Towards Designing AI Systems that Help Improve Both Short-Term and Long-Term Mental Well-Being

Tony Wang
yw2567@cornell.edu
Cornell University
USA

Qian Yang
qianyang@cornell.edu
Cornell University
USA

Abstract

Prior research shows that artificial intelligence (AI) systems are poor at handling trade-offs in short-term and long-term outcomes, implying potential risks with the use of optimization algorithms in mental health applications. A tool that evaluates an AI system's ability to optimize for both short-term and long-term outcomes and allows management of potential trade-offs between these metrics could empower health and human-computer interaction (HCI) researchers to better design AI-powered longitudinal user experiences. In this poster for the Envisioning the Future of Interactive Health Workshop at CHI '25, we introduce a research artifact that focuses on AI journaling as a case study. The artifact is an experimental platform for evaluating the impact of suggestions provided by AI on the writing process. This platform can be used for building AI writing assistants for mental well-being, bootstrapping research around designing interventions to improve both short-term and long-term outcomes in mental health, and serving as a boundary object between technology and health experts.

CCS Concepts

• **Human-centered computing** → **Interactive systems and tools**.

Keywords

longitudinal user experience, journaling, mental health, artificial intelligence, bandit algorithms

ACM Reference Format:
Tony Wang and Qian Yang. 2025. Towards Designing AI Systems that Help Improve Both Short-Term and Long-Term Mental Well-Being. In *Proceedings of (CHI '25 Workshop on Envisioning the Future of Interactive Health).* ACM, New York, NY, USA, 6 pages.

1 Introduction

Artificial intelligence (AI) has the potential to optimize and tailor mental health interventions to user needs, but research shows that algorithmic optimization leads to trade-offs in short-term and long-term outcomes. For example, optimizing for short-term engagement is known to have long-term effects including both addiction and disengagement [24]. Would similar risks exist in AI systems for mental health, which often are only evaluated in short-term studies prior to lengthier clinical trials? Answering this question could be critical to human-computer interaction (HCI) contributions in prototyping and evaluating new AI interventions for mental health.

A short-term proximal outcome may be an important construct to measure over time to support users taking steps towards a long-term health outcome. In the case of digital therapeutics or well-being apps, researchers may want to design the interface to increase short-term engagement (e.g., time spent in a single usage session) if the intervention is meant to help individuals acquire habits such as journaling, meditation, or breathing exercises. It is also critical to show that engagement with the intervention connects to desirable long-term outcomes (e.g., improved mental health status, emotion regulation skills) [26, 36]. Thus, if we can evaluate an AI system's ability to optimize for short-term and long-term outcomes, and manage trade-offs thoughtfully and explicitly when designing AI systems, health and HCI researchers may gain a new set of tools for designing and prototyping for longitudinal user experience (UX).

To improve the design process of AI systems that improve both short-term and long-term well-being, we developed a research artifact that focuses on AI journaling as a case study. The artifact is an experimental platform for evaluating users' reaction to AI-powered suggestions from systems such as language models or recommender systems. It can be used to build AI writing assistants for mental well-being, support research around the use of short-term and long-term outcomes in health HCI research, and serve as a boundary object between researchers, health professionals, and users. In the workshop poster, we present a specific instantiation focused on journaling for well-being that employs bandit algorithms with the following features:

1. A front-end system that provides writing prompts and features that support journaling.
2. A curated dataset of writing prompts for journaling.
3. A description of logged pilot user data and how we can use logged data as input into recommender systems.
4. Two baseline models for designers and researchers to study the impact of optimization on writing recommendations.

In addition, we propose a longitudinal user study to evaluate how this system can be used to support longitudinal care and invite feedback on the potential of this system from workshop attendees.

2 Related Work

2.1 Longitudinal UX in Mental Health

Researchers studying human-computer interaction (HCI), computer-supported cooperative work, and ubiquitous computing have long investigated longitudinal UX in health contexts. Personal informatics, data tracking, and lived experience models suggest that individuals experience multistage, iterative, and cyclic processes

CHI '25 Workshop on Envisioning the Future of Interactive Health, Yokohama, Japan

when using technology to enhance their health. Given a long-term goal of improving one's health with technology, an individual must break down their goal into smaller decisions about which tools are best for tracking health data and how those tools fit into their lives [10], and then undergo a process of preparing, collecting, and reflecting on their data [21]. HCI designers and researchers have sought to make these steps easier to accomplish, empowering patient self-care and the sharing of data with clinicians [29].

In mental health and HCI research, a longitudinal perspective is critical to designing technologies that support cognitive change and care work. Mental health providers help patients identify actionable short-term goals that are meant to improve patients' long-term well-being [2]. Technology also has the potential to augment their decision-making on how to steer patients' treatments over time [30]. Although digital interventions are designed to supplement professional care and make acquisition of cognitive or behavioral skills more engaging, many commonly available digital interventions have low retention rates [3, 28] and providers are often wary of their efficacy [40]. In contrast to mental health practitioners' expert ability to tailor treatment over long time horizons, the uncertain efficacy of commonly available interventions in fostering desirable long-term outcomes reveals a possible sociotechnical gap [1, 25] in what technology can do for mental health and well-being.

2.2 Designing with Recommender Systems for Mental Health

Exploring novel technical capabilities that meet the standard of expert care can help bridge this gap. Designers and researchers have sought to augment expert-led treatment by developing adaptive interventions that make recommendations to users to encourage behavior change [26]. In particular, optimization techniques from reinforcement learning (RL) algorithms may enable digital interventions to tailor interactions in a way that helps users achieve better long-term outcomes [44]. However, a key challenge in designing such systems lies in identifying the right outcomes to optimize for, particularly in light of concerns around the use of engagement as a proximal outcome in for-profit health apps [34].

Recent research on recommender systems has examined two approaches to balancing short-term and long-term metrics optimization using RL-based approaches such as bandits and controllers. One approach proposes using constraints to steer AI [4, 5] while the other proposes using joint optimization strategies during the model training process [22, 35]. For example, hierarchical bandit models can jointly optimize short-term and long-term metrics so long as there is a relationship between proximal metrics like engagement and long-term ones like retention [33]. These algorithmic approaches to managing temporal tradeoffs in user outcomes may help tackle the challenges of optimizing interventions for mental health and well-being.

However, the choice of what to implement requires careful justification of design decisions regarding dataset and model features. Traditional HCI techniques that exclude training full AI models might help designers focus on end user experiences, but these techniques do not accommodate the stochastic nature of AI systems, optimization of metrics, and interaction with their use over time [42]. Developing working AI systems require a significant number of design choices that are often left up to technical experts at the exclusion of other stakeholders [8].

2.3 Journaling and AI for Mental Health

Journaling helps patients process their thoughts and emotions, reducing the need for acute mental health care [32, 38], and supports the development of resilience by fostering the ability to regulate emotional experiences through self-reflection over time [13, 23]. Psychological studies have shown that journaling's effectiveness as an intervention may need to be monitored and tailored by a clinical expert as excessive rumination and over-immersion may result from the writing process [31, 43]. On the other hand, journaling habits can be difficult to acquire [12] and off-the-shelf journaling apps have low retention rate [3]. To tackle such problems, AI-powered interactive journaling has recently received interest as a way to improve traditional journaling by making it more engaging with AI-tailored prompts [15, 27]. Such systems can foster increased user engagement while serving as a data collection and empathy-building tool for providers to better understand their clients [14].

Journaling is a valuable testbed for studying AI systems that can optimize short-term and long-term outcomes because journaling requires engagement through writing to improve well-being over time. However, research has not addressed the issue of managing tradeoffs between these outcomes. AI compounds existing concerns about potential harms of journaling by making it even easier for users to over-engage with the writing process, limiting the system's ability to help users achieve healthier long-term outcomes. For example, Kim et al. [14] report that clinicians are worried that AI journaling's highly interactive capabilities can lead to rumination, which would further worsen mental illness by leaving patients in low-mood states without techniques to regulate those emotions. How to mitigate possible harms for writers while creating an effective intervention for mental health is a challenging question that AI journaling designers and researchers have yet to solve.

3 Designing a Recommender System for Journaling

3.1 Research Method

We use Research through Design (RtD) [45] to create an AI-powered journaling system that demonstrably balances tradeoffs in short- and long-term metrics in its recommendations. Our approach documents the complexities of choosing various hyperparameters, metrics, and algorithmic components of AI journaling systems to demonstrate how careful selection of short-term and long-term metrics, detailed collection of user input data, the training process for the AI, and choice of algorithm impact longitudinal UX. In doing so, we propose this initial platform as a boundary object that encourages collaboration between HCI, AI, and mental health researchers and practitioners in designing adaptive interventions. Lastly, such a system serves as an intervention, and thus must require a user evaluation to help us understand how effectively it learns to jointly optimize from the perspective of users.

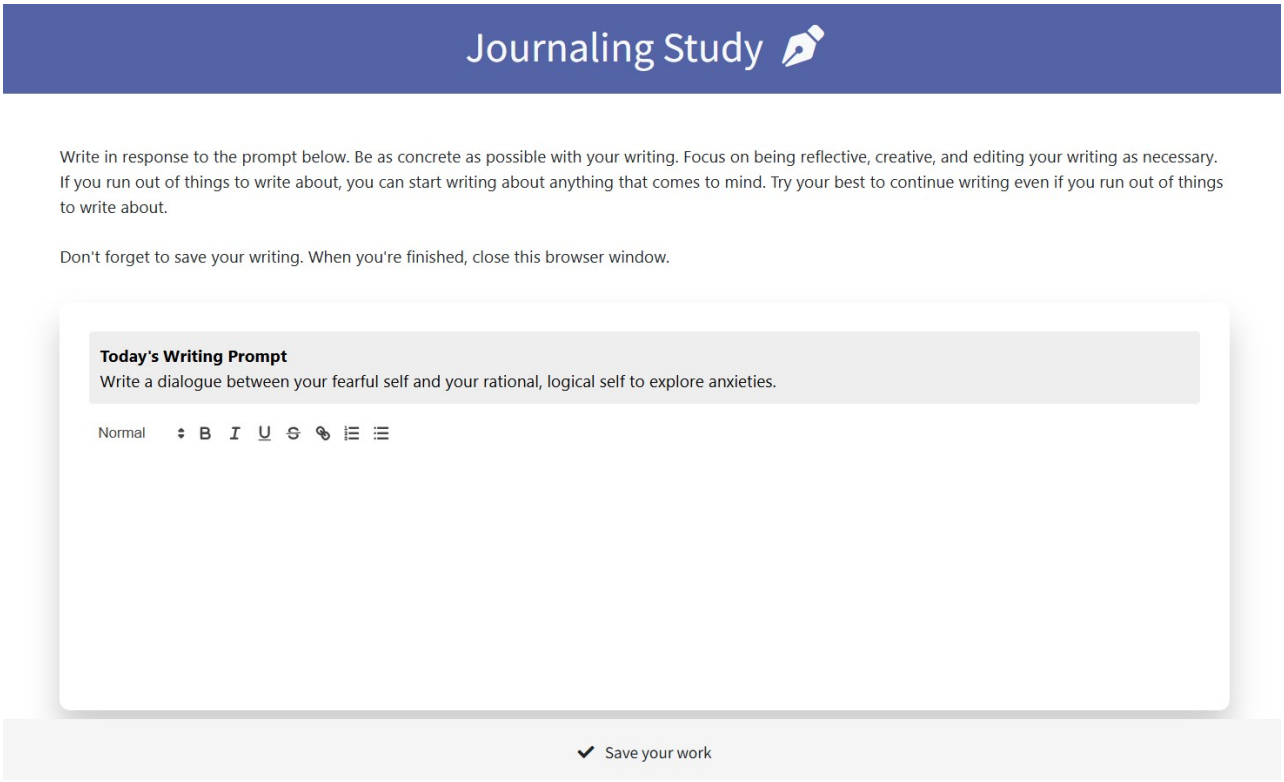

Figure 1: Front-end interface for our early-stage journaling system with a writing prompt, reminders on how to complete a writing task, and an editor for users to write in.

3.2 Prototype Design and Implementation

We use CoAuthor [19] as the base for our front end interface. It comes with a writing editor and logging capabilities that automatically capture what users insert, delete, and edit in their writing. CoAuthor also allows language model prompting capabilities by enabling writers to ask for AI writing suggestions through a back-end call to OpenAI's GPT models for writing suggestions. We modify the front-end to output a writing prompt from a recommender system aimed at helping users journal. Pressing the "Save your work" button surfaces a study completion code that can be used for compensation in longitudinal studies. The system is built using a Flask back-end deployed to Heroku and can be leveraged in both in-person and remote studies.

We choose a recommender system to power this longitudinal journaling intervention for several reasons. First, it can tailor recommendations to individuals based on user history, mirroring the work that mental health practitioners do in tailoring treatment for their patients. Second, recommender systems can commonly be used with online learning algorithms such as contextual bandits, allowing HCI and health researchers a way to investigate how optimization changes longitudinal UX. Third, they have been used in prior research on adaptive interventions for health contexts and are likely of interest to domain experts [41].

3.3 Recommender System Design

3.3.1 State Space. The state space of the recommender system is the state of the user denoted as a *context vector*. The context vector includes self-report and behavioral data. Information about a user's prior experience with journaling and psychotherapy is collected in a screener survey. Logs data from a user's journal entries are added with each writing session.

3.3.2 Action Space. The action space of the system is a set of writing prompts that users respond to when journaling. The system recommends only one prompt at a time, drawn from the larger set of prompts. We curated a set of 247 writing prompts through an iterative search and reflective process starting with eight examples of writing strategies from academic journaling literature such gratitude journaling [37], expressive writing [17, 32], and best possible self writing [16]. We conducted a Google search to find variations of prompts posted on the internet, adding new strategies or techniques while cross-referencing them with mental health literature. To expand the diversity of writing instructions, we also add writing prompts aimed at improving creative flow or sparking ideas. The motivation for these items is in increasing the diversity of recommendations and including items that may be fun or engaging for users, but not relevant to mental health or well-being prima facie. This is not meant to be a representative set of all writing prompts, but a starting set for this research.

3.3.3 Algorithms. The current prototype only supports variations of contextual bandits. We choose this approach for several reasons. Contextual bandits have been used in prior HCI research on self-experimentation tools that fostering user engagement with behavior change interventions [7, 18]. They are relatively simple and interpretable compared to other optimization techniques from reinforcement learning or neural networks, making it a strong choice for HCI design research where our aim is not cutting edge performance of the system but to identify how the design of various features of the system influence its behavior. Furthermore, contextual bandits were originally developed as a way to map patient features to interventions in clinical trials for sequential decision-making problems [39], making them an excellent algorithm for examining longitudinal UX and optimization in health contexts.

In addition to simple contextual bandits, we also implement hierarchical or nested bandit algorithms. Rastogi et al. [33] showed that a nested contextual bandit trained on rewards for both short-term and long-term metrics is capable of video recommendations that balance click-through engagement and user retention. By implementing techniques from recommendation system literature, our system provides a testbed for designing interventions optimized on short-term and long-term outcomes for mental health.

3.3.4 Reward Functions. A reward function must be defined in order for contextual bandits to recommend the right writing prompt that maximizes or minimizes an objective. In this section, we provide a sample of how two types of evaluative metrics can be converted into reward functions in our system.

1. **Short-term (S) metrics.** Proximal outcomes generally fall under short-term metrics and can be operationalized from a single interactive session.
   - **Word-based reward.** The use of pronouns and affective words in journaling is correlated with a reduction in mental health visits [6]. A short-term metric for journaling and well-being may be increasing the amount of affective self-disclosure in a single journal entry.
   - **Interaction logs-based reward.** Difficulty with writing may be reflected in the number of deletions, edits, or pauses during the writing process. A short-term metric for journaling habit development may be a reduction in the number of deletions, edits, and pauses during a session.
2. **Long-term (L) metric.** Long-term outcomes are those that reflect goals for longer time horizons and can be operationalized from multiple sessions.
   - **Word-based reward.** The Coleman-Liau Index (CLI) which provides readability assessment based on character and word structure within a sentence. A greater CLI measure indicates a better writing quality, and an increase of CLI can indicate psychosocial improvement [11]. A long-term metric for journaling and well-being may be to increase the average CLI score over time.
   - **Interaction logs-based reward.** One measure of long-term habit development is resistance to lapses [9]. A long-term metric for journaling habit development may be a smaller number of days missed over time.
3. **Linked reward function.** This is a function critical to studies that optimize short-term and long-term outcomes together. It describes the trade-off relationship between metrics that allows the algorithm to learn which journaling prompts to recommend. Borrowing from [33], a simple parameterization for this function might be

   $$r = w(metric_S \mid prompt) + (1 - w)(metric_L \mid prompt)$$

   where w is a learned weight that helps decide the influence of the long-term metric. This linking function allows a nested contextual bandit to adjust how much emphasis to place on the short-term and long-term metrics.

3.4 Ongoing Data Collection and Baseline User Evaluation

To train the bandits, journaling data and user contexts need to be collected. We are currently conducting a study with crowdworkers (n=15) where each worker is asked to participate in a four week study where they write for 10 minutes daily. This deployment uses a fully random policy such that each writing prompt has an equal chance of being recommended. In the first two weeks, users familiarize themselves with the system and are incentivized to use it daily so that they build a journaling routine. In the last two weeks, users are paid a lump sum and asked to journal at will so long as they meet a minimum amount of entries. We collect data from all four weeks as initial user data for bandit training. Surveys are conducted at the end of the study to understand participants' experience with the system. We analyze writing logs and exit suvery data to investigate the impact of a completely random policy on journaling habits. This study design draws from prior longitudinal studies on AI-powered journaling interventions that separate practice phases with free-will journaling phases [20, 27].

3.5 Future Evaluation of Trained Algorithms with Experts

After collecting initial data and training the system, a longitudinal user study will be conducted to evaluate the performance of trained bandits and collect feedback on the system. To do so, we will conduct a study similar to the one discussed in 3.4, but recruit in-person study participants. We will ask participants to fill out a survey at the beginning and end of the study, in addition to recording their interaction logs data. We plan to compare the results of this study with the baseline random recommendation policy to examine if there is a difference in journaling with a trained contextual bandit recommendation system. Furthermore, we plan to conduct post-study interviews in-person to gain additional qualitative insight into participants' experience with the system.

Finally, we also plan a round of interview studies with mental health and writing experts to gather their feedback. By developing this artifact and collecting some data first, this prototype serves as a boundary object that can support conversations with domain experts who may not have expertise in AI or digital intervention design but do have an interest in the choice of metrics and optimization techniques. This will allow us to collect feedback on our design choices and improve the system such that it can better support practitioners' work in tailoring care.

Acknowledgments

Research reported in this publication was supported by the National Library of Medicine of the National Institutes of Health under award number 1R01LM014306-01. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. This research is also partially supported by the Multi-Investigator Seed Grant "Improving the Robustness of Mobile Sensing and AI Systems for Mental Health Care" from Weill Cornell Medicine.

References

[1] Mark S Ackerman. 2000. The intellectual challenge of CSCW: the gap between social requirements and technical feasibility. *Human–Computer Interaction* 15, 2-3 (2000), 179–203.

[2] Elena Agapie, Patricia A Areán, Gary Hsieh, and Sean A Munson. 2022. A longitudinal goal setting model for addressing complex personal problems in mental health. *Proceedings of the ACM on Human-Computer Interaction* 6, CSCW2 (2022), 1–28.

[3] Amit Baumel, Frederick Muench, Stav Edan, and John M Kane. 2019. Objective user engagement with mental health apps: systematic search and panel-based usage analysis. *Journal of medical Internet research* 21, 9 (2019), e14567.

[4] Craig Boutilier, Martin Mladenov, and Guy Tennenholtz. 2023. Modeling recommender ecosystems: Research challenges at the intersection of mechanism design, reinforcement learning and generative models. *arXiv preprint arXiv:2309.06375* (2023).

[5] Kianté Brantley, Zhichong Fang, Sarah Dean, and Thorsten Joachims. 2024. Ranking with Long-Term Constraints. In *Proceedings of the 17th ACM International Conference on Web Search and Data Mining*. 47–56.

[6] R Sherlock Campbell and James W Pennebaker. 2003. The secret life of pronouns: Flexibility in writing style and physical health. *Psychological science* 14, 1 (2003), 60–65.

[7] Nediyana Daskalova, Jina Yoon, Yibing Wang, Cintia Araujo, Guillermo Beltran Jr, Nicole Nugent, John McGeary, Joseph Jay Williams, and Jeff Huang. 2020. SleepBandits: Guided flexible self-experiments for sleep. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–13.

[8] Fernando Delgado, Stephen Yang, Michael Madaio, and Qian Yang. 2021. Stakeholder Participation in AI: Beyond" Add Diverse Stakeholders and Stir". *arXiv preprint arXiv:2111.01122* (2021).

[9] Daniel A Epstein, Monica Caraway, Chuck Johnston, An Ping, James Fogarty, and Sean A Munson. 2016. Beyond abandonment to next steps: understanding and designing for life after personal informatics tool use. In *Proceedings of the 2016 CHI conference on human factors in computing systems*. 1109–1113.

[10] Daniel A Epstein, An Ping, James Fogarty, and Sean A Munson. 2015. A lived informatics model of personal informatics. In *Proceedings of the 2015 ACM international joint conference on pervasive and ubiquitous computing*. 731–742.

[11] Sindhu Kiranmai Ernala, Asra F Rizvi, Michael L Birnbaum, John M Kane, and Munmun De Choudhury. 2017. Linguistic markers indicating therapeutic outcomes of social media disclosures of schizophrenia. *Proceedings of the ACM on Human-Computer Interaction* 1, CSCW (2017), 1–27.

[12] Annaleis K Giovanetti, JUlia C Revord, Maria P Sasso, and Gerald J Haeffel. 2019. Self-distancing may be harmful: Third-person writing increases levels of depressive symptoms compared to traditional expressive writing and no writing. *Journal of social and clinical psychology* 38, 1 (2019), 50–69.

[13] Debra Jackson, Angela Firtko, and Michel Edenborough. 2007. Personal resilience as a strategy for surviving and thriving in the face of workplace adversity: a literature review. *Journal of advanced nursing* 60, 1 (2007), 1–9.

[14] Taewan Kim, Seolyeong Bae, Hyun Ah Kim, Su-woo Lee, Hwajung Hong, Chanmo Yang, and Young-Ho Kim. 2024. MindfulDiary: Harnessing Large Language Model to Support Psychiatric Patients' Journaling. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. 1–20.

[15] Taewan Kim, Donghoon Shin, Young-Ho Kim, and Hwajung Hong. 2024. DiaryMate: Understanding User Perceptions and Experience in Human-AI Collaboration for Personal Journaling. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. 1–15.

[16] Laura A King. 2001. The health benefits of writing about life goals. *Personality and social psychology bulletin* 27, 7 (2001), 798–807.

[17] Laura A King and Kathi N Miner. 2000. Writing about the perceived benefits of traumatic events: Implications for physical health. *Personality and social psychology bulletin* 26, 2 (2000), 220–230.

[18] Geza Kovacs, Zhengxuan Wu, and Michael S Bernstein. 2018. Rotating online behavior change interventions increases effectiveness but also increases attrition. *Proceedings of the ACM on Human-Computer Interaction* 2, CSCW (2018), 1–25.

[19] Mina Lee, Percy Liang, and Qian Yang. 2022. Coauthor: Designing a human-ai collaborative writing dataset for exploring language model capabilities. In *Proceedings of the 2022 CHI conference on human factors in computing systems*. 1–19.

[20] Yi-Chieh Lee, Naomi Yamashita, and Yun Huang. 2021. Exploring the effects of incorporating human experts to deliver journaling guidance through a chatbot. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW1 (2021), 1–27.

[21] Ian Li, Anind Dey, and Jodi Forlizzi. 2010. A stage-based model of personal informatics systems. In *Proceedings of the SIGCHI conference on human factors in computing systems*. 557–566.

[22] Ziru Liu, Shuchang Liu, Zijian Zhang, Qingpeng Cai, Xiangyu Zhao, Kesen Zhao, Lantao Hu, Peng Jiang, and Kun Gai. 2024. Sequential recommendation for optimizing both immediate feedback and long-term retention. In *Proceedings of the 47th International ACM SIGIR Conference on Research and Development in Information Retrieval*. 1872–1882.

[23] Max S Lohner and Carmela Aprea. 2021. The resilience journal: Exploring the potential of journal interventions to promote resilience in university students. *Frontiers in Psychology* 12 (2021), 702683.

[24] Smitha Milli, Luca Belli, and Moritz Hardt. 2021. From optimizing engagement to measuring value. In *Proceedings of the 2021 ACM conference on fairness, accountability, and transparency*. 714–722.

[25] Sean A Munson, Hasan Cavusoglu, Larry Frisch, and Sidney Fels. 2013. Sociotechnical challenges and progress in using social media for health. *Journal of medical Internet research* 15, 10 (2013), e226.

[26] Inbal Nahum-Shani, Shawna N Smith, Bonnie J Spring, Linda M Collins, Katie Witkiewitz, Ambuj Tewari, and Susan A Murphy. 2018. Just-in-time adaptive interventions (JITAIs) in mobile health: key components and design principles for ongoing health behavior support. *Annals of Behavioral Medicine* (2018), 1–17.

[27] Subigya Nepal, Arvind Pillai, William Campbell, Talie Massachi, Michael V Heinz, Ashmita Kunwar, Eunsol Soul Choi, Xuhai Xu, Joanna Kuc, Jeremy F Huckins, et al. 2024. MindScape Study: Integrating LLM and Behavioral Sensing for Personalized AI-Driven Journaling Experiences. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 8, 4 (2024), 1–44.

[28] Michelle M Ng, Joseph Firth, Mia Minen, and John Torous. 2019. User engagement in mental health apps: a review of measurement, reporting, and validity. *Psychiatric Services* 70, 7 (2019), 538–544.

[29] Francisco Nunes, Nervo Verdezoto, Geraldine Fitzpatrick, Morten Kyng, Erik Grönvall, and Cristiano Storni. 2015. Self-care technologies in HCI: Trends, tensions, and opportunities. *ACM Transactions on Computer-Human Interaction (TOCHI)* 22, 6 (2015), 1–45.

[30] Bruna Oewel, Nadia Azizan, Patricia A Arean, and Elena Agapie. 2024. Technology's Role in Fostering Therapist-Client Collaboration and Engagement with Goals. *Proceedings of the ACM on Human-Computer Interaction* 8, CSCW2 (2024), 1–28.

[31] Antonio Pascual-Leone, Nikita Yeryomenko, Orrin-Porter Morrison, Robert Arnold, and Ueli Kramer. 2016. Does feeling bad, lead to feeling good? Arousal patterns during expressive writing. *Review of General Psychology* 20, 3 (2016), 336–347.

[32] James W Pennebaker and Sandra K Beall. 1986. Confronting a traumatic event: toward an understanding of inhibition and disease. *Journal of abnormal psychology* 95, 3 (1986), 274.

[33] Richa Rastogi, Yuta Saito, and Thorsten Joachims. 2024. MultiScale Policy Learning for Alignment with Long Term Objectives. In *ICML 2024 Workshop on Models of Human Feedback for AI Alignment*.

[34] Marijn Sax. 2021. Optimization of what? For-profit health apps as manipulative digital environments. *Ethics and Information Technology* 23, 3 (2021), 345–361.

[35] Ashudeep Singh, Yoni Halpern, Nithum Thain, Konstantina Christakopoulou, E Chi, Jilin Chen, and Alex Beutel. 2020. Building healthy recommendation sequences for everyone: A safe reinforcement learning approach. In *FAccTRec Workshop*.

[36] Petr Slovak and Sean A Munson. 2024. HCI Contributions in Mental Health: A Modular Framework to Guide Psychosocial Intervention Design. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. 1–21.

[37] Joshua M Smyth, Jillian A Johnson, Brandon J Auer, Erik Lehman, Giampaolo Talamo, and Christopher N Sciamanna. 2018. Online positive affect journaling in the improvement of mental distress and well-being in general medical patients with elevated anxiety symptoms: A preliminary randomized controlled trial. *JMIR mental health* 5, 4 (2018), e11290.

[38] Monika Sohal, Pavneet Singh, Bhupinder Singh Dhillon, and Harbir Singh Gill. 2022. Efficacy of journaling in the management of mental illness: a systematic review and meta-analysis. *Family medicine and community health* 10, 1 (2022).

[39] Ambuj Tewari and Susan A Murphy. 2017. From ads to interventions: Contextual bandits in mobile health. *Mobile health: sensors, analytic methods, and applications* (2017), 495–517.

[40] John Torous, John Luo, and Steven R Chan. 2018. Mental health apps: what to tell patients. *Current Psychiatry* 17, 3 (2018), 21–25.

[41] Thi Ngoc Trang Tran, Alexander Felfernig, Christoph Trattner, and Andreas Holzinger. 2021. Recommender systems in the healthcare domain: state-of-the-art and research issues. *Journal of Intelligent Information Systems* 57, 1 (2021), 171–201.

[42] Qian Yang, Aaron Steinfeld, Carolyn Rosé, and John Zimmerman. 2020. Re-examining whether, why, and how human-AI interaction is uniquely difficult to design. In *Proceedings of the 2020 chi conference on human factors in computing systems*. 1–13.

[43] Carly Yasinski, Adele M Hayes, and Jean-Philippe Laurenceau. 2016. Rumination in everyday life: The influence of distancing, immersion, and distraction. *Journal of experimental psychopathology* 7, 2 (2016), 225–245.

[44] Chao Yu, Jiming Liu, Shamim Nemati, and Guosheng Yin. 2021. Reinforcement learning in healthcare: A survey. *ACM Computing Surveys (CSUR)* 55, 1 (2021), 1–36.

[45] John Zimmerman, Jodi Forlizzi, and Shelley Evenson. 2007. Research through design as a method for interaction design research in HCI. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. 493–502.