the therapeutic relationship may become compromised. Users may question the motives of the recommendations provided by the app, which can erode the sense of authenticity and trust in the therapeutic process. Moreover, the recognition of an individual's autonomy by their therapist is crucial to the therapeutic process (Kinsella [60], [61]). Hence, the substitution of therapists for digital mental health tools raises the question of whether a user's

autonomy will be supported and facilitated similarly. Non-Western approaches such as Ubuntu, Hinduism, or Japanese techno-animism also bring a different relationship not only to humans but also to nonhumans. We are in relationships not only with people but also with technology. This disrupts Western concepts of autonomy, agency, or personhood.

Another concern relates to the impact of continued and long-term use of mental health apps on user autonomy, a question that remains underexplored. In the short term, mental health apps may improve negative symptoms of mental illnesses; however, continued use over long periods may negatively impact one's self-esteem and confidence. Continued use of digital mental health apps can sometimes foster dependency among users, which can have implications for their autonomy (Bakker, et al. [7]). While these apps are designed to provide support and tools for managing mental well-being, prolonged reliance on them may inadvertently diminish user's confidence in their growing abilities outside of the app's framework, even resulting in feelings of worthlessness (Laestadius, et al. [64], Parsakia [76]). This dependency can undermine their autonomy by limiting their willingness or ability to seek alternative sources of support or develop their own coping strategies. Dependency can also alienate users and impact their connectedness to others, and they may withdraw from interpersonal interactions or neglect opportunities to connect with friends, family, or mental health professionals. Connectedness is considered one of the essential dimensions for recovery from mental health conditions (Bergamin, et al. [12]). However, relations are often neglected in the brain, cognition, and neuro-focused AI discourse. Moreover, connectedness also relates to autonomy, particularly relational autonomy, which may also be impacted by mental health apps. Relational autonomy is a concept that expands upon traditional notions of autonomy, emphasizing the influence of an individual's sociocultural environment on their behavior and actions (Mackenzie and Stoljar [68]). While technology has the potential to provide individuals with a wider range of opportunities which, in turn, enhance their relational autonomy, mental health apps must be developed in a way that retains a focus on relational autonomy instead of overemphasizing individualistic imposed set of values (Burr & Morley [17]). For example, in contrast to individualistic approaches, non-Western traditions, such as certain Ubuntu-based African approaches to mental health support often endorse relational autonomy, emphasizing the significance of community and social interconnectedness in shaping well-being (Herrera-Ferrá, Salles, and Cabrera [53]).

Moreover, non-Western approaches also offer a stronger focus on body and embodiment, thus enabling a more holistic approach to mental health. This is important because the discourse on mental health tends to forget the connection and inseparability of psyche and physique.

It then becomes clear that mental health apps should be built with autonomy being a central focus. The concept of autonomy should then include the following requirements: authenticity, competence, independence, and availability of options. The latter allows individuals, users of mental health apps, to make decisions in line with their values, desires, and goals ultimately influencing their behavior, from meaningful options. This, in turn, raises questions about the autonomy of individuals from low- or middle-income countries, where access to human therapists is not an option, and mental health app access is the only option (Deb, et al. [25]).

This, of course, warrants scrutiny of the quality of mental health apps in preserving user autonomy. Autonomy and personal agency are key factors in the management of a patient's mental illness (Drake and Whitley [28]). Mental illness can undermine an individual's personal autonomy, and so can mental health apps (Bergamin, et al. [12]). Mental health apps that are being developed should ensure that autonomy is not further disrupted by ensuring the preservation of an individual's competence and authenticity [12].

2.2. SOCIAL JUSTICE AND SOCIAL HARMONY

As mental health treatment moves beyond a "one-size-fits-all" approach, there is a recognition that it must be more ambitious still and include the full context of the recipient (Gómez-Carrillo, et al. [46]). However, digital self-service condition management tools, often positioned as optional entry points or augments to therapist-led treatment, can inadvertently transform recipients of care into users of technology (Baldwin [8]), with unclear accountability for outcomes (Hamdoun, et al.[48]) and little recognition that these are determined by the wider context (Kohn, et al. [62]) and conditions that may not reflect the intended setting (Kim, et al. [58]). Furthermore, these data-driven tools are imperfect (Bergin, et al. [13]), reliant on the accuracy of sometimes out-of-date electronic health records, seen as the only version of the "truth" with no recourse to contest them by the individual or families they are meant to represent. Along with this Epistemic Injustice, issues such as embedded bias (Friedman and Nissenbaum [35], Gebru [39]) and the realized potential of innovations to widen existing inequalities and power imbalances are often distant concepts to those privileged not to experience them (Benjamin [11]).[5]

Schukking et al. [82] highlight how the neoliberal paradigm has pushed other non-Western ethical perspectives to the periphery, but, as Ziesche's 2023 review [104] demonstrates, the non-WEIRD global traditions can enrich our understanding of global governance needs. For example, while evaluation frameworks for mental health apps evolve (Ramos et al., [79]), the Hindu principle of "Dharma" reminds us that the original motive and how

[5] Faissner et al. [25] identify three types of epistemic injustice in mental healthcare: "testimonial, hermeneutical, and contributory Injustice."

these ends are achieved are also important. In contrast with the liberal, individualist tradition that has heavily influenced Western technology ethics (Berman [14]), the emphasis in some Asian cultures on community and harmony over personal achievement (Kim, Yang, and Hwang [59]) suggests different designs, functionality, and measures of success in digital mental health applications. Given that responsibility for ethics in this field is distributed among developers, commissioners, patients, and practitioners (Wykes, Lipshitz, and Schueller [102]), collectivist frames such as the African "Ubuntu" which acknowledge connections, the Buddhist focus on reduction in suffering, and the Islamic "Magasid" concept of focusing on meeting essential human needs first add to the richer, complex picture. Similarly, recognizing the global externalities of innovation (e.g., Crawford [24]), the Maori concept of "mauri," which includes environmental sustainability and complexity, also seems highly relevant. The Recovery approach in mental health, which avoids a reductionist, medical focus on symptom management (Deegan [26]), shares important commonalities with these non-Western perspectives. Its emphasis on a personal journey toward living a meaningful life (Slade [88]) aligns in interesting ways with the Confucian idea of cultivating harmonious social relationships and fulfilling one's role-based responsibilities in the community (Tu and Wei-Ming [95]). The added complexity, for example, of relational autonomy (Walter and Ross [97]) in the context of health conditions that feature the varying capacity to consent (Appelbaum and Grisso [4]), surely enriches discussions of data privacy and risk, particularly when considered through the power-imbalance lens of Feminist ethics (Figueroa, et al. [31]).

Apparent consistency in the proliferation of digital health standards is shown to be cosmetic when the operational characteristics of stated values are detailed [37]. This suggests that, instead of assuming that everyone agrees on the meaning of a particular guiding principle, tools such as the ACT Matrix [78], which encourage group convergence on behavioral definitions of what "good" does and does not look like, might usefully be employed. However, by curating a collage of diverse perspectives, we risk assembling a "diffuse pluralism" [100] that contradicts and confuses and cannot compete with the neoclassical paradigm. Furthermore, by imposing standards, we perpetuate the disempowerment of the hardly reached, underserved, and unheard. The principle of coproduction in digital mental health has gained traction in the design of services and innovations [75], but what about involvement across the full lifecycle and engendering a sense of ownership, and, therefore, values [72] in the ethical frameworks that govern? One possible solution is to embrace the complexity evolution paradigm that can bring the non-neoclassical approaches back into the center and allow them to coexist. If the landscape of digital mental health can be considered one of the "Commons," a shared public good with multiple stakeholders negotiating access to its resources, then perhaps the ProSocial method,

with its application of Common Design Principles (Wilson, Ostrom, and Cox [86]), can facilitate convergence on shared definitions to guide moral decision-making for each solution in the field.

2.3. ACCESSIBILITY

The wide accessibility of mental health apps may help address challenges caused by the rising incidence of mental health problems (WHO, 2013 [101]) and the substantial treatment gap globally (Moitra et al., [69]; Moitra et al. [70]). In Western countries, psychological treatments are often provided by highly educated professionals (e.g., psychiatrists and psychologists) in one-to-one sessions at hospitals and private clinics (Kazdin et al. [57]). While this treatment model is evidence-based and widespread, it also creates structural barriers to care-seeking, such as economic costs, waiting lists, and geographical distances (Andrade et al., [3]). New initiatives to improve access to mental health care are, therefore, of utmost importance, such as interventions that can be upscaled to reach a larger number of individuals (*scalability*), reach individuals whom contemporary services struggle to support (*reachability*), and affordable (*affordability*) (Kazdin et al. [57]). Mental health apps can be designed to be all three of these things.

However, the high accessibility of mental health apps may also (1) foster inequalities in access to care and (2) care dependency. Structural barriers to seeking care, such as economic costs, waiting lists, and geographical distances, are often linked to sociodemographic factors. For example, low-income earners in rural areas face long waiting lists associated with public or insurance-based care systems and, in some cases, long travel distances to reach mental health providers. In contrast, high-income earners in urban areas can access a broader range of conveniently located practitioners and can bypass waiting lists by seeking private treatment. Mental health apps can reach individuals who might otherwise not receive care but also create incentives to deprioritize investment in high-quality public health infrastructure aimed at less resourceful individuals. Consequently, the accessibility of mental health apps can solidify existing structural barriers and even create new ones caused by unequal access to digital technologies or individual differences in technological literacy.

The second drawback is the potential creation of dependence or overreliance on someone or something to sustain mental health (Geurtzen et al. [41]; Glanert et al. [44]). Patients may believe that the therapist/treatment represents the only option to improve mental health (*lack of perceived alternatives*), actively seek the emotional support and physical presence of the therapist (*emotional dependency*), and assume a helpless stance wherein they rely on the therapist to act on their behalf to manage mental health (*submissive dependency*) (Geurtzen et al., [42]). While some degree of dependency is essential to establishing a strong therapeutic alliance, and in

interhuman relationships, in general, if dependency undermines patients' confidence in their ability to sustain mental health, then the patient may be reluctant to end treatment. The 24/7, on-demand accessibility of mental health apps, coupled with stakeholders' financial interests, has the potential to perpetuate care dependency. Instead of attending therapy sessions once a week, the "therapist" essentially "moves into the patient's house" and is never more than a couple of clicks away during times of distress.

Dependency may be particularly troubling in the context of recently developed AI conversational technologies capable of engaging in human-sounding conversations and responding in empathetic manners. Considering human tendencies to attribute mental traits to inanimate entities (Epley, et al. [29])and form attachments to chatbots (Brandtzaeg, et al., [15] Skjuve, et al. [86],[87]), apps that take advantage of such AI technologies pose significant risks for fostering dependencies, particularly among vulnerable individuals. This creates financial incentives for application developers, who can design their apps to foster strong therapeutic alliances with users to retain customers. Given this, it is crucial to develop and implement regulations that pay attention to the need to have apps that, on the one hand, foster patients' confidence in managing their conditions but, on the other hand, avoid fostering therapeutic alliances aimed at retaining users.

2.4. EFFICACY

As noted in Section 1, the regulatory regime for mental health apps is diverse and still developing. Even for those interventions that have a broad consensus in the scientific and clinical community, such as CBT or psychiatric treatments that use dopamine reuptake inhibitors, to name a few, the epistemic basis is significantly lower than in the strictly medical domain. There are several reasons for this situation, but the one that may be particularly important to mention in the context of mental health apps is the abundance of possibly confounding factors in any one intervention and the methodological difficulties that face any attempt at identifying the "active ingredient" in any successful therapy (maybe cite the relevant paper where this term of art is introduced). In some cases, the best possible explanation of success lies in the relationship that forms between the therapist and the patient, rather than the specific therapeutic techniques of the intervention. In this context, more experimental approaches, including interventions with apps, face the additional difficulty of having no identifying baseline that can distinguish those apps that work because of the way they embody insights from psychiatry or psychology from those that work for unknown reasons or because of the equivalent of the placebo effect. This lack of a baseline for assessing efficacy can arguably make some of the ethical issues worse and it can be linked to at least two distinct problems for reconceptualizing the ethical guidelines for mental health apps. First, it is unclear what may be called therapy or a mental health intervention if we have no way of being

specific about the causal chain that leads from an intervention to an outcome. This introduces affordances for abuse, lobbying, and short-term interests overriding duties toward patients, especially those who are vulnerable in all the ways that were previously discussed. People looking to help themselves via a mental health app will be left at the mercy of the good faith of app developers and business interests who have obligations (both legal and ethical) primarily to investors and their employees. At worst, users of apps, in cases when their conditions may turn worse, will have little or no recourse to sue, complain, or generally hold anyone responsible. Seeking the assistance of a professional therapist reassures individuals of being covered by established normative frameworks that regulate the profession, assuming these frameworks are present in their location. These standards ensure a level of efficacy and governance over the patient-therapist relationship, as outlined, for example, in the APA code of conduct. In short, the lack of standards for the efficacy of mental health apps will have the downstream effect of complicating the normative landscape significantly in the context of certifications, classifications, and standards, without any prospect of resolution before negative outcomes affect users. A second, and perhaps more severe, issue stemming from unclear standards of efficacy for mental health apps is the potential to keep users engaged with the app longer than beneficial—a problem outlined previously in Section 2.3. While developers of mental health apps have clear financial motives for retained use, the users of such apps may have a clear goal of discontinued use. This needs to be kept in mind when developing mental health apps so that there is clear regulation that protects users from undesired dependence on mental health apps and clear strategies on how to ensure that users can easily discontinue the app when their individual health goals are met, even if those are not well defined for them.

This last issue that changes in target symptomatology (be it anxiety, depression, mood, etc.) can be traced over time in ways that are convergent with standard measures and that there is a clear, clinically significant goal that is targeted, even if it is not clear to the user. Thus, when the predefined goal is met, used or, in cases where the app is a replacement for clinical intervention, treatment is ended. In a similar vein, if there is no improvement in symptomatology over a predefined period despite continued use, the user should be alerted that the mental health app does not have the expected effect—however, that is defined by standards of care for the mental health concern at stake—and referred to other mental health resources. Here, it is important to stress that the use of an app cannot constitute a measure of effect (the logic “users use this; thus it must be helpful”), since interaction with features of the app (e.g., tracking and goal setting) may not directly correlate with improvements in mental health outcomes (Zhang [103]).

3. CONCLUSION

In conclusion, this white paper emphasizes the urgent need for a more inclusive and diverse approach to the ethical regulation of mental health apps. It argues that the current regulatory landscape is heavily influenced by Western ethical frameworks and does not adequately address the needs and perspectives of non-Western users, who represent a significant and growing portion of the global user base.

The white paper proposes two key strategies to address these gaps. First, it advocates for the inclusion of non-Western ethical traditions in the development and regulation of mental health apps. This approach recognizes the cultural diversity of users and helps to ensure that the apps are not solely guided by Western ethical notions. Second, it calls for the integration of insights from feminist and disability studies, which have been historically marginalized in ethical theorizing. These perspectives can provide valuable insights into the experiences of vulnerability, social relationships, and diverse embodiments, which are crucial in the context of mental health.

This white paper also highlights the limitations of the current regulatory status quo, which is geographically and culturally limited and heavily biased toward Western perspectives. It calls for a more comprehensive and inclusive approach to regulation, which considers the diverse needs and perspectives of users from around the world.

By broadening the ethical frameworks that inform the design and regulation of mental health apps, this white paper argues that we can create more effective, inclusive, and ethically sound digital solutions for mental health care. This approach not only benefits users from diverse cultural backgrounds but also contributes to the overall quality and effectiveness of mental health apps. The white paper concludes with a call to action for developers and regulators to think more widely and critically about the ethical frameworks from which they draw to better serve the growing global audience of mental health app users.

Another aspect is the influence of big tech. Their values and power dynamics, which are inscribed in technology, play a significant role in shaping mental health. Nosthoff and Maschweski [73] describe how companies like Google and Apple, with their wearables and apps, undertake a “mapping of human health” (as has already been done with Google Earth or Google Maps)—where maps serve not only for orientation in the world but also make it navigable and manageable.

(This leads to the neglect of the needs of minority groups that are underrepresented in big tech and power dynamics, etc.)

4. REFERENCES

The following sources have either been referenced within this paper or may be useful for additional reading:

[1] Aizenberg, E., and J. van den Hoven, "Designing for human rights in AI," *Big Data Soc.*, vol. 7, no. 2, Jul. 2020, Art. no. 205395172094956.

[2] Alqahtani, F., S. Meier, and R. Orji, "Personality-based approach for tailoring persuasive mental health applications," *User Model. User-Adapted Interact.*, vol. 32, no. 3, pp. 253–295, Jul. 2022.

[3] Andrade, L. H., Alonso, J., Mneimneh, Z., Wells, J. E., Al-Hamzawi, A., Borges, G., ... & Kessler, R. C. (2014). Barriers to mental health treatment: results from the WHO World Mental Health surveys. Psychological medicine, 44(6), 1303-1317.

[4] Appelbaum, P. S. and T. Grisso, "The MacArthur treatment competence study. I: Mental illness and competence to consent to treatment," in *Clinical Forensic Psychology and Law*. Evanston, IL, USA: Routledge, 2019, pp. 129–156.

[5] Arora, P., *The Next Billion Users: Digital Life Beyond the West*. Cambridge, MA, USA: Harvard Univ. Press, 2019.

[6] Austin, A. and R. Goodman, "Perceptions of transition-related health and mental health services among transgender adults," *J. Gay Lesbian Social Services*, vol. 30, no. 1, pp. 17–32, Jan. 2018, doi: 10.1080/10538720.2017.1408515.

[7] Bakker, D., et al., "Mental health smartphone apps: Review and evidence-based recommendations for future developments," *JMIR Mental Health*, vol. 3, no. 1, Mar. 2016, Art. no. e4984.

[8] Baldwin, C., "Technology, dementia and ethics: Rethinking the issues," *Disab. Stud. Quart.*, vol. 25, no. 3, Jun. 2005, Art. no. 3, doi: 10.18061/dsq.v25i3.583.

[9] Bannon, L. J. and P. Ehn, "Design: Design matters in participatory design," in *Routledge International Handbook of Participatory Design*. Evanston, IL, USA: Routledge, 2012, pp. 37–63.

[10] Bassil, K., "Balancing the double-edged implications of AI in psychiatric digital phenotyping," *Amer. J. Bioethics*, vol. 24, no. 2, pp. 113–115, Feb. 2024.

[11] Benjamin, R., *Race After Technology: Abolitionist Tools for the New Jim Code*. Hoboken, NJ, USA: Wiley, 2019.

[12] Bergamin, J., et al., “Defining autonomy in psychiatry,” *Frontiers Psychiatry*, vol. 13, May 2022, Art. no. 801415.

[13] Bergin, A. D. G., et al., “Identifying and categorizing adverse events in trials of digital mental health interventions: Narrative scoping review of trials in the international standard randomized controlled trial number registry,” *JMIR Mental Health*, vol. 10, Feb. 2023, Art. no. e42501, doi: 10.2196/42501.

[14] Berman, B. J., “Artificial intelligence and the ideology of capitalist reconstruction,” *AI Soc.*, vol. 6, no. 2, pp. 103–114, Apr. 1992, doi: 10.1007/bf02472776.

[15] Brandtzaeg, P. B., Skjuve, M., & Følstad, A. (2022). My AI friend: How users of a social chatbot understand their human–AI friendship. *Human Communication Research, 48(3),* 404-429.

[16] Brey, P., “Design for the value of human well-being,” in *Handbook of Ethics and Values in Technological Design: Sources, Theory, Values, and Application Domains*, J. van den Hoven, P. Vermaas, and I. van de Poel, Eds. Dordrecht, The Netherlands: Springer, 2014, pp. 1–14.

[17] Burr, C. and J. Morley, “Empowerment or engagement? Digital health technologies for mental healthcare,” in *The 2019 Yearbook of the Digital Ethics Lab*, 2019, pp. 67–88.

[18] Burr, C. and R. Powell, “Trustworthy assurance of digital mental healthcare,” The Alan Turing Inst., London, U.K., Tech. Rep., Nov. 2022, doi: 10.5281/zenodo.7107200.

[19] Burr, C., M. Taddeo, and L. Floridi, “The ethics of digital well-being: A thematic review,” *Sci. Eng. Ethics*, vol. 26, pp. 2313–2343, Jan. 2020, doi: 10.1007/s11948-020-00175-8.

[20] Burr, C., et al., “Digital psychiatry: Risks and opportunities for public health and wellbeing,” *IEEE Trans. Technol. Soc.*, vol. 1, no. 1, pp. 21–33, Mar. 2020.

[21] Chhabra B. and Kapadia, S. (2023). Mental health policies in queer community: are we doing enough?. Journal of Loss and Trauma, 28(2), 182-186.

[22] Coleman, E., et al., “Standards of care for the health of transgender and gender diverse people, version 8,” *Int. J. Transgender Health*, vol. 23, no. 1, pp. S1–S259, Aug. 2022, doi: 10.1080/26895269.2022.2100644.

[23] Costanza-Chock, S., “Design justice: Towards an intersectional feminist framework for design theory and practice,” in *Proc. DRS, Catalyst*, Jun. 2018, pp. 1–13.

[24] Crawford,K., *The Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence*. New Haven, CT, USA: Yale Univ. Press, 2021.

[25] Deb, K. S., et al., "Is India ready for mental health apps (MHApps)? A quantitative-qualitative exploration of caregivers' perspective on smartphone-based solutions for managing severe mental illnesses in low resource settings," *PLoS ONE*, vol. 13, no. 9, Sep. 2018, Art. no. e0203353.

[26] Deegan, P. E., "Recovery: The lived experience of rehabilitation.," *Psychosocial Rehabil. J.*, vol. 11, no. 4, pp. 11–19, Apr. 1988, doi: 10.1037/h0099565.

[27] Dennis, M. J. and R. Clancy (2022). Intercultural Ethics for Digital Well-Being- Identifying Problems and Exploring Solutions. Digital Society. Online first.

[28] Drake, R. E. and R. Whitley, "Recovery and severe mental illness: Description and analysis," *Can. J. Psychiatry*, vol. 59, no. 5, pp. 236–242, 2014.

[29] Epley, N., Waytz, A., & Cacioppo, J. T. (2007). On seeing human: a three-factor theory of anthropomorphism. *Psychological review*, 114(4), 864.

[30] Faissner, M., et al., "Detecting your depression with your smartphone?—An ethical analysis of epistemic injustice in passive self-tracking apps," *Ethics Inf. Technol.*, vol. 26, no. 2, p. 28, Jun. 2024, doi: 10.1007/s10676-024-09765-7.

[31] Figueroa, C. A., et al., "The need for feminist intersectionality in digital health," *Lancet Digit. Health*, vol. 3, no. 8, pp. e526–e533, Aug. 2021.

[32] Forbes, M. K., et al., "Reconstructing psychopathology: A data-driven reorganization of the symptoms in DSM-5," Tech. Rep., 2023, doi: 10.31234/osf.io/7um9a.

[33] Fricker, M., *Testimonial Injustice. Contemporary Epistemology: An Anthology*, 2019, pp. 149–163.

[34] Friedman, B., D. G. Hendry, and A. Borning, "A survey of value sensitive design methods," *Found. Trends Hum.-Comput. Interact.*, vol. 11, no. 2, pp. 63–125, 2017, doi: 10.1561/1100000015.

[35] Friedman, B. and H. Nissenbaum, "Bias in computer systems," *ACM Trans. Inf. Syst.*, vol. 14, no. 3, pp. 330–347, Jul. 1996.

[36] Fuchs, T., *Das Gehirn—Ein Beziehungsorgan. Eine Phänomenologisch-Ökologische Konzeption*, 5th ed. Stuttgart, Germany: Kohlhammer, 2017.

[37] Fung, P and H. Etienne, “Confucius, cyberpunk and Mr. Science: Comparing AI ethics principles between China and the EU,” *AI Ethics*, vol. 3, no. 2, pp. 505–511, May 2023, doi: 10.1007/s43681-022-00180-6.

[38] Garry, A., S. J. Khader, and A. Stone, *The Routledge Companion to Feminist Philosophy*. New York, NY, USA: Routledge, 2017.

[39] Gebru, T., “Oxford handbook on AI ethics book chapter on race and gender,” 2019, *arXiv:1908.06165*.

[40] Gebru, T., “Race and Gender,” in *The Oxford Handbook of Ethics of AI*, M. D. Dubber, F. Pasquale, and S. Das, Eds. New York, NY, USA: Academic, 2020, doi: 10.1093/oxfordhb/9780190067397.013.16.

[41] Geurtzen, N., Karremans, J. C., Keijsers, G. P., & Hutschemaekers, G. J. (2023). "I Need You!" Patients’ Care Dependency Patterns During Psychotherapy for Personality Disorders and Its Association with Symptom Reduction and Wish for Treatment Continuation. *Journal of Contemporary Psychotherapy*, 53(3), 257-267.

[42] Geurtzen, N., Keijsers, G. P., Karremans, J. C., & Hutschemaekers, G. J. (2018). Patients’ Care Dependency in Mental Health Care: Development of a Self-Report Questionnaire and Preliminary Correlates. *Journal of Clinical Psychology*, 74(7), 1189-1206.

[43] Gilligan, C., *In a Different Voice: Psychological Theory and Women’s Development*. Cambridge, MA, USA: Harvard Univ. Press, 1993.

[44] Glanert, S., Sürig, S., Grave, U., Fassbinder, E., Schwab, S., Borgwardt, S., & Klein, J. P. (2021). Investigating care dependency and its relation to outcome (ICARE): Results from a naturalistic study of an intensive day treatment program for depression. Frontiers in Psychiatry, 12, 644972.

[45] Goldsmith, S., *Universal Design*. Evanston, IL, USA: Routledge, 2007.

[46] Gómez-Carrillo A., et al., “Restoring the missing person to personalized medicine and precision psychiatry,” *Frontiers Neurosci.*, vol. 17, Feb. 2023, Art. no. 1041433, doi: 10.3389/fnins.2023.1041433.

[47] Hagendorff, T., “Blind spots in AI ethics,” *AI Ethics*, vol. 2, no. 4, pp. 851–867, Nov. 2022, doi: 10.1007/s43681-021-00122-8.

[48] Hamdoun, S., et al., “AI-based and digital mental health apps: Balancing need and risk,” *IEEE Technol. Soc. Mag.*, vol. 42, no. 1, pp. 25–36, Mar. 2023, doi: 10.1109/MTS.2023.3241309.

[49] Hamraie, A. and Fritsch, K. (2019). Crip technoscience manifesto. Catalyst: Feminism, theory, technoscience, 5(1), 1–33.

[50] Hendl, T., B. Jansky, and V. Wild, "From design to data handling. Why mHealth needs a feminist perspective," in *Feminist Philosophy of Technology* (Techno: Phil—Aktuelle Herausforderungen Der Technikphilosophie), J. Loh and M. Coeckelbergh, Eds., 2019, pp. 77–103, doi: 10.1007/978-3-476-04967-4_5.

[51] Henrich, J., *The WEIRDest People in the World: How the West Became Psychologically Peculiar and Particularly Prosperous*. Farrar, Straus and Giroux, 2020.

[52] Henrich, J., S. J. Heine, and A. Norenzayan, "The weirdest people in the world?" *Behav. Brain Sci.*, vol. 33, nos. 2–3, pp. 61–83, Jun. 2010, doi: 10.1017/s0140525x0999152x.

[53] Herrera-Ferrá, K., A. Salles, and L. Cabrera. (Apr. 10, 2018). *Global Neuroethics and Cultural Diversity: Some Challenges to Consider*. The Neuroethics Blog. [Online]. Available: http://www.theneuroethicsblog.com/2018/04/global-neuroethics-and-cultural.html

[54] Imrie, R. and P. Hall, *Inclusive Design: Designing and Developing Accessible Environments*. New York, NY, USA: Taylor & Francis, 2003.

[55] Jacobs, N., "Capability sensitive design for health and wellbeing technologies," *Sci. Eng. Ethics*, vol. 26, no. 6, pp. 3363–3391, Dec. 2020.

[56] Jacobs, N., "Two ethical concerns about the use of persuasive technology for vulnerable people," *Bioethics*, vol. 34, no. 5, pp. 519–526, Jun. 2020.

[57] Kazdin, A. E., and Rabbitt, S. M. (2013). Novel models for delivering mental health services and reducing the burdens of mental illness. Clinical Psychological Science, 1(2), 170-191.

[58] Kim, J. Y., et al., "Development and preliminary testing of health equity across the AI lifecycle (HEAAL): A framework for healthcare delivery organizations to mitigate the risk of AI solutions worsening health inequities," *PLOS Digit. Health*, vol. 3, no. 5, May 2024, Art. no. e0000390, doi: 10.1371/journal.pdig.0000390.

[59] Kim, U., G. Yang, and K.-K. Hwang, *Indigenous and Cultural Psychology: Understanding People in Context*. Springer, 2006.

[60] Kinsella, M., "Fostering the therapeutic alliance: Recognizing autonomy's dialogical antecedents," *Phil. Psychol.*, vol. 32, no. 3, pp. 332–356, Apr. 2019.

[61] Kinsella, M., "Exercising leadership within the therapeutic alliance: An autonomy-grounded perspective," *J. Humanistic Psychol.*, vol. 63, no. 1, pp. 103–125, 2023.

[62] Kohn, R., et al., “The treatment gap in mental health care,” *Bull. World Health Org.*, vol. 82, no. 11, pp. 858–866, 2004.

[63] Kurosu, M., *Human Centered Design*. Berlin, Germany: Springer, 2009, pp. 94–103.

[64] Laestadius, L., et al., “Too human and not human enough: A grounded theory analysis of mental health harms from emotional dependence on the social chatbot Replika,” *New Media Soc.*, Dec. 2022, Art. no. 146144482211420.

[65] Lin, J. and H. Baumeister, “Internet- und mobilebasierte interventionen in der psychotherapie,” *Public Health Forum*, vol. 23, no. 3, pp. 176–179, Sep. 2015, doi: 10.1515/pubhef-2015-0063.

[66] Lu, L., “Chinese well-being,” in *Oxford Handbook of Chinese Psychology*, M. H. Bond, Ed. New York, NY, USA: Oxford Univ. Press, 2012, pp. 327–342.

[67] Mackenzie, C., W. Rogers, and S. Dodds, “Introduction: What is vulnerability and why does it matter for moral theory,” in *Vulnerability: New Essays in Ethics and Feminist Philosophy*, 2014, pp. 1–29.

[68] Mackenzie, C. and N. Stoljar, *Relational Autonomy: Feminist Perspectives on Autonomy, Agency, and the Social Self*. London, U.K.: Oxford Univ. Press, 2000.

[69] Moitra, M., Santomauro, D., Collins, P. Y., Vos, T., Whiteford, H., Saxena, S., & Ferrari, A. J. (2022). The global gap in treatment coverage for major depressive disorder in 84 countries from 2000–2019: A systematic review and Bayesian meta-regression analysis. PLoS medicine, 19(2), e1003901.

[70] Moitra, M., Owens, S., Hailemariam, M., Wilson, K. S., Mensa-Kwao, A., Gonese, G., ... & Collins, P. Y. (2023). Global mental health: Where we are and where we are going. Current Psychiatry Reports, 25(7), 301–311.

[71] Nickel, P.J., et al., “Justice and empowerment through digital health: Ethical challenges and opportunities,” *Digit. Soc.*, vol. 2, no. 3, p. 42, Dec. 2023.

[72] Norton, M. I., D. Mochon, and D. Ariely, “The IKEA effect: When labor leads to love,” *J. Consum. Psychol.*, vol. 22, no. 3, pp. 453–460, 2012.

[73] Nosthoff, A. V., and F. Maschewski, *Die Gesellschaft Der Wearables. Digitale Verführung Und Soziale Kontrolle*, 2019.

[74] Oosterlaken, I., Inappropriate artefact, unjust design? Human diversity as a key concern in the capability approach and inclusive design," in *The Capability Approach, Technology and Design*. Dordrecht, The Netherlands: Springer, 2012, pp. 223–244.

[75] Palmer, V. J., "The participatory zeitgeist in health care: It is time for a science of participation," *J. Participatory Med.*, vol. 12, no. 1, Jan. 2020, Art. no. e15101, doi: 10.2196/15101.

[76] Parsakia, K., "The effect of chatbots and AI on the self-efficacy, self-esteem, problem-solving and critical thinking of students," *Health Nexus*, vol. 1, no. 1, pp. 71–76, Jan. 2023.

[77] Perez, C. C., *Invisible Women: Data Bias in a World Designed for Men*. Harry N. Abrams, 2019.

[78] Polk, K., "The matrix, evolution, and improving work-group functioning with Ostrom's eight design principles," in *The ACT Matrix: A New Approach to Building Psychological Flexibility Across Settings and Populations*, 2014, pp. 235–250.

[79] Ramos, G., et al., "Considerations of diversity, equity, and inclusion in mental health apps: a scoping review of evaluation frameworks," *Behav. Res. Therapy*, vol. 147, Dec. 2021, Art. no. 103990, doi: 10.1016/j.brat.2021.103990.

[80] Rubeis, G., "E-mental health applications for depression: An evidence-based ethical analysis," *Eur. Arch. Psychiatry Clin. Neurosci.*, vol. 271, no. 3, pp. 549–555, Apr. 2021, doi: 10.1007/s00406-019-01093-y.

[81] Sax, M., N. Helberger, and N. Bol, "Health as a means towards profitable ends: mHealth apps, user autonomy, and unfair commercial practices," *J. Consum. Policy*, vol. 41, pp. 103–134, Jun. 2018.

[82] Schukking, F., T. Veloz, and S. Leijnen, "Broadening the idea of intelligence in AI and the Noosphere," Tech. Rep., 2023, doi: 10.13140/RG.2.2.22830.32321.

[83] Shew, A., "Ableism, technoableism, and future AI," *IEEE Technol. Soc. Mag.*, vol. 39, no. 1, pp. 40–85, Mar. 2020.

[84] Shumanov, M. and L. Johnson, "Making conversations with chatbots more personalized," *Comput. Hum. Behav.*, vol. 117, Apr. 2021, Art. no. 106627.

[85] Singh, S. and R. Sagar, "Time to have effective regulation of the mental health apps market: Maximize gains and minimize harms," *Indian J. Psychol. Med.*, vol. 44, no. 4, pp. 399–404, Jul. 2022.

[86] Skjuve, M., Følstad, A., Fostervold, K. I., & Brandtzaeg, P. B. (2022). A longitudinal study of human–chatbot relationships. *International Journal of Human-Computer Studies*, *168*, 102903.

[87] Skjuve, M., Følstad, A., Fostervold, K. I., & Brandtzaeg, P. B. (2021). My chatbot companion-a study of human-chatbot relationships. *International Journal of Human-Computer Studies*, *149*, 102601.

[88] Slade, M., "Peer relationships," in *Personal Recovery and Mental Illness: A Guide for Mental Health Professionals*. Cambridge, U.K.: Cambridge Univ. Press, 2009, pp. 103–113, doi: 10.1017/cbo9780511581649.012.

[89] Smith, P. B., "On the distinctiveness of Chinese psychology; or: Are we all Chinese?" in *Oxford Handbook of Chinese Psychology*, M. H. Bond, Ed. London, U.K.: Oxford Univ. Press, 2012, pp. 699–710.

[90] Spencer-Rodgers, J., et al., "Dialectical self-esteem and east-west differences in psychological Well-Being," *Personality Social Psychol. Bull.*, vol. 30, no. 11, pp. 1416–1432, Nov. 2004.

[91] Stappenbelt, B., "Ethics in engineering: Student perceptions and their professional identity development," *J. Technol. Sci. Educ.*, vol. 3, no. 1, pp. 86–93, Feb. 2013, doi: 10.3926/jotse.51.

[92] Suh, E. M., "Self, the hyphen between culture and subjective well-being," in *Culture and Subjective Well-Being*. Cambridge, MA, USA: MIT Press, 2000, pp. 63–86.

[93] Talbert, T. L. (2023). Transgender Health: The Present and the Future. In Cultural Issues in Healthcare: Emerging Challenges and Opportunities (pp. 117-132). Cham: Springer International Publishing.

[94] Tiberius, V., "Cultural differences and philosophical accounts of well-being," *J. Happiness Stud.*, vol. 5, no. 3, p. 293, 2004.

[95] Tu, W. and H.-P. Y. Wei-Ming, *Confucian Thought: Selfhood as Creative Transformation*. SUNY Press, 1985.

[96] Varma, T., et al., "Diversity in clinical research: Public health and social justice imperatives," *J. Med. Ethics*, vol. 49, no. 3, pp. 1–4, 2022, doi: 10.1136/medethics-2021-108068.

[97] Walter, J.K. and L. F. Ross, "Relational autonomy: Moving beyond the limits of isolated individualism," *Pediatrics*, vol. 133, no. 1, pp. S16–S23, Feb. 2014.

[98] Wilson, C. and L. A. Cariola, "LGBTQI+ youth and mental health: A systematic review of qualitative research," *Adolescent Res. Rev.*, vol. 5, no. 2, pp. 187–211, Jun. 2020, doi: 10.1007/s40894-019-00118-w.

[99] Wilson, D.S., E. Ostrom, and M. E. Cox, "Generalizing the core design principles for the efficacy of groups," *J. Econ. Behav. Org.*, vol. 90, pp. S21–S32, Jun. 2013, doi: 10.1016/j.jebo.2012.12.010.

[100] Wilson, D.S. and D. J. Snower, "Rethinking the theoretical foundation of economics I: The multilevel paradigm," *Economics*, vol. 18, no. 1, Feb. 2024, Art. no. 20220070, doi: 10.1515/econ-2022-0070.

[101] World Health Organization. SIXTY-SIXTH WORLD HEALTH ASSEMBLY WHA66.8 Agenda item 13.3 27 May 2013 “Comprehensive mental health action plan 2013–2020.” URL: https://apps.who.int/gb/ebwha/pdf_files/WHA66/A66_R8-en.pdf.

[102] Wykes, T., J. Lipshitz, and S. M. Schueller, “Towards the design of ethical standards related to digital mental health and all its applications,” *Current Treatment Options Psychiatry*, vol. 6, no. 3, pp. 232–242, Sep. 2019.

[103] Zhang. (2019). Clinically Meaningful Use of Mental Health Apps and its Effects on Depression: Mixed Methods Study. Available: https://www.jmir.org/2019/12/e15644/

[104] Ziesche, S., “Stimuli from selected non-Western approaches to AI ethics,” Tech. Rep., 2023.

[105] *Zum Einfluss Bildgebender Verfahren Der Hirnforschung Auf Erziehungswissenschaftliche Diskurse*, Sci. Stud., Bielefeld, Germany, 2012.

RAISING THE WORLD'S STANDARDS

3 Park Avenue, New York, NY 10016-5997 USA http://standards.ieee.org

Tel.+1732-981-0060 Fax+1732-562-1571