ORIGINAL PAPER

Examining Ethical and Social Implications of Digital Mental Health Technologies Through Expert Interviews and Sociotechnical Systems Theory

Jonathan Adams[1] [ID]

Abstract

This paper aims to understand how science and technology experts working in the digital mental health field interpret the ethical and social implications of its technologies, combining an 'expert interview' methodology with insights from sociotechnical systems theory. Following recruitment of experts in science and technology fields who had experience of supporting the development of DMH interventions, 11 semi-structured interviews were conducted and analyzed in accordance with the Framework Method. A single theme of 'complexity of implications' is presented here and divided into the categories of 'implications for users', 'implications for healthcare professionals and systems', and 'implications for society'. Participants identified a range of ethical and social implications of digital mental health technologies at the three different levels, which this discussion relates to three key aspects of complex sociotechnical systems identified in existing theoretical work. These are 'heterogeneity', 'interdependence' and 'distribution', each of which raises important questions for future research about how complex values, relationships and responsibilities should be negotiated in digital mental health. The paper concludes that this study's approach provides a model for understanding the implications of digital health more broadly, with participants' combined experience and knowledge shedding light on key interventions at the forefront of digitalization in healthcare.

Keywords Digital health · Mental health · Expert interviews · Ethical and social issues · Sociotechnical systems

✉ Jonathan Adams
   jonathan.adams@medisin.uio.no

1   Centre for Medical Ethics, Institute of Health and Society, Faculty of Medicine, University of Oslo, Kirkeveien 166, Fredrik Holsts hus, 0450 Oslo, Norway

🍂 Springer

1 Introduction

> [T]he machine has no feelings, it feels no fear and no hope, which only disturb, it has no wishes with regard to the result, it operates according to the pure logic of probability. For this reason I assert that the robot perceives more accurately than man […] (Frisch, 1959, p. 76).

The above quotation from Max Frisch's unsentimental engineer in the classic novel *Homo Faber* represents, in caricature, what is sometimes thought to be the world-view of those seeking to usher new technologies further into emotionally fraught areas of human life. The aim of this study is to move beyond such broad strokes and delve empirically into the views of science and technology experts who work on digital technologies specifically designed to improve mental health. The research will advance our understanding of the ethical and social implications of digital mental health (DMH) technologies from a hands-on perspective and generate theoretical linkages with literature from ethics and the social sciences.

Drawing on qualitative interviews with 11 experts from a range of science and technology disciplines who have engaged in the development of DMH interventions, this paper relies on sociotechnical systems theory to discuss challenges posed by complexity in DMH systems. The study shows how science and technology experts identify and conceptualize the ethical and social implications of DMH in thinking through potential risks and failures of DMH at the individual, organizational, and societal levels. Specifically, this paper argues that participants identified a range of implications that can be related to features of heterogeneity, interdependence, and distributed control in DMH systems and suggest future directions for debating values, relationships, and responsibilities in DMH.

2 Background

2.1 Digital Mental Health as a Complex Sociotechnical System

DMH represents a growing area of technology development that is characterized by the use of information and communication technology for "mental health assessment, support, prevention, and treatment" (Wies et al., 2021, p. 2). Hereafter, these uses will be described collectively as 'DMH interventions'. With DMH approaches ranging from the relatively low-tech to the sophisticated (Lederman & D'Alfonso, 2021), the most popular routes, according to Baños et al. (2022), fall into three categories:

1. *Internet-based* interventions provide therapeutic programs to meet specific needs. Often broken up into online modules or lessons with homework assignments, they can be accessed using web-enabled devices such as mobile or computer devices (Andersson & Titov, 2014).
2. *Smartphone apps* can perform tasks throughout clinical care by tracking mental health symptoms passively (via movement sensors, speech detectors, etc.),

supporting patient self-management (using medication reminders or in-app feedback) or transferring information between a user and their healthcare provider (National Institute of Mental Health, 2019).

3. *Virtual reality* (VR) and *augmented reality* technologies for mental health provide a sophisticated way of treating phobias and other mental disorders through exposure therapy using three-dimensional simulations (Emmelkamp & Meyerbröker, 2021).

This list of applications is not exhaustive (for example, apps and systems for hospital-based monitoring and decision support appear difficult to place within it), but the categories provide a rough guide to the major types of DMH intervention, albeit only with the necessary acknowledgement of one overarching technological development. Cutting across these three categories, artificial intelligence (AI) can be used to enhance treatment options using internet-based, app-based or VR methods or to enable potentially more precise diagnostic and monitoring tools (Bickman, 2020). With mental healthcare dominated by subjective assessments and outpatient care, Lovejoy (2019, p. 1) suggests, AI may offer "greater objectivity and may have better predictive value" while allowing "increased monitoring in the community" via wearables and mobile devices.

DMH technologies are partly important to scrutinize because recent years have seen a concerted push for investment and innovation in DMH by actors such as the World Economic Forum, the National Health Service, and private investors (Hollis et al., 2015; Jankovic et al., 2020; World Economic Forum, 2021). This trend has been accelerated by the impacts of COVID-19 and an increasing recognition of the wide gap between those who need mental healthcare and those who receive it via conventional means (Balcombe & Leo, 2021; Carter et al., 2021). Research in DMH has changed dramatically in the last decade, with a recent analysis showing that in the period from 2008 to 2010, when the field was beginning to accelerate, its focus was on promoting "self-management and the mitigation of mental health impacts from physical sickness" (Timakum et al., 2022, p. 13). Yet more recently, the authors report, the scope of research papers has expanded to cover DMH approaches to lifestyle interventions, mental illness prevention, and the use of AI tools to reduce the overall mental health burden. One way of thinking about large-scale transformations in DMH is through the small but growing number of contributions that see digital health interventions as "essentially sociotechnical systems" (Mohr et al., 2017, p. 2) dependent on patient interactions and human support for their technologies' performance. As this paper will explore, the lens of sociotechnical systems theory (SST) is useful for understanding a range of ethical and social aspects of the complex and shifting DMH domain but first requires some explanation in terms of the range of systems and interactions that it handles.

SST originated with Trist and Bamforth (1951) at the Tavistock Institute in the United Kingdom (UK) and studies the "synergistic combination of humans, machines, environments, work activities and organisational structures and processes that comprise a given enterprise" (Carayon et al., 2015, p. 550). Within SST, these diverse components are organized into two interwoven strands that make up a sociotechnical system: the social sub-system and the technical sub-system (Mumford, 2006). The

former consists of "the people, their attitudes, values, norms, histories and competencies" and the latter "human artefacts such as physical structures, buildings and other pieces of technology, but also policies, strategies and procedures" (Troxler & Lauche, 2003, p. 3). These subsystems are highly interactive, and overall system performance requires 'joint optimization', whereby the demands of the technical match the demands of the social and vice versa (Leonardi, 2012). When combined with research on DMH, an SST approach reveals a range of social and technical components. Sittig and Singh's (2015) influential model identifies eight dimensions that make up a health technology system's "design, development, use, implementation, and evaluation" (p. 63). These are:

1. Hardware and software infrastructure
2. Clinical content
3. The human-computer interface
4. People
5. Workflows and communication
6. Organizational policies and cultures
7. External pressures and regulations
8. Measurement and monitoring

Clearly, DMH technologies incorporate different varieties of hardware (e.g., laptops, mobile phones and sensors) and software (on the internet or apps) to store *clinical content* (e.g., educational text, activities, or patient data). These facilitate different types of interaction with a *human-computer interface* by a user—often a patient—who is one of a number of relevant *people* who design, develop, use, implement, and evaluate the technology (Balcombe & De Leo, 2022). The 'workflow and communication' dimension of DMH involves coordination of people and tasks across organizations, including integration of DMH's technical aspects into "the workflows of clinicians who are expected to deliver them" (Graham et al., 2019, p. 1223). Tightly bound to the 'people' dimension are the *organizational policies and cultures* that set priorities for all other dimensions and even dictate "readiness for DMH adoption" (Ganapathy et al., 2021, p. 1339) in terms of resources and procedures. *External* forces may then support or constrain the performance and growth of a system through increased restrictions or incentives that individuals and organizations will navigate (Armontrout et al., 2016). Finally, *measurement and monitoring* practices assess the availability, use, performance, and consequences of a system, with studies of failed projects "[providing] important insights into how digital interventions should be planned and conducted" (Muuraiskangas et al., 2016, p. 11).

Sittig and Singh's model allows one to identify individual components within their dimensions and sub-systems, but the authors suggest that they cannot be understood by analyzing their individual elements as they are "complex adaptive systems" (2015, p. 63). They thereby draw on a body of knowledge that has been incorporated into SST from the physical and biological sciences and is known as 'complexity theory' (Bauer & Herder, 2009). This approach identifies complexity as a property of systems that consist of multiple, heterogenous, interdependent components that interact in dynamic ways and, moreover, "come to be defined by those interactions"

(Cohn et al., 2013, p. 42). What distinguishes complexity from the mere synergy considered above is the way in which internal relationships in a complex system give rise to emergent features via 'self-organization' (De Wolf & Holvoet, 2005). These causal effects are crucially non-linear, allowing emergent properties to arise at the macro level due to small changes that occur elsewhere in the system under a limited set of conditions. Where novel features emerge to ensure the survival or flourishing of a complex system in response to changes in its environment, the system qualifies as not merely complex but what is known as a 'complex adaptive system' (Holland, 2006). In health systems research, Greenhalgh and Papoutsi (2018) suggest that this picture of causality forms part of a new paradigm informed by complexity that is distinct from the paradigm given by traditional models of cause and effect.

Though DMH interventions are mentioned in studies of complexity (Rosen et al., 2020), the core features of complex systems above have yet to be applied to DMH in depth, and while this paper alone cannot fill this gap, it can provide some reasons to think of DMH interventions as complex systems. First, as a multidisciplinary enterprise bridging psychological, medical, and computing specialisms and the education, healthcare, and technology sectors, DMH involves significant cultural and conceptual heterogeneity (Nordgreen et al., 2021). In SST, different sources of expertise and organizational commitments are thought to mean divergence on crucial values or procedures (Carayon, 2006), and it is only "[i]n the best of cases" that "different participants combine their expertise in creative, effective ways, often compromising goals and principles for the greater good" (Norman & Stappers, 2015, p. 87). Second, interdependence plays a crucial role in DMH, with Iorfino et al.'s (2021) systems modelling identifying that the use of technology to coordinate interactions within a mental health system can have greater benefits than introducing separate digital alternatives. This is due partly to mental health interventions' general reliance on "collaboration among practitioners, other components of the healthcare service, communities, caregivers, patients, and their dependents" (Aryana et al., 2019, p. 407) but also the social infrastructure that Shaw and Donia (2021, p. 4) identify as laying "material foundations that make everyday life possible". Finally, through efforts to advance co-design and co-production of DMH technologies, control over the design and implementation of DMH interventions is increasingly distributed among stakeholders (Bevan Jones et al., 2020), with "order, organisation and control that is distributed and locally generated […] rather than centrally produced" (Ladyman et al., 2013, p. 38). This decreases reliance on centralized strategies for coping with disturbances (Weijnen et al., 2008), but when combined with non-linearity, it also means that any part of the system can produce "dramatic, widespread, unpredictable effects" (Effken, 2002, p. 61).

2.2 Analyzing the Ethical and Social Implications of Digital Mental Health

As well as the rapid growth of DMH as a field, there are significant concerns about ethical and social implications that raise the importance of critically investigating its technologies. The twin categories of *ethical* and *social* implications are difficult to define and demarcate (Otto et al., 2021), but fields such as health technology assessment have necessitated some practical distinctions to prize the two apart. On Lysdahl

et al.'s (2016) approach, ethical implications are defined as those engaging "moral values and norms, i.e., what is good or bad (what is a good life for humans?) and what is right and wrong (what is the right way for a human to act in a given situation?)" (p. 60). Social implications, by comparison, are those "related to the interpersonal organization of human cohabitation […] represented in social norms" (ibid., p. 77) at various levels of society, from the family or workplace to patterns of wealth distribution and status. Bioethicists, naturally focused on the former category, have stressed their distinctive ability to "cut through the hype" (Skorburg & Friesen, 2021, p. 24) and highlight potential ethical risks and failures of DMH, with inspiration being derived from Beauchamp and Childress's classic bioethical principles, for example (Coghlan et al., 2023; Lederman et al., 2020). While these principles of beneficence, nonmaleficence, autonomy, and justice are certainly helpful when unpacked and applied to DMH, their resonance is not immediately clear to non-bioethicists: thus, the brief review of ethical values below will use equity, safety, privacy, and relationality as a simpler match for the DMH literature cited. These more specific ethical values for DMH can be seen as building on bioethicists' traditional concerns of justice, wellbeing, and autonomy in digital contexts with the additional advantage of incorporating the increasing focus on ethical relationships (Jennings, 2016). The norms of interpersonal organization relevant to DMH, meanwhile, are at least as varied and overlap considerably with the ethical values; but responsibility, stigmatization, and bias provide some key pillars around which we can arrange an overview of DMH's social implications.

The value of equity is often advanced in favor of DMH interventions as means of increasing the accessibility of mental health services (Tekin, 2020) but is challenged by the 'digital divide' in access to and engagement with DMH, which many believe will only compound existing inequities (Skorburg & Yam, 2022). Moreover, the threat to equity from DMH may come from biases embedded in the technologies themselves, with underrepresentation of marginalized populations in the data that AI models are trained on meaning that the accuracy of their results may depend on human biases (Park et al., 2022). Safety is another key threat, according to DMH's critics, because of the absence of mechanisms for "accountability and oversight" (Martinez-Martin et al., 2020, p. 3) in a field where patient outcomes are so variable, patients so vulnerable and expansion so rapid. The crucial value of privacy appears to conflict with the use of devices to collect data on "global positioning system (GPS), voice, keyboard usage, photos, video and overall phone usage behavior" (Torous et al., 2019, p. 97), with regulatory mechanisms currently seeming insufficient for data protection. Studies suggest that even health and wellbeing apps "certified as clinically safe and trustworthy by the UK NHS Health Apps Library" (Huckvale et al., 2015, p. 1) commonly place personal information at risk of interception (Parker et al., 2019). Finally, the idea of the 'therapeutic alliance', a model of working that involves "an agreement on goals, an assignment of task or a series of tasks, and the development of bonds" (Bordin, 1979, p. 253), represents a concept of relationality that could be threatened by DMH. Indeed, Hollis et al.'s (2018) study shows that central focuses of people with lived experience and care providers were DMH's "effect on the therapeutic alliance between therapist and patient, and concerns about the potential impact of removing the human aspects of face-to-face treatments" (p. 852).

When it comes to social implications, responsibility is popularly reduced to a simple moral burden acquired by individuals, whereas social scientists have emphasized how "understandings of responsibility are flexibly and dynamically transformed" (Trnka & Trundle, 2014, p. 149) based on interpersonal ties and expectations. Thus, in the literature on DMH, scholars have frequently scrutinized how it contributes to an excessive framing of one's mental well-being as an individual responsibility (Parker et al., 2018), though one must also consider how responsibilities for *technologies* and their impacts are assigned (Lenk & Maring, 2001). AI has been said to bring 'responsibility gaps' in medicine whereby no agent can be held accountable for the outputs (Bleher & Braun, 2022), and in mental healthcare specifically it "introduces a level of uncertainty around standards of correct diagnoses that might require redefining the scope of responsibility" (Minerva & Giubilini, 2023, p. 815). Regarding stigma, DMH tools often promise to reduce that which is involved in seeking mental healthcare by presenting private and non-threatening treatment options (Tekin, 2020); however, critics have alleged that they might in fact amplify stigmatizing social norms (Wies et al., 2021). One possible route to stigmatization is via the arguably related social process of medicalization, to which DMH interventions have been said to contribute by shifting attention to a quantified understanding of individual well-being that discounts patients' social experiences of suffering (Cosgrove et al., 2020; Maturo & Gibin, 2022). Finally, bias in DMH systems has been identified as a result of assumptions in human society broadly, as well as psychiatry more specifically, that tend to reveal themselves in algorithmic functioning (Pendse et al., 2022; Straw & Callison-Burch, 2020). Though the biases in DMH may themselves be either desirable or undesirable, consequences for the scope and effectiveness of diagnosis and treatment are significant for stakeholders from data scientists and clinicians to regulators and policy makers (Mellino et al., 2022; Timmons et al., 2022).

Clearly, substantial attention has been given to the ethical and social implications of DMH already, but little research has looked at the related views of experts with technical knowledge who are involved in developing DMH interventions. "Academics," as Arigo et al. (2019, p. 75) note, "regularly develop apps and devices for use in behavioral interventions, which they then test in controlled studies". Moreover, there are many university-based DMH projects that are driven by the domain-specific knowledge possessed by academics, clinicians, and software professionals who work in partnership (Austin et al., 2020). With a combination of practical experience and advanced knowledge, these groups seem well-placed to address the wider ethical and social implications of DMH technologies but have rarely been consulted for their views on such issues. The closest body of research is that which investigates barriers and facilitators to engagement, implementation, or optimization in DMH, which features perspectives from users and patients as well as some technical experts that are relevant to ethical and social challenges (Borghouts et al., 2021; Tönnies et al., 2021). For example, Nicholas et al.'s (2017) analysis of a survey conducted at a specialist DMH conference in Australia identified that experts, including researchers and health professionals, believed that problems such as stigma and data privacy issues could stall technology adoption. Yet, perceptions of the unacceptability of DMH technologies are cashed out in these studies not as sources of ethical or social insight but as barriers "that will likely affect implementation of such systems" (Bucci et al., 2019,

p. 1). They represent, therefore, a distinct approach from that which considers expert perspectives on ethical and social implications with an aim not to facilitate DMH's introduction but to build more sophisticated conceptualizations of what is at stake in its development and use.

One study that seeks substantive ethical recommendations from experts is Martinez-Martin et al.'s (2021) attempt to forge consensus among a sample of experts, yet its sample was notably broad, including not just computer scientists and psychiatrists but also ethicists and lawyers. Of course, consultation with professional experts on ethical, legal, and social issues can often be illuminating, and it is important to recognize that there are many forms of expertise in mental healthcare, including 'expertise by experience' (Schleider, 2023). Indeed, the recommendations of people with lived experience are increasingly being sought within commercial and academic DMH development settings (Fortuna et al., 2020; Veldmeijer et al., 2023), contributing to a blurring of the boundaries between scientific and experiential forms of expertise (Egher, 2022). Yet it remains the case that, for the most part, those involved in developing DMH interventions arrive at the domain not due to lived experience but rather technical background, and therefore to approach DMH technologies through the eyes of others typically means approaching them from 'the outside'. A problem with this kind of external inquiry is that it is liable to fall into reductionist approaches including the reduction of what De Vries (2006, p. 120) calls "the human aspects in technology" to a numbers game where consequences are seen as objectively calculable risks. As a remedy to such reductionism, De Vries suggests "treating ethical problems as design problems [which] means: taking into account the full complexity of the situation" and "seeking creative solutions that overcome the conflict in that dilemma" (2006, p. 121). Aligned with the complexity already identified in socio-technical systems for DMH interventions, this view of technologies' ethical implications (and, similarly, social implications) as design problems supposes not that they can be solved but that they can be reckoned with creatively. Thus, in the spirit of Whitbeck's (2011) similar suggestion that in ethics, as in design, "there is rarely, if ever, a uniquely correct solution or response, or indeed, any predetermined number of correct responses" (p. 139), this paper turns to the perceptions of experts involved in designing DMH interventions. Though clinicians and academics may not ordinarily be perceived as 'designers', understanding the term to encompass technology and service design serves to emphasize the role of creative engagement with ethical and social complexities in the perceptions of DMH's technical experts.

3 Materials and Methods

3.1 Aims and Methodology

Due to the limited empirical work that has been devoted to experts' perceptions of ethical and social implications of DMH technologies, this study's primary research aims to address the following broad, exploratory research question:

How do science and technology experts working in DMH conceptualize the ethical and social implications of DMH technologies?

It does so by using semi-structured interviews undertaken in the methodological mold of the 'expert interview' that has been theorized and employed by social scientists since the early 1990s (Döringer, 2020; Flick, 2009). Offering a simple way of gathering good quality data, as experts are generally familiar with presenting information and opinions clearly and professionally (Bogner et al., 2009), expert interviews have also been used in similar studies on digital and data-driven technologies in healthcare (Buhr & Schicktanz, 2022; Viberg Johansson et al., 2022). Within this approach, Bogner and Menz (2009) define experts not just as people who possess specific knowledge but as those whose "action orientations, knowledge and assessments decisively structure, or help to structure, the conditions of action of other actors" (p. 54). Such a conceptualization of expert knowledge as being tied to a certain kind of social influence informed this study's use of the "theory-generating expert interview" (Bogner & Menz, 2009, p. 48) as a specific form of investigation described by these authors. Accordingly, the interviews conducted for this study approached science and technology experts not simply as sources of technical information or even of 'insider knowledge' about systems and processes (though such knowledge claims were also of interest). The primary role of expert interviews in this study was to build an understanding of participants' attitudes and practices that would answer the exploratory research question above about ethical and social conceptualizations in relation to their projects.

3.2 Participants

This study used purposive sampling techniques to identify the professional web pages of potential participants who met the following inclusion criteria:

- advanced degree (master's or doctoral) in a science or technology discipline or equivalent professional experience
- experience of supporting the development or application of a DMH technology

'Science and technology' disciplines were understood to include the natural sciences, medical sciences, psychological sciences, and information technology but not straightforward social sciences such as sociology or empirical approaches to law or bioethics. As a behavioral science discipline, the case of psychology notoriously bridges the gap between natural and social sciences (Bunge, 1990); but given the prominence of *clinical* psychologists in DMH, it was considered sufficiently technical to qualify under the first inclusion criterion. Potential participants' professional web pages were identified by combining relevant Google search terms such as "mental health", "psychiatr*", "psycholog*", "digital" and "artificial intelligence". Although the criteria would have permitted recruitment of experts outside academia, it was hypothesized that individuals in the same sector as the researcher would be more responsive to requests for participation. As a result, only experts working on UK university-based projects were recruited to comprise the study's sample. Inclusion

of professional experience equivalent to an advanced degree was deemed important to recognize the technical knowledge of software professionals, though ultimately all participants had a doctoral education except for one software developer from a humanities background.

Participant recruitment took place between June and July 2022, with a total of 29 people being contacted on their public university email addresses with variations on a standardized invitation email and an attached information sheet. A small sample size of 10–12 participants was considered reasonable given the depth of the interviews as well as the narrow range of eligible experts who had not only researched or published on but taken part in the development of DMH interventions (Baker & Edwards, 2012). This sample size is in the range used by recent expert interview studies in similar topic areas, which have frequently selected around 12 participants (Hangl et al., 2023; Weidener & Fischer, 2023), though some samples have been far smaller (Minkkinen et al., 2022; Mirbabaie et al., 2021). Thus, recruitment stopped once 11 interviews had been confirmed with participants and consent forms had been distributed.

As part of the purposive sampling strategy, the researcher attempted to achieve diversity across the demographic characteristics on which data were collected in interviews: gender, age, professional role and educational (i.e., disciplinary) background (see Table 1). Diversity was generally achieved in these respects, but the gender balance was disappointing, with 4 female and 7 male participants in the sample. In addition to demographic considerations, the sample purposefully contained participants who had worked on a wide variety of types of DMH technologies, including internet- and app-based therapeutics, remote monitoring tools, VR for exposure therapy and AI-enabled assessment of mental health.

Table 1 Participant characteristics

Participant	Gender	Age	Role	Education	DMH Project
1	Male	40–50	Research software developer	Health informatics	Remote monitoring
2	Male	50–60	Professor	Medicine	AI and electronic health records (EHRs)
3	Male	30–40	Research fellow	Computer science and economics	AI and EHRs
4	Female	30–40	Lecturer	Psychology	Internet-based therapeutics
5	Female	60–70	Professor	Clinical psychology	App-based therapeutics
6	Male	50–60	Virtual reality developer	Humanities	VR for exposure therapy
7	Male	40–50	Lecturer	Engineering	AI and speech data
8	Female	40–50	Senior lecturer	Computer science	Remote monitoring and AI
9	Male	30–40	Postdoctoral researcher	Physics	AI and neuroimaging
10	Male	40–50	Clinical academic	Clinical psychology	Remote monitoring and app-based therapeutics
11	Female	40–50	Lecturer	Clinical psychology	App-based therapeutics

3.3 Data Collection and Ethical Approval

To gather rich data about the science and technology experts' views, the researcher conducted semi-structured individual interviews with all participants between July and August 2022, with each interview being scheduled to last one hour. Interviews were audio recorded in most cases by the researcher's laptop microphone and in others using Microsoft Teams, with recordings lasting an average of 51 min. All interviews took place via Microsoft Teams except for one, which took place in a quiet room at the participant's workplace.

The interviews followed a topic guide that began with a brief introduction to the project's aims and two questions to establish the participant's experience with DMH technologies, i.e., what they had done and how the relevant technologies could be used. There were then four main section titles drawn from Dean et al.'s (2021) tool to facilitate sociotechnical inquiry, which was selected on the basis that it offered clear and distinct vantage points on ethical and social implications: value, optimization, consensus, and failure. Within each section, generally open-ended questions were intended to elicit both how the participants' conceptualized and how they might respond to ethical and social implications of DMH technologies. Optional probes were indicated beneath the core questions to guide the researcher in improvising requests for more information, clarification, or examples from the participants (Bearman, 2019). The four-section structure was maintained across interviews, although section titles and questions were rephrased and reordered depending on what was found to be successful, resulting in a final version that was used in the last two interviews (see Appendix 1).

This study involved human subjects and therefore required ethical approval via submission of a Minimal Risk Self-Registration Form and, as a postgraduate dissertation project, authorization from the dissertation supervisor. Ethical clearance was granted on 16 June 2022.

3.4 Coding and Analysis

The interviews were analyzed according to the conventions of the Framework Method, and the process consisted of interview transcription, familiarization with the data, coding, developing an analytical framework, charting data into a coding matrix and interpretation of results (Gale et al., 2013). This method of analysis was selected due to its applicability to studies aiming to interpret participants' viewpoints as well as the researcher's familiarity with the technique from previous work.

At the transcription stage, the researcher listened to the recordings in full and used an 'intelligent verbatim' mode of transcription such that the text generally followed participants' speech, but extraneous hedges, pauses and repetitions were omitted for the sake of readability (McMullin, 2021). The resulting Microsoft Word transcripts were then uploaded to NVivo for familiarization via repeated close reading. In the next stage, open coding allowed the researcher to consider each phrase, line, or sentence individually and use a broadly inductive approach to remain open to a wide range of codes emerging as theoretically significant or descriptively rich. Within the first few transcripts, the researcher identified a split between codes concerning the

implications of DMH technologies (i.e., the possible impacts they will have in the future) and those concerning the value of DMH technologies (i.e., the inherent qualities that inspire its advocates in the present). Given the richness of the data and the limits of the present study, only the *implications* of DMH can be dealt with in this paper, and the contents are deliberately limited to suit this focus. One of the advantages of the Framework Method, along with other 'codebook' approaches to thematic analysis is its flexibility in allowing a researcher to focus on one 'frame' of the analysis for the purposes of a publication or discussion (Brooks et al., 2015; Gibbs, 2012). Thus, codes relating to the value of DMH have been left aside by the researcher to receive sufficient attention in light of relevant literature in future work.

In the next stage, the researcher reviewed the codes assigned and devised categories to group them together through an iterative process of trial and error that culminated when no new codes emerged and, therefore, data saturation had been achieved. Nine codes were summarized individually in the analytical framework under three broad descriptive categories: 'implications for users, 'implications for healthcare professionals and systems' and 'implications for society' (see Table 2). All transcripts were then carefully coded using the analytical framework's nine codes and data were manually charted into a coding matrix on Microsoft Word, providing a concise guide to the key data points. Close interpretation of the results, combined with the researcher's interest in complex sociotechnical systems, gave rise to one overarching theme: 'complexity of implications'. This theme is presented as this study's core finding in the next section and broken up according to the three categories of implications, with results being addressed in what follows in relation to the overall theoretical notion of complexity.

Table 2 Analytical framework for the theme 'implications of DMH'

Category	Code	Description
Implications for users	"Maybe it just makes things worse"	DMH negatively affecting users' mental or physical health
	"What happens when we miss people?"	DMH failing to benefit certain users as it is intended to
	"Information security"	DMH technologies threatening privacy or protection of users' data
Implications for healthcare professionals and systems	"Is the computer going to replace me?"	DMH technologies replacing humans within current systems
	Changing relationships	DMH technologies altering the interpersonal relations involved in care
	"Reluctance"	DMH technologies deterring potential engagement in wider systems
Implications for society	Social inequities	DMH technologies creating or exacerbating inequitable distributions or relations in society
	"Sustainability and green issues"	DMH technologies impacting the natural environment
	Public mistrust	DMH technologies causing fear, mistrust, or panic in wider society

4 Results

4.1 Implications for Users

At the individual level, several participants saw a risk of DMH applications being counterproductive and posing threats to patients' well-being, either through action or omission. One alluded to the possibility of remote therapy enabling certain behaviors that a patient may wish to combat, such as "[not] getting out of the house" and "[not] seeing people in the flesh" (P10).[1] As with the other results in this section, this perception of DMH inevitably rests on some degree of speculation about users' likely responses to interventions, but for the purposes of this study it tells us something important about how experts perceive their technologies. Other possibilities for harm that participants noted were unnecessary anxiety due to the "false positives" (P5) produced by digital monitoring systems and insensitive communication of any negative results: "how do you tell vulnerable people that they're getting worse in such a way that's not going to make them worse?" (P7).

An example twice used to illustrate the idea that digital tools could threaten patients' well-being was that of suicide risk, with P4 identifying the possibility of someone taking their own life while being remotely monitored, which face-to-face care might help to avoid. Another participant gave the more speculative example of conversational agents recommending suicide to individuals who are depressed, with existing work on OpenAI's GPT-3 (Generative Pre-trained Transformer 3) being used as evidence:

> [GPT-3 has] been fed, you know, billions of texts, and actually [an AI practitioner]… basically, he started the conversation, he did an adversarial example and in one of the questions he asked GPT-3, well, he said, 'I'm depressed. Should I commit suicide?'. And GPT-3's answer was, 'yeah, I think you should' (P8).[2]

Beyond the risk of neglecting or even increasing suicide risk, one participant flagged the possibility of digital tools failing to protect well-being by conveying that prediction models will "inevitably" (P7) miss people who are at risk. Another suggested that a mental health diagnosis could "contribute to being stigmatized", which was partly a concern because "there are many ethical issues with diagnosis itself" (P9) but particularly significant if diagnoses are simply taken from a computer. These views suggest that participants' perceptions of risks to users gave weight to the role of unintentional harms that emerge from the relationships between clinical judgement, social stigmatization, computational accuracy, and sensitivity to human wishes.

By engaging this web of varied sources, participants also depicted risks to users as raising almost insurmountable questions for the assignment of responsibility. In some

---

[1] Participants are referred to in abbreviated form according to the numbers assigned in Table 1, e.g., Participant 10 is here called 'P10'.

[2] Interviews were conducted before OpenAI's release of GPT 4, but the concern seems to apply to any similar large language model.

cases, participants even added a further layer of social and normative complexity by presenting the matter in legal terms:

> [I]f something does support decisions and that something gets it wrong, who carries the legal can? Who carries responsibility for it? Is it the person who made the algorithm? Or is it the trust that bought the algorithm? Or is it still the clinician who used the algorithm? So, there's a medico-legal, sort of, difficult area for that (P2).

Explicit mention was sometimes made of a "complex world with lots of interacting pieces that are moving, changing" (P1), for example, that could be related to the indeterminacy of responsibility in DMH: "I don't think anyone's cracked it yet fully" (P1). Even without this clear appeal to a complex causal network, however, participants' discourse of indeterminate responsibility took hold in a context of actions distributed widely across a heterogeneous system.

Threats to data protection were also conceptualized as risks to users, and the majority of participants said that factors such as information security (P9), the General Data Protection Regulation (P4) and information governance frameworks (P3) featured in their daily work. Participants used examples to illustrate what could happen to individuals as a result of "risks around identification" (P2), data being "seen by people that should not see them" (P9) or—a particular concern—data being passed on to other parties. For instance, one participant mentioned the worry that "you may find you do a job application and then somebody knows that you have a certain medical condition that you haven't disclosed or things like that" (P10). Similarly, two participants drew on contemporary debates around data concerning individual health risk being sold to insurance companies, although "probably more in the US than here" (P10), so that they can provide targeted quotes: "You know, 'we'll cover you for 'X, Y and Z because you're obviously at risk of A, B and C'" (P7). These discussions frequently started with a concern for privacy and moved on to such issues of exploitation or manipulation, with participants appealing to the idea that data collected for DMH applications would be "very sensitive" (P8).

Discussing the use of speech data to detect mental illness, one elucidated why this particular case may be especially sensitive:

> [S]peech data is so personal. There's a lot of us in our speech, and part of why I love it is there's so much of us in there. There's all this richness of information in there of what they're saying, how they're saying that, why are they saying it, you know, all these things we can infer and learn. And that then means, yeah, we have some responsibilities to look after that data properly (P7).

This coheres with the general agreement among participants that personal information intended for the delivery of mental health services must be guarded safely if used by DMH interventions.

4.2 Implications for Healthcare Professionals and Systems

At the level of professionals and systems, participants discussed the potential disruption to existing roles and relationships that could come as a consequence of DMH. In terms of the patient perspective, participants suggested that wariness around personal data usage could weaken engagement with the system and have a ripple effect through research and practice:

> Any data breach… will create an absolute scare for people and then you might get, you know, a kind of general aversion from people to opt-in for their data to be available. And if there has been identification or data breaches or so forth you may get less data being made available and this will impede research as a result (P8).

Participants also noted healthcare professionals' concerns, with organization-level impacts being articulated primarily through the lens of 'replacement', which is based on the idea that "the machine is trying to be a human and trying to be better than a human" (P2). They expressed that, at one level, "healthcare professionals may feel threatened in terms of their job role" (P4), with one participant contextualizing professionals' fear of "'is the computer going to replace me?'" within the wider societal fear of "is the computer going to come in and take away the role I have?'" (P5). Since none of the participants had an active clinical role, it is important not to take these comments as first-person testimonies, yet the notion of role replacement did feature particularly in interviews with those who had worked on digital therapeutics and qualified in psychology. These participants were particularly likely to tie the fear of replacement to a deeper concern about the loss of valuable interactions that could result from broader changes in healthcare:

> [T]here's also 'will the computer be damaging for the relationship that I have with the patient?'… you know, all of us, I have, anyway, experience of sitting with the GP [General Practitioner] who doesn't look at me but just looks at the computer screen and types away and doesn't seem to notice I'm in the room! (P5).

Two participants related the threat of relationship disruption to the classic therapeutic alliance and suggested that colleagues had been concerned about the impact of DMH technologies on this partnership-oriented approach to working with clients. Though anecdotal, this bears out Hollis et al.'s (2018) finding cited earlier that care providers fear DMH's impacts on the therapeutic alliance. Even where participants thought their own interventions would not fundamentally alter the nature of the therapeutic relationship, they voiced concern about a possible future where "the AI would take care of the interaction with the patient [by] digitizing the algorithms that the therapist has in their heads" (P6).

To counter concerns about 'replacement', participants tended to develop a competing vision of DMH technologies accommodating themselves to the existing system:

> I think the key thing with digital interventions is to try to fit in with existing practice and enhance it, rather than trying to change practice too much just by telling people they've got to change practice (P2).

The idea of 'fitting in' occurred at several levels, being used to refer to the technological aim to "integrate the new technology with the existing system" (P3) but also the preservation of the role of the mental health professional. According to one participant's account of her own intervention team, "we wouldn't want the digital mental health space to ever take the place of a therapist" (P4). This narrative of accommodation was supported by the limitation of a DMH technology's role to some clear use that was seen as complementary, rather than threatening, to that of a human professional in the existing system. For example, an AI tool may be used "to help clinicians make better decisions rather than to make it make decisions for them" (P2), or perhaps, in a therapeutic environment, "the software acts as a coach or supervisor that's holding the more junior therapist's hand as they go through the therapy." (P11).

Within participants' delineation of the role of DMH in mental health systems, there was also a recurring suggestion of optionality:

> Tools, that's all they are. They're just tools to enrich their practice. That's a really key thing, and that's how we think of the technology. It's a tool to help the clinician work more effectively, but some people just don't need that (P6).

This message that DMH technologies are merely one possible means of reaching standard clinical goals seemed to function as a way of emphasizing that the introduction of digital tools into mental healthcare need not be thought of as a call for system change. Participants involved in digital therapeutics stressed that the therapeutic relationship was merely mediated via their particular DMH technologies:

> It's something called a 'blended approach'. So, the therapy is not entirely delivered by the digital technology, but it's sort of supported. So, there's still human contact and sessions, but there are extra things that happen and that are supported by digital technology (P10).

The idea of preservation goes hand in hand with that of integration, with the abiding impression from participants being that DMH technologies can preserve the human element in mental healthcare while introducing new practices into professional practices and systems.

4.3 Implications for Society

Although one participant mentioned the potential environmental impact of DMH technologies that he attributed to the general nature of "high-performance computing" (P7), by far the major consequence raised at the societal level was inequity. Participants seemed broadly cognizant of the idea that the use of DMH technologies could, under certain circumstances, replicate or exacerbate existing inequities and

saw the issue through two lenses: the digital divide and bias. These were each mentioned by around half of the participants.

The digital divide was of concern to participants in light of the possibility that DMH interventions are "repeating the same old failures of access for tools which are good and which are beneficial for people but there isn't any serious attempt to distribute them equitably" (P5). Various participants also explained different dimensions of this divide. These included differences in access to standard digital devices, e.g., a "substantial number of individuals… do not own technology, computer, laptop, mobile phones" (P10); access to standard devices of varying quality, e.g., "using an old, cheap Android phone versus using a brand-new Apple phone" (P7); and access to novel devices such as virtual reality headsets. As well as identifying the financial elements of the digital divide, participants also referred to intersecting aspects of "sociodemographic reach" (P4), exploring how ethnicity, age, education, and gender would affect engagement with user research and interventions: "in terms of psychological therapies, younger black men are less likely to access the therapies, less likely to be referred, less likely to access, [less] likely to stay in the therapies and maybe [they] show poor outcomes" (P5). This was tied to the issue of bias, with one participant emphasizing that her own team's composition was relevant:

> [T]he intervention development team are all of a white demographic, and we're all female. So, there are going to be biases in how we've written the intervention and how well are we engaging with other demographics. And I think that links to my point earlier that we've recognized that this is a bias, [that] this does have potential downstream effects on perhaps further enhancing that digital divide (P4).

This example illustrates a complex relationship between human biases and the design of digital interventions that was also discussed by participants in relation to algorithmic models and differences in the performance of an assessment tool, for example: "something like depression affects the pitch of your voice, which is affected by if you're male or female" (P7). Further complicating factors were identified, including skewed datasets, which could be derived from intervention teams' own research or large-scale data sources. For example, one participant suggested that the student populations often favored by researchers would not be representative in terms of lifestyle and biobanking tends to capture a "very specific" (P9) range of people. The complexity associated with identifying whether biases emerge out of human error or from existing social disparities was also considered: "sometimes you think about the bias maybe [it] is from the practice like how people are doing the work but sometimes you will say, 'Oh no, this is maybe from some lack of resources in some area'" (P3).

When thinking about DMH technologies' consequences from a broad societal perspective, participants were keen to highlight the responsibilities of developers to think about and report on biases in their applications: "we need to test our data from different angles, different viewpoints and report different factors about how we're doing the analysis" (P7). However, a variety of actors across social domains were also engaged, from mental health trusts, who could support "digital inclusion" (P10) initiatives to overcome financial barriers to accessing care, to the "commissioner

level" (P4), with service commissioners being seen as capable of designing services to combat health inequities.

Another major societal implication that participants suggested was the threat of widespread mistrust and misunderstanding across society, which relates back to the potential threats to individual users considered above. One participant suggested, for example, that privacy breaches like the Cambridge Analytica scandal are "always going to reflect badly on anyone sort of working in technology and working in digital health" (P7). In these cases, the risks to public trust were seen as especially profound and the net of responsibility was cast even wider, with participants allocating certain responsibilities to media institutions:

> So, the data governance committees and, you know, the gatekeepers for the data… they have a responsibility to ensure safe usage of the data and safe utilization of the data within digital health frameworks. There's also a huge responsibility for the media and, unfortunately, the media in the UK and around the world has not done a great job, kind of, reporting things as they are (P8).

As with other issues raised by participants, this was linked to the competing interests and incentives that different sectors involved in shaping opinion about DMH technologies have: "Journalists are sometimes very keen to do the 'AI replacing doctors'-style headline, which is not helpful" (P7). It was also recognized that scientists must play a role in reaching out to the public and "need to, you know, be engaging with everybody to get that broader, wider understanding of academia in general, let alone the value of health research" (P7).

Responsibility for societal consequences was therefore presented as involving actors across society, in the case of public mistrust, as well as a range of technical factors such as bias and the quality of devices in the case of inequity.

5 Discussion

It was suggested earlier that viewing DMH interventions as complex sociotechnical systems emphasizes several important features: heterogeneity, interdependence, and distribution. Now that the ethical and social implications of DMH have been considered in detail, this discussion can revisit these dimensions of complexity in ethical and social terms before considering the above results in their light. Each of these concepts is itself open to interpretation and thus brings new challenges for scholars and theorists of DMH; but collectively they represent a step towards achieving analytical clarity on sociotechnical DMH systems by breaking up the general notion of complexity into several parts. Given the "tokenistic manner" (Carroll et al., 2023, p. 1) in which complexity has often been discussed in health and social care research, this refinement of the category should be of value to future analyses of DMH, though the features' specific applications will depend on the ethical and social focuses of such scholarship. For example, researchers pragmatically interested in the implementation of particular DMH interventions may benefit from acknowledging the distributed nature of control in their research contexts and looking for specific points of

potential influence (Long et al., 2018). Meanwhile, scholarship concerned primarily with seeking explanations of social phenomena might look to the interdependence of heterogenous parts as a way of understanding how a plurality of local rules gives rise to the seemingly unintelligible order of complex systems (Paley & Eva, 2011). While other scholars can take on more case-based applications of the three key features of complexity and further demonstrate their utility, what is most important for this paper is that they shed light on ethical and social insights offered by participants involved in DMH interventions from science and technology backgrounds.

Through the notion of *heterogeneity*, theorists see complex sociotechnical systems in ideal terms as "hybrid collectives" (Callon, 2004, p. 3) while questioning the sustainability of "normative heterogeneity" and "countervailing centres of power" (Silbey, 2009, pp. 343–344). Accordingly, while this study's participants voiced support for collaboration among diverse stakeholders, they also presented threats to individual privacy and safety in terms of challenging relationships between heterogeneous actors, including those with commercial and public focuses. This raises questions for future research about what kind of values should be adopted and how power could be allocated to hold such a diverse group of actors together in any given DMH system. Sharon (2018) suggests that when private companies are involved in digital health projects, with less stringent requirements imposed upon them than on researchers, "data may flow between medical, social and commercial contexts governed by different privacy norms" (p. 2). In this context, a series of scandals bridging the public and private domains relating to health data and technology companies remains fresh in the public imagination, with examples from the UK being the care.data and DeepMind-Royal Free scandals (Samuel et al., 2021). Thus, the participants' emphasis on data stewardship speaks to the fragility of individual rights within a heterogeneous assemblage of public-private partnerships and invites consideration of how complex systems may fail due to "internal contradictions" (Ivory & Alderman, 2005, p. 5) but adapt via reorganization. The crucial normative questions for practitioners and policymakers are whether commercial entities represent a coherent and desirable feature of DMH systems and whether barriers should be placed between these presently porous domains.

Secondly, the presence of *interdependence* in complex systems adds credence to Respess's (2022, p. 125) suggestion that "limitations [of DMH] regarding therapeutic relationships, experiences, and settings have gone undertheorized" in comparison to ethical concerns such as privacy and equity. It is important, therefore, to move from a simple observation of interdependence to asking what kinds of dependencies between humans or technologies are socially desirable or ethically permissible. One element of participants' depictions of disruption to existing professional roles was their emphasis on affective aspects of system failure that are transmitted through a sense of threat to conventional positions and relationships. In previous theorizations of resistance to new digital health technologies, considerable emphasis has been placed on fears, such as healthcare professionals' "fear of change", "fear of losing power or control" and "fear of losing moral or professional integrity" (Nilsen et al., 2016, p. 14). While wariness of change itself is not necessarily tied to interpersonal interactions, the latter two fears are closely intertwined with the relationships that afford professionals power and social or ethical standing. The participants' sugges-

tion that such reluctance would lead to weakened engagement with health systems conveys a distinctive way in which interdependence between technology, people and organizational culture could lead to DMH systems' failure due to perceptions of social and ethical esteem. Yet, in response to the perceived threat to therapeutic relationality, participants offered a counterpoint to the narrative of new technologies being socially disruptive (Hopster, 2021), suggesting that DMH technologies may, in fact, be integrated with existing components, resulting not in system failure but in a form of system adaptation.

Thirdly, the distribution of *control* in a complex sociotechnical system and resultant problems for the allocation of responsibility arose in the context of inequity as a potential implication of DMH technologies due to the digital divide and biases inherent in their design and functioning. The 'problem of many hands', i.e., there are many people who could be responsible for harms, is well-known in computing and engineering ethics (van de Poel et al., 2011), but complex systems face the additional problem of "'many things' that […] causally contribute to the technological action" (Coeckelbergh, 2019, p. 2051). When it came to inequity, participants did not endorse a simplistic model of division between 'the haves and the have-nots' (Qureshi, 2014; Rich et al., 2019) but rather recognized multiple sources of inequity, including material access to devices, human and automated biases in design, and wider societal conditions. This range of aspects of society that participants saw as feeding into inequities highlights not only how widely control is distributed in DMH systems but also its dispersion to components, such as computers, that are not morally responsible in the way we take humans to be. The implication is that challenging inequities requires not only calling on human actors to discharge their duties but also reaching into domains where biases and disparities are mechanized to exert human control. In practical terms, participants recognized that because no single locus of control exists in a complex system, reversing inequities in DMH "will require approaches at multiple levels […] as well as for different stakeholders" (Banerjee, 2021, p. 78).

5.1 Limitations

Whatever this study's contributions, it is important to note some of its limitations. First, the sample was composed exclusively of science and technology experts working in university-based DMH research and development, which will have its own ethical and social culture not shared by other sites of expertise. This is especially significant given the hybrid role of academics and their colleagues as both observers and creators of the social meaning attached to innovations as well as participants in both 'academia' and 'industry', which could yield dynamic and unstable identities (Jain et al., 2009; Lam, 2010). Future research with experts who are more firmly based in commercial activities may contribute by finding little appreciation of system complexity or the ethical and social implications identified in this study. Second, to allow for a richer discussion, this study included only negative forms of ethical and social implications within its frame of analysis. This is even though previous research in AI and data ethics has used the term 'implications' as a deliberately broad choice not only capturing negative impacts "as alternative terms like 'issues', 'risks', or 'challenges' might suggest" (Whittlestone et al., 2019, p. 6). It would be wise for research

in the future to reflect a bigger picture of scholarly critique and advocacy, including defenses of DMH that suggest refocusing on "where DMH holds value" (Bautista & Schueller, 2022, p. 191) and caution against "dismissing [DMH apps] in their nascent stages" (Stroud et al., 2022, p. 210). Third, this study focused on faithfully unearthing, rather than debating, experts' subjective perceptions of what is ethically and socially at stake in DMH and therefore included more limited interrogation of their assumptions than perhaps a more oppositional approach to interview data (Langley & Meziani, 2020). Nevertheless, this study's conception of its purpose and its findings could be complemented by research purposefully recruiting critical experts from science and technology disciplines or adversarial collaborations where critical scholars work with DMH developers to forge consensus (Clark & Tetlock, 2023).

6 Conclusion

Some scholarship suggests that DMH "represents an 'extreme form' of the legal and ethical issues of digital medicine more generally" (Gooding, 2019, p. 2) due to the distinctive feasibility of remote delivery in mental healthcare and the sensitivity of mental health information. For these and other reasons, the experts interviewed for this study can thus be seen as both actors at the more feasible stage of a healthcare revolution and observers with special proximity to a range of sensitive implications. This dual function has given this small study of eleven participants added depth and allowed the researcher to access first-hand observations on dynamic interactions in socio-technical systems through the eyes of experts involved in DMH development. By adopting an expert interview approach, the study has made explicit science and technology experts' interpretations of DMH with an aim to add to existing knowledge about perceptions of its ethical and social implications. Through a single theme of 'complexity of implications', it has highlighted the ways in which experts understood the interactions between social and technical aspects of DMH systems and touched on issues concerning users, healthcare professionals and organizations, and society at large.

Returning to the idea with which this paper began, the study described here should encourage greater dialogue between researchers interested in the ethical and social implications of DMH and science and technology experts working on developments in the field. There are, no doubt, many interlocking barriers to such cross-pollination, not least DMH's tendencies towards sparse consideration of ethical and social issues in scientific publications, narrow conceptions of research risks and ethical review, and exclusion of service user perspectives (Gooding & Kariotis, 2021). What this study underlines is that it is possible for research focused on ethical and social aspects, while using established methods and theories appropriate to such enquiry, to take active steps to engage the voices of participants in DMH development from science and technology disciplines and thereby learn from their interpretations. With other fields having struggled to make ethical and social analysis specific enough to guide practical developments (Shumpert et al., 2014), this paper frames the current deficits in DMH as providing an opportunity for the entire digital health field to enrich existing debates using insights from the scientific and technological community.

Appendix 1: Interview Topic Guide

Introduction

- The aim of this study is to understand how technology experts conceptualize and address the ethical and social implications of digital mental health technologies.
- My final report will seek partly to identify common themes in how technology experts think about digital mental health and find links with expert insights from disciplines like philosophy and sociology.
- This is a semi-structured interview, so I'll be asking you some questions for a few different sections that might seem a little rigid, but after you lay out your general answer, we can exchange ideas more freely so that I can understand your views and experiences.

Section 1: Getting started

a. What have you done that involves digital mental health technologies?
b. How could tools like [participant's technology/technologies] could be used in practice, either in the short-term or the long-term?

Section 2: Value

a. How do you imagine tools like [participant's technology/technologies] to be of ethical or social value in the future?
b. How might you measure the value of tools like [participant's technology/technologies] to society or individuals?

Section 3: Barriers

a. What are the barriers to realising the value of digital mental health tools?
b. What can be done to ensure the potential value of digital mental health tools is realised in practice?

Section 4: Risks

a. What do you think are the major ethical and social risks of digital mental health tools?
b. Whose responsibility is it to mitigate potential risks of digital mental health tools?

Section 5: Consensus

a. How do people in your field reach agreement about the future of digital mental health tools?
b. Who should be involved in discussing the future of digital mental health tools?

Biographical data:

- Age
- Gender
- Occupation
- Education

Closing section:

- Perfect. Those were the last questions that I had planned.
- Thank you very much for your time.
- Is there anything else you would like to say?

Data Availability The data collected during this study are not publicly available as participants did not give written consent for their data to be shared.

Declarations

Ethical Approval This manuscript is composed of original material, derived from my Master of Science dissertation, that is not under review elsewhere. Ethical approval was granted by the King's College London Research Ethics Committee with the reference MRSU-21/22-29807.

Conflict of Interest The author declares that there are no conflicts of interest.

References

Andersson, G., & Titov, N. (2014). Advantages and limitations of internet-based interventions for common mental disorders. *World Psychiatry*, *13*(1), 4–11. https://doi.org/10.1002/wps.20083.
Arigo, D., Jake-Schoffman, D. E., Wolin, K., Beckjord, E., Hekler, E. B., & Pagoto, S. L. (2019). The history and future of digital health in the field of behavioral medicine. *Journal of Behavioral Medicine*, *42*(1), 67–83. https://doi.org/10.1007/s10865-018-9966-z.

Armontrout, J., Torous, J., Fisher, M., Drogin, E., & Gutheil, T. (2016). Mobile mental health: Navigating new rules and regulations for digital tools. *Current Psychiatry Reports*, *18*(10), 1–7. https://doi.org/10.1007/s11920-016-0726-x.

Aryana, B., Brewster, L., & Nocera, J. A. (2019). Design for mobile mental health: An exploratory review. *Health and Technology*, *9*(4), 401–424. https://doi.org/10.1007/s12553-018-0271-1.

Austin, D., May, J., Andrade, J., & Jones, R. (2020). Delivering digital health: The barriers and facilitators to university-industry collaboration. *Health Policy and Technology*, *10*(1), 104–110. https://doi.org/10.1016/j.hlpt.2020.10.003.

Baker, S. E., & Edwards, R. (2012). *How many qualitative interviews is enough?* National Centre for Research Methods. Retrieved November 24, 2023, from https://eprints.ncrm.ac.uk/id/eprint/2273/4/how_many_interviews.pdf.

Balcombe, L., & De Leo, D. (2021). Digital mental health amid COVID-19. *Encyclopedia*, *1*(4), 1047–1057. https://doi.org/10.3390/encyclopedia1040080.

Balcombe, L., & De Leo, D. (2022). Human-computer interaction in digital mental health. *Informatics*, *9*(1), 14. https://doi.org/10.3390/informatics9010014.

Banerjee, A. (2021). Digital health interventions and inequalities: The case for a new paradigm. *BMJ Evidence-Based Medicine*, *26*(3), 77–78. https://doi.org/10.1136/bmjebm-2019-111282.

Baños, R. M., Herrero, R., & Vara, M. D. (2022). What is the current and future status of digital mental health interventions? *The Spanish Journal of Psychology*, *25*(e5), 1–8. https://doi.org/10.1017/sjp.2022.2.

Bauer, J. M., & Herder, P. M. (2009). Designing socio-technical systems. In A. Meijers (Ed.), *Philosophy of technology and engineering sciences* (pp. 601–630). Elsevier.

Bautista, J., & Schueller, S. M. (2022). Digital mental health deserves investment but the questions are which interventions and where? *AJOB Neuroscience*, *13*(3), 191–193. https://doi.org/10.1080/21507740.2022.2082596

Bearman, M. (2019). Focus on methodology: Eliciting rich data: A practical approach to writing semi-structured interview schedules. *Focus on Health Professional Education: A Multi-Professional Journal*, *20*(3), 1–11. https://doi.org/10.11157/fohpe.v20i3.387.

Bevan Jones, R., Stallard, P., Agha, S. S., Rice, S., Werner-Seidler, A., Stasiak, K., Kahn, J., Simpson, S. A., Alvarez-Jimenez, M., Rice, F., Evans, R., & Merry, S. (2020). Practitioner review: Co-design of digital mental health technologies with children and young people. *Journal of Child Psychology and Psychiatry*, *61*(8), 928–940. https://doi.org/10.1111/jcpp.13258.

Bickman, L. (2020). Improving mental health services: A 50-year journey from randomized experiments to artificial intelligence and precision mental health. *Administration and Policy in Mental Health and Mental Health Services Research*, *47*(5), 795–843. https://doi.org/10.1007/s10488-020-01065-8.

Bleher, H., & Braun, M. (2022). Diffused responsibility: Attributions of responsibility in the use of AI-driven clinical decision support systems. *AI and Ethics*, *2*(4), 747–761. https://doi.org/10.1007/s43681-022-00135-x.

Bogner, A., & Menz, W. (2009). The theory-generating expert interview: Epistemological interest, forms of knowledge, interaction. In A. Bogner, B. Littig, & W. Menz (Eds.), *Interviewing experts* (pp. 43–80). Palgrave Macmillan.

Bogner, A., Littig, B., & Menz, W. (2009). Introduction: Expert interviews—an introduction to a new methodological debate. *Interviewing experts* (pp. 1–13). Palgrave Macmillan.

Bordin, E. S. (1979). The generalizability of the psychoanalytic concept of the working alliance. *Psychotherapy: Theory Research & Practice*, *16*(3), 252. https://doi.org/10.1037/h0085885.

Borghouts, J., Eikey, E., Mark, G., De Leon, C., Schueller, S. M., Schneider, M., Stadnick, N., Zheng, K., Mukamel, D., & Sorkin, D. H. (2021). Barriers to and facilitators of user engagement with digital mental health interventions: Systematic review. *Journal of Medical Internet Research*, *23*(3), e24387. https://doi.org/10.2196/24387.

Brooks, J., McCluskey, S., Turley, E., & King, N. (2015). The utility of template analysis in qualitative psychology research. *Qualitative Research in Psychology*, *12*(2), 202–222. https://doi.org/10.1080/14780887.2014.955224.

Bucci, S., Berry, N., Morris, R., Berry, K., Haddock, G., Lewis, S., & Edge, D. (2019). "They are not hard-to-reach clients. We have just got hard-to-reach services." Staff views of digital health tools in specialist mental health services. *Frontiers in Psychiatry*, *10*. https://doi.org/10.3389/fpsyt.2019.00344

Buhr, L., & Schicktanz, S. (2022). Individual benefits and collective challenges: Experts' views on data-driven approaches in medical research and healthcare in the German context. *Big Data and Society*, *9*(1). https://doi.org/10.1177/20539517221092653.

Bunge, M. (1990). What kind of discipline is psychology: Autonomous or dependent, humanistic or scientific, biological or sociological? *New Ideas in Psychology*, *8*(2), 121–137. https://doi.org/10.1016/0732-118x(90)90002-j.

Callon, M. (2004). The role of hybrid communities and socio-technical arrangements in the participatory design. *Journal of the Center for Information Studies*, *5*(3), 3–10.

Carayon, P. (2006). Human factors of complex sociotechnical systems. *Applied Ergonomics*, *37*(4), 525–535. https://doi.org/10.1016/j.apergo.2006.04.011.

Carayon, P., Hancock, P., Leveson, N., Noy, I., Sznelwar, L., & van Hootegem, G. (2015). Advancing a sociotechnical systems approach to workplace safety– developing the conceptual framework. *Ergonomics*, *58*(4), 548–564. https://doi.org/10.1080/00140139.2015.1015623.

Carroll, Á., Collins, C., McKenzie, J., Stokes, D., & Darley, A. (2023). Application of complexity theory in health and social care research: A scoping review. *British Medical Journal Open*, *13*(3), e069180. https://doi.org/10.1136/bmjopen-2022-069180.

Carter, H., Araya, R., Anjur, K., Deng, D., & Naslund, J. A. (2021). The emergence of digital mental health in low-income and middle-income countries: A review of recent advances and implications for the treatment and prevention of mental disorders. *Journal of Psychiatric Research*, *133*, 223–246. https://doi.org/10.1016/j.jpsychires.2020.12.016.

Clark, C. J., & Tetlock, P. E. (2023). Adversarial collaboration: The next science reform. In L. Frisby, R. E. Redding, W. T. O'Donohue, & S. O. Lilienfeld (Eds.), *Ideological and political bias in psychology* (pp. 905–927). Springer Nature.

Coeckelbergh, M. (2019). Artificial intelligence, responsibility attribution, and a relational justification of explainability. *Science and Engineering Ethics*, *26*(4), 2051–2068. https://doi.org/10.1007/s11948-019-00146-8.

Coghlan, S., Leins, K., Sheldrick, S., Cheong, M., Gooding, P., & D'Alfonso, S. (2023). To chat or bot to chat: Ethical issues with using chatbots in mental health. *Digital Health*, *9*. https://doi.org/10.1177/20552076231183542.

Cohn, S., Clinch, M., Bunn, C., & Stronge, P. (2013). Entangled complexity: Why complex interventions are just not complicated enough. *Journal of Health Services Research and Policy*, *18*(1), 40–43. https://doi.org/10.1258/jhsrp.2012.012036.

Cosgrove, L., Karter, J. M., McGinley, M., & Morrill, Z. (2020). Digital phenotyping and digital psychotropic drugs: Mental health surveillance tools that threaten human rights. *Health and Human Rights*, *22*(2), 33–39.

De Vries, M. J. (2006). Ethics and the complexity of technology: A design approach. *Philosophia Reformata*, *71*(2), 118–131. https://doi.org/10.1163/22116117-90000383.

De Wolf, T., & Holvoet, T. (2005). Emergence versus self-organisation: Different concepts but promising when combined. In: S.A. Brueckner, G. Di Marzo Serugendo, A. Karageorgos, R. Nagpal (Eds.), *Engineering self-organising systems* (pp. 1–15). ESOA 2004. Lecture Notes in Computer Science. Springer.

Dean, S., Gilbert, T. K., Lambert, N., & Zick, T. (2021). Axes for sociotechnical inquiry in AI research. *IEEE Transactions on Technology and Society*, *2*(2), 1–9. https://doi.org/10.1109/tts.2021.3074097.

Döringer, S. (2020). The problem-centred expert interview. Combining qualitative interviewing approaches for investigating implicit expert knowledge. *International Journal of Social Research Methodology*, *24*(3), 1–14. https://doi.org/10.1080/13645579.2020.1766777.

Effken, J. A. (2002). Different lenses, improved outcomes: A new approach to the analysis and design of healthcare information systems. *International Journal of Medical Informatics*, *65*(1), 59–74. https://doi.org/10.1016/s1386-5056(02)00003-5.

Egher, C. (2022). *Digital healthcare and expertise: Mental health and new knowledge practices*. Palgrave Macmillan.

Emmelkamp, P. M. G., & Meyerbröker, K. (2021). Virtual reality therapy in mental health. *Annual Review of Clinical Psychology*, *17*(1), 495–519. https://doi.org/10.1146/annurev-clinpsy-081219-115923.

Flick, U. (2009). *An introduction to qualitative research*. Sage.

Fortuna, K. L., Naslund, J. A., LaCroix, J. M., Bianco, C. L., Brooks, J. M., Zisman-Ilani, Y., Muralidharan, A., & Deegan, P. (2020). Digital peer support mental health interventions for people with a lived experience of a serious mental illness: Systematic review. *JMIR Mental Health*, *7*(4), e16460. https://doi.org/10.2196/16460.

Frisch, M. (1959). *Homo Faber: A report*. Harcourt Brace Jovanovich.

Gale, N. K., Heath, G., Cameron, E., Rashid, S., & Redwood, S. (2013). Using the framework method for the analysis of qualitative data in multi-disciplinary health research. *BMC Medical Research Methodology*, *13*(1), 1–8. https://doi.org/10.1186/1471-2288-13-117.

Ganapathy, A., Clough, B. A., & Casey, L. M. (2021). Organizational and policy barriers to the use of digital mental health by mental health professionals. *Telemedicine and e-Health*, *27*(12), 1332–1343. https://doi.org/10.1089/tmj.2020.0455.

Gibbs, G. (2012). *Template analysis Part 7, How many themes? Video interview with Professor Nigel King*. Retrieved November 25, 2023, from https://www.youtube.com/watch?v=TfXgvr-DMrc&list=PL2EF5188800C558D0&ind ex=7, accessed 9 November 2023.

Gooding, P. (2019). Mapping the rise of digital mental health technologies: Emerging issues for law and society. *International Journal of Law and Psychiatry*, *67*. https://doi.org/10.1016/j.ijlp.2019.101498.

Gooding, P., & Kariotis, T. (2021). Ethics and law in research on algorithmic and data-driven technology in mental health care: Scoping review. *JMIR Mental Health*, *8*(6), e24668. https://doi.org/10.2196/24668.

Graham, A. K., Lattie, E. G., & Mohr, D. C. (2019). Experimental therapeutics for digital mental health. *JAMA Psychiatry*, *76*(12), 1223–1224. https://doi.org/10.1001/jamapsychiatry.2019.2075.

Greenhalgh, T., & Papoutsi, C. (2018). Studying complexity in health services research: Desperately seeking an overdue paradigm shift. *BMC Medicine*, *16*(1), 1–6. https://doi.org/10.1186/s12916-018-1089-4.

Hangl, J., Krause, S., & Behrens, V. J. (2023). Expert interviews of drivers, barriers and social considerations for AI adoption in SCM. *Technology in Society*, *74*. https://doi.org/10.1016/j.techsoc.2023.102299.

Holland, J. H. (2006). Studying complex adaptive systems. *Journal of Systems Science and Complexity*, *19*(1), 1–8. https://doi.org/10.1007/S11424-006-0001-Z.

Hollis, C., Morriss, R., Martin, J., Amani, S., Cotton, R., Denis, M., & Lewis, S. (2015). Technological innovations in mental healthcare: Harnessing the digital revolution. *British Journal of Psychiatry*, *206*(4), 263–265. https://doi.org/10.1192/bjp.bp.113.142612.

Hollis, C., Sampson, S., Simons, L., Davies, E. B., Churchill, R., Betton, V., Butler, D., Chapman, K., Easton, K., Gronlund, T. A., Kabir, T., Rawsthorne, M., Rye, E., & Tomlin, A. (2018). Identifying research priorities for digital technology in mental health care: Results of the James Lind Alliance Priority Setting Partnership. *The Lancet Psychiatry*, *5*(10), 845–854. https://doi.org/10.1016/s2215-0366(18)30296-7

Hopster, J. (2021). What are socially disruptive technologies? *Technology in Society*, *67*, 1–8. https://doi.org/10.1016/j.techsoc.2021.101750.

Huckvale, K., Prieto, J. T., Tilney, M., Benghozi, P. J., & Car, J. (2015). Unaddressed privacy risks in accredited health and wellness apps: A cross-sectional systematic assessment. *BMC Medicine*, *13*(1), 1–13. https://doi.org/10.1186/s12916-015-0444-y.

Iorfino, F., Occhipinti, J. A., Skinner, A., Davenport, T., Rowe, S., Prodan, A., Sturgess, J., & Hickie, I. B. (2021). The impact of technology-enabled care coordination in a complex mental health system: A local system dynamics model. *Journal of Medical Internet Research*, *23*(6), p.e25331. https://doi.org/10.2196/25331.

Ivory, C., & Alderman, N. (2005). Can project management learn anything from studies of failure in complex systems? *Project Management Journal*, *36*(3), 5–16. https://doi.org/10.1177/875697280503600302.

Jain, S., George, G., & Maltarich, M. (2009). Academics or entrepreneurs? Investigating role identity modification of university scientists involved in commercialization activity. *Research Policy*, *38*(6), 922–935. https://doi.org/10.1016/j.respol.2009.02.007.

Jankovic, D., Bojke, L., Marshall, D., Saramago Goncalves, P., Churchill, R., Melton, H., Brabyn, S., & Gega, L. (2020). Systematic review and critique of methods for economic evaluation of digital mental health interventions. *Applied Health Economics and Health Policy*, *19*(1), 17–27. https://doi.org/10.1007/s40258-020-00607-3.

Jennings, B. (2016). Reconceptualizing autonomy: A relational turn in bioethics. *Hastings Center Report*, *46*(3), 11–16. https://doi.org/10.1002/hast.544.

Ladyman, J., Lambert, J., & Wiesner, K. (2013). What is a complex system? *European Journal for Philosophy of Science*, *3*, 33–67. https://doi.org/10.1007/s13194-012-0056-8.

Lam, A. (2010). From ivory tower traditionalists to entrepreneurial scientists? Academic scientists in fuzzy university—industry boundaries. *Social Studies of Science*, *40*(2), 307–340. https://doi.org/10.1177/0306312709349963.

Langley, A., & Meziani, N. (2020). Making interviews meaningful. *The Journal of Applied Behavioral Science*, *56*(3), 370–391.

Lederman, R., & D'Alfonso, S. (2021). The digital therapeutic alliance: Prospects and considerations. *JMIR Mental Health*, *8*(7), 1–4. https://doi.org/10.2196/31385.

Lederman, R., D'Alfonso, S., Rice, S., Coghlan, S., Wadley, G., & Alvarez-Jimenez, M. (2020, June 15). Ethical issues in online mental health interventions. Proceedings of the 28th European Conference on Information Systems (ECIS). https://aisel.aisnet.org/ecis2020_rp/66/.

Lenk, H., & Maring, M. (2001). Responsibility and technology. In A. E. Auhagen, & H. W. Bierhoff (Eds.), *Responsibility: The many faces of a social phenomenon* (pp. 93–108). Routledge.

Leonardi, P. M. (2012). Materiality, sociomateriality, and socio-technical systems: What do these terms mean? How are they different? Do we need them? *Materiality and Organizing: Social Interaction in a Technological World*, *25*(10), 1093. https://doi.org/10.1016/j.apergo.2006.04.011.

Long, K. M., McDermott, F., & Meadows, G. N. (2018). Being pragmatic about healthcare complexity: Our experiences applying complexity theory and pragmatism to health services research. *BMC Medicine*, *16*(1). https://doi.org/10.1186/s12916-018-1087-6.

Lovejoy, C. A. (2019). Technology and mental health: The role of artificial intelligence. *European Psychiatry*, *55*, 1–3. https://doi.org/10.1016/j.eurpsy.2018.08.004.

Lysdahl, K. B., Mozygemba, K., Burns, J., Chilcott, J. B., Brönneke, J. B., & Hofmann, B. (Eds.). (2016). *Guidance for assessing effectiveness, economic aspects, ethical aspects, socio-cultural aspects and legal aspects in complex technologies*. INTEGRATE-HTA. Retrieved November 25, 2023, from https://validatehta.eu/wp-content/uploads/2022/07/3.-Guidance-for-assessing-effectiveness-economic-aspects-ethical-aspects-socio-cultural-aspects-and-legal-aspects-in-complex-technologies.pdf.

Martinez-Martin, N., Dasgupta, I., Carter, A., Chandler, J., Kellmeyer, P., Kreitmair, K., Weiss, A., & Cabrera, L. (2020). Ethics of digital mental health during COVID-19: Crisis and opportunities. *JMIR Mental Health*, *7*, 1–9. https://doi.org/10.2196/23776.

Martinez-Martin, N., Greely, H. T., & Cho, M. K. (2021). Ethical development of digital phenotyping tools for mental health applications: Delphi study. *JMIR mHealth and uHealth*, *9*(7), e27343. https://doi.org/10.2196/27343.

Maturo, A. F., & Gibin, M. (2022). The digital forces of medicalization: The role of apps for mental health. In M. Elliott (Ed.), *Research handbook on society and mental health* (pp. 53–65). Edward Elgar Publishing.

McMullin, C. (2021). Transcription and qualitative methods: Implications for third sector research. *VOLUNTAS: International Journal of Voluntary and Nonprofit Organizations*, *34*, 140–153. https://doi.org/10.1007/s11266-021-00400-3.

Mellino, S., Morey, C., & Rohner, C. (2022). Biases in digital health measures. In D. Cirillo, S. Catuara-Solarz, & E. Guney (Eds.), *Sex and gender bias in technology and artificial intelligence* (pp. 95–112). Academic.

Minerva, F., & Giubilini, A. (2023). Is AI the future of mental healthcare? *Topoi*, *42*(3), 809–817. https://doi.org/10.1007/s11245-023-09932-3.

Minkkinen, M., Niukkanen, A., & Mäntymäki, M. (2022). What about investors? ESG analyses as tools for ethics-based AI auditing. *AI and Society*. https://doi.org/10.1007/s00146-022-01415-0.

Mirbabaie, M., Hofeditz, L., Frick, N. R. J., & Stieglitz, S. (2021). Artificial intelligence in hospitals: Providing a status quo of ethical considerations in academia to guide future research. *AI and Society*, *37*(4), 1361–1382. https://doi.org/10.1007/s00146-021-01239-4.

Mohr, D. C., Lyon, A. R., Lattie, E. G., Reddy, M., & Schueller, S. M. (2017). Accelerating digital mental health research from early design and creation to successful implementation and sustainment. *Journal of Medical Internet Research*, *19*(5), e7725. https://doi.org/10.2196/jmir.7725.

Mumford, E. (2006). The story of socio-technical design: Reflections on its successes, failures and potential. *Information Systems Journal*, *16*(4), 317–342. https://doi.org/10.1111/j.1365-2575.2006.00221.x.

Muuraiskangas, S., Harjumaa, M., Kaipainen, K., & Ermes, M. (2016). Process and effects evaluation of a digital mental health intervention targeted at improving occupational well-being: Lessons from an intervention study with failed adoption. *JMIR Mental Health*, *3*(2), e4465. https://doi.org/10.2196/mental.4465.

National Institute of Mental Health (2019). *Technology and the future of mental health treatment*. Retrieved November 25, 2023, from https://www.nimh.nih.gov/health/topics/technology-and-the-future-of-mental-health-treatment.

Nicholas, J., Huckvale, K., Larsen, M. E., Basu, A., Batterham, P. J., Shaw, F., & Sendi, S. (2017). Issues for eHealth in psychiatry: Results of an expert survey. *Journal of Medical Internet Research*, *19*(2), e6957. https://doi.org/10.2196/jmir.6957.

Nilsen, E. R., Dugstad, J., Eide, H., Gullslett, M. K., & Eide, T. (2016). Exploring resistance to implementation of welfare technology in municipal healthcare services– A longitudinal case study. *BMC Health Services Research*, *16*(1). https://doi.org/10.1186/s12913-016-1913-5.

Nordgreen, T., Rabbi, F., Torresen, J., Skar, Y. S., Guribye, F., Inal, Y., Flobakk, E., Wake, J. D., Mukhiya, S. K., Aminifar, A., Myklebost, S., Lundervold, A. J., Kenter, R., Hammar, Å., Nordby, E., Kahlon, S., Sekse, T., Griffin, R. J., Jakobsen, K. F., & Lamo, P., Y (2021). Challenges and possible solutions in cross-disciplinary and cross-sectorial research teams within the domain of e-mental health. *Journal of Enabling Technologies*, *15*(4), 241–251. https://doi.org/10.1108/jet-03-2021-0013.

Norman, D. A., & Stappers, P. J. (2015). DesignX: Complex sociotechnical systems. *She Ji: The Journal of Design Economics and Innovation*, *1*(2), 83–106. https://doi.org/10.1016/j.sheji.2016.01.002.

Otto, I., Kahrass, H., & Mertz, M. (2021). Same same but different? On the questionable but crucial differentiation between ethical and social aspects in health technology assessment. *Zeitschrift für Evidenz Fortbildung Und Qualität Im Gesundheitswesen*, *164*, 1–10. https://doi.org/10.1016/j.zefq.2021.05.007.

Paley, J., & Eva, G. (2011). Complexity theory as an approach to explanation in healthcare: A critical discussion. *International Journal of Nursing Studies*, *48*(2), 269–279. https://doi.org/10.1016/j.ijnurstu.2010.09.012.

Park, J., Arunachalam, R., Silenzio, V., & Singh, V. K. (2022). Fairness in mobile phone–based mental health assessment algorithms: Exploratory study. *JMIR Formative Research*, *6*(6), e34366. https://doi.org/10.2196/34366.

Parker, L., Bero, L., Gillies, D., Raven, M., Mintzes, B., Jureidini, J., & Grundy, Q. (2018). Mental health messages in prominent mental health apps. *The Annals of Family Medicine*, *16*(4), 338–342. https://doi.org/10.1370/afm.2260.

Parker, L., Halter, V., Karliychuk, T., & Grundy, Q. (2019). How private is your mental health app data? An empirical study of mental health app privacy policies and practices. *International Journal of Law and Psychiatry*, *64*, 198–204. https://doi.org/10.1016/j.ijlp.2019.04.002.

Pendse, S. R., Nkemelu, D., Bidwell, N. J., Jadhav, S., Pathare, S., De Choudhury, M., & Kumar, N. (2022). From treatment to healing: Envisioning a decolonial digital mental health. *CHI Conference on Human Factors in Computing Systems*, 1–23. https://doi.org/10.1145/3491102.3501982.

Qureshi, S. (2014). Overcoming technological determinism in understanding the digital divide: Where do we go from here? *Information Technology for Development*, *20*(3), 215–217. https://doi.org/10.1080/02681102.2014.930981.

Respess, S. (2022). Going telemental: Contact and intimacy in digital mental health. In I. Vasiliu-Feltes, & J. Thomason (Eds.), *Applied Ethics in a Digital World* (pp. 125–139). IGI Global.

Rich, E., Miah, A., & Lewis, S. (2019). Is digital health care more equitable? The framing of health inequalities within England's digital health policy 2010–2017. *Sociology of Health and Illness*, *41*(S1), 31–49. https://doi.org/10.1111/1467-9566.12980.

Rosen, A., Gill, N. S., & Salvador-Carulla, L. (2020). The future of community psychiatry and community mental health services. *Current Opinion in Psychiatry*, *33*(4), 375–390. https://doi.org/10.1097/YCO.0000000000000620.

Samuel, G., Roberts, S. L., Fiske, A., Lucivero, F., McLennan, S., Phillips, A., Hayes, S., & Johnson, S. B. (2021). COVID-19 contact tracing apps: UK public perceptions. *Critical Public Health*, *32*(1), 1–13. https://doi.org/10.1080/09581596.2021.1909707.

Schleider, J. L. (2023). The fundamental need for lived experience perspectives in developing and evaluating psychotherapies. *Journal of Consulting and Clinical Psychology*, *91*(3), 119–121. https://doi.org/10.1037/ccp0000798.

Sharon, T. (2018). When digital health meets digital capitalism, how many common goods are at stake? *Big Data & Society*, *5*(2), 1–12. https://doi.org/10.1177/2053951718819032.

Shaw, J. A., & Donia, J. (2021). The sociotechnical ethics of digital health: A critique and extension of approaches from bioethics. *Frontiers in Digital Health*, *127*. https://doi.org/10.3389/fdgth.2021.725088.

Shumpert, B. L., Wolfe, A., Bjornstad, D. J., Wang, S. W., & Campa, M. F. (2014). Specificity and engagement: Increasing ELSI's relevance to nano–scientists. *Nanoethics*, *8*(2), 193–200. https://doi.org/10.1007/s11569-014-0194-x.

Silbey, S. S. (2009). Taming Prometheus: Talk about safety and culture. *Annual Review of Sociology*, *35*(1), 341–369. https://doi.org/10.1146/annurev.soc.34.040507.134707.

Sittig, D. F., & Singh, H. (2015). A new socio-technical model for studying health information technology in complex adaptive healthcare systems. In V. L. Patel, T. G. Kannampallil, & D. R. Kaufman (Eds.), *Cognitive informatics for biomedicine* (pp. 59–80). Springer International Publishing.

Skorburg, J. A., & Friesen, P. (2021). Mind the gaps: Ethical and epistemic issues in the digital mental health response to Covid-19. *Hastings Center Report*, *51*(6), 23–26. https://doi.org/10.1002/hast.1292.

Skorburg, J. A., & Yam, J. (2022). Is there an app for that? Ethical issues in the digital mental health response to COVID-19. *AJOB Neuroscience*, *13*(3), 177–190. https://doi.org/10.1080/21507740.2021.1918284.

Straw, I., & Callison-Burch, C. (2020). Artificial Intelligence in mental health and the biases of language based models. *PLOS ONE*, *15*(12), e0240376. https://doi.org/10.1371/journal.pone.0240376.

Stroud, A. M., Richardson, J., & Sharp, R. R. (2022). Assessing digital mental health apps: The importance of patient-centric measures of utility. *AJOB Neuroscience*, *13*(3), 209–211. https://doi.org/10.1080/21507740.2022.2082591.

Tekin, Ş. (2020). Is big data the new stethoscope? Perils of digital phenotyping to address mental illness. *Philosophy and Technology*, *34*(3), 447–461. https://doi.org/10.1007/s13347-020-00395-7.

Timakum, T., Xie, Q., & Song, M. (2022). Analysis of E-mental health research: Mapping the relationship between information technology and mental healthcare. *Bmc Psychiatry*, *22*(1). https://doi.org/10.1186/s12888-022-03713-9.

Timmons, A. C., Duong, J. B., Fiallo, S., Lee, N., Vo, T., Ahle, H. P. Q., Comer, M. W., Brewer, J. S., Frazier, L. C., S. L., & Chaspari, T. (2022). A call to action on assessing and mitigating bias in artificial intelligence applications for mental health. *Perspectives on Psychological Science*, *18*(5), 1062–1096. https://doi.org/10.1177/17456916221134490.

Tönnies, J., Oeljeklaus, L., Wensing, M., Hartmann, M., Friederich, H. C., & Haun, M. W. (2021). Health policy experts' perspectives on implementing mental health specialist video consultations in routine primary care– a qualitative interview study. *BMC Health Services Research*, *21*(1), 1–12. https://doi.org/10.1186/s12913-021-06676-x.

Torous, J., Andersson, G., Bertagnoli, A., Christensen, H., Cuijpers, P., Firth, J., Haim, A., Hsin, H., Hollis, C., Lewis, S., Mohr, D. C., Pratap, A., Roux, S., Sherrill, J., & Arean, P. A. (2019). Towards a consensus around standards for smartphone apps and digital mental health. *World Psychiatry*, *18*(1), 97–98. https://doi.org/10.1002/wps.20592.

Trist, E. L., & Bamforth, K. W. (1951). Some social and psychological consequences of the longwall method of coal-getting: An examination of the psychological situation and defences of a work group in relation to the social structure and technological content of the work system. *Human Relations*, *4*(1), 3–38. https://doi.org/10.1177/001872675100400101.

Trnka, S., & Trundle, C. (2014). Competing responsibilities: Moving beyond neoliberal responsibilisation. *Anthropological Forum*, *24*(2), 136–153. https://doi.org/10.1080/00664677.2013.879051.

Troxler, P., & Lauche, K. (2003). Knowledge management and learning culture in distributed engineering. In *DS 31: Proceedings of ICED 03, the 14th International Conference on Engineering Design*, Stockholm.

van de Poel, I., Nihlén Fahlquist, J., Doorn, N., Zwart, S., & Royakkers, L. (2011). The problem of many hands: Climate change as an example. *Science and Engineering Ethics*, *18*(1), 49–67. https://doi.org/10.1007/s11948-011-9276-0.

Veldmeijer, L., Terlouw, G., Van Os, J., Van Dijk, O., Van 't Veer, O., J., & Boonstra, N. (2023). The involvement of service users and people with lived experience in mental health care innovation through design: Systematic review. *JMIR Mental Health*, *10*, e46590–e46590. https://doi.org/10.2196/46590.

Viberg Johansson, J., Bentzen, H. B., & Mascalzoni, D. (2022). What ethical approaches are used by scientists when sharing health data? An interview study. *BMC Medical Ethics*, *23*(1). https://doi.org/10.1186/s12910-022-00779-8.

Weidener, L., & Fischer, M. (2023). Artificial intelligence teaching as part of medical education: A qualitative analysis of expert interviews. *JMIR Medical Education*, *9*(1). https://doi.org/10.2196/46428.

Weijnen, M. P. C., Herder, P. M., & Bouwmans, I. (2008). Designing complex systems: A contradiction in terms. In M. Eekhout, R. Visser, & T. Tomiyama (Eds.), *Delft science in design* (pp. 235–254). IOS.

Whitbeck, C. (2011). *Ethics in engineering practice and research*. Cambridge University Press.

Whittlestone, J., Nyrup, R., Alexandrova, A., Dihal, K., & Cave, S. (2019). *Ethical and societal implications of algorithms, data, and artificial intelligence: A roadmap for research*. Nuffield Foundation. Retrieved November 25, 2023, from https://www.nuffieldfoundation.org/sites/default/files/files/Ethical-and-Societal-Implications-of-Data-and-AI-report-Nuffield-Foundat.pdf.

Wies, B., Landers, C., & Ienca, M. (2021). Digital mental health for young people: A scoping review of ethical promises and challenges. *Frontiers in Digital Health*, *3*. https://doi.org/10.3389/fdgth.2021.697072.

World Economic Forum (2021). *Global governance toolkit for digital mental health*. Retrieved November 25, 2023, from https://www3.weforum.org/docs/WEF_Global_Governance_Toolkit_for_Digital_Mental_Health_2021.pdf.